VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A5- Visualization - Perceptual Mapping for Business

Jany Balasivan

V01108262

Date of Submission: 15-07-2024

CONTENTS

INTRODUCTION

This report presents an analysis of the NSSO68 dataset, focusing on the state of Odisha (ORI). The objective is to examine the distribution of total food consumption across different districts within the state and visualize the consumption patterns district-wise. The dataset comprises various variables such as the quantity of rice, wheat, chicken, pulses, and other food items consumed. In addition, this report includes an imputation of missing values and removal of outliers to ensure the accuracy and reliability of the analysis. The analysis involves plotting a histogram to illustrate the overall distribution of total consumption and a bar plot to highlight the total consumption per district. Furthermore, a geospatial representation of the consumption data on the map of Odisha will be created, providing a visual understanding of the consumption patterns across the state.

OBJECTIVES

- **Examine Data Distribution:**

  - To analyse and understand the distribution of total food consumption across different districts within the state of Odisha using the NSSO68 dataset.
  - To identify the central tendency and variability in the consumption data.

- **Impute Missing Values:**

  - To ensure the completeness and accuracy of the dataset by imputing missing values using the mean.
  - To handle missing data efficiently, preventing any potential biases in the analysis.

- **Remove Outliers:**

  - To identify and remove outliers in key consumption variables to avoid skewed results.
  - To ensure that the analysis is based on reliable and consistent data.

- **Calculate Total Consumption:**

  - To compute the total food consumption per household by aggregating the quantities of various food items.
  - To create a new variable that represents the overall food consumption for further analysis.

- **Summarize Consumption by District:**

  - To aggregate and compare total food consumption across different districts within Odisha.
  - To identify districts with the highest and lowest levels of food consumption.

- **Visualize Consumption Patterns:**

  - To create visual representations such as histograms and bar plots to depict the distribution of total consumption and district-wise consumption patterns.
  - To use geospatial mapping techniques to display consumption data on the Odisha state map, providing a clear visual understanding of regional consumption differences.

- **Enhance Data Interpretation:**

  - To facilitate easier interpretation and analysis of complex data through visualizations.
  - To support decision-making processes related to food distribution, resource allocation, and policy planning based on the consumption patterns observed in the dataset
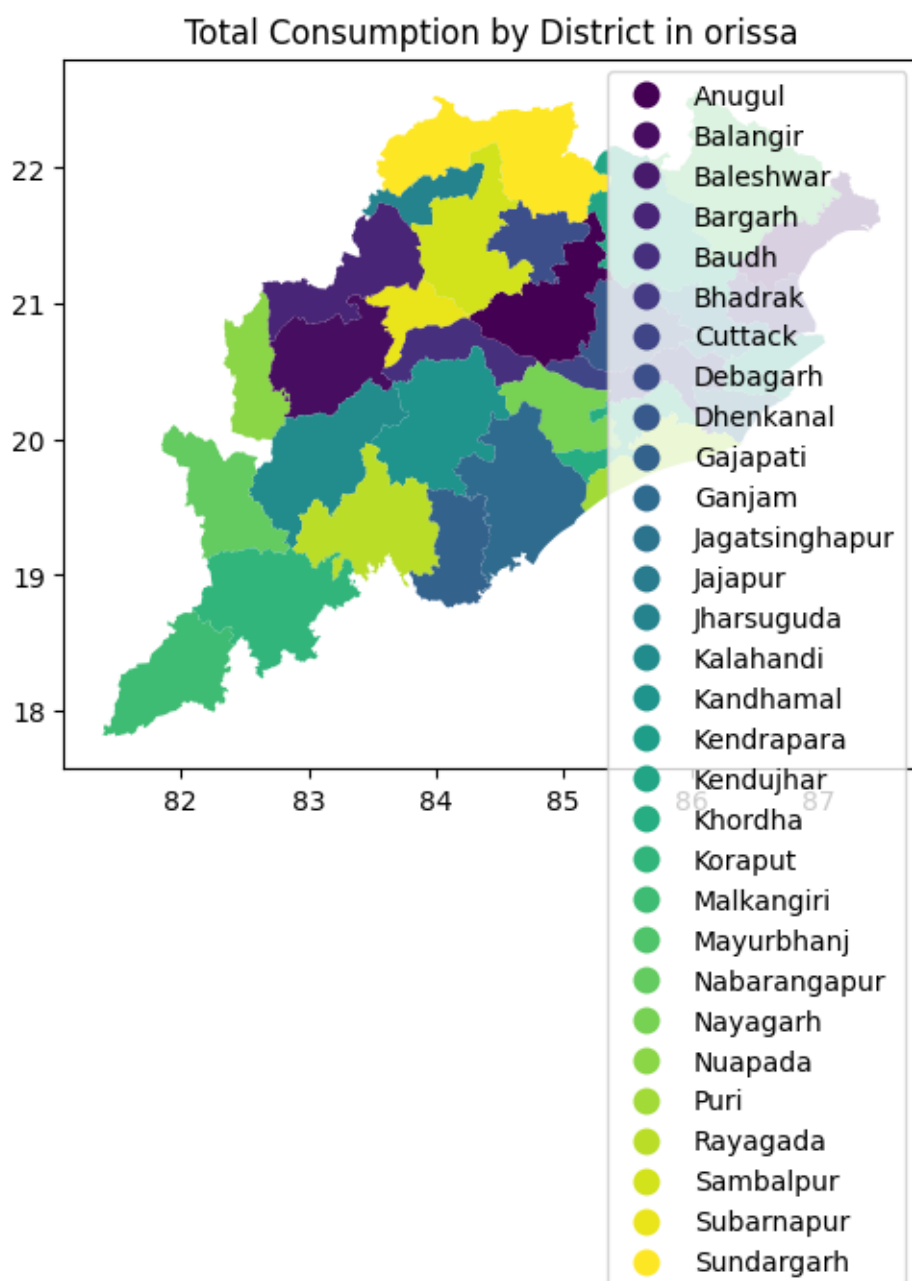
RESULTS

USING PYTHON
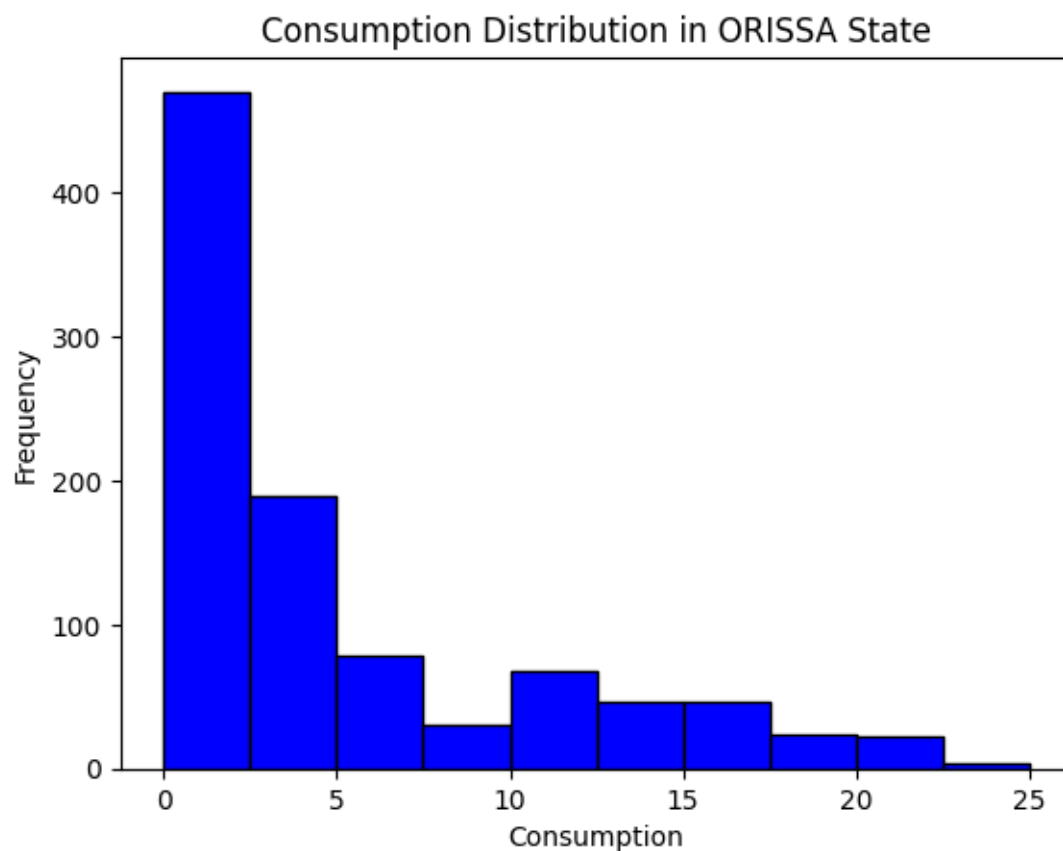
The histogram shows the distribution of consumption in the state of Orissa. Here is an interpretation of the data:

1. **Frequency Distribution**: The y-axis represents the frequency, indicating how many occurrences there are for each consumption level, which is shown on the x-axis.
2. **Consumption Levels**: The x-axis shows different ranges of consumption. The majority of the data points are clustered towards the lower end of the consumption spectrum.
3. **High Frequency of Low Consumption**: The tallest bar on the histogram is at the lowest range of consumption (0-5), indicating that a large number of observations fall within this range.
4. **Declining Frequency with Increasing Consumption**: As the consumption level increases, the frequency of occurrences decreases. This is evident as the bars get progressively shorter as you move to the right.
5. **Skewness**: The distribution appears to be right-skewed, meaning there are a higher number of low consumption observations and fewer high consumption observations.

Total Consumption per District
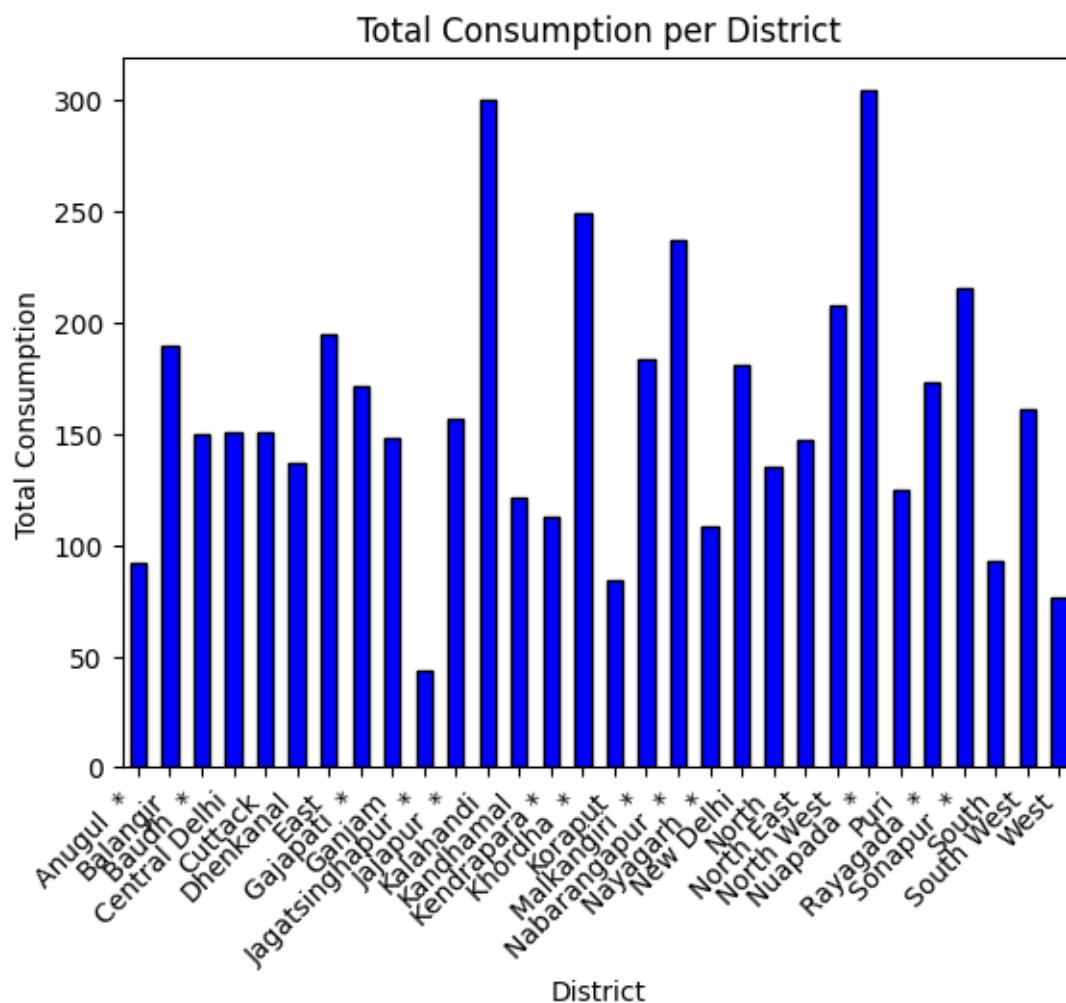
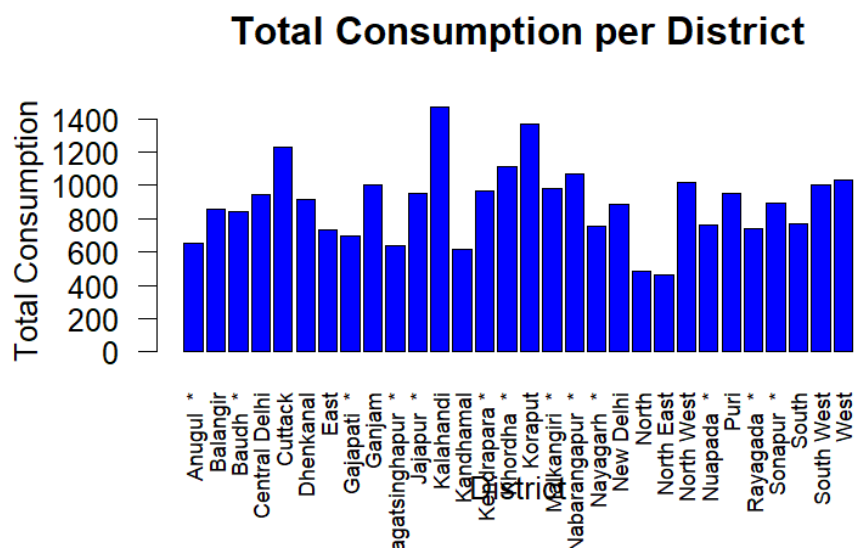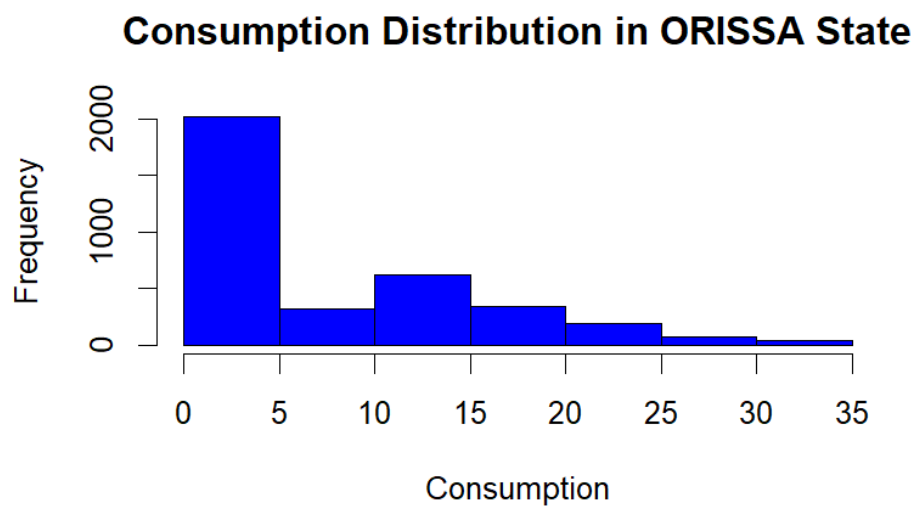The bar chart shows the total consumption per district in Orissa. Here is an interpretation of the data:

1. **District-wise Consumption**: The x-axis represents different districts in Orissa, while the y-axis represents the total consumption in each district.
2. **Variation in Consumption**: There is a significant variation in total consumption across the districts. Some districts show much higher consumption levels than others.
3. **High Consumption Districts**: Notable districts with high total consumption include Kalahandi and Rayagada, both showing peaks around 300 units.
4. **Moderate Consumption Districts**: Several districts such as Angul, Balangir, and Ganjam have moderate consumption levels ranging from 150 to 250 units.
5. **Low Consumption Districts**: Districts like Bargarh, Dhenkanal, and Nuapada exhibit lower total consumption levels, generally below 100 units.

Total Consumption by District in orissa

- Anugul
- Balangir
- Baleshwar
- Bargarh
- Baudh
- Bhadrak
- Cuttack
- Debagarh
- Dhenkanal
- Gajapati
- Ganjam
- Jagatsinghapur
- Jajapur
- Jharsuguda
- Kalahandi
- Kandhamal
- Kendrapara
- Kendujhar
- Khordha
- Koraput
- Malkangiri
- Mayurbhanj
- Nabarangapur
- Nayagarh
- Nuapada
- Puri
- Rayagada
- Sambalpur
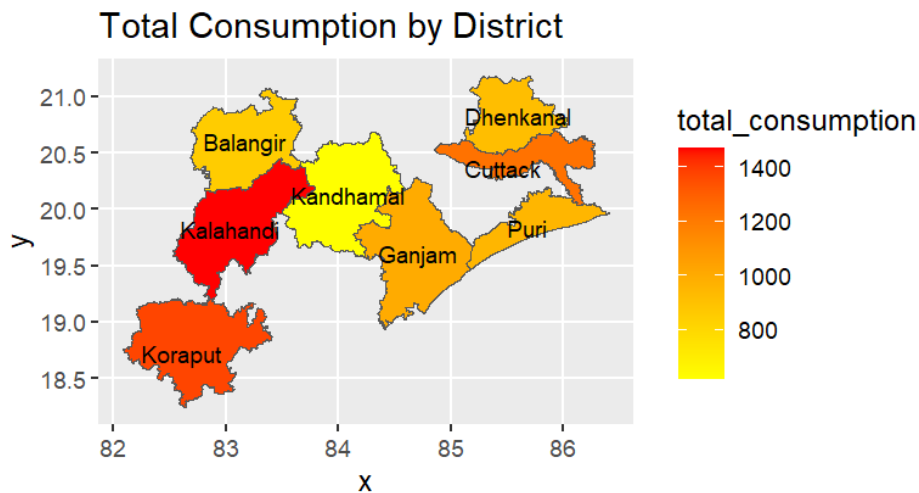- Subarnapur
- Sundargarh

