Module 8 Assignment

– *Gregory Foster* –

Step 1

Importing the Dataset with the read.table() function

```
x= read.table("Assignment 6 Dataset.txt", header=TRUE, sep=",")
x
```

```
##          Name Age    Sex Grade
## 1        Raul  25   Male    80
## 2     Booker  18   Male    83
## 3      Lauri  21 Female    90
## 4     Leonie  21 Female    91
## 5    Sherlyn  22 Female    85
## 6    Mikaela  20 Female    69
## 7    Raphael  23   Male    91
## 8       Aiko  24 Female    97
## 9   Tiffaney  21 Female    78
## 10    Corina  23 Female    81
## 11 Petronila  23 Female    98
## 12    Alecia  20 Female    87
## 13   Shemika  23 Female    97
## 14    Fallon  22 Female    90
## 15   Deloris  21 Female    67
## 16    Randee  23 Female    91
## 17     Eboni  20 Female    84
## 18   Delfina  19 Female    93
## 19 Ernestina  19 Female    93
## 20      Milo  19   Male    67
```

Step 2

Using the plyr package to add a Grade.Average column to the data-frame

```
library(plyr)
y = ddply(x, "Sex", transform, Grade.Average=mean(Grade))
y
```

```
##          Name Age    Sex Grade Grade.Average
## 1       Lauri  21 Female    90       86.9375
## 2      Leonie  21 Female    91       86.9375
## 3     Sherlyn  22 Female    85       86.9375
## 4     Mikaela  20 Female    69       86.9375
## 5        Aiko  24 Female    97       86.9375
## 6    Tiffaney  21 Female    78       86.9375
## 7      Corina  23 Female    81       86.9375
## 8   Petronila  23 Female    98       86.9375
## 9      Alecia  20 Female    87       86.9375
## 10    Shemika  23 Female    97       86.9375
## 11     Fallon  22 Female    90       86.9375
## 12    Deloris  21 Female    67       86.9375
## 13     Randee  23 Female    91       86.9375
## 14      Eboni  20 Female    84       86.9375
## 15    Delfina  19 Female    93       86.9375
## 16 Ernestina  19 Female    93       86.9375
## 17       Raul  25   Male    80       80.2500
## 18     Booker  18   Male    83       80.2500
## 19    Raphael  23   Male    91       80.2500
## 20       Milo  19   Male    67       80.2500
```

Step 3

Printing the updated dataset to a local file

```
write.table(y, "Sorted_Average")
```

Converting the file to CSV with an additional function argument

```
write.table(y, "Sorted_Average", sep=",")
```

Step 4

Filtering the dataset to only include names that contain the letter “i”, using a combination of the subset() and grepl() functions

```
newx = subset(x, grepl("[iI]", x$Name))
newx
```

```
##          Name Age    Sex Grade
## 3       Lauri  21 Female    90
## 4      Leonie  21 Female    91
## 6     Mikaela  20 Female    69
## 8        Aiko  24 Female    97
## 9    Tiffaney  21 Female    78
## 10     Corina  23 Female    81
## 11 Petronila  23 Female    98
## 12     Alecia  20 Female    87
## 13    Shemika  23 Female    97
## 15    Deloris  21 Female    67
## 17      Eboni  20 Female    84
## 18    Delfina  19 Female    93
## 19 Ernestina  19 Female    93
## 20       Milo  19   Male    67
```

Writing the 'newx' subset data to a new local file

```
write.table(newx,"DataSubset", sep=",")
```