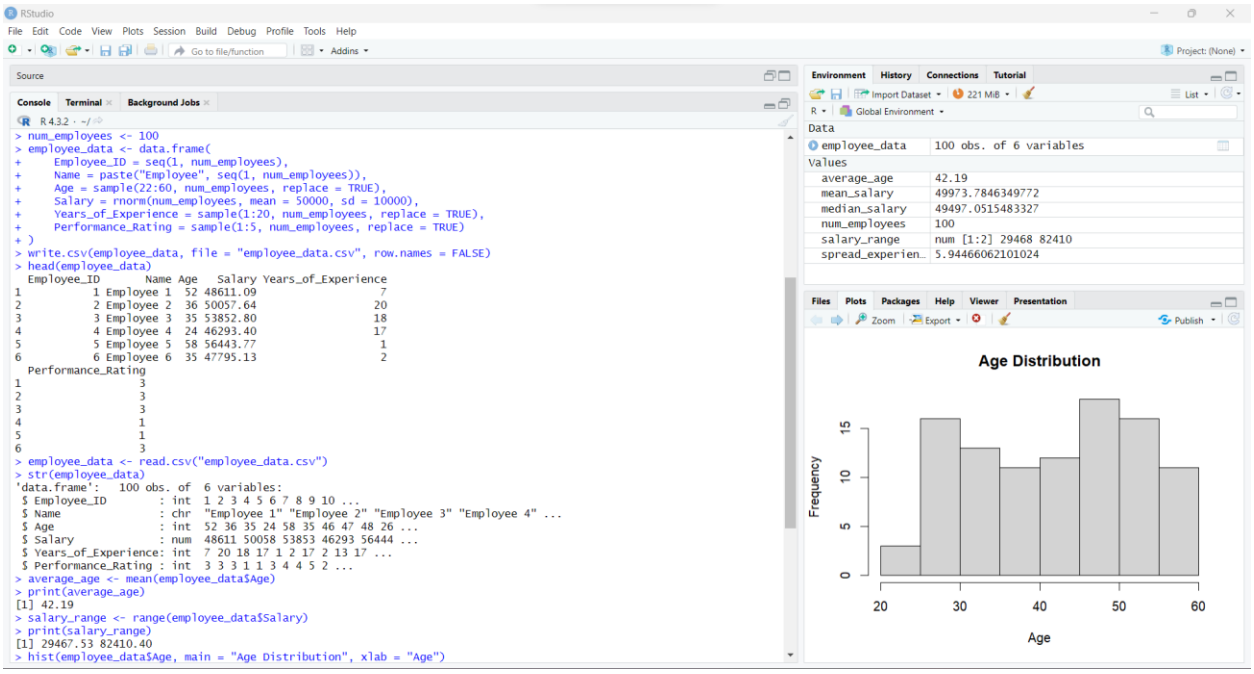
Pehesara A.J.P.G - 21001316

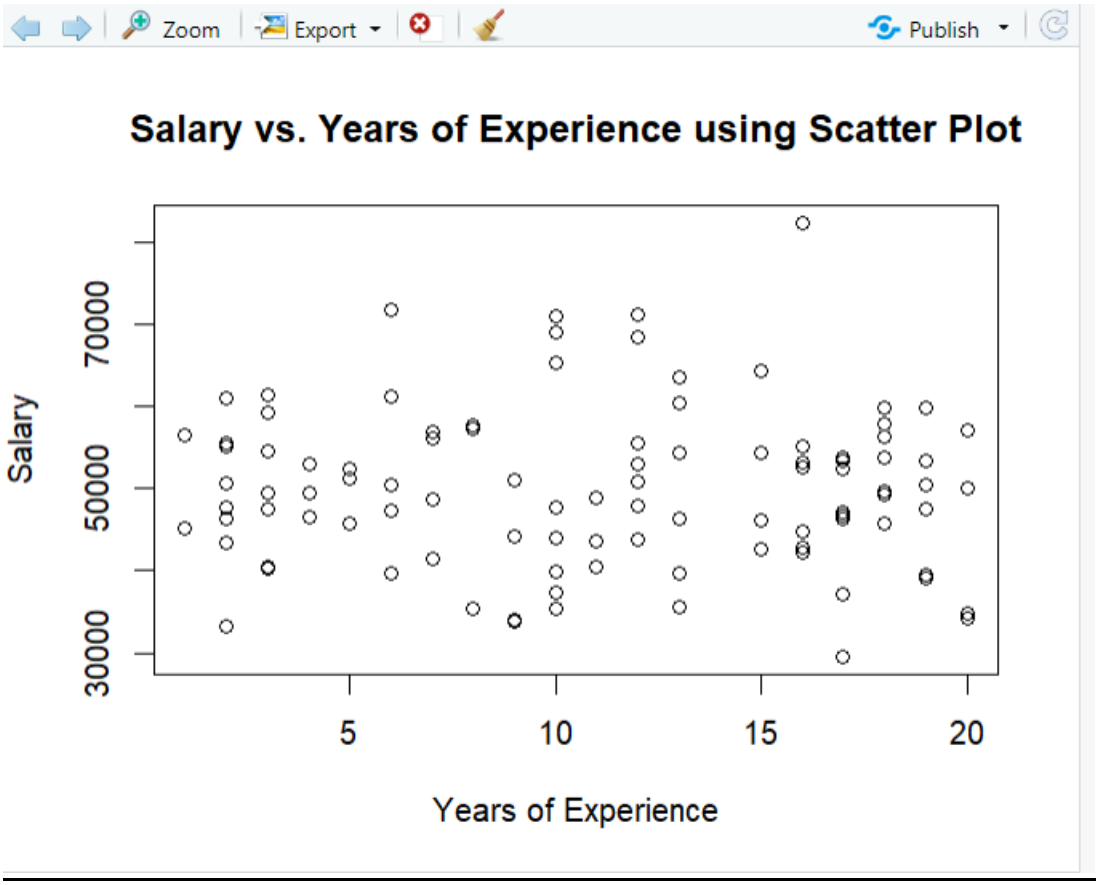
SCS2211 – Laboratory ||

In Class Assingment 1

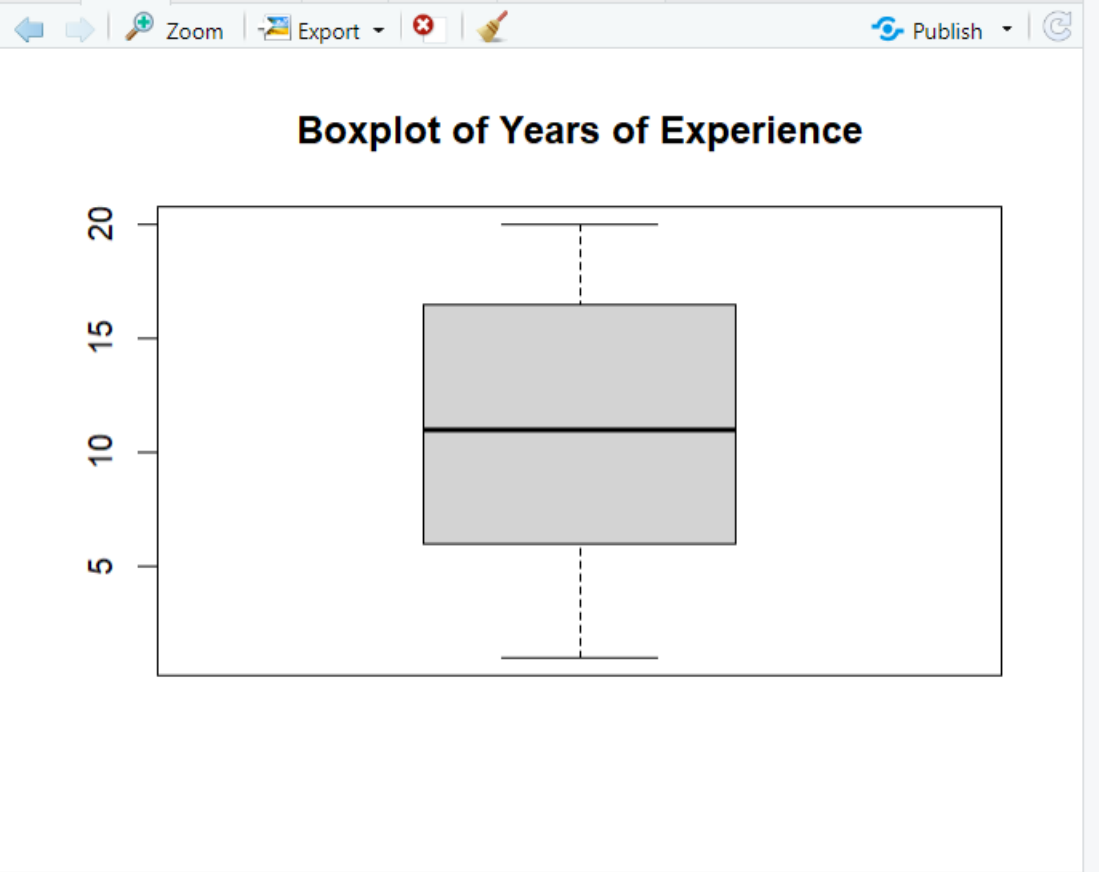
Question 1



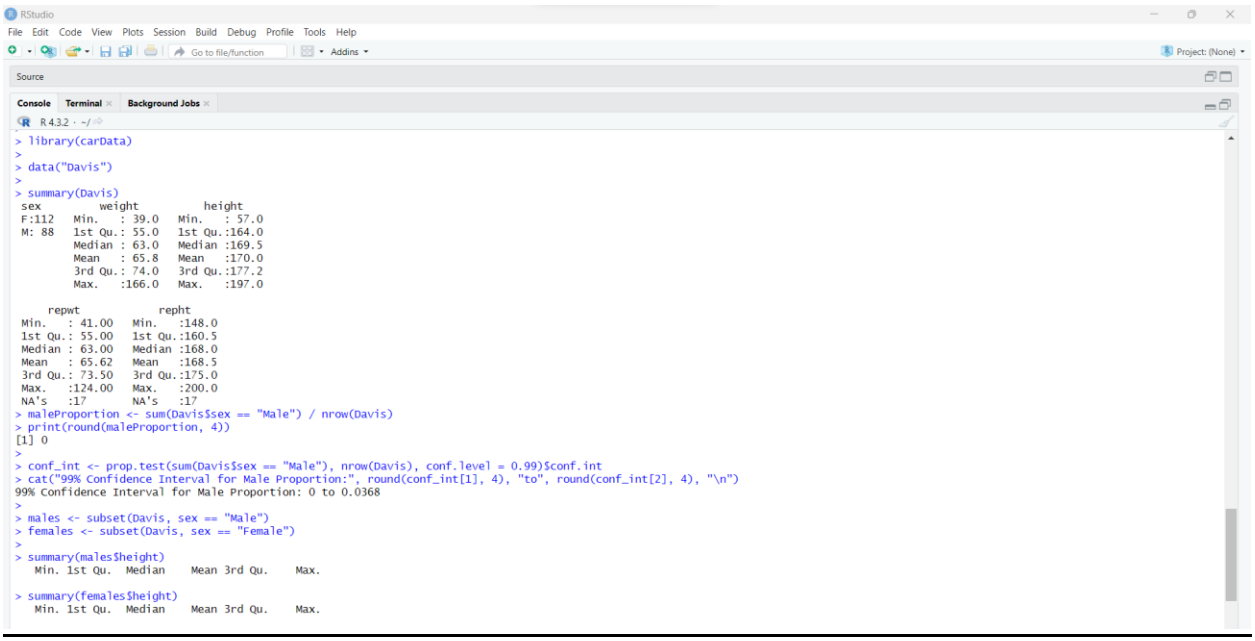

```
> plot(employee_data$Years_of_Experience, employee_data$Salary,
+      main = "Salary vs. Years of Experience using Scatter Plot",
+      xlab = "Years of Experience", ylab = "Salary")
> mean_salary <- mean(employee_data$Salary)
> print(mean_salary)
[1] 49973.78
> median_salary <- median(employee_data$Salary)
> print(median_salary)
[1] 49497.05
```




```
> spread_experience <- sd(employee_data$Years_of_Experience)
>
> print(spread_experience)
[1] 5.944661
> boxplot(employee_data$Years_of_Experience, main = "Boxplot of Years of Experience")
>
```


Zoom
Export
Publish
Boxplot of Years of Experience
20
15
10
5

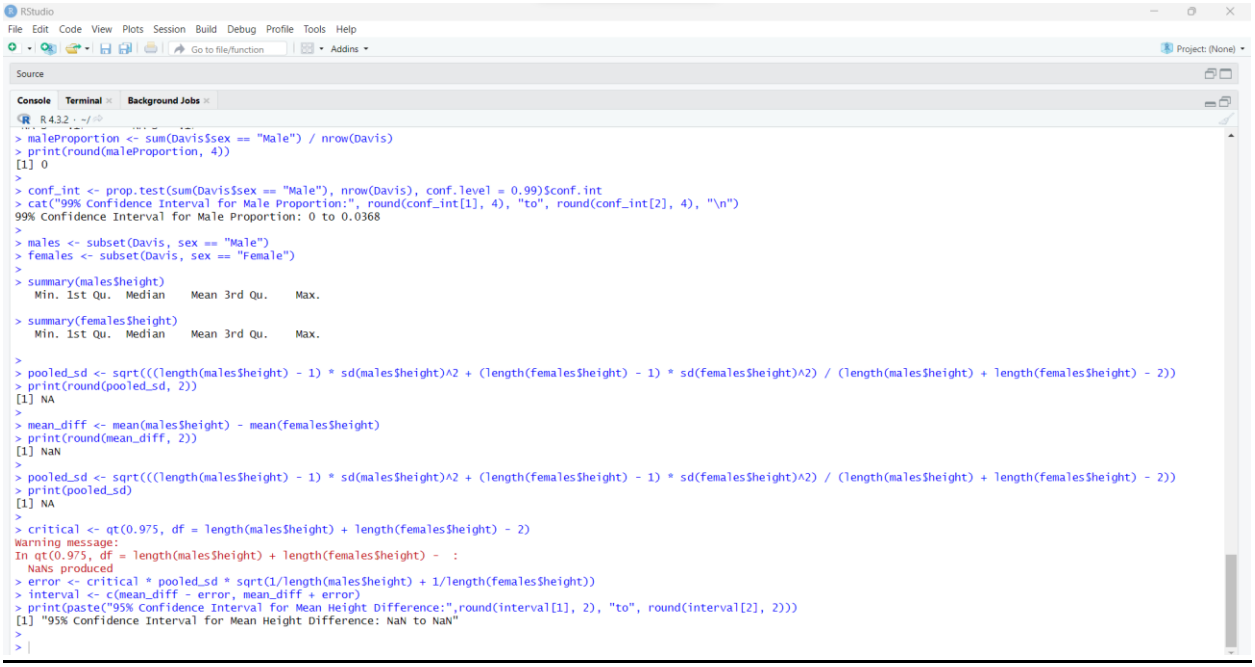
Question 2



```
> library(carData)
>
> data("Davis")
>
> summary(Davis)
 sex        weight          height
 F:112   Min.   : 39.0   Min.   : 57.0
 M: 88   1st Qu.: 55.0   1st Qu.:164.0
         Median : 63.0   Median :169.5
         Mean   : 65.8   Mean   :170.0
         3rd Qu.: 74.0   3rd Qu.:177.2
         Max.   :166.0   Max.   :197.0

     repwt            repht
 Min.   : 41.00   Min.   :148.0
 1st Qu.: 55.00   1st Qu.:160.5
 Median : 63.00   Median :168.0
 Mean   : 65.62   Mean   :168.5
 3rd Qu.: 73.50   3rd Qu.:175.0
 Max.   :124.00   Max.   :200.0
 NA's   :17       NA's   :17
> maleProportion <- sum(Davis$sex == "Male") / nrow(Davis)
> print(round(maleProportion, 4))
[1] 0
>
> conf_int <- prop.test(sum(Davis$sex == "Male"), nrow(Davis), conf.level = 0.99)$conf.int
> cat("99% Confidence Interval for Male Proportion:", round(conf_int[1], 4), "to", round(conf_int[2], 4), "\n")
99% Confidence Interval for Male Proportion: 0 to 0.0368
>
> males <- subset(Davis, sex == "Male")
> females <- subset(Davis, sex == "Female")
>
> summary(males$height)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.

> summary(females$height)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
```



```
> maleProportion <- sum(Davis$sex == "Male") / nrow(Davis)
> print(round(maleProportion, 4))
[1] 0
>
> conf_int <- prop.test(sum(Davis$sex == "Male"), nrow(Davis), conf.level = 0.99)$conf.int
> cat("99% Confidence Interval for Male Proportion:", round(conf_int[1], 4), "to", round(conf_int[2], 4), "\n")
99% Confidence Interval for Male Proportion: 0 to 0.0368
>
> males <- subset(Davis, sex == "Male")
> females <- subset(Davis, sex == "Female")
>
> summary(males$height)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.

> summary(females$height)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.

>
> pooled_sd <- sqrt(((length(males$height) - 1) * sd(males$height)^2 + (length(females$height) - 1) * sd(females$height)^2) / (length(males$height) + length(females$height) - 2))
> print(round(pooled_sd, 2))
[1] NA
>
> mean_diff <- mean(males$height) - mean(females$height)
> print(round(mean_diff, 2))
[1] NaN
>
> pooled_sd <- sqrt(((length(males$height) - 1) * sd(males$height)^2 + (length(females$height) - 1) * sd(females$height)^2) / (length(males$height) + length(females$height) - 2))
> print(pooled_sd)
[1] NA
>
> critical <- qt(0.975, df = length(males$height) + length(females$height) - 2)
Warning message:
In qt(0.975, df = length(males$height) + length(females$height) -  :
  NaNs produced
> error <- critical * pooled_sd * sqrt(1/length(males$height) + 1/length(females$height))
> interval <- c(mean_diff - error, mean_diff + error)
> print(paste("95% Confidence Interval for Mean Height Difference:",round(interval[1], 2), "to", round(interval[2], 2)))
[1] "95% Confidence Interval for Mean Height Difference: NaN to NaN"
>
>
```