

CPS Abstract

CPS Abstract

A Bayesian Approach to estimate the completeness of death registration

Author: Prof. Jairo Alberto Fuquene Patino

Submission ID: 1805

Reference Number: 1805

Brief Description

This is a joint work with Tim Adair (University of Melbourne) we propose a new Bayesian approach to estimate the completeness of death registration.

The use of our approach can allow National Statistical Offices improve model parameter estimates and more accurately predict completeness of death registration.

Abstract

Civil registration and vital statistics (CRVS) systems should be the primary source of mortality data for governments. Accurate and timely measurement of the completeness of death registration helps inform interventions to improve CRVS systems and to generate reliable mortality indicators. In this work we propose the use of hierarchical Bayesian linear mixed models with Global-Local (GL) priors to estimate the completeness of death registration at global, national and subnational levels. The use of GL priors in this paper is motivated for situations where demographic covariates can explain much of the observed completeness but where unexplained within-country (i.e. by year) and between-country variability also play an important role. The use of our approach can allow institutions improve model parameter estimates and more accurately predict completeness of death registration. Our models are based on a dataset which uses Global Burden of Disease (GBD) death estimates based on the GBD 2019 and comprises 120 countries and 2,748 country-years from 1970-2019. To illustrate the effectiveness of our proposal we consider the completeness of death registration in the departments of Colombia in 2017.

CPS Abstract

A Cohort Study of COVID-19 Patients: Hawassa University Referral Hospital, Ethiopia

Author: Mr Anteneh Bezabih Ali

Submission ID: 1848

Reference Number: 1848

Brief Description

Foundational work in statistics and data science: Better understanding and innovation in statistical research and decision-making

Abstract

Abstract: The COVID-19 pandemic is caused by the global spread of the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and has resulted from 622,119,701 confirmed cases and 6,546,118 deaths to date. Most studies of the clinical features of hospitalized COVID-19 patients are from high-income countries, with little information available for developing countries such as Ethiopia. This study aimed to assess the clinical characteristics of hospitalized COVID-19 patients and risk factors for in-hospital mortality in Hawassa, Ethiopia. This cohort included 804 consecutive laboratory-confirmed hospitalised COVID-19 patients from 24, September 2020 - 26, November 2021 at Hawassa University Comprehensive Specialized Hospital. The patient characteristics at hospital admission were characterized. In-hospital mortality was modelled using Cox regression. The median age was 45[$\min$ 0.006, $\max$ 100] years and 64.1% were male. Severe cases accounted for 16.0% of the study population. The most common symptoms were general pain (68.9%), cough (68.7%), shortness of breath (65.7%), and fever (57.6%). Comorbidities were common, with 49.1% having at least one comorbidity, and 12.6% having two or more comorbidities. The most prevalent comorbidity was diabetes (15.9%), followed by hypertension (15.2%). More than half of patients (53.0%) had received treatment before being admitted to the hospital, for example, antibiotics. In univariable Cox regression analysis, sex, age, source of referral, severity status, various symptoms, and comorbidities were associated with an increased risk of death in COVID-19 patients. This study characterised a large cohort of hospitalised COVID-19 patients in Ethiopia and showed that similar to high-income countries, age, sex, comorbidities, and symptoms upon admission were associated with an increased risk of in-hospital death. The average age was remarkably low.

Keywords: COVID-19, Clinical characteristics, date of admission, Univariate cox-ph,

CPS Abstract

A distribution regression approach to estimate the rural-urban well-being in Mexico.

Author: Flor Martinez

Submission ID: 2014

Reference Number: 2014

Brief Description

In surveys, samples are chosen to be representative for a given population. In case we are interested in a particular subpopulation, in general it is not valid to work with the subsample obtained by limiting the original sample to this subpopulation as there is no guarantee that the subsample will be representative for the subpopulation of interest.

In this work, we present a novel approach to the problem of estimating population variables of a survey for non-representative groups. Our approach relies on learning a distribution regression model from the representative distribution (of the group) of some variables in another related survey to the variable of interest. Using the aforementioned model, a rough estimate of the variable of interest is obtained for different subgroups of the sample. We apply our methodology to estimate the well-being of the Mexican rural and urban population in a separate way using data from the Household Income and Expenditure survey. In 2021, the National Institute of Statistics and Geography (INEGI) conducted the National of Survey of Self-Reported Well-being (ENBIARE), but the results are only representative at the National and State level. Using the distribution of specific variables from the Household Income and Expenditure survey (ENIGH), a distribution regression model to the subjective well-being is estimated. The model is used to estimate and compare the well-being of the rural and urban population

Abstract

In surveys, samples are chosen to be representative for a given population. In case we are interested in a particular subpopulation, in general it is not valid to work with the subsample obtained by limiting the original sample to this subpopulation as there is no guarantee that the subsample will be representative for the subpopulation of interest.

In this work, we present a novel approach to the problem of estimating population variables of a survey for non-representative groups. Our approach relies on learning a distribution regression model from the representative distribution (of the group) of some variables in another related survey to the variable of interest. Using the aforementioned model, a rough estimate of the variable of interest is obtained for different subgroups of the sample. We apply our methodology to estimate the well-being of the Mexican rural and urban population in a separate way using data from the Household Income and Expenditure survey. In 2021, the National Institute of Statistics and Geography (INEGI) conducted the National of Survey of Self-Reported Well-being (ENBIARE), but the results are only representative at the National and State level. Using the distribution of specific variables from the Household Income and Expenditure survey (ENIGH), a distribution regression model to the subjective well-being is estimated. The model is used to estimate and compare the well-being of the rural and urban population

CPS Abstract

A polynomial regression model for excess mortality in Mexico 2020-2022 due to the COVID-19 pandemic

Author: Dr Arturo Erdely

Submission ID: 1906

Reference Number: 1906

Files/Uploads

[excessmortmx](#)

Brief Description

Based on the comprehensive national death registry of Mexico spanning from 1998 to 2022 a point and interval estimation method for the excess mortality in Mexico during the years 2020-2022 is proposed based on illness-induced deaths only, using a polynomial regression model. The results obtained estimate that the excess mortality is around 788,000 people (39.3%) equivalently to a rate of 626 per 100,000 inhabitants. The male/female ratio is estimated to be 1.7 times. As a reference for comparison, for the whole period 2020-2020 Mexico's INEGI estimated an excess of mortality between 673,000 with a quasi-Poisson model and 808,000 using endemic channels estimation.

Abstract

Based on the comprehensive national death registry of Mexico spanning from 1998 to 2022 a point and interval estimation method for the excess mortality in Mexico during the years 2020-2022 is proposed based on illness-induced deaths only, using a polynomial regression model. The results obtained estimate that the excess mortality is around 788,000 people (39.3%) equivalently to a rate of 626 per 100,000 inhabitants. The male/female ratio is estimated to be 1.7 times. As a reference for comparison, for the whole period 2020-2020 Mexico's INEGI estimated an excess of mortality between 673,000 with a quasi-Poisson model and 808,000 using endemic channels estimation.

CPS Abstract

Adoption of GSBPM to manage Geospatial Information Production Processes

Author: Mr Manuel Cuéllar-Río

Submission ID: 2082

Reference Number: 2082

Brief Description

To standardize production processes, INEGI uses the Statistical and Geographical Process Model (MPEG, by its Spanish acronym), our version of the GSBPM, since we are also responsible to produce geographical information. The implementation of the MPEG helps to mitigate risks, clearly assign responsibilities, and achieve an explicit division of labor. Likewise, the MPEG provides a common language that facilitates communication between the various actors involved in the information production process.

The text discusses the management of the Mexican Geospatial Data Cube through the MPEG framework, as an illustrative example of the standardized description and documentation of activities carried out by the geospatial information production processes.

Elaborated by: Manuel Cuellar Rio, Jose de Jesus Lujan Salazar, Marco Gutierrez Romero.

Abstract

As part of adhering to the Fundamental Principles of Official Statistics, commitment to international standards is essential to promoting consistency and compatibility with global practices. In this regard, the Generic Statistical Business Process Model (GSBPM) has been incorporated by INEGI through its own Statistical and Geographical Process Model (MPEG, by its Spanish acronym), outlining standardized activities for information production. Additionally, the Technical Standard for the Production Process of Statistical and Geographical Information is the normative instrument currently regulating MPEG's implementation in INEGI, effective since September 2018.

Thus, MPEG allows for the description and documentation of activities for both statistical and geospatial information production processes. An illustrative example of the latter is the Mexican Geospatial Data Cube (CDGM, by its Spanish acronym). The CDGM started as an initiative to explore the potential of satellite images due to their increasing availability, objectivity, and timeliness.

Drawing on the CDGM case, this paper exemplifies the management of a geospatial information production process through the MPEG framework. One of the main challenges was transitioning the CDGM from a research project to a Production Process. For this, it was necessary to identify the key activities of each sub-process that needed to be designed for its continuous execution and the generation of results consistent with identified information needs.

This experience will further support the standardization and systematization of production processes and contribute to future cases involving innovative methods.

CPS Abstract

Africa SDG Progress Index

Author: Eman Aboaldahab Elsayed

Submission ID: 2069

Reference Number: 2069

Files/Uploads

[IAOS-ISI 2024-Abstract](#)

Brief Description

Authors

- Eman Aboaldahab Elsayed
Associate Statistician- the African Center of Statistics of UNECA
- Oliver Chinganya
Director of the African Center of Statistics of UNECA

Strand: Foundational work in statistics and data science: Better understanding and innovation in statistical research and decision-making

Paper Title: Africa SDG Progress Index

Abstract: The Sustainable Development Goals (SDGs) are a universal call for action to end poverty, protect the planet, and ensure that all people enjoy peace and prosperity by 2030. These goals are interconnected and are designed to balance social, economic, and environmental sustainability. While all 17 SDGs are important, there are five key areas often considered transformative due to their broad impact across multiple goals. These areas are including People, Planet, Prosperity, Peace, and Partnership. These five areas highlight the holistic and interconnected nature of the SDGs.

With the increasing number of composite indicators as a useful tool in policy analysis and public communication, it has become exciting and necessary to calculate a simplified quantitative indicator to measure progress of the 2030 agenda which has several and interrelated dimensions. Accordingly, this study aims to construct an SDG index that measures the Africa's progress from the perspective of the 5 transformative areas. The proposed index provides a scientific measure and policy tool to advocate for SDGs transformative areas in Africa. In addition, identifying the disparities between countries and Africa Subregions that would allow paving the way to support decision-makers. The index will help countries to highlight challenges and key enablers to achieve the SDG agenda.

The study uses the SDG monitoring framework to identify set of indicators to measure each dimension. Essential procedures for constructing the composite index will be applied including A) Normalization, B) Assigning the weight and C) Aggregating the indicators within and across dimensions. Moreover, to ensure the statistical validation and reliability of the Africa SDGs index, tests will be employed such as testing collinearity/redundancy between indicators. In addition to testing the sensitivity and robustness of the index. The index will use the SDGs global database as a main data source.

Abstract

Abstract

The Sustainable Development Goals (SDGs) are a universal call to action to end poverty, protect the planet, and ensure that all people enjoy peace and prosperity by 2030. These goals are interconnected and are designed to balance social, economic, and environmental sustainability. While all 17 SDGs are important, there are five key areas often considered transformative due to their broad impact across multiple goals. These areas are including People, Planet, Prosperity, Peace, and Partnership. These five areas highlight the holistic and interconnected nature of the SDGs.

With the increasing number of composite indicators as a useful tool in policy analysis and public communication, it has become exciting and necessary to calculate a simplified quantitative indicator to measure progress of the 2030 agenda which has several and interrelated dimensions. Accordingly, this study aims to construct an SDG index that measures the Africa's progress from the perspective of the 5 transformative areas. The proposed index provides a scientific measure and policy tool to advocate for SDGs transformative areas in Africa. In addition, identifying the disparities between countries and Africa Subregions that would allow paving the way to support decision-makers. The index will help countries to highlight challenges and key enablers to achieve the SDG agenda.

The study uses the SDG monitoring framework to identify set of indicators to measure each dimension. Essential procedures for constructing the composite index will be applied including A) Normalization, B) Assigning the weight and C) Aggregating the indicators within and across dimensions. Moreover, to ensure the statistical validation and reliability of the Africa SDGs index, tests will be employed such as testing collinearity/redundancy between indicators. In addition to testing the sensitivity and robustness of the index. The index will use the SDGs global database as a main data source.

CPS Abstract

An Exploration of Statistical Models on Planned and Unplanned Survey Reporting Domains**Author:** Mr Caiphus Mashaba**Submission ID:** 1883**Reference Number:** 1883**Files/Uploads**[ISI 2024 Contributed paper_Oral](#)**Brief Description**

The aim of this paper is to provide detailed exploration of statistical models that provide estimates for planned and unplanned reporting domains during survey design. It is a global norm that National Statistical Offices depends on surveys to respond to the needs of stakeholders, however survey cost remains a challenge during survey design, estimation, and reporting stages. The paper reviewed literature relating to statistical approaches that may be used to provide estimates at lower (unplanned or planned) reporting domains using statistical models, data fragmentation and other methods that relates to big data. Several techniques such as Small Area Estimation⁰, have been outlined in various literature, however most national statistical offices (NSOs) do not rely on modelled data to respond to user-requests. The base of discussion for this paper focuses on models such as area level models (Fay-Herriot model, model with Correlated Sampling Errors, Time Series and Cross-Sectional Models and Spatial Models), unit level models (Multivariate Nested Error Regression Model¹, Random Error Variance Linear Model², Two level Model³, and Logistic Regression Models⁴). This exploration is paramount to any NSO to evaluate the most suitable models that can be used for unplanned reporting domains. The paper will provide a summary of the recommendations of feasible models that can be utilized efficiently in the production environment.

Abstract

T

ABSTRACT

he aim of this paper is to provide detailed exploration of statistical models that provide estimates for planned and unplanned reporting domains during survey design. It is a global norm that National Statistical Offices depends on surveys to respond to the needs of stakeholders, however survey cost remains a challenge during survey design, estimation, and reporting stages. The paper reviewed literature relating to statistical approaches that may be used to provide estimates at lower (unplanned or planned) reporting domains using statistical models, data fragmentation and other methods that relates to big data. Several techniques such as Small Area Estimation⁰, have been outlined in various literature, however most national statistical offices (NSOs) do not rely on modelled data to respond to user-requests. The base of discussion for this paper focuses on models such as area level models (Fay-Herriot model, model with Correlated Sampling Errors, Time Series and Cross-Sectional Models and Spatial Models), unit level models (Multivariate Nested Error Regression Model¹, Random Error Variance Linear Model², Two level Model³, and Logistic Regression Models⁴). This exploration is paramount to any NSO to evaluate the most suitable models that can be used for unplanned reporting domains. The paper will provide a summary of the recommendations of feasible models that can be utilized efficiently in the production environment.

Key words: Small area estimation⁰, Fay-Herriot model, Correlated Sampling Errors, Time series and Cross-sectional models, Spatial Models, Unit Level Models, Random Error Variance Linear Model², Two Level Model³, Logistic Regression Models⁴.

CPS Abstract

An in-depth assessment of the impact of metrology on the economy: recommendations for official statistics systems.

Author: Dr Monika Rozkrut

Coauthors: Malgorzata Tarczynska-Luniewska

Submission ID: 2091

Reference Number: 2091

Brief Description

The article aims to present a proposal for advanced modelling of the impact of metrology on the economy, integrating the macroeconomic and microeconomic approaches. The article proposes filling information gaps for in-depth modelling and provides recommendations for improving the official statistics on science, technology, and innovation to facilitate impact assessments. The integration of official statistics and data obtained from a supplementary survey enabled novel analyses that prove that metrology acts as a catalyst for innovation.

Abstract

The socio-economic benefits of the functioning of quality infrastructure, including its essential element, i.e. metrology, are the result of a wide range of impact mechanisms and concern areas such as technology diffusion, increase in innovation, reduction of market imperfections, reduction of transaction costs, increase in productivity, building competitive advantage, or increasing economic efficiency. The impact of metrology on the socio-economic environment has already been discussed in the literature, but only a few studies address this issue. This problem is related to the unavailability of appropriate data for in-depth analysis of complex problems such as measuring impact, results, and outcomes. As a result, most attempts are limited to simple macroeconomic approaches.

The article aims to present a proposal for advanced modelling of the impact of metrology on the economy, integrating the macroeconomic approach with in-depth microeconomic modelling based on the use of data provided by official statistics. The article proposes filling information gaps for applying in-depth modelling. The analysis conducted indicates that the official statistics system plays a crucial role in addressing information gaps. However, this requires an appropriate scope of statistical research. Most official statistics systems conduct research on science, technology, and innovation, which is necessary for providing comprehensive information in this area. The authors have identified the required range of information and have highlighted the research issues that require further supplementation. The research conducted has revealed that by slightly expanding the scope of existing research, it is possible to analyse the impact of metrology on the economy. The article provides recommendations for improving the official statistics on science, technology, and innovation to facilitate deeper analyses and impact assessments.

Next, the authors present the impact analyses that were made possible thanks to the use of official statistics data and data obtained from a supplementary survey of enterprises. The integration of these sources enabled novel and extended analyses, which showed that the industry sectors that use metrology most intensively are the most productive. This research proves that metrology acts as a catalyst for innovation.

CPS Abstract

Analysis of Correlated Outcomes of Anthropometric Measurements for Under-Five Children in Tanzania

Author: Mr Edgar Pallangyo

Submission ID: 1800

Reference Number: 1800

Files/Uploads

[Fig conceptual](#)
[Fig2 Prevalence](#)
[Table 1](#)
[Table 2](#)
[Table 3](#)
[Table 4](#)

Brief Description

The title "Analysis of correlated outcomes of anthropometric measurements for under five children in Tanzania" appears to be a specific research topic related to the field of statistics and data science.

It likely focuses on applying statistical methods to understand the relationships between different anthropometric measurements in young children in Tanzania.

In the context of foundational work in statistics and data science, this research can contribute to "Better understanding and innovation in statistical research and decision making" by providing practical insights into how statistical analysis can be applied to address specific public health issues in the context of child nutrition and growth.

It showcases the real-world application of statistical methods, which is essential for advancing the field and making data-driven decisions.

Abstract

Malnutrition is a leading cause of morbidity and mortality among children under the age of five worldwide. In Africa, the prevalence of wasting is 6.4 percent, which is lower than the global average of 6.9 percent and there has been a slow progress to meeting global nutrition goals, and children under the age of five are at the highest risk of all types of malnutrition, with 40% stunting and 27% wasting (WHO, 2020). However, The study aimed at applying Multivariate Generalized Linear Mixed Models to examine factors associated with correlation outcomes, in particular, anthropometric measurements among under-five children in Tanzania. Three anthropometric measurements include Weight for Age, Height for Age and Weight for Height among under-five children in Tanzania were jointly modelled to identify common factors associated with childhood malnutrition. A total of 9,052 children with valid measures of Height and Weight were processed and analyzed. The results indicate that Weight for Age was correlated with Height for Age ($p\text{-value} < 2e-16$) and Weight for Height ($p\text{-value} < 2e-16$). The Multivariate Ordered Logit Model has lower AIC=53213.92 and BIC=52727.95, indicating better model fit than the Multivariate Ordered Probit Model. In Tanzania, the age of the child, birth order, mother education level, child gender, mother working status, wealth index, marital status, and mother body mass index are important determinants of malnutrition among children under the age of five. Moreover, the common factors were child's age, Birth order, Mother's education attainment, child's sex, Mother working status, wealth index, Marital status and Mother's Body Mass Index. As a result, emphasis should be placed on analyzing correlated health outcomes in order to draw conclusions about the factors that may have a mutual effect on anthropometric measurements.

CPS Abstract

Analysis of Genotype performance in Grain yield of Sorghum via Generalized Additive Mixed Model**Author:** Mr Mulugeta Tesfa**Submission ID:** 1991**Reference Number:** 1991**Files/Uploads**[Analysis of Genotype performance in Grain yield of Sorghum via Generalized Additive Mixed Model](#)**Brief Description**

The paper identified the application of Generalized Additive mixed model on genotype selection of sorghum, which shows the non-linear dependence between grain yield and yield related traits of sorghum as well as select the best genotype performing better under stressed conditions and non-stress conditions.

Abstract

Sorghum is the leading cereal crop worldwide and the top five cereal crops in production annually. Yields of sorghum are affected by different factors from which genotype and genotype by environmental interaction are the main sources of variability of yield. The study aims to evaluate genotype performance and examine the nonlinear dependence of grain yield and yield-related traits of sorghum using a generalized additive mixed model and also identify the impact of treatment on grain yield. The data was collected using the lattice square design of the experiment to characterize the yield and yield-related traits. The study considered grain yield as a response and incorporated yield-related traits as fixed, genotype, genotype by treatment and replication within the treatment as random effects and smooth terms for the analysis. The analysis was performed through GAMM. The result shows that there was a significant difference in grain yield among the treatments and that most of the variability of the grain yield (more than 84%) was explained by genotype, genotype by treatment interaction whereas the variability of grain yield (less than 2%) due to smooth term and replication within treatment was negligible. The study found that grain yield and plant height had significant non-linear dependence and also grain yield and panicle width was non-linearly associated where grain yield and other continuous yield-related traits were linearly associated. The relationship between grain yield and the independent variables was non-linear. Genotypes that are recommended for mass production in a cultivate are G196, G149 and G187 respectively in order of performance, Genotypes which enabled to perform better under irrigated condition are G145, G149 and G142 whereas genotype that have better performance under stress condition are G196, G22 and G19 in order of performance. In general, the study recommends using GAMM for the genotype selection of crops and identifying linear or non-linear associations between grain yield and continuous covariates of yield related traits in crop production. The future work will focus on multivariate analysis to perform genotype selection for crop production.

CPS Abstract

Application of Ai to Bridge the Teacher to Pupil Ratio in Uganda

Author: Mr Alban Manishimwe

Submission ID: 2081

Reference Number: 2081

Brief Description

This session will showcase opportunities of application of Artificial Intelligence(AI) technology to policy making and decision processes encountered by education leaders to allocate teachers in various schools in Uganda. The outcomes of this session will be a novel approach of using AI as a policy maker , hence overcoming & gaps that exist therein, the teacher to ratio in Uganda.

Abstract

ABSTRACT

This paper explores the potential of Artificial Intelligence (AI) models in addressing the high teacher-to-pupil ratio in Uganda's education system. The study suggest that AI algorithms can be leveraged to improve the quality of education by reducing the teacher-to-pupil ratio, thereby enhancing the individual attention students receive.

The research focuses on the application of prediction algorithms, which has the potential to re-allocate teachers to schools based on need, quality and scarcity, rather than the existing non-automated system which eventually leads to high teacher to pupil ration in Uganda. This is achieved by integrating AI models into distribution and allocation of teachers nationwide.

The study also emphasizes the importance of data privacy when using AI tools in the educational sector. It underscores the need for institutions to ensure that the use of AI complies with data protection legislation and existing data privacy policies.

The model suggests that the acceptance of AI as a policymaker can be influenced by factors such as the AI's ability to adapt to the pupils' and teachers' demography, and expertise.

The paper further discusses the potential of AI to work as a partner with policy and decision makers in distribution and allocation of teachers. It also highlights the need for policy makers to become familiar with the use of AI tools and to incorporate them as an integral part of their day-to-day work.

The study concludes by suggesting that AI technology can become a policy maker's partner or an administrative assistant, and it can allocate and re-distribute new and existing labor while policy makers maintain responsibility for planning, communication, and coordination.

In summary, the paper argues that AI models can significantly bridge the teacher-to-pupil ratio, thereby enhancing the individual attention pupils receive. However, it also emphasizes the need for careful consideration of data privacy issues and the importance of teacher acceptance of AI-based policy making.

CPS Abstract

Assessing Multidimensional Poverty Measures in Sri Lanka on an innovative approach to tracking progress towards the SDGs

Author: Dr Diana Dilshanie Deepawansa

Submission ID: 1886

Reference Number: 1886

Files/Uploads

[AOS2024_Abstract_DDDDeepawansa_2024](#)

Brief Description

Poverty is a complex social issue affecting millions globally. Poverty is more than the lack of wealth but the circle of deprivations due to limited opportunities and social exclusions influenced by various factors. Like many countries, Sri Lanka adopted to multidimensional approach to measure poverty as complementary to the traditional methods focusing solely on wealth metrics. The innovative approach employed in this research provides more accurate and up-to-date poverty in recognizing the strengths and limitations of existing methods to achieve SDGs on poverty effectively.

Abstract

Assessing Multidimensional Poverty Measures in Sri Lanka on an innovative approach to tracking progress towards the SDGs

Abstract

Poverty is a multifaceted and prevalent social issue that affects millions of people worldwide. Poverty is not merely a lack of wealth but a cycle of disadvantages with limited opportunities and social exclusion influenced by various factors. As in many countries, Sri Lanka has been emphasizing measuring poverty through a multidimensional approach. The traditional tools used to measure poverty often focus solely on wealth and conditions set in the past to measure current status. The innovative approach called the Synthesis method applied for this analysis leverages more accurate and up-to-date poverty measures in a multidimensional approach. This study aims to recognize the accuracy of poverty measures to recognize the respective strengths and limitations of existing methods. To achieve these objectives the application of the method was employed to the Household Income and Expenditure Survey data-2019 conducted by the Department of Census and Statistics (DCS) in Sri Lanka. This analysis includes three main dimensions; education, health and living standards with ten comprehensive indicators. The analysis shows the national multidimensional poverty headcount at 32.8 percent with an average deprivation of 45.6 percent among the poor. Comparatively, the official multidimensional poverty headcount of 16 per cent and average deprivation of 41.6 per cent experienced by poor persons published by the DCS shows a difference emerge. Further, the official monetary poverty headcount of 14 per cent paints a less comprehensive picture of poverty. This comparison highlighted the significance of the innovative approach of multidimensional poverty measures providing a more comprehensive picture of holistic measures to address poverty effectively. This innovative lens equips policymakers to monitor and achieve Sustainable Development Goals by assessing and tracking poverty progressively.

Keywords: Multidimensional approach, Fuzzy set, Poverty, Deprivation
JEL Classification C38 C43, I31, I32

CPS Abstract

Assessing desertification and vegetation loss in Natural Protected Areas of northern Mexico

Author: Mr Ranyart Rodrigo Suarez Ponce de Leon

Submission ID: 1938

Reference Number: 1938

Brief Description

One of the impacts triggered by climate change is the alteration of soil composition. Fluctuations in temperatures, whether rising or falling, can influence the landscape in a way that results in changes to the type and quantity of vegetation. This, in turn, can lead to vegetation loss and eventual desertification of the environment. In this study, we examine the changes undergone in soil coverage in 11 control zones corresponding to Natural Protected Areas in northern Mexico to analyze landscape desertification over the past 30 years. To conduct this analysis, we employed the libraries of Google Earth Engine due to their easy access to cloud data and computational capabilities. The Trends-Earth library was used to measure changes in soil type. Additionally, in-situ data from meteorological stations were incorporated to identify changes in the historical trend in accumulated precipitation over time. Results reveal transitions in soil type, indicating shifts from higher to lower vegetation coverage in some of the analyzed areas and also a trend of cumulative precipitation loss over the same period of time.

This study was carried out in 11 protected natural areas, governed by special legislation designed to minimize human impact. This unique circumstance provides an excellent opportunity for the observed changes in the control zones to more accurately reflect the influence of climate change rather than human activities.