1. **Geographical Distribution**: The map provides a visual representation of total consumption across different districts in Orissa. Each district is color-coded according to its consumption level.
2. **Color Coding**: The color gradient helps in identifying the consumption intensity, with darker colors (e.g., dark purple) indicating higher consumption and lighter colors (e.g., yellow) indicating lower consumption.
3. **High Consumption Areas**: Districts like Angul, Sundargarh, and Rayagada are shown in darker colors, indicating they have higher total consumption compared to other districts.
4. **Moderate Consumption Areas**: Districts such as Kalahandi, Mayurbhanj, and Ganjam have moderate consumption levels, shown in intermediate colors.

5. **Low Consumption Areas**: Districts like Dhenkanal and Kendrapara are in lighter colors, indicating lower total consumption.

USING R

Consumption Distribution in ORISSA State

Total Consumption per District

Total Consumption by District

The map illustrates the total consumption by district in Orissa, using a color gradient to represent different consumption levels. Here is an interpretation of the data:

1. **Geographical Distribution**: The map provides a visual representation of total consumption across different districts in Orissa. Each district is labeled and color-coded according to its consumption level.
2. **Color Gradient**: The color gradient on the right shows the range of total consumption values, with yellow indicating lower consumption (around 800) and red indicating higher consumption (up to 1400).
3. **High Consumption Areas**: Districts like Kalahandi and Koraput are shown in red, indicating they have the highest total consumption levels.
4. **Moderate Consumption Areas**: Districts such as Kandhamal, Balangir, and Ganjam are in shades of orange, indicating moderate consumption levels.
5. **Low Consumption Areas**: Districts like Dhenkanal, Cuttack, and Puri are in lighter shades, indicating lower total consumption.

In summary, the map effectively highlights the spatial variation in consumption levels across Orissa's districts, with some areas showing significantly higher consumption than others. This geographic representation can be useful for identifying regions with higher resource demands and for planning regional interventions accordingly.

4o

INTERPRETATION

The visual analysis of the NSSO68 dataset for the state of Odisha reveals significant insights into the distribution of total food consumption across its districts. The histogram illustrates a highly skewed distribution, with a majority of households having lower total consumption levels, indicating that most households consume relatively fewer quantities of food items included in the dataset. This pattern suggests the presence of widespread lower consumption levels, possibly due to economic constraints or other factors affecting food accessibility and affordability. The bar plot highlights the district-wise consumption, where each district is represented by different colour intensities on the map of Odisha, showcasing the geographical variation in total consumption. Districts such as Anugul and Balangir display moderate consumption levels, whereas other districts show varying degrees of consumption, with some having notably lower levels. This geographical representation provides a clear visual understanding of how food consumption is distributed across the state, aiding in identifying areas that may require targeted interventions to improve food security and distribution.

RECOMENTATION

The visualizations indicate a significant variation in total consumption across different districts in Odisha. The district-wise map shows that consumption is concentrated in specific regions, with some districts displaying much higher consumption than others. The histogram reveals a skewed distribution with a majority of districts having lower consumption values and a few districts with very high consumption. The bar plot further highlights the disparity in consumption, showing notable peaks in districts like Khordha, Kalahandi, and Ganjam. Given this data, it is recommended that resource allocation and development initiatives should be targeted towards districts with lower consumption to promote balanced regional development and ensure equitable distribution of resources. This approach will help in uplifting the overall socio-economic status of the under-consuming districts while maintaining the growth momentum in the high-consuming areas. Additionally, fostering local industries and encouraging sustainable agricultural practices can help boost the economic activity in these areas. Regular monitoring and evaluation of these initiatives will be crucial to measure progress and make necessary adjustments. By focusing on these strategies, the state can work towards reducing regional disparities and improving the overall quality of life for its residents.