

# Sheth L.U.J. & Sir M.V. College

## Practical 10: Creating Graphical Reports using ggplot2 ®

**Aim :-** To create graphical reports using ggplot2 in R for better data visualization and interpretation. The following graphs are generated:

1. Scatter Plot
2. Pie Chart
3. High–Low Chart

### Dataset Description

The dataset is loaded from an external CSV file and contains the following important variables:

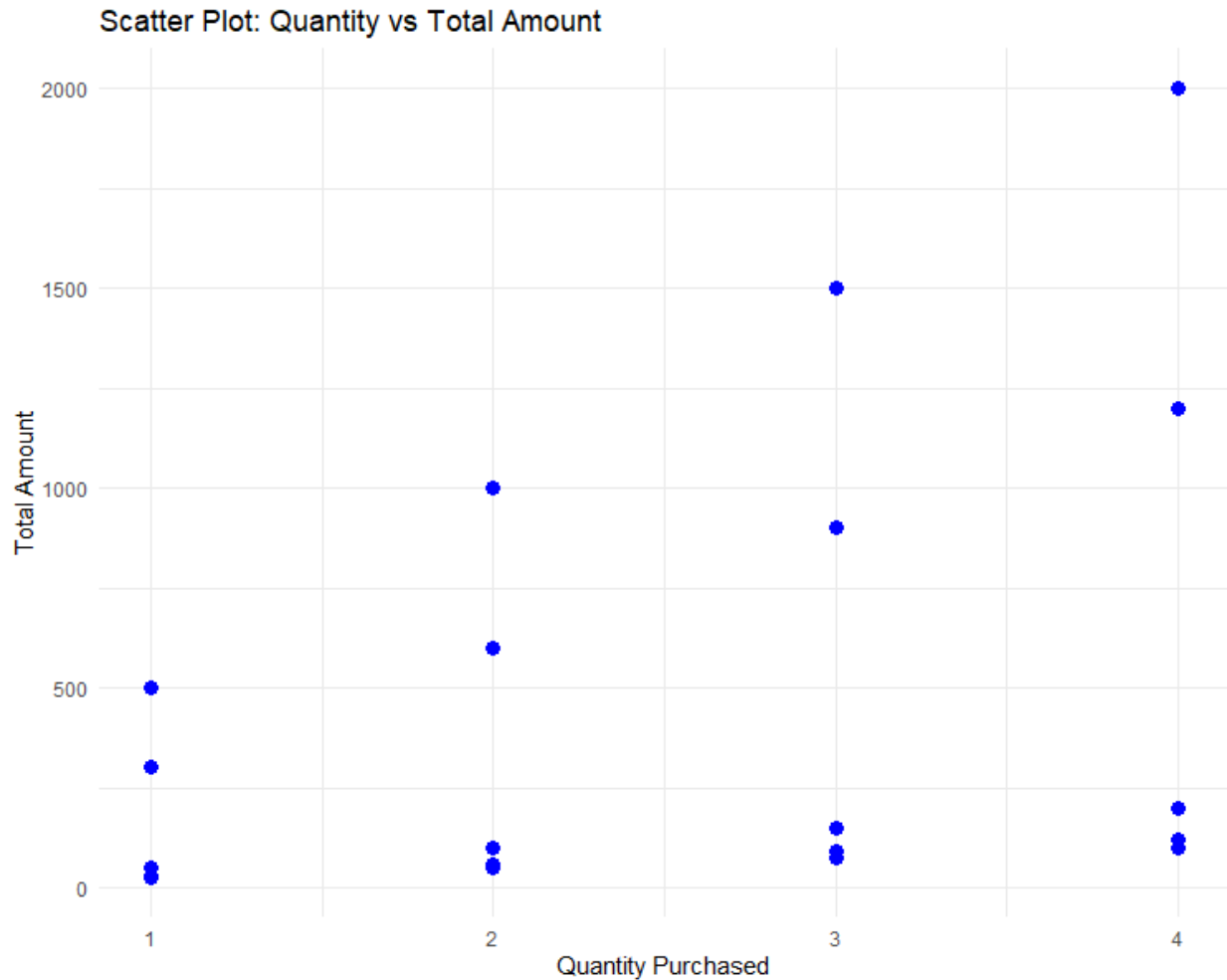
Variable Name	Description
Date	Date of transaction
Product.Category	Category of product (Electronics, Clothing, Beauty)
Quantity	Number of items purchased
Total.Amount	Total sales amount