Abstract

One of the impacts triggered by climate change is the alteration of soil composition. Fluctuations in temperatures, whether rising or falling, can influence the landscape in a way that results in changes to the type and quantity of vegetation. This, in turn, can lead to vegetation loss and eventual desertification of the environment. In this study, we examine the changes undergone in soil coverage in 11 control zones corresponding to Natural Protected Areas in northern Mexico to analyze landscape desertification over the past 30 years. To conduct this analysis, we employed the libraries of Google Earth Engine due to their easy access to cloud data and computational capabilities. The Trends-Earth library was used to measure changes in soil type. Additionally, in-situ data from meteorological stations were incorporated to identify changes in the historical trend in accumulated precipitation over time. Results reveal transitions in soil type, indicating shifts from higher to lower vegetation coverage in some of the analyzed areas and also a trend of cumulative precipitation loss over the same period of time.

CPS Abstract

Beyond Compliance: The Impact of Data Protection on Statistical Excellence - A View from the 2021 Population-Housing Census in Greece

Author: Mr Apostolos Kaspis

Submission ID: 1981

Reference Number: 1981

Brief Description

This session, part of Strand 5 - "Living the Fundamental Principles of Official Statistics: Experiences from NSOs," delves into the intricate relationship between data protection and the fundamental principles governing official statistics. Focused on the 2021 Hellenic Population and Housing Census, the presentation will offer a comprehensive examination of how adherence to data protection principles is instrumental in shaping and advancing best practices in statistical processes. Attendees can expect insights into successful cases where the Fundamental Principles are seamlessly integrated into the daily work of National Statistical Offices (NSOs), providing a valuable exploration of real-world applications.

Abstract

In this paper we attempt to delve into the implications of personal data legislation on Censuses. The primary goal is to highlight the most effective ways to integrate this legislation into the core fabric of the fundamental principles of official statistics that predominantly govern Censuses. For a practical illustration, we utilize the Greek Population and Housing Census of 2021 as a representative example. More specifically, the paper addresses the personal data related issues that arose during the 2021 Census in Greece and the approaches adopted by the Hellenic Statistical Authority (ELSTAT) within the scope of the General Data Protection Regulation (GDPR) of the European Union (EU). The analysis covers a range of aspects that include the particular types of data processed, the specific processing activities across the different phases of the Census and the utilization of data processors. Furthermore, we provide an account of the purposes of use and the legal bases upon which National Statistical Offices (NSOs) can rely on and explore the intricate link between the prevailing conditions under which the Census was deployed and the various data protection considerations that arose. We then elucidate how the Census was built on the 'Privacy by Design' principle and how this approach was crafted taking into account existing operational limitations. Finally, our findings highlight how this strategy not only adhered to but also enhanced best practices that strengthen and complement the fundamental principles of official statistics.

CPS Abstract

Building a National Data Infrastructure for Malaysia's Future: A Focus on Data Production and Innovation

Author: Ms MALATHI PONNUSAMY

Submission ID: 1988

Reference Number: 1988

Brief Description

The aim of this paper is to highlight the importance of developing a robust national data infrastructure in Malaysia to support the country's future growth and innovation. By focusing on data production and innovation, Malaysia can effectively harness the potential of data-driven technologies and capitalise on new opportunities. This study highlights the Department of Statistics Malaysia (DOSM)'s initiatives to strengthen and institutionalize modern practices in producing and disseminating reliable, accurate, and timely statistical information. DOSM examines the innovation phase with a particular focus on incorporating sophisticated methodologies such as leveraging the Price Intelligence (PI) module to produce innovative price statistics, and examining social media data for the purpose of evaluating the Public Maturity of Official Statistics (PMAOS). The PMAOS platform assesses the public's perception and response to the dissemination of DOSM products, provides key insights for strategic planning and improvement. DOSM's commitment to promoting transparency and accessibility is exemplified through the use of platforms such as OpenDOSM, StatsDW, and eStatistics. These platforms facilitate the accessibility and availability of statistics, enabling the public, policymakers, and the global community to make informed decisions in a more effective manner. OpenDOSM functions as an up-to-date open data platform that aims to promote transparency, engagement, and innovation leading to data centric citizens.

Abstract

Abstract: The aim of this paper is to highlight the importance of developing a robust national data infrastructure in Malaysia to support the country's future growth and innovation. By focusing on data production and innovation, Malaysia can effectively harness the potential of data-driven technologies and capitalise on new opportunities. This study highlights the Department of Statistics Malaysia (DOSM)'s initiatives to strengthen and institutionalize modern practices in producing and disseminating reliable, accurate, and timely statistical information. DOSM examines the innovation phase with a particular focus on incorporating sophisticated methodologies such as leveraging the Price Intelligence (PI) module to produce innovative price statistics, and examining social media data for the purpose of evaluating the Public Maturity of Official Statistics (PMAOS). The PMAOS platform assesses the public's perception and response to the dissemination of DOSM products, provides key insights for strategic planning and improvement. DOSM's commitment to promoting transparency and accessibility is exemplified through the use of platforms such as OpenDOSM, StatsDW, and eStatistics. These platforms facilitate the accessibility and availability of statistics, enabling the public, policymakers, and the global community to make informed decisions in a more effective manner. OpenDOSM functions as an up-to-date open data platform that aims to promote transparency, engagement, and innovation leading to data centric citizens.

CPS Abstract

CAUSAL IMPACT OF ARMED CONFLICT ON PRIMARY EDUCATION IN MALI

Author: Mr Franck MIGONE

Submission ID: 1810

Reference Number: 1810

Brief Description

This session will be led by Franck MIGONE (Statistician) and Josias Sadia (Student, Undergraduate).

We will present the main results of a paper which analyzes the causal effect of conflict in a west African country Mali on Education for primary school student.

Abstract

The aim of this article is to assess the impact of armed conflict on the number of years of education of primary school pupils in Mali. To do this, we formed two cohorts from the Enquete Harmonisée sur les Conditions de Vie des Ménages (EHCVM), adopting an identification strategy similar to that of Dabalen and Paul (2012). A first cohort (12-22years) or young cohort is composed of all children likely to be directly affected by the conflict (being in primary school over the period 2012-2019) and a second cohort (23-32years) composed of those no longer in primary school at the time of the conflict. We then merged this database with the conflict database (ACLED) to identify affected individuals on the basis of the occurrence of a conflict within the department in which they reside. By adopting various identification strategies (binary or continuous treatment) and evaluation modes (Difference in Difference, Propensity Score Matching) and taking individual heterogeneities into account, it emerges that the affected cohort has fewer years of education than the unaffected cohort. Affected individuals have fewer years of education than their unaffected peers, and this conflict has more pronounced impacts in rural areas (value) and among girls (value). The main mechanisms highlighted in this study are the deterioration of the learning environment (insecurity and absenteeism) and the erosion of household purchasing power. The main findings tend to sound the alarm about the need for policies to (i) reduce the intensity of conflict, (ii) set up training programs to enable children to continue schooling even in conflict zones, (iii) set up intensive schooling programs in the worst-affected areas to reduce disparities within departments, place of residence (rural and urban) and by gender (girls and boys).

CPS Abstract

Challenges and innovations in data production. Graduates monitoring - an innovative statistical product for decision-making processes in education.

Author: Baiba Zukula

Submission ID: 1885

Reference Number: 1885

Brief Description

Quality education is one of the basic conditions for the development of the individual and the country. Data-driven decisions are needed, so that students can make informed choices about what to study, or government to develop educational programs and policies.

The alignment of graduates with the labour market is crucial for the socio-economic advancement of the country. The government must enhance the quality of higher and professional education, hence it is imperative to conduct regular assessments of the performance of educational programs implemented by institutions. Surveys can be used to assess the status of graduates, though more consistent and timely data can also be derived from administrative data sources through data matching.

Abstract

Quality education is one of the basic conditions for the development of the individual and the country. The purpose of education in its broadest sense is to prepare an individual for life, which includes general knowledge and skills, the individual's ability to engage in social and political processes, as well as professional training - the ability to participate in economic activities as an employer or employee*. Data-driven decisions are needed, so that students can make informed choices about what to study, or government to develop educational programs and policies. The alignment of graduates with the labour market is crucial for the socio-economic advancement of the country. The government must enhance the quality of higher and professional education, hence it is imperative to conduct regular assessments of the performance of educational programs implemented by institutions. Surveys can be used to assess the status of graduates, though more consistent and timely data can also be derived from administrative data sources through data matching.

In Latvia, comprehensive and high-quality administrative data encompassing information on graduates of educational institutions, are available. Use of a single personal identifier facilitates matching the data on graduates with their income, employment, including changes over time. Graduates monitoring - the database includes data for each year's graduates, starting from 2017, providing comprehensive account of their employment and income for ten years period after graduation.

In partnership with the Ministry of Education and Science, Central Statistical Bureau of Latvia has actively pursued the possibilities for data linking. This endeavour aims to yield an innovative statistical product tailored to the specific requirements of decision-making processes. Monitoring data are extensively used both, at policy level and by universities for their strategic planning and assessment purposes. Summary data is published as open data, and the CSB provides also access to anonymized data for researchers.

*Gingell, J., Winch, C. (2004). *Philosophy and Educational Policy*. London: Routledge.

CPS Abstract

Classification using probabilistic models and other methods for mixed data

Author: GONZALO PEREZ DE LA CRUZ

Submission ID: 1969

Reference Number: 1969

Files/Uploads

[Perez&Eslava_Classification_IAOS-ISI_30-11-2023](#)

Brief Description

In this work we present comparative results of some classification methods for the case of two populations and a set of binary and continuous variables. Methods used include classical statistical methods like discriminant analysis and logistic regression, as well as discriminant analysis based on probabilistic graphical models; algorithmic methods like support vector machines and random forests; and Deep neural networks. The performance of these methods is compared in terms of classification error rates on both simulated and real data. The results in terms of error rates showed that i) additive models or linear methods performed poorly, ii) models with interactions and selection of variables or non-linear methods performed better than the linear ones, and iii) Deep neural networks did perform as well as the non-linear methods though they did not show the best performance, this might be due to the small-medium size of the tabular data.

Abstract

Classification of labelled observations into two or more well-defined classes based on a set of measurements is a problem that has been developed within Statistics. Methodology for classification is continuously being developed within the area of Statistical Learning. The problem has also been considered for a long time in the area of Artificial Intelligence where recently Deep learning methods are becoming computationally more feasible and have achieved important results on classification or prediction problems. Probabilistic models for Deep learning are also potentially useful and much research activity is in progress, these models combine probability distributions with mathematical concepts of graphs and are known as graphical models, when the graphs are directed they are known as Bayesian networks and when undirected as Markov networks.

When all measurements, features or predictors, can be represented as continuous variables, most known methods apply, but when the variables are both categorical and continuous the performance of the methods might diminish predictivity or have errors with less ability to generalize.

In this work we present comparative results of some classification methods for the case of two populations and a set of binary and continuous variables. Methods used include classical statistical methods like discriminant analysis and logistic regression, as well as discriminant analysis based on probabilistic graphical models; algorithmic methods like support vector machines and random forests; and Deep neural networks. The performance of these methods is compared in terms of classification error rates on both simulated and real data.

The results in terms of error rates showed that i) additive models or linear methods performed poorly, ii) models with interactions and selection of variables or non-linear methods performed better than the linear ones, and iii) Deep neural networks did perform as well as the non-linear methods though they did not show the best performance, this might be due to the small-medium size of the tabular data.

CPS Abstract

Computer Assisted Web interviews on mobile phones

Author: Norberts Talers

Submission ID: 2061

Reference Number: 2061

Brief Description

The presentation topic will be Computer Assisted Web Interviews on mobile phones (CAWI mobile) at Central Statistical Bureau of Latvia. CSB of Latvia uses electronic data collection technology via web forms since 2002 for legal persons and since 2010 for natural persons. This presentation will focus on surveys that are intended for collecting data from natural persons. Although electronic data collection is not new for CSB of Latvia the massive increase of usage of mobile devices has led to the point where a lot of respondents are receiving and reading emails sent to them stating that they have become respondents on mobile devices. These letters contain a hyperlink to the data collection system and naturally the link would firstly be opened on a mobile device just to understand that the system is not suitable for working on such devices and a personal computer should be used instead. The project consisted of several parts. The first part was an internal project where subject matter experts from social statistics domain, data collection experts and IT experts were brought together to elaborate the set of requirements on what transformations must be made from classical CAWI surveys system where a computer must be used to a solution where the surveys can be filled in on a mobile device. The next part was tendering an outsource company to develop the solution. And the last part was the implementation of the solution in production. The presentation will cover the whole process.

Abstract

The presentation topic will be Computer Assisted Web Interviews on mobile phones (CAWI mobile) at Central Statistical Bureau of Latvia. CSB of Latvia uses electronic data collection technology via web forms since 2002 for legal persons and since 2010 for natural persons. This presentation will focus on surveys that are intended for collecting data from natural persons. Although electronic data collection is not new for CSB of Latvia the massive increase of usage of mobile devices has led to the point where a lot of respondents are receiving and reading emails sent to them stating that they have become respondents on mobile devices. These letters contain a hyperlink to the data collection system and naturally the link would firstly be opened on a mobile device just to understand that the system is not suitable for working on such devices and a personal computer should be used instead. The project consisted of several parts. The first part was an internal project where subject matter experts from social statistics domain, data collection experts and IT experts were brought together to elaborate the set of requirements on what transformations must be made from classical CAWI surveys system where a computer must be used to a solution where the surveys can be filled in on a mobile device. The next part was tendering an outsource company to develop the solution. And the last part was the implementation of the solution in production. The presentation will cover the whole process.

CPS Abstract

Contribution of sociodemographic and economic determinants to the gendered division of unpaid work and trade-offs between unpaid work and employment

Author: Mrs khady BA

Submission ID: 1808

Reference Number: 1808

Files/Uploads

[Proposal for IAOS-ISI conference mexico_khady BA](#)

Brief Description

As "Achieving gender equality and empowering all women and girls" (SDG 5) is one of the top priorities of the current decade, particularly in Africa, the main objective of this study is to compare the contribution of men and women to the country's economy and the costs of opportunities. The results allowed as to analyze the role of sociodemographic and economic determinants of the Senegalese population to the gendered division of unpaid work.

Abstract

Unpaid work is seen as a barrier to women's economic empowerment. The objective of this research is to study the contribution of the sociodemographic and economic determinants of the Senegalese population to the gendered division of unpaid work but also to value it in order to have an overview of its weight on the Senegalese economy.

The analysis of Senegal Time Use Survey (ENETS) results by gender highlighted gender disparities in unpaid labour participation between men and women in Senegal. Indeed, 63.1% of men against 11.9% of women do not participate in unpaid activities. In terms of time, women perform an average of 4.3 hours of unpaid work per day compared to 36 minutes for men.

It is also noted that from one group to another, participation times could be impacted by some socio-demographic and economic characteristics. For men, income levels and the presence of biological children in the household did not influence participation in unpaid activities, while the result was the opposite for women. Analyzing from the married perspective only, the results showed that all variables influenced participation in unpaid activities except place of residence. Also, women were 17.4 times more likely to engage in unpaid work than men.

By valuing the working time not paid by the minimum wage (SMIC), we obtain a weight corresponding to 27.2% of Gross Domestic Product (GDP) in 2021, of which 24.5% is produced by women. This indicates the importance of considering unpaid work.

CPS Abstract

Democracy Dies in Darkness without Official Statistics

Author: Dr Luca Di Gennaro

Submission ID: 1897

Reference Number: 1897

Brief Description

I am finishing an article on "Democracy and Official Statistics" with historical and quantitative insights: Democracy dies in darkness without official statistics. I cited Keynes, Schumpeter and Popper, among others. And I use data from the Statistical Performance Indicators (World Bank) and the Democracy Index (Economist).

Abstract

Is there a quantitative relationship between democracy and official statistics? What are the National Statistics Offices (NSOs)? Why does the NSO exist? What is its role? Why do the NSO disseminate their statistics free of charge? How can a democratic country organize an election without official statistics? How close are the NSO staff close to politicians? What qualifications do NSO staff possess? In a world increasingly shaped by data, the importance of official statistics often remains in the shadows, even within NSOs. Yet, these statistics (like employment rates, gross domestic product and pandemic data) are the lifeblood of democracy, influencing policymaking, media narratives and electoral choices. In this work, the vital nexus between official statistics and democracy is highlighted by quantitative data and historical prospective. The correlation between democracy and official statistics is high, and there is good country performance of statistical systems that are more likely in a democratic country. I found out that the correlation between democracy and official statistics is seventy per cent. NSOs' responsibility should be high and abuse should not be tolerated. This paper not only underscores the global significance of official statistics, but also aligns with the United Nations Fundamental Principles of Official Statistics and the Sustainable Development Goals.

CPS Abstract

Driving innovative and ambitious statistical change with user needs at the core

Author: Ms Fiona Dawe

Submission ID: 1935

Reference Number: 1935

Brief Description

The advance in accessibility, volume and quality of digital administrative data offers exciting opportunities to explore the potential to enhance and replace traditional population statistical methodology. Since the 2011 Census the UK Office for National Statistics has been actively exploring the viability of moving from a census based to administrative based method for producing more timely and frequent population estimates. A proof of concept for the statistical methodology to underpin an administrative data centred design has been demonstrated alongside a suite of published research and evidence.

For changes to any statistical system to be successfully adopted user needs must be at the centre. Without understanding the needs of data users the development of new innovative statistical outputs will fail to deliver evidence relevant to impactful decision making. Without the support from data users of the opportunities that fundamental changes to statistical methodology offer, the development of innovative solutions is threatened.

In 2023 ONS conducted a large scale public consultation to discuss the proposed new administrative data centred approach for population statistics with users to understand whether their current and future data needs would be met. The current census-based system is highly regarded and extensively used by decision-makers across a diverse user base: central and local government, businesses, third sector and academia. Using the experience of the public consultation and subsequent delivery of the Recommendations, Fiona Dawe will explore the importance of understanding user needs as both (a) a fundamental prerequisite to guide research plans, and (b) the perception and environment in which new and innovative solutions may land and then be managed.

CPS Abstract

Effectiveness of selected methods of data integration in tourism statistics

Author: Prof. Marek Cierpial-Wolan

Submission ID: 2066

Reference Number: 2066

Brief Description

Foundational work in statistics and data science: Better understanding and innovation in statistical research and decision-making

Abstract

Effectiveness of selected methods of data integration in tourism statistics

Marek Cierpial-Wolan, Assoc. Prof., Statistics Poland, University of Rzeszow
Dominik Rozkrut, PhD, Statistics Poland, University of Szczecin

The modern world is being determined by many threats, both globally and locally. World economies are facing growing problems such as: instability caused by numerous armed conflicts, the unprecedented scale of global migration, and since 2020 the world is facing a coronavirus pandemic. All these circumstances cause various types of effects of a socio-economic nature, and in particular significantly affect the tourism industry. The consequence of these events is the emergence of huge information needs regarding tourism. This is related to both the specifics of the tourism market (short-term and sometimes incidental business) and activity in the gray zone. The demand for high-quality information, available in real mode, requires the integration of data from various sources, i.e. administrative records, databases created on the basis of full and sample surveys, and especially big data. Thus, the development of innovative methods of data integration is practically becoming an imperative for academia and institutional statistics today. This article evaluates the usefulness of the following probabilistic methods of data linkage and deduplication: Natural Language Processing (NLP), Machine learning algorithm (K Nearest Neighbors (K-NN) using TF-IDF techniques) and Fuzzy matching. The paper used these methods to combine data from web scraping of booking portals (Booking.com, Hotels.com and AirBnB.com) with tourism survey frame. The effective use of multiple portals was made possible not only by text algorithms but also by photo comparison algorithms such as comparing the similarity of color histograms, comparing visuals using feature descriptors and digital fingerprints. This method is based on analyzing the visual features of the images that are assigned to each offer. The paper evaluates the quality of the tourism statistics frame through the acquisition of data from web scraping of tourism portals, and how this affected the results of monthly surveys of accommodation facilities. The conducted surveys showed that the most useful among the tested methods was Fuzzy matching based on Levenshtein's algorithm combined with Vincenty's formula. In addition, as a result of the integration, it was possible to significantly improve the quality of the tourism survey frame (an increase in the number of new facilities by 1.1%) and thus correct the results of the survey of the supply side of tourism in Poland.

Keywords: probabilistic record linkage, web scraping, comparing visuals

JEL: C1, C81, Z32

CPS Abstract

Engaging young minds: Analyzing the importance of bringing official statistics to children

Author: Dr Alfredo Mendoza-Gonzalez

Submission ID: 1875

Reference Number: 1875

Brief Description

Today, children increasingly show a real interest in what is happening in their environment, including social, economic, and environmental problems. This early awareness allows them to take an active interest in understanding the challenges facing society, of which they consider themselves active members. This quality not only provides them with valuable knowledge but also fosters their sense of social responsibility, empathy towards their environment and the need to take action to promote change.

Countries are including the opinions of young citizens in decision-making. Sessions, forums, consultations, and assemblies are means through which young citizens exercise their representativeness before official decision-making bodies. Over the years, the door has been opened for them to be heard.

National Statistical Offices can generate in children a more complete and balanced perspective of the world, allowing them to develop greater empathy and understanding towards everyone's experiences and challenges through knowledge of statistical information with a clear, contextualized design, depending on their development. Through the internet and digital platforms, knowledge is now easily accessible, fostering accessibility and inclusivity, and empowerment across diverse populations. National statistical offices therefore have the capacity to provide the tools and data that these knowledge-hungry young citizens can consume.

This research collects the experiences of countries that have involved future decision makers, the techniques and means used to present information to be consumed by them, and carries out an analysis of the advances, points of opportunity and new perspectives that revolve around this matter.

CPS Abstract

Enhancing Data Integration and Utilization: An Open Source Platform for Consolidating Statistical Office Outputs

Author: Dr Elio Villaseñor

Submission ID: 1943

Reference Number: 1943

Brief Description

This paper presents a case study of the collaborative efforts between the National Institute of Statistics and Geography (INEGI) in Mexico and the National Institute of Statistics (INE) in Chile in integrating data science methodologies into their public statistical systems. The focus is on the synergistic relationship developed between these institutions to enhance their statistical capacities using big data, machine learning, and artificial intelligence. The study outlines the joint initiatives, including the development of shared data science labs and cross-border data projects, aimed at improving data collection, analysis, and dissemination. The challenges encountered during this collaboration, such as harmonizing methodologies across different statistical frameworks and ensuring data security, are explored in depth. Additionally, the study examines the impact of this collaboration on the quality of public statistics and its implications for policy-making. This paper underscores the value of international cooperation in the realm of data science for the evolution of national statistical systems, offering insights that are pertinent for statistical offices worldwide seeking to modernize their approaches through collaborative frameworks.

Abstract

This research paper presents a comprehensive case study of the collaboration between the National Institute of Statistics and Geography (INEGI) in Mexico and the National Institute of Statistics (INE) in Chile. This collaboration is at the forefront of integrating advanced data science methodologies into public statistical systems. It focuses on the synergistic relationship developed between these institutions to augment their statistical capacities using big data, machine learning, and artificial intelligence.

A central element of this collaboration is INEGI's development of a sophisticated technological platform, featuring a versatile data lake capable of accommodating a wide range of statistical and geographic digital data formats. This platform is pivotal in facilitating data collection, secure storage, and controlled access. It allows for the application of unconventional data analysis techniques and the development of innovative data product prototypes.

Key features of INEGI's technological platform include:

Workflows for systematic gathering, storage, processing, analyzing, and presenting information from both internal and external data repositories. Models for statistical and geographic information, maintaining data lineage and integrity for seamless analysis and exploitation based on specific needs.

Tools that support visual representations of analysis results, enhancing interpretation and utilization of the information.

An environment fostering ongoing development and collaborative efforts, enabling continuous evolution and improvement of the platform.

This system seamlessly integrates statistical and geographic information from various open-data sources. Data from these diverse sources are harmonized within a unified storage environment, adaptable for both temporary and permanent storage needs.

The paper details the joint initiatives undertaken by INEGI and INE, including shared data science laboratories and cross-border data projects. These initiatives aim to improve data collection, analysis, and dissemination, leveraging the capabilities of the technological platform.

Challenges such as harmonizing methodologies across different statistical frameworks and ensuring data security are explored in depth. The study also assesses the impact of this collaboration on public statistics quality and policy-making implications.

The value of international cooperation in the data science realm for evolving national statistical systems is underscored. The case study provides insights for global statistical offices looking to modernize through collaborative frameworks.

CPS Abstract

Enhancing the Quality of Online Self-Paced Training on the SEEA Using Large Language Models and Interactive Knowledge Graphs**Author:** Mr Issoufou Seidou Sanda**Submission ID:** 2075**Reference Number:** 2075**Brief Description**

The burgeoning field of online education has presented unique opportunities and challenges, particularly in resource-constrained regions like Africa. This paper delves into an innovative approach to bolstering online statistical training programs, addressing the quintessential issue of limited funding while ensuring high-quality training, with an application to an online training course on the System of Environmental-Economic Accounting. The crux of our research revolves around integrating cutting-edge large language models (LLMs), chatbots, and interactive knowledge graphs to simulate the dynamic and responsive experience of traditional tutor-led instruction. Our study is anchored in the context of Africa, where the demand for statistical education is on the rise. Still, the availability of resources and access to expert guidance remains constrained.

Our proposed methodology involves leveraging the capabilities of LLMs and language models to create an interactive, AI-driven educational environment. These models are designed to comprehend and respond to student queries, facilitate problem-solving, and provide personalized feedback, thereby emulating the nuanced interaction of a human tutor. Additionally, using interactive knowledge graphs in our system allows for a structured and visually engaging presentation of statistical concepts, enhancing the learning experience.

In this paper, we present proof of concept and discuss the efficacy of AI tutors in bridging the gap caused by the absence of human instructors in online, self-paced learning environments.

Furthermore, we explore the model's scalability, highlighting its potential to reduce the cost of online statistical training delivered by organizations in charge of statistical development in the continent with a minimal impact on quality.

Strand: Foundational work in statistics and data science

Abstract

The burgeoning field of online education has presented unique opportunities and challenges, particularly in resource-constrained regions like Africa. This paper delves into an innovative approach to bolstering online statistical training programs, addressing the quintessential issue of limited funding while ensuring high-quality training, with an application to an online training course on the System of Environmental-Economic Accounting. The crux of our research revolves around integrating cutting-edge large language models (LLMs), chatbots, and interactive knowledge graphs to simulate the dynamic and responsive experience of traditional tutor-led instruction. Our study is anchored in the context of Africa, where the demand for statistical education is on the rise. Still, the availability of resources and access to expert guidance remains constrained.

Our proposed methodology involves leveraging the capabilities of LLMs and language models to create an interactive, AI-driven educational environment. These models are designed to comprehend and respond to student queries, facilitate problem-solving, and provide personalized feedback, thereby emulating the nuanced interaction of a human tutor. Additionally, using interactive knowledge graphs in our system allows for a structured and visually engaging presentation of statistical concepts, enhancing the learning experience.

In this paper, we present proof of concept and discuss the efficacy of AI tutors in bridging the gap caused by the absence of human instructors in online, self-paced learning environments.

Furthermore, we explore the model's scalability, highlighting its potential to reduce the cost of online statistical training delivered by organizations in charge of statistical development in the continent with a minimal impact on quality.

