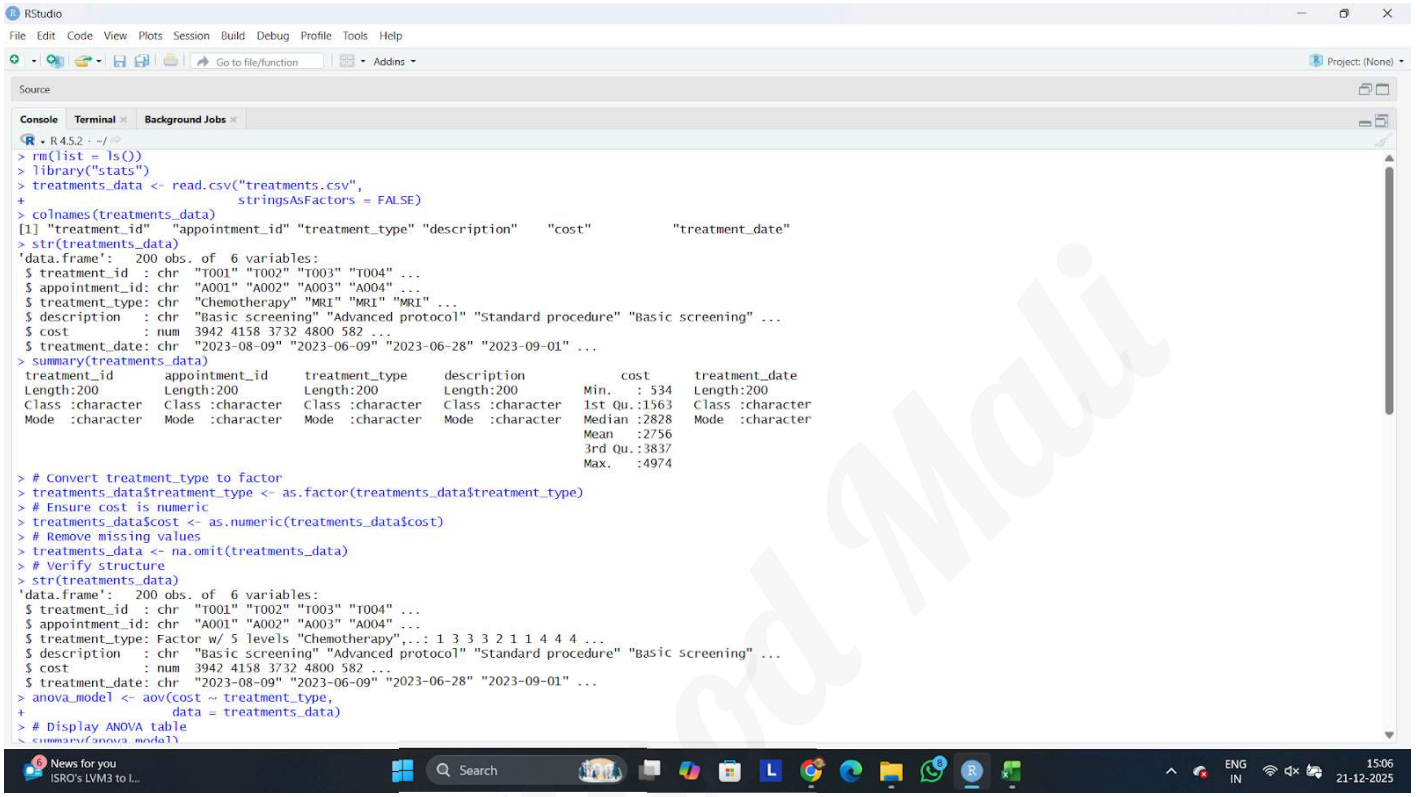
Sheth L.U.J & Sir M.V College Of Science

