VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A5 : Visualisation – Perceptual Mapping for Business

VARUN YOGANANDA

V01107265

Date of Submission: 15-07-2024

CONTENTS

INTRODUCTION

1.1. Histogram and Barplot to indicate the consumption district-wise for Assam

The focus of this study is on the state of Assam, using data from the National Sample Survey Office (NSSO) to analyze district-wise consumption patterns. We aim to visualize the distribution of total consumption across districts with a histogram and provide a detailed view of consumption per district using a Bar-plot. The NSSO68 dataset includes comprehensive consumption-related data for both rural and urban sectors.

The analysis involves handling missing values, identifying and removing outliers, and standardizing district and sector names. By summarizing the consumption data regionally and district-wise, we aim to provide valuable insights into consumption patterns within Assam. These visualizations will help policymakers and stakeholders understand how consumption varies across different districts, facilitating targeted interventions and promoting equitable development across the state.

1.2. Plotting the variable 'Non-total_v' on Karnataka state map

The focus of this study is on visualizing meat consumption patterns across the state of Karnataka using data from the National Sample Survey Office (NSSO). We aim to plot the variable 'Non-total_v,' which indicates total meat consumption, on the Karnataka state map using the NSSO68 dataset. This visualization will help illustrate the distribution and intensity of meat consumption across different districts in Karnataka.

The NSSO68 dataset provides comprehensive data on various consumption-related aspects for both rural and urban sectors. By mapping the 'nontotal_v' variable, we aim to identify and highlight areas with high and low meat consumption. This analysis will provide valuable insights for policymakers and stakeholders, enabling them to understand regional consumption patterns better and make informed decisions for targeted interventions and resource allocation in Karnataka.

OBJECTIVES

2.1 Histogram and Bar-plot to indicate the consumption district-wise for Assam

- Visualize Consumption Distribution
- Detailed District-wise Analysis
- Identify Consumption Patterns

2.2 Plotting the variable 'Non-total_v' on Karnataka state map Identify Distinct Respondent Groups

2.2.1 Visualize Meat Consumption

2.2.2 Identify Regional Consumption Patterns

BUSINESS SIGNIFICANCE

3.1 Histogram and Barplot Indicating Consumption District-wise for Assam

Creating histograms and barplots to showcase consumption patterns across districts in Assam holds significant business implications. These visualizations provide critical insights for resource allocation, enabling efficient planning of food supplies and infrastructure development based on district-specific consumption needs. Businesses can leverage this data to tailor marketing strategies and product offerings, effectively targeting consumer preferences across different districts. Additionally, understanding consumption disparities supports optimized supply chain management, ensuring products are distributed efficiently to meet varying demand levels. Policymakers can benefit by formulating targeted interventions aimed at promoting balanced consumption and improving public health outcomes. These visualizations also facilitate competitive analysis, helping businesses identify market opportunities and gaps for strategic market entry and positioning. Overall, analyzing consumption patterns aids in socioeconomic impact assessment and supports initiatives aimed at fostering healthier eating habits and lifestyle choices across Assam.

3.2 Plotting the Variable 'Nontotal_v' on the Karnataka State Map

Mapping the variable 'nontotal_v' on Karnataka's state map carries substantial business significance. It provides valuable insights into regional consumption patterns, aiding businesses in strategic decision-making and resource allocation. By visualizing meat consumption on a map, businesses can identify areas with high demand and adjust supply chains accordingly, ensuring efficient distribution and minimizing logistical costs. This data also supports market segmentation efforts, enabling businesses to tailor marketing strategies and product offerings based on district-specific preferences. Moreover, the map helps policymakers assess dietary trends and health implications, guiding initiatives to promote balanced nutrition and public health awareness. Overall, the geo map task facilitates informed business strategies, enhances market competitiveness, and supports efforts towards sustainable development and consumer welfare in Karnataka.