Strand: Foundational work in statistics and data science

CPS Abstract

Enterprises' ESG practices and sustainable development: challenges and opportunities for official statistics.**Author:** Dr Monika Rozkrut**Coauthors:** Małgorzata Tarczynska-Luniewska**Submission ID:** 2093**Reference Number:** 2093**Brief Description**

Malgorzata Tarczynska-Luniewska, Monika Rozkrut The Sustainable Development Goals (SDGs) and Environmental, Social, and Corporate Governance (ESG) are interconnected frameworks that address similar aspects of sustainable development and responsible business conduct. Both frameworks, SDGs and ESG, share common goals and overlap in focus areas, however at different levels of aggregation. This overlap forms possible opportunity for official statistics, to deepening the scope of statistical production around sustainable development. The purpose of this study is to identify challenges and opportunities for official statistics.

Abstract

Official statistics play a crucial role in effective socio-economic policy-making. The adoption of policies related to the sustainable development has led to an adjustment and adaptation process for official statistics. The Sustainable Development Goals (SDGs) and Environmental, Social, and Corporate Governance (ESG) are interconnected frameworks that address similar aspects of sustainable development and responsible business conduct. The relationship between the SDGs and business is significant, and there is a growing recognition that businesses play a crucial role in achieving sustainable development. Both frameworks, SDGs and ESG, share common goals and overlap in focus areas, however at different levels of aggregation. While SDGs cover a wide range of social, economic, and environmental issues, aiming to address global challenges, ESG are criteria used by businesses and other stakeholders to evaluate a company's impact on society and the environment. As more and more companies try to align their strategies and operations with the SDGs to contribute to global sustainability efforts, ESG guides businesses in adopting responsible practices.

Investors and stakeholders increasingly evaluate companies based on their ESG performance, which can influence investment decisions and corporate reputation. This overlap, where on one hand, companies are encouraged to report on their contributions to the SDGs through reporting that helps to demonstrate their commitment to sustainable development, and on the other hand, companies disclose ESG information in their reports, enabling investors and stakeholders to assess their impact on the environment and society, forms possible opportunity for official statistics, to deepening the scope of statistical production in the area of sustainable development.

The purpose of this study is to identify challenges and opportunities for official statistics when considering jointly both the Sustainable Development Goals (SDGs) and Environmental, Social, and Governance (ESG) practices of enterprises. The analysis conducted revealed many areas where official statistics and sustainable development of enterprises and ESG practices intersect and require mutual functioning or synergy. This includes the necessity of transparent communication and standardized reporting of data with comparability and interoperability. Identifying these challenges is a crucial step towards developing good practices and effective tools for monitoring and assessing the sustainable development and ESG practices of enterprises. This also highlights the growing importance of and opportunities for official statistics in this field.

CPS Abstract

Estimation of Life-Space Assessment Score for Brazilian Elderly Population Combining Probability and Nonprobability Samples

Author: Dr Maria Luiza Guerra de Toledo

Submission ID: 1987

Reference Number: 1987

Brief Description

Social restriction measures due to pandemic caused substantial reduction in older adults' life-space mobility in Brazil. Social inequalities strongly affected vulnerable groups. To investigate the immediate impact of COVID-19 pandemic on life-space mobility of community-dwelling Brazilian older adults and examine the social determinants of health associated with life-space mobility, it was conducted an online and phone survey using a convenience snowball sample of participants 60 years and older ($n = 1,482$) living in 22 states in Brazil. It was used an adapted version of the Life-Space Assessment (LSA), and the goal was to investigate social determinants of health in LSA score. Preliminary analysis showed that the sample was biased when representing Brazilian elderly people in terms of sociodemographic characteristics: Compared to the target population, a higher proportion of women, white, from the Northeast region, singles and with higher income people appeared than would be expected. Corrected point and variance estimators for LSA measures were obtained with quasi-randomization methods. Pseudo selection probabilities were estimated by using the sample in combination with an external data set from a reference survey that covers the target population (a national household survey from IBGE – the official statistics agency in Brazil). Logistic regression was used to estimate the pseudo-inclusion probabilities for the nonprobability cases, using common covariates in both surveys. Some results were that Life-space mobility was lower among black individuals, those living alone and aged between 70 and 79. Other variables associated with life-space mobility, to a lesser extent, were sex, education, and income. Appropriate actions should be put in place to overcome the deterioration in life-space mobility amongst these groups.

Abstract

Social restriction measures due to pandemic caused substantial reduction in older adults' life-space mobility in Brazil. Social inequalities strongly affected vulnerable groups. To investigate the immediate impact of COVID-19 pandemic on life-space mobility of community-dwelling Brazilian older adults and examine the social determinants of health associated with life-space mobility, it was conducted an online and phone survey using a convenience snowball sample of participants 60 years and older ($n = 1,482$) living in 22 states in Brazil. It was used an adapted version of the Life-Space Assessment (LSA), and the goal was to investigate social determinants of health in LSA score. Preliminary analysis showed that the sample was biased when representing Brazilian elderly people in terms of sociodemographic characteristics: Compared to the target population, a higher proportion of women, white, from the Northeast region, singles and with higher income people appeared than would be expected. Corrected point and variance estimators for LSA measures were obtained with quasi-randomization methods. Pseudo selection probabilities were estimated by using the sample in combination with an external data set from a reference survey that covers the target population (a national household survey from IBGE – the official statistics agency in Brazil). Logistic regression was used to estimate the pseudo-inclusion probabilities for the nonprobability cases, using common covariates in both surveys. Some results were that Life-space mobility was lower among black individuals, those living alone and aged between 70 and 79. Other variables associated with life-space mobility, to a lesser extent, were sex, education, and income. Appropriate actions should be put in place to overcome the deterioration in life-space mobility amongst these groups.

CPS Abstract

Experience of Statistics Poland in the production of experimental statistics on road and maritime transport using innovative data sources

Author: Mr Michal Bis

Submission ID: 1952

Reference Number: 1952

Brief Description

The abstract is initially assigned to the submitted session under Strand 3: The new normal: Alternative data sources and new technologies, proposal by Nikolaos Roubanis from European Commission, Eurostat.

Session title: Using new and innovative data sources for transport statistics and indicators.

The development of digital technologies and increased availability of big data volumes has enabled the use of Big Data to produce statistics on road and maritime transport in shorter time. Moreover, the use of innovative data can lead to the development of new transport indicators that cannot be produced through traditional surveys. Identifying potential innovative and new data sources is leading to the production of experimental transport statistics that enable to improve the quality of indicators, that can be delivered through the benchmarking of new information with available statistical data. Furthermore, the use of new sources of data allows to reduce the burden on respondents and the costs of surveys. This session will present projects using innovative data and technologies to produce transport indicators in the European context.

Abstract

The development of digital technologies, increasing availability of Big Data and advanced processing techniques have enabled Statistics Poland to modernize the system for the production of road and maritime transport statistics. Actions taken to adapt modern Big Data technologies and data from sensors, e.g. from the Automatic Identification System (AIS) or the e-TOLL electronic toll collection system, enabled obtaining new statistics and accelerating the data dissemination.

The main goal of this speech is to present the innovative TranStat system (<https://transtat.stat.gov.pl>) that enables the production of road and maritime transport statistics using large volumes of data for shaping the country's transport policy, developed under the GOSPOSTRATEG programme.

The implementation of the TranStat project in the field of road and maritime statistics has enriched the current statistical production carried out by Statistics Poland through:

- access to streaming innovative data source related to maritime (AIS) and road (e-TOLL) transport;
- implementation of the necessary Big Data technology for data from sensors enabling an automatic process of data flow, validation and processing;
- development of traffic intensity, transport performance and emissions models in maritime transport with the use of data from sensors;
- development of algorithms enabling generation of new statistics and obtaining new knowledge in the field of maritime transport statistics by using the combination of multiple data sources;
- reduction of research costs thanks to the use of modern technology in the acquisition and processing of non-statistical sources

CPS Abstract

Filling the data gap in Italy on household multidimensional poverty by income, consumption and wealth

Author: Dr Gabriella Donatiello

Submission ID: 2009

Reference Number: 2009

Files/Uploads

[Abstract_Istat_Bank of Italy](#)

Brief Description

Strand: Challenges and innovations in data production: Address emerging data needs and stakeholders; enhance citizen participation

Authors:

Gabriella Donatiello, Marcello D'Orazio, Andrea Neri*. David Lo Schiavo*, Federico Tullio*,
Italian National Institute of Statistics - ISTAT and Bank of Italy*

Abstract

The availability of multidimensional indicators on the economic resources and material deprivation of households is considered an indispensable tool for policies to combat poverty and inequality. However, the production of micro data on the joint distribution of household Income, Consumption and Wealth (ICW) continues to present many challenges. This paper highlights the production of experimental data on the joint distribution of ICW in Italy to fill the missing data gap through statistical matching of the surveys "EU Statistics on Income and Living Conditions - EU-SILC", "Household Budget Survey - HBS" and "Survey on Household Income and Wealth - SHIW" of the Bank of Italy.

Since 2018, ISTAT and Bank of Italy started a close cooperation with the aim of creating a synthetic dataset in order to analyse the propensity to consume, to save, and asset-based poverty and wealth inequality. The paper describes the statistical matching methods applied by the Bank of Italy and ISTAT to produce experimental joint distributions of income, consumption and wealth. The Bank of Italy survey collects data on income, wealth, and information on consumption expenditures and after an adjustment of the collected income with fiscal data applies the hot deck procedures to impute the total consumption from HBS into SHIW.

The ISTAT method is more complex because it essentially proceeds with a double imputation: first, the total consumption from HBS and then it imputes the net wealth from SHIW, using a mixed method that takes into account the sampling weights. This is a well-known approach to the statistical matching of data from complex sample surveys modified to perform integration at the micro level and impute a continuous variable. Finally, we present the results of the experimental ICW distributions and the quality assessment, which continues to be a crucial point for the estimates produced with statistical matching techniques.

CPS Abstract

First combined (register-and-questionnaire-based) census in Latin America and the Caribbean? Innovations of the Uruguayan 2023 population census.

Author: Mr Federico Segui

Submission ID: 1904

Reference Number: 1904

Brief Description

In a groundbreaking move, the National Institute of Statistics (NSI) of Uruguay is at the forefront of a significant transition from traditional census methods to a pioneering register-based census, marking a historic shift in data collection strategies within Latin America and the Caribbean. The NSI's innovative approach involves a pilot project running in parallel with the traditional 2023 population census, aiming to compare and analyze data from both sources to identify gaps and enhance the overall accuracy of population statistics.

Central to this transformative initiative is the System of Integrated Statistical Registers and Surveys (SIREE), a comprehensive and dynamic system that goes beyond being a mere central database or information system. It represents a holistic framework with interconnected elements, prominently featuring a data warehouse as its most visible component. The SIREE system is underpinned by several key elements, including Master Data and Metadata Management, a tool named MasterData adhering to the ISO 11.179 standard, the GSRBPM model (adapted from the original GSBPM model by UNECE), and a Conceptual and Methodological framework.

SIREE is a system of integrated statistical registers but also censuses and surveys, drawing inspiration from the successful model developed by the Nordic countries. At its core are three base registers – population, business, and real estate. The architecture of the SIREE data warehouse comprises different layers, demonstrating a sophisticated blend of information technology and streamlined organizational processes.

The results of the pilot register-based census have been very auspicious and promising. Thus, the NSI is combining data from administrative registers with data from the traditional census to produce the final census outputs.

This combined census initiative and the implementation of the SIREE system represent a paradigm shift in census methodologies in Latin America and the Caribbean, showcasing Uruguay's commitment to advancing data accuracy, efficiency, and comprehensiveness. The comparative analysis between traditional and register-based approaches promises valuable insights, paving the way for future advancements in statistical practices across the region. The SIREE system, with its innovative architecture and methodology, stands as a beacon for other countries seeking to modernize their census processes and enhance the reliability of demographic data.

Abstract

In a groundbreaking move, the National Institute of Statistics (NSI) of Uruguay is at the forefront of a significant transition from traditional census methods to a pioneering register-based census, marking a historic shift in data collection strategies within Latin America and the Caribbean. The NSI's innovative approach involves a pilot project running in parallel with the traditional 2023 population census, aiming to compare and analyze data from both sources to identify gaps and enhance the overall accuracy of population statistics.

Central to this transformative initiative is the System of Integrated Statistical Registers and Surveys (SIREE), a comprehensive and dynamic system that goes beyond being a mere central database or information system. It represents a holistic framework with interconnected elements, prominently featuring a data warehouse as its most visible component. The SIREE system is underpinned by several key elements, including Master Data and Metadata Management, a tool named MasterData adhering to the ISO 11.179 standard, the GSRBPM model (adapted from the original GSBPM model by UNECE), and a Conceptual and Methodological framework.

SIREE is a system of integrated statistical registers but also censuses and surveys, drawing inspiration from the successful model developed by the Nordic countries. At its core are three base registers – population, business, and real estate. The architecture of the SIREE data warehouse comprises different layers, demonstrating a sophisticated blend of information technology and streamlined organizational processes.

The results of the pilot register-based census have been very auspicious and promising. Thus, the NSI is combining data from administrative registers with data from the traditional census to produce the final census outputs.

This combined census initiative and the implementation of the SIREE system represent a paradigm shift in census methodologies in Latin America and the Caribbean, showcasing Uruguay's commitment to advancing data accuracy, efficiency, and comprehensiveness. The comparative analysis between traditional and register-based approaches promises valuable insights, paving the way for future advancements in statistical practices across the region. The SIREE system, with its innovative architecture and methodology, stands as a beacon for other countries seeking to modernize their census processes and enhance the reliability of demographic data.

CPS Abstract

Generating Synthetic People and Households from Incomplete Aggregated Counts Using Mathematical Programming

Author: Prof. Gonzalo Peraza

Submission ID: 1929

Reference Number: 1929

Brief Description

Population synthesis has a long tradition within the micro-simulation and agent-based simulation domains. Traditional synthesis methods like Iterative Proportional Fitting (IPF) and Generalized Raking (GR) have been applied numerous times to generate agent-based populations required for simulation applications. Here we show such methods fail to generate adequate populations when trying to match aggregated counts with an incomplete set of marginals, such as those provided by the Mexican census. We discuss how and why these methods fail. We do so by studying the properties of linear systems that describe the census constraints. We then propose a new method to generate a synthetic population with integer counts based on a Mathematical Programming approach that satisfies both people and household constraints simultaneously and exactly. Our method can handle equality and inequality constraints, the latter of such being commonplace in the Mexican census. By being able to fit inequality constraints, we can naturally account for uncertainties in count statistics. Unlike methods based on IPF and GR, sample weights are optional, but can be accounted for during the fit. Furthermore, the method can generate a family of solutions, instead of a single best-fit solution, ranked by the fit statistic, a useful property for performing sensibility analysis. We demonstrate the method by generating a synthetic population for transport simulations in the Mexican city of Monterrey, analyze the properties of the fitted solutions, and perform a comparison with solutions obtained from IPF and GR.

Abstract

Population synthesis has a long tradition within the micro-simulation and agent-based simulation domains. Traditional synthesis methods like Iterative Proportional Fitting (IPF) and Generalized Raking (GR) have been applied numerous times to generate agent-based populations required for simulation applications. Here we show such methods fail to generate adequate populations when trying to match aggregated counts with an incomplete set of marginals, such as those provided by the Mexican census. We discuss how and why these methods fail. We do so by studying the properties of linear systems that describe the census constraints. We then propose a new method to generate a synthetic population with integer counts based on a Mathematical Programming approach that satisfies both people and household constraints simultaneously and exactly. Our method can handle equality and inequality constraints, the latter of such being commonplace in the Mexican census. By being able to fit inequality constraints, we can naturally account for uncertainties in count statistics. Unlike methods based on IPF and GR, sample weights are optional, but can be accounted for during the fit. Furthermore, the method can generate a family of solutions, instead of a single best-fit solution, ranked by the fit statistic, a useful property for performing sensibility analysis. We demonstrate the method by generating a synthetic population for transport simulations in the Mexican city of Monterrey, analyze the properties of the fitted solutions, and perform a comparison with solutions obtained from IPF and GR.

CPS Abstract

Harnessing the Power of Artificial Intelligence and Machine Learning – Creating a Formidable Multidisciplinary Research Collaboration to aid Human Cap**Author:** Prof. Oluwakemi Okuwa**Submission ID:** 1818**Reference Number:** 1818**Brief Description**

Collaboration across boundaries: Data community and statistics research community coming together.

Abstract

1O.B. Okuwa and 2O.J. Akintande

1Nigerian Institute of Social and Economics Research, Ibadan, Oyo state, Nigeria.

2Computational Statistical Unit, Department of Statistics, University of Ibadan, Oyo state, Nigeria.

Abstract

Data plays a crucial role in development planning of any nation. The dearth, sparsity and inaccessibility of data in African nations have created huge gaps in the development process and human capital development of these countries. Lack of data harmony from various data-generating institutions complicated the problem leading to the inconsistent repository of data for planning purposes and human capital development. Consequently, there is need to integrate and harmonized available data using rigorous models that will address major development challenges – such as unemployment, uncoordinated commerce, high inflation, societal unrest, insecurity/terrorism, natural disasters, infrastructural deficiencies, brain drain, maternal and child mortality, literacy and life expectancy issues in Africa region. Africa is endowed with great human resources that can drive structural transformation; however, these huge human resources and its potentials have not been fully tapped and explore due to lack of harmony in data coordination and collaborations. Countries like China and India have explored their human capital resources to drive economic growth and development in the last few decades, setting them on the path of structural and economic transformation among global economic powerhouses. This paper explores the potential of Artificial intelligence and Machine learning in harmonizing the gaps in data coordination in Africa. The paper also explores how modern tools can be harnessed to promote formidable multidisciplinary collaboration within and among nations in Africa. Thus, the study systematically explores how AI/ML can help African nations achieve human capital development by creating a likely correlation of variables derived from the developed nations. And using AI/ML to measure the key drivers of human capital development across these powerhouses - to promote data-driven policy modelling in Africa and Nigeria.

CPS Abstract

INEGI's Advancements in Remote Sensing for a Comprehensive View of Mexico's Agricultural Landscape Using Data Science Techniques

Author: Dr Abel Coronado

Submission ID: 1946

Reference Number: 1946

Brief Description

This research by INEGI focuses on the advanced analysis of Mexico's agricultural zones using remote sensing, particularly necessary due to the challenges in conducting agricultural censuses. It leverages Landsat imagery and the Mexican Geospatial Data Cube, integrating data from the 2007 Agricultural Census with modern remote sensing techniques, including GEOBIA, which processes images at a 30-meter resolution across over 9 million polygons. The study also incorporates SENTINEL-2 imagery, achieving up to 91.24% accuracy in certain regions using the "random forest" algorithm, a notable improvement over the Landsat approach. The research underscores the use of open-source Python tools and cutting-edge techniques like the geometric median and GEOBIA, offering a detailed, current view of Mexico's agricultural areas.

Abstract

INEGI recognizes the significance of studying Mexico's agricultural zones, particularly in light of the financial and budgetary challenges posed by conducting agricultural censuses. The delineation and identification of lands with agricultural activity is a complex task, and remote sensing has emerged as a crucial instrument for this purpose. Especially when traditional methods face economic and logistical barriers, the amalgamation of conventional techniques with innovative data science approaches is of paramount importance.

Utilizing Landsat imagery and the Mexican Geospatial Data Cube, data from the 2007 Agricultural Census (updated to 2019) were integrated with state-of-the-art remote sensing methodologies. Embracing GEOBIA has been instrumental, as it processes images on a national scale at a 30-meter resolution, resulting in over 9 million distinct polygons.

Moreover, SENTINEL-2's remote sensing capabilities have recorded impressive advancements in accuracy, reaching up to 91.24% in select regions when employing the "random forest" algorithm. This method notably surpasses the Landsat approach, attributed to the enhanced resolution of SENTINEL-2.

Highlighting both studies is the emphasis on open-source Python tools and avant-garde techniques such as the geometric median and GEOBIA. When paired with high-resolution images and contemporary databases, these resources provide an in-depth and updated depiction of the agricultural regions in Mexico.

CPS Abstract

Impact of the first COVID-19 lockdown on HIV diagnosis in England**Author:** Feysal Muhammed**Submission ID:** 2017**Reference Number:** 2017**Files/Uploads**[Feysal_STRAND 6_Mexico_Abstract](#)**Brief Description**

This paper is the collaborative work of different researchers across the world, the main author (Feysal Kemal Muhammed from Ethiopia, and other authors from the UK and Italy).

Abstract

STRAND 6: Collaboration across boundaries: Data community and statistics research community coming together

Title: Impact of the first COVID-19 lockdown on HIV diagnosis in England

Feysal Kemal Muhammed¹, Peter Kirwan², Tasnuva Tabassum², James Lester³, Pantelis Samartsidis², Alison Brown³, Daniela De Angelis², Chiara Chiavenna⁴, and Anne Presanis²

¹Department of Statistics, Hawassa University, Ethiopia

²MRC Biostatistics Unit, University of Cambridge, UK

³UK Health Security Agency, UK

⁴Bocconi Institute for Data Science and Analytics, Bocconi University, Italy

STI and HIV prevention and care services in England were unavoidably disrupted by the COVID-19 lockdowns, and it is likely incidence was also affected. We estimated the magnitude of the effect of the first UK COVID-19 lockdown in March 2020 on subsequent HIV diagnoses, using the quarterly time series of observed diagnoses during the period 2004 to 2020. We used a Bayesian Structural Time Series model to estimate the counterfactual number of new HIV diagnoses in the post-lockdown period that would have occurred in the absence of lockdown. This counterfactual was constructed using synthetic controls (time series correlated with HIV diagnoses unaffected by pre-lockdown) based on Google Trend data. The effect was estimated as the difference between the counterfactual and the observed number of diagnoses, for strata defined by gender, sexual orientation, age, and whether or not the diagnosis was late. All strata had greater than 50% posterior probability that the first COVID-19 lockdown had an effect on HIV diagnosis, with the highest posterior probability of an effect (99%) in heterosexual men aged 50+ with late diagnosis. The estimated overall number of missed diagnoses was 97 (95% prediction interval [-688, 850]). We found that the effect size varied by stratum and by time since the lockdown, with the largest effect in the first quarter after the lockdown. The effect then reduced, as HIV testing, and consequent diagnosis, caught up during the summer of 2020. As the number of new HIV diagnoses made in England and estimated incidence fell consistently in the five years pre-pandemic, we hypothesise that the effect of the lockdown may have been less pronounced than it could have been if the epidemic had not already been in decline. Public health campaigns which encouraged HIV home testing during the lockdown may also have helped to mitigate the effect.

CPS Abstract

Implementation of the Open Data Cube for Earth Observation in Mexico: Challenges and Prospects in Generating Statistical and Geographical Information

Author: Dr Abel Coronado

Submission ID: 1944

Reference Number: 1944

Brief Description

This research focuses on Mexico's National Institute of Statistics and Geography's (INEGI) implementation of the Open Data Cube (ODC) for managing Earth Observation (EO) data, highlighting challenges and benefits. The study emphasizes the creation of the Mexican Geospatial Data Cube (CDGM), integrating Landsat and Sentinel-2 data for local storage and data ingestion. The CDGM's effectiveness in monitoring national resources and agricultural activities through spatiotemporal analysis is showcased. Finally, the paper discusses the lessons learned, future expansion plans, and the importance of international collaboration to maximize EO data utilization.

Abstract

Efficient management and detailed analysis of Earth Observation (EO) data are pivotal in the current era, marked by significant environmental changes and steady population growth. This research underscores the experience of Mexico's National Institute of Statistics and Geography (INEGI) in implementing the Open Data Cube (ODC) within its local infrastructure, focusing on generating advanced analyses and facilitating access to vital information through standardized and consistent products.

Leveraging tools provided by the Open Data Cube (ODC) initiative and the Swiss Data Cube, this study presents the challenges, costs, and benefits associated with deploying an "on premises" infrastructure for the ODC. It highlights its applicability in supervised classification of natural and economic phenomena. The analysis emphasizes how the ODC can navigate the intricacies posed by the vast and intricate spectrum of EO data, allowing for a more effective exploitation of these resources.

The paper elaborates on the creation of the Mexican Geospatial Data Cube (CDGM), integrating data from Landsat and Sentinel-2, emphasizing the local storage strategy and data ingestion capabilities that enhance the system's expansion to accommodate various EO data types. Through specific case studies, such as monitoring water resources and agricultural activities, the efficacy of the CDGM in delivering relevant spatiotemporal analyses at the national level is demonstrated.

The research concludes by spotlighting lessons learned, future endeavors necessary for the expansion and optimization of the CDGM, and the paramount importance of ongoing collaboration between national and international entities to fully harness the potentialities offered by Earth Observation data.

CPS Abstract

Innovations and Strategies in the Mexican Electoral Registry through Data Science

Author: Miguel David Alvarez Hernández

Submission ID: 2099

Reference Number: 2099

Brief Description

The electoral registry, which comprises both the Electoral Roll and the Nominal List of Voters, plays a central role in the Mexican Electoral System by determining the eligibility of individuals to participate as voters or candidates in elections. The accuracy and comprehensiveness of this instrument are crucial, since it is imperative to ensure the inclusion of all citizens entitled to vote and to be elected as public representatives. This implies guaranteeing that no social group is excluded or disadvantaged, thus allowing all citizens to fully exercise their political-electoral rights.

However, the demographic dynamics of the country and the inherent processes of incorporation and purging of the electoral registry present challenges that make it difficult to achieve a complete and updated inclusion of all citizens at a specific time. Factors such as the interval between the moment a person reaches the age of majority and his or her registration in the voter registry, the time lapses between the change of address and its notification, as well as the death of citizens and the corresponding notification, contribute to make the construction and updating of the voter registry a complex and continuous task.

In this context, this presentation explores various proposals and strategies aimed at modernizing the capture, management and analysis of Mexican electoral registry information. This includes the adaptation and transformation of administrative entities in the current context of data science.

Abstract

The electoral registry, which comprises both the Electoral Roll and the Nominal List of Voters, plays a central role in the Mexican Electoral System by determining the eligibility of individuals to participate as voters or candidates in elections. The accuracy and comprehensiveness of this instrument are crucial, since it is imperative to ensure the inclusion of all citizens entitled to vote and to be elected as public representatives. This implies guaranteeing that no social group is excluded or disadvantaged, thus allowing all citizens to fully exercise their political-electoral rights.

However, the demographic dynamics of the country and the inherent processes of incorporation and purging of the electoral registry present challenges that make it difficult to achieve a complete and updated inclusion of all citizens at a specific time. Factors such as the interval between the moment a person reaches the age of majority and his or her registration in the voter registry, the time lapses between the change of address and its notification, as well as the death of citizens and the corresponding notification, contribute to make the construction and updating of the voter registry a complex and continuous task.

In this context, this presentation explores various proposals and strategies aimed at modernizing the capture, management and analysis of Mexican electoral registry information. This includes the adaptation and transformation of administrative entities in the current context of data science.

CPS Abstract

Innovative methods of producing and communicating SDG data disaggregation in Poland

Author: Ms Renata Bielak

Submission ID: 1857

Reference Number: 1857

Brief Description

Improving SDG data disaggregation is essential for making the global goals a reality in cities and regions. So as to bridge SDG data gaps, in particular with regard to territorial disaggregation, Statistics Poland has engaged in exploring innovative methods for producing and communicating SDG statistics for the local and regional level. In response to the need for the ready-made analyses and accessible ways of data presentation, since 2020, Statistics Poland has also issued digital SDG reports.

Abstract

Strand: Leaving no one behind: Greater disaggregation and counting oversights.

Title: Innovative methods of producing and communicating SDG data disaggregation in Poland.

