

VIRGINIA COMMONWEALTH UNIVERSITY

Statistical Analysis and Modelling (SCMA 632)

**A5: Visualization - Perceptual Mapping for
Business**

ANJELINA SAJAN

V01107258

Date of Submission: 15-07-2024

CONTENTS

Sl. No.	TITLE	Page No.
I.	Introduction	3
II.	Objective	3
III.	Business Significance	4
IV.	Results and Interpretation	5
V.	Recommendations	12
VI.	Conclusion	13

I. INTRODUCTION

Perceptual mapping is a powerful visualization tool used in business analytics to depict the relative positioning of various entities or variables within a defined space. It allows businesses to understand patterns, trends, and relationships in their data, providing insights that can drive strategic decisions. In this assignment, we aim to utilize perceptual mapping techniques to analyze consumption data from the NSSO68 dataset for the state of Maharashtra. By plotting a histogram and a bar plot, we will illustrate the distribution and magnitude of total consumption across different districts, enabling us to identify patterns and outliers in consumption behavior. These visualizations will serve as a foundation for understanding regional consumption disparities and can inform policy decisions and targeted interventions.

Furthermore, we will enhance our analysis by mapping a chosen variable on the geographical map of Maharashtra, utilizing the NSSO68 dataset. This geospatial visualization will offer a clear, intuitive representation of how the selected variable varies across the state's districts, providing additional context to our consumption analysis. By integrating both statistical and geospatial visualizations, this assignment aims to deliver a comprehensive overview of consumption patterns in Maharashtra, thereby offering valuable insights for stakeholders looking to address regional disparities and optimize resource allocation.

II. OBJECTIVE

- **Plot Histogram of Total Consumption:**
Create a histogram to display the distribution of total consumption across different districts in Maharashtra.
- **Create Bar plot of Consumption per District:**
Generate a bar plot to visualize the total consumption in each district of Maharashtra, clearly labelling each district.
- **Geospatial Mapping:**
Plot a selected variable from the NSSO68 dataset on the map of Maharashtra to illustrate its distribution across the state's districts.

III. BUSINESS SIGNIFICANCE

- **Targeted Marketing Strategies:**

Understanding consumption patterns across different districts in Maharashtra allows businesses to tailor their marketing strategies. Companies can focus their advertising efforts on high-consumption areas to maximize impact and sales.

- **Resource Allocation:**

By identifying districts with higher or lower consumption levels, businesses can optimize their resource allocation. This ensures efficient distribution of products and services, reducing wastage and meeting demand effectively.

- **Product Development and Customization:**

Insights into regional consumption behavior help businesses develop and customize products to suit the specific needs and preferences of different districts. This can lead to higher customer satisfaction and loyalty.

- **Investment Decisions:**

Businesses can make informed investment decisions by analyzing consumption trends. Districts with rising consumption might be targeted for new store openings, expansion projects, or increased inventory.

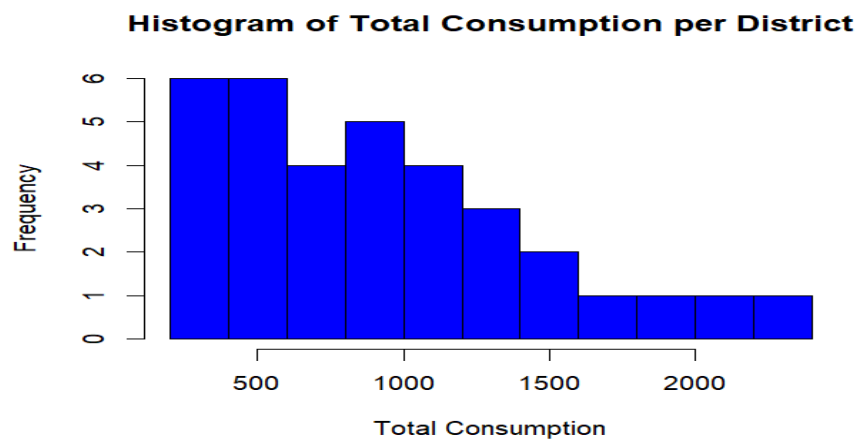
- **Policy and Planning:**

For policymakers, understanding the consumption patterns at a district level can guide the creation of targeted social and economic policies. This can help in addressing regional disparities and promoting balanced economic development across Maharashtra.

IV. DATA ANALYSIS AND INTERPRETATION

1. Histogram of Total Consumption

Analysis based on R Studio:



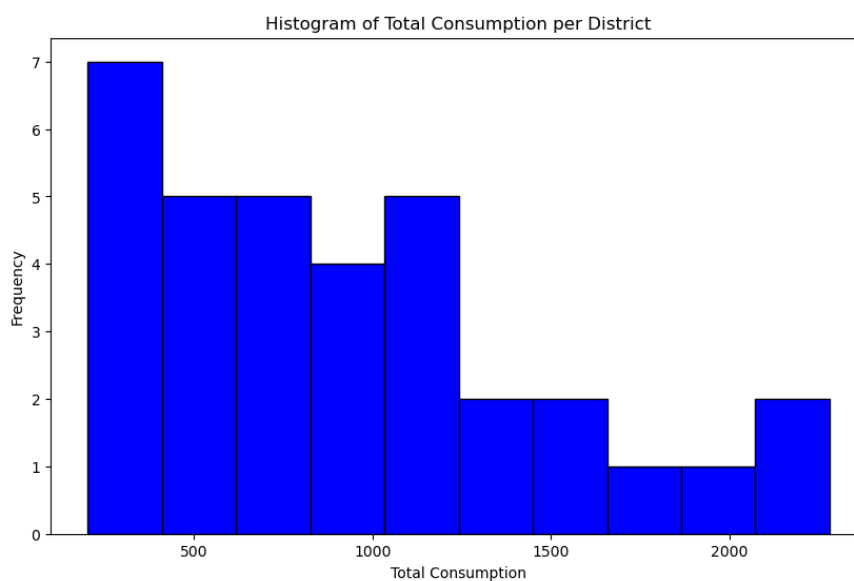
```
mhnew$District <- as.character(mhnew$District)
mhnew$Sector <- as.character(mhnew$Sector)
mhnew$District <- ifelse(mhnew$District %in% names(district_mapping), district_mapping[mhnew$District], mhnew$District)
mhnew$Sector <- ifelse(mhnew$Sector %in% names(sector_mapping), sector_mapping[mhnew$Sector], mhnew$Sector)

# Aggregate total consumption per district
MH_consumption <- aggregate(total_consumption ~ District, data = mhnew, sum)

# Data for histogram
consumption_values <- MH_consumption$total_consumption

# Histogram of total consumption values
hist(consumption_values, breaks = 10, col = 'blue', border = 'black',
      xlab = "Total Consumption", ylab = "Frequency", main = "Histogram of Total Consumption per District")
```

Analysis based on Python:



```
In [18]: # Data for histogram
consumption_values = MH_consumption['total_consumption']

# Histogram of total consumption values
plt.figure(figsize=(10, 6))
plt.hist(consumption_values, bins=10, color='blue', edgecolor='black')
plt.xlabel("Total Consumption")
plt.ylabel("Frequency")
plt.title("Histogram of Total Consumption per District")
plt.show()
```

INTERPRETATION:

The histogram shown represents the distribution of total consumption across different districts in Maharashtra, as provided in the NSSO68.csv data.

Distribution Shape

- Both histograms (Python and R) show a right-skewed distribution, indicating that most districts have lower total consumption, while a few districts have significantly higher consumption levels.

Key Observations

- **Low Consumption Districts:**
 - Districts like Gadchiroli (201.98), Gondiya (204.45), and Bhandara (212.41) have the lowest total consumption values.
 - These districts appear in the left-most bins of the histogram, contributing to the higher frequency of lower consumption values.
- **Moderate Consumption Districts:**
 - Districts like Hingoli (582.15), Dhule (598.31), and Latur (683.83) fall in the middle consumption range.
 - These districts are represented in the middle bins of the histogram, indicating a moderate frequency of districts with average consumption values.
- **High Consumption Districts:**
 - Mumbai (Suburban) (2281.49), Pune (2156.71), and Thane (1919.44) are among the districts with the highest total consumption values.
 - These districts contribute to the lower frequency bins on the right side of the histogram, highlighting the fewer number of districts with high consumption.

Combined Insights

- **Right Skewness:** The right-skewed nature of the histogram suggests that while most districts have lower total consumption, a few districts significantly drive up the average with their high consumption levels.
- **High Frequency of Low Consumption:** The majority of districts, such as Gadchiroli, Gondiya, and Bhandara, have low total consumption, which can be seen by the taller bars on the left side of the histogram.
- **Low Frequency of High Consumption:** A few districts, like Mumbai (Suburban), Pune, and Thane, have much higher total consumption, contributing to the shorter bars on the right side of the histogram.

Numerical Summary (Total Consumption per District)

- **Lowest Consumption:** Gadchiroli (201.98)
- **Moderate Consumption:** Dhule (598.31), Latur (683.83)

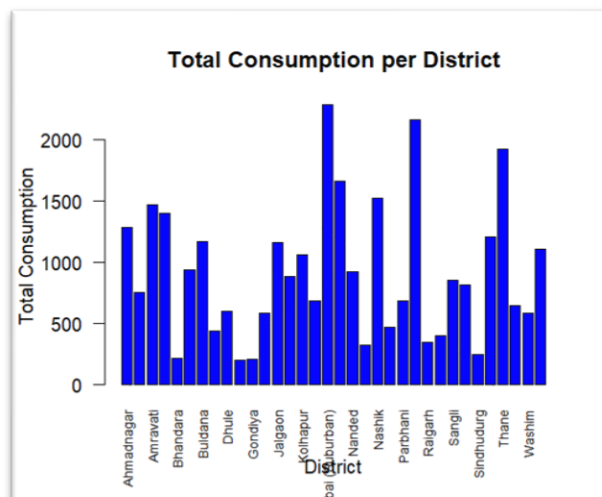
- **Highest Consumption:** Mumbai (Suburban) (2281.49), Pune (2156.71)

Conclusion

The histograms effectively depict the distribution of total consumption across districts in Maharashtra, illustrating a common pattern where a few districts account for a substantial portion of the total consumption. This skewness highlights potential areas for targeted economic interventions and resource allocation.

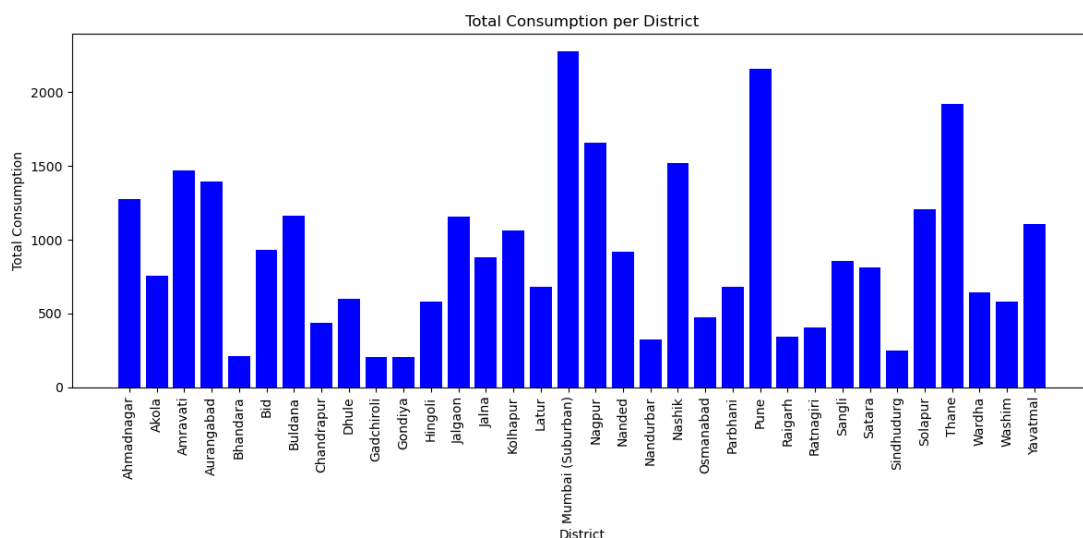
2. Bar plot of Consumption per District

Analysis based on R Studio:



```
# Additional Plot: Bar plot of total consumption per district
barplot(MH_consumption$total_consumption,
        names.arg = MH_consumption$District,
        las = 2, # Makes the district names vertical
        col = 'blue',
        border = 'black',
        xlab = "District",
        ylab = "Total Consumption",
        main = "Total Consumption per District",
        cex.names = 0.7) # Adjust the size of district names if needed
```

Analysis based on Python:



INTERPRETATION:

The bar plots provided illustrate the total consumption per district in Maharashtra using the NSSO68 dataset. The first plot was generated using Python, while the second was created using R. Despite the difference in tools, both visualizations convey the same information regarding consumption patterns across districts.

Key Observations:

1. High Consumption Districts:

- **Mumbai (Suburban):** Both plots indicate that Mumbai (Suburban) has the highest total consumption. This reflects the high population density, urbanization, and economic activity in the region.
- **Nagpur:** Similarly, Nagpur shows a notably high level of consumption, indicating its importance as a major urban center and economic hub in Maharashtra.
- **Pune:** Pune is another district with high consumption, emphasizing its status as a significant educational and industrial center.

2. Moderate Consumption Districts:

- **Kolhapur, Nashik, Thane:** These districts have moderate consumption levels. This suggests a balanced mix of urban and rural areas, with substantial economic activities.
- **Latur, Solapur:** These districts also show moderate consumption, possibly reflecting their agricultural and industrial activities.

3. Low Consumption Districts:

- **Bhandara, Gondiya, Hingoli:** These districts exhibit lower total consumption, which might be due to their smaller population size, less industrialization, or predominantly rural nature.
- **Ratnagiri, Sindhudurg:** These coastal districts have lower consumption, potentially reflecting less urbanization and more reliance on agriculture and fishing.

Analysis:

1. Urban vs. Rural Divide:

- Urban districts like Mumbai (Suburban), Nagpur, and Pune exhibit higher consumption due to larger populations, better infrastructure, and more economic activities.
- Rural districts, particularly those in the Vidarbha and Marathwada regions, show lower consumption, highlighting the disparity in economic development and population density.

2. Economic Activities:

- High consumption in districts like Mumbai (Suburban), Nagpur, and Pune is likely driven by diverse economic activities, including industry, services, and commerce.
- Districts with moderate consumption may have mixed economies with significant agricultural output alongside industrial activities.

3. Implications for Policy and Planning:

- The significant variation in consumption across districts suggests a need for targeted economic policies to address regional disparities.
- Enhancing infrastructure and economic opportunities in lower consumption districts could help balance economic development across the state.

Visualization Insights:

- **Python Plot:**

- Offers a wide aspect ratio which might make it easier to differentiate between districts with high and low consumption due to more space between bars.
- Suitable for detailed analysis when comparing a large number of categories (districts).

- **R Plot:**

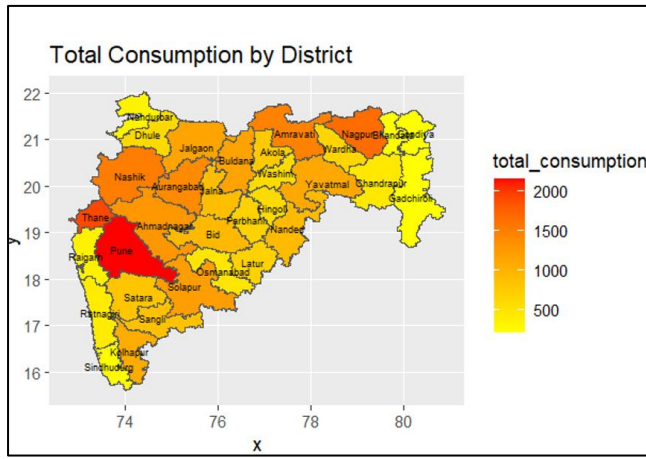
- Presents a more compact view, which could be beneficial for quick overviews.
- Useful for presentations where space is limited but a comprehensive view of all districts is required.

Conclusion:

Both plots effectively highlight the disparities in total consumption across various districts in Maharashtra. The high consumption in urban centers versus the lower consumption in rural areas underscores the economic divide within the state. These insights can guide policymakers in addressing regional imbalances and promoting inclusive economic growth.

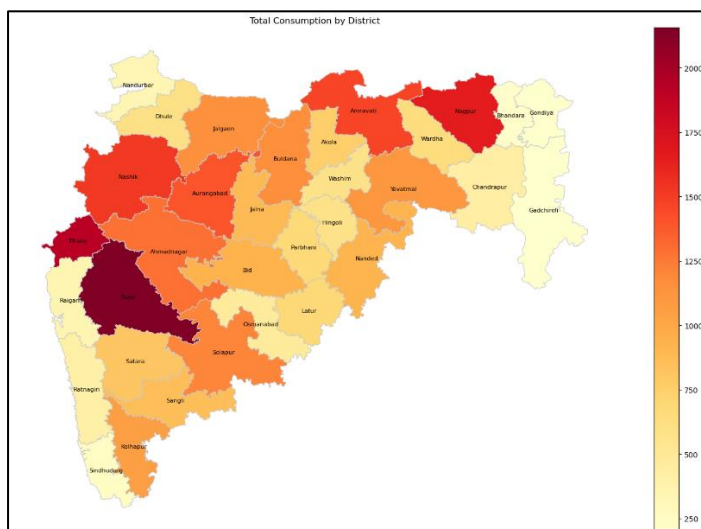
3. Geospatial Mapping

Analysis based on R Studio



```
# Plotting the map
ggplot(data_map_data) +
  geom_sf(aes(fill = total_consumption, geometry = geometry)) +
  scale_fill_gradient(low = "yellow", high = "red") +
  ggtitle("Total Consumption by District") +
  geom_sf_text(aes(label = District, geometry = geometry), size = 2, color = "black")
```

Analysis based on Python:



```
# Plotting the map
fig, ax = plt.subplots(1, 1, figsize=(15, 10))
data_map_data.plot(column='total_consumption', cmap='YlOrRd')

# Add colorbar
sm = plt.cm.ScalarMappable(cmap='YlOrRd', norm=plt.Normalize(
    sm._A = []
))
cbar = fig.colorbar(sm)

# Add labels
ax.set_title('Total Consumption by District')
for idx, row in data_map_data.iterrows():
    ax.annotate(text=row['District'], xy=(row.geometry.centroid.x, row.geometry.centroid.y))

plt.axis('off')
plt.tight_layout()
plt.show()
```

INTERPRETATION:

Both visualizations created using Python and R Studio provide insights into the distribution of total consumption across different districts in Maharashtra.

Key Observations:

1. High Consumption Areas:

- **Pune:** Marked in dark red in both visualizations, indicating it has the highest consumption.
- **Thane:** Also highlighted in dark shades, showing significant consumption levels.

- **Nagpur:** Appears in dark red in the Python visualization, indicating high consumption.
- 2. **Moderate Consumption Areas:**
 - **Nashik, Aurangabad, Ahmednagar, Amravati:** These districts show moderate levels of consumption in both visualizations, depicted in shades of orange.
 - **Solapur, Kolhapur:** Display moderate consumption levels, contributing to regional economic activity.
- 3. **Low Consumption Areas:**
 - **Gadchiroli, Sindhudurg, Ratnagiri:** Consistently marked in light yellow across both visualizations, indicating lower consumption levels.

Combined Analysis:

1. **Urban vs. Rural Divide:**
 - Urban districts like Pune, Thane, and Nagpur exhibit the highest consumption levels due to higher population densities, greater economic activity, and better infrastructure.
 - Rural districts, especially in the eastern part of Maharashtra, show lower consumption, which may be due to lower population density, less economic activity, and possibly higher levels of poverty.
2. **Economic Development:**
 - The high consumption areas reflect regions with better economic development, infrastructure, and access to markets.
 - Lower consumption areas indicate regions that might benefit from targeted economic development programs to improve infrastructure, access to goods and services, and overall economic activity.
3. **Policy Implications:**
 - Policymakers should focus on developing infrastructure and economic opportunities in lower consumption areas to boost overall consumption and economic growth.
 - Targeted programs to improve access to goods and services in rural and underdeveloped areas could help balance the consumption levels across districts.

Recommendations:

1. **Focused Economic Development:**
 - Implement targeted development programs in low-consumption districts to improve economic conditions and infrastructure.
 - Encourage investments in these regions to enhance economic opportunities and boost consumption.

2. Sustainable Urban Growth:

- In high-consumption urban centers, focus on sustainable urban planning to manage resources efficiently and support continuous growth.
- Address issues related to high consumption in urban areas, such as resource management and sustainable development.

3. Comprehensive Policy Approach:

- Combine economic, social, and infrastructural policies to ensure balanced regional development.
- Monitor the impact of implemented policies through periodic data analysis and adjust strategies as needed.

By leveraging insights from both visualizations, policymakers and stakeholders can make informed decisions to foster balanced economic growth and improve the overall well-being of Maharashtra's population.

V. RECOMMENDATIONS

• Targeted Economic Development in Low-Consumption Districts:

Implement specific development programs in districts with low consumption levels, such as Gadchiroli, Gondiya, and Bhandara. These programs should focus on improving infrastructure, economic opportunities, and access to basic services to boost local consumption and economic activity.

• Sustainable Urban Growth Management:

For high-consumption urban centers like Mumbai (Suburban), Pune, and Thane, prioritize sustainable urban planning. This includes efficient resource management, infrastructure development, and policies that support continuous, balanced growth while minimizing environmental impacts.

• Balanced Regional Policy Approaches:

Develop a comprehensive policy framework that combines economic, social, and infrastructural initiatives to ensure balanced regional development. Regularly monitor the impact of these policies through data analysis and make necessary adjustments to address emerging challenges and disparities.

• Investment in Agricultural and Industrial Diversification:

Encourage diversification in moderate-consumption districts such as Kolhapur, Nashik, and Thane by investing in both agricultural and industrial sectors. This can help

stabilize the local economy, increase employment opportunities, and enhance overall consumption levels.

- **Enhanced Data-Driven Decision Making:**

Utilize advanced data analytics and visualization techniques to continuously monitor consumption patterns and trends across districts. This can help policymakers and businesses make informed decisions, anticipate future needs, and proactively address regional economic disparities.

VI. CONCLUSION

The analysis of consumption data across different districts in Maharashtra using perceptual mapping techniques has provided valuable insights into regional disparities. The histogram and bar plot visualizations revealed a significant right-skewed distribution, indicating that while most districts exhibit lower consumption, a few districts, such as Mumbai (Suburban) and Pune, show significantly higher consumption levels. These patterns highlight the need for targeted economic interventions to address the imbalances and promote inclusive growth.

Geospatial mapping further emphasized the urban-rural divide, with urban centers like Pune, Thane, and Nagpur demonstrating high consumption due to better infrastructure and economic activities, whereas rural districts like Gadchiroli and Sindhudurg exhibited lower consumption. These insights suggest that enhancing infrastructure and economic opportunities in lower consumption districts can help in achieving balanced economic development. Additionally, focusing on sustainable urban planning in high-consumption areas can ensure the efficient management of resources and support ongoing growth.

In summary, the integration of statistical and geospatial visualizations in this analysis provides a comprehensive overview of consumption patterns in Maharashtra. Policymakers and businesses can leverage these insights to implement strategies that address regional disparities, optimize resource allocation, and enhance economic development across the state. By focusing on targeted development programs, customized marketing strategies, and sustainable planning, Maharashtra can achieve more balanced and inclusive economic growth.