**Name :- Priya Gupta**

**Roll no. :- So81**

# Sheth L.U.J. & Sir M.V. College

**Graph 1: Scatter Plot (Quantity vs Total Amount)**



## Explanation

- This scatter plot shows the relationship between quantity purchased and total sales amount.
- As the quantity increases, the total amount generally increases.

**Name :- Priya Gupta**

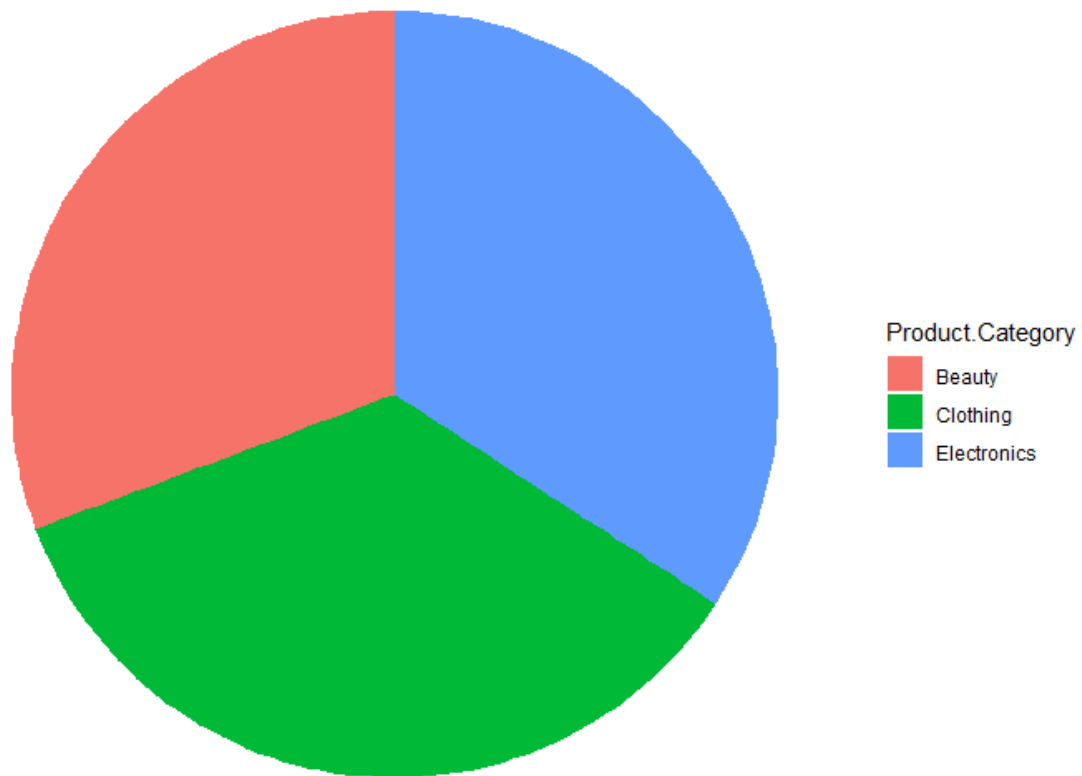
**Roll no. :- So81**

## Sheth L.U.J. & Sir M.V. College

- Higher quantities (3 and 4) show greater variation, indicating bulk purchases lead to higher sales values.
- The plot helps identify patterns and outliers in sales data.

