ALY 6000, Winter 2022

Module 1: Executive Summary Report 1

Submitted by: Abhinav Jain

NUID: 002938209

Submitted to: Richard Zhi

Date: 01/18/2022

Introduction:

In the report below we are going to solve questions related to R. This assignment consisting of 14 problems given in the module 1 assignment. We have used R-studio to solve this assignment.
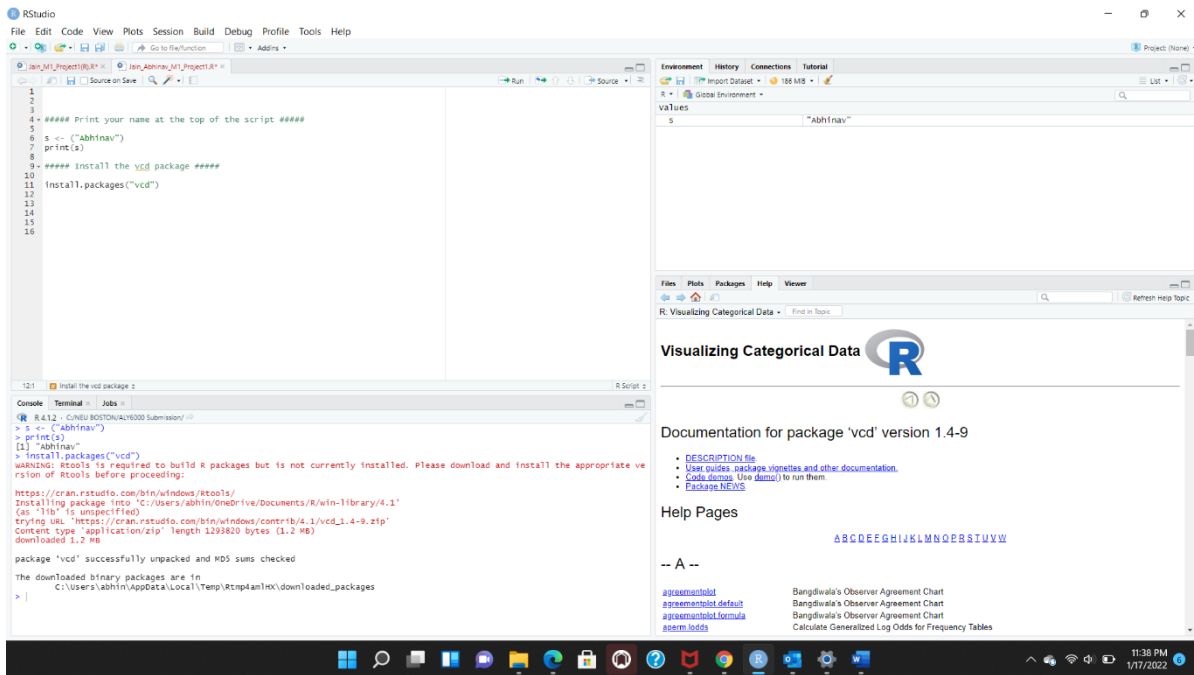
Assignment Problems:

Below are the step-by-step solutions for the given problems.

Task 1: Syntax to Print name on the top of the script [4]

