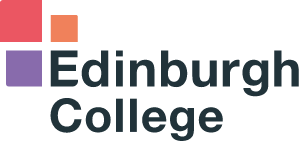
Data Science: ePortfolio 

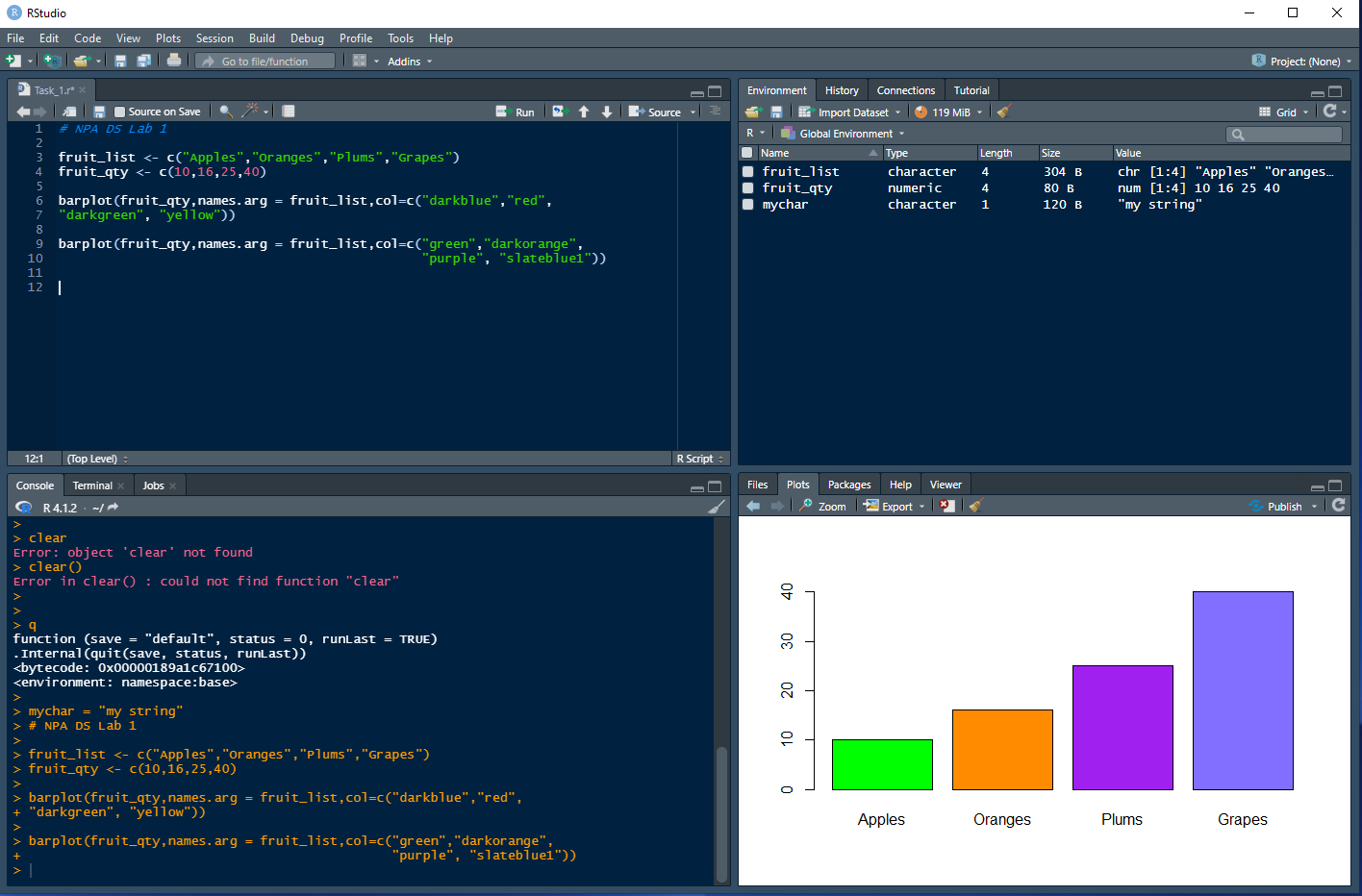
|  |  |
| --- | --- |
| First Name (s) | Kevin |
| Last Name (s) | Gemson |
| EC Email Address | EC2111447@edinburghcollege.ac.uk |

# Guidance - How to complete this document

* Check the course Moodle page to see what evidence is required.
* All screenshots should be cropped and labelled to show the key information.
* The lab title and date of completion should be added above each section.
* Text should be in 11pt or 12pt, any code should be in a monospaced font.
* Each week complete the summary with an overview of the tasks completed.

# Lab 1: Data In Society - Date Completed: 05/02/2022

1. A screenshot of RStudio Running on your computer



|  |
| --- |
| 1. Your plot from step 1.11 of the lab 2. Your plot from step 1.12 of the lab |

## Weekly summary, complete the following:

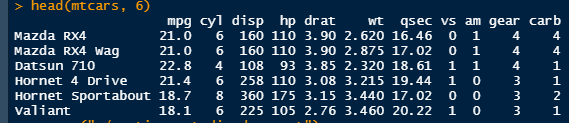
This week I learned about ... principles of Data Citizenship, and how data is collected and used (and can be misused) in Society. Also, learned the definition of Data Science and its relationship to machine learning and artificial intelligence.

In the lab I … learned how to install R and R Studio, and run a number of tasks to understand how the basic syntax works. Also followed a number of YouTube tutorials including the extra learning module, to try to get a better understanding of what R can do.

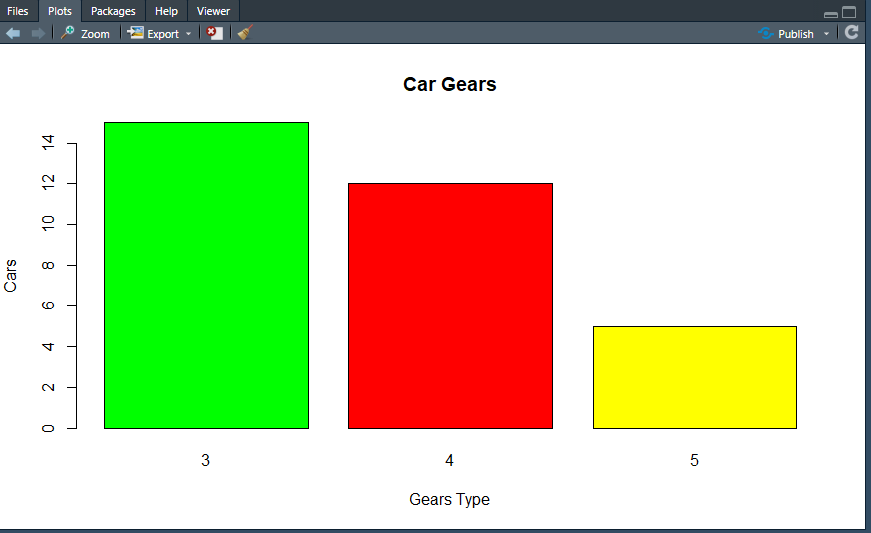
I enjoyed.. getting to see the possibilities of what R can do, being able to produce graphs with some fairly simple commands

I found it challenging when… working through some of the different components of the syntax (vectors matrices, datasets, etc), especially as the syntax is different to other languages I am more familiar with.

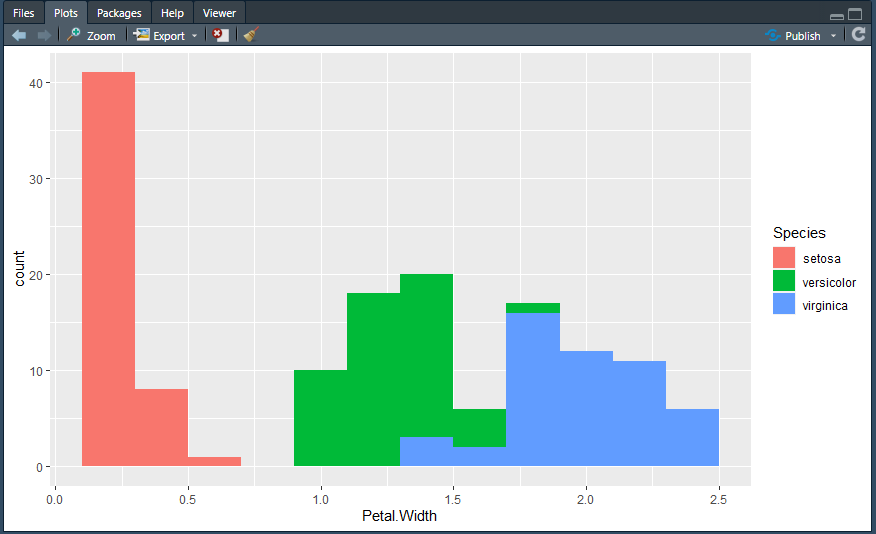
# Lab 2: Working With Data - Date Completed: 10/02/2022