Improving SDG data disaggregation is essential for making the global goals a reality in cities and regions, while ensuring that no one is left behind. So as to bridge SDG data gaps in particular with regard to territorial disaggregation, Statistics Poland has engaged in exploring innovative methods for producing and communicating SDG statistics for the local and regional level.

Thanks to the application of i.a. Earth Observation methods, Statistics Poland has succeeded in providing greater disaggregation for urban and rural areas as well as regions, which is crucial for better tailored and evidence-based policy formulation at all levels. To enhance communication and impact of these new disaggregations, Statistics Poland is developing a new product, which integrates various means of data presentation (maps, charts and analytical commentary) in an interactive and user-friendly way. The product contains data covering SDG 9. Industry, innovation and infrastructure and SDG 11. Sustainable cities and communities, and will be further expanded with other indicators and disaggregation categories from innovative sources.

Catering for the need for ready-made data analyses and accessible ways of data presentation, since 2020, Statistics Poland has issued digital reports "Poland on the way to SDGs" on an annual basis. The interactive reports provide data users with a detailed overview of selected issues connected to sustainable development, so far taking into account: inclusive economic growth, environmentally sustainable development and gender equality in Poland.

CPS Abstract

Integrating AI into Household Survey Data Encoding Workflows

Author: Alejandro Pimentel

Submission ID: 1920

Reference Number: 1920

Brief Description

Statistical offices regularly undertake the arduous and resource-intensive task of coding records for various data products. This paper delves into the strategies employed to develop a system based on deep learning and natural language processing techniques, with the goal of automating aspects of this encoding process. The experiments and results presented herein demonstrate the potential for these methodological innovations to partially replace the manual coding processes traditionally carried out by the Mexican Statistical Institute (INEGI) through the application of deep learning.

In this proposal, we introduce a novel phase into the data encoding workflows utilized by INEGI for the household surveys, incorporating an Artificial Intelligence (AI) model. The primary objective of this new phase is to reduce the volume of records that currently require manual coding. To achieve this, our proposal involves the selection of a subset of records for which the AI algorithm exhibits a sufficient level of confidence in assigning the correct code.

The effectiveness of this proposal was assessed by implementing a production line that closely resembled the existing process, but with the integration of the AI-based phase. Our results indicate that the utilization of deep learning algorithms enables a substantial 50% reduction in manual coding while maintaining a quality level on par with the original workflow.

CPS Abstract

LMIS: Better data, better jobs

Author: Mr Edgardo Greising

Submission ID: 2018

Reference Number: 2018

Brief Description

The identification of labour market problems in both developed and developing economies relies heavily on the availability of data, information, and analysis. Labour market indicators (LMIs) provide an essential basis for employment and labour market policies, and facilitates the design, implementation, monitoring and evaluation of better targeted policies.

The International Labour Organization (ILO) provides assistance to countries interested in implementing Labour Market Information Systems (LMIS) through a structured roadmap building capacities to:

- the calculation of LMIs from available data sources, according to international recommendations and standards;
- the harmonization of data obtained from heterogeneous sources, mostly relying on the SDMX information model, and standard classifications (international, regional and/or national);
- the correct interpretation of the LMIs and the application of statistical and econometric analysis techniques;
- the establishment of a governance model for the system that promotes the participation of all producers and users, information transparency, and increasing the exchange of data and metadata between institutions.

ILO also provides the software tools (open source and standards-based) and technical support, as well as the necessary knowledge for their use.

Components of a LMIS

LMIS consists of four main components: data and information collection and compilation, information repository, analytical capacity and tools, and institutional arrangements and networks.

Data collection consists not only of data on the labour market, but also on the economy as a whole. The main sources of labour statistics are household surveys, population censuses, establishment surveys and administrative registers. Each of them has advantages and limitations, are heterogeneous and not always suitable for obtaining statistical indicators. Therefore, an effective LMIS must be based on all sources, but the information must be harmonized.

This information collected and/or derived from such primary sources is stored and made accessible through a single repository of labour market information.

Analytical skills are essential to identify and interpret labour market development and trends, and to relate these trends to policies or other factors that influence labour market outcomes.

The institutional arrangements and governance of LMIS facilitate the creation of networks of users and producers, enabling access to data, as well as to effective dissemination of information and analysis.

LMIS and the Principles

The implementation of a LMIS is an example of application of the Fundamental Principles of Official Statistics. The information on the Labour Market produced by multiple institutions is made available to the public in a fully transparent way, honouring citizens' entitlement to public information (Principle 1), with the highest professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data (Principle 2), based on scientific standards (Principle 3). The information to feed a LMIS is drawn from multiple data sources (Principle 5), keeping strict confidentiality on the individual's data (Principle 6).

As already mentioned, the establishment of a LMIS promotes the participation of all producers and users, increasing the coordination and exchange of data and metadata between institutions (Principle 8). The system relies heavily on the use of international standards for concepts, classifications, methods, and tools as enablers for data harmonization (Principle 9).

The application across multiple countries of a structured work plan with a well-defined set of activities composing the roadmap for the implementation, has contributed to the establishment of bilateral collaborations among countries, for the development of their LMIS (Principle 10).

CPS Abstract

Leitmotifs of the international debate – how do official statistics can leverage global topics to define and provide trusted evidence?

Author: Ms Olga Swierkot-Struzewska

Submission ID: 2064

Reference Number: 2064

Brief Description

New global phenomena often result in new topics of debate in the world of official statistics.

Over the last number of years we have witnessed the world changing rapidly. New challenging issues emerged, such as the Post Truth world, the Digital Revolution and the growing emphasis on ethics and official statistics. These challenges have generated high level discussion in the world of official statistics around themes such as: "Living in a post-truth world where data can make a difference"; "The Digital Revolution – new data sources impacting our work methods and products"; "Statistics are key for decision-making and fighting disinformation"; "We must strive to redefine our roles in a changing data ecosystem, acting ethically, leveraging our values and principles"; "SDGs are there to leave no one behind"; "Official statistics is a public good".

This intervention aims at tackling the recent leitmotifs of the international debate from the perspective of communication. What do our key audiences really understand of these topics? How well do we translate their impact into explaining the value and relevance of official statistics? To what extent do we use their potential in order to leverage the role and mission of national statistical offices?

The presentation will explore the increasing need for trust in official statistics which can only be instilled if the adequate message is conveyed. How best do we go about building trust and ensuring social acceptability for our work in an increasingly complex world and how do we adapt our communications practices to help us achieve this?

The presentation also examines powerful global topics in the context of data and official statistics. How can we leverage their profile as a trigger to enhance the usage of official statistics, develop appropriate tools and rethink how we reach target audiences. How can we use global themes as a spring board to transform and modernise our traditional communication channels in NSOs to better explain why official statistics and data really matter.

Abstract

New global phenomena often result in new topics of debate in the world of official statistics.

Over the last number of years we have witnessed the world changing rapidly. New challenging issues emerged, such as the Post Truth world, the Digital Revolution and the growing emphasis on ethics and official statistics. These challenges have generated high level discussion in the world of official statistics around themes such as: "Living in a post-truth world where data can make a difference"; "The Digital Revolution –new data sources impacting our work methods and products"; "Statistics are key for decision-making and fighting disinformation"; "We must strive to redefine our roles in a changing data ecosystem, acting ethically, leveraging our values and principles"; "SDGs are there to leave no one behind"; "Official statistics is a public good".

This intervention aims at tackling the recent leitmotifs of the international debate from the perspective of communication. What do our key audiences really understand of these topics? How well do we translate their impact into explaining the value and relevance of official statistics? To what extent do we use their potential in order to leverage the role and mission of national statistical offices?

The presentation will explore the increasing need for trust in official statistics which can only be instilled if the adequate message is conveyed. How best do we go about building trust and ensuring social acceptability for our work in an increasingly complex world and how do we adapt our communications practices to help us achieve this?

The presentation also examines powerful global topics in the context of data and official statistics . How can we leverage their profile as a trigger to enhance the usage of official statistics, develop appropriate tools and rethink how we reach target audiences. How can we use global themes as a spring board to transform and modernise our traditional communication channels in NSOs to better explain why official statistics and data really matter.

CPS Abstract

Mapping Migration Trends: Leveraging Data Science and Web-Scraped News for Comprehensive Analysis of Transitory Populations

Author: Dr Elio Villaseñor

Submission ID: 1934

Reference Number: 1934

Brief Description

This paper addresses the complex challenge of monitoring the foreign population transiting through a country en route to the United States, a task complicated by the predominantly undocumented and adaptive nature of the migrants. The primary focus of the study is the integration of data from official sources, analyzed both temporally and spatially, to effectively visualize the patterns of migrant mobility. Central to this research is the development of a methodology for estimating the annual flow of migrants through the country. This involves the application of data science techniques not only to estimate migration routes but also to visualize both territorial and seasonal patterns of these movements. A significant innovation of this study is the utilization of web-scraped news as a tool to measure the intensity of migrant-related discourse in various regions. The implementation of a specialized algorithm capable of recognizing named locations within news content plays a pivotal role in this analysis, enhancing the understanding of migration patterns. Furthermore, the paper explores the congruence between alternative data sources, such as online news, and traditional data in terms of the intensity and spatial distribution of migration. This exploration is crucial for evaluating the efficacy of these sources in tracking the evolution of migration trends in a timely manner and for providing in-depth insights into specific migrant flows. The findings of this study contribute to a more nuanced understanding of migration patterns and offer a methodological framework that can be applied to similar contexts globally.

Abstract

The research presented in this paper tackles the intricate challenge of tracking the movement of foreign populations transiting through Mexico on their way to the United States. This task is particularly complex due to the largely undocumented status of these migrants, who often change their routes and strategies. The study's main objective is to integrate and analyze data from official sources in a temporal and spatial context to effectively map the mobility patterns of these migrants.

In this research we look for the creation of a novel methodology aimed at estimating the yearly influx of migrants traversing the country. This method leverages advanced data science techniques, not just for mapping out potential migration routes, but also for depicting the territorial and seasonal variances in these movements. Such analysis is crucial for understanding the dynamics of migration and for planning and policy-making purposes.

A key aspect of this study is its use of web-scraped news sources as a means to quantify the intensity of migrant-related discourse across different regions. The study introduces a specialized algorithm designed to identify specific named locations within news content. This algorithm is critical in enhancing the precision of migration pattern analysis, offering a more detailed and nuanced understanding of how and where migrant flows are concentrated.

Moreover, the paper delves into the relationship between alternative data sources, such as online news, and traditional data, particularly in terms of the intensity and spatial distribution of migration patterns. This investigation evaluates the reliability and effectiveness of these sources in capturing the ongoing changes in migration trends. By doing so, the study looks for deeper insights into particular migrant flows.

The findings from this research are not only pivotal for comprehending migration patterns more thoroughly but also for offering a methodological framework that can be applied in other global contexts. The framework has the potential to assist policymakers, researchers, and humanitarian organizations in making informed decisions and strategies related to migration management and support.

CPS Abstract

Mapping indicators on Forcibly Displaced Persons (FDPs): the UNHCR journey.

Author: Andrea Pellandra

Submission ID: 1967

Reference Number: 1967

Brief Description

The paper will describe UNHCR's journey towards the elaboration of a framework for the selection and prioritization of key statistically representative indicators of particular relevance for Forcibly Displaced Populations based on main international frameworks and recommendations, and on the challenges involved in deriving them from existing data sources. This work's findings can be generalized to some extent and can be of use for other organizations interested in embarking on a similar journey, in a flexible way, either to inform the selection process of key indicators or to add further indicators to their analysis.

Abstract

To ensure that the promise of the 2030 Agenda to "leave no one behind" is effectively met, several vulnerable population groups are identified for attention including migrants, refugees, and other forcibly displaced persons (FDPs), and to achieve this goal, enhancing the availability of better data and evidence is essential. In 2018, the Expert Group on Refugee and IDP Statistics (EGRIS) developed a common set of indicators to measure refugees' well being and inclusion in host communities, identifying identified key policy areas that are most relevant for forcibly displaced populations. In 2022, UNHCR embarked in a large project aiming at developing methodological standard to produce curated socioeconomic and protection indicators on forced displacement. The purpose of this paper is to provide a review of the challenges involved in the production of statistically solid thematic indicators on forcibly displaced people, including the presence of FDPs in sampling frames, sampling size requirements necessary to produce statistically representative indicators, survey questions for the identification of FDPs in datasets, and the availability of variables required for the computation of indicators in existing data. More specifically, the paper touches upon the journey towards the elaboration of a framework for the selection and prioritization of key indicators of particular relevance for FDPs based on main frameworks and recommendations, and on how they can be derived from existing data sources. This work's findings can be generalized to some extent and can be of use for other organizations interested in embarking on a similar journey, in a flexible way, either to inform the selection process of key indicators or to add further indicators will be included in their analysis.

CPS Abstract

Measuring the scale of envelope wages: Evidence from Poland

Author: Prof. Marek Kosny

Submission ID: 1996

Reference Number: 1996

Brief Description

We develop a methodology for measuring the scale of the unregistered wages paid to staff by employers. We show that the differences in earnings distribution between public- and private-sector employees are not reflected in the distribution of expenditure. After showing that the distribution of income from non-wage sources does not systematically depend on the sector, we argue that the differences may be attributed to envelope wages. We estimate that 10-11% of employees in Poland receive envelope wages, which account for around 5% of total earnings.

Abstract

The paper addresses the phenomenon of paying employees some of their earnings outside formal employment relationships. This involves paying them in two parts. One is officially set out in their employment contract. The other results from an unwritten and informal arrangement between the employee and the employer. This part is handed over in cash and not included in records (unreported). This phenomenon is referred to as "envelope wages". In cases when employees earn some of their remuneration informally, we refer to their reported (formal) income or to their actual (true) income (the latter is the sum of their formal and informal income).

Using data from Structure of Earnings Survey and Household Budget Survey, we develop a methodology for measuring the scale of the unregistered wages paid to staff by employers. We show that the differences in earnings distribution between public- and private-sector employees are not reflected in the distribution of expenditure. Assuming that public-sector employees do not receive unreported income, this points to undisclosed income in the private sector. After showing that the distribution of income from non-wage sources does not systematically depend on the sector, we argue that the differences may be attributed to envelope wages. We estimate that 10-11% of employees in Poland receive envelope wages, which account for around 5% of total earnings.

CPS Abstract

Modelling local variations across small geographical areas of COVID-19 Pandemic: A machine learning approach

Author: Prof. Oyelola Adegboye

Submission ID: 2016

Reference Number: 2016

Brief Description

Foundational work in statistics and data science: Better understanding and innovation in statistical research and decision-making

This strand aims to provide the platform for statistical researchers to exchange ideas and disseminate results and to promote innovative research in statistics and data science.

Abstract

Estimating health and welfare at the neighbourhood level from surveillance data could have many social and economic applications, such as assisting local policymakers in organizing targeted activities to improve the health and well-being of their residents. However, health outcomes data often exhibit spatial correlation, rendering traditional statistical modelling methods impractical. While several spatial statistical models have been proposed to address this issue, they are prone to computational complexities. Machine learning techniques present attractive alternatives with potential benefits over traditional models, gaining recognition as solutions to small-area estimation challenges. Yet, they may not inherently recognize spatial context. Applying machine learning directly to health outcomes data without accounting for potential spatial autocorrelation could lead to biased outcomes. In this study, spatial context was integrated into machine learning techniques to efficiently capture the spatial variation of COVID-19 in contagion US and its associated risk factors. Beyond addressing spatial heterogeneity in COVID-19 data, the study utilized multiple indicators to create spatially weighted composite indicators to explain the disproportionate COVID-19 prevalence in small areas of the US in terms of demographic, socioeconomic, health access, vaccination, meteorological, and environmental risks.

CPS Abstract

Monthly statistics on self-employed - based on algorithms which process administrative data

Author: Miss Tove Henning

Submission ID: 1870

Reference Number: 1870

Brief Description

Challenges and innovations in data production: Address emerging data needs and stakeholders; enhance citizen participation. This strand will examine how new data demands offer opportunities for current and future statistical programs.

Abstract

Introduction

In an era when sample surveys are expensive and respondents are hard to find and extract information from, administrative data sources become more attractive. In addition, due to the process of digitalization, individuals leave both more and more frequent administrative footprints after their activities. Monthly statistics on population's statuses become now possible for Statistics Sweden to produce. This unfolds opportunities to follow e.g. cyclical labor market phenomena on a granular level.

To obtain a comprehensive picture of the population's status of employment, it is necessary to develop methods which identify individuals' who are self-employed and classify their status of employment. Therefore, Statistics Sweden has developed two separate algorithms to identify: i) self-employed persons in sole proprietorships, trading or limited partnerships and ii) owners of limited companies. As far as we know, these algorithms are in an international context novel methods of identifying self-employed individuals on monthly basis.

Novel algorithms

The algorithm to determine self-employment is divided into several steps. In a first step, administrative data is obtained from the Swedish Tax Agency and the Swedish Companies Registration Office, which is delimited at the individual and company level with the help of the Register of Total Population (RTB) and the Business Register (FDB). Based on this set of data an algorithm determines who are self-employed.

The algorithm for identifying the owners of limited companies is more straight forward. It uses the information of which individuals are in the register of beneficial ownership at the Swedish Companies Registration Office and delimited at the individual and company level with the help of the Register of Total Population (RTB) and the Business Register (FDB).

Result

The results from these algorithms satisfy users with adequate statistics on the development of self-employed on monthly basis. When producing this new statistics, monthly development of self-employed, in its different status of employment, can hence follow the current development with a lag of two months.

CPS Abstract

Multidimensional poverty in Senegal: design of a household vulnerability assessment tool

Author: Mr Jean Pierre Adiouma NDIAYE

Submission ID: 1882

Reference Number: 1882

Brief Description

For over a decade, poverty-related issues have been at the heart of Senegal's public of Senegal's public authorities. Thus, the Senegalese government Senegal, through the two PRSPs, the SNDS and now the Plan Sénégal Emergeant (PSE), has made the fight against poverty its top priority. However, the unidimensional approach to poverty, based on the observation of income, is proving to be restrictive, insofar as it fails to capture the phenomenon in all its complexity. complexity of the phenomenon. Hence the urgent need to develop a measurement approach that can the phenomenon in all its complexity. This report is a response to this need. that need. Through a multidimensional poverty index for Senegal, it aims to equip Senegal, to equip decision-makers in their policies to combat poverty. The multidimensional multidimensional analysis of poverty, following the methodology of Alkire and Foster (2007) indispensable insofar as it enables poverty to be assessed within a socio-economic framework economic framework, such as health, education, living conditions, food insecurity insecurity, etc. As such, it provides a better understanding of the notion of poverty as a whole. The analyses were carried out on data from the Enquête Harmonisée sur les Conditions de Vie des Ménages au Sénégal (EHCVM 2018/2019). The results from these analyses show that in Senegal, almost one household in two (50.19%) is multidimensionally poor, and that on average a poor household suffers from almost 54% of deprivations. In addition, households identified as poor suffer 27.22% of all possible deprivations that all households in Senegal could experience. This level of deprivation is essentially attributable to the precarious situation of less urbanized or peripheral regions, but also to adult illiteracy and the level of schooling nationwide; to the food insecurity that prevails in urban areas; and to the lack of comfort facilities, difficulties in accessing electricity, quality sanitation and cleanliness services in rural areas. As with monetary poverty (ANDS, EHCVM 2018/2019), households headed by women are less exposed to multidimensional poverty than those headed by men. On the other hand, households with married heads are more (3.4 times) likely to be multidimensionally poor than those under the authority of a single head. As a result of these analyses, a household vulnerability assessment tool has been designed to regularly monitor the impact of anti-poverty policies.

Abstract

For over a decade, poverty-related issues have been at the heart of Senegal's public of Senegal's public authorities. Thus, the Senegalese government Senegal, through the two PRSPs, the SNDS and now the Plan Sénégal Emergeant (PSE), has made the fight against poverty its top priority. However, the unidimensional approach to poverty, based on the observation of income, is proving to be restrictive, insofar as it fails to capture the phenomenon in all its complexity. complexity of the phenomenon. Hence the urgent need to develop a measurement approach that can the phenomenon in all its complexity. This report is a response to this need. that need. Through a multidimensional poverty index for Senegal, it aims to equip Senegal, to equip decision-makers in their policies to combat poverty. The multidimensional multidimensional analysis of poverty, following the methodology of Alkire and Foster (2007) indispensable insofar as it enables poverty to be assessed within a socio-economic framework economic framework, such as health, education, living conditions, food insecurity insecurity, etc. As such, it provides a better understanding of the notion of poverty as a whole. The analyses were carried out on data from the Enquête Harmonisée sur les Conditions de Vie des Ménages au Sénégal (EHCVM 2018/2019). The results from these analyses show that in Senegal, almost one household in two (50.19%) is multidimensionally poor, and that on average a poor household suffers from almost 54% of deprivations. In addition, households identified as poor suffer 27.22% of all possible deprivations that all households in Senegal could experience. This level of deprivation is essentially attributable to the precarious situation of less urbanized or peripheral regions, but also to adult illiteracy and the level of schooling nationwide; to the food insecurity that prevails in urban areas; and to the lack of comfort facilities, difficulties in accessing electricity, quality sanitation and cleanliness services in rural areas. As with monetary poverty (ANDS, EHCVM 2018/2019), households headed by women are less exposed to multidimensional poverty than those headed by men. On the other hand, households with married heads are more (3.4 times) likely to be multidimensionally poor than those under the authority of a single head. As a result of these analyses, a household vulnerability assessment tool has been designed to regularly monitor the impact of anti-poverty policies.

CPS Abstract

Nowcasting industrial production index with high-frequency highway toll data

Author: Mr Peter Knížat

Submission ID: 1956

Reference Number: 1956

Files/Uploads

[Nowcasting industrial production index with high-frequency highway toll data_abstract](#)

Brief Description

Nowcasting is a phenomenon that has recently become widely researched area in the context of official statistics. The main reason is an exploration of new data sources that can be used in the early estimation of the economic activity. A digital system that records passages of vehicles through highways generates such big data of early available information. It assembles high-frequency data, which can be collected relatively shortly afterwards, that holds information related to transport of goods, commodities and population through roads. A transportation of goods and commodities, usually carried out by trucks, through roads can provide an early indicator of the economic activity for industries that rely on the road transport. The objective of this paper is to estimate an index, which measures monthly changes of the number of vehicles on highways. The truck-passage index is able to provide a nowcasting estimate of the economic output for industries that heavily rely on the road transport. In the empirical study, we use toll data from the Slovak digital highway system that capture in-and-out passages of all passing trucks through monitored sections of highways. The truck-passage index is then compared to a monthly industrial production index that aims to validate its accuracy, and, potentially, provide its early nowcasting indicator. Furthermore, we use an empirical mode decomposition to extract a business cycle and de-trend time series of both truck-passage and industrial production indices that allows for the in-depth analysis of fluctuations of economic time series.

Abstract

Nowcasting is a phenomenon that has recently become widely researched area in the context of official statistics. The main reason is an exploration of new data sources that can be used in the early estimation of the economic activity. A digital system that records passages of vehicles through highways generates such big data of early available information. It assembles high-frequency data, which can be collected relatively shortly afterwards, that holds information related to transport of goods, commodities and population through roads. A transportation of goods and commodities, usually carried out by trucks, through roads can provide an early indicator of the economic activity for industries that rely on the road transport. The objective of this paper is to estimate an index, which measures monthly changes of the number of vehicles on highways. The truck-passage index is able to provide a nowcasting estimate of the economic output for industries that heavily rely on the road transport. In the empirical study, we use toll data from the Slovak digital highway system that capture in-and-out passages of all passing trucks through monitored sections of highways. The truck-passage index is then compared to a monthly industrial production index that aims to validate its accuracy, and, potentially, provide its early nowcasting indicator. Furthermore, we use an empirical mode decomposition to extract a business cycle and de-trend time series of both truck-passage and industrial production indices that allows for the in-depth analysis of fluctuations of economic time series.

CPS Abstract

Opportunities And Challenges Faced By National Statistical Offices In The Digital Era: The Recent Experiences From China

Author: Ms ZHIHUA TANG

Submission ID: 2104

Reference Number: 2104

Brief Description

Authors: LYU Haiqi, XU Yipin, ZHAO Junli and TANG Zhihua.

National Bureau of Statistics of China(NBS).

Email address: tangzhihua198@163.com

Abstract: The rapid development of the digitalization of society brings both opportunities and challenges for national statistical offices (NSOs). On one hand, the application of modern information technologies provides NSOs with valuable data resources, bringing new possibilities for technical methods as well as new releasing channels for us. On the other hand, the public's demand for official data is rising in both quantity and quality, therefore, NSOs face more competition from a growing number of data providers. To address these challenges, the National Bureau of Statistics of China (NBS) has undertaken a series of explorations and implemented many innovative practices in recent years, such as remote sensing application, digital economy accounting, statistical cloud programs, and the application of e-commerce data in retail sales statistics. The experiences of China will hopefully be of benefit to the NSOs of the other countries. The National Bureau of Statistics of China is willing to work with other countries to jointly embrace opportunities and challenges, and contribute wisdom to strengthening the statistical capacity building of national statistical institutions in the digital era.

Abstract

The rapid development of the digitalization of society brings both opportunities and challenges for national statistical offices (NSOs). On one hand, the application of modern information technologies provides NSOs with valuable data resources, bringing new possibilities for technical methods as well as new releasing channels for us. On the other hand, the public's demand for official data is rising in both quantity and quality, therefore, NSOs face more competition from a growing number of data providers. To address these challenges, the National Bureau of Statistics of China (NBS) has undertaken a series of explorations and implemented many innovative practices in recent years, such as remote sensing application, digital economy accounting, statistical cloud programs, and the application of e-commerce data in retail sales statistics. The experiences of China will hopefully be of benefit to the NSOs of the other countries. The National Bureau of Statistics of China is willing to work with other countries to jointly embrace opportunities and challenges, and contribute wisdom to strengthening the statistical capacity building of national statistical institutions in the digital era.

CPS Abstract

Particulate Matter drives Acute Respiratory Infection among Under-Five Children across sub-Saharan African Countries: Machine Learning Approaches**Author:** Dr Haile Mekonnen Fenta**Submission ID:** 1990**Reference Number:** 1990**Brief Description**

Background: Symptoms of Acute Respiratory infections (ARIs) among under-five children are a global health challenge. We aimed to train and evaluate the machine learning (ML) classification approaches in predicting symptoms of ARIs reported by mothers (caregivers) among children younger than five years in sub-Saharan African (sSA) countries.

Methods: We used the most recent (2012-2022) nationally representative Demographic and Health Surveys (DHS) data of 33 sSA countries. Machine learning algorithms such as logistic regression, LASSO, Elastic Net, Random Forest (RF), Artificial Neural Network (ANN), Naïve Bayes (NB), Boosting, and Bagged Tree were used for predicting the symptoms of ARIs among under-five children. We split the dataset into two datasets randomly, in which 80% of the data was used to train the model, and the remaining 20% was used to test the trained model. The model performance was evaluated using sensitivity, specificity, accuracy, and the area under the receiver operating characteristic curve (AU-ROC).

Results: A total of 327,507 under-five children were included in the study. The prevalence of ARI was highest in Mozambique (15.4%), Uganda (15.2%), Togo (14.2%), and Namibia (13.8 %), whereas Uganda (40.4%), Burundi (38.1%), Zimbabwe (37.4%), and Namibia (32.0%) had the highest prevalence of cough. The results of the random forest plot revealed that longitude, latitude, enhanced vegetation index, particulate matter, land surface temperature, nitrogen dioxide, and the number of cattle in the houses are the most important features in predicting the diagnosis of symptoms of ARIs among under-five children in sSA. The RF algorithm was selected as the best ML model (AUC=0.72, Accuracy=0.68) to predict the ARIs among children aged less than five.

