Course Exercises

Data analysis for weather and climate research

These exercises are designed to help you to understand the concepts introduced in the lectures and to learn how to conduct, interpret and present basic statistical analyses. They are intended to be completed in order (later exercises assume knowledge acquired in earlier ones) and each exercise assumes that you have studied the corresponding chapter of the lecture notes. The exercises can be completed according to the following schedule:

Day	1	2	3	4	5
Exercise	1, 2	3, 4	5, 6	7, 8	9

Example solutions are available on the course web-page.

Chris Ferro (Tutor) `c.a.t.ferro@reading.ac.uk`

Exercise 1. Introduction

1. Visit the course web-site via
   `www.met.rdg.ac.uk/cag/courses`
   and download the datasets `atldjf.txt` etc.

2. Open `R` and familiarise yourself with the help pages.
   ```
   > help.start()
   ```

3. Read the data in the file `temp.txt` into `R`.
   ```
   > data <- read.table(‘‘temp.txt’’, TRUE)
   > attach(data)
   ```

4. Produce a boxplot and a histogram of the daily mean temperatures.
   ```
   > boxplot(Tmean)
   > hist(Tmean)
   ```
   What do these plots tell you about the shape of the distribution? Are there any unusual values? If so, can you explain their occurrence? How will this affect your analysis?

5. Reproduce the histogram with the width of the intervals equal to 5 degrees.
   ```
   > hist(Tmean, breaks = seq(-5, 30, 5))
   ```
   Does this affect your view of the distribution? What implications does this have for plotting data?
   P.T.O.

6. Compute the mean and standard deviation of the daily mean temperatures.

   ```
   > mean(Tmean, na.rm = TRUE)
   > sqrt(var(Tmean, na.rm = TRUE))
   ```

   Compute the median, quartiles, range and inter-quartile range.

   ```
   > summary(Tmean)
   ```

   What do these statistics tell you about the data? There are some missing values in the data: why do you think they occurred? How does this affect your analysis?

7. Write a short summary of your analysis so far, including your answers to the earlier questions and any other conclusions you have drawn. To export figures from `R` see

   ```
   > ?Devices
   ```

8. Visit the `stats@met` web-site

   `www.met.rdg.ac.uk/cag/stats`

   and follow the link to the 'Rice Virtual Lab in Statistics'. Try out the demonstrations on 'Mean and Median', 'Histograms, Bin Widths, and Cross Validation' and 'Comparing distributions'. Browse some of the other links on the `stats@met` web-site.

   P.T.O.

9. Use a calculator, not `R`, to complete this question. For each of the age, height and weight datasets in the lectures notes (also available in the file `rdgmorph.txt`)

   (a) compute the mean and standard deviation;

   (b) rank the data in increasing order;

   (c) compute the median, quartiles, range and inter-quartile range; and

   (d) describe the shape of the distribution.

   Record your answers below.

	Mean	SDev	Q1	Q2	Q3	IQR	Range
Age							
Height							
Weight							

The age distribution...

The height distribution...

The weight distribution...