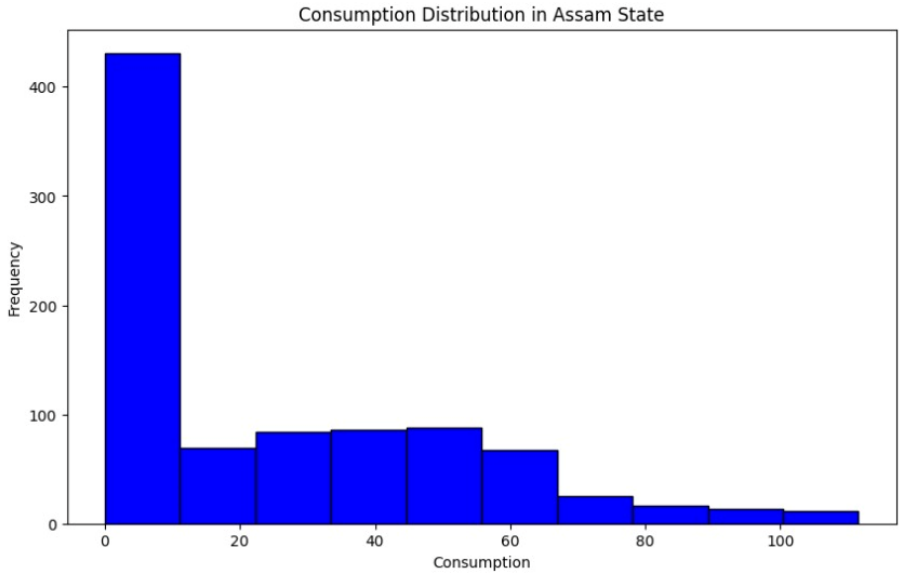
RESULTS AND INTERPRETATIONS

4.1 Histogram and Barplot to indicate the consumption district-wise for Gujarat

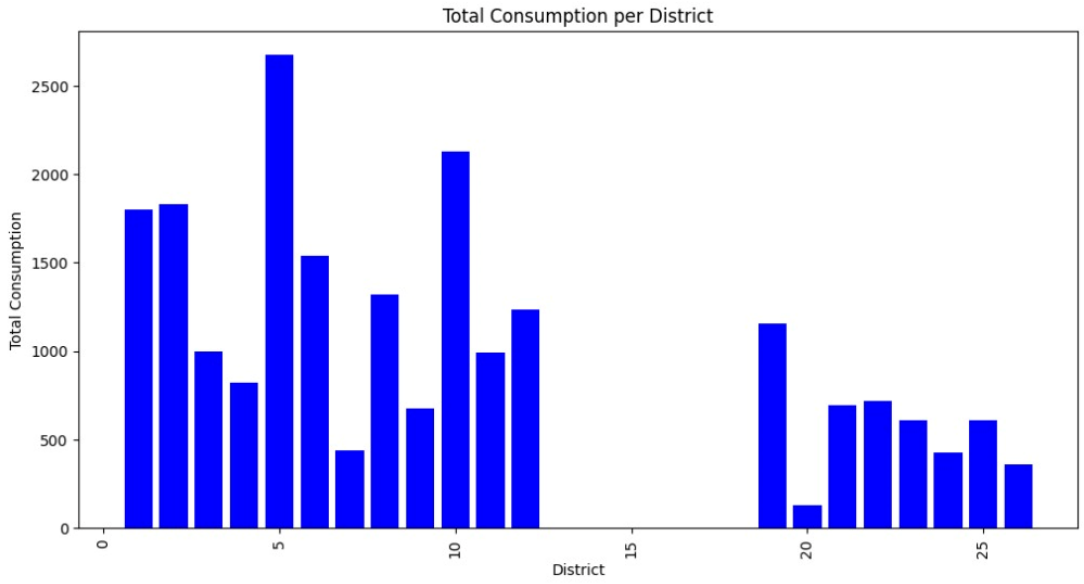
Python CODES:

```
Index(['slno', 'grp', 'Round_Centre', 'FSU_number', 'Round', 'Schedule_Number',
       'Sample', 'Sector', 'state', 'State_Region',
       ...
       'pickle_v', 'sauce_jam_v', 'Othrprocessed_v', 'Beveragestotal_v',
       'foodtotal_v', 'foodtotal_q', 'state_1', 'Region', 'fruits_df_tt_v',
       'fv_tot'],
      dtype='object', length=384)
    slno         grp  Round_Centre  FSU_number  Round  Schedule_Number  \
0  45494  7.100000e+31             1       70974     68               10
1  45495  7.100000e+31             1       70974     68               10
2  45496  7.100000e+31             1       70974     68               10
3  45497  7.100000e+31             1       70974     68               10
4  45498  7.100000e+31             1       70974     68               10
```

```
   Sample  Sector  state  State_Region  ...  pickle_v  sauce_jam_v  \
0       1       1     33           334  ...   0.00000          0.0
1       1       1     33           334  ...   0.00500          0.0
2       1       1     33           334  ...   0.00500          0.0
3       1       1     33           334  ...   0.00375          0.0
4       1       1     33           334  ...   0.00000          0.0

   Othrprocessed_v  Beveragestotal_v  foodtotal_v  foodtotal_q  state_1  \
0              0.0         32.500000  1021.508000    55.553185       TN
1              0.0          6.671667   316.547667    21.170447       TN
2              0.0          6.805000   319.173000    16.952281       TN
3              0.0          7.503750   299.116250    19.945296       TN
4              0.0         25.000000   546.134250    26.125330       TN

   Region  fruits_df_tt_v      fv_tot
0       4          116.50  266.090000
1       4            9.50   49.206667
2       4            3.60   36.224000
3       4            7.50   56.787500
4       4           55.25  133.295000

[5 rows x 384 columns]
(944, 384)
Missing Values Information:
slno             0
grp              0
Round_Centre     0
FSU_number       0
Round            0
                ..
foodtotal_q      0
state_1          0
Region           0
fruits_df_tt_v   0
fv_tot           0
Length: 384, dtype: int64
Top Consuming Districts:
   District  total_consumption
4         5        2676.124951
9        10        2129.976667
1         2        1830.343640
0         1        1803.854881
```