1. A screenshot of running 1.4, how many mpg's was the Mazda RX-4? **- 21.0 mpg**
2. Add an image of, and the associated R code to create the bar plot from step 1.9.

barplot(mtcarsgear, main = "Car Gears", ylab = 'Cars', xlab = "Gears Type", col=c('green','red','yellow'))

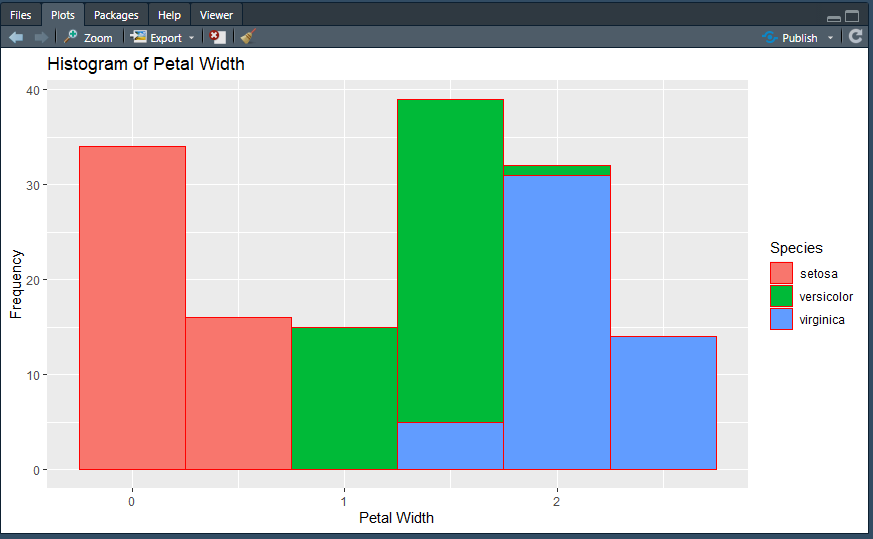


1. Add an image of the histogram from step 2.8.



1. Add an image of the chart from step 2.11, and add the associated code to create this. Add an explanation of what each part of the code does.

histogram + geom\_histogram(binwidth=0.5,color="red",aes(fill=Species)) + xlab("Petal Width") + ylab("Frequency") + ggtitle("Histogram of Petal Width")



Explanation of code -

histogram – variable defined earlier in code which is assigned a ‘ggplot’ graph based on iris’s PetalWidth values

geom\_histogram(binwidth=0.5,color="red",aes(fill=Species)) - displays the counts with bars. Width of bins is 0.5, outline of each bin is to be red. ‘aes’ = aesthetics; using the ‘Species’ data from ‘iris’ as the fill value groups the data together based on the values in that field and overlays each set of values.

+ xlab("Petal Width") - sets X-axis label

+ ylab("Frequency") - sets Y-axis label

+ ggtitle("Histogram of Petal Width") - sets title

## Weekly summary, complete the following:

This week I learned about ...

In the lab I …

I enjoyed ...

I found it challenging when ...

# Lab 3: Management And Security - Date Completed: 24/02/2022

* A screenshot of running 1.4, the head command running in part 1.4
* The finalised plot from part 1.8, with the code used to create it
* One of the word clouds created in step 2.5
* Find a news or magazine article, convert the article to a text file and use the skills from this lab to create a wordcloud of this.
* Add an explanation of why you chose the text that you have chosen, what are the most common words, was this expected?

## Weekly summary, complete the following:

This week I learned about ...

In the lab I …

I enjoyed ...

I found it challenging when ...

# Lab 4: Title - Date Completed: DD/MM/YYYY

## Weekly summary, complete the following:

This week I learned about ...

In the lab I …

I enjoyed ...

I found it challenging when ...