

SHETH L.U.J. AND SIR M.V. COLLEGE

SUBJECT:Data Analysis with R

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

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

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

7TH OUTPUT:

```

R S103 - RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
+ Untitled1* anova_result students
Source on Save Run Source
1 library(dplyr)
2 students <- read.csv("StudentsPerformance.csv")
3 # Rename column for easy usage
4 colnames(students)[7] <- "reading_score"
5
6 # Convert test preparation course to factor
7 students$test.preparation.course <- as.factor(students$test.preparation.course)
12:1 (Top Level) R Script

Console Terminal Background Jobs
> library(dplyr)
> students <- read.csv("StudentsPerformance.csv")
> # Rename column for easy usage
> colnames(students)[7] <- "reading_score"
> # Convert test preparation course to factor
> students$test.preparation.course <- as.factor(students$test.preparation.course)
> anova_result <- aov(reading_score ~ test.preparation.course, data = students)
> # ANOVA table
> summary(anova_result)
   Df Sum Sq Mean Sq F value Pr(>F)
test.preparation.course  1 12449  12449  61.96 9.08e-15 ***
Residuals                998 200504     201
---
signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
> View(anova_result)
> View(students)
>

```

The screenshot shows the RStudio interface with two main panes. The left pane displays the 'students' data frame in a grid format. The right pane shows the R environment with objects 'anova_result' and 'students' listed.

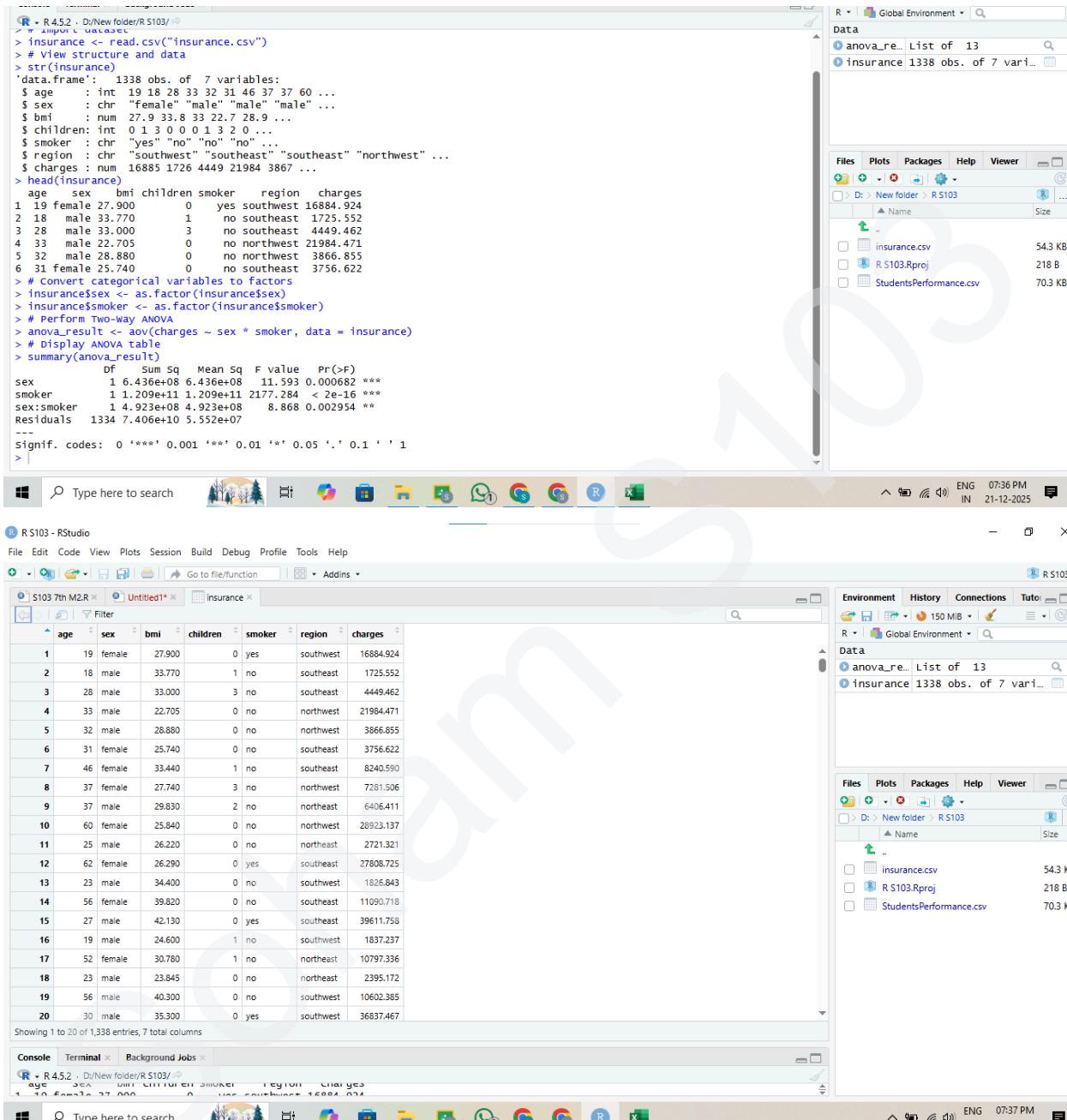
	gender	race.ethnicity	parental.level.of.education	lunch	test.preparation.course	math.score	reading_score	writing.score
1	female	group B	bachelor's degree	standard	none	72	72	74
2	female	group C	some college	standard	completed	69	90	88
3	female	group B	master's degree	standard	none	90	95	93
4	male	group A	associate's degree	free/reduced	none	47	57	44
5	male	group C	some college	standard	none	76	78	75
6	female	group B	associate's degree	standard	none	71	83	78
7	female	group B	some college	standard	completed	88	95	92
8	male	group B	some college	free/reduced	none	40	43	39
9	male	group D	high school	free/reduced	completed	64	64	67
10	female	group B	high school	free/reduced	none	38	60	50
11	male	group C	associate's degree	standard	none	58	54	52
12	male	group D	associate's degree	standard	none	40	52	43
13	female	group B	high school	standard	none	65	81	73
14	male	group A	some college	standard	completed	78	72	70
15	female	group A	master's degree	standard	none	50	53	58
16	female	group C	some high school	standard	none	69	75	78
17	male	group C	high school	standard	none	88	89	86
18	female	group B	some high school	free/reduced	none	18	32	28