Conclusions: Random forest machine learning algorithm (MLA) was classified as the best classifier to be employed for the prediction of the symptoms of ARI among under-five children.

Abstract

Background: Symptoms of Acute Respiratory infections (ARIs) among under-five children are a global health challenge. We aimed to train and evaluate the machine learning (ML) classification approaches in predicting symptoms of ARIs reported by mothers (caregivers) among children younger than five years in sub-Saharan African (sSA) countries.

Methods: We used the most recent (2012-2022) nationally representative Demographic and Health Surveys (DHS) data of 33 sSA countries. Machine learning algorithms such as logistic regression, LASSO, Elastic Net, Random Forest (RF), Artificial Neural Network (ANN), Naïve Bayes (NB), Boosting, and Bagged Tree were used for predicting the symptoms of ARIs among under-five children. We split the dataset into two datasets randomly, in which 80% of the data was used to train the model, and the remaining 20% was used to test the trained model. The model performance was evaluated using sensitivity, specificity, accuracy, and the area under the receiver operating characteristic curve (AU-ROC).

Results: A total of 327,507 under-five children were included in the study. The prevalence of ARI was highest in Mozambique (15.4%), Uganda (15.2%), Togo (14.2%), and Namibia (13.8 %), whereas Uganda (40.4%), Burundi (38.1%), Zimbabwe (37.4%), and Namibia (32.0%) had the highest prevalence of cough. The results of the random forest plot revealed that longitude, latitude, enhanced vegetation index, particulate matter, land surface temperature, nitrogen dioxide, and the number of cattle in the houses are the most important features in predicting the diagnosis of symptoms of ARIs among under-five children in sSA. The RF algorithm was selected as the best ML model (AUC=0.72, Accuracy=0.68) to predict the ARIs among children aged less than five.

Conclusions: Random forest machine learning algorithm (MLA) was classified as the best classifier to be employed for the prediction of the symptoms of ARI among under-five children.

CPS Abstract

Perception of insecurity in municipalities in Mexico. A Small Area Estimation approach.

Author: Mr MARIO SANTILLANA

Submission ID: 1970

Reference Number: 1970

Brief Description

In this work, the percentage of the population aged 18 years and over with perception of insecurity during March and April 2021 is estimated for each municipality in Mexico using Small Area Estimation techniques. Two methods are considered: the Empirical Best Linear Unbiased Predictor (EBLUP) and the Spatial Empirical Best Linear Unbiased Predictor (SEBLUP), both based on the Fay-Herriot area-level model. The National Survey of Victimization and Perception of Public Safety 2021 (ENVIPE 2021) is the base survey from which the variable object of estimation is obtained.

Abstract

The perception of insecurity is a paramount element in the study of crime, it measures the number of people who experience fear of being a victim of a crime; it is an important measure in decision-making involving public safety policies that allow the design, monitoring and evaluation of these programs. In Mexico, the National Survey of Victimization and Perception of Public Safety (ENVIPE) is an annual survey conducted by the National Subsystem of Information on Government, Public Safety and Law Enforcement (SNIGSPIJ) coordinated by the National Institute of Statistics and Geography (INEGI) that is aimed to collect information that allows the estimation of victimization and perception of insecurity in the place of residence at both national and state levels. However, local governments have increased their demand for reliable and official information at local levels, such as the municipal level. This disaggregation level is not considered in the survey design, which implies that in some municipalities the sample is null, or insufficient to provide estimates with acceptable coefficients of variation according to the reliability criteria considered by the INEGI. This, coupled with the lack of other information sources that satisfy this demand, leads to the implementation of methods to obtain this reliable information in such a way that the planned costs and resources are not altered. In this paper, the percentage of the population aged 18 years and over with perception of insecurity during March and April 2021 is estimated for each municipality in Mexico using Small Area Estimation techniques. Two methods are considered: the Empirical Best Linear Unbiased Predictor and the Spatial Empirical Best Linear Unbiased Predictor, both based on the Fay-Herriot area-level model.

CPS Abstract

Population Register

Author: Madam Lidija Brković

Submission ID: 1839

Reference Number: 1839

Files/Uploads

[2023-10-abstract-Population Registrar](#)

Brief Description

One of the prerequisites for creating a modern, digitised, and efficient public administration that provides quality and efficient services to both the economy and citizens is a system of interoperable databases and registers. In Croatia we are in the process of creating and establishing a Register of Population, Families and Households. This will be a single register in which data on population, families, and households as well as data on income and assets will be integrated. Once this register is operational, we will no longer need to carry out population censuses every decade and we will continually have access to up to date information on our citizens, which will provide relevant information for guiding policy and making good decisions. Bodies governed by public law will benefit from the Register within the scope of their work, as will citizens, in terms of having fewer administrative procedures when exercising their rights and fulfilling their obligations.

Before establishing the Population Register, a detailed analysis of the data to be included in the Register and how it would be organised, as well as a breakdown of who would manage the data was carried out. Work still needs to be done with respect to making appropriate amendments to the relevant acts and adopting subordinate legislation, as well as establishing a register of buildings and dwellings, so that an individual can be linked to a household.

The establishment of a high-quality and modern population register is a long-term project that requires resources and political will. However, once established, it will provide an invaluable resource for monitoring macroeconomic, socioeconomic, and demographic indicators in real time, which would enable public policies to be defined and implemented better.

Abstract

One of the prerequisites for creating a modern, digitised, and efficient public administration that provides quality and efficient services to both the economy and citizens is a system of interoperable databases and registers. In Croatia we are in the process of creating and establishing a Register of Population, Families and Households. This will be a single register in which data on population, families, and households as well as data on income and assets will be integrated. Once this register is operational, we will no longer need to carry out population censuses every decade and we will continually have access to up to date information on our citizens, which will provide relevant information for guiding policy and making good decisions. Bodies governed by public law will benefit from the Register within the scope of their work, as will citizens, in terms of having fewer administrative procedures when exercising their rights and fulfilling their obligations.

Before establishing the Population Register, a detailed analysis of the data to be included in the Register and how it would be organised, as well as a breakdown of who would manage the data was carried out. Work still needs to be done with respect to making appropriate amendments to the relevant acts and adopting subordinate legislation, as well as establishing a register of buildings and dwellings, so that an individual can be linked to a household.

The establishment of a high-quality and modern population register is a long-term project that requires resources and political will. However, once established, it will provide an invaluable resource for monitoring macroeconomic, socioeconomic, and demographic indicators in real time, which would enable public policies to be defined and implemented better.

CPS Abstract

Prevalence and Associated Factors of Preterm Birth Among Reproductive-Aged Women in Zambia: Examining Socioeconomic Inequalities.**Author:** Ms Mutale Sampa**Submission ID:** 1789**Reference Number:** 1789**Brief Description**

Leaving no one behind: Greater disaggregations and counting oversights

Abstract**Background**

Preterm birth, occurring between 20 and 37 weeks of pregnancy, poses a significant global public health challenge. An alarming 15 million babies are born prematurely each year, with over 90% of these occurrences afflicting Low- and Middle-Income Countries (LMICs). Preterm birth rates exhibit regional disparities, ranging from a mere 5% in developed nations to a staggering 18% in LMICs, peaking in Sub-Saharan Africa (SSA). Zambia annually witnesses around 77,600 preterm births. While interventions have targeted rural areas to narrow health disparities, urban areas have often been overlooked. This study seeks to quantify preterm birth distribution in both rural and urban settings, categorized by wealth quintiles, and evaluate disparities to pinpoint and address inequalities. To determine factors associated with preterm births in Zambia.

Methods

This cross-sectional study utilized 2018 Zambia Demographic and Health Survey data, focusing on women aged 15-49 who had given birth within five years before the survey (n=10,962). Employing survey analysis to account for the complex design, equity plots revealed regional and wealth quintile disparities. The survey's logistic regression identified factors associated with preterm births.

Results

Generally, urban areas had higher proportions of preterm birth than rural areas when desegregated by wealth quintile, highest among the poorest. In rural areas, the wealthiest had a higher proportion of preterm births. History of a terminated pregnancy, having ten children or more significantly increased the odds of preterm birth, while attending ≥ 4 ANC, and being older than 15-19 years significantly reduced the odds of preterm birth.

Conclusion

This study found that while improvements are seen in the rural areas, the disparities have reversed, and the urban areas are being left behind. The study also reviews that within the rural areas, the wealthiest are being left behind.

CPS Abstract

Prevalence for the Disclosure of HIV Status to Sexual Partners and Its Determinants among Adults under cART in Amhara Region, Northwest Ethiopia

Author: Prof. AWOKE SEYOUM TEGEGNE

Submission ID: 1830

Reference Number: 1830

Brief Description

Disclosure of the HIV status facilitates other behaviors that may improve the management of HIV.

Disclosing one's HIV status to sexual partners means speaking fairly about sexual alignment, possible drug-taking practice, and results of tests.

These are often unthinkable issues that are hard to talk about openly and honestly in most communities.

Abstract

Background: Globally, the transmission of HIV from one individual to another causes 1.8 million new infections each year, 36.7 million people living with HIV, and one million people died from HIV-related illnesses. The objective of this study was to determine the prevalence of the disclosure of HIV status to sexual partners and its determinants among adults under cART in the Amhara Region, northwest Ethiopia.

Methods: A retrospective study design was conducted on 792 randomly selected samples. The study was conducted in the Amhara Region, from 2015 to 2020. A binary logistic regression modeling was used for data analysis. The data were collected using a stratified random sampling technique where the residential areas were considered strata. Data were collected by trained health practitioners in the ART section in Felege-Hiwot Teaching and Specialized Hospital. The hospital is a referral in which many patients from different districts and zonal hospitals in the region are referred to this hospital.

Results: The rate of disclosure of HIV status to sexual partners in this study was 21%, which is very low compared to the average rate of disclosure in developing countries. Among the predictors, age of patients (AOR = 1.02, 95% CI: (1.001, 1.120); -value = 0.004); number of baseline CD4 cell count (AOR = 0.980; 95% CI: (0.764, 0.991); -value < 0.01); number of hospital visits (AOR = 1.01; 95% CI: (1.001, 1.034); -value < 0.01); marital status (living with partner) (AOR = 1.01; 95% CI: (1.003, 1.112); -value = 0.006); female HIV-positive adults (AOR = 1.01; 95% CI: (1.001, 1.021); -value = 0.007); rural residence (AOR = 0.98; 95% CI: (0.96, 0.99); -value = 0.004); non-educated adult patients (AOR = 0.950, 95% CI: (0.92, 0.98); -value = 0.003); cART non-adherent adult patients (AOR = 0.940, 95% CI: (0.61, 0.97); -value < 0.001); non-opportunistic infectious diseases (AOR = 1.062, 95% CI: (1.049, 1.191); -value = 0.002); and non-existence of social violence (AOR = 1.012, 95% CI: (1.008, 1.); -value < 0.01) significantly affected the variable of interest. Of these, the number of CD4 cell count, male HIV-positive adults, rural residence, and existence of social violence negatively affected the variable of interest.

Conclusions: Some groups of HIV patients did not disclose their level of HIV status to their sexual partners. Health-related education is recommended for patients who did not disclose their HIV status to sexual partners. This helps to reduce the transmission of HIV from infected individuals to non-infected ones and from mother-to-child HIV transmission.

CPS Abstract

Quality reports for statistical products - towards relevant and trustworthy official statistics in Kenya

Author: Geoffrey Kariuki

Submission ID: 2059

Reference Number: 2059

Brief Description

The term "official statistics" indicates some statistics should have an important position compared to other statistics within the national statistical system. Official statistics should be of high quality, objective, impartial and accessible to all users. However, there must be a deliberate effort to promote the use of official Statistics. Official Statistics are the basis of evidence-based planning, decision making and monitoring of policies and interventions. To promote the status of official statistics the National Statistics Office in Kenya has started to produce quality reports for statistical products classified as official statistics. A quality report is a documentation describing the quality of a product concerning all relevant quality dimensions and enables users to understand if the statistics are fit for purpose.

During the past couple of years, the Kenya National Bureau of Statistics (KNBS) has worked in cooperation with Statistics Sweden to achieve this outcome. In this presentation, KNBS will share their insights and experiences on the work undertaken to develop quality reports and templates and how these can be implemented among all producers of official statistics.

Through these efforts, the expectation is that official statistics in Kenya should improve its status as high-quality and relevant statistics for decision-makers to use but also to increase the possibilities to hold them to account.

Abstract

This presentation delves into the pivotal role of "official statistics" within the national statistical system, emphasizing their significance in evidence-based planning, decision-making, and policy monitoring. Official statistics, distinguished by their high quality, objectivity, impartiality, and accessibility, serve as the cornerstone for informed governance.

Kenya National Bureau of Statistics (KNBS), acknowledges the necessity of deliberate efforts to elevate the use of official statistics. Trust in official statistics is key for all users. In collaboration with Statistics Sweden, KNBS has undertaken a comprehensive initiative to produce quality reports for statistical products classified as official statistics. These reports serve as comprehensive documentation, outlining the quality dimensions of each product, ensuring fitness for purpose.

This presentation sheds light on the Kenyan experiences and insights in collaboration, focusing on the development of quality reports and templates. By sharing these practices, KNBS aims to foster a culture of high-quality reporting among all producers of official statistics in Kenya.

The envisioned outcome is an elevated status for official statistics, ensuring their relevance for decision-makers while enhancing accountability through transparent and reliable data. This experience will be of great value to other National Statistics Offices on how they can ensure official statistics remain relevant and trustworthy for all users.

CPS Abstract

Recent challenges in Labour Market Statistics and the role of the EU Labour Force Survey in user needs satisfaction

Author: Mrs Hanna Strzelecka

Submission ID: 1958

Reference Number: 1958

Brief Description

The proposal of the paper concerns the strand "Challenges and innovation in data production"

Abstract

The past years have been challenging for producers of Labour Market Statistics. New phenomena have appeared and gained in importance which should be taken into account in data production by official statistics. There are users who look for data such as digital platform employment, that is difficult to be captured by using traditional survey methods. COVID-19 measures have impacted statistical data collection, while at the same time the interest in high quality and timely labour market statistics has risen, and new indicators appropriate for the situation at hand were requested. The EU Labour Force Survey has been the most important source of data on the labour market, but the users still would like to add more topics to the EU LFS or receive more granular data. Moreover there are new ILO resolutions which should be implemented within the EU LFS. From the other hand response rate in the LFS in many countries is decreasing which only confirms that it is more difficult to conduct the survey.

The presentation will discuss above mentioned issues by trying to answer two below questions:

- 1) Is the EU Labour Force Survey able to deliver data on changing user needs?
- 2) How to find a balance between high quality data and response burden/LFS costs?

CPS Abstract

Respondent Driven Sampling strategies for hard-to-reach populations

Author: PROF. DR. Giorgio Alleva

Submission ID: 2106

Reference Number: 2106

Files/Uploads

[Abstract Alleva Falorsi ISI 2024](#)

Brief Description

The paper focuses on respondent-driven sampling (RDS), which is a valuable survey methodology to estimate the size and characteristics of hidden (e.g., homeless people, undocumented immigrants) or hard-to-measure population groups (e.g., minorities, indigenous people). The RDS methodology makes it possible to gather information on these populations by exploiting the relationships between their components. The estimation problem is addressed and three unbiased estimation methods are proposed.

Abstract

In this paper, we focus on respondent-driven sampling (RDS), which is a valuable survey methodology for both national and international organisations to estimate the size and characteristics of hidden (e.g., homeless people, undocumented immigrants) or hard-to-measure population groups (e.g., minorities, indigenous people).

The principle of "leaving no one behind" is at the heart of the 2030 Agenda, and a key requirement for many Sustainable Development Goals (SDG) indicators is to be available for the most vulnerable and marginalised population groups. Nevertheless, halfway through the implementation of the 2030 Agenda, most SDG indicators are still not available at the needed level of disaggregation to monitor the socioeconomic conditions of hidden and hard-to-count population groups. As a result, it is neither possible to produce reliable structural data on the needed disaggregation dimensions nor to monitor the developments of emerging phenomena that need to be approached with targeted evidence-based policy interventions.

The RDS methodology makes it possible to gather information on these populations by exploiting the relationships between their components. Moreover, the effectiveness of the RDS can be further increased by employing an integrated approach in which the RDS is used in conjunction with other information sources, such as administrative or geographical data.

In this paper, we address the estimation problem, and by approaching the RDS methodology as a particular indirect sampling technique, we propose three unbiased estimation methods as possible solutions. In particular, the first method assumes a random sampling of the initial individuals. In contrast, the second method, which considers purposive sample selection, creates a nonbiased estimation if the initial sample of respondents falls into all the clusters of networks that characterise the population of interest. Finally, leveraging the generalised capture-recapture estimation approach, we propose an estimator that accounts for the noncoverage of two independent indirect samplings.

CPS Abstract

Rethinking official statistics: a sociological perspective

Author: Dr Arman Bidarbakhtnia

Submission ID: 1948

Reference Number: 1948

Brief Description

Despite many progressive and people-centered initiatives for the development of official statistics (such as data values, inclusive data charter, data for now, and citizen data collaborative), the key global action plans for the development of statistics in the past two decades (Marrakech in 2004, Busan in 2011, and Cape Town in 2017) do not propose a single action aiming to increase citizens' participation, including their cognitive inputs, into the production of official statistics. This paper argues that the current language of "evidence-based decision-making" provides inadequate scope for the international statistical community to discuss solutions that give agency to people and non-state institutions to influence "what" statistics and "how" should be produced and disseminated on the social realities that they actively and persistently construct. The paper provides a sociological model for official statistics based on the social reality construction theory and describes how the new understanding of official statistics as social stabilizer can influence our approach towards the development and assessment of statistical systems.

Abstract

Despite many progressive and people-centered initiatives for the development of official statistics (such as data values, inclusive data charter, data for now, and citizen data collaborative), the key global action plans for the development of statistics in the past two decades (Marrakech in 2004, Busan in 2011, and Cape Town in 2017) do not propose a single action aiming to increase citizens' participation, including their cognitive inputs, into the production of official statistics. This paper argues that the current language of "evidence-based decision-making" provides inadequate scope for the international statistical community to discuss solutions that give agency to people and non-state institutions to influence "what" statistics and "how" should be produced and disseminated on the social realities that they actively and persistently construct. The paper provides a sociological model for official statistics based on the social reality construction theory and describes how the new understanding of official statistics as social stabilizer can influence our approach towards the development and assessment of statistical systems.

CPS Abstract

Risk estimation of bi-index stock portfolio using bivariate copula models with APTL marginals.

Author: Dr Charles Iwuji

Submission ID: 1983

Reference Number: 1983

Files/Uploads

[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 1](#)
[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 2](#)
[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 3](#)
[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 4](#)
[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 5](#)
[IAOS-ISI Conference 2024 _ Abstract 1 _ Table 6](#)

Brief Description

COLLABORATION ACROSS BOUNDARIES.

This strand seeks to demonstrate the needs and best practices in how the statistical researchers and the data community, in the public and private sectors, work together to improve statistical literacy and decision-making.

Abstract

Risk measurement and optimization remains an essential aspect of stock market investment as the risk averse investor always seek to know the risk of investment. In this study, the value-at-risk (VaR) is estimated for bi-index stock portfolio investment. The VaR is estimated using bivariate copula models with Alpha power transformed logistic (APTL) marginals in comparison to normal, student's t, Cauchy and logistic distributed marginals. The APTL distribution applied in this work for bi-index portfolio copula-VaR estimation has been used to estimate quantile value-at-risk for single stocks in the US stock markets but the bi-index portfolio case has not been considered. Copula is used to model the bi-index data. Copula bivariate APTL distribution and density functions that can be used to model different bivariate data sets in different fields like in economics and meteorology are derived using the Gaussian, Student's t, Clayton, Gumbel and frank copulas. In estimating the copula-value-at-risk for bi-index stock portfolio, some of the world's biggest stock indices are used. This includes the New York stock Exchange(NYSE), National Association of Securities Dealers Automated Quotations (NASDAQ) stock Exchange, Japan Stock exchange (JPX) and Shanghai stock exchange (SSE). The result of the copula model selection using AIC and BIC showed that the student's t copula is the most appropriate copula to model the NYSE-NASDAQ and SSE-JPX portfolios. The copula-VaR estimates of the NYSE-NASDAQ and SSE-JPX portfolios were then obtained using the student's t copula with APTL, APTL-logistic, normal, student's t, Cauchy and logistic marginal at 99%, 99.5% and 99.9% confidence levels. The results obtained from the data analysis showed that the student's t copula model with APTL marginal gives the minimum expected loss for equally weighted investment in the NYSE-NASDAQ and SSE-JPX portfolios at 99.5% and 99.9% confidence levels in comparison to the other marginals.

CPS Abstract

Spatial Analysis of Reading Culture Among Higher Education Students in Southwestern Nigeria

Author: Dr Isaac Ajao

Submission ID: 1795

Reference Number: 1795

Brief Description

In order to achieve the objective of the research, We wish to follow the following steps:

1.

Definition of Research Objectives:

Clearly statement of the objectives of the research is made.

What we want to discover or investigate regarding the reading culture among students of higher learning in Nigeria, what specific questions do we aim to answer

2.

Literature Review

We wish to conduct a comprehensive literature review to understand the existing research on reading culture among students in Nigeria, and identify gaps in the literature that our research can address.

3.

Data Collection

Determining the data collection methods to use.

This could involve surveys, interviews, observations, or a combination of these, and ensure that the data collection tools are well-structured and appropriate for the research.

4.

Hypothesis Testing

Formulate hypotheses related to how geographical location impacts reading culture among students.

For example, we might hypothesize that students in urban areas read more than those in rural areas.

5.

Data Visualization

Creation of maps, charts and graphs to present our spatial analysis results effectively.

Visualization can help convey the geographical differences in reading culture.

6.

Data Analysis

Decide on the spatial analysis techniques to use to analyze the data.

We will explore geographic information systems (GIS) or other spatial statistical methods to assess reading culture in different locations.

The following may be needed:

- Spatial autocorrelation to identify spatial patterns in reading culture.

- Spatial regression to assess the relationship between geographical location and reading habits.
- Spatial mapping to visually represent variations in reading culture across different locations.

7.

Interpretation and Discussion

We will present and discuss our findings.

Explain the significance of the results in the context of reading culture among Nigerian higher education students.

8.

Recommendations and Implications

We will provide recommendations for educational institutions, policymakers, and other stakeholders based on our findings.

9.

Conclusion

We shall summarize the key findings and the contributions of the research.

10.

Citations and References

We shall ensure we properly cite all the sources used in the literature review and throughout the paper.

11.

Conference Paper Structure

We will follow the guidelines provided by the conference organizers regarding paper lengthformatting, and any specific requirements.

Abstract

This study presents a spatial analysis of reading culture among students of higher learning in the southwestern region of Nigeria. The research aims to investigate how geographical location influences students' reading habits and, subsequently, their academic performance. By examining data from universities, polytechnics, and colleges of education, we assess factors such as the number of hours devoted to reading, book preferences, and student performance. Our findings reveal intriguing spatial patterns in reading culture, indicating variations across urban and rural areas and their impact on students' academic outcomes. We employ Geographic Information Systems (GIS) and spatial statistical methods to visualize these patterns and identify statistically significant relationships. The research not only offers an original perspective on reading culture but also holds practical implications for educational institutions and policymakers. By pinpointing areas with distinct reading culture characteristics, we provide insights that can inform targeted interventions to enhance literacy and academic achievements in higher education. This study showcases the interdisciplinary potential of merging geography and education, shedding light on regional disparities that may otherwise remain unexplored. It is belief that our research contributes to the international dialogue on educational development and provides a model for similar studies in diverse geographic contexts.

CPS Abstract

Spatial pattern and predictors of malaria in Ethiopia: Application of auto logistics regression

Author: Melkamu A Zeru

Submission ID: 1820

Reference Number: 1820

Brief Description

Malaria is a serious health problem in Ethiopia, affecting the socio-economic and health status of the country at large.

Malaria outpatient caseloads may vary several-fold from year to year and season to season in unstable epidemic-prone transmission pattern.

Peak malaria transmission occurs season between September and December in most Ethiopia, after the main summer season from June to August.

Abstract

Introduction: Malaria is a severe health threat in the World, mainly in Africa. It is the major cause of health problems in which the risk of morbidity and mortality associated with malaria cases are characterized by spatial variations across the county. This study aimed to investigate the spatial patterns and predictors of malaria distribution in Ethiopia.

Methods: A weighted sample of 15,239 individuals with rapid diagnosis test obtained from the Central Statistical Agency and Ethiopia malaria indicator survey of 2015. Global Moran's I and Moran scatter plots were used in determining the distribution of malaria cases, whereas the local Moran's I statistic was used in identifying exposed areas. The auto logistics spatial binary regression model was used to investigate the predictors of malaria.

Results: The final auto logistics regression model was reported that male clients had a positive significant effect on malaria cases as compared to female clients [AOR = 2.401, 95% CI: (2.125– 2.713)]. The distribution of malaria across the regions was different. The highest incidence of malaria was found in Gambela [AOR = 52.55, 95%CI: (40.54–68.12)] followed by Bene shangul [AOR = 34.95, 95%CI: (27.159–44.963)]. Similarly, individuals in Amhara [AOR = 0.243, 95% CI:(0.195–0.303], Oromiya [AOR = 0.197, 95% CI:(0.158–0.244)], Dire Dawa [AOR = 0.064, 95%CI(0.049–0.082)], Addis Ababa[AOR = 0.057,95%CI:(0.044–0.075)], Somali[AOR = 0.077,95%CI:(0.059–0.097)], SNNPR[OR = 0.329, 95%CI: (0.261–0.413)] and Harari [AOR = 0.256, 95%CI:(0.201–0.325)] were less likely to had low incidence of malaria as compared with Tigray. Furthermore, for one meter increase in altitude, the odds of positive rapid diagnostic test (RDT) decrease by 1.6% [AOR = 0.984, 95% CI: (0.984– 0.984)]. The use of a shared toilet facility was found as a protective factor for malaria in Ethiopia [AOR = 1.671, 95% CI: (1.504–1.854)]. The spatial autocorrelation variable changes the constant from AOR = 0.471 for logistic regression to AOR = 0.164 for auto logistics regression.

Conclusions: This study found that the incidence of malaria in Ethiopia had a spatial pattern which is associated with socio-economic, demographic, and geographic risk factors. Spatial clustering of malaria cases had occurred in all regions, and the risk of clustering was different across the regions. The risk of malaria was found to be higher for those who live in soil floor-type houses as compared to those who lived in cement or ceramics floor type. Similarly, households with thatched, metal and thin, and other roof-type houses have a higher risk of malaria than ceramics tiles roof houses. Moreover, using a protected anti-mosquito net was reducing the risk of malaria incidence.

CPS Abstract

Statistical Population Register: from needs to goals and implementation

Author: Dr Inga Masiulaityte-Sukevic

Submission ID: 1856

Reference Number: 1856

Brief Description

Statistics Lithuania (State Data Agency) has been producing statistical information using administrative data for more than two decades. 2021 Lithuanian Population and Housing Census was conducted based on administrative sources. Users nowadays require to provide more frequent, more timely and more detailed statistics about population.

The need to harmonise and standardise demographic and social statistics, to satisfy user needs for having more data, achieved good quality of the administrative data had led to have an individualised statistical data system as repository of up-to-date information of characteristics of all residents of the country. The goal of creation of Statistical Population Register is very important for the future statistics and modernisation of registers of the country.

