Part A: R preliminaries

Some quick remarks about R

1. Inbuilt data sets

As part of the R installation, many inbuilt datasets are available for assisting R learners develop their knowledge using real data. Some of the more commonly used data sets are *iris* and *mtcars.*

2. Dataframes

The conventional structure for a dataset is called a data frame. It is a table of rows and columns. Each row corresponds to a "case" and each column corresponds to a variable. Most packages including ggplot2 work best when the data is structured as a data frame.

A data frame can be accessed simply by typing in the name.

```
mtcars
```

3. Inspecting a data set
Quite often a data set will be very large and it is not easy to get a sense of it by looking at it directly. There are some useful commands that give us important information about the data frame.
a. The summary() command
The summary() command is a very versatile command that gives the user a brief summary of the data object specified.

```
summary(mtcars)
```

b. The head() command

The head command allows the user to read the column names and the data from the first six cases, allowing the user to develop a sense of the structure of the data set.

```
head(mtcars)
```

    c. **The dim() command**

The dim() command yields the numbers of rows and columns (i.e. the dimensions) of a data object.

```
dim(mtcars)
```

    d. **The names() command**

The names() command returns the columns names (i.e. variable names) of a data object.

```
dim(mtcars)
```

4. Base R graphics

R has a default graphics packages installed automatically. Try out the following command to get a sense of it.

```
plot(iris)
```

It is felt by many that the standard of graphical output required for professional use requires too much programming skill to be practicable for most users.
ggplot2 yields much better visualisations with much simpler code.

5. Installing and loading a package.

Packages can be added to the R workspace to provide greater functionality. Packages have to be downloaded from a mirror and then loaded into the workspace using the `library()` command.

```
install.packages("ggplot2")
library(ggplot2)
```

6. Colours

R supports a very wide range of colours. Type in the command colours() to see what there is available.

```
colours()
```

-