Page 1

**Capstone Project**

**Preliminary Stage Assignment 2**

**Course code:** CSA1643

**Course :** Data warehousing and Data Mining for Data Science

**S.No**: 37

**Name** : S.Devendra

**Reg. No**: 192211393

**Slot** : c

**Title :** : Supply Chain Risk Management With Data mining Techniques

**Assignment Release Date** : 14/02/2024

**Assignment** **Preliminary Stage ( Assignment 2 ) submission Date** : 16/02/2024  
**Mentor Name :** Dr.R.Rajakumari

**Mentor Phone number and Department** : 98418 82726 and Department of Placement

**R PROGRAMMING :**# Load required libraries

library(cluster)

# Generate synthetic data

set.seed(123)

# Number of products and suppliers

num\_products <- 100

num\_suppliers <- 50

# Generate product data

products <- data.frame(

ProductID = 1:num\_products,

ProductName = paste0("Product\_", 1:num\_products),

Demand = rnorm(num\_products, mean = 1000, sd = 200),

Price = rnorm(num\_products, mean = 50, sd = 10)

)

# Generate supplier data

suppliers <- data.frame(

SupplierID = 1:num\_suppliers,

SupplierName = paste0("Supplier\_", 1:num\_suppliers),

LeadTime = sample(1:30, num\_suppliers, replace = TRUE),

QualityRating = sample(1:5, num\_suppliers, replace = TRUE)

)

# Generate risk factors

risk\_factors <- data.frame(

ProductID = sample(1:num\_products, num\_products \* 0.2, replace = TRUE),

SupplierID = sample(1:num\_suppliers, num\_suppliers \* 0.2, replace = TRUE),

Delay = rbinom(n = num\_products \* 0.2, size = 1, prob = 0.1),

QualityIssue = rbinom(n = num\_suppliers \* 0.2, size = 1, prob = 0.1)

)

# Combine risk factors with product and supplier data

risk\_data <- merge(risk\_factors, products, by = "ProductID")

risk\_data <- merge(risk\_data, suppliers, by = "SupplierID")

# Perform hierarchical clustering

hc <- hclust(dist(risk\_data[, c("Delay", "QualityIssue", "Demand", "Price", "LeadTime", "QualityRating")]))

plot(hc, main = "Dendrogram of Risk Factors")

# Cut the dendrogram to form clusters

clusters <- cutree(hc, k = 3)

# Add cluster information to the risk data

risk\_data$Cluster <- clusters

# Summary of clusters

summary(clusters)

# Output the risk data with cluster information

head(risk\_data)  
  
  
  
 **OUTPUT :**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Supplier id** | **Product id** | **Delay** | **Quality Issue** | **Product Name** | **Demand** | **Price** | **Supplier Name** | **Lead Time** |
| 3 | 5 | 0 | 0 | Product\_5 | 1025.8575 | 40.48381 | Supplier\_3 | 3 |
| 3 | 76 | 0 | 0 | Product\_76 | 1205.1143 | 39.04004 | Supplier\_3 | 3 |
| 4 | 11 | 0 | 0 | Product\_11 | 1244.8164 | 44.24653 | Supplier\_4 | 9 |
| 4 | 53 | 0 | 0 | Product\_53 | 991.4259 | 53.32203 | Supplier\_4 | 9 |
| 6 | 70 | 0 | 0 | Product\_70 | 1410.0169 | 53.68965 | Supplier\_6 | 7 |
| 6 | 56 | 0 | 0 | Product\_56 | 1303.2941 | 47.19605 | Supplier\_6 | 7 |

|  |  |
| --- | --- |
| **Quality Rating** | **Cluster** |
| 4 | 1 |
| 4 | 2 |
| 1 | 2 |
| 1 | 1 |
| 1 | 2 |
| 1 | 2 |

