

MVLU COLLEGE

R PROGRAMMING

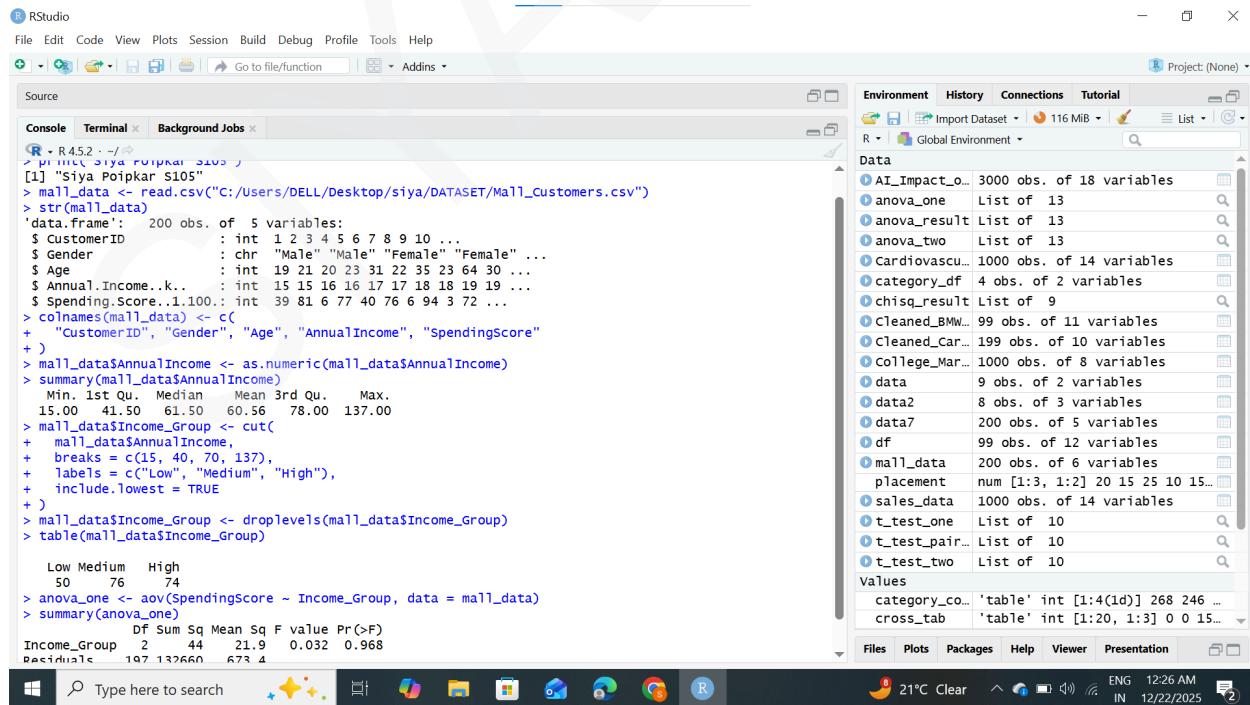
PRACTICAL NO. 7

AIM: 7 Performing one-way ANOVA using aov() (R).

#CODE:

```
# Practical 7: One-Way ANOVA
print("Siya Poipkar S105")
mall_data <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/Mall_Customers.csv")
str(mall_data)
colnames(mall_data) <- c(
  "CustomerID", "Gender", "Age", "AnnualIncome", "SpendingScore"
)
mall_data$AnnualIncome <- as.numeric(mall_data$AnnualIncome)
summary(mall_data$AnnualIncome)
mall_data$Income_Group <- cut(
  mall_data$AnnualIncome,
  breaks = c(15, 40, 70, 137),
  labels = c("Low", "Medium", "High"),
  include.lowest = TRUE
)
mall_data$Income_Group <- droplevels(mall_data$Income_Group)
table(mall_data$Income_Group)
anova_one <- aov(SpendingScore ~ Income_Group, data = mall_data)
summary(anova_one)
```

#OUTPUT:



The screenshot shows the RStudio interface with the following details:

- File menu:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Console tab:** Displays the R session history with the code and its output.
- Environment tab:** Shows the global environment with various objects listed, such as AI_Impact, anova_one, anova_result, Cardiovascu, category_df, chisq_result, cleaned_BMW, cleaned_Car, College_Mar..., data, data2, data7, df, mall_data, placement, sales_data, t_test_one, t_test_pair..., and t_test_two.
- Plots tab:** Not visible in the screenshot.
- Packages tab:** Not visible in the screenshot.
- Help tab:** Not visible in the screenshot.
- Viewer tab:** Not visible in the screenshot.
- Presentation tab:** Not visible in the screenshot.
- System tray:** Shows the Windows taskbar with icons for File Explorer, Task View, Start, and others. The system clock shows 12:26 AM on 12/22/2025.

