Exercises on Experimental Design and Analysis of Variance

1) The retailing manager of a supermarket chain wants to determine whether product location has an effect on the sales of fluffy toys. Three different aisle locations are considered: front, middle and rear. A random sample of 18 stores is selected with 6 stores randomly assigned to each aisle location. The size of the display area and the price of the product are constant for all stores. At the end of a one-month trial period, the sales volumes (in 1000s of dollars) of the product in each store are given in the table below.

Aisle Location		
Front	**Middle**	**Rear**
8.6	3.2	4.6
7.2	2.4	6.0
5.4	2.0	4.0
6.2	1.4	2.8
5.0	1.8	2.2
4.0	1.6	2.8

The rearranged data is in the file ***sales.csv.***

(a) Generate boxplots of the sales for the aisle locations and comment on any similarities and differences.
(b) Is there evidence of a treatment effect, i.e., differences in the mean sales between the various aisle locations? Use the 5% level of significance.
(c) If the null hypothesis in Part (b) is rejected, perform post-hoc analysis, using Tukey's multiple comparisons tests.
(d) Perform diagnostic checks on the assumptions of normality of sales and homogeneity of variances of sales across the three locations.

2) A tyre company has developed a new tyre and has conducted tread-wear tests to determine whether or not there is a significant difference in the tread-wear of the tyre if the average speed at which a car is driven, varies. The company set up an experiment with three treatment levels: slow speed (car is driven at 30 km/h), medium speed (car is driven at 60 km/h) and high speed (car is driven at 90 km/h). Company researchers realised that several possible variables could confound the study. One of these variables is supplier of rubber from which the tyres are made. The company uses five different suppliers of rubber. The researchers used supplier as a blocking variable in the analysis. Fifteen sets of tyres were selected for the study, three sets of which were made from the rubber of each of the five suppliers. Each of the 15 cars fitted with these tyres, was driven over different periods of time, until 10,000 kilometres was covered within a given time frame. Thereafter, the tyre tread-wear

depth was measured in millimetres. Note that the deeper the thread, the less the wear on the tyres. The data from the experiment are given below.

	Speed		
Supplier	Slow	Medium	Fast
1	2.3	2.8	1.9
2	2.1	2.4	1.8
3	2.2	2.6	1.9
4	2.0	2.2	1.6
5	2.4	3.0	2.1

This data has been rearranged in the file ***tyre.csv***

a) Generate boxplots of tyre thread-wear for the different speeds and suppliers the and comment on any similarities and differences.
b) Test at the 5% level of significance if there is a treatment effect, i.e., if there is evidence of differences in mean tread-wear depth across the three speed categories.
c) Test at the 5% of significance level if there a blocking effect, i.e. if there is evidence of differences in mean tread-wear depth across the five suppliers.
d) Perform diagnostic checks on the assumptions of normality of tread-wear and on homogeneity of variances of tread-wear across the three speeds.
e) Explain why the design used in this experiment is preferable to a design that does not take into account the blocking effect and explain how we may confirm this.

3) A student team in a data analysis course performed a factorial experiment to investigate the sales of a new product (number of items sold). The two factors were marketing technique (X, Y, Z) and geographical location of the market (City 1, City 2 and City 3). The data are presented in the table below.

	Marketing Technique		
Market	**X**	**Y**	**Z**
City 1	87	45	89
	97	56	67
	78	65	85
	102	33	56
City 2	67	32	75
	56	23	63
	67	42	56
	75	18	67
City 3	68	35	90
	78	37	68
	95	36	80
	70	54	58

The rearranged data is in the file ***products.csv***

(a) Generate boxplots taking into account both marketing techniques and comment on differences and similarities
(b) Test if there is an interaction effect between brand marketing technique and location of the market. Generate and comment on the interaction plot of the mean sales for each marketing technique for the three locations.
(c) Test if there is an effect due to marketing techniques, i.e., if sales differ as a result of different techniques.
(d) If appropriate, determine which marketing techniques differ with regards to sales.
(e) Test if there is an effect due to the location of the market, i.e., if sales differ between market locations.
(f) If appropriate, determine which locations differ with regards to sales.
(g) Perform diagnostic checks to assess the assumption of normality of the underlying population distribution of sales and and on homogeneity of variances of sales for the treatment combinations.