Statistical Population Register (hereinafter referred to SPR) was started to create in 2021. The establishment of SPR was influenced by number of different factors: starting from having register-based Census database, State Data Governance Information System and ending capability to involve of high-quality specialists in establishment process of SPR.

Theoretical and practical opportunities and challenges of creation of Statistical population register will be presented.

Abstract

Statistics Lithuania (State Data Agency) has been producing statistical information using administrative data for more than two decades. 2021 Lithuanian Population and Housing Census was conducted based on administrative sources. Users nowadays require to provide more frequent, more timely and more detailed statistics about population. The need to harmonise and standardise demographic and social statistics, to satisfy user needs for having more data, achieved good quality of the administrative data had led to have an individualised statistical data system as repository of up-to-date information of characteristics of all residents of the country. The goal of creation of Statistical Population Register is very important for the future statistics and modernisation of registers of the country.

Statistical Population Register (hereinafter referred to SPR) was started to create in 2021. The establishment of SPR was influenced by number of different factors: starting from having register-based Census database, State Data Governance Information System and ending capability to involve of high-quality specialists in establishment process of SPR.

Theoretical and practical opportunities and challenges of creation of Statistical population register will be presented.

CPS Abstract

Statistical support for citizens as decision-makers in Poland

Author: Ms Dominika Rogalińska

Submission ID: 1884

Reference Number: 1884

Brief Description

In an era flooded with information, access to reliable and accurate data and understanding how to interpret and apply statistical data is crucial for every citizen. This presentation highlights innovative approaches and best practices in providing statistical support. I will navigate the evolving landscape of official statistics in Poland, empowering citizens and shaping a more data-informed society. Join to explore how statistical insights drive informed decisions for Polish citizens.

Abstract

In an era flooded with information, access to reliable and accurate data and understanding how to interpret and apply statistical data is crucial for every citizen. This presentation highlights innovative approaches and best practices in providing statistical support. I will navigate the evolving landscape of official statistics in Poland, empowering citizens and shaping a more data-informed society. Join to explore how statistical insights drive informed decisions for Polish citizens.

An example of dedicated tool for local governments is Public Services Monitoring System (SMUP). The purpose of SMUP is to provide local government units, businesses and the public with information necessary for a comprehensive assessment of services provided at the local level. The base of this information are the resources of public statistics as well as the data collected by the administration in many dispersed sources. The system is intended to organise them and enable access to high quality data on public services for all interested users of the system in the form of indicators presented at gmina level. The long-term goal of building the SMUP is optimising the supply of services that public administrations are responsible for delivering, based on integrated, high-quality data.

CPS Abstract

THE USE OF SCANNER DATA IN OFFICIAL STATISTICS

Author: Mr TOMAS RUDYS

Submission ID: 2105

Reference Number: 2105

Files/Uploads

[Rudys_Tomas_LT_abstract](#)

Brief Description

During past years, availability of new or alternative data sources showed wide range of possible applications in different official statistics topics. In price statistics, scanner data from private data owners is of high importance for future improvements of HICP calculations. Nevertheless, the application or integration of these kind of data into production process requires detailed analysis and preparation work. In the presentation we will give a short overview of process integrating scanner data, automated quality checks, classification process. We will present some results on application of machine learning methods for classification of data.

Abstract

During past years, availability of new or alternative data sources showed wide range of possible applications in different official statistics topics. In price statistics, scanner data from private data owners is of high importance for future improvements of HICP calculations. Nevertheless, the application or integration of these kind of data into production process requires detailed analysis and preparation work. In the presentation we will give a short overview of process integrating scanner data, automated quality checks, classification process. We will present some results on application of machine learning methods for classification of data.

CPS Abstract

Testing Enablement with Paper Forms versus Internet Code in New Zealand's 2023 Census

Author: Dr Betsy Williams

Submission ID: 1918

Reference Number: 1918

Brief Description

New Zealand's 2023 Census tailored its field collection strategies to different areas. Using a cluster-randomised controlled trial, we test response mode, timeliness, and quality after initially mailing internet access codes versus paper forms with codes. We also test whether our measures of paper response and non-response, based on 2018 Census patterns, sent paper forms to the right places. We briefly discuss how a cyclone affected planned collections.

Abstract

New Zealand's 2023 Census tailored its field collection strategies to different areas. Using a cluster-randomised controlled trial, we test response mode, timeliness, and quality after mailing letters with internet access codes (IACs) versus paper forms with IACs. We test whether our measures of paper response and non-response, based on 2018 Census patterns, deployed paper correctly. Note: results too disclosive of unreleased census data will be redacted.

Context: About 70% of private dwellings in the country are in the mailout stream. We mail each dwelling a letter with an IAC, with or without a pack of paper forms. Anyone could request paper forms, and later processes gave a mixture of paper and IACs.

Enablement mode: In mailout, 20% were enabled with paper; 80%, without (IAC-only). To assign meshblock areas to an enablement mode, we developed a "paper response index" and "non-response index": areas above a cutoff on either scale were assigned to paper enablement.

Randomised enablement: Did paper enablement affect response mode (paper or online), self-response, or response quality? An observational study would lack statistical power. Instead, after ethical analysis, we randomised about 3% of mailout areas, those with scores near the cutoff on one or both indices.

Analysis 1: We use GLMMs to test the effects of paper enablement on responding on paper, self-response received before non-response follow-up, and response quality. We use a Holm-Bonferroni correction for multiplicity.

Analysis 2: We use a multi-rating regression discontinuity design to estimate those three effects while incorporating information from the wider mailout stream. These analyses also show how well the paper response index predicted response on paper and how well the non-response index predicted non-response, controlling for enablement mode.

Complications: A devastating cyclone hit the week before initial mailout. Census was postponed in the hardest hit regions, which are removed from these analyses.

Strand: As this deals with public capacity for online survey response in a post-pandemic world, this may be most appropriate in strand 1 (Challenges and innovations in data production) or 3 (The new normal).

This is joint work from a group at Statistics New Zealand.

CPS Abstract

The Community and Individual Well-Being Interaction in Alternative Modelling Approaches Using Spatial Data.

Author: Prof. Włodzimierz Okrasa

Submission ID: 2100

Reference Number: 2100

Brief Description

The problem of mutual influence between the levels of local community welfare and individual well-being (including measures of subjective well-being) - which is interesting from the point of view of local development policies - poses several challenges to public statistics and to the task of selection the appropriate, problem-specific analytical strategy. The first concerns the availability of spatial data from various sources (administrative and survey) – such as used in this study: Local Data Bank, EU-SILC, Time Use Survey, Social Diagnosis, Social Cohesion Survey – which have been integrated in a multi-source analytical database. using a territorial code as 'integrator'. The latter includes a review of available models motivated by the need to select a specific (relatively 'best') analytical strategy, making the selection criterion - formally its methodological effectiveness - dependent on substantively important aspects of the modeled interaction. Hence, the 'evaluation' of each of the rich set of approaches demonstrated in the empirical part of this study - from descriptive ones, such as Benefit Incident Analysis (in the Marginal BIA version) and Correspondence Analysis through Mahalanobis Distance and Compensating Variations to the analysis of spatial dependence and multilevel spatial modeling - is supplemented with the specification of the included to the appropriate equations contextual factors which are (hypothetically) modifying the modeled relationship between individual and community measures of well-being. Such as the profile of local risk factors, the presence (activity) of the third sector entities (NGOs and 'social capital'), homogeneity-heterogeneity of the community (including differences in the wealth of residents), etc. In summary, an analytical framework is proposed in the form of a matrix approach to the types of determinants of interaction (taken into account in the appropriate hypotheses) and the types of analytical strategies suitable for its modeling in specific problem contexts.

Abstract

The problem of mutual influence between the levels of local community welfare and individual well-being (including measures of subjective well-being) - which is interesting from the point of view of local development policies - poses several challenges to public statistics and to the task of selection the appropriate, problem-specific analytical strategy. The first concerns the availability of spatial data from various sources (administrative and survey) – such as used in this study: Local Data Bank, EU-SILC, Time Use Survey, Social Diagnosis, Social Cohesion Survey – which have been integrated in a multi-source analytical database. using a territorial code as 'integrator'. The latter includes a review of available models motivated by the need to select a specific (relatively 'best') analytical strategy, making the selection criterion - formally its methodological effectiveness - dependent on substantively important aspects of the modeled interaction. Hence, the 'evaluation' of each of the rich set of approaches demonstrated in the empirical part of this study - from descriptive ones, such as Benefit Incident Analysis (in the Marginal BIA version) and Correspondence Analysis through Mahalanobis Distance and Compensating Variations to the analysis of spatial dependence and multilevel spatial modeling - is supplemented with the specification of the included to the appropriate equations contextual factors which are (hypothetically) modifying the modeled relationship between individual and community measures of well-being. Such as the profile of local risk factors, the presence (activity) of the third sector entities (NGOs and 'social capital'), homogeneity-heterogeneity of the community (including differences in the wealth of residents), etc. In summary, an analytical framework is proposed in the form of a matrix approach to the types of determinants of interaction (taken into account in the appropriate hypotheses) and the types of analytical strategies suitable for its modeling in specific problem contexts.

CPS Abstract

The SORS Decision-Making Support System

Author: Mr Branko Josipovic

Submission ID: 1939

Reference Number: 1939

Brief Description

The paper explores Decision-Making Support System (DMSS), a strategic response to complex data challenges. DMSS, encompassing government modules and the Labour Market Intelligence System, enhances decision-makers' capacity through tailored information. The system ensures data stewardship, legal compliance, and promotes statistical literacy, demonstrated by its adaptability during the COVID-19 pandemic and the 2022 Population Census.

Abstract

In the dynamic landscape of a data-rich world, the Statistical Office of the Republic of Serbia (SORS) has developed the Decision-Making Support System (DMSS) as a strategic response to the challenges presented by a complex data ecosystem. This paper delves into how DMSS plays a pivotal role in addressing the multifaceted interactions within the official statistics system, with a specific focus on capacity building, data dissemination, and statistical literacy.

DMSS, comprising government and local-level modules along with the Labour Market Intelligence System, actively contributes to the capacity building of decision-makers by providing tailored information and indicators. The system's innovative approach ensures efficient data stewardship and compliance with legal frameworks, navigating the intricacies of data sharing.

Moreover, the paper underscores DMSS's commitment to promoting statistical literacy, empowering users to make informed decisions. Success stories and practical applications during the COVID-19 pandemic and the Population Census of 2022 demonstrate how DMSS adapts to post-pandemic conditions, utilizing alternative data sources and cutting-edge technologies.

The DMSS aligns with the evolving landscape of a data-driven society, ensuring adaptability and optimal decision-making. This paper provides an in-depth exploration of the system's components, functionality, and practical applications, showcasing its role in advancing statistical support for better policy outcomes in the Republic of Serbia and fostering a new era of statistical literacy and collaboration.

CPS Abstract

The elimination of infectious diseases caused by climate change in the United States

Author: Mr Said Mohammed

Submission ID: 2062

Reference Number: 2062

Brief Description

This session will explore the best practices for fulfilling the 2030 Agenda for Sustainable Development.

Abstract

For over two decades, researchers have delved into understanding the consequences of observed and anticipated shifts in weather and climate on the extent and nature of health outcomes that are sensitive to climate change. Warmer temperatures, mild winters, and early springs are providing mosquitoes and ticks with extended periods for reproduction, leading to the expansion of their habitats across the United States. This phenomenon is contributing to the increased prevalence of diseases such as Lyme disease, West Nile virus, and Valley fever. Between 2004 and 2018, a notable surge in infectious diseases has been observed, with these illnesses spreading to new regions of the United States. The reported cases of illnesses resulting from mosquito, tick, and flea bites more than doubled during this period, surpassing 760,000 reported cases in the United States. During this period, nine new pathogens transmitted by mosquitoes and ticks were identified or introduced into the United States. This paper has three primary objectives. Firstly, it seeks to conceptualize the potential direct and indirect health impacts of climate change. Additionally, it offers an overview of factors that amplify the health effects associated with climate change. Secondly, the paper compiles and reviews existing literature concerning the correlation between infectious diseases and climate change. It delves into potential strategies for the eradication of these diseases by the year 2030. Lastly, the paper explores the impact of climate change on the dissemination of the West Nile virus outbreak in the United States in 1999. This paper yields two significant findings. Firstly, it demonstrates that climate change is a causal factor for various diseases and intensifies their impact. Secondly, it underscores the most adverse health consequences related to the increased spread of infectious diseases. Specifically, the occurrence of hot summers is identified as a precursor to the outbreak of West Nile virus disease in the United States, leading to widespread illnesses and fatalities across the nation.

CPS Abstract

The employability of the immigration from South and Central America to Mexico in 2023

Author: Miss Berenice Rodríguez Tovar

Submission ID: 1964

Reference Number: 1964

Brief Description

The paper has two chapters: first, a brief historical context about the immigration from South and Central America to Mexico, and the theoretical framework. The second one, is the proposal of the paper that is the factors that determine the employability of the immigration. In this part is included the statistical method with the results and interpretation.

Abstract

This paper describes the employability of immigration from different countries in South and Central America to Mexico and the sociodemographic, socio-occupational, and territorial heterogeneity. The primary data source of information is the 2023 National Survey of Occupation and Employment (ENOE). The quantitative technique used to analyze the data is logistical regression models. The results demonstrate an increase in the volume of immigration from South and Central America to Mexico, differentiated patterns of employability for the birth country and region, and emerging demographic changes in the recent arrival of immigrants that detonate a significant heterogeneity. Although the concentration of the immigrants from Guatemala, Honduras, and El Salvador is, in relative terms, on no modern subsectors of the tertiary activities, there is a certain grade of specialization according to the country of origin: 1) more concentration of the Guatemalans and Hondurans in the agriculture; 2) and the Salvadorans in the tertiary sector (mainly in no modern subsectors). Conversely, South American immigrants in Mexico are concentrated in the modern subsector of tertiary activities. The specialization according to the country of origin is Colombians, Cubans, and Venezuelans in modern subsectors, such as professional services. The adjustment of the logistical regression models for the population and differentiated group reaffirms the importance of the local labor markets for increasing the possibility of getting a formal job, as well as the branches of economic activity, for instance, residing in the northern part of the country, in the first place, or the center, in second place. In addition, if these regions have an incidence of manufacturing industry or modern services, improve the opportunity to get a better job. This paper aims to shed light to the extensive field of immigration and provide findings that help to understand its dynamic better in the frame of the principal two of the universal values: leave no one behind.

CPS Abstract

The evolution of the spatial data production model in Istat. New perspectives for the analysis of socio-economic phenomena of the population.

Author: Dr Giancarlo Carbonetti

Submission ID: 1984

Reference Number: 1984

Files/Uploads

[IAOS_ISI_2024_Abstract_Giancarlo_Carbonetti](#)

Brief Description

The work represents the National Statistical Institute's (Istat) response to new challenges in the production of spatial data. There is a growing need for accurate, timely and geo-referenced data at small spatial domains on demographic, social and economic issues. Such information is essential to study the main phenomena characterising the dynamics of resident population and households at the highest level of spatial detail.

Methodological advances and the development of IT solutions have facilitated the integration of administrative data with those of sample surveys, georeferencing down to the address of residence and the construction of the integrated system of statistical registers.

The new production processes of Istat's official statistics are therefore characterised by highly innovative elements.

The results will offer the possibility of producing spatial indicators with a high information content useful for conducting spatial and temporal analyses as never before.

Abstract

Over the years, official statistics have shown increasing attention to the strong need for statistical information referring to sub-municipal levels. These data are useful for spatial analyses, business objectives and social, economic and environmental decision-making processes. In particular, users demand data with greater frequency, consistency and richness of information detail.

In this context, the Italian National Statistical Institute (Istat) has been engaged in a modernisation process that includes a significant revision of the production processes for official statistics. In particular, two important pillars of Istat's new organisational and production structure should be mentioned: the construction of statistical registers and the Integrated Register System (IRS); the design and implementation of the Permanent Population and Housing Census (PPHC).

The development of statistical methodologies and the adoption of advanced IT solutions represented the main topic to integrate administrative data with the results of sample surveys and to geocode information at a very fine level over the territory (enumeration areas or addresses). This has made it possible to define an integrated and georeferenced database for the territory, rich in annually renewable information on the main demographic, social and economic variables of individuals, household types and some characteristics of dwellings.

The new process meets high quality standards, allows for a high timeliness in the release of outputs and guarantees maximum data consistency. The results offer new perspectives for analysing the main socio-economic phenomena of the population at sub-municipal scale, both on spatial and temporal dimensions. It is possible to focus attention on specific issues (e.g. "household deprivation") through the definition and calculation of appropriate indicators.

Istat is also planning the release of sub-municipal data and the design of a user-friendly platform for the dissemination and representation of results using GIS tools.

CPS Abstract

The impact of new technologies on remote work and productivity in post pandemic era

Author: Dr Hope Ifeyinwa Mbachu

Submission ID: 2013

Reference Number: 2013

Brief Description

The Impact of New Technologies on Remote Work and Productivity in the Post-Pandemic Era

This strand will review how statistical offices and data users have adjusted to post pandemic condition.

Alternative data sources and new technology advancement to post Covid 19.

Abstract

Title: The Impact of New Technologies on Remote Work and Productivity in the Post-Pandemic Era

Authors: 1) Mbachu Hope I.; Dept of Statistics, Imo State University Owerri, Imo State Nigeria. Email: ifyjoe100@yahoo.com Phone +2348034732597

2) Onuoha Petrinus E.; Dept of Math/Statistics, Federal Polytechnic Nekede Owerri, Imo State Nigeria Email: petrinus@fida-ng.com Phone +2347037092612

Abstract:

The onset of the COVID-19 pandemic reshaped work dynamics worldwide, accelerating the adoption of remote work and ushering in an era where technological advancements play a pivotal role in shaping work environments. This study delves into the multifaceted impact of new technologies on remote work and productivity in the post-pandemic landscape. First this paper explores the fundamental shifts in workplace structures catalyzed by the pandemic, leading to a widespread implementation of remote work models. It then navigates through an intricate analysis of the technological tools and platforms that have emerged or evolved, examining their influence on enhancing connectivity, collaboration, and task management in remote settings. The study investigates the varying implications of these technologies on productivity across different industries, considering the challenges and opportunities they present. It scrutinizes the socio-psychological impacts on employees, evaluating the effects of technology-driven remote work on work-life balance, mental health, and job satisfaction. This research presents case studies and empirical evidence showcasing successful adoptions of innovative technologies and best practices, highlighting their role in augmenting productivity and efficiency in remote work setups. This study aims to provide a comprehensive understanding of the symbiotic relationship between new technologies, remote work dynamics, and productivity in the post-pandemic era, offering insights, decisions and recommendations for businesses, policymakers, and individuals navigating this evolving landscape

Keywords: Remote work, Technological advancement, Productivity, Post-pandemic, Job

CPS Abstract

The role of official statistics in individual decision making: A pathway to the public good

Author: Ms Sofi Nickson

Submission ID: 1835

Reference Number: 1835

Brief Description

Official statistics are essential tools for government, policy makers, and organisations across civil society.

This presentation, however, looks beyond these widely-explored uses, focussing on the often underappreciated yet still valuable role of official statistics in the hands of the public.

Based on research exploring the role of statistics in decision making by individual citizens, we investigate how statistics empower individuals to make informed choices.

These findings can help statistics and data become a powerful resource for all.

Abstract

Official statistics have long been recognised as essential tools for government, policy makers, and organisations across civil society. However, when considering how to improve decision making for all we cannot stop there. This presentation looks beyond the widely-explored government and organisational uses of official statistics. Here, we focus on the often underappreciated yet still valuable role of official statistics in the hands of the public.

At the UK's Office for Statistics Regulation, our vision is that official statistics will serve the public good. As such, we have a research programme seeking to deepen understanding in this area. To truly serve the public good, we have seen that official statistics need to be for everyone, which is why we are sharing in this presentation our research project exploring the role of statistics in decision making by individual citizens. We investigate how statistics empower individuals to make informed choices in their personal lives, and the findings of this work can be used widely to help official statistics become a powerful resource for all.

We invite you to join our presentation and explore the relationship between official statistics and individual decision making, ushering in a new area of data-driven empowerment and serving the public good.

CPS Abstract

Transforming and Modernizing National Statistics Systems: Leaving No One Behind

Author: Mr Kabeli Mefane

Submission ID: 1896

Reference Number: 1896

Brief Description

In today's data-driven world, national statistics systems play a crucial role in shaping policy decisions and addressing socio-economic challenges. The abstract will highlight the importance of transforming and modernizing these systems to ensure that no one is left behind in the pursuit of sustainable development and inclusive governance.

Abstract

As societies evolve and become more complex, the traditional methodologies of data collection and analysis are often insufficient to capture the diverse and dynamic needs of their populations. This transformation entails a multifaceted approach, involving the integration of new technologies, the adoption of innovative data collection methods, and a renewed commitment to inclusivity and accessibility. As the landscape of data collection evolves, leveraging the potential of big data, artificial intelligence, and data analytics can enhance the accuracy, timeliness, and relevance of statistical information. Furthermore, the integration of geospatial and satellite data can provide valuable insights into the geographical distribution of resources and needs. Ensuring that no one is left behind in the context of national statistics systems implies a focus on vulnerable and marginalized populations, such as rural communities, ethnic minorities, and people with disabilities. This calls for the development of inclusive data collection strategies that are sensitive to the specific challenges these groups face. These strategies should respect privacy, consent, and ethical considerations. Modernization efforts should also emphasize the importance of data dissemination and communication. Making statistical information readily available and easily understandable to the public can foster informed community and accountability. Open data platforms and user-friendly interfaces can bridge the gap between statistics producers and consumers, facilitating data-driven decision-making at all levels of governance.

CPS Abstract

UK Public Services Productivity Review –early achievements, key challenges and plans for shaping the future of Public Service measurement in the UK Na

Author: Ms Jessica Barnaby

Coauthors: Debra Prestwood

Submission ID: 1891

Reference Number: 1891

Brief Description

The ONS is reviewing how we measure public service productivity. We are partnering with Government departments to develop methodology and data sources to drive forward improved metrics. We would like to take this opportunity to introduce the review to the international community and begin the discussion around the key questions we are trying to answer.

Abstract

The ONS is reviewing how we measure public service productivity. We are partnering with Government departments to develop methodology and data sources to drive forward improved metrics. We would like to take this opportunity to introduce the review to the international community and begin the discussion around the key questions we are trying to answer.

Measuring public services productivity is challenging because there is no direct charge for services that are delivered – this makes measuring productivity and any changes to it tricky. In addition to reviewing inputs and outputs, the key factors in determining productivity growth, we are also taking into account the quality of the service through making quality adjustments. Making quality adjustments to the methodology will help produce more robust and consistent outputs and outcomes of the services provided. Better methods will make it easier to answer questions like:

- How much improvement to pupil outcomes are schools delivering?
- How has the quality of social security administration changed over time?
- What impact has the Coronavirus (COVID-19) pandemic had on health services?

In this session, we will tell the story of the review. We will explain the history of measuring productivity in the UK and the steps we are taking to improve this. We will go into detail on the challenges we have faced up to now and how we plan to take the project forward, through implementation of new methods in the UK National Accounts.

CPS Abstract

Understanding Data Stewardship and role of NSS in Africa**Author:** Dr Anjana Dube**Submission ID:** 2067**Reference Number:** 2067**Brief Description**

In 2021 the UN working group on data stewardship was established and is lead by UNSD. The work plan of the working group covers the development of a conceptual framework for data stewardship and its various dimensions. The draft report of the working group has developed and circulated questionnaire to examine the data stewardship developments from the NSOs perspectives on the context, scope and priorities of data stewardship to further develop the conceptual framework on data stewardship. The working group report will be submitted to the 55th Statistical Commission in its meeting in Feb-Mar2024.

As far as Africa is concerned various initiatives to use technology and build the digital data collection mechanism have been undertaken by the countries supported by their own governments, stakeholders and donors. While UNECA has a roadmap for transformation and modernisation of official statistical systems, what is lacking perhaps is an overall national level framework to enhance data driven mechanism and governance. The debate is on the exact role and definitions of data stewards, which is best agency/agencies to take the role of data steward. The data steward has responsibility throughout the data life cycle, from collection to dissemination and preservation of the data. Thus, data governance models become fundamental and depend on the role that NSOs envision and the complexity of the national statistical system (NSS) and other data producers. This paper will examine these developments in the context of the working groups report and look at the way forward in developing National Statistical Systems in Africa.

Some of the stewardship roles of NSO's relate to the Management of integrated databases, the quality framework for administrative registers, standardization of meta-data, development of common data handling protocols, access to open data, common data dissemination platforms, integrating datasets by creating common identifiers common identifiers and classification for the use of data for governance purposes.

The paper aims at helping the participants to understand the developments in UNSC in building framework including definitions and concepts for data stewardship and the role to be played by the NSOs within the larger data ecosystem as the data steward and understanding the scope of data stewardship and extent to which these can be embedded in the National data strategy and way forward for NSS specially in Africa.

Presenters: Ian Rutherford, UNSD and Anjana Dube, UNECA

Abstract

In 2021 the UN working group on data stewardship was established and is lead by UNSD. The work plan of the working group covers the development of a conceptual framework for data stewardship and its various dimensions. The draft report of the working group has developed and circulated questionnaire to examine the data stewardship developments from the NSOs perspectives on the context, scope and priorities of data stewardship to further develop the conceptual framework on data stewardship. The working group report will be submitted to the 55th Statistical Commission in its meeting in Feb-Mar2024.

As far as Africa is concerned various initiatives to use technology and build the digital data collection mechanism have been undertaken by the countries supported by their own governments, stakeholders and donors. While UNECA has a roadmap for transformation and modernisation of official statistical systems, what is lacking perhaps is an overall national level framework to enhance data driven mechanism and governance. The debate is on the exact role and definitions of data stewards, which is best agency/agencies to take the role of data steward.

This paper will examine these developments in the context of the working groups report and look at the way forward in developing National Statistical Systems in Africa.

The paper aims at helping the participants to understand the developments in UNSC in building framework including definitions and concepts for data stewardship and the role to be played by the NSOs within the larger data ecosystem as the data steward and understanding the scope of data stewardship and extent to which these can be embedded in the National data strategy and way forward for NSS specially in Africa.

CPS Abstract

Unified Robust Estimation

Author: PROF. DR. Zhu Wang

Submission ID: 1892

Reference Number: 1892

Brief Description

Foundational work in statistics and data science: Better understanding and innovation in statistical research and decision-making

This strand aims to provide the platform for statistical researchers to exchange ideas and disseminate results and to promote innovative research in statistics and data science.

Abstract

Robust estimation is primarily concerned with providing reliable parameter estimates in the presence of outliers. Numerous robust loss functions have been proposed in regression and classification, along with various computing algorithms. In modern penalized generalized linear models (GLM), however, there is limited research on robust estimation that can provide weights to determine the outlier status of the observations. This article proposes a unified framework based on a large family of loss functions, a composite of concave and convex functions (CC-family). Properties of the CC-family are investigated, and CC-estimation is innovatively conducted via the iteratively reweighted convex optimization (IRCO), which is a generalization of the iteratively reweighted least squares in robust linear regression. For robust GLM, the IRCO becomes the iteratively reweighted GLM. The unified framework contains penalized estimation and robust support vector machine and is demonstrated with a variety of data applications.