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 7

AIM: 8 Performing two-way ANOVA using `aov()` (R).

#CODE:

```
# Practical 8: Two-Way ANOVA
print("Siya Poipkar S105")
ai_data <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/AI_Impact_on_Jobs_2030.csv")

str(ai_data)

colnames(ai_data) <- gsub(" ", "_", colnames(ai_data))

ai_data$Average_Salary <- as.numeric(ai_data$Average_Salary)

ai_data$Education_Level <- as.factor(ai_data$Education_Level)
ai_data$Risk_Category <- as.factor(ai_data$Risk_Category)

ai_data$Education_Level <- droplevels(ai_data$Education_Level)
ai_data$Risk_Category <- droplevels(ai_data$Risk_Category)

table(ai_data$Education_Level)
table(ai_data$Risk_Category)

anova_two <- aov(Average_Salary ~ Education_Level * Risk_Category, data = ai_data)

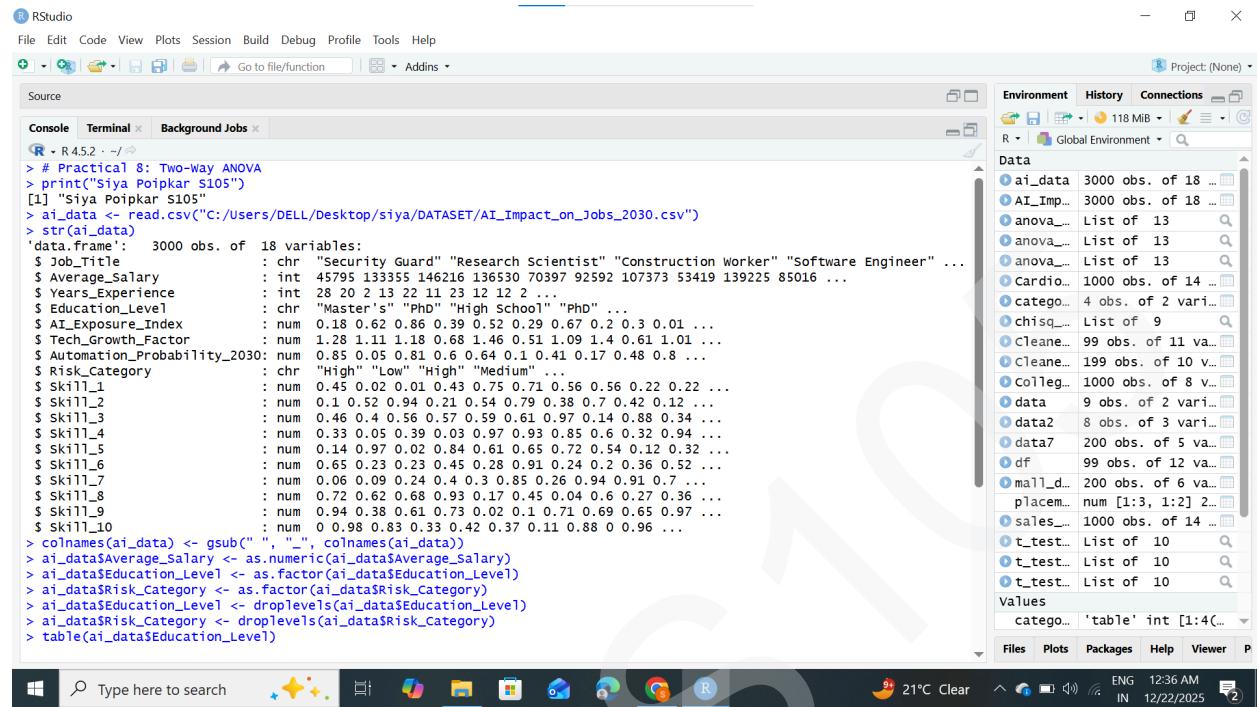
summary(anova_two)
```