Defined variable name 'x' then print function, print (), name displayed in console
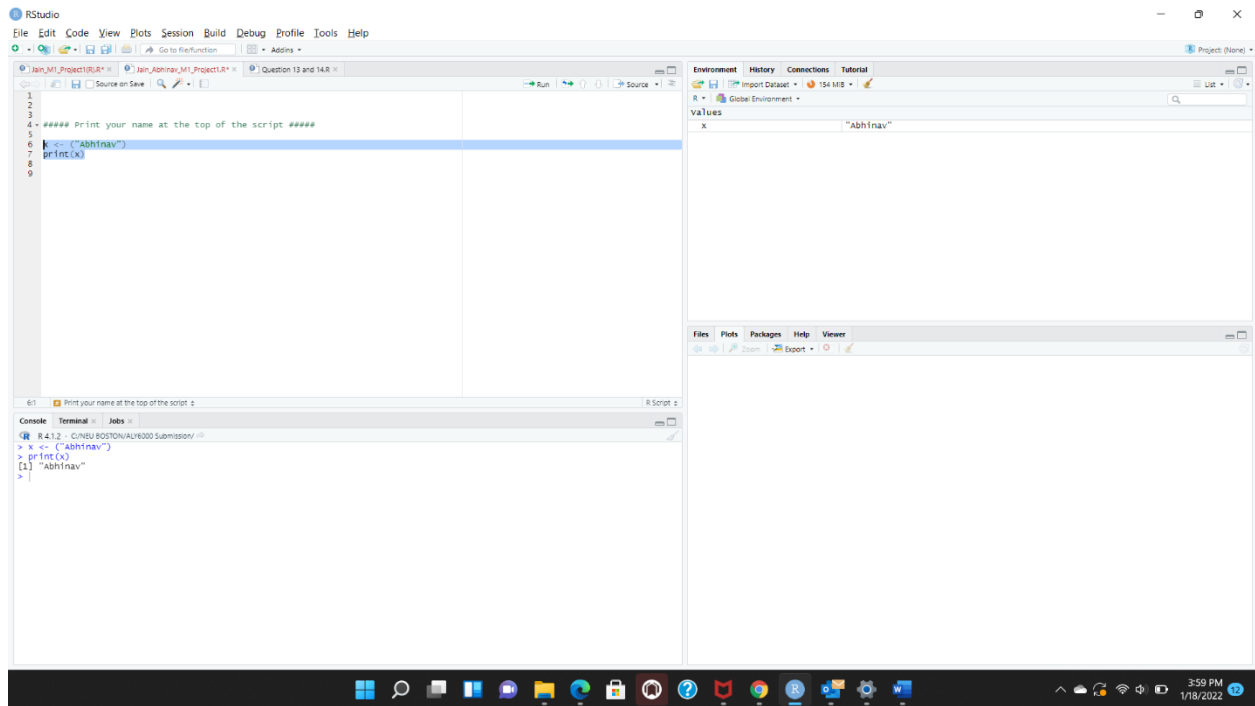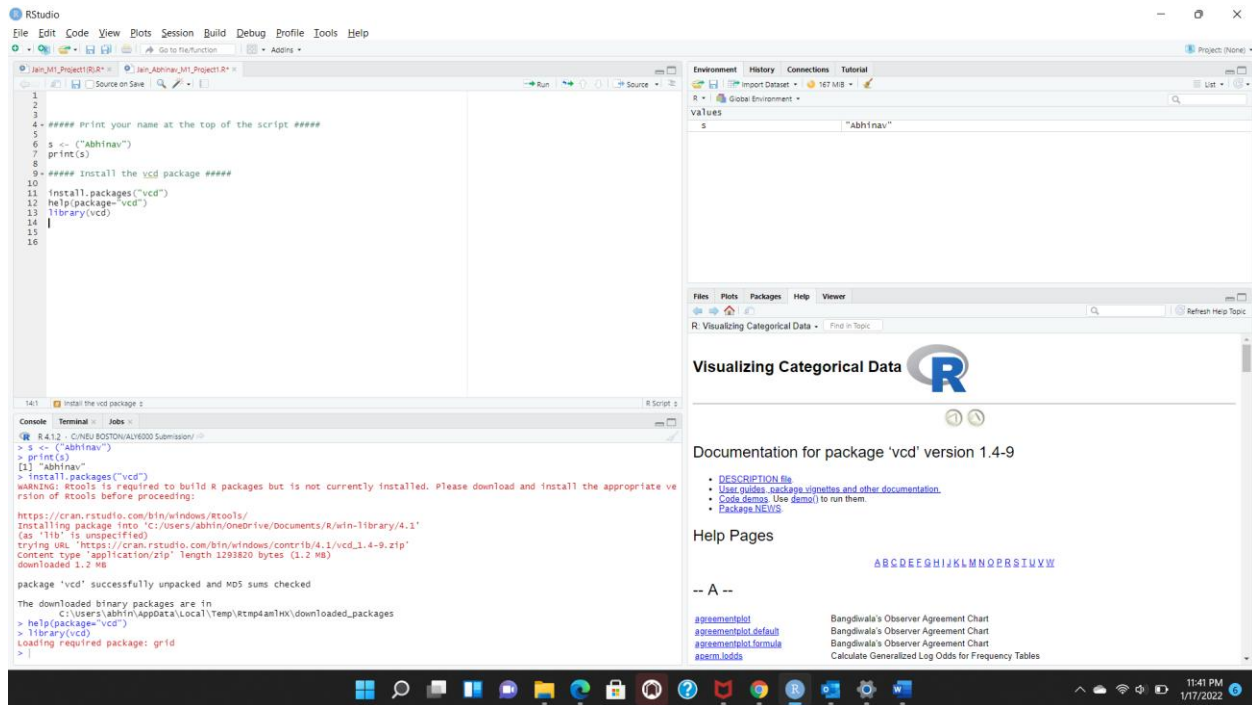
Task 2: Install vcd Package [2] In the task below we have installed "vcd" package with function, install.package(" vcd").
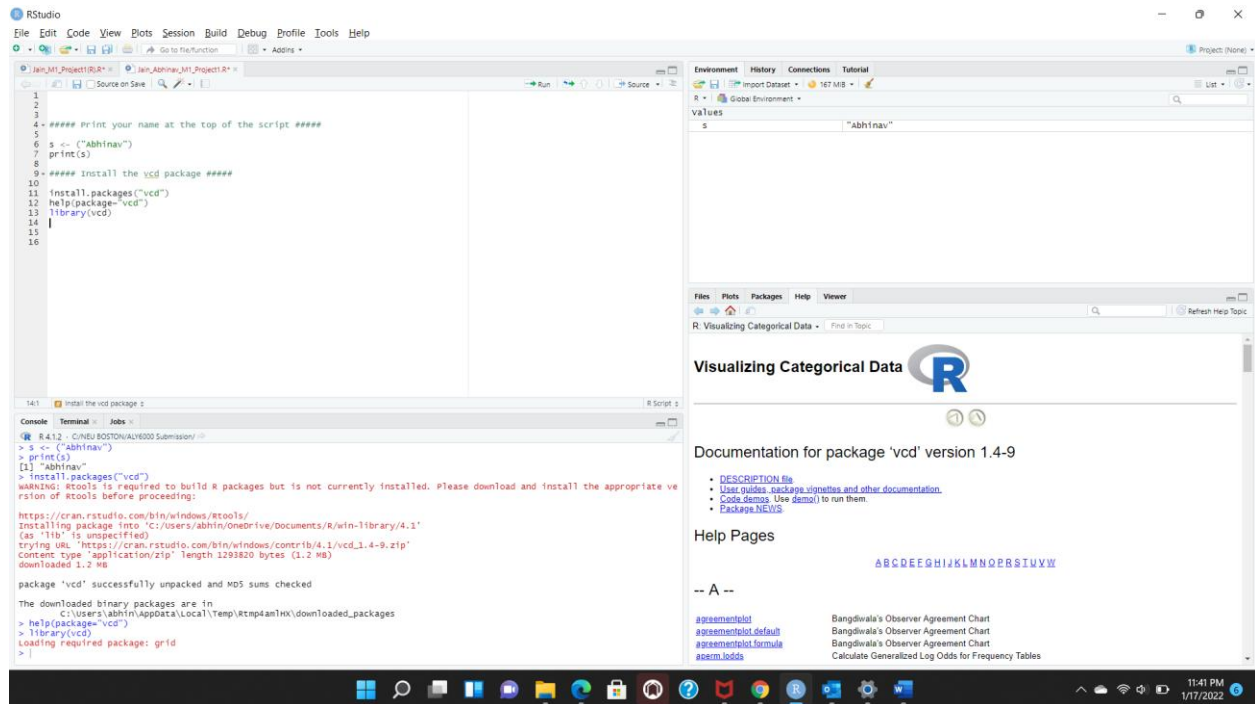

Task 3: Import the vcd library [2]