### Graph 2: Pie Chart (Product Category Distribution)

Pie Chart: Product Category Distribution



**Name :- Priya Gupta**

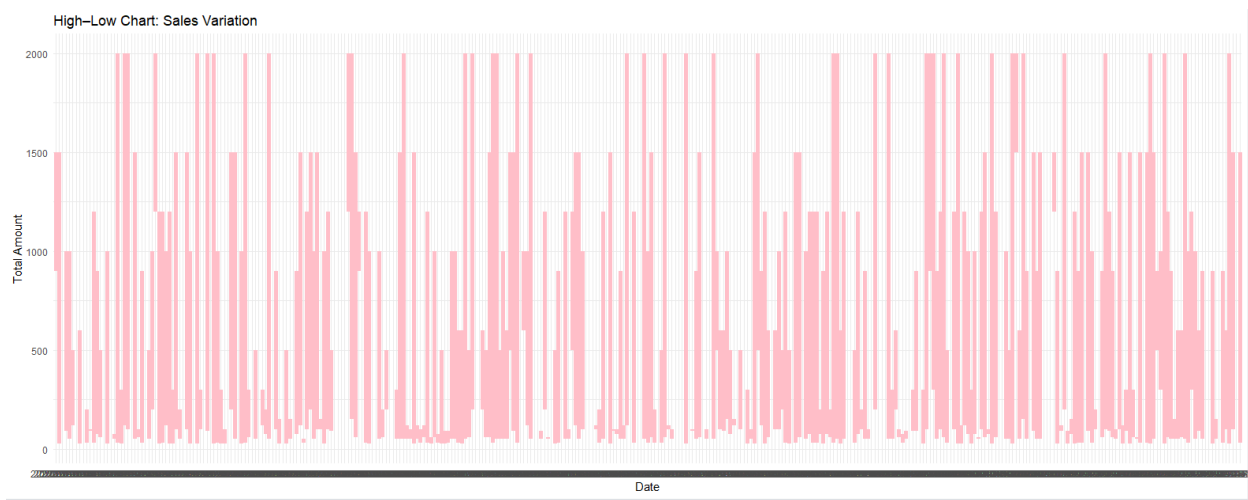
**Roll no. :- So81**

# Sheth L.U.J. & Sir M.V. College

## Explanation

- The pie chart represents the distribution of sales across product categories.
- Electronics, Clothing, and Beauty categories contribute almost equally.
- This indicates balanced sales performance across all categories.
- Useful for understanding which category dominates the market share.

## Graph 3: High–Low Chart (Sales Variation Over Time)



## Explanation

- The High–Low chart shows daily sales variation.

**Name :- Priya Gupta**

**Roll no. :- So81**

## **Sheth L.U.J. & Sir M.V. College**

- The lower end represents minimum sales, and the upper end shows maximum sales per day.
- Light color improves readability and avoids visual clutter.
- This chart helps identify volatile and stable sales periods.

### **Overall Observations**

- Scatter plot confirms a positive relationship between quantity and total amount.
- Pie chart shows even distribution across product categories.
- High–Low chart highlights sales fluctuations over time, useful for trend analysis.

### **Conclusion**

This practical demonstrates the use of ggplot2 for creating meaningful visualizations.

Graphical representation improves understanding of data patterns, trends, and distributions, making analysis more effective and interpretable.