MVLU COLLEGE

R PROGRAMMING

PRACTICAL NO. 7

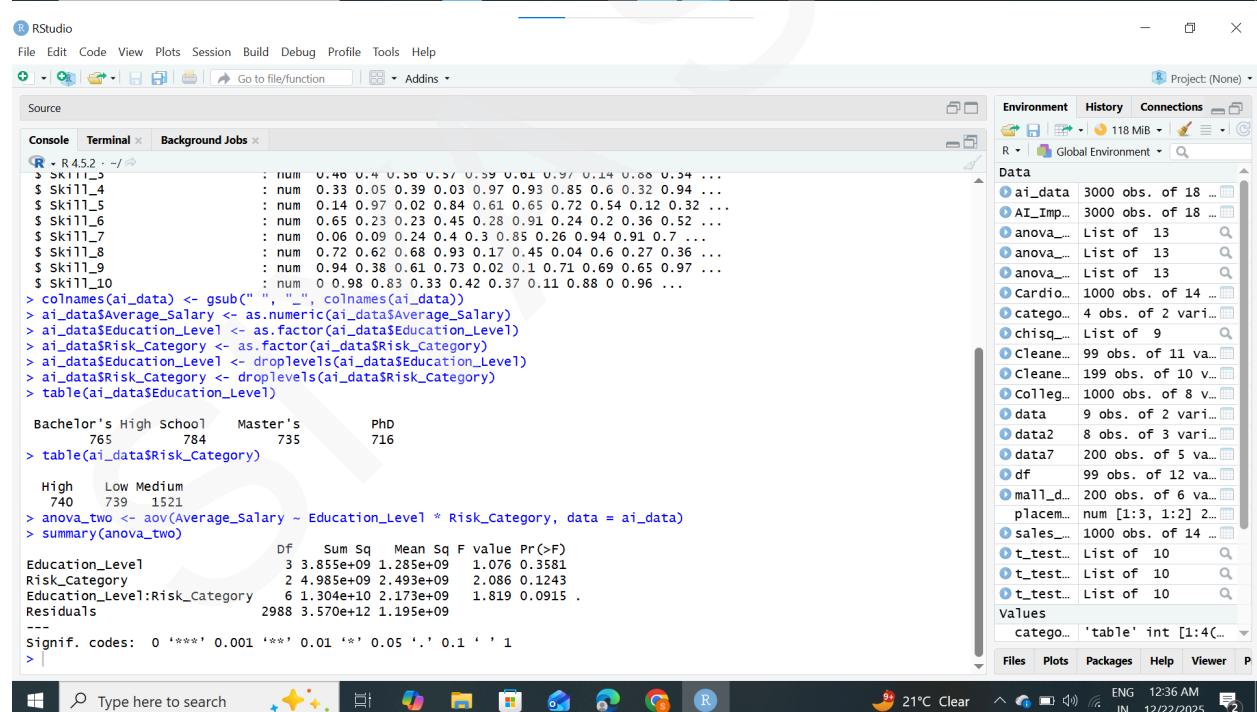
#OUTPUT:



```

> # Practical 8: Two-Way ANOVA
> print("Siya Poipkar S105")
[1] "Siya Poipkar S105"
> ai_data <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/AI_Impact_on_Jobs_2030.csv")
> str(ai_data)
'data.frame': 3000 obs. of 18 variables:
 $ Job_Title      : chr "Security Guard" "Research Scientist" "Construction Worker" "Software Engineer" ...
 $ Average_Salary : int 45795 133355 146216 136530 70397 92592 107373 53419 139225 85016 ...
 $ Years_Experience: int 28 20 2 13 22 11 23 12 12 2 ...
 $ Education_Level : chr "Master's" "PhD" "High School" "PhD" ...
 $ AI_Exposure_Index: num 0.18 0.62 0.86 0.39 0.52 0.29 0.67 0.2 0.3 0.01 ...
 $ Tech_Growth_Factor: num 1.28 1.11 1.18 0.68 1.46 0.51 1.09 1.4 0.61 1.01 ...
 $ Automation_Probability_2030: num 0.85 0.05 0.81 0.6 0.64 0.1 0.42 0.17 0.48 0.8 ...
 $ Risk_Category   : chr "High" "Low" "High" "Medium" ...
 $ skill_1         : num 0.45 0.02 0.01 0.43 0.75 0.56 0.56 0.22 0.22 ...
 $ skill_2         : num 0.1 0.52 0.94 0.21 0.54 0.79 0.38 0.7 0.42 0.12 ...
 $ skill_3         : num 0.46 0.4 0.56 0.57 0.59 0.61 0.97 0.14 0.88 0.34 ...
 $ skill_4         : num 0.33 0.05 0.39 0.03 0.97 0.93 0.85 0.6 0.32 0.94 ...
 $ skill_5         : num 0.14 0.97 0.02 0.84 0.61 0.65 0.72 0.54 0.12 0.32 ...
 $ skill_6         : num 0.65 0.23 0.23 0.45 0.28 0.91 0.24 0.2 0.36 0.52 ...
 $ skill_7         : num 0.06 0.09 0.24 0.4 0.3 0.85 0.26 0.94 0.91 0.7 ...
 $ skill_8         : num 0.72 0.62 0.68 0.93 0.17 0.45 0.04 0.6 0.27 0.36 ...
 $ skill_9         : num 0.94 0.38 0.62 0.73 0.02 0.1 0.71 0.69 0.65 0.97 ...
 $ skill_10        : num 0 0.98 0.83 0.33 0.42 0.37 0.11 0.88 0 0.96 ...
> colnames(ai_data) <- gsub(" ", "", colnames(ai_data))
> ai_data$Average_Salary <- as.numeric(ai_data$Average_Salary)
> ai_data$Education_Level <- as.factor(ai_data$Education_Level)
> ai_data$Risk_Category <- as.factor(ai_data$Risk_Category)
> ai_data$Education_Level <- droplevels(ai_data$Education_Level)
> ai_data$Risk_Category <- droplevels(ai_data$Risk_Category)
> table(ai_data$Education_Level)

```

```

Bachelor's High School    Master's      PhD
          765           784           735       716
> table(ai_data$Risk_Category)

High     Low Medium
740     739    1521
> anova_two <- aov(Average_Salary ~ Education_Level * Risk_Category, data = ai_data)
> summary(anova_two)
             Df  Sum Sq Mean Sq F value Pr(>F)
Education_Level 3 3.855e+09 1.285e+09  1.076 0.3581
Risk_Category   2 4.985e+09 2.493e+09  2.086 0.1243
Education_Level:Risk_Category 6 1.304e+10 2.173e+09  1.819 0.0915 .
Residuals      2988 3.570e+12 1.195e+09
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
>

```

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 7

AIM: 9 Conducting Chi-square tests using chisq.test() (R)

```
#CODE:  
# Practical 9: Chi-Square Test  
print("Siya Poipkar S105")  
heart_data <-  
read.csv("C:/Users/DELL/Desktop/siya/DATASET/Cardiovascular_Disease_Dataset.csv")  
  
str(heart_data)  
  
colnames(heart_data) <- gsub(" ", "_", colnames(heart_data))  
  
heart_data$gender <- as.factor(heart_data$gender)  
heart_data$target <- as.factor(heart_data$target)  
  
heart_data$gender <- droplevels(heart_data$gender)  
heart_data$target <- droplevels(heart_data$target)  
  
cont_table <- table(heart_data$gender, heart_data$target)  
  
cont_table  
  
chi_result <- chisq.test(cont_table)  
  
chi_result
```

MVLU COLLEGE

R PROGRAMMING

PRACTICAL NO. 7

#OUTPUT:

RStudio interface showing the code for Practical 9: Chi-Square Test. The code reads a CSV file 'Cardiovascular_Disease_Dataset.csv' and performs a Chi-square test on the gender variable.

```

> # Practical 9: Chi-Square Test
> print("Siya Poipkar S105")
[1] "Siya Poipkar S105"
> heart_data <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/Cardiovascular_Disease_Dataset.csv")
> str(heart_data)
'data.frame': 1000 obs. of 14 variables:
 $ patientid : int 103368 119250 119372 132514 146211 148462 168686 170498 188225 192523 ...
 $ age        : int 53 40 49 43 31 24 79 52 62 61 ...
 $ gender     : int 1 1 1 1 1 1 1 1 0 ...
 $ chestpain   : int 2 0 2 0 1 1 2 0 0 0 ...
 $ restingBP    : int 171 94 133 138 199 173 130 127 121 190 ...
 $ serumcholesterol : int 0 229 142 295 0 0 240 345 357 181 ...
 $ fastingbloodsugar: int 0 0 0 1 0 0 0 0 0 0 ...
 $ restingelectro : int 1 1 0 1 2 0 2 0 1 1 ...
 $ maxheartrate  : int 147 115 202 153 136 161 157 192 138 150 ...
 $ exerciseangia : int 0 1 0 0 0 0 0 1 0 0 ...
 $ oldpeak      : num 5.3 3.7 5.3 2.5 5.3 4.7 2.5 4.9 2.8 2.9 ...
 $ slope        : int 3 1 1 2 3 3 2 1 0 2 ...
 $ noofmajorvessels: int 3 1 0 2 2 2 1 0 0 0 ...
 $ target       : int 1 0 0 1 1 1 1 0 0 1 ...
> colnames(heart_data) <- gsub(" ", "_", colnames(heart_data))
> heart_data$gender <- as.factor(heart_data$gender)
> heart_data$target <- as.factor(heart_data$target)
> heart_data$gender <- droplevels(heart_data$gender)
> heart_data$target <- droplevels(heart_data$target)
> cont_table <- table(heart_data$gender, heart_data$target)
> cont_table

0 1
0 102 133
1 318 447
> chi_result <- chisq.test(cont_table)

Chi-squared test with Yates' continuity correction
```

RStudio interface showing the continuation of the Chi-Square Test code. The code prints the results of the Chi-square test, including the X-squared value, degrees of freedom, and p-value.

```

> chi_result <- chisq.test(cont_table)
> chi_result

Pearson's Chi-squared test with Yates' continuity correction

data: cont_table
X-squared = 0.17902, df = 1, p-value = 0.6722
>
```