This library can be used to visualize categorial data to plot graphs and tables.
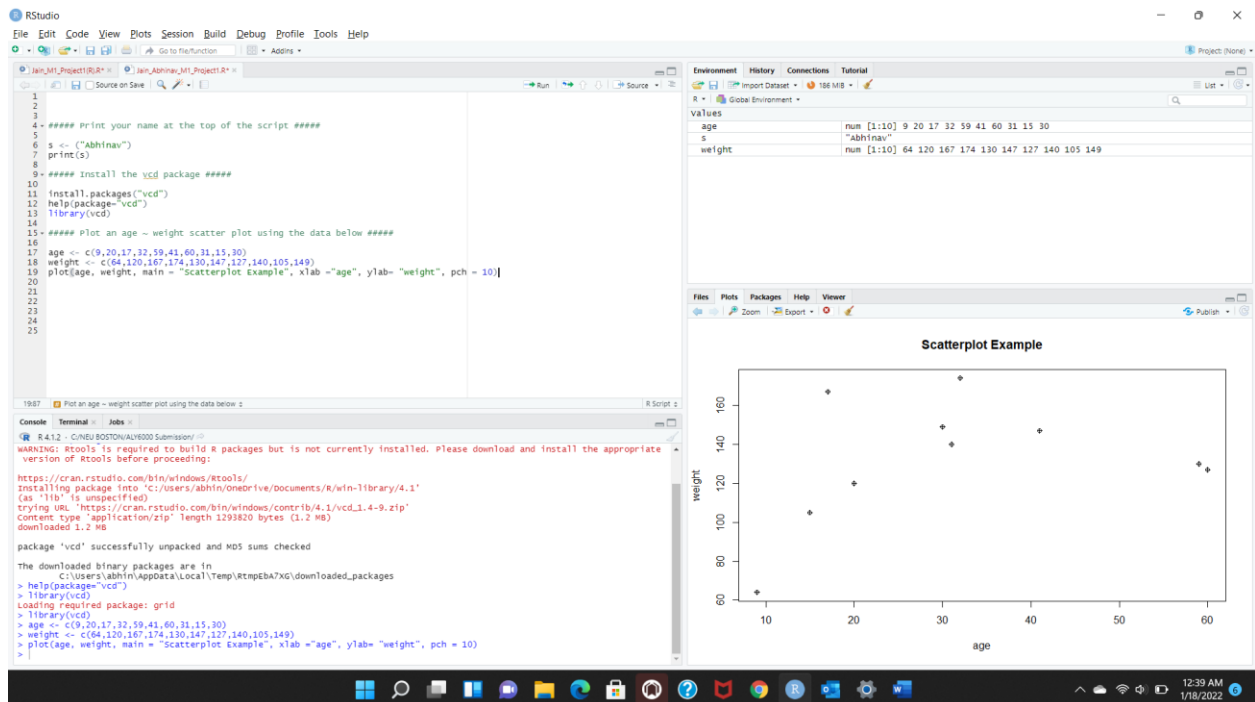
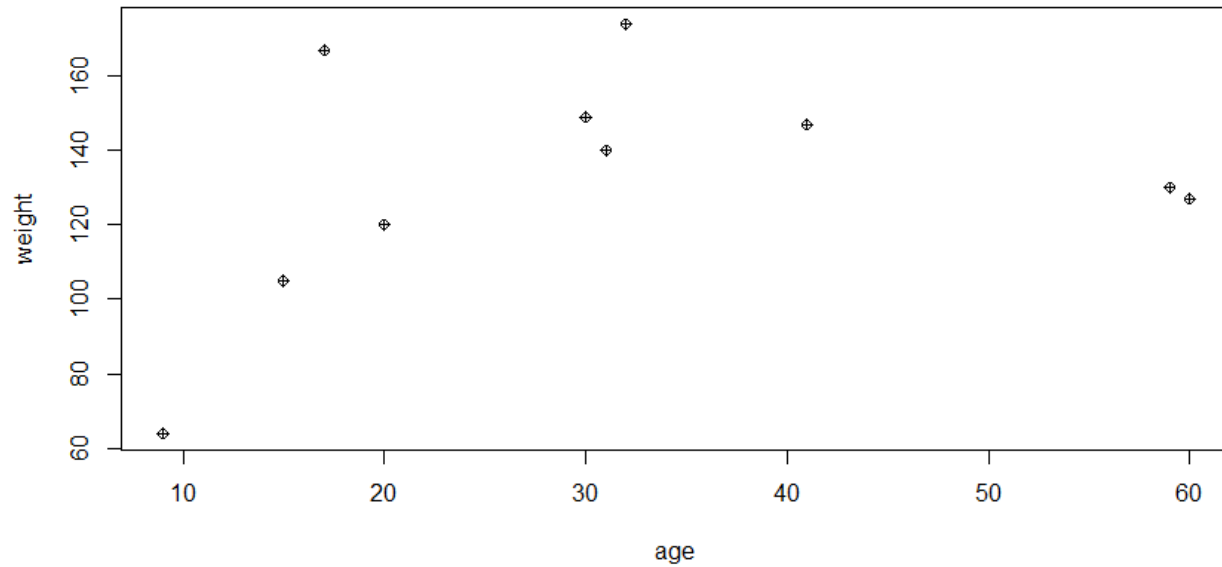Screenshot - After the installation of vcd package

Task 4: Plot an age ~weight scatter plot using the data [11]

We found the Correlation between age and weight by defining the variable x and y which represent the horizontal and vertical axis of the graph respectively.
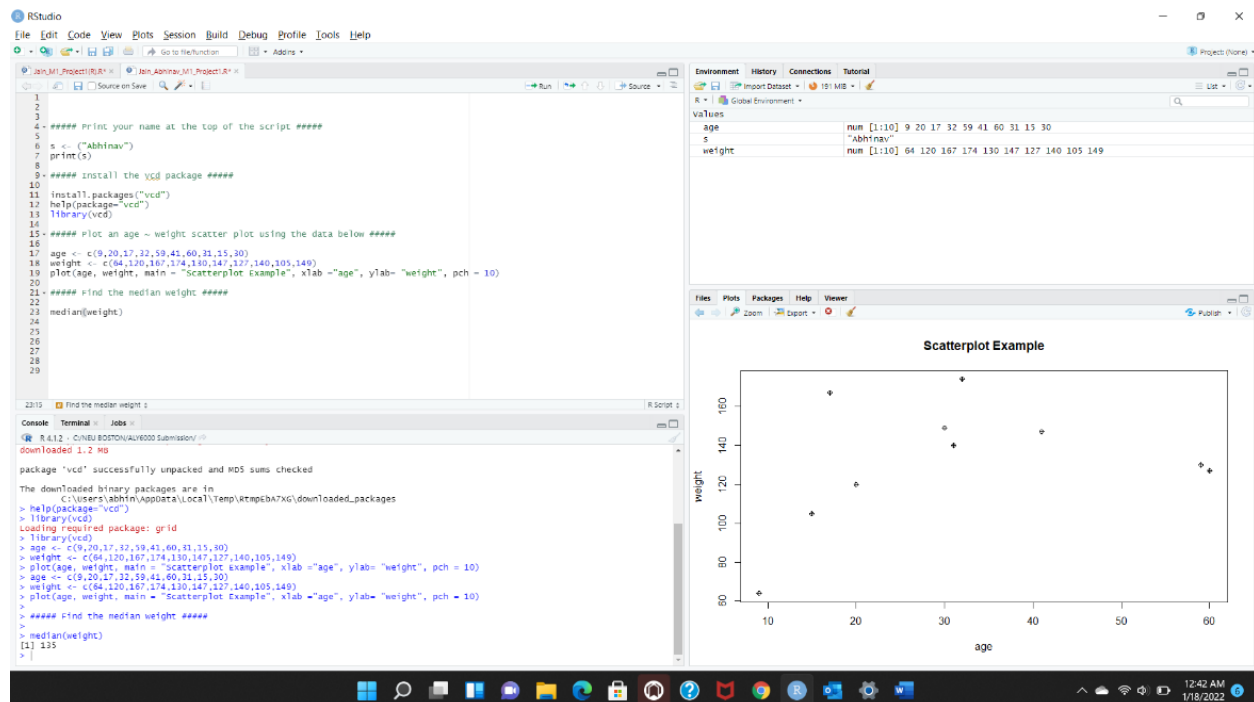
Scatterplot Example

Task 5: Find Median of Weight

In this task we are trying to find the middle value of the dataset with the function median (). Below is the solution to find the median of weight variable.