**Name :- Priya Gupta**

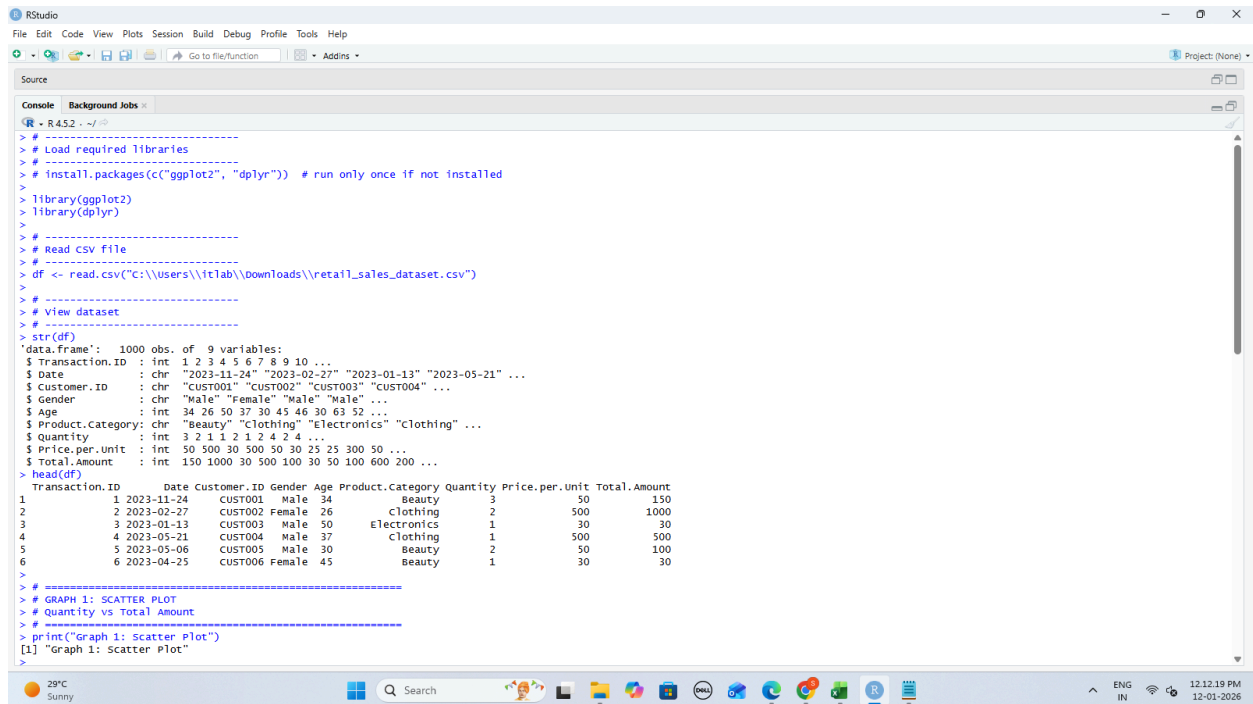
**Roll no. :- So81**

# Sheth L.U.J. & Sir M.V. College

## Result

Scatter Plot, Pie Chart, and High–Low Chart were successfully created using ggplot2, and the sales data was analyzed visually.

## Screenshots



```
R - R 4.5.2 - ~/RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins
Project: (None)
Source
Console Background Jobs
> # -----
> # Load required libraries
> #
> # install.packages(c("ggplot2", "dplyr")) # run only once if not installed
>
> library(ggplot2)
> library(dplyr)
>
> # -----
> # Read CSV file
> # -----
> df <- read.csv("c:\\Users\\ftlab\\Downloads\\retail_sales_dataset.csv")
>
> # -----
> # View dataset
> # -----
> str(df)
'data.frame': 1000 obs. of  9 variables:
 $ Transaction.ID : int  1 2 3 4 5 6 7 8 9 10 ...
 $ Date           : chr  "2023-11-24" "2023-02-27" "2023-01-13" "2023-05-21" ...
 $ Customer.ID    : chr  "CUST001" "CUST002" "CUST003" "CUST004" ...
 $ Gender         : chr  "Male" "Female" "Male" "Male" ...
 $ Age           : int   34 26 50 37 30 45 46 30 63 52 ...
 $ Product.Category: chr  "Beauty" "Clothing" "Electronics" "Clothing" ...
 $ Quantity       : int   3 2 1 1 2 1 2 4 2 4 ...
 $ Price.per.unit : int   50 500 30 500 50 30 25 25 300 50 ...
 $ Total.Amount   : int  150 1000 30 500 100 30 50 100 600 200 ...
> head(df)
  Transaction.ID      Date Customer.ID Gender Age Product.Category Quantity Price.per.unit Total.Amount
1      1 2023-11-24    CUST001   Male   34      Beauty              3             50             150
2      2 2023-02-27    CUST002  Female   26      Clothing             2            500            1000
3      3 2023-01-13    CUST003   Male   50      Electronics            1             30             30
4      4 2023-05-21    CUST004   Male   37      Clothing             1            500             500
5      5 2023-05-06    CUST005   Male   30      Beauty              2             50             100
6      6 2023-04-25    CUST006  Female   45      Beauty              1             30             30
>
> # -----
> # GRAPH 1: SCATTER PLOT
> # Quantity vs Total Amount
> # -----
> print("Graph 1: Scatter Plot")
[1] "Graph 1: Scatter Plot"
```