Showing 1 to 19 of 1,000 entries, 8 total columns

SHETH L.U.J. AND SIR M.V. COLLEGE

SUBJECT:Data Analysis with R

8th output:



The screenshot shows the RStudio interface with the following details:

- Console:**

```
R > R.4.5.2 - D:\New folder\R S103\>
> insurance <- read.csv("insurance.csv")
> # View structure and data
> str(insurance)
'data.frame': 1338 obs. of 7 variables:
 $ age      : int 19 18 28 33 32 31 46 37 37 60 ...
 $ sex       : chr "female" "male" "male" "male" ...
 $ bmi       : num 27.9 33.8 33 22.7 28.9 ...
 $ children: int 0 1 3 0 0 0 1 3 2 0 ...
 $ smoker    : chr "yes" "no" "no" "no" ...
 $ region    : chr "southwest" "southeast" "southeast" "northwest" ...
 $ charges   : num 16885 1726 4449 21984 3867 ...
> head(insurance)
#> # Convert categorical variables to factors
#> insurance$sex <- as.factor(insurance$sex)
#> insurance$smoker <- as.factor(insurance$smoker)
#> # Perform Two-Way ANOVA
#> anova_result <- aov(charges ~ sex * smoker, data = insurance)
#> # Display ANOVA table
#> summary(anova_result)
#>
```
- Data View:** Shows the first 6 rows of the insurance dataset.
- File Explorer:** Shows files in the R S103 folder, including insurance.csv, R S103.Rproj, and StudentsPerformance.csv.
- Global Environment:** Shows the objects available in the environment, including anova_result and insurance.
- Plots:** No plots are visible.
- Packages:** No packages are visible.
- Help:** No help is visible.
- Viewer:** No viewer content is visible.
- Bottom Status Bar:** Shows the date (21-12-2025), time (07:36 PM), and system status (ENG IN).

SHETH L.U.J. AND SIR M.V. COLLEGE

SUBJECT:Data Analysis with R

9th output:

The screenshot shows the RStudio interface with the following details:

- Header:** R S103 - RStudio
- Menu Bar:** File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help
- Toolbar:** Source, Console, Terminal, Background Jobs, Go to file/function, Addins
- Console:** Displays R code and its output. The code includes loading the dplyr library, importing the titanic dataset from a CSV file, viewing the structure and data, and printing the first few rows of the titanic dataset.
- Data View:** Shows the titanic dataset as a table with columns: PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin, Embarked.
- Environment:** Shows the global environment with objects like chi_resu, titanic, and values.
- Files:** Shows files in the current directory: insurance.csv, R S103.Rproj, S103 8TH M2.R, StudentsPerformance.csv, tested.csv, and S103 9TH M2.R.

The screenshot shows the RStudio interface with the following components:

- Header:** R S103 - RStudio, File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help.
- Toolbar:** Go to file/function, Addins.
- Source Editor:** Shows the R code used for the analysis, including the loading of the titanic dataset, conversion of sex and survived to factors, creation of a contingency table, and performance of a Chi-square test.
- Console:** Displays the output of the R code, including the contingency table, the Chi-square test results (X-squared = 413.69, df = 1, p-value < 2.2e-16), and summaries of the titanic and chi_result objects.
- Environment:** Shows the global environment with objects like chi_resum, titanic, and table_da.
- Data View:** Shows the titanic dataset with columns: Pclass, Name, Sex, Age, SibSp, Parch, Fare, and Survived.
- Files:** Shows the project structure with files like insurance.csv, R.S103.Rproj, S103 8TH M2.R, StudentsPerformance.csv, tested.csv, and S103 9TH M2.R.
- Plots:** Shows a small preview of a plot.
- Packages:** Shows available packages.
- Help:** Help documentation.
- Viewer:** Shows the titanic dataset as a table.

SHETH L.U.J. AND SIR M.V. COLLEGE

SUBJECT:Data Analysis with R

The screenshot shows the RStudio interface with the following details:

- Environment Pane:** Shows objects like "chi_resu..." (List of 9), "titanic" (418 obs. of 12 vari...), and "Values".
- Global Environment:** Shows the "titanic" object as a "table" with dimensions 1:2, 1:2..
- Data View:** A grid view of the "titanic" dataset with columns: PassengerId, Survived, Pclass, Name, Sex, Age, SibSp, Parch, Ticket, Fare, Cabin.
- Console:** Placeholder "Type here to search" and a row of system icons.
- Bottom Status Bar:** Shows "ENG 07:45 PM IN 21-12-2025".