Task 6: Delete the 5th element from the age vector [1]

In the task below we have used age[-5] to delete 5th element from the age vector.
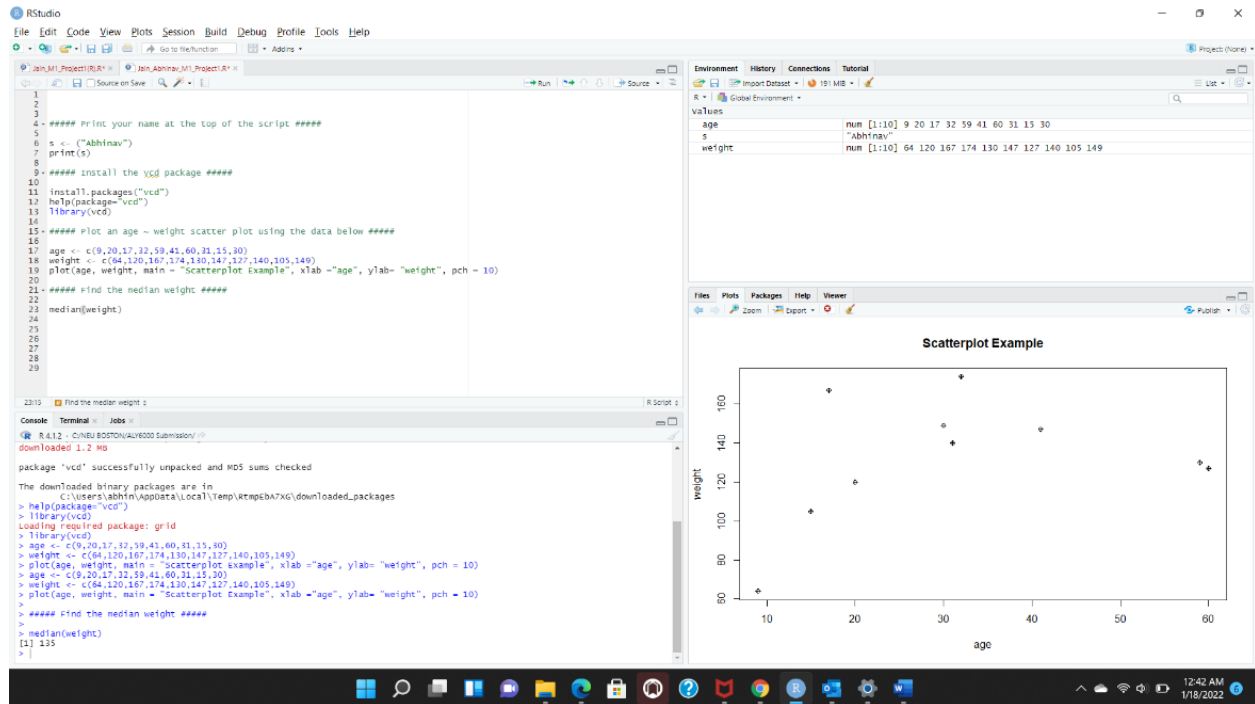
Task 7: Insert 78 as the 5th element into the age vector [7]

In the task below we have used append (), function to insert number 78 at 5 positions of the vector.
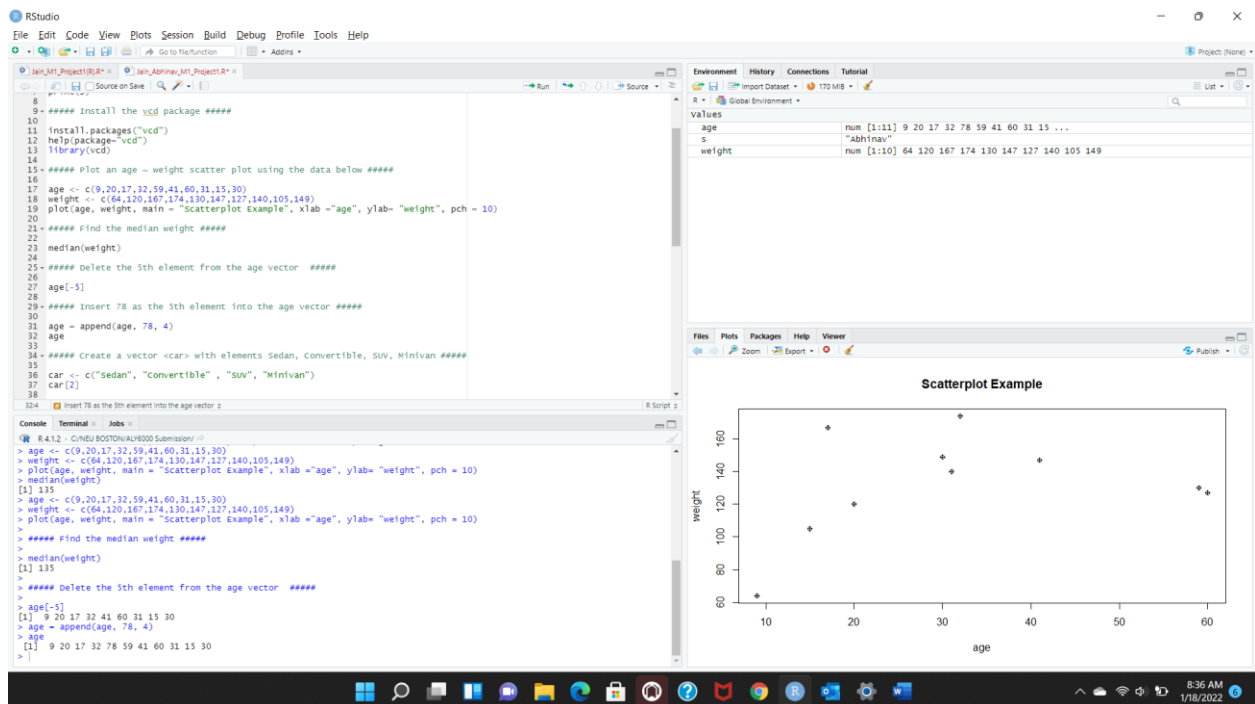
Task 8: Create a vector <car> with elements Sedan, Convertible, SUV, Minivan

