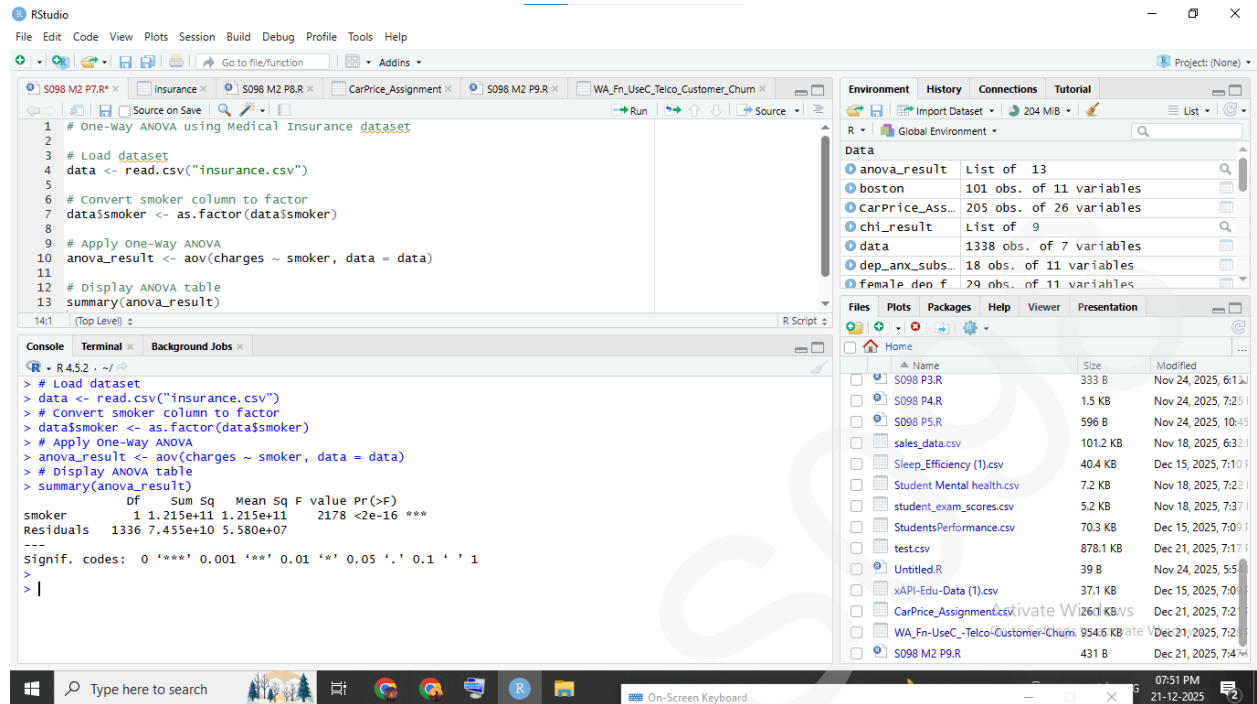


MVLU COLLEGE

Subject:-Data Analysis with SAS / SPSS /R

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



The screenshot displays the RStudio interface with a script editor, console, and environment pane. The script editor contains the following R code:

```
1 # One-way ANOVA using Medical Insurance dataset
2
3 # Load dataset
4 data <- read.csv("insurance.csv")
5
6 # Convert smoker column to factor
7 data$smoker <- as.factor(data$smoker)
8
9 # Apply One-Way ANOVA
10 anova_result <- aov(charges ~ smoker, data = data)
11
12 # Display ANOVA table
13 summary(anova_result)
```

The console shows the output of the code execution:

```
> # Load dataset
> data <- read.csv("insurance.csv")
> # Convert smoker column to factor
> data$smoker <- as.factor(data$smoker)
> # Apply One-way ANOVA
> anova_result <- aov(charges ~ smoker, data = data)
> # Display ANOVA table
> summary(anova_result)
```

The ANOVA table output is as follows:

	df	Sum Sq	Mean Sq	F value	Pr(>F)
smoker	1	1.215e+11	1.215e+11	2178	<2e-16 ***
Residuals	1336	7.455e+10	5.580e+07		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The environment pane on the right shows the following objects:

- anova_result: List of 13
- boston: 101 obs. of 11 variables
- carPrice_Ass...: 205 obs. of 26 variables
- chi_result: List of 9
- data: 1338 obs. of 7 variables
- dep_anx_subs...: 18 obs. of 11 variables
- female den f...: 79 obs. of 11 variables

MVLU COLLEGE

Subject:-Data Analysis with SAS / SPSS /R

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

The screenshot displays the RStudio interface with the following components:

