CS 422 Section 01

Josh Bowden

1.1. By taking this course, I foremost hope to learn how to use machine learning to create models and make sense of larger datasets. I also hope to learn how to use R more since I do have some experience with the Python side of things. I'd also hope to get a better understanding of how to use fundamental statistics for understanding your data.

Problem 1

1-A: Load the *cars* dataset and print it.

```
data(cars)
cars
```

```
##     speed dist
## 1       4    2
## 2       4   10
## 3       7    4
## 4       7   22
## 5       8   16
## 6       9   10
## 7      10   18
## 8      10   26
## 9      10   34
## 10     11   17
## 11     11   28
## 12     12   14
## 13     12   20
## 14     12   24
## 15     12   28
## 16     13   26
## 17     13   34
## 18     13   34
## 19     13   46
## 20     14   26
## 21     14   36
## 22     14   60
## 23     14   80
## 24     15   20
## 25     15   26
## 26     15   54
## 27     16   32
## 28     16   40
## 29     17   32
## 30     17   40
## 31     17   50
## 32     18   42
## 33     18   56
## 34     18   76
## 35     18   84
## 36     19   36
```

```
## 37      19    46
## 38      19    68
## 39      20    32
## 40      20    48
## 41      20    52
## 42      20    56
## 43      20    64
## 44      22    66
## 45      23    54
## 46      24    70
## 47      24    92
## 48      24    93
## 49      24   120
## 50      25    85
```

1-B: Plot the data in the *cars* dataset. Provide a title for the graph. The title should be "The Cars Dataset".

```r
plot(cars,
     main = "The Cars Dataset",
     xlab = "Speed (mph)",
     y = "Stopping Distance (ft)")
```
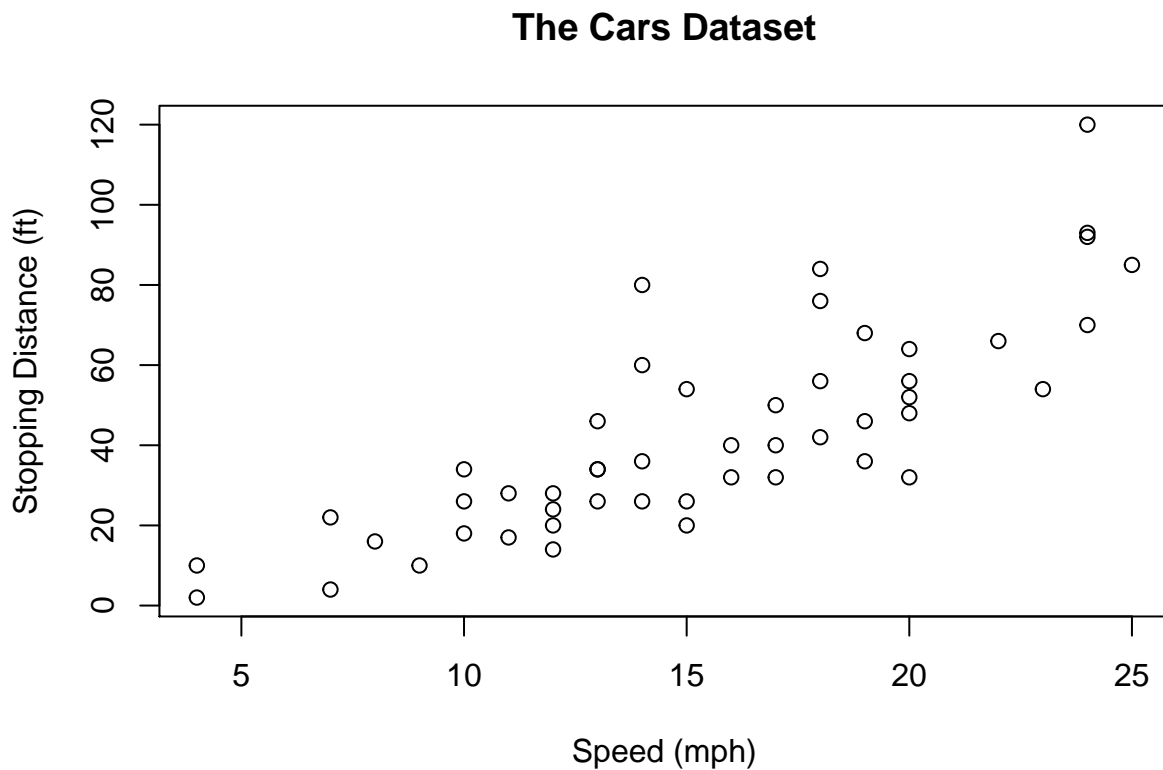
The Cars Dataset

1-C: Print a summary of the cars dataset using the R `summary()` command.

```r
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
```

```
##   1st Qu.:12.0    1st Qu.: 26.00
##   Median :15.0    Median : 36.00
##   Mean   :15.4    Mean   : 42.98
##   3rd Qu.:19.0    3rd Qu.: 56.00
##   Max.   :25.0    Max.   :120.00
```

1-D: What is the maximum speed and minimum distance as shown in the `summary()` command?

The minimum distance is 2ft while the maximum distance is 120ft.

Problem 2

2-A: Read the file into a data frame and add a new row consisting of the following data to the existing rows:

name: Brad Pitt
id: 40051
gpa: 2.21
grade: C

Before adding the student:

```
students <- read.csv("student.csv", sep=";", header=T)
students
```

```
##                    name    id  gpa grade
## 1       Robert DeNiro  9711 3.44     B
## 2 Robert Downey, Jr 33771 2.00     C
## 3       Samantha Bee 61094 4.00     A
## 4       Jason Bateman 45003 1.00     E
## 5        Alicia Keyes 55101 3.92     B
## 6 Denzel Washington 71005 4.00     A
## 7   Joaquin Phoenix 41180 2.89     C
```

After adding the new student:

```
brad_pitt <- data.frame(name = "Brad Pitt",
                        id = 40051,
                        gpa = 2.21,
                        grade = "C")
students.new <- rbind(students, brad_pitt)
students.new
```

```
##                    name    id  gpa grade
## 1       Robert DeNiro  9711 3.44     B
## 2 Robert Downey, Jr 33771 2.00     C
## 3       Samantha Bee 61094 4.00     A
## 4       Jason Bateman 45003 1.00     E
## 5        Alicia Keyes 55101 3.92     B
## 6 Denzel Washington 71005 4.00     A
## 7   Joaquin Phoenix 41180 2.89     C
## 8           Brad Pitt 40051 2.21     C
```