In the task below we have created vector "car" with the given elements
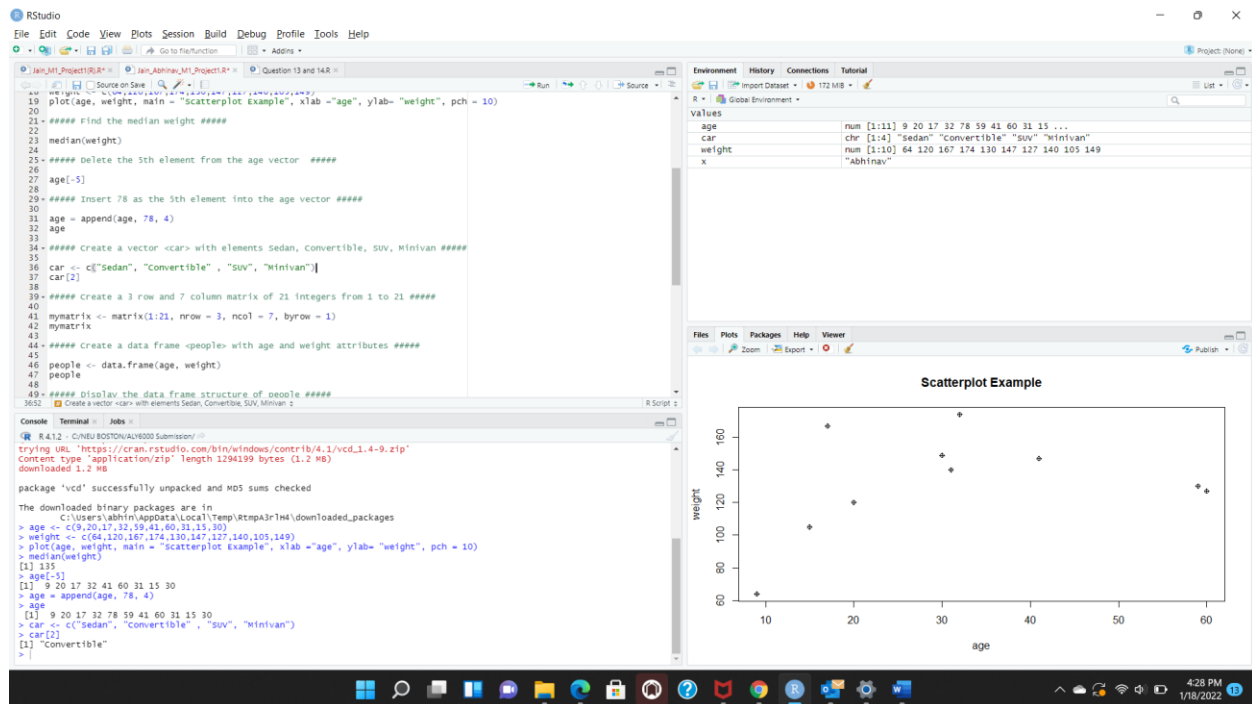
Task 9: In the given task we have created a matrix of 3 Row and 7 Column with 21 integers from 1 to 21 [6]
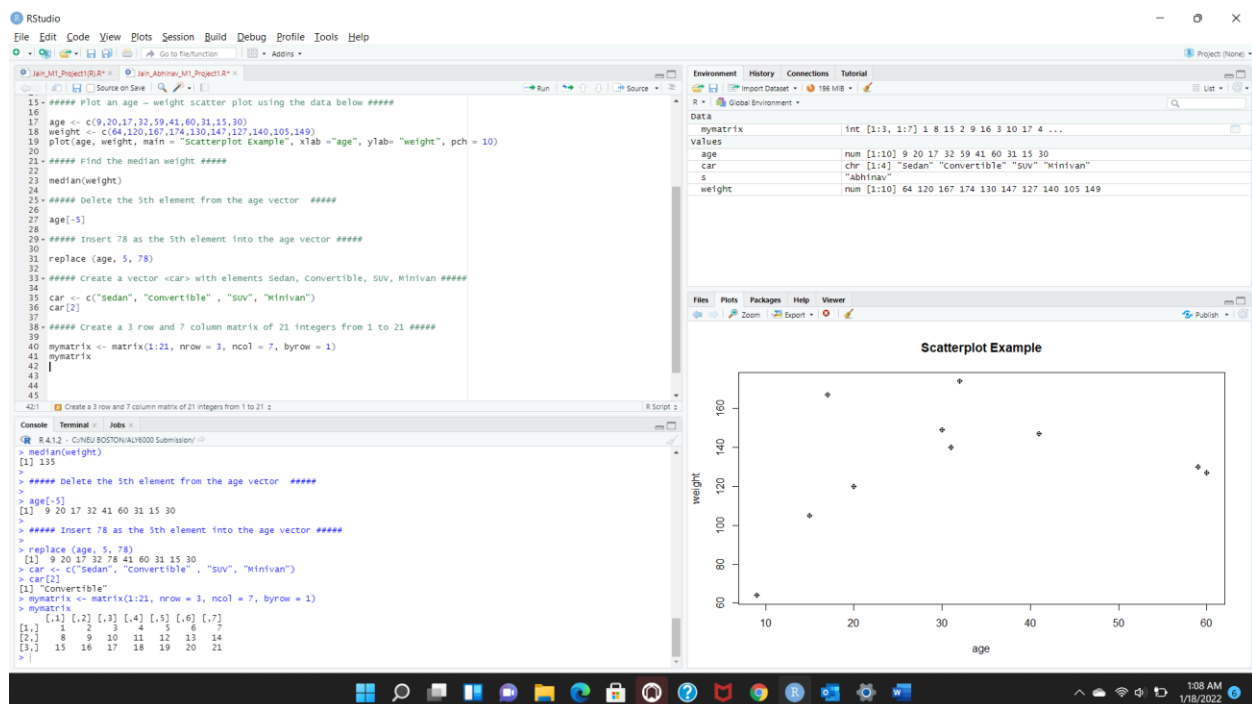
Task 10: In the given task we have created a data frame "people" with the attributes named as age and weight