- Script Editor:** Contains R code for loading the 'CarPrice_Assignment.csv' dataset, converting categorical variables to factors, and performing a two-way ANOVA with interaction using the `aov()` function.
- Console:** Shows the execution of the R code, resulting in the following ANOVA table:

	DF	Sum Sq	Mean Sq	F value	Pr(>F)
fueltype	1	1.454e+08	145405324	2.595	0.109
carbody	4	1.716e+09	428973871	7.656	9.44e-06 ***
fueltype:carbody	3	1.759e+08	58645132	1.047	0.373
Residuals	196	1.098e+10	56032669		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
- Environment:** Lists the objects created in the global environment, including `anova_result`, `boston`, `CarPrice_Ass...`, `chi_result`, `data`, `dep_anx_subs...`, and `female den f`.
- Files:** Shows a list of files in the current directory, including `S098 P3.R`, `S098 P4.R`, `S098 P5.R`, `sales_data.csv`, `Sleep_Efficiency (1).csv`, `Student Mental health.csv`, `student_exam_scores.csv`, `StudentsPerformance.csv`, `test.csv`, `Untitled.R`, `xAPI-Edu-Data (1).csv`, `CarPrice_Assignment.csv`, `WA_Fn-UseC_-Telco-Customer-Churn`, and `S098 M2 P9.R`.

Subject:-Data Analysis with SAS / SPSS /R

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

The screenshot displays the RStudio interface with the following components:

- Source Editor:** Contains R code for loading a dataset and inspecting its structure.


```
1 # Practical 9: chi-Square Test of Independence
2
3 # Load the dataset
4 data <- read.csv("WA_Fn-UseC_-Telco-Customer-Churn.csv")
5
6 # View first few rows
7 head(data)
8
9 # Create contingency table between Gender and Churn
450
```
- Console:** Shows the execution of the code, displaying the first 6 rows of the dataset and the first 5 rows of the contingency table.


```
R > R 4.5.2 - ~/
> # Load the dataset
> data <- read.csv("WA_Fn-UseC_-Telco-Customer-Churn.csv")
> # View first few rows
> head(data)
customerID gender SeniorCitizen Partner Dependents tenure Phoneservice Multiplelines
1 7590-VHVEG Female 0 Yes No 1 No No phone service
2 5575-GNVEI Male 0 No 34 Yes No No
3 3668-QPYBK Male 0 No No 2 Yes No No
4 7795-CFOCW Male 0 No 45 No No No phone service
5 9237-WGTU Female 0 No 2 Yes No No
6 9305-CDSKC Female 0 No No 8 Yes Yes
InternetService OnlineSecurity OnlineBackup DeviceProtection TechSupport StreamingTV StreamingMovies
1 DSL DSL Yes No Yes No No No
2 DSL Yes No Yes No No No No
3 DSL Yes Yes No No No No No
4 DSL Yes No Yes Yes No No No
5 Fiber optic No No No No No No No
6 Fiber optic No No Yes No Yes No Yes
Contract PaperlessBilling PaymentMethod MonthlyCharges TotalCharges Churn
1 Month-to-month Yes Electronic check 29.85 29.85 No
2 One year No Mailed check 56.95 1889.50 No
3 Month-to-month Yes Mailed check 53.85 108.15 Yes
4 One year No Bank transfer (Automatic) 42.30 1840.75 No
5 Month-to-month Yes Electronic check 70.70 151.65 Yes
```
- Environment Pane:** Lists the objects in the Global Environment.

Object	Type	Size
chi_result	List of 9	
data	7043 obs. of 21 variables	
dep_anx_subs...	18 obs. of 11 variables	
female_dep_f...	29 obs. of 11 variables	
high_age_sub...	36 obs. of 11 variables	
high_cgpa_fi...	91 obs. of 11 variables	
insurance	1338 obs. of 7 variables	
lanna_sales	303 obs. of 14 variables	

```
# Practical 9: Chi-Square Test of Independence
# Load the dataset
data <- read.csv("WA_Fn-UseC_-Telco-Customer-Churn.csv")
# View first few rows
head(data)
# Create contingency table between Gender and Churn
# Display the table
print(contingency_table)
# Apply Chi-Square Test
chi_result <- chisq.test(contingency_table)
# Display test result
chi_result
```

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

```
data: contingency_table
X-squared = 0.48408, df = 1, p-value = 0.4866
```

Environment: Global Environment

Object	Class	Attributes
chi_result	list	of 9
data	data.frame	7043 obs. of 21 variables
dep_anx_subs...	data.frame	18 obs. of 11 variables
female_dep_f...	data.frame	29 obs. of 11 variables
high_age_sub...	data.frame	36 obs. of 11 variables
high_cgpa_fi...	data.frame	91 obs. of 11 variables
insurance	data.frame	1338 obs. of 7 variables
large_sales	data.frame	393 obs. of 14 variables