VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A4.3: Multidimensional Scaling

SAYA SANTHOSH

V01101901

Date of Submission: 08-07-2024

CONTENTS

Introduction

Multidimensional Scaling:
In order to assess and comprehend links between the several ice cream parameters in the "icecream.csv" dataset—which include Brand, Price, Availability, Taste, Flavour, Consistency, and Shelflife—this assignment requires employing Multidimensional Scaling (MDS). The aim is to identify trends and similarities between various ice cream companies by visualizing their proximity based on these features. Data preprocessing, feature standardization, missing value checks, and exploratory data analysis (EDA) employing visualizations such as pairplots, correlation heatmaps, histograms, and scatter plots are all steps in the process. MDS is used to compute brand dissimilarities and reduce dimensionality to a 2D space for visual representation after the data has been prepared. The plot's conclusion demonstrates brand ties.

Objectives :

- Conduct data preprocessing to prepare the dataset for analysis.
- Perform exploratory data analysis (EDA) to uncover initial insights and patterns within the data.
- Apply multidimensional scaling (MDS) to reduce dimensionality and visualize the similarities or dissimilarities in the data.
- Create visualizations and interpret the results to gain meaningful insights and inform decision-making.

Business Significance :

There are important business ramifications when using Multidimensional Scaling (MDS) to examine the ice cream properties in the "icecream.csv" dataset. Businesses can obtain important insights about consumer preferences and market positioning by visualizing and analyzing the proximity of various ice cream brands based on characteristics like Price, Taste,

and Flavour. Strategic product differentiation and the development of marketing tactics catered to specific customer categories are made possible by an understanding of the relative distances between brands. For example, product development activities might be guided by insights from MDS to match consumer expectations and preferences with flavors or pricing strategies. Additionally, companies can improve their market segmentation strategies and customize promotional activities to successfully target particular customer categories by discovering clusters or groupings of products with similar attribute profiles. In the end, using MDS makes it easier to make knowledgeable decisions about marketing and product management, which improves competitiveness and responsiveness to changing market conditions.

Results and Interpretation using R

```
> # Display the dataset to confirm it's loaded correctly
> head(icecream)
```

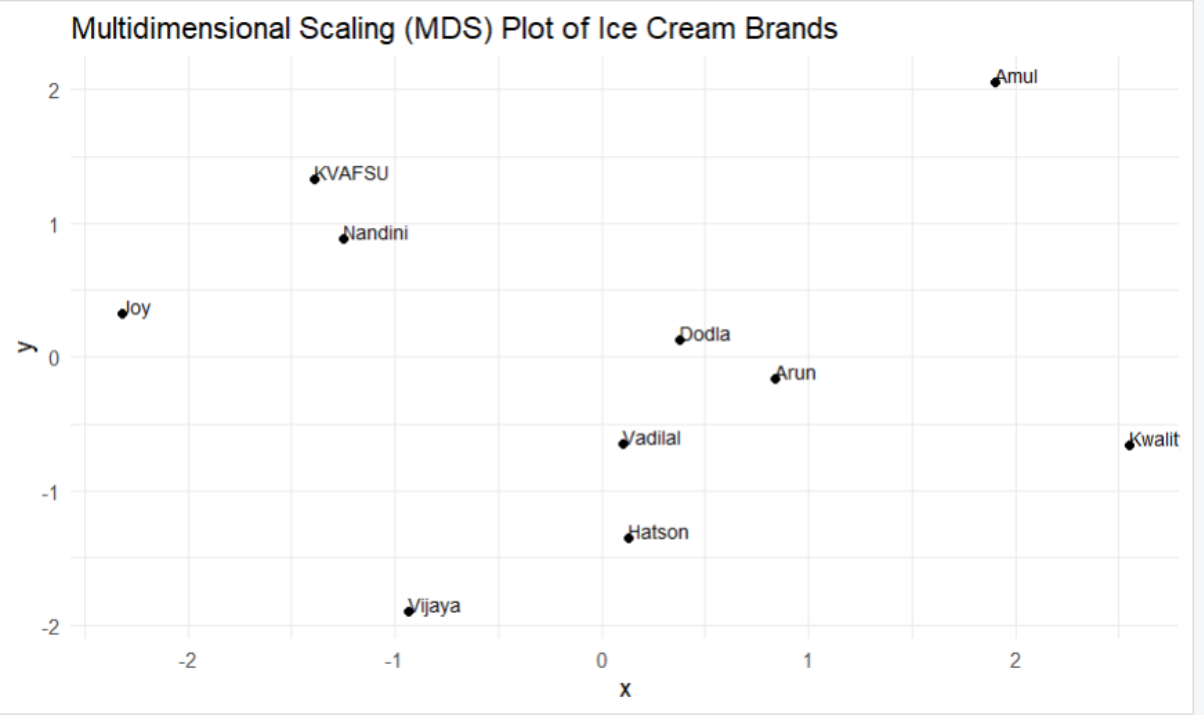
	Brand	Price	Availability	Taste	Flavour	Consistency	Shelflife
1	Amul	4	5	4	3	4	3
2	Nandini	3	2	3	2	3	3
3	Vadilal	2	2	4	3	4	4
4	Vijaya	3	1	3	5	3	4
5	Dodla	3	3	3	4	4	3
6	Hatson	2	2	4	4	3	4

Interpretation:

The dataset appears to contain information on various ice cream brands, evaluated across several attributes: Price, Availability, Taste, Flavour, Consistency, and Shelf Life. Each attribute seems to be rated on a scale, possibly from 1 to 5, with higher numbers indicating better ratings. For instance, the first row shows that the Amul brand has relatively high ratings across most attributes, especially Availability and Taste, while the Nandini brand has lower ratings for Availability and Flavour. This dataset could be used to compare and analyze the strengths and weaknesses of different ice cream brands based on these characteristics.

```
> ggplot(plot_data, aes(x, y, label = brand)) +
+   geom_point() +
```

```
+   geom_text(size = 3, hjust = 0, vjust = 0) +
+   labs(title = "Multidimensional Scaling (MDS) Plot of Ice Cream Brands"
) +
+   theme_minimal()
```



Interpretation:

The Multidimensional Scaling (MDS) plot of ice cream brands displays the relative similarities and dissimilarities among different brands based on multiple attributes. Brands positioned closer together on the plot are more similar to each other in terms of the evaluated attributes, while brands further apart are more dissimilar. For instance, Amul and Kwality are positioned on the far right, indicating they share similarities and potentially high ratings across multiple attributes. In contrast, brands like Vijaya and Hatson are located further apart from these brands, suggesting notable differences in their attribute ratings. This visualization helps in understanding the relative positioning and competitive landscape of various ice cream brands based on the evaluated criteria.

Results and Interpretation using Python

```
# Display the dataset to confirm it's loaded correctly
print(icecream.head())
```

	Brand	Price	Availability	Taste	Flavour	Consistency	Shelflife
0	Amul	4	5	4	3	4	3
1	Nandini	3	2	3	2	3	3
2	Vadilal	2	2	4	3	4	4
3	Vijaya	3	1	3	5	3	4
4	Dodla	3	3	3	4	4	3

Interpretation:

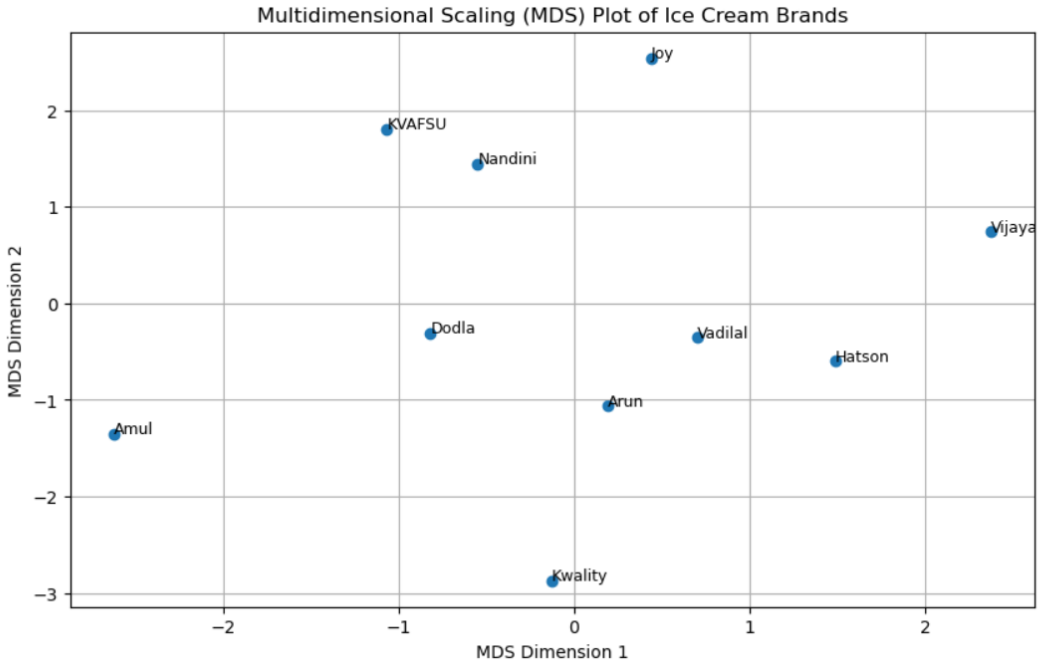
The dataset provided contains ratings for five ice cream brands—Amul, Nandini, Vadilal, Vijaya, and Dodla—across six attributes: Price, Availability, Taste, Flavour, Consistency, and Shelf Life. Each attri bute is rated on a scale, likely from 1 to 5. Amul stands out with high ratings for Availability (5) and Ta ste (4), indicating strong market presence and favorable taste perception. Nandini, with lower ratings i n Availability (2) and Flavour (2), might struggle in those areas. Vadilal shows a consistent performan ce with solid ratings across most attributes. Vijaya excels in Flavour (5) but has low Availability (1). Do dla has balanced ratings with a notable score in Consistency (4). This dataset helps in comparing the performance and market positioning of these brands based on their attribute ratings.

```
# Create a DataFrame for the MDS results
plot_data = pd.DataFrame({
    'x': mds_results[:, 0],  # X-axis coordinates
    'y': mds_results[:, 1],  # Y-axis coordinates
    'brand': icecream['Brand']  # Brand names
})

# Plot the MDS results
plt.figure(figsize=(10, 6))
plt.scatter(plot_data['x'], plot_data['y'])

for i in range(plot_data.shape[0]):
    plt.text(plot_data['x'][i], plot_data['y'][i], plot_data['brand'][i], fontsize=9)

plt.title('Multidimensional Scaling (MDS) Plot of Ice Cream Brands')
plt.xlabel('MDS Dimension 1')
plt.ylabel('MDS Dimension 2')
plt.grid(True)
plt.show()
```