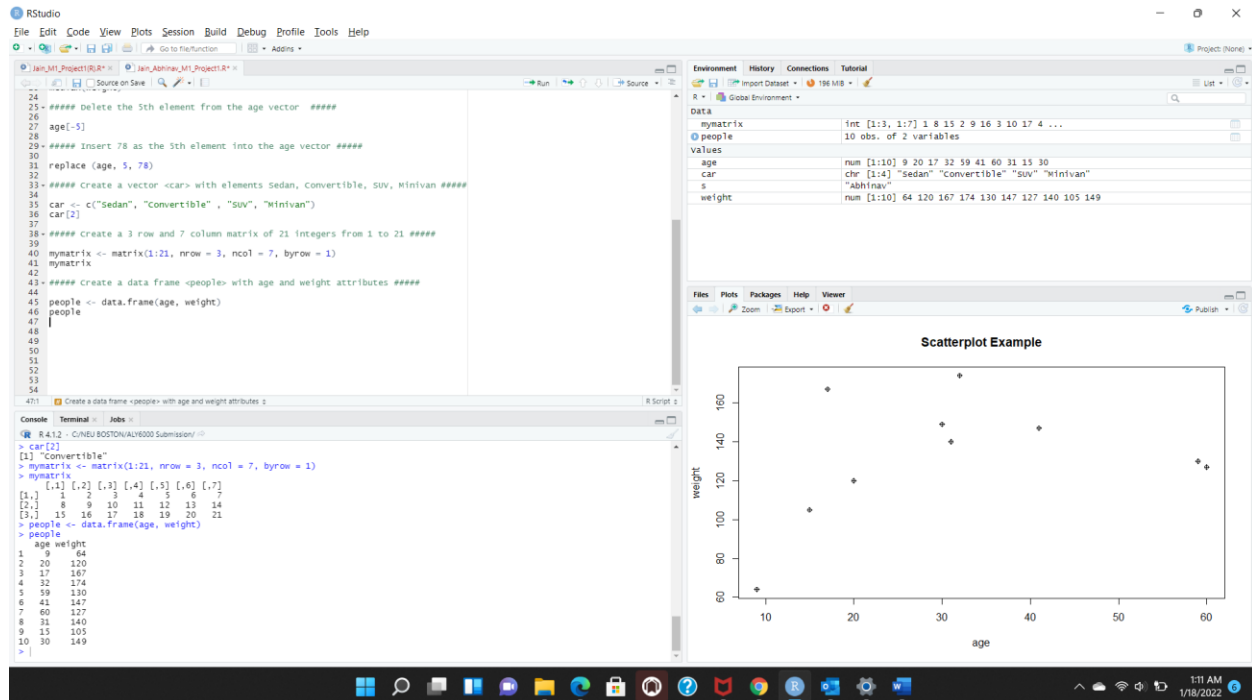
Task 11: In the given task we have represented the structure of the people's data frame[5]
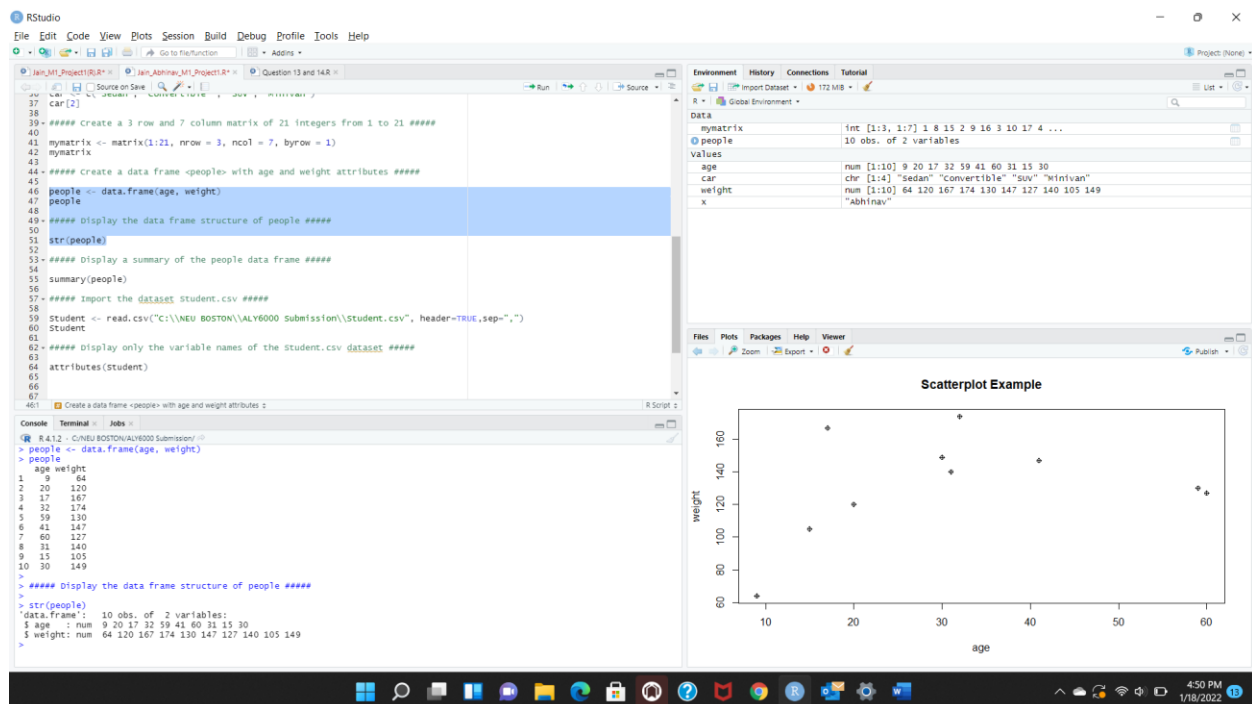
Task 12: In the given task we have extract the summary of the people' data frame [5]
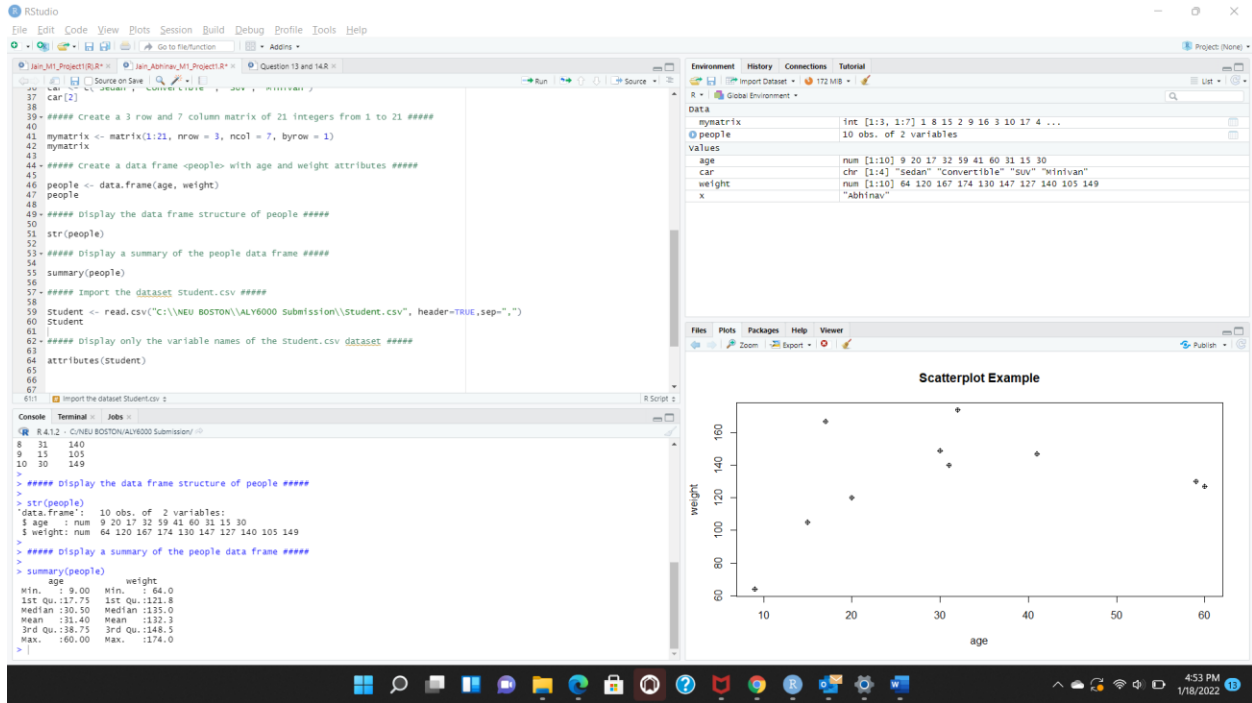
Task 13: In the given task we have Imported dataset from Student.csv [3]