HISTOGRAM



BAR GRAPH

Total Consumption per District

- The bar chart depicts the total consumption across various districts.
- District 5 has the highest consumption, exceeding 2500 units.
- Districts 10, 20, and 25 also show significant consumption, though notably lower than District 5.
- Consumption varies widely across districts, indicating uneven distribution of food consumption.

Consumption Distribution in Assam State

- The histogram illustrates the frequency distribution of consumption levels within Assam State.
- A majority of the data points have low consumption values, clustered around the 0-20 range.
- There are fewer instances of higher consumption, with frequencies decreasing as consumption values increase.
- This indicates that most households or individuals in Assam have low consumption levels, with only a small fraction consuming larger amounts.

INTERPRETATIONS:

The dataset contains 944 rows and 384 columns, each representing various attributes related to food consumption and demographic data across different regions and states in India. Here's a detailed interpretation based on the information provided:

Data Structure and Attributes:

1. **Key Identifiers**:
   - `slno`, `grp`, `Round_Centre`, `FSU_number`, `Round`, `Schedule_Number`, and `Sample` are unique identifiers and administrative variables indicating the specific sample and survey round.
2. **Geographic and Demographic Information**:
   - Variables such as `Sector`, `state`, `State_Region`, `state_1`, and `Region` categorize the data by urban/rural sectors, state codes, and broader regional divisions.
3. **Food Consumption Metrics**:
   - Columns like `pickle_v`, `sauce_jam_v`, `Othrprocessed_v`, `Beveragestotal_v`, `foodtotal_v`, `foodtotal_q`, `fruits_df_tt_v`, and `fv_tot` detail the consumption volumes (in appropriate units) of various food items and categories.
   - `foodtotal_v` and `foodtotal_q` provide total food consumption in value and quantity respectively, indicating overall dietary intake.

Missing Values:

The dataset has no missing values across all 384 columns, ensuring a complete dataset for analysis.

Consumption Insights:

From the sample data provided:

- The total food consumption (`foodtotal_v`) varies significantly among different entries, indicating diverse dietary patterns.
- The top consuming districts based on total food consumption are:
  1. District 5 with a total consumption of 2676.12 units.
  2. District 10 with 2129.98 units.
  3. District 2 with 1830.34 units.
  4. District 1 with 1803.85 units.

Regional Patterns:

- The state labeled `TN` (likely Tamil Nadu) appears frequently, suggesting significant representation in the sample.

