**Module # 8 Input/Output, string manipulation and plyr package**

# Load required packages

install.packages("plyr")

library(data.table)

library(plyr)

# Import dataset

Data\_Student <- read.table("C:/Users/kc290/OneDrive/Desktop/Intro to R/Assignment 6 Dataset.txt", header = TRUE, sep = ",")

# Calculate mean using Sex as the category

Student\_Average <- ddply(Data\_Student, "Sex", summarise, Grade.Average = mean(Grade))

# Write output to a file

write.table(Student\_Average, "StudentAverage.txt", sep = ",", row.names = FALSE)

# Filter the original data set to include only data for which the student name contains the letter "i"

i\_students <- subset(Data\_Student, grepl("i", Data\_Student$Name, ignore.case = TRUE))

# Write names containing "i" to a CSV file

write.csv(i\_students$Name, "i\_students\_names.csv", row.names = FALSE)

# Write the filtered dataset to a CSV file using file.choose() and subset()

write.csv(i\_students, "Datasubset.csv",sep = ",")

Adding the Necessary Packages:  
The required packages are loaded and installed in this section. The commands install.packages("plyr") and library(plyr) are used, respectively, to install and load the plyr package. The library is loaded into the data.table package (data.table).

Bringing in the Dataset:  
The read.table() function is used to import the dataset. It is specified that the file includes a header and that commas are used as separators, along with the file path. A variable called Data\_Student holds the dataset.

Finding the Mean with Sex as the Category:  
To find the mean of the 'Grade' variable classified by the 'Sex' category, utilize the ddply() method from the plyr package. The outcome is kept in Student\_Average, a data frame.

Composing Results to a File:  
The Student\_Average data frame is written to a text file called "StudentAverage.txt" using the write.table() function. Row names are excluded (row.names = FALSE) and the data is separated by commas (sep = ",").

Data Filtering to Remove 'i' from Names:  
The original dataset (Data\_Student) is filtered using the subset() method to only include rows where the letter 'i' appears in the 'Name' column. An whole new data frame called i\_students contains this portion.

Transferring Names with a 'i' to a CSV File:  
The 'Name' column of the i\_students data frame is written to a CSV file called "i\_students\_names.csv" using the write.csv() function. Row names are not included in the result.

Writing a CSV File with a Filtered Dataset Making use of file.Subset() and choose():  
The i\_students data frame is written to a CSV file called "Datasubset.csv" using the write.csv() function. The user is prompted to select the location where the file will be saved using the file.choose() function. Commas are used to separate the data (sep = ",").