Task 14: In this given task we have shown the variable name of the student datasets [8] [10]

Summary:

Here, through this assignment want to share with my experience that based on the dataset and instructions given I would like to summarize what is achievable. Through this assignment, I want to summarize the key take away after completing the given task. In this Project, the study suggested that dataset is the most significant information through which we can analyze the stat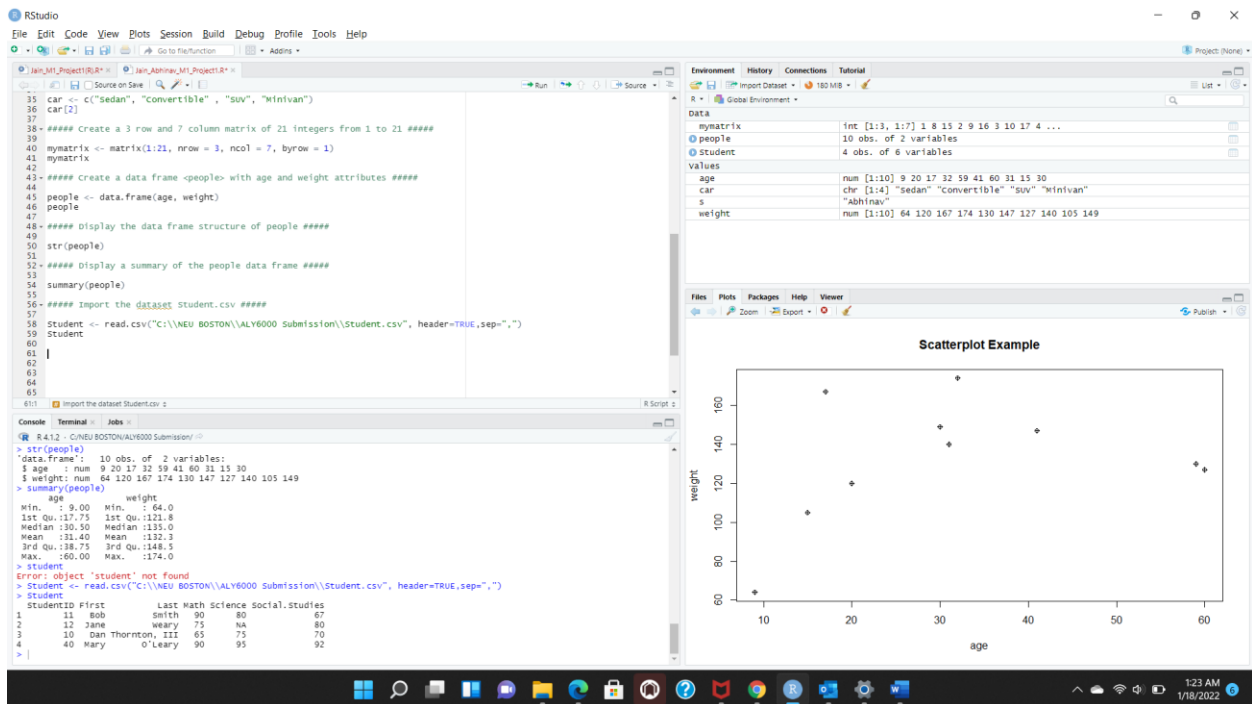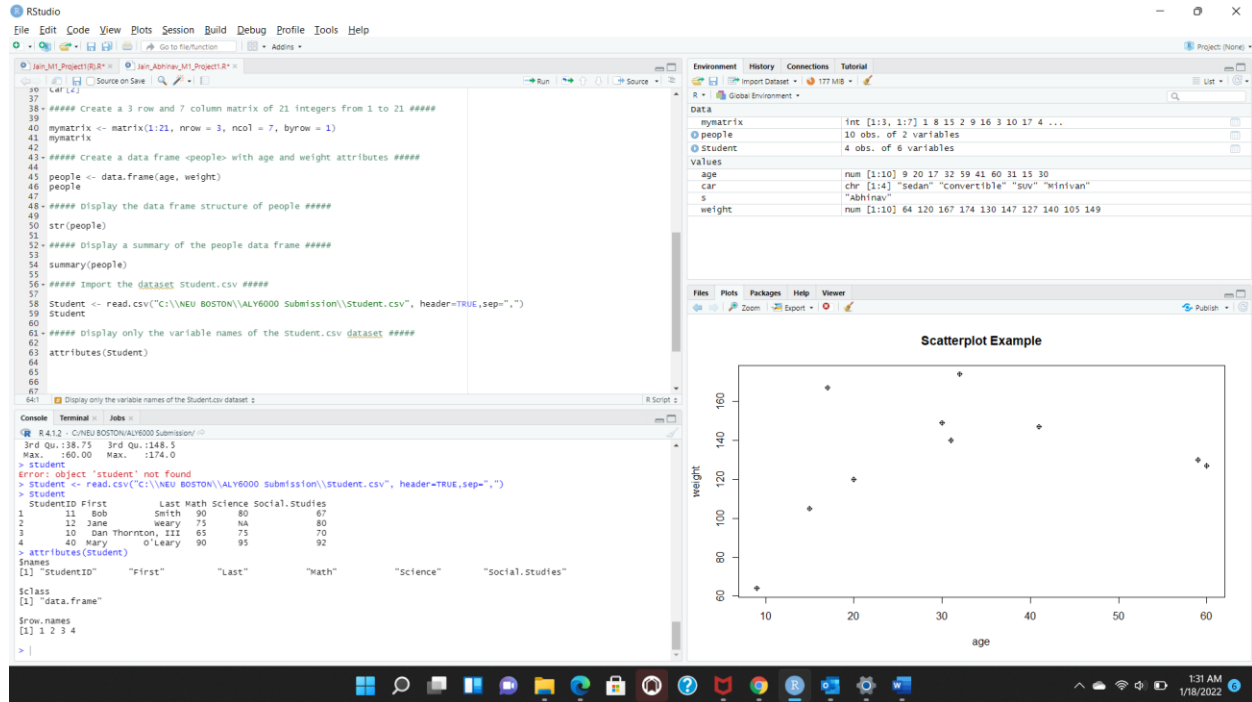istical information. I started with the basics like print command with my name, understanding about how to install package and library (like vcd package). These libraries help in plotting Graphs and functions which act like a tool. Furthermore, the age and weight were corelated, which shows how we can compare weight with the age. Let's understand it in a most simplified way. If we take the data from population of a country, through the scatter plot we can analyze the weight of a particular age group. In 'R' programming, here I performed many functions like median of weight, deleted elements from dataset, insert elements between the dataset or vector, likewise we can perform many functions. Apart from this, to understand about how to create a vector from the data, I created the car vector to analyze the elements of the convertible car. Moreover, created 3 by 7 matrix of 21 integers form to display the data in "myMatrix". Applied data frame function to build data frame by using given data points of age and weight. Used summary () to find the summary of the data frame. In the given task, the file was Imported by using read.csv ("file path"). Lastly, I have imported the file of the dataset and displayed variables of the "student" dataset. Performing these tasks gave me the hands-on experience in learning R.

