

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 10 TO 12 (MODULE 2)

AIM : 10: Creating graphical reports using ,ggplot2 (R).

```
library(ggplot2)
library(dplyr)
print("Sya Poipkar S105")
data <- read.csv("~/S105/DATASET/Cardiovascular_Disease_Dataset.csv")
names(data)
# 1. Scatter Plot
# Age vs Serum Cholesterol
ggplot(data, aes(x = age, y = serumcholesterol)) +
  geom_point(color = "pink") +
  labs(
    title = "Scatter Plot of Age vs Serum Cholesterol",
    x = "Age",
    y = "Serum Cholesterol"
  ) +
  theme_minimal()

# 2. Pie Chart
# Gender Distribution
gender_data <- data %>%
  count(gender)

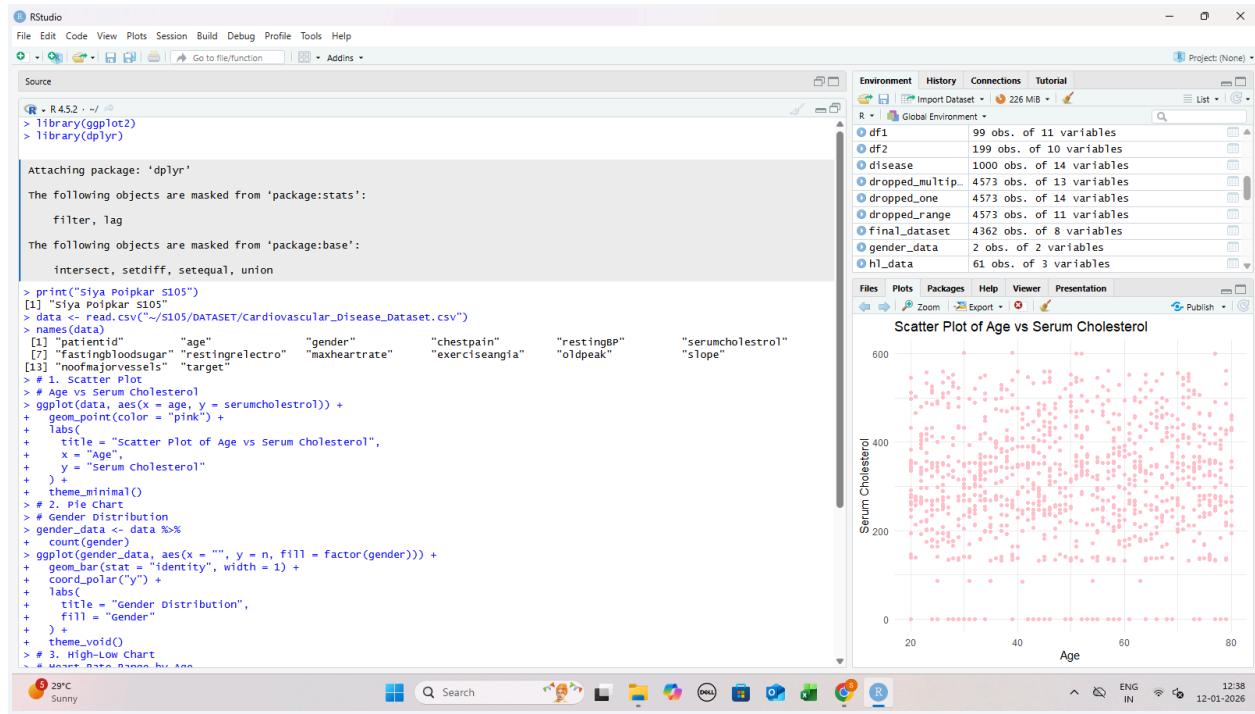
ggplot(gender_data, aes(x = "", y = n, fill = factor(gender))) +
  geom_bar(stat = "identity", width = 1) +
  coord_polar("y") +
  labs(
    title = "Gender Distribution",
    fill = "Gender"
  ) +
  theme_void()

# 3. High-Low Chart
# Heart Rate Range by Age
hl_data <- data %>%
  group_by(age) %>%
  summarise(
    min_hr = min(maxheartrate),
    max_hr = max(maxheartrate)
  )

ggplot(hl_data, aes(x = age)) +
```

