Overview of the activity

This activity demonstrates how to perform Market Basket Analysis using the Apriori algorithm in R.

Using the **Groceries** dataset from the arules package, I performed:

- Load and explore the dataset
- Convert it into transactions format
- Generate frequent itemsets
- Mine association rules
- Analyzed and visualized the rules

1.) Install Libraries and dataset loading

```r
install.packages("arules")
install.packages("arulesViz")

# load libraries
library(arules)
library(arulesViz)

# load dataset
data("Groceries")

# verify
summary(Groceries)

# inspect sample/transactions
inspect(Groceries[1:5])
```

2.) Explored top 20 most frequent items

```r
#itemFrequency
itemFrequencyPlot(Groceries, topN = 20, type = "absolute")
```

Key findiings:

- Whole milk
- Other vegetables
- Rolls/buns
- Soda
- Yogurt

3.) Generating data rules

```
# apply Apriori
rules <- apriori(Groceries, parameter = list(supp = 0.005, conf = 0.5))

rules
```

Algorithm parameters:

- Support = 0.005 => itemsets appears in atleast 0.5% of transactions
- Confidence = 0.5 => rule must be >=50% accurate

4.) Inspect and sort rules

```
# inspect rules
inspect(head(rules))

rules_sorted <- sort(rules, by = "lift", decreasing = TRUE)
inspect(head(rules_sorted))
```

5.) Sample rule Interpretaition

{ yogurt } => { whole milk }

Support = 0.056, confidence = 0.40, lift = 1.57

Interpretation:

40% of the baskets containing yogurt also contains whole milk.

Lift = 1.57 , the relationship is 57% stronger than random chance

6.) Visualization of the rules

```
# visualize
plot(rules, method = "scatterplot")
plot(rules, method = "graph", engine = "htmlwidget")
plot(rules, method = "grouped")
```