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 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.

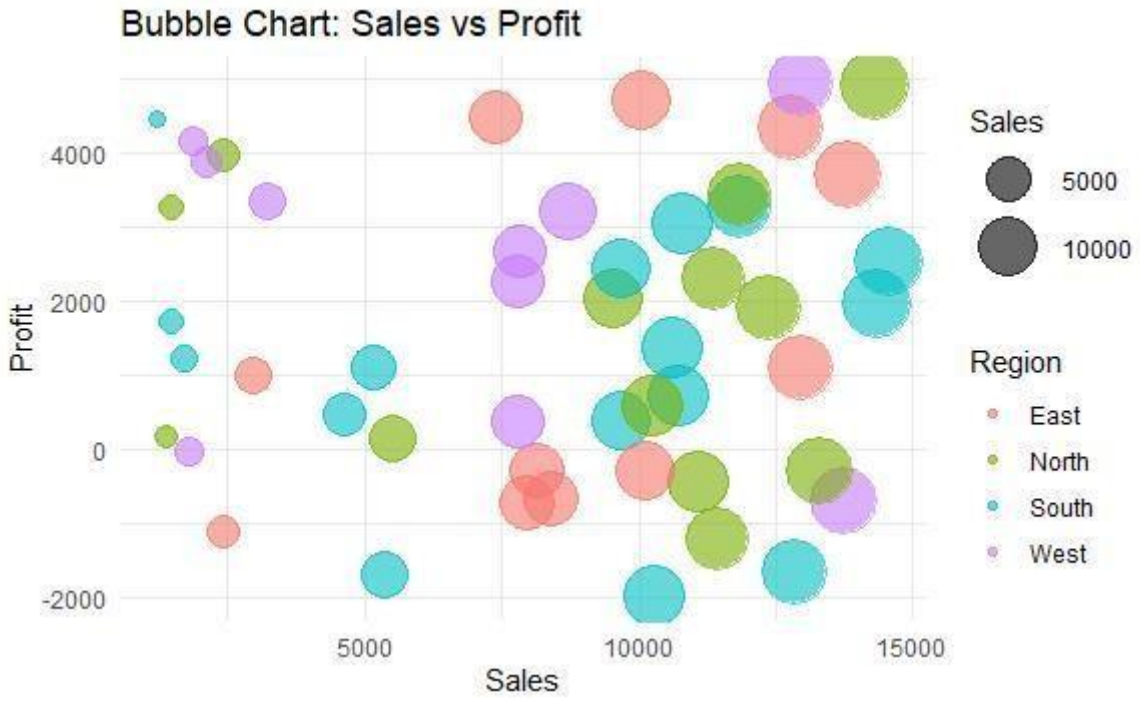
IN-LAB TASKS (Using R Language) • Create parallel coordinate plots • Generate bubble charts • Implement trellis displays by region

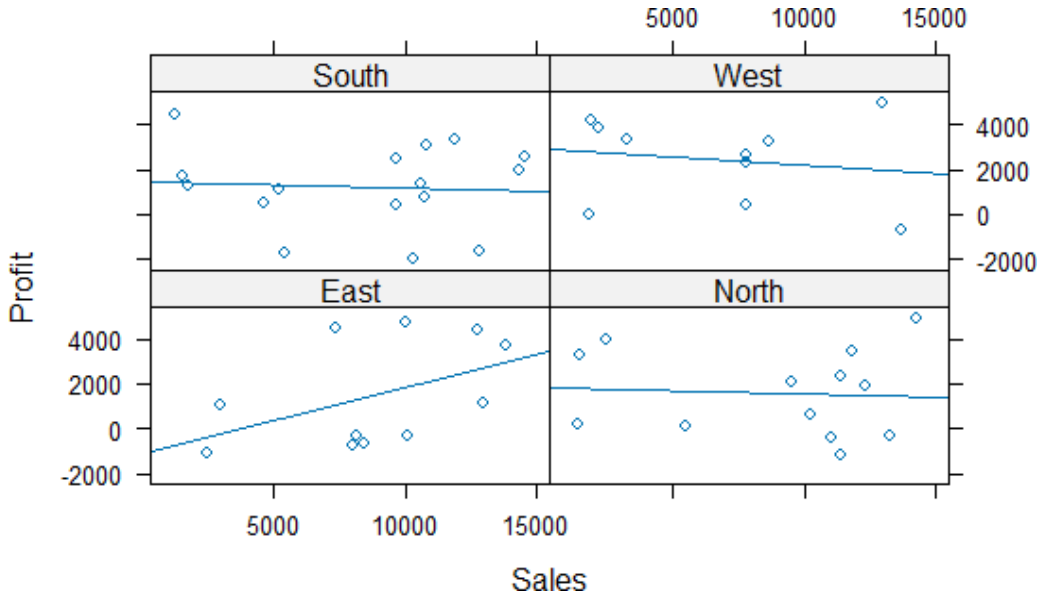
Code:

```
# ---------------------------------------------
# EXPERIMENT 6: MULTIVARIATE DISPLAYS
# ---------------------------------------------
("Sai Vaishnavi R 23BAD094")
# Load required libraries
library(ggplot2)
library(GGally)
library(lattice)

# Set file path
file_path <- "C:/Users/student/Downloads/6.retail_business.csv"

# Load dataset
retail_data <- read.csv(file_path)

# Display dataset details
head(retail_data)
str(retail_data)

# ---------------------------------------------
# PARALLEL COORDINATE PLOT
# ---------------------------------------------

# Select only numeric columns (Base R)
numeric_data <- retail_data[sapply(retail_data, is.numeric)]

ggparcoord(
  data = numeric_data,
  columns = 1:ncol(numeric_data),
  scale = "uniminmax",
  alphaLines = 0.4
) +
  ggtitle("Parallel Coordinate Plot of Retail Data")

# ---------------------------------------------
# BUBBLE CHART
# Sales vs Profit (bubble size = Sales)
# ---------------------------------------------

ggplot(retail_data, aes(
  x = Sales,
  y = Profit,
  size = Sales,
  color = Region
)) +
  geom_point(alpha = 0.6) +
  scale_size_continuous(range = c(3, 12)) +
  ggtitle("Bubble Chart: Sales vs Profit") +
  theme_minimal()

# ---------------------------------------------
# TRELLIS DISPLAY BY REGION
# ---------------------------------------------

xyplot(
  Profit ~ Sales | Region,
  data = retail_data,
  layout = c(2, 2),
  type = c("p", "r"),
  xlab = "Sales",
  ylab = "Profit"
)
```

Output:


Parallel Coordinate Plot of Retail Data
1.00
0.75
0.50
0.25
0.00
value
Order_ID
Sales
Profit
Discount
variable


Bubble Chart: Sales vs Profit
Profit
4000
2000
0
-2000
5000
10000
15000
Sales
Sales
5000
10000
Region
East
North
South
West


5000
10000
15000
South
West
East
North
Profit
4000
2000
0
-2000
Sales

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	
Total	**30**	
	Faculty Signature	