- The `Region` column value '4' is consistent across the sample data provided, possibly indicating a specific area within the state or a classification used in the study.

Food Item Specifics:

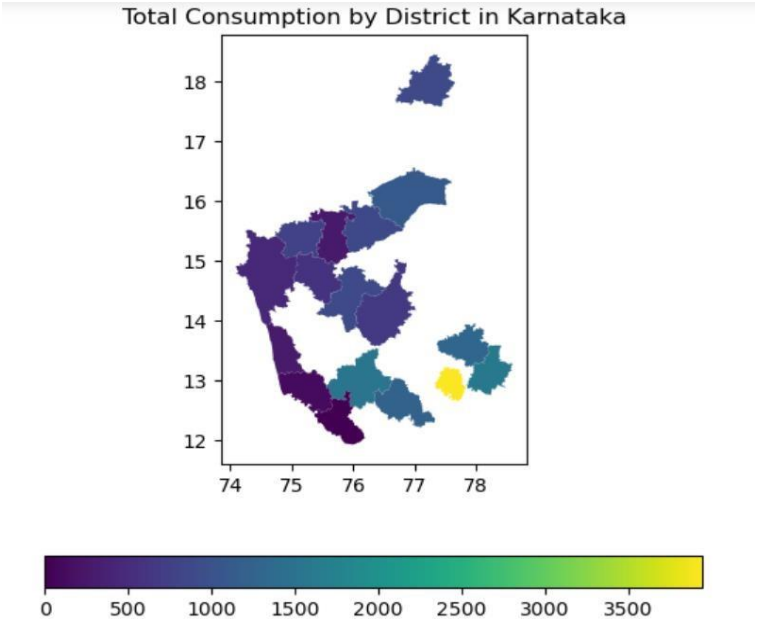
- Items like `pickle_v`, `sauce_jam_v`, and `Othrprocessed_v` have low or zero values, indicating lesser consumption of these processed items in the sample.
- `Beveragestotal_v` shows a wide range of values, reflecting diverse beverage consumption patterns.

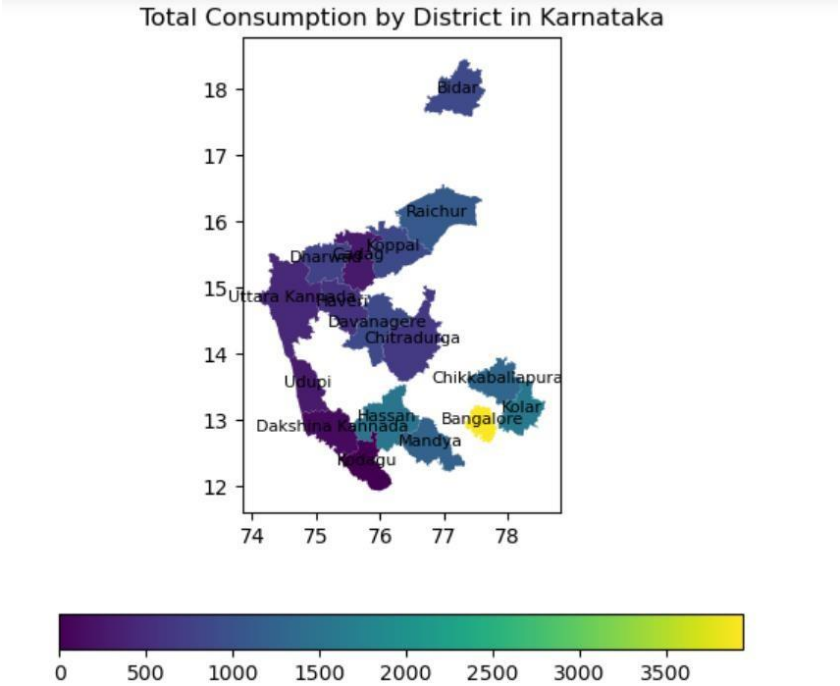
Overall:

The dataset provides a comprehensive view of food consumption across different regions and demographic segments in India, offering valuable insights for dietary studies, policy-making, and regional nutritional planning.