CPS Abstract

Unmasked Realities: Unraveling the Impact of COVID-19 on SDG 3 Health and Wellbeing Indicators in Africa

Author: Mr Kossi Edem Kludza

Submission ID: 1911

Reference Number: 1911

Brief Description

Introduction: Amidst the COVID-19 pandemic, which severely impacted healthcare systems in developed countries and strained the unprepared and fragile health systems of many African nations, considerable concern emerged regarding the potential consequences for SDG 3 (Health and Wellbeing) indicators. This study aims to investigate whether the COVID-19 pandemic has indeed significantly impacted SDG 3 indicators in Africa.

Methods: This paper used a dual approach of desk review and empirical analysis. It synthesized information from scientific journals and renowned institutions working on SDG implementation, such as the United Nations System. The empirical analysis utilized a statistical inference (Wilcoxon signed-rank test) to assess any significant differences in the rate of progress pre- and post-COVID-19. Data was sourced from the United Nations Global SDG database.

Results: Out of the 19 indicators examined, the findings reveal that only five indicators were significantly impacted: Maternal mortality (SDG 3.1.1), Neonatal mortality (SDG 3.2.2), Prevalence of hepatitis B (SDG 3.3.4), Vaccination Programmes (SDG 3.b.1), and Health worker density (SDG 3.c.1). It is noteworthy that the situation is not homogenous across countries.

Conclusion: Contrary to initial expectations, the available data suggest that the COVID-19 pandemic did not have the anticipated devastating impact on SDG 3. In the short term, the continent seemed equipped to respond to the pandemic, resulting in the limited impact observed in this study. A wider availability of data across the countries and indicators would have provided a deeper understanding of the situation.

Keywords: SDG 3, Health and wellbeing, Africa, COVID-19, SDG Progress

Abstract

Introduction: Amidst the COVID-19 pandemic, which severely impacted healthcare systems in developed countries and strained the unprepared and fragile health systems of many African nations, considerable concern emerged regarding the potential consequences for SDG 3 (Health and Wellbeing) indicators. This study aims to investigate whether the COVID-19 pandemic has indeed significantly impacted SDG 3 indicators in Africa.

Methods: This paper used a dual approach of desk review and empirical analysis. It synthesized information from scientific journals and renowned institutions working on SDG implementation, such as the United Nations System. The empirical analysis utilized a statistical inference (Wilcoxon signed-rank test) to assess any significant differences in the rate of progress pre- and post-COVID-19. Data was sourced from the United Nations Global SDG database.

Results: Out of the 19 indicators examined, the findings reveal that only five indicators were significantly impacted: Maternal mortality (SDG 3.1.1), Neonatal mortality (SDG 3.2.2), Prevalence of hepatitis B (SDG 3.3.4), Vaccination Programmes (SDG 3.b.1), and Health worker density (SDG 3.c.1). It is noteworthy that the situation is not homogenous across countries.

Conclusion: Contrary to initial expectations, the available data suggest that the COVID-19 pandemic did not have the anticipated devastating impact on SDG 3. In the short term, the continent seemed equipped to respond to the pandemic, resulting in the limited impact observed in this study. A wider availability of data across the countries and indicators would have provided a deeper understanding of the situation.

Keywords: SDG 3, Health and wellbeing, Africa, COVID-19, SDG Progress

CPS Abstract

Use of modern statistical methods in evaluating the impact of public policy programs.

Author: Humberto Vaquera Huerta

Submission ID: 1868

Reference Number: 1868

Brief Description

Public policies in the agricultural sector have the purpose of reducing rural poverty. Evaluating the impact of public policies is a topic that allows government offices to make adjustments to improve and citizens to have evidence of the effectiveness of the policy. It is important to incorporate modern statistical methods (machine learning models) to improve inference and predictive capacity. This work discusses the incorporation of learning models in the evaluation of the impact of public policies. A case study is presented with an evaluation study of the income of the beneficiaries of the 2014-2015 period of the rural extension program in Mexico (FAO/SAGARPA), the data is given at the farm level for groups of beneficiaries and not beneficiaries. In particular, the use of a generalized additive nonlinear model with Pareto dependent variable, the Bayesian additive regression tree model and linear models is shown.

CPS Abstract

Using New Technologies to Leverage Alternative Data in the Production of Official Statistics

Author: Dr Linda J. Young

Submission ID: 2002

Reference Number: 2002

Brief Description

Adopting new technologies leads to opportunities to leverage alternative data more effectively. In this presentation, new products and processes the USDA National Agricultural Statistics Service has developed based on new technologies and administrative or remotely sensed data, two forms of alternative data, are presented. Current and future efforts to use new technologies to further leverage alternative data are highlighted. The potential impact of these efforts on data quality is discussed.

Abstract

National Statistical Offices (NSOs), including the United States Department of Agriculture (USDA) National Agricultural Statistics Service (NASS), are increasingly turning to alternative data to obtain needed information, thereby reducing survey and census respondent burden. In addition, many NSOs are adopting modern technologies that lead to improved efficiencies in the production of official statistics. Adopting new technologies leads to opportunities to leverage alternative data more effectively. In this presentation, new products and processes NASS has developed based on new technologies and administrative or remotely sensed data, two forms of alternative data, are presented. As examples, advances in technology have improved the use of administrative and remotely sensed data in list frame development, resulting in a more complete list frame. Alternative data and previously reported data are being used in surveys and censuses to reduce respondent burden. Editing and imputation based on alternative data have been improved using modern technology, leading to improved data quality and a reduction in staff time devoted to these processes. Analysis and estimation of official statistics have improved and the time staff devote to the review process has been reduced for several surveys through effective use of alternative data and new technology. NASS is continuing to evolve its processes. Current and future efforts to use new technologies to further leverage alternative data are highlighted. The potential impact of these efforts on data quality is discussed.

CPS Abstract

Using record linkage of administrative records to improve federal justice statistics in the United States

Author: Dr Mark Motivans

Submission ID: 1995

Reference Number: 1995

Brief Description

Mark Motivans, Ph.D., Statistician Bureau of Justice Statistics, U.S. Department of Justice To improve the analytical capability of the federal data, the Bureau of Justice Statistics has sought the best approach to link records across the federal criminal case process. Linked or longitudinal data can effectively address analytic and policy questions concerning system efficiencies, workload issues, and resource allocation decisions. Linked data can also be used to augment information that may be missing or incomplete in one agency with data from another agency. Examples include records linked between arrest and prosecution, prosecution and adjudication, adjudication and sentencing, and sentencing and prison admission. This paper describes the methods used by BJS to improve link rates by using both person- and case-centric information and highlights the utility of this approach for informing policy makers and the public.

Abstract

Mark Motivans, Ph.D., Statistician

Bureau of Justice Statistics, U.S. Department of Justice

The Federal Justice Statistics Program is managed by the Bureau of Justice Statistics (BJS) and serves as the national clearinghouse of administrative federal criminal case processing data. Under this program, data are received from six federal justice agencies each year and are standardized, maintained, linked, analyzed, and archived. BJS uses identifiers from these data sources to create link files that offer researchers interested in studying the case processing of federal defendants a new way to investigate a range of issues affecting policy, theory, and practice. This presentation describes the linked files—a set of cross-walks that provide the researcher with the ability to track suspects forward (e.g., from arrest to subsequent processing stages) and backward (e.g., from sentencing to earlier stages); the link rates, and potential applications for enhancing federal justice statistics. The goal is to promote the use of the BJS FJSP data and its linking system for external researchers and to share our experiences with others who have similar issues linking records across diverse data.

See: "Using Linked Data from the Federal Justice Statistics Program" Mark Motivans (National Academy of Sciences: <https://vimeo.com/370915615>),

CPS Abstract

“Social and economic indicators from transactional banking data.”

Author: Alejandro Ruiz

Submission ID: 1927

Reference Number: 1927

Brief Description

We present the initial outcomes of a collaboration between INEGI and three major financial institutions in Mexico—Banorte, BBVA, and Santander—for the provision of monthly statistical information based on their transactional data. In this initial stage, we will be publishing information on payroll disbursement. To achieve this and by standardized algorithmic and computational processes, each banking institution processes statistics by deciles of nominees from their servers. These statistics are calculated for various subpopulation groups formed by the intersection of demographics—gender and age—and residence geography—national, state, and municipal. Finally, the resulting statistics are sent to INEGI, where we validate, merge, and represent them in various interactive visual elements that facilitate access for all types of users. The frequency and disaggregation of the information will enable its use in studies and analyses of complex issues, as well as in making timely decisions regarding public policy on the evolution of labor market conditions and income distribution. The outcomes of such collaborations reflect the willingness of private institutions and the potential for leveraging new sources of information to complement the current data from National Statistical Offices. However, we must discuss and work on the different challenges to achieve sustainable collaborations.

Abstract

We present the initial outcomes of a collaboration between INEGI and three major financial institutions in Mexico—Banorte, BBVA, and Santander—for the provision of monthly statistical information based on their transactional data. In this initial stage, we will be publishing information on payroll disbursement. To achieve this and by standardized algorithmic and computational processes, each banking institution processes statistics by deciles of nominees from their servers. These statistics are calculated for various subpopulation groups formed by the intersection of demographics—gender and age—and residence geography—national, state, and municipal. Finally, the resulting statistics are sent to INEGI, where we validate, merge, and represent them in various interactive visual elements that facilitate access for all types of users. The frequency and disaggregation of the information will enable its use in studies and analyses of complex issues, as well as in making timely decisions regarding public policy on the evolution of labor market conditions and income distribution. The outcomes of such collaborations reflect the willingness of private institutions and the potential for leveraging new sources of information to complement the current data from National Statistical Offices. However, we must discuss and work on the different challenges to achieve sustainable collaborations.

CPS Poster

CPS Poster

Adjusting Sample Weights to Integers: Effects on Mean and Variance Estimators in Simple Random Sampling

Author: Mr Alberto Manuel Padilla

Submission ID: 1850

Reference Number: 1850

Brief Description

In probabilistic sampling, sample weights are used to construct estimates of means and variances, among other quantities. These weights may include adjustments for non-response, adjustments to known population totals, and calibration with information on auxiliary variables positively correlated with the survey variables of interest; as well as truncation or rounding of sampling weights to integers. The adjustment to integer numbers of the sample weights is carried out in relevant surveys for decision-making such as the ENOE (National Survey of Occupation and Employment) carried out by the INEGI or the monthly survey of the IMEF (Mexican Institute of Finance Executives). This survey is used to construct the manufacturing and non-manufacturing IMEF Indicators. The methodological documents of these surveys do not describe the method used to adjust numbers to integers or the reasons for its use. In this poster it will be shown that the mean estimator under simple random sampling is biased when adjusting the weights to integers. Some examples of these effects will be shown. The effect on systematic sampling will also be discussed.

Abstract

In probabilistic sampling, sample weights are used to construct estimates of means and variances, among other quantities. These weights may include adjustments for non-response, adjustments to known population totals, and calibration with information on auxiliary variables positively correlated with the survey variables of interest; as well as truncation or rounding of sampling weights to integers. The adjustment to integer numbers of the sample weights is carried out in relevant surveys for decision-making such as the ENOE (National Survey of Occupation and Employment) carried out by the INEGI or the monthly survey of the IMEF (Mexican Institute of Finance Executives). This survey is used to construct the manufacturing and non-manufacturing IMEF Indicators. The methodological documents of these surveys do not describe the method used to adjust numbers to integers or the reasons for its use. In this poster it will be shown that the mean estimator under simple random sampling is biased when adjusting the weights to integers. Some examples of these effects will be shown. The effect on systematic sampling will also be discussed.

CPS Poster

Analyzing Nationwide Urban Service Accessibility for Women through Computational Analysis.

Author: Ms Alejandra Figueroa

Submission ID: 1947

Reference Number: 1947

Brief Description

Accessibility to education and health services emerges as a matter of crucial importance for the overall well-being of the population, particularly for the female subpopulation. While it is essential to have services that encompass the entire population, their location is equally critical. These services must be easily accessible, situated at a reasonable distance to ensure that emergency access is viable and that education does not become a daily sacrifice.

The availability of an approach that provides us with information on distance is a significant step towards improving accessibility to education and health services for the population. This approach not only allows us to understand the distance and time required by women in Mexico to travel from their homes to crucial service centers but also lays the groundwork for a comprehensive analysis. The analysis will focus on the precise identification of the coverage of these centers, highlighting those of special importance to women. Measurement is carried out considering the volume of the population of interest and the maximum acceptable time for travel.

Presenting these results on a poster will not only provide a clear and effective visualization of the gathered information but will also facilitate the communication of crucial findings. This analysis has the potential to inform and guide strategic decisions to improve accessibility and address existing barriers, thereby contributing to the well-being and equal opportunities for the vulnerable female population in the context of essential services.

Abstract

Accessibility to education and health services emerges as a matter of crucial importance for the overall well-being of the population, particularly for the female subpopulation. While it is essential to have services that encompass the entire population, their location is equally critical. These services must be easily accessible, situated at a reasonable distance to ensure that emergency access is viable and that education does not become a daily sacrifice.

The availability of an approach that provides us with information on distance is a significant step towards improving accessibility to education and health services for the population. This approach not only allows us to understand the distance and time required by women in Mexico to travel from their homes to crucial service centers but also lays the groundwork for a comprehensive analysis. The analysis will focus on the precise identification of the coverage of these centers, highlighting those of special importance to women. Measurement is carried out considering the volume of the population of interest and the maximum acceptable time for travel.

Presenting these results on a poster will not only provide a clear and effective visualization of the gathered information but will also facilitate the communication of crucial findings. This analysis has the potential to inform and guide strategic decisions to improve accessibility and address existing barriers, thereby contributing to the well-being and equal opportunities for the vulnerable female population in the context of essential services.

CPS Poster

Citizen Science and VGI to integrate and improve the scale and disaggregation of data related to GBV on mapping focus groups, case: street harassment.**Author:** Prof. Sandra Hernandez-Zetina**Submission ID:** 2094**Reference Number:** 2094**Brief Description**

Strand: Challenges and innovations in data production: address emerging data needs and stakeholders; enhance citizen participation. Citizen Science and VGI to integrate and improve the scale and disaggregation of data related to GBV on mapping focus groups, case: street harassment. Sandra Lucía Hernández Zetina (1) Ana Belen Anquela Julian (2) Angel Esteban Martin Furones (2) Lidia Alejandra Gonzalez Becerril (1) (1) Autonomous University of the State of Mexico

(2) Polytechnic

University of Valencia Violence against women and girls is one of the most widespread human rights violations; studies carried out by the World Health Organization (WHO) determine that this situation has already become a public health problem. Among the types of violence against this sector is street harassment, which threatens the integrity of one out of every three women; its main characteristic is that it is a momentary personal interaction that, nevertheless, restricts the personal integrity of women while they are moving in public spaces. However, there is a lack of detailed or disaggregated information on the behavior of this activity. The information available is generally at the municipality level, which does not allow for a more specific analysis of the urban areas or zones in which these events occur and the specific characteristics of the environment where street harassment occurred. This work presents a proposal that uses non-traditional data sources, such as those generated by citizens, through Voluntary Geographic Information, with the use of focus groups as a research technique to collect data through group interaction from mapping sessions and collaborative mapping within a web application designed to conduct an audit of public space through the participation of users of these spaces to determine their perception of fear from several variables, which give guidelines for the analysis of information from geospatial indicators, in open data schemes.

Abstract

Violence against women and girls is one of the most widespread human rights violations; studies carried out by the World Health Organization (WHO) determine that this situation has already become a public health problem. Among the types of violence against this sector is street harassment, which threatens the integrity of one out of every three women; its main characteristic is that it is a momentary personal interaction that, nevertheless, restricts the personal integrity of women while they are moving in public spaces. However, there is a lack of detailed or disaggregated information on the behavior of this activity. The information available is generally at the municipality level, which does not allow for a more specific analysis of the urban areas or zones in which these events occur and the specific characteristics of the environment where street harassment occurred. This work presents a proposal that uses non-traditional data sources, such as those generated by citizens, through Voluntary Geographic Information, with the use of focus groups as a research technique to collect data through group interaction from mapping sessions and collaborative mapping within a web application designed to conduct an audit of public space through the participation of users of these spaces to determine their perception of fear from several variables, which give guidelines for the analysis of information from geospatial indicators, in open data schemes.

CPS Poster

Determining Levels of Health Literacy in South Africa: evidence from the Demographic and Health Survey 2016**Author:** Ms Thembelihle Luthuli**Submission ID:** 2083**Reference Number:** 2083**Brief Description**

Thembelihle Luthuli, Student, University of KwaZulu-Natal, South Africa

Abstract

Sexual and reproductive health remains a critical concern for public health and development in South Africa, where elevated rates of teenage pregnancy, unplanned pregnancies, and unsafe abortions pose significant threats to women's well-being. Increasing health literacy in a population can reduce poor sexual and reproductive health outcomes aligning with the Sustainable Development Goals (SDGs) related to health and well-being. Health literacy, a critical determinant of population health, is an essential tool for preventing and managing health challenges. Recognising its association with health outcomes, there is a pressing need to assess health literacy levels in diverse contexts.

Methods: Utilizing the South African Demographic and Health Survey, factor analysis was employed to extract a singular health literacy indicator, reflecting multiple dimensions. A logistic regression model was fitted using the health literacy indicator.

Results: The developed health literacy indicator demonstrated reliable internal consistency. Results revealed marked differences in health literacy levels based on gender, education, and geographical residence. Additionally, regional disparities in health literacy were identified.

Conclusion: This research underscores the urgent need to address health literacy as a strategic tool for mitigating poor sexual and reproductive health outcomes in South Africa. The insights gained from this study may inform targeted interventions and policies to enhance health literacy and, consequently, improve sexual and reproductive health across diverse populations.

CPS Poster

ESG Criteria to produce integrated sustainability statistics

Author: Paola Casciotti

Submission ID: 2055

Reference Number: 2055

Brief Description

ESG Criteria to produce integrated sustainability statistics

Abstract

Implementing the complex Agenda 2030, with its high global Sustainable Development Goals (SDGs) requires, in the by now short time horizon of reference, an extraordinary effort, at institutional, private and individual levels, to converge effectively on the whole system of intermediate and interrelated targets. At the same time, it is necessary to ensure financial long term balances in public and private sectors. Sustainable Finance (ESG criteria, Impact Investing,..) is absolutely central to achieve Sustainability Goals, net of deviant greenwashing phenomena.

The comparative study of the main institutional frameworks and enterprises reporting ones, concerning the multiple and complex dimensions of Sustainability, demonstrates how ESG criteria are suitable to allow the development of an integrated analysis model for goals, targets and indicators currently in use. The model may help to converge effectively on the SDGs and to implement and evaluate coherent public and enterprise's policies paying attention to realistic sustainability reports at all levels.

The study shows the results achieved by comparing these frameworks through the proposed classification, based on the individual E-S-G criteria and on their possible combinations (ES-EG-SG-ESG) and quantifies the weight of the singular E-S-G drivers and of their combinations for each framework considered and through them altogether, highlighting interesting areas of correlations between drivers.

The adopted methodology consists in defining multiple analytical partitions of the frameworks examined, to produce a classification of the individual elements identified (goals-dimensions-targets-855 indicators), based on ESG criteria and to represent them into multidimensional matrixes.

The need to estimate a correlation matrix of all elements considered is the main challenge for collecting new Data and new Measurement Systems at institutional level, in order to correctly estimate a correlation ESG matrix, for subsequently defining an integrated materiality matrix comprehensive of risk sustainability factors and financial ones, suitable for global statistical production and decision making processes.

CPS Poster

Enhance Participation in Survey

Author: Hui Yi Yeo

Submission ID: 1917

Reference Number: 1917

Presentation File

abstracts/iaos-isi-2024_b1d055000f3019fbed60ad7906f06cba.pdf

Brief Description

ILO's strand: Challenges and innovations in data production:

Address emerging data needs and stakeholders; enhance citizen participation.

This strand will examine how new data demands offer opportunities for current and future statistical programs.

Decision-making and policy formation in Singapore is data-based. This is no different in the area of labour statistics. To this end, the Manpower Research and Statistics Department (MRSD, Ministry of Manpower, Singapore) is constantly enhancing our data inventory. We are, however, wary of respondent burden and look to administrative data wherever possible. This submission on the proposed poster will describes the strategies undertaken by MRSD, to minimise the cost of response to the respondent and motivate respondent participation, when surveys are our only recourse.

Abstract

>>Minimizing the Cost of Response to Businesses and Improving Response Quality – Automated Data Transfer Program (ADT): The conventional mode of eliciting responses has revolved around business providing answers to questions posed in questionnaire (mainly online in today's internet age). The ADT goes a step further and transfers relevant records from the business's system to MRSD's database. Alternatively, businesses may authorize transfer through approved HR solutions providers. Doing so obviates the need for businesses to take on the burden of collating the required responses.

>>Bringing Statistics Closer to Public - Returning Value to Businesses for Their Responses: The ability to benchmark their labour market performance against the industry average is a valuable resource to businesses. To motivate businesses to participate in our surveys, MRSD provides our business respondents with a corporate benchmarking scorecard so that they may evaluate how they compare against their industry counterparts in wages, employment conditions and staff turnover. For those not selected for responses, calculators are provided. In doing so, they will be motivated to participate in future rounds when they are selected.

CPS Poster

Evaluation of the quality of mortality information before and during the Covid-19 pandemic: Mexico 2000-2021**Author:** Leslie Guzman**Submission ID:** 2101**Reference Number:** 2101**Brief Description**

The vital data records of each country such as Mexico contribute to the description of the national and state demographic and social distribution. In addition, they provide very valuable basic statistical information at different levels of disaggregation and thematic diversity for free, with easy access and use. Therefore, it is required that said information be truthful, timely, relevant and of high quality, in order to contribute to the development of public health strategies to improve health promotion for the entire population and vulnerable groups.

Objective: The objective of this study is to evaluate the quality of the information of the INEGI mortality data in Mexico from 2000 to 2021; and compare the pre-Covid-19 pandemic periods (2000-2019) and during the Covid-19 pandemic (2020-2021) to identify impacts on the accuracy and usefulness of this information.

Material and methods: Study to evaluate the quality of information from the INEGI general mortality databases at the state and national level, using the information quality index of the Ministry of Health of Mexico.

Results: The highest result of the quality index was in 2005 with 99.97 and the lowest in 2021 with 86.85, the quality of the information ranged from 99.84 to 90.01. In the comparative analysis pre-Covid and during-Covid they showed an average of 95.29 and 90.65 respectively.

Conclusions: The quality of the death registry data in Mexico prior to the Covid-19 pandemic improved and remained at good quality. However, the exponential increase in deaths from Covid-19 impacted the components of coverage, consistency, and integrity; due to the absence of data reported by federal entities, low percentage between the comparison of the data reported and estimated by CONAPO, as well as timely registration.

CPS Poster

Factors associated with Women's timing of first antenatal care visit during their last pregnancy: evidence from 2016 Uganda demographic health survey**Author:** Miss Rebecca Namyalo**Submission ID:** 1900**Reference Number:** 1900**Brief Description**

Utilization of antenatal care services (ANC) during pregnancy has been recognized as a major public health intervention to abate maternal morbidity and mortality. Uganda has experienced high levels of maternal morbidity and mortality over the past two decades. This study aimed at investigating the factors associated with timing of first ANC visit by women in Uganda.

This study used secondary data from the 2016 Uganda Demographic and Health Survey (UDHS). The study population comprises of women aged 15–49 who reported to have given their last birth during the five years preceding the 2016 UDHS survey. The outcome variable for this study was the timing of first ANC visit. Univariate, bivariate, and multilevel binary logistic regression analysis was used to determine the factors associated with the utilization of timing of first ANC visit.

Findings show that only 30% [95%CI; 0.28–0.31] of women utilized ANC during the first trimester. Women of higher parity (4+) were less likely to utilize ANC in the first trimester compared to the lower parity (1) (AOR, 0.74, CI; 0.60–0.92). Women who reside in communities with good access to health facility were more likely to utilize ANC during the first trimester as compared to women residing in communities inaccessible to health facility (AOR, 1.36, CI; 1.04–1.77). Women who reside in less diverse ethnic communities were less likely to utilize ANC in the first trimester compared to their counterparts (AOR, 0.15, CI; 0.11–0.22).

This study demonstrated that contextual factors are important predictors of utilization of ANC during the first trimester apart from individual, factors. It is thus important for maternal health programme interventions to consider both individual and contextual factors when encouraging women to utilize ANC services during the first trimester.

Abstract**Abstract**

Utilization of antenatal care services (ANC) during pregnancy has been recognized as a major public health intervention to abate maternal morbidity and mortality. Uganda has experienced high levels of maternal morbidity and mortality over the past two decades. This study aimed at investigating the factors associated with timing of first ANC visit by women in Uganda.

This study used secondary data from the 2016 Uganda Demographic and Health Survey (UDHS). The study population comprises of women aged 15–49 who reported to have given their last birth during the five years preceding the 2016 UDHS survey. The outcome variable for this study was the timing of first ANC visit. Univariate, bivariate, and multilevel binary logistic regression analysis was used to determine the factors associated with the utilization of timing of first ANC visit.

Findings show that only 30% [95%CI; 0.28–0.31] of women utilized ANC during the first trimester. Women of higher parity (4+) were less likely to utilize ANC in the first trimester compared to the lower parity (1) (AOR, 0.74, CI; 0.60–0.92). Women who reside in communities with good access to health facility were more likely to utilize ANC during the first trimester as compared to women residing in communities inaccessible to health facility (AOR, 1.36, CI; 1.04–1.77). Women who reside in less diverse ethnic communities were less likely to utilize ANC in the first trimester compared to their counterparts (AOR, 0.15, CI; 0.11–0.22).

This study demonstrated that contextual factors are important predictors of utilization of ANC during the first trimester apart from individual, factors. It is thus important for maternal health programme interventions to consider both individual and contextual factors when encouraging women to utilize ANC services during the first trimester.

CPS Poster

Impact of access to sustainable energy and cooking solutions – does it change people's lives?

Author: Ole Sandvik

Submission ID: 1863

Reference Number: 1863

Brief Description

The proposal is submitted under stand 2: Leaving no one behind: Greater disaggregation and counting oversights

Having access to energy is a critical issue with huge impacts on individuals, households and communities, as well as economies and the environment. Energy is today the backbone of development, and it is essential for meeting the basic needs of people. Never the less, millions of people lack access to reliable and affordable energy, something that hinders their social and economic growth. To address this issue, governments, organizations, and communities around the world are working to increase access to sustainable energy.

The Sustainable Development Goal (SDG) number 7 is to assure "affordable, reliable, sustainable and modern energy for all within 2030". Access to electricity is one dimension of the goal. Energy is needed in the form of electricity for a number of issues, but usually energy for transport, heating and cooking is provided in other ways. At household level, the main non-electrical energy consumption is for cooking and in some countries also for heating. The energy carriers for cooking are usually solid fuel, gas (liquid gas and biogas), or liquid fuel but may also be covered by solar energy or electricity. The amount of emissions and toxic fumes associated with cooking, varies greatly with both the type of fuel and oven used.

For countries to develop a national energy strategy, it is essential to document the access to electricity and cooking solutions in a joined manner, using evidence to guide policies for having both a sustainable mix and use of energy. This session will discuss the impact of having access to energy and improved cooking solutions.

Moderator:

Mr. Dag Roll-Hansen, Statistics Norway

Abstract

One Sustainable Development Goals (SDGs), is to guarantee inclusion of gender and children, these being considered one of the main pillars for the development of a society. Education is an essential instrument for development, then increasing the supply of quality education is one of the main strategic measures proposed by Mozambique. The opportunity cost of participating in school, considering the child's role within the family, can be a determining factor, since by enrolling in school the child may have to give up carrying out other domestic tasks for the family.