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 10 TO 12 (MODULE 2)

```
geom_linerange(aes(ymin = min_hr, ymax = max_hr), color = "purple") +
  labs(
    title = "High-Low Chart of Heart Rate by Age",
    x = "Age",
    y = "Heart Rate"
  ) +
  theme_minimal()
```



MVLU COLLEGE

R PROGRAMMING

PRACTICAL NO. 10 TO 12 (MODULE 2)

RStudio

```

[R - R 4.5.2 - ~/]
[1] "patientid"      "age"          "gender"        "chestpain"     "restingBP"    "serumcholesterol"
[7] "fastingbloodsugar" "restingelectro" "maxheartrate"   "exerciseangia" "oldpeak"     "slope"
[13] "noofmajorvessels" "target"

# 1. Scatter Plot
> Age vs Serum Cholesterol
> ggplot(data, aes(x = age, y = serumcholesterol)) +
+   geom_point(color = 'pink') +
+   labs(
+     title = "Scatter Plot of Age vs Serum Cholesterol",
+     x = "Age",
+     y = "Serum Cholesterol"
+   ) +
+   theme_minimal()
> # 2. Pie Chart
> # Gender Distribution
> gender_data <- data %>%
+   count(gender)
> ggplot(gender_data, aes(x = "", y = n, fill = factor(gender))) +
+   geom_bar(stat = "identity", width = 1) +
+   coord_polar("y") +
+   labs(
+     title = "Gender Distribution",
+     fill = "Gender"
+   ) +
+   theme_void()
> # 3. High-Low Chart
> # Heart Rate Range by Age
> h1_data <- data %>%
+   group_by(age) %>%
+   summarise(
+     min_hr = min(maxheartrate),
+     max_hr = max(maxheartrate)
+   )
> ggplot(h1_data, aes(x = age)) +
+   geom_linerange(aes(ymin = min_hr, ymax = max_hr), color = "purple") +
+   labs(
+     title = "High-Low Chart of Heart Rate by Age",
+     x = "Age",
+     y = "Heart Rate"
+   ) +
+   theme_minimal()
>

```

29°C Sunny 12:39 12-01-2026

RStudio

```

[R - R 4.5.2 - ~/]
[1] "patientid"      "age"          "gender"        "chestpain"     "restingBP"    "serumcholesterol"
[7] "fastingbloodsugar" "restingelectro" "maxheartrate"   "exerciseangia" "oldpeak"     "slope"
[13] "noofmajorvessels" "target"

# 1. Scatter Plot
> Age vs Serum Cholesterol
> ggplot(data, aes(x = age, y = serumcholesterol)) +
+   geom_point(color = 'pink') +
+   labs(
+     title = "Scatter Plot of Age vs Serum Cholesterol",
+     x = "Age",
+     y = "Serum Cholesterol"
+   ) +
+   theme_minimal()
> # 2. Pie Chart
> # Gender Distribution
> gender_data <- data %>%
+   count(gender)
> ggplot(gender_data, aes(x = "", y = n, fill = factor(gender))) +
+   geom_bar(stat = "identity", width = 1) +
+   coord_polar("y") +
+   labs(
+     title = "Gender Distribution",
+     fill = "Gender"
+   ) +
+   theme_void()
> # 3. High-Low Chart
> # Heart Rate Range by Age
> h1_data <- data %>%
+   group_by(age) %>%
+   summarise(
+     min_hr = min(maxheartrate),
+     max_hr = max(maxheartrate)
+   )
> ggplot(h1_data, aes(x = age)) +
+   geom_linerange(aes(ymin = min_hr, ymax = max_hr), color = "purple") +
+   labs(
+     title = "High-Low Chart of Heart Rate by Age",
+     x = "Age",
+     y = "Heart Rate"
+   ) +
+   theme_minimal()
>

```

29°C Sunny 12:39 12-01-2026

Environment History Connections Tutorial

Global Environment

- df1 99 obs. of 11 variables
- df2 199 obs. of 10 variables
- disease 1000 obs. of 14 variables
- dropped_multip... 4573 obs. of 13 variables
- dropped_one 4573 obs. of 14 variables
- dropped_range 4573 obs. of 11 variables
- final_dataset 4362 obs. of 8 variables
- gender_data 2 obs. of 2 variables
- h1_data 61 obs. of 3 variables