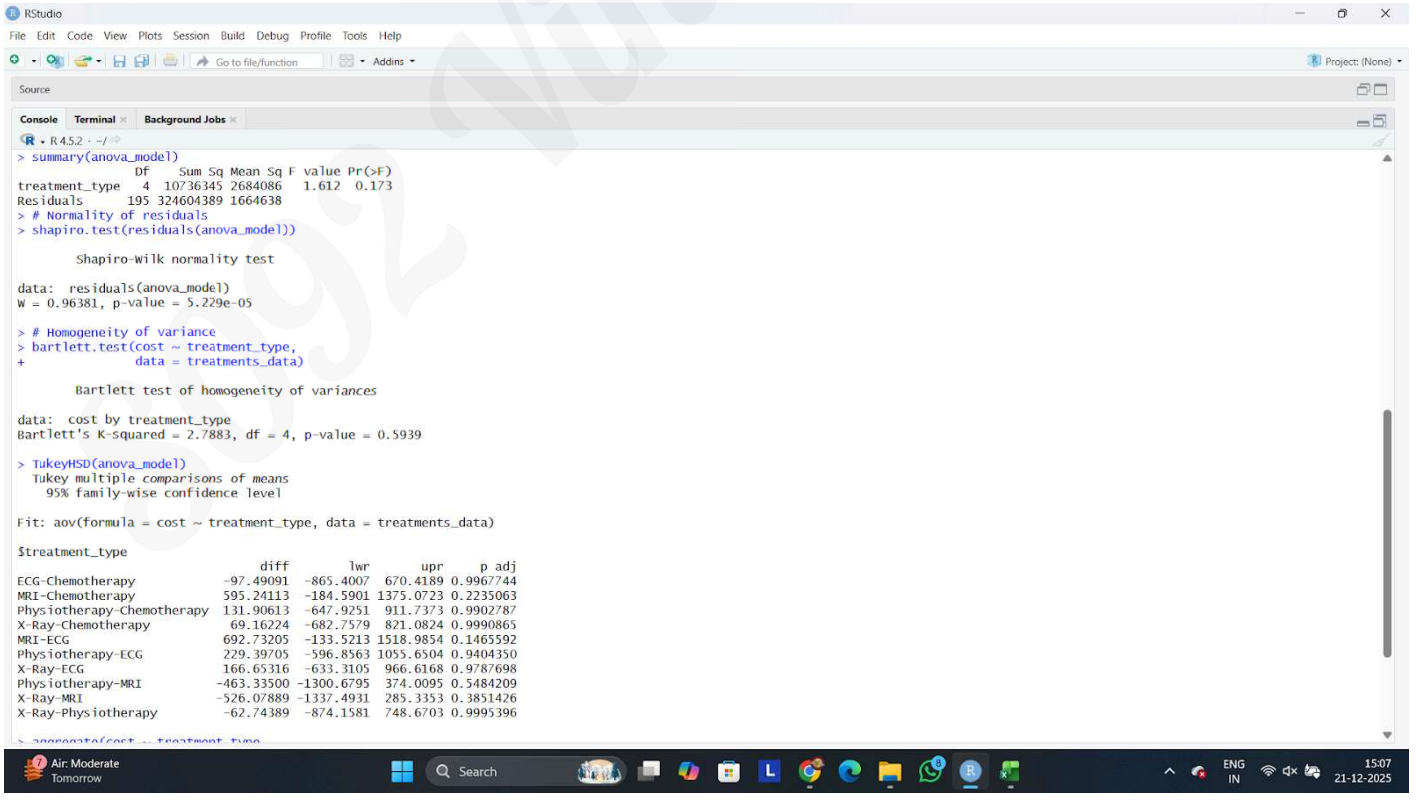
Subject :- Data Analysis With SASS/SPSS/R

