

|                         |  |
|-------------------------|--|
| <b>EXPT NO:6</b>        | <b>IMPLEMENTATION OF MULTIVARIATE DISPLAYS</b> |
| <b>DATE: 24.01.2026</b> |  |

### **PRE-LAB QUESTIONS:**

#### **1. Why are multivariate displays important in AI analytics?**

Multivariate displays allow simultaneous visualization of multiple variables, helping AI analysts understand relationships, patterns, correlations, and anomalies in high-dimensional data. They improve feature understanding, model interpretability, and data-driven decision-making.

#### **2. How do parallel coordinates differ from scatter plots?**

Scatter plots typically show relationships between **two variables**, whereas parallel coordinate plots display **many variables at once**, with each observation represented as a line crossing multiple axes.

#### **3. What challenges exist in interpreting multivariate plots?**

- Visual clutter with large datasets
- Overlapping data points
- Difficulty identifying individual patterns
- Requires careful scaling and color encoding

#### **4. Where are trellis displays commonly used?**

Trellis displays are commonly used in:

- Retail analytics
  - Time series analysis
  - Geographic comparisons
  - Customer segmentation
- They allow comparisons across subsets of data using small multiples.

#### **5. How does multivariate visualization aid model evaluation?**

It helps in:

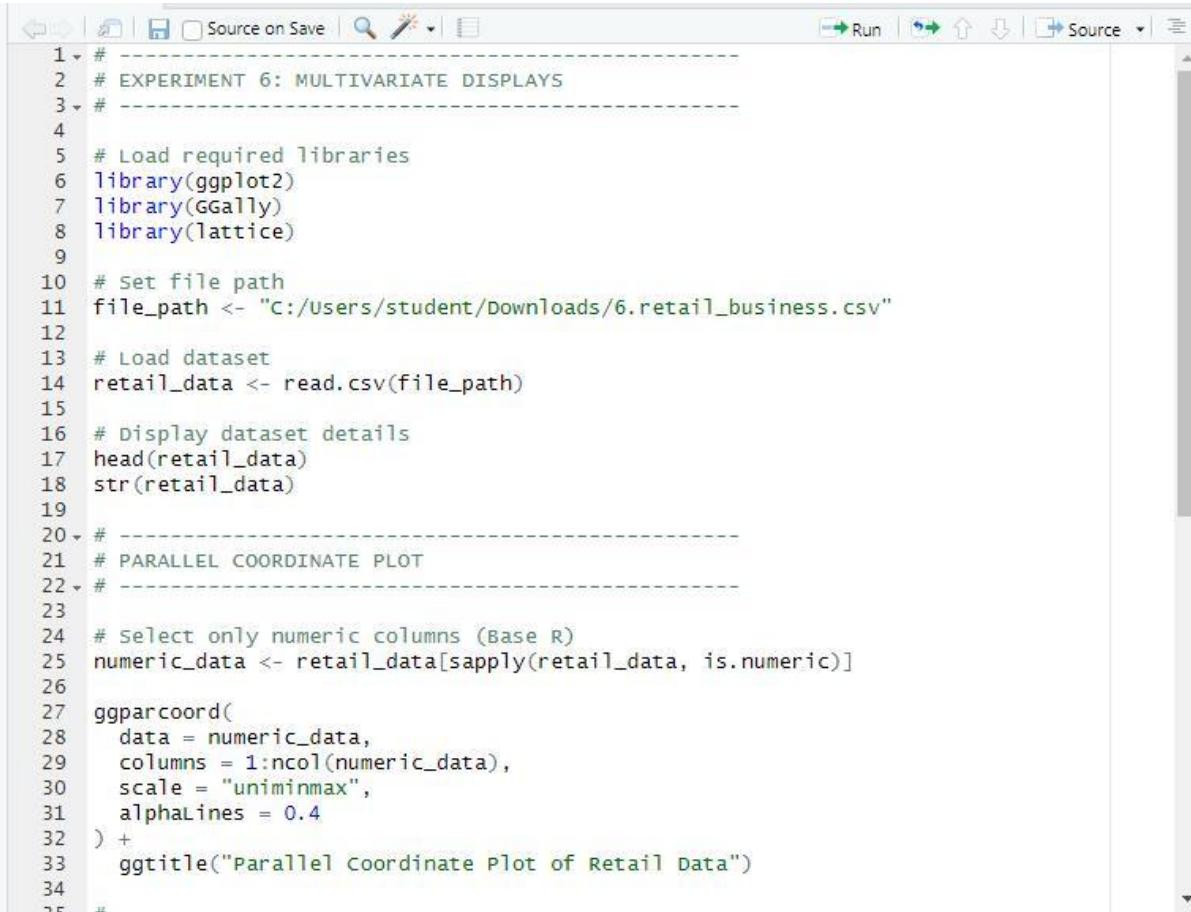
- Identifying feature importance
- Detecting overfitting and outliers
- Comparing model performance across categories
- Understanding prediction behavior across segments

**OBJECTIVE :** To implement advanced multivariate displays for complex data analysis.

**SCENARIO** A retail analytics firm studies sales, profit, customer segment, and region to optimize business strategy.

**IN-LAB TASKS (Using R Language)** • Create parallel coordinate plots • Generate bubble charts • Implement trellis displays by region

**Code:**



The screenshot shows an RStudio interface with the following code in the script pane:

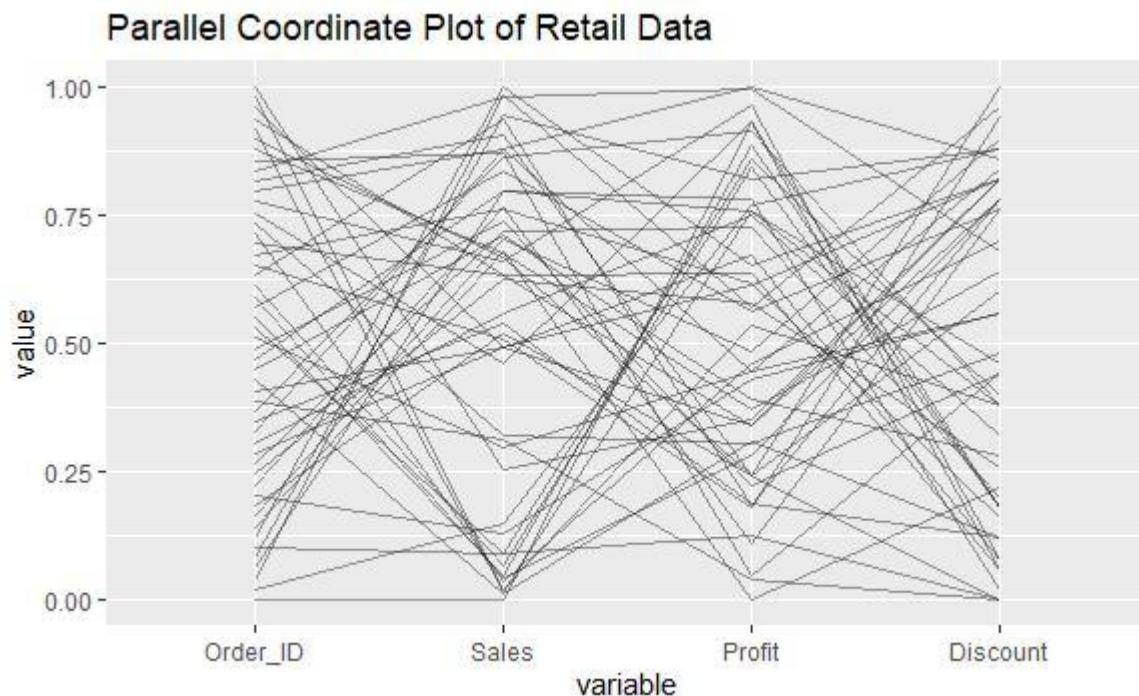
```
1 # -----
2 # EXPERIMENT 6: MULTIVARIATE DISPLAYS
3 # -----
4
5 # Load required libraries
6 library(ggplot2)
7 library(GGally)
8 library(lattice)
9
10 # Set file path
11 file_path <- "c:/users/student/Downloads/6.retail_business.csv"
12
13 # Load dataset
14 retail_data <- read.csv(file_path)
15
16 # Display dataset details
17 head(retail_data)
18 str(retail_data)
19
20 # -----
21 # PARALLEL COORDINATE PLOT
22 # -----
23
24 # Select only numeric columns (Base R)
25 numeric_data <- retail_data[sapply(retail_data, is.numeric)]
26
27 ggparcoord(
28   data = numeric_data,
29   columns = 1:ncol(numeric_data),
30   scale = "uniminmax",
31   alphaLines = 0.4
32 ) +
33   ggtitle("Parallel Coordinate Plot of Retail Data")
34
35 #
```

```

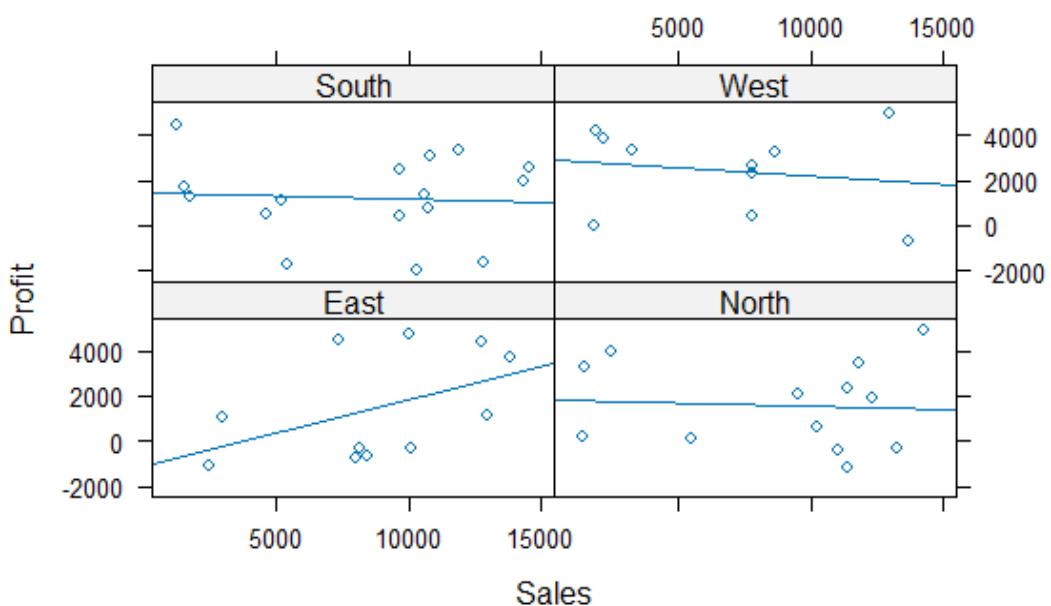
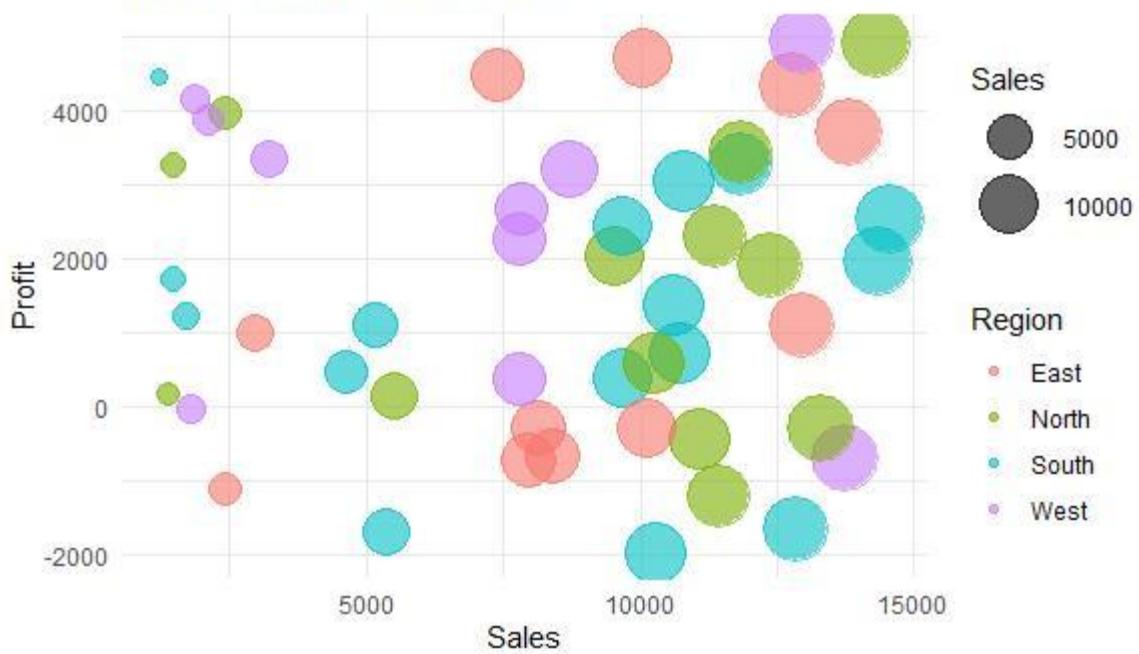
 30   scale = "uniminmax",
 31   alphaLines = 0.4
 32 ) +
 33   ggtitle("Parallel Coordinate Plot of Retail Data")
 34
 35 # -----
 36 # BUBBLE CHART
 37 # Sales vs Profit (Bubble size = Sales)
 38 # -----
 39
 40 ggplot(retail_data, aes(
 41   x = Sales,
 42   y = Profit,
 43   size = Sales,
 44   color = Region
 45 )) +
 46   geom_point(alpha = 0.6) +
 47   scale_size_continuous(range = c(3, 12)) +
 48   ggtitle("Bubble chart: sales vs Profit") +
 49   theme_minimal()
 50
 51 # -----
 52 # TRELLIS DISPLAY BY REGION
 53 # -----
 54
 55 xyplot(
 56   Profit ~ Sales | Region,
 57   data = retail_data,
 58   layout = c(2, 2),
 59   type = c("p", "r"),
 60   xlab = "Sales",
 61   ylab = "Profit"
 62 )
 63

```

**Output:**



Bubble Chart: Sales vs Profit



## **POST-LAB QUESTIONS**

### **1. What insights are gained from parallel coordinates?**

Parallel coordinate plots help identify patterns, trends, and outliers across multiple variables simultaneously. They reveal relationships between variables such as sales, profit, and discount, and show how different records vary across dimensions.

### **2. How does faceting simplify complex data?**

Faceting divides large datasets into smaller, category-based subsets. This makes it easier to compare patterns across groups such as regions and improves clarity in understanding multivariate relationships.

### **3. What limitations exist in bubble charts?**

Bubble charts can become cluttered when many data points overlap. Differences in bubble sizes may also be hard to interpret accurately, and large bubbles can visually dominate smaller ones.

### **4. How can these displays support AI-driven recommendations?**

Multivariate displays help identify key trends, correlations, and anomalies in data, which AI models can use to generate accurate, data-driven business recommendations and strategic decisions.

### **5. Suggest improvements for large multivariate datasets.**

For large datasets, interactive visualizations, data filtering, dimensionality reduction techniques, clustering, and sampling methods can be used to reduce complexity and improve interpretability.

**LEARNING OUTCOME: Students apply multivariate visualization for business intelligence.**

## **ASSESSMENT**

| Description              | Max Marks | Marks Awarded |
|--------------------------|-----------|---------------|
| Pre Lab Exercise         | 5         |               |
| In Lab Exercise          | 10        |               |
| Post Lab Exercise        | 5         |               |
| Viva                     | 10        |               |
| <b>Total</b>             | <b>30</b> |               |
| <b>Faculty Signature</b> |           |               |