Interpretation:

The Multidimensional Scaling (MDS) plot of ice cream brands illustrates the relative similarities and dissimilarities among different brands based on multiple attributes. Brands closer to each other on the plot are more similar in terms of their attribute ratings, while those farther apart are more dissimilar. For example, Nandini and KVAFSU are positioned closely together, indicating they share similar attributes. In contrast, brands like Amul and Kwality are placed far from each other, suggesting significant differences in their ratings. Vijaya and Vadilal are also positioned at a distance from the majority of other brands, indicating unique attribute profiles. This visualization aids in understanding the competitive landscape and the relative positioning of these ice cream brands.

Recommendations

Based on the Multidimensional Scaling (MDS) analysis of ice cream brands, the paper suggests segmenting and focusing on particular consumer segments. In order to stand out in the market, brands that are grouped together should think about differentiating tactics like

distinctive tastes or creative product formulas. Marketing messaging can be modified and favorable brand qualities reinforced to strengthen market placement. When developing new products, market gaps and emerging trends should be taken into consideration as well as customer preferences. Opportunities for strategic alliances or acquisitions can be found by regularly observing changes in the competitive landscape and evaluating the brand's positioning against rivals. It's also advised to improve consumer engagement tactics using MDS insights. It is advised to regularly reevaluate consumer perceptions and brand positioning using MDS or related analytical tools in order to keep a competitive edge in the ever-evolving ice cream industry.

R Codes

```
# Load required libraries
library(MASS)
library(ggplot2)

# Load the dataset
icecream <- read.csv('C:\\Users\\sayas\\OneDrive\\New folder\\python projects\\icecre
am.csv')

# Display the dataset to confirm it's loaded correctly
head(icecream)

# Extract the attributes for MDS (excluding the Brand column)
icecream_mds <- icecream[, -1]

# Perform Multidimensional Scaling (MDS)
mds <- cmdscale(dist(icecream_mds))

# Plot the MDS results
plot_data <- data.frame(
```

```
  x = mds[, 1],  # X-axis coordinates
  y = mds[, 2],  # Y-axis coordinates
  brand = icecream$Brand  # Brand names
)

ggplot(plot_data, aes(x, y, label = brand)) +
  geom_point() +
  geom_text(size = 3, hjust = 0, vjust = 0) +
  labs(title = "Multidimensional Scaling (MDS) Plot of Ice Cream Brands") +
  theme_minimal()
```

Python Codes

```
# Import required libraries
import pandas as pd
import numpy as np
from scipy.spatial.distance import pdist, squareform
from sklearn.manifold import MDS
import matplotlib.pyplot as plt
# Load the dataset
icecream = pd.read_csv('C:\\Users\\sayas\\OneDrive\\New folder\\python projects\\ice
cream.csv')
# Display the dataset to confirm it's loaded correctly
print(icecream.head())
# Extract the attributes for MDS (excluding the Brand column)
icecream_mds = icecream.iloc[:, 1:]  # Assuming 'Brand' is the first column

# Calculate the distance matrix
dist_matrix = squareform(pdist(icecream_mds))

# Perform Multidimensional Scaling (MDS)
mds = MDS(n_components=2, dissimilarity='precomputed', random_state=42)
```

```
mds_results = mds.fit_transform(dist_matrix)
# Create a DataFrame for the MDS results
plot_data = pd.DataFrame({
    'x': mds_results[:, 0],  # X-axis coordinates
    'y': mds_results[:, 1],  # Y-axis coordinates
    'brand': icecream['Brand']  # Brand names
})

# Plot the MDS results
plt.figure(figsize=(10, 6))
plt.scatter(plot_data['x'], plot_data['y'])

for i in range(plot_data.shape[0]):
    plt.text(plot_data['x'][i], plot_data['y'][i], plot_data['brand'][i], fontsize=9)

plt.title('Multidimensional Scaling (MDS) Plot of Ice Cream Brands')
plt.xlabel('MDS Dimension 1')
plt.ylabel('MDS Dimension 2')
plt.grid(True)
plt.show()
```

References

1. www.github.com