Given that 45% of the country's population is aged between 0 and 14 years old, and that 47.4% of these are in rural areas, and as Mozambique is a mostly rural country (70%) with access to electricity from the electricity grid in 13% in rural areas. Figures from a Survey on the impact of access to energy in 2022. Therefore, the study aims to verify the extent to which access to energy and the existence of improved kitchen solutions can influence the development and well-being of gender and impact the educational lives of children.

Access to improved sources for cooking considerably reduces the time spent preparing food and purchasing or acquiring cooking fuels, impacting the socioeconomic and educational development of gender and children. The higher the level of access to energy, the better the forms of kitchen convenience, reducing wasted time in the processes of preparing the stove for cooking, as well as cooking time, favoring the availability of time to carry out other tasks. activities that can translate into improving the socioeconomic quality of life.

The results will highlight the importance of policies and programs that promote universal access to energy and the adoption of cleaner stove technologies. This will not only benefit gender equality by empowering women.

CPS Poster

Institutional arrangements to strengthen administrative data: the case of performance contracting in Kenya and Uganda.

Author: Ms Sarah Omache

Submission ID: 2063

Reference Number: 2063

Brief Description

The session relates to a technique that can be used by government to provide high level support for the National Statistical Office to implement its mandate of coordination the National Statistical System in complementarity with the statistics law . This is through compelling the government Ministries, Departments, agencies and local governments to submit administrative data to the NSO as per guidelines provided by the NSO. This also includes submission of reports from these entities on how official statistics have been used in the government entities and the impact it has created. This can be achieved through the performance contract process.

The main objective of the session is to create a forum for relevant stakeholders to share knowledge and experiences, pros and cons of this technique of strengthening administrative data. It will also allow countries like Uganda that use this technique to share their experiences and allow sharing of best experiences.

Abstract

The background of the presentation will be in form of the challenges/gaps that are faced by NSOs in obtaining quality administrative data for some statistics especially disaggregated by all required aspects. It will also touch on the level of engagement of the NSO in processing of microdata within government entities. The presentation will also describe the status of data use and impact within the government and beyond and if there is a means of measuring data use. It will outline the performance contracting process in Kenya and some of the performance indicators currently used to assess performance of government entities. I will make a proposal of how data provision by state entities and data use can be integrated as indicators or part of indicators within the performance contracting process in Kenya and as a best practice for other countries.

CPS Poster

NSOs in a Data-Rich World: Capacity Building, Data Dissemination, and Statistical Literacy

Author: Mr Kabeli Mefane

Submission ID: 1902

Reference Number: 1902

Brief Description

As we navigate through an increasingly data-rich world, National Statistical Offices (NSOs) play a pivotal role in ensuring the effective utilization of information for informed decision-making. The poster presentation will explore the responsibilities of NSOs in this dynamic landscape, emphasizing three key pillars: capacity building, data dissemination, and statistical literacy.

In conclusion, this paper underscores the indispensable role of NSOs in navigating the complexities of a data-rich world. By prioritizing capacity building, implementing innovative data dissemination strategies, and promoting statistical literacy, NSOs contribute significantly to building a foundation for evidence-based decision-making and fostering a data-informed society.

The presentation invites discussions on best practices, challenges, and future directions in the context of NSOs in our data-rich world.

Abstract

In an era characterized by unprecedented data abundance, the pivotal role of National Statistical Offices (NSOs) becomes increasingly significant. This poster presentation explores the multifaceted responsibilities of NSOs in navigating a data-rich landscape, focusing on three critical aspects: Capacity Building, Data Dissemination, and Statistical Literacy.

Capacity Building: NSOs are at the forefront of enhancing the analytical capabilities of individuals and organizations. This section delves into the strategies employed by NSOs to build and strengthen the skills of statisticians, policymakers, and the broader public. From training programs to collaborative initiatives, the poster elucidates how NSOs contribute to a skilled workforce capable of harnessing the potential of diverse datasets.

Data Dissemination: Efficient data dissemination is essential for informed decision-making and public engagement. The poster examines the evolving methods employed by NSOs in disseminating data to various stakeholders. Whether through traditional publications, interactive online platforms, or innovative data visualization techniques, NSOs play a pivotal role in ensuring that data is accessible, understandable, and impactful.

Statistical Literacy: As data permeates every aspect of society, statistical literacy becomes imperative for individuals to critically interpret and evaluate information. This segment of the presentation explores the initiatives undertaken by NSOs to promote statistical literacy among diverse demographic groups. From educational outreach programs to user-friendly communication strategies, NSOs contribute to fostering a data-literate society.

CPS Poster

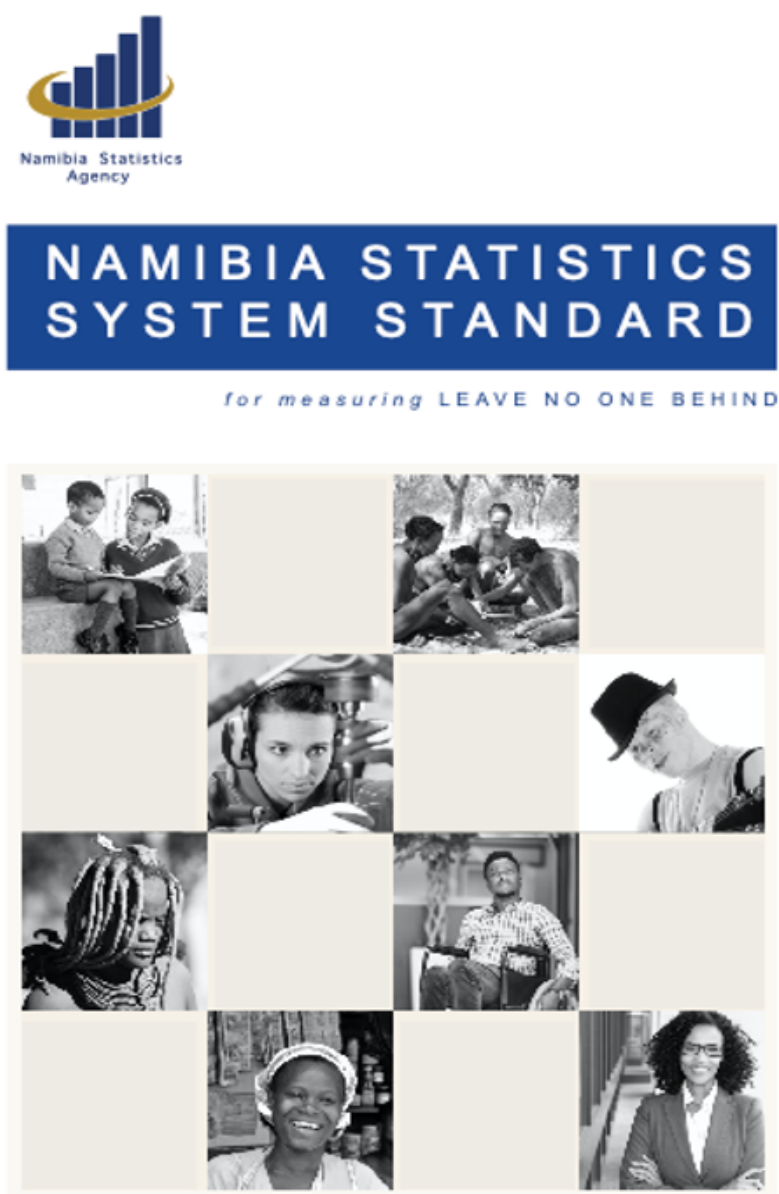
Namibia Statistical System Standard for measuring “Leave No One Behind”

Author: Miss Otja Tjipetekera

Submission ID: 1865

Reference Number: 1865

Presentation Image



Brief Description

The Namibia Statistics Agency (NSA) was established in terms of the Statistics Act, (No.9 of 2011), with the mandate among others to be the central repository for official statistics and to coordinate the National Statistical System. This is to ensure that the produced statistics are harmonized and of high quality for planning, decision-making, and monitoring of national and global development agendas.

It is against this background that the "Leave No One Behind" (LNOB) standard was developed to prescribe the core set of variables which are mandatory to all data producers in the NSS, in order to measure the "leave no one behind" concept. These key variables (age; sex; citizenship; disability; ethnicity; education; geographical location; marital status; orphan-hood) shall form the basis of all data collection tools for surveys, censuses and administrative data in Namibia, where possible. The LNOB concept underpins the Sustainable Development Goals (SDGs) which aims to understand who are left behind and their reasons, in order to end extreme poverty, inequality and to fast-track progress for the people who are left behind. In addition, the LNOB echoes the African Agenda 2063, "the Africa we want", our national agendas such as the 5th National development plans (NDP5) as well as Vision 2030, that aims to achieve prosperity for all Namibians.

In order to achieve goals as set out in these agendas and for the ease of monitoring and tracking progress made thereof, data should be disaggregated to an extent that populations that are left behind are easily identifiable for further intervention. The purpose of this standard is therefore to present a core set of variables required to measure the "leave no one behind" concept in Namibia.

Abstract

The Namibia Statistics Agency (NSA) was established in terms of the Statistics Act, (No.9 of 2011), with the mandate among others to be the central repository for official statistics and to coordinate the National Statistical System. This is to ensure that the produced statistics are harmonized and of high quality for planning, decision-making, and monitoring of national and global development agendas.

It is against this background that the "Leave No One Behind" (LNOB) standard was developed to prescribe the core set of variables which are mandatory to all data producers in the NSS, in order to measure the "leave no one behind" concept. These key variables (age; sex; citizenship; disability; ethnicity; education; geographical location; marital status; orphan-hood) shall form the basis of all data collection tools for surveys, censuses and administrative data in Namibia, where possible. The LNOB concept underpins the Sustainable Development Goals (SDGs) which aims to understand who are left behind and their reasons, in order to end extreme poverty, inequality and to fast-track progress for the people who are left behind. In addition, the LNOB echoes the African Agenda 2063, "the Africa we want", our national agendas such as the 5th National development plans (NDP5) as well as Vision 2030, that aims to achieve prosperity for all Namibians.

In order to achieve goals as set out in these agendas and for the ease of monitoring and tracking progress made thereof, data should be disaggregated to an extent that populations that are left behind are easily identifiable for further intervention. The purpose of this standard is therefore to present a core set of variables required to measure the "leave no one behind" concept in Namibia.

CPS Poster

PREVALENCE OF RESPIRATORY SYNCYTIAL VIRUS CHILDREN WITH INFLUENZA-LIKE ILLNESS SYMPTOMS FROM SELECTED HOSPITALS IN UGANDA

Author: Ms Elizabeth Nampewo

Submission ID: 1893

Reference Number: 1893

Brief Description

As the emergence of pandemic prone respiratory viruses increases in Uganda, it is important to establish baseline information on the prevalence of respiratory Syncytial Virus (RSV) which is one of the leading viral causative agents of Influenza like illness worldwide. Therefore, the objective of this study was to determine the prevalence of RSV in Uganda.

A retrospective study was carried out on archived throat/nasopharyngeal samples obtained from five hospitals during an Influenza surveillance study done by Makerere University Walter Reed Project (MUWRP) in period 2008-2014. Samples were tested by Nested- reverse transcription polymerase chain reaction assay. Out of the 800 throat/nasopharyngeal samples, 102 samples (12.75%) were positive for RSV. The prevalence of RSV was highest among children aged 5 years and below (PR=1.76, 95% C.I 0.81-3.8) with most cases occurring in the dry season (PR=1.00, 95% C.I 0.5-1.1). By hospital, Bugiri hospital had the highest prevalence (15%) followed by Mulago hospital (13.55%), Gulu hospital (13.02%), Jinja hospital (10.88%), and lastly Kayunga hospital (3.57%).

In conclusion, these findings are useful in supporting the existing data from other countries that RSV is still a global issue which requires effective measures like vaccines to be enforced.

Key words: RSV- Respiratory syncytial virus.

Abstract

PREVALENCE OF RESPIRATORY SYNCYTIAL VIRUS CHILDREN WITH INFLUENZA-LIKE ILLNESS SYMPTOMS FROM SELECTED HOSPITALS IN UGANDA

Nampewo Elizabeth

Email: nampewozabetha@gmail.com

Department of Population Studies, School of Statistics and Applied Economics, Makerere University, P.O.Box 7062, Kampala, Uganda.

Theme: Leaving no one behind

ABSTRACT

As the emergence of pandemic prone respiratory viruses increases in Uganda, it is important to establish baseline information on the prevalence of respiratory Syncytial Virus (RSV) which is one of the leading viral causative agents of Influenza like illness worldwide. Therefore, the objective of this study was to determine the prevalence of RSV in Uganda.

A retrospective study was carried out on archived throat/nasopharyngeal samples obtained from five hospitals during an Influenza surveillance study done by Makerere University Walter Reed Project (MUWRP) in period 2008-2014. Samples were tested by Nested- reverse transcription polymerase chain reaction assay. Out of the 800 throat/nasopharyngeal samples, 102 samples (12.75%) were positive for RSV. The prevalence of RSV was highest among children aged 5 years and below (PR=1.76, 95% C.I 0.81-3.8) with most cases occurring in the dry season (PR=1.00, 95% C.I 0.5-1.1). By hospital, Bugiri hospital had the highest prevalence (15%) followed by Mulago hospital (13.55%), Gulu hospital (13.02%), Jinja hospital (10.88%), and lastly Kayunga hospital (3.57%).

In conclusion, these findings are useful in supporting the existing data from other countries that RSV is still a global issue which requires effective measures like vaccines to be enforced.

Key words: RSV- Respiratory syncytial virus.

CPS Poster

Statistical and geographical Information Infrastructure

Author: Juan Eduardo Rioja Romo

Submission ID: 2088

Reference Number: 2088

Brief Description

This poster will illustrate the National Institute of Statistics and Geography of Mexico (INEGI, by its Spanish acronym) and the National System of Statistical and Geographical Information (SNIEG) advancements and challenges to promote and benefit from the common use of the information infrastructure in the production processes.

Abstract

There is a close link among different stages of the data life cycle (production, integration, conservation, exchange). Statistical and geographical production processes can significantly improve and facilitate its exchange and subsequent integration with the use of standards. The information infrastructure is the set of data and methodologies that support the information production process to facilitate its interoperability and it is made up of catalogs, classifications, statistical and geographical registries, and methodologies.

The National System of Statistical and Geographical Information (SNIEG, by its Spanish acronym) Law has defined the minimum information infrastructure for each Information Subsystem: demographic and social; economic; geographical, environmental, territorial, and urban planning; and government, public safety, and justice. The use of a common information infrastructure facilitates the integration or linkage of information from different statistical and geographical production processes.

The National Institute of Statistics and Geography of Mexico (INEGI, by its Spanish acronym), with its dual function as an information producer and as SNIEG'S coordinator, has driven actions to generate statistical and geographical information infrastructure and outlining data governance activities, to clearly identify its components and facilitate data standardization and interoperability.

This poster will illustrate INEGI and SNIEG'S advancements and challenges to promote and benefit from the common use of the information infrastructure.

CPS Poster

Statistical literacy: It's all in the communication

Author: Ms Sofi Nickson

Submission ID: 1836

Reference Number: 1836

Brief Description

The UK Office for Statistics Regulation is proud to present insights from an extensive literature review on statistical literacy, with emphasis on effectively communicating official statistics to the public.

Statistical literacy is fundamental in the modern world, enabling individuals to interpret, evaluate, and utilise data for informed decisions.

We present existing literature, then flip traditional perspectives – rather than increasing levels of statistical literacy, we focus on effective communication to meet the needs of a range of audiences.

Abstract

The UK Office for Statistics Regulation is proud to present a poster offering insights derived from an extensive literature review on statistical literacy, with a particular emphasis on effective methods of communicating official statistics to the public. Statistical literacy is a fundamental skill in the modern world, enabling individuals to interpret, evaluate, and utilise data for informed decision making. Our research examines the existing body of literature, then turns traditional perspectives of statistical literacy on their head – rather than aiming to increase levels of statistical literacy in the public, we see a way forward in focussing on effective communication to meet the needs of a range of audiences.

Our findings underscore the importance of understanding the target audience, providing relevant contextual information, establishing trust, aiming for language that is easy to understand, considering format or framing, and appropriately communicating uncertainty in statistical estimates.

By synthesising latest evidence in this field, our poster provides practical guidance for those involved in the dissemination of statistics and data. Join this poster presentation to discover how enhancing statistical communication can be a solution to varying levels of statistical literacy, leading to a more informed and empowered society, and ultimately aligning with the overarching conference aim of improving decision making for all. We invite attendees to engage with our findings and explore the transformational potential of effective communication in the world of official statistics.

CPS Poster

The role of using the new technologies and programs in improving the quality of official statistical data (comparative study)

Author: Dr Abdulla Fathi

Submission ID: 2079

Reference Number: 2079

Brief Description

There are great relationships between statistics and information technology in the information era that we are living now, where there is a huge and fast flow of data and information. Statistics consider the first step of the development in all sectors in our life which providing data and presenting it appropriately to the decision maker.

The importance of the study is that the enormous increase in the production of statistical information necessitates the use of the latest technology in the fields of data collection, processing and analysis and display faster, simpler and in high quality shape. The Paper present a comparative analysis between the latest two censuses in Egypt, traditional Egyptian census which based on paper and Egyptian technology census which based on using the new technology and electronic tablets which reflected on enhancing the quality of data.

Using technology in all phases and procedures of the first Egyptian technology census has resulted in increasing the quality and accuracy of statistical data and present the census data more rapidly and modernly in line with the speed of the changes we are living now thus increasing the accuracy of decisions based on these data and information at various levels.

Abstract

There are great relationships between statistics and information technology in the information era that we are living now, where there is a huge and fast flow of data and information. Statistics consider the first step of the development in all sectors in our life which providing data and presenting it appropriately to the decision maker.

The importance of the study is that the enormous increase in the production of statistical information necessitates the use of the latest technology in the fields of data collection, processing and analysis and display faster, simpler and in high quality shape. The Paper present a comparative analysis between the latest two censuses in Egypt, traditional Egyptian census which based on paper and Egyptian technology census which based on using the new technology and electronic tablets which reflected on enhancing the quality of data.

Using technology in all phases and procedures of the first Egyptian technology census has resulted in increasing the quality and accuracy of statistical data and present the census data more rapidly and modernly in line with the speed of the changes we are living now thus increasing the accuracy of decisions based on these data and information at various levels.

CPS Poster

Trends and correlates of sexually transmitted infections among sexually active Ugandan female youths: Case: 2006–2016(Uganda demographic surveys)

Author: Mrs Irene Nakazzi

Submission ID: 1899

Reference Number: 1899

Brief Description

Abstract

Female adolescents and young women have the highest risk of curable sexually transmitted infections (STIs) globally. Data on the prevalence of STIs among young women in Uganda are limited. In this study, Time trends and correlates of STIs among adolescent girls and young women (15–24 years) in Uganda were investigated in Uganda who reported suffering from genital sores, and or genital discharge or any other vaginal complaints acquired after sexual intercourse within 12 months of the studies and examined the changes over time. A pooled multivariable logistic regression was used to examine the correlates of reporting an STI in the last 12 months preceding the study. Stata was used to cater for the survey sample design. Among these young women, 36.3, and 23.1% reported a sexually transmitted infection in 2006, 2011, and 2016 respectively. Between 2006 and 2011, there was evidence of change (+ 14.3%, $p < 0.001$) in STI prevalence before a significant reduction (– 12.0%, $p < 0.001$) in 2016. Youths aged 20–24 years reported a higher STI prevalence (27.3%) compared to young participants (23.6%). Correlates of reporting an STI among rural and urban young women were: having multiple total lifetime partners (adjusted odds ratio (aOR) 1.6, 95% CI 1.4–1.6), being sexually active in the last 4 weeks (aOR 1.3, 95% CI 1.1–1.6), and being affiliated to Muslim faith (aOR 1.3, 95% CI 1.1–1.6) or other religions (aOR 1.8, 95% CI 1.1–2.9) as compared to Christian were more likely to report an STI. Living in Northern Uganda compared to living in Kampala city was found protective against STIs (aOR 0.5, 95% CI 0.3–0.7). The prevalence of STIs was high among female youths, 15–24 years. This highlights the need for a comprehensive STIs screening, surveillance, and treatment programme to potentially reduce the burden of STIs in the country.

Abstract

Abstract

Female adolescents and young women have the highest risk of curable sexually transmitted infections (STIs) globally. Data on the prevalence of STIs among young women in Uganda are limited. In this study, Time trends and correlates of STIs among adolescent girls and young women (15–24 years) in Uganda were investigated in Uganda who reported suffering from genital sores, and or genital discharge or any other vaginal complaints acquired after sexual intercourse within 12 months of the studies and examined the changes over time. A pooled multivariable logistic regression was used to examine the correlates of reporting an STI in the last 12 months preceding the study. Stata was used to cater for the survey sample design. Among these young women, 36.3, and 23.1% reported a sexually transmitted infection in 2006, 2011, and 2016 respectively. Between 2006 and 2011, there was evidence of change (+ 14.3%, $p < 0.001$) in STI prevalence before a significant reduction (– 12.0%, $p < 0.001$) in 2016. Youths aged 20–24 years reported a higher STI prevalence (27.3%) compared to young participants (23.6%). Correlates of reporting an STI among rural and urban young women were: having multiple total lifetime partners (adjusted odds ratio (aOR) 1.6, 95% CI 1.4–1.6), being sexually active in the last 4 weeks (aOR 1.3, 95% CI 1.1–1.6), and being affiliated to Muslim faith (aOR 1.3, 95% CI 1.1–1.6) or other religions (aOR 1.8, 95% CI 1.1–2.9) as compared to Christian were more likely to report an STI. Living in Northern Uganda compared to living in Kampala city was found protective against STIs (aOR 0.5, 95% CI 0.3–0.7). The prevalence of STIs was high among female youths, 15–24 years. This highlights the need for a comprehensive STIs screening, surveillance, and treatment programme to potentially reduce the burden of STIs in the country.

CPS Poster

Unlocking Economic Insights: Enhancing GRDP Predictions through Temporal Disaggregation

Author: Mr Hilman Hanivan

Submission ID: 1963

Reference Number: 1963

Brief Description

Unlocking Economic Insights: Enhancing GRDP Predictions through Temporal Disaggregation

Abstract

In the pursuit of localized economic understanding, lower-level governments increasingly demand a more dynamic portrayal of their economic landscape, prompting a shift from annual to quarterly GRDP assessments. However, this transition is hindered by limited data availability at the lower regional echelons, necessitating innovative approaches.

This paper draws on provincial data from Indonesia to address the challenge by advocating for the use of statistics, specifically employing the Denton-Chollette model for temporal disaggregation. Recognizing the potential of this model to enhance data granularity, the paper scrutinizes the predicted quarterly series of GRDP. It unravels the intricacies of this process, highlighting how various factors, including the quality of quarterly supplement indicators, contribute to the divergence in predictions.

The findings underscore the importance of statistical tools in overcoming data limitations and providing a more nuanced understanding of sectoral growth at the local level. By shedding light on the potential disparities in quarterly GRDP predictions, this research equips lower-level governments with valuable insights for more informed economic governance.

CPS Poster

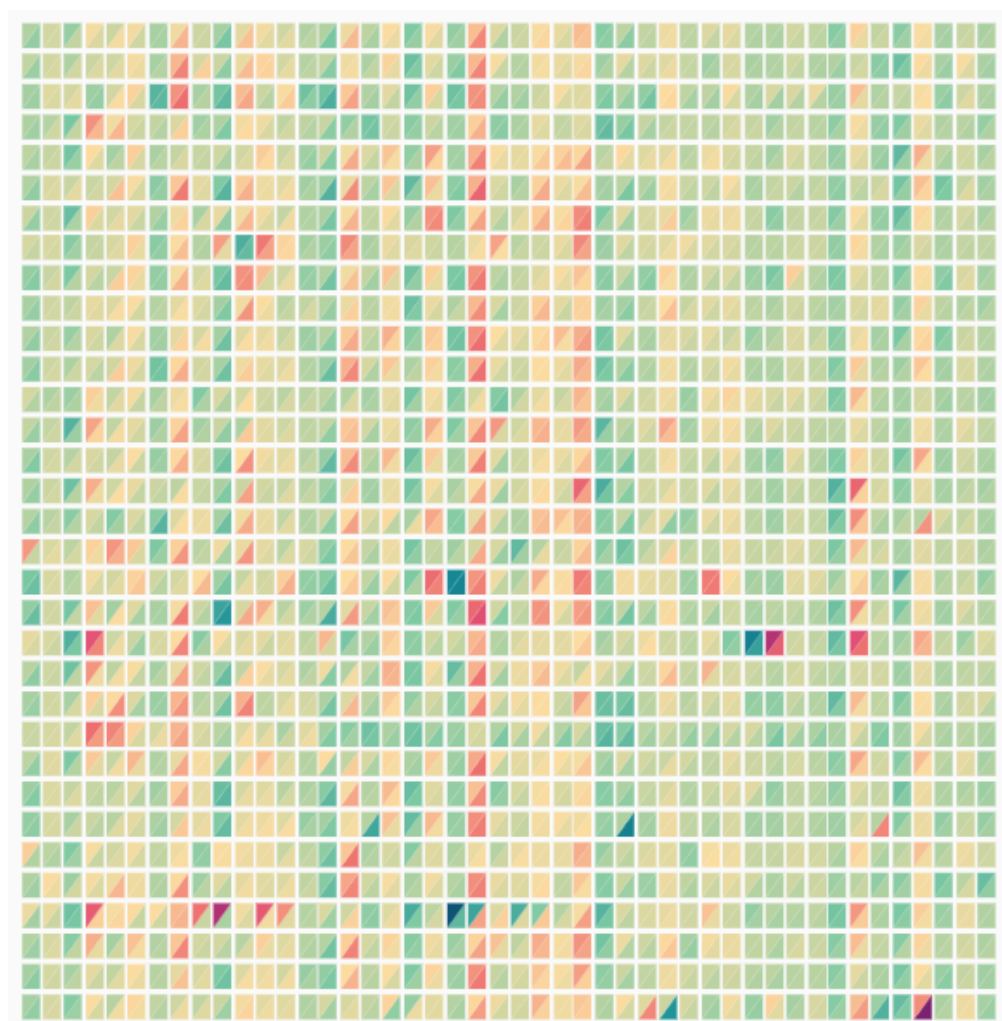
Visualization of disaggregated annual economic activity from banking information

Author: Mr Amado Esquer

Submission ID: 1933

Reference Number: 1933

Presentation Image



Brief Description

As a result of the collaboration between INEGI and the three major financial institutions in Mexico -Banorte, BBVA and Santander-, the publication and sharing of very granular information regarding payroll disbursement became a challenge.

Using interactive tools and modern visualization techniques, several dashboards were created to analyze the information gathered in a cross-section and longitudinal way. In the poster are presented the most striking, enlightening and accessible visualizations created with this data, along with relevant socioeconomic events that could have had an effect in the trends and values of the information shown.

Abstract

As a result of the collaboration between INEGI and the three major financial institutions in Mexico -Banorte, BBVA and Santander-, the publication and sharing of very granular information regarding payroll disbursement became a challenge. Using interactive tools and modern

visualization techniques, several dashboards were created to analyze the information gathered in a cross-section and longitudinal way. In the poster are presented the most striking, enlightening and accessible visualizations created with this data, along with relevant socioeconomic events that could have had an effect in the trends and values of the information shown.