Name :- Priya Gupta

Roll no. :- So81

# Sheth L.U.J. & Sir M.V. College

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)

Source
Console Background Jobs
R • R 4.52.2 ~ /
> # QUANTITY vs TOTAL AMOUNT
> # Quantity vs Total Amount
> # =====
> print("Graph 1: Scatter Plot")
[1] "Graph 1: Scatter Plot"
>
> ggplot(df, aes(x = Quantity, y = Total.Amount)) +
+   geom_point(color = "blue", size = 3) +
+   labs(
+     title = "Scatter Plot: Quantity vs Total Amount",
+     x = "Quantity Purchased",
+     y = "Total Amount"
+   ) +
+   theme_minimal()
>
> # =====
> # GRAPH 2: PIE CHART
> # Product Category Distribution
> # =====
> print("GRAPH 2: PIE CHART")
[1] "GRAPH 2: PIE CHART"
>
> product_data <- df %>%
+   count(Product.Category)
>
> print("Graph 2: Pie chart")
[1] "Graph 2: Pie chart"
>
> ggplot(product_data, aes(x = "", y = n, fill = Product.Category)) +
+   geom_bar(stat = "identity", width = 1) +
+   coord_polar("y") +
+   labs(title = "Pie chart: Product Category Distribution") +
+   theme_void()
>
> # =====
> # GRAPH 3: HIGH-LOW CHART
> # Sales Variation by Date
> # =====
> print("GRAPH 3: HIGH-LOW CHART")
[1] "GRAPH 3: HIGH-LOW CHART"
>
```

```
RStudio
File Edit Code View Plots Session Build Debug Profile Tools Help
Go to file/function Addins Project: (None)

Source
Console Background Jobs
R • R 4.52.2 ~ /
+   count(Product.Category)
>
> print("Graph 2: Pie chart")
[1] "Graph 2: Pie chart"
>
> ggplot(product_data, aes(x = "", y = n, fill = Product.Category)) +
+   geom_bar(stat = "identity", width = 1) +
+   coord_polar("y") +
+   labs(title = "Pie chart: Product Category Distribution") +
+   theme_void()
>
> # =====
> # GRAPH 3: HIGH-LOW CHART
> # Sales Variation by Date
> # =====
> print("GRAPH 3: HIGH-LOW CHART")
[1] "GRAPH 3: HIGH-LOW CHART"
>
> hl_data <- df %>%
+   group_by(Date) %>%
+   summarise(
+     High = max(Total.Amount, na.rm = TRUE),
+     Low = min(Total.Amount, na.rm = TRUE)
+   )
>
> print("Graph 3: High-Low Chart")
[1] "Graph 3: High-Low Chart"
>
> ggplot(hl_data, aes(x = Date)) +
+   geom_linerange(
+     aes(ymin = Low, ymax = High),
+     color = "pink",
+     size = 2
+   ) +
+   labs(
+     title = "High-Low chart: Sales Variation",
+     x = "Date",
+     y = "Total Amount"
+   ) +
+   theme_minimal()
>
```

Name :- Priya Gupta

Roll no. :- So81