Files Plots Packages Help Viewer Presentation

Gender Distribution

Gender

Gender	Percentage
0	~20%
1	~80%

High-Low Chart of Heart Rate by Age

Heart Rate

Age

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 10 TO 12 (MODULE 2)

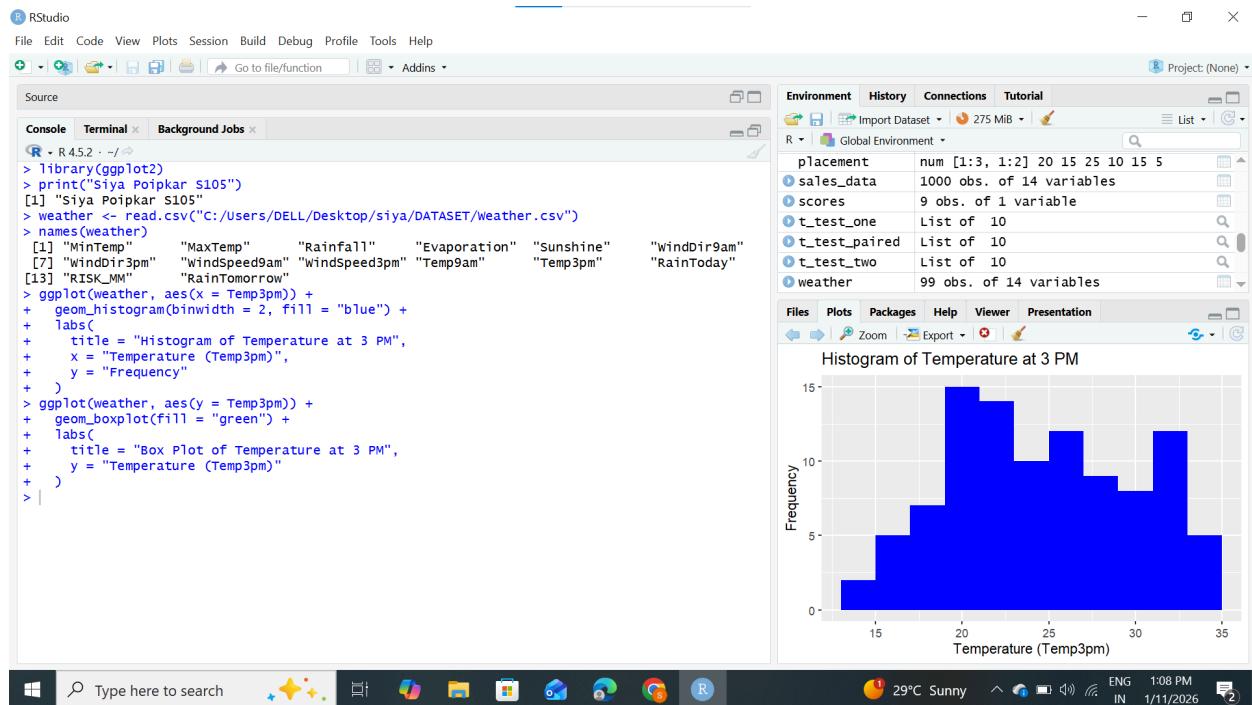
AIM 11: Generating histograms and box plots using ggplot2 (R).

```
library(ggplot2)
print("Sya Poipkar S105")
weather <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/Weather.csv")
names(weather)

ggplot(weather, aes(x = Temp3pm)) +
  geom_histogram(binwidth = 2, fill = "blue") +
  labs(
    title = "Histogram of Temperature at 3 PM",
    x = "Temperature (Temp3pm)",
    y = "Frequency"
  )

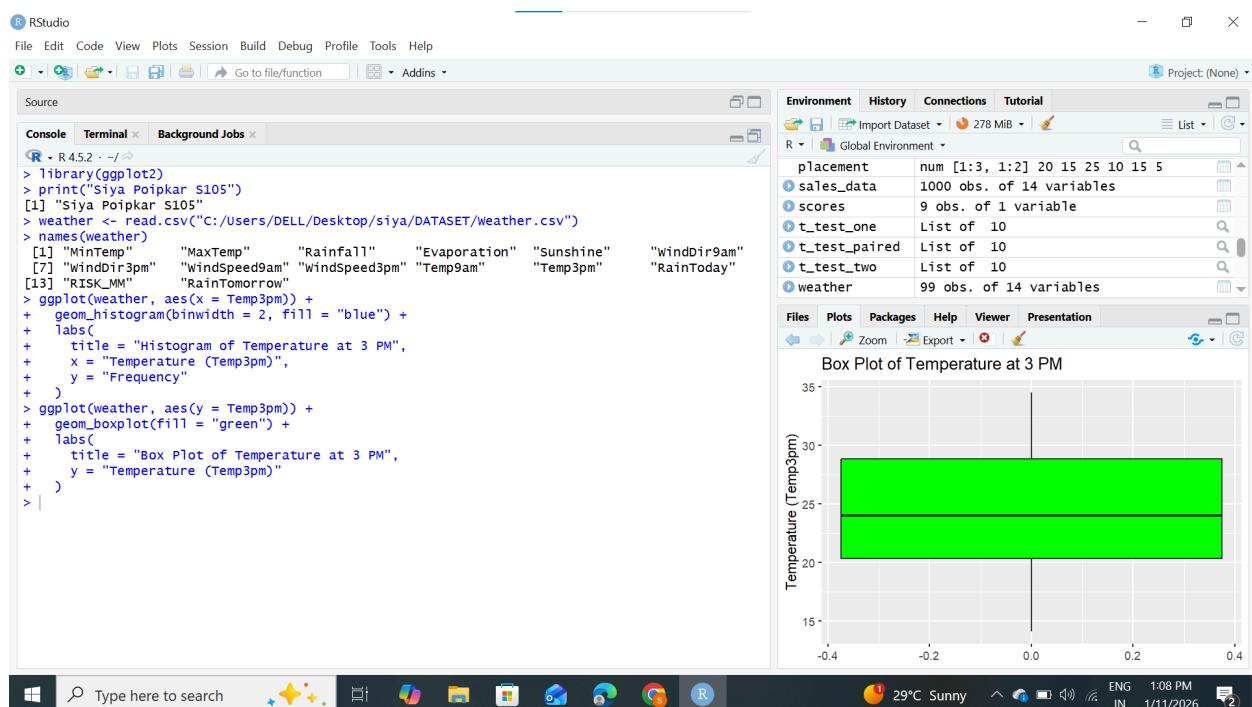
ggplot(weather, aes(y = Temp3pm)) +
  geom_boxplot(fill = "green") +
  labs(
    title = "Box Plot of Temperature at 3 PM",
    y = "Temperature (Temp3pm)"
  )
```

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 10 TO 12 (MODULE 2)



The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays R code for generating histograms and box plots of temperature data.
- Environment Tab:** Shows objects like placement, sales_data, scores, t_test_one, t_test_paired, t_test_two, and weather.
- Plots Tab:** Displays a histogram titled "Histogram of Temperature at 3 PM". The x-axis is "Temperature (Temp3pm)" ranging from 15 to 35, and the y-axis is "Frequency" ranging from 0 to 15. The distribution is roughly bell-shaped, peaking around 20-22.
- System Taskbar:** Shows system icons, a search bar, and a notification for 29°C Sunny.



The screenshot shows the RStudio interface with the following details:

- Console Tab:** Displays R code for generating histograms and box plots of temperature data.
- Environment Tab:** Shows objects like placement, sales_data, scores, t_test_one, t_test_paired, t_test_two, and weather.
- Plots Tab:** Displays a box plot titled "Box Plot of Temperature at 3 PM". The y-axis is "Temperature (Temp3pm)" ranging from 15 to 35. The box plot shows a median of approximately 23, an interquartile range from about 20 to 27, and whiskers extending from approximately -0.4 to 0.4.
- System Taskbar:** Shows system icons, a search bar, and a notification for 29°C Sunny.

MVLU COLLEGE
R PROGRAMMING
PRACTICAL NO. 10 TO 12 (MODULE 2)

AIM 12: Generating correlation matrices using `cor()` (R).

```
print("Sya Poipkar S105")
mall <- read.csv("C:/Users/DELL/Desktop/siya/DATASET/Mall_Customers.csv")

mall_numeric <- mall[, c("Age", "Annual.Income..k..", "Spending.Score..1.100.")]

cor_matrix <- cor(mall_numeric)

print(cor_matrix)
```

