Assignment 2: Feature Engineering, Statistical Analysis and Machine Learning

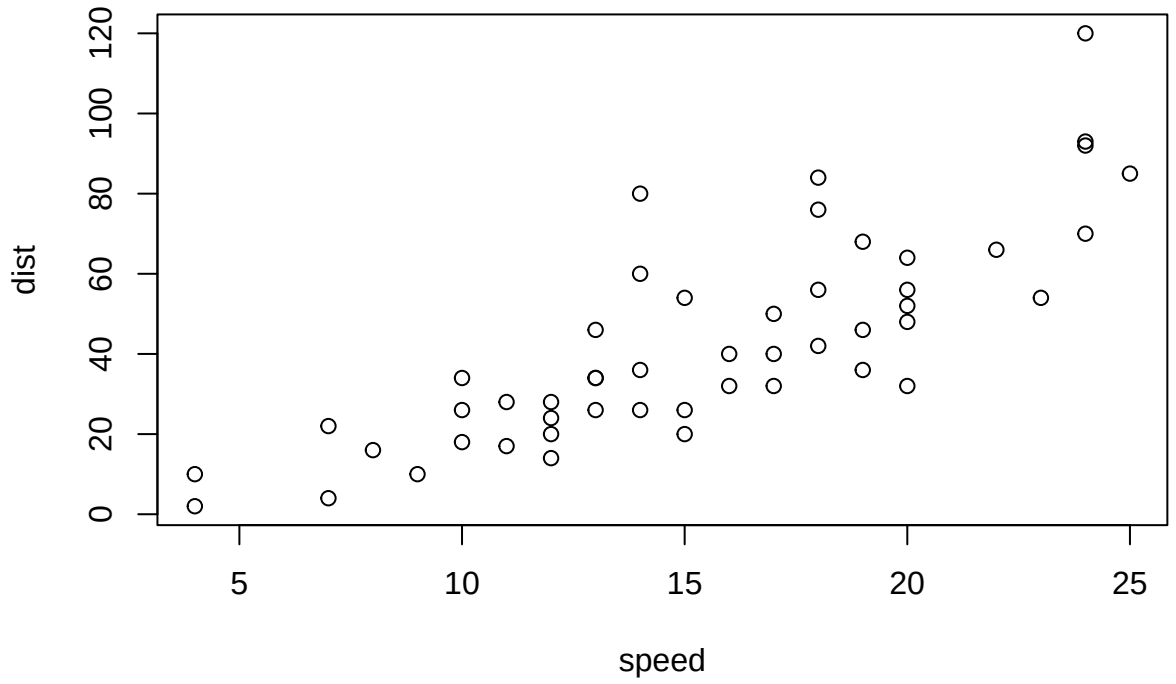
Student Name (Student Number)

Introduction

This is an R Markdown Notebook. When you execute code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by placing your cursor inside it and pressing *Ctrl + Shift + Enter*.

```
plot(cars)
```



Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing *Ctrl + Alt + I*.

When you save the notebook, an HTML file containing the code and output will be saved alongside it (click the *Preview* button or press *Ctrl + Shift + K* to preview the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor. Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the output of the chunk when it was last run in the editor is displayed.

Personalising the notebook

In the notebook header and in the preamble.tex file, change "Student Name" to your name. This should be **exactly** as it appears on Canvas. Replace "Student Number" with your student number. Make sure the file is named with *your* student name and *your* student number as follows: `csc2062_a2_barry_devereux_4012345.Rmd`.

When you submit your work, ensure that the latest `.html` and `.pdf` renderings of the notebook are present.

Importing code

In general, you can just present your code in the R chunks in this notebook. In a some cases where you have more complex pieces of code (e.g. function definitions) you may wish to have that code implemented in a seperate R file. The code can then be imported into the notebook using the *source* command. You do not need more than at most one ancillary code file per section.

In the example R imported below file below, there is a number of functions, including *squarenumbers*, which gives the squares of the numbers 1 to a, for any positive integer a.

```
source("./section1_code.r")
```

Let's calculate the result of the function for the numbers 5 to 10:

```
numbers = c(5:10)
for (number in numbers) {
  print(squarenumbers(number))
}
```

```
## [1]  1  4  9 16 25
## [1]  1  4  9 16 25 36
## [1]  1  4  9 16 25 36 49
## [1]  1  4  9 16 25 36 49 64
## [1]  1  4  9 16 25 36 49 64 81
##  [1]   1   4   9  16  25  36  49  64  81 100
```

In general, as you develop your code, you may find it convenient to write and test it in a standard `.r` file, before adding the working code to this notebook.

Word count

If you want to check the wordcount of your notebook (excluding code), you can install the wordcount Rstudio addin by Ben Marwick.

Follow the installation and use instructions at the link above.

```
wordcountaddin:::text_stats()
```

```
## For information on available language packages for 'koRpus', run
##
##   available.koRpus.lang()
##
## and see ?install.koRpus.lang()
```

Method	koRpus	stringi
Word count	511	485
Character count	2880	2909
Sentence count	42	Not available
Reading time	2.6 minutes	2.4 minutes

Introduction

In your actual report, you can delete this introduction. Replace it with a few sentences describing the goals of the assignment, in your own words.

Section 1

Your work for this section here.

Section 2

Your work for this section here.

Section 3

Section 3.1

Your work for this subsection here.

Section 3.2

Your work for this subsection here.

Section 3.3

Your work for this subsection here.

Section 3.4

Your work for this subsection here.

Section 4

Section 4.1

Your work for this subsection here.

Section 4.2

Your work for this subsection here.

Section 4.3

Your work for this subsection here.

Section 4.4

Your work for this subsection here.