Module 2 Practical no 8

Aim :- Performing two-way ANOVA using aov() (R).



```
> rm(list = ls())
> library("stats")
> treatments_data <- read.csv("treatments.csv",
+                             stringsAsFactors = FALSE)
> colnames(treatments_data)
[1] "treatment_id"   "appointment_id" "treatment_type" "description"    "cost"           "treatment_date"
> str(treatments_data)
'data.frame':   200 obs. of  6 variables:
 $ treatment_id  : chr  "T001" "T002" "T003" "T004" ...
 $ appointment_id: chr  "A001" "A002" "A003" "A004" ...
 $ treatment_type: chr  "Chemotherapy" "MRI" "MRI" "MRI" ...
 $ description   : chr  "Basic screening" "Advanced protocol" "Standard procedure" "Basic screening" ...
 $ cost          : num  3942 4158 3732 4800 582 ...
 $ treatment_date: chr  "2023-08-09" "2023-06-09" "2023-06-28" "2023-09-01" ...
> summary(treatments_data)
 treatment_id       appointment_id     treatment_type     description            cost      treatment_date
 Length:200         Length:200         Length:200         Length:200         Min.   : 534   Length:200
 Class :character   Class :character   Class :character   Class :character   1st Qu.:1563   Class :character
 Mode  :character   Mode  :character   Mode  :character   Mode  :character   Median :2828   Mode  :character
                                                                             Mean   :2756
                                                                             3rd Qu.:3837
                                                                             Max.   :4974
> # Convert treatment_type to factor
> treatments_data$treatment_type <- as.factor(treatments_data$treatment_type)
> # Ensure cost is numeric
> treatments_data$cost <- as.numeric(treatments_data$cost)
> # Remove missing values
> treatments_data <- na.omit(treatments_data)
> # Verify structure
> str(treatments_data)
'data.frame':   200 obs. of  6 variables:
 $ treatment_id  : chr  "T001" "T002" "T003" "T004" ...
 $ appointment_id: chr  "A001" "A002" "A003" "A004" ...
 $ treatment_type: Factor w/ 5 levels "Chemotherapy",..: 1 3 3 3 2 1 1 4 4 4 ...
 $ description   : chr  "Basic screening" "Advanced protocol" "Standard procedure" "Basic screening" ...
 $ cost          : num  3942 4158 3732 4800 582 ...
 $ treatment_date: chr  "2023-08-09" "2023-06-09" "2023-06-28" "2023-09-01" ...
> anova_model <- aov(cost ~ treatment_type,
+                    data = treatments_data)
> # Display ANOVA table
> summary(anova_model)
```



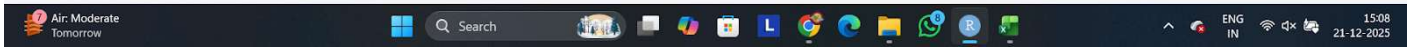
```
> summary(anova_model)
               Df    Sum Sq Mean Sq F value Pr(>F)
treatment_type   4  10736345 2684086   1.612  0.173
Residuals      195 324604389 1664638
> # Normality of residuals
> shapiro.test(residuals(anova_model))

        Shapiro-Wilk normality test

data:  residuals(anova_model)
W = 0.96381, p-value = 5.229e-05

> # Homogeneity of variance
> bartlett.test(cost ~ treatment_type,
+               data = treatments_data)

        Bartlett test of homogeneity of variances

data:  cost by treatment_type
Bartlett's K-squared = 2.7883, df = 4, p-value = 0.5939

> TukeyHSD(anova_model)
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = cost ~ treatment_type, data = treatments_data)

$treatment_type
                                diff        lwr       upr     p adj
ECG-Chemotherapy           -97.49091  -865.4007  670.4189 0.9967744
MRI-Chemotherapy           595.24113 -184.5901 1375.0723 0.2235063
Physiotherapy-Chemotherapy 131.90613 -647.9251  911.7373 0.9902787
X-Ray-Chemotherapy          69.16224 -682.7579  821.0824 0.9990865
MRI-ECG                    692.73205 -133.5213 1518.9854 0.1465592
Physiotherapy-ECG          229.39705 -596.8563 1055.6504 0.9404350
X-Ray-ECG                  166.65316 -633.3105  966.6168 0.9787698
Physiotherapy-MRI         -463.33500 -1300.6795  374.0095 0.5484209
X-Ray-MRI                 -526.07889 -1337.4931  285.3353 0.3851426
X-Ray-Physiotherapy        -62.74389  -874.1581  748.6703 0.9995396

> aggregate(cost ~ treatment_type
```

Name :- Vinod Mali

Roll No :- S092

Sheth L.U.J & Sir M.V College Of Science
Subject :- Data Analysis With SASS/SPSS/R
Module 2 Practical no 8

```
> aggregate(cost ~ treatment_type,
+           data = treatments_data,
+           mean)
  treatment_type     cost
1   Chemotherapy 2629.708
2            ECG 2532.217
3            MRI 3224.949
4  Physiotherapy 2761.614
5          X-Ray 2698.870
>
>
```



Name :- Vinod Mali
Roll No :- S092