The key takeaways from this assignment are as follows:

- install packages
- import library
- plot graphs
- median
- insert the element
- delete the element
- create vector-matrix-data frame
- extract the summary

Bibliography

This includes YouTube videos, instruction materials, google search results, texts that informed your study of statistics and R. Adhere to APA standards.

[0]STAT 4772/5772 Fall 2015 - University of Northern Iowa
Computational Statistics - STAT 4772/STAT 5772 - Fall 2015. Normam Matloff The Art of R Programming: A Tour of Statistical Software Design. R/SPSS: Albums assignment due on December 28th. SPSS and R and Linaar and Multiple Regression along with Scatterplots. Predicting record album sales.

www.cs.uni.edu/~jacobson/4772

[1]R Vector: Create, Modify and Access Vector Elements

https://www.datamentor.io/r-programming/vector/

[2]How to install HH package?
Sky1014
https://community.rstudio.com/t/how-to-install-hh-package/68765

[3]Importing Data
robk@statmethods.net

[4]Print

https://www.statmethods.net/input/importingdata.html

[5]Data Frames

https://www.tutorialspoint.com/r/r_data_frames.htm

[6]What's the difference between integer class and numeric class in R
KeonKeon 1 et al.

https://stackoverflow.com/questions/23660094/whats-the-difference-between-integer-class-and-numeric-class-in-r#:~:text=R%20handles%20the%20differences%20between,for%20you%20in%20the%20background.&text=(Putting%20capital%20'L'%20after,a%20subset%20of%20%22numeric%22.&text=Integers%20only%20go%20to%20a,numerics%20can%20be%20much%20bigger.

[7]Here we go again: append an element to a list in R
user443854user443854 6 et al.
https://stackoverflow.com/questions/17046336/here-we-go-again-append-an-element-to-a-list-in-r

[8]CSV Files
https://www.tutorialspoint.com/r/r_csv_files.htm

[9]How to extract the text description of variables under variable name in R?
Hao et al.
https://stackoverflow.com/questions/55266645/how-to-extract-the-text-description-of-variables-under-variable-name-in-r

[10] 'Incomplete final line' warning when trying to read a .csv file into R
KateKate 1 et al.

https://stackoverflow.com/questions/5990654/incomplete-final-line-warning-when-trying-to-read-a-csv-file-into-r

[11] Scatter plots in R Language

https://www.geeksforgeeks.org/scatter-plots-in-r-language/#:~:text=A%20scatter%20plot%20is%20a,reveals%20a%20correlation%20between%20them.

Appendix: The R Script you wrote and executed

```r
##### Print your name at the top of the script #####

x <- ("Abhinav")

print(x)

##### Install the vcd package #####

install.packages("vcd")

help(package="vcd")

library(vcd)

##### Plot an age ~ weight scatter plot using the data below #####

age <- c(9,20,17,32,59,41,60,31,15,30)

weight <- c(64,120,167,174,130,147,127,140,105,149)

plot(age, weight, main = "Scatterplot Example", xlab ="age", ylab= "weight", pch = 10)

##### Find the median weight #####

median(weight)

##### Delete the 5th element from the age vector  #####

age[-5]

##### Insert 78 as the 5th element into the age vector #####

age = append(age, 78, 4)

age

##### Create a vector <car> with elements Sedan, Convertible, SUV, Minivan #####

car <- c("Sedan", "Convertible" , "SUV", "Minivan")

car[2]

##### Create a 3 row and 7 column matrix of 21 integers from 1 to 21 #####

mymatrix <- matrix(1:21, nrow = 3, ncol = 7, byrow = 1)

mymatrix
```

Create a data frame <people> with age and weight attributes

people <- data.frame(age, weight)

people

Display the data frame structure of people

str(people)

Display a summary of the people data frame

summary(people)

Import the dataset Student.csv

Student <- read.csv("C:\\NEU BOSTON\\ALY6000 Submission\\Student.csv", header=TRUE,sep=",")

Student

Display only the variable names of the Student.csv dataset

attributes(Student)