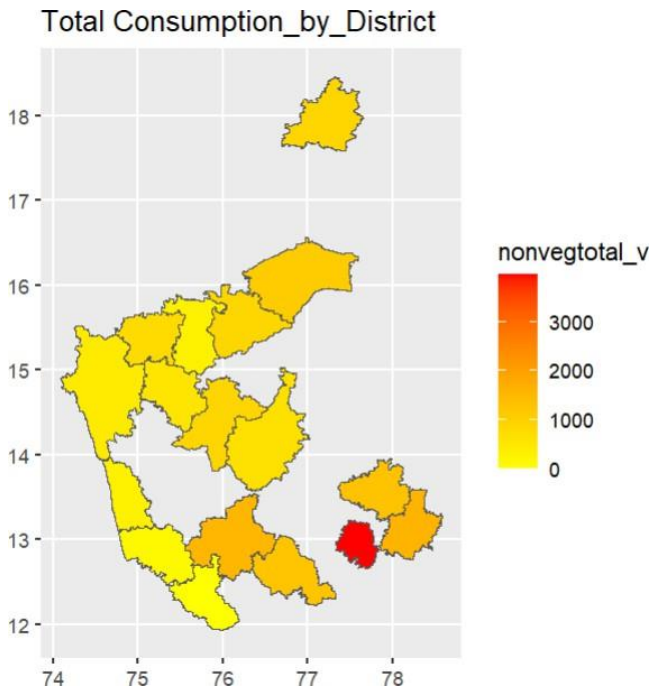
4.2 Plotting the variable ‘Non-total_v’ on Karnataka state map

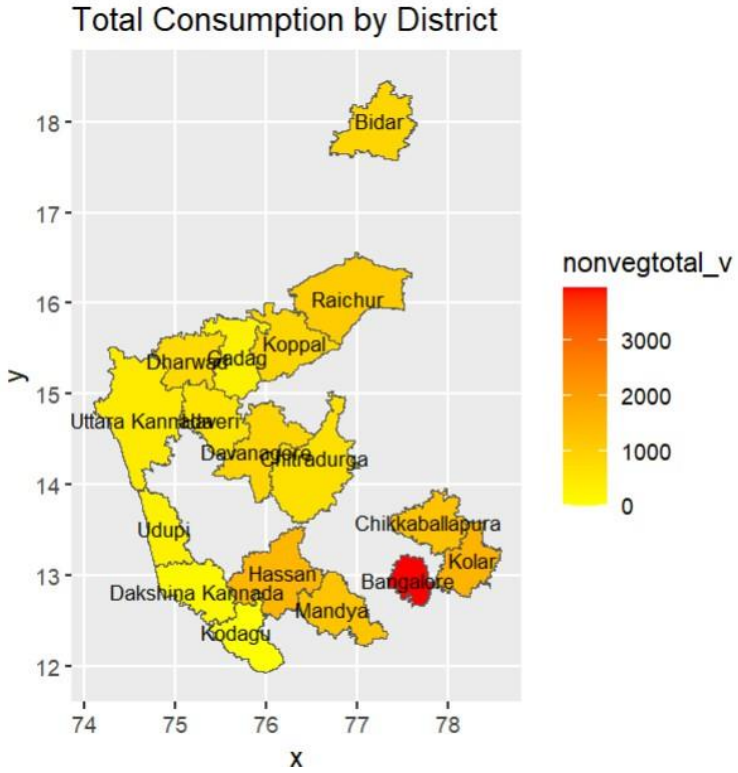
- Python





⇨ R


Total Consumption_by_District
nonvegtotal_v
3000
2000
1000
0
18
17
16
15
14
13
12
74
75
76
77
78


Total Consumption by District
nonvegtotal_v
3000
2000
1000
0
Bidar
Raichur
Koppal
Dharwad
Gadag
Uttara Kannada
Haveri
Davanagere
Chitradurga
Udupi
Chikkaballapura
Kolar
Hassan
Bangalore
Dakshina Kannada
Mandya
Kodagu
y
x
18
17
16
15
14
13
12
74
75
76
77
78

The output provided performs an analysis on consumption data by districts and regions in Karnataka, focusing on non-vegetarian food items such as eggs, fