1 MAT013 - Example Sheet

1.1 Chapter 1 + 2

1. Create the data set "first_data_set" from the notes and export to csv.

2. Download the files JJJ.csv and MMM.csv import them in to R.

3. View both the MMM and JJJ data sets.

4. View the structure of the MMM and JJJ data sets using the str command.

5. Obtain the mean age, height in metres and weight in kg for the observations of the MMM and JJJ datasets.

6. Obtain the mean age, height in metres and weight in kg for the observations of the MMM and JJJ datasets compartmentalising your output by sex (you'll need to use the by function).

7. Download the file math_tests.csv, import it in to R and output a frequency table of teachers against pass_fail.

8. Obtain correlation tables for all the numerical values in JJJ and MMM (using 2 separate approaches (one of which will require you to download a package).

9. Do a regression analysis of the variable height_in_metres against weight_in_kg and savings_in_poundsfor the data sets JJJ and MMM.

10. Download the data set math.csv and run an ANOVA test to see if the grades depends on the professor.

11. Obtain a histogram for the variables weight_in_kg for the data sets JJJ and MMM.

12. Obtain a scatter plot of weight_in_kg against height_in_metres for the data sets JJJ and MMM

13. Output all of the above graphs to a pdf file.

The relevant data can be found here:

- JJJ.csv
- MMM.csv
- math_tests.csv
- math.csv