DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
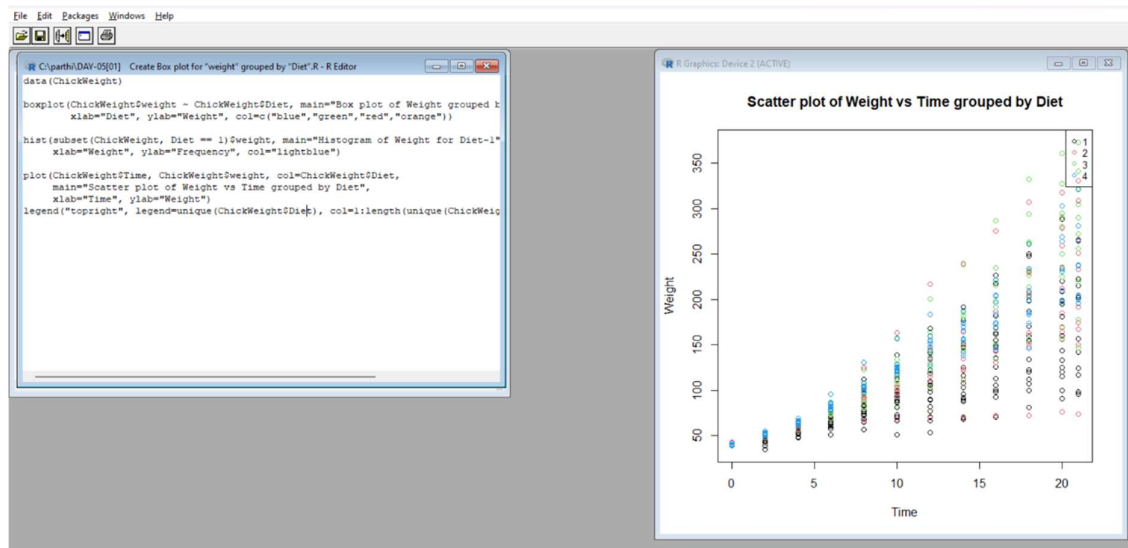## LIST OF EXPERIMENTS
ITA04- STATISTICS WITH R PROGRAMMING
DAY-05

1. a. Create Box plot for "weight" grouped by "Diet"
   b. Create a Histogram for "weight" features belong to Diet- 1 category
   c. Create Scatter plot for "weight" vs "Time" grouped by Diet.

2.a. Create multi regression model to find a weight of the chicken, by
      "Time" and "Diet" as predictor variables
   b. Predict weight for Time=10 and Diet=1
   c. Find the error in model for same.

```
R Console
> cat("Error in the model:", error, "\n")
Error in the model: -5.429309 4.570691 0.5706915 -11.42931 7.570691 25.57069 13⦶
> cat("Error in the model:", error, "\n")
Error in the model: -5.429309 4.570691 0.5706915 -11.42931 7.570691 25.57069 13⦶
> # Load ChickWeight dataset
> data(ChickWeight)
>
> # Convert 'Diet' variable to factor
> ChickWeight$Diet <- as.factor(ChickWeight$Diet)
>
> # a. Create multi regression model
> model <- lm(weight ~ Time + Diet, data = ChickWeight)
>
> # b. Predict weight for Time=10 and Diet=1
> new_data <- data.frame(Time = 10, Diet = factor(1, levels = levels(ChickWeight$
> predicted_weight <- predict(model, newdata = new_data)
> cat("Predicted weight for Time = 10 and Diet = 1:", predicted_weight, "\n")
Predicted weight for Time = 10 and Diet = 1: 98.42931
>
> # c. Find the error in model
> actual_weight <- subset(ChickWeight, Time == 10 & Diet == "1")$weight
> error <- actual_weight - predicted_weight
> cat("Error in the model:", error, "\n")
Error in the model: -5.429309 4.570691 0.5706915 -11.42931 7.570691 25.57069 13⦶
> |
```

```
C:\parthi\DAY-05[02]  Create multi regression model to find a weight of the chicken.R - R Editor
data(ChickWeight)

ChickWeight$Diet <- as.factor(ChickWeight$Diet)

model <- lm(weight ~ Time + Diet, data = ChickWeight)

new_data <- data.frame(Time = 10, Diet = factor(1, levels = levels(ChickWeight$Di
predicted_weight <- predict(model, newdata = new_data)
cat("Predicted weight for Time = 10 and Diet = 1:", predicted_weight, "\n")

actual_weight <- subset(ChickWeight, Time == 10 & Diet == "1")$weight
error <- actual_weight - predicted_weight
cat("Error in the model:", error, "\n")
```

3. For this exercise, use the (built-in) dataset Titanic.
   a. Draw a Bar chart to show details of "Survived" on the Titanic based on passenger Class
   b. Modify the above plot based on gender of people who survived
   c. Draw histogram plot to show distribution of feature "Age"

```
C:\parthi\DAY-05[03]  this exercise, use the (built-in) dataset Titanic.R - R Editor
data(Titanic)

# a. Bar chart for Survived based on passenger Class
barplot(rowSums(Titanic[,,,"Yes"]), beside=TRUE, col=rainbow(2), legend=TRUE,
        main="Survival Count by Passenger Class", xlab="Passenger Class", ylab="C
legend("topright", legend=c("Survived", "Did not survive"), fill=rainbow(2))

# b. Bar chart for Survived based on passenger Class and Gender
barplot(as.matrix(Titanic[,,c("Male","Female"),"Yes"]), beside=TRUE, col=c("light
        main="Survival Count by Passenger Class and Gender", xlab="Passenger Clas
legend("topright", legend=c("Male", "Female"), fill=c("lightblue", "pink"))

# c. Histogram plot for Age distribution
hist(Titanic$Age, breaks=20, col="skyblue", main="Age Distribution", xlab="Age",
```


Survival Count by Passenger Class

ANALYTICAL QUESTIONS

1.a. Create a data frame based on below table.

Month	1	2	3	4	5	6	7	8	9	10	11	12
Spends	1000	4000	5000	4500	3000	4000	9000	11000	15000	12000	7000	3000
Sales	9914	40487	54324	50044	34719	42551	94871	118914	158484	131348	78504	36284
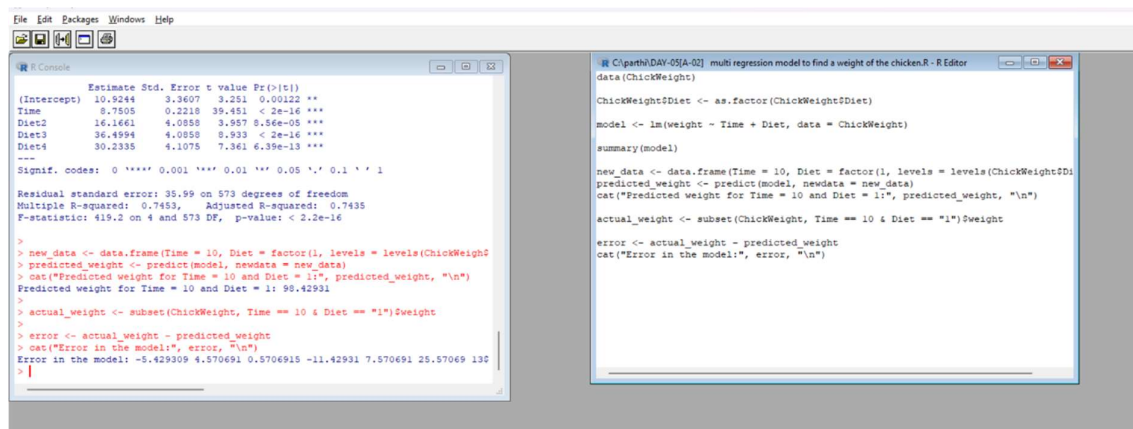
b. Create a regression model for that data frame table to show the amount of sales (Sales) based on the how much the company spends (Spends) in advertising
c. Predict the Sales if Spend=1350

2.a. Create multi regression model to find a weight of the chicken, by "Time" and "Diet" as predictor variables
 b. Predict weight for Time=10 and Diet=1
c. Find the error in model for same

.Explore the USArrests dataset, contains the number of arrests for murder, assault, and rape for each of the 50 states in 1973. It also contains the percentage of people in the state who live in an urban area.
(i) a. Explore the summary of Data set, like number of Features and its type. Find the number of records for each feature. Print the statistical feature of data

b. Print the state which saw the largest total number of rape
c. Print the states with the max & min crime rates for murder
(ii).a. Find the correlation among the features