EXPT NO:6	IMPLEMENTATION OF MULTIVARIATE DISPLAYS
DATE: 24.01.2026	

PRE-LAB QUESTIONS

1. **Why are multivariate displays important in AI analytics?**
   Multivariate displays are important because AI datasets contain many variables/features. These plots help us visualize relationships, patterns, trends, and outliers across multiple variables together, which improves decision-making.

2. **How do parallel coordinates differ from scatter plots?**
   A scatter plot shows the relationship between only two variables at a time. Parallel coordinate plots can represent many variables at once, where each data record is drawn as a line across multiple parallel axes.

3. **What challenges exist in interpreting multivariate plots?**
   Multivariate plots can become cluttered when there are many data points. Overlapping lines/points make patterns hard to detect, and proper scaling and grouping are required to interpret the plot correctly.

4. **Where are trellis displays commonly used?**
   Trellis displays are commonly used in business intelligence, retail analytics, customer segmentation, regional analysis, and dashboards to compare patterns across multiple groups easily.

5. **How does multivariate visualization aid model evaluation?**
   It helps in model evaluation by showing feature relationships, correlations, separability of classes, bias across groups, and outliers. This helps improve model accuracy, fairness, and performance.

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.

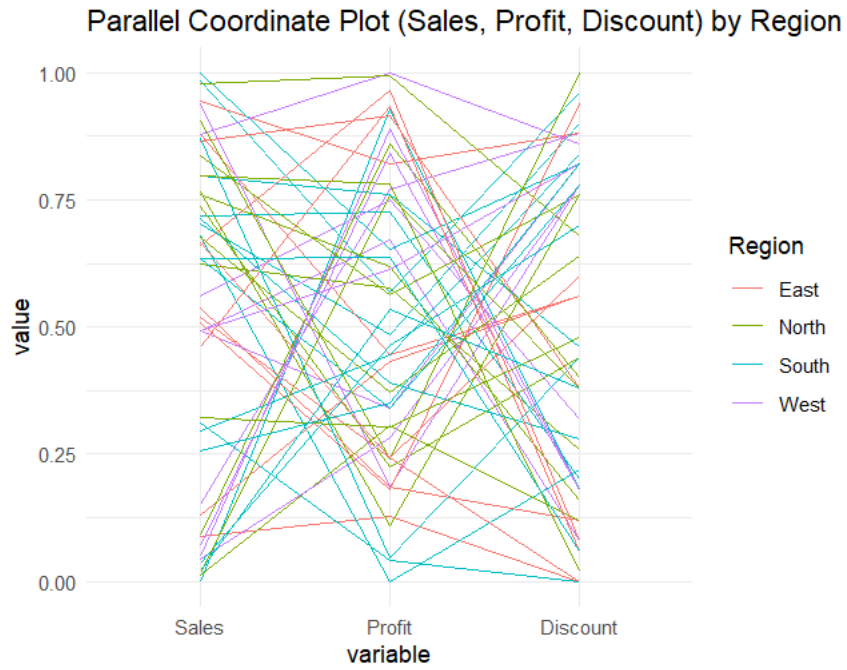
IN-LAB TASKS (Using R Language) • Create parallel coordinate plots • Generate bubble charts • Implement trellis displays by region

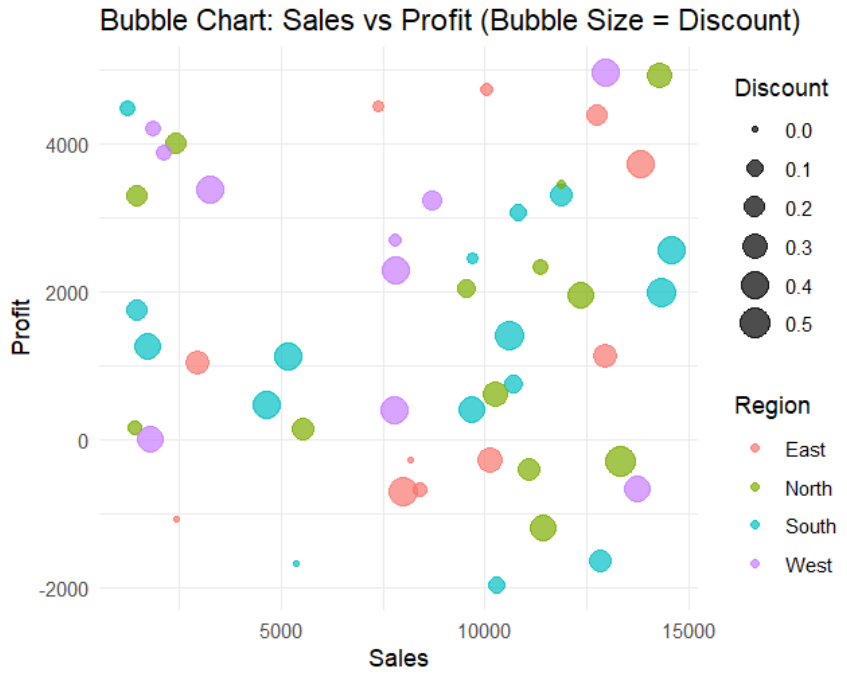
CODE:

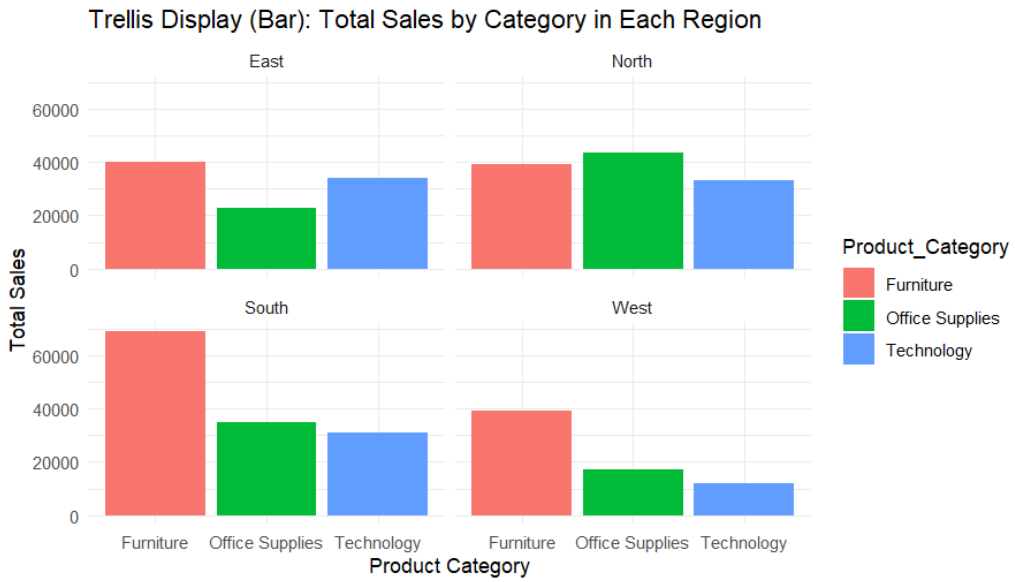
```
# ============================================================
# EXPT NO: 6 - IMPLEMENTATION OF MULTIVARIATE DISPLAYS
# ROLL NO: 23BAD076
# ============================================================

# Load Required Libraries

library(ggplot2)
library(GGally)
```

```
# Load Dataset

retail <- read.csv("6.retail_business.csv")

# 1) PARALLEL COORDINATE PLOT

# Numeric columns + Region for grouping
pc_data <- retail[, c("Sales", "Profit", "Discount", "Region")]

ggparcoord(data = pc_data,
        columns = 1:3,
        groupColumn = 4,
        scale = "uniminmax") +
 ggtitle("Parallel Coordinate Plot (Sales, Profit, Discount) by Region") +
 theme_minimal()

# 2) BUBBLE CHART

ggplot(retail, aes(x = Sales, y = Profit, size = Discount, color = Region)) +
 geom_point(alpha = 0.7) +
 ggtitle("Bubble Chart: Sales vs Profit (Bubble Size = Discount)") +
 xlab("Sales") +
 ylab("Profit") +
 theme_minimal()

# 3) TRELLIS DISPLAY BY REGION (BAR FORMAT)

# Summarize Total Sales by Region and Product Category
sales_summary <- aggregate(Sales ~ Region + Product_Category, data = retail, sum)

ggplot(sales_summary, aes(x = Product_Category, y = Sales, fill = Product_Category)) +
 geom_bar(stat = "identity") +
 facet_wrap(~ Region) +
 ggtitle("Trellis Display (Bar): Total Sales by Category in Each Region") +
 xlab("Product Category") +
 ylab("Total Sales") +
 theme_minimal()
```

OUTPUT:


Parallel Coordinate Plot (Sales, Profit, Discount) by Region
1.00
0.75
0.50
0.25
0.00
value
Sales
Profit
Discount
variable
Region
East
North
South
West


Bubble Chart: Sales vs Profit (Bubble Size = Discount)
4000
2000
0
-2000
Profit
5000
10000
15000
Sales
Discount
0.0
0.1
0.2
0.3
0.4
0.5
Region
East
North
South
West



POST-LAB QUESTIONS

1. **What insights are gained from parallel coordinates?**
   Parallel coordinate plots help us compare multiple variables (Sales, Profit, Discount) at the same time. They clearly show patterns such as high sales with low profit, the impact of discount on profit, and differences between regions.

2. **How does faceting simplify complex data?**
   Faceting divides a large plot into multiple smaller plots based on a category (like Region). This makes the data easier to interpret because each region is analyzed separately, reducing clutter and confusion.

3. **What limitations exist in bubble charts?**
   Bubble charts become difficult to interpret when there are many points. Bubbles may overlap, exact bubble size comparison is hard, and small differences in discount values are not easily visible.

4. **How can these displays support AI-driven recommendations?**
   These visualizations help AI systems identify high-performing regions and customer segments, detect loss-making discount strategies, and recommend better pricing, product focus, and region-wise marketing strategies.

5. **Suggest improvements for large multivariate datasets.**
   For large datasets, we can use sampling, clustering, interactive visualizations (Plotly/Shiny), filtering, dimensionality reduction methods (PCA), and better aggregation techniques to reduce clutter and improve readability.

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**	
Total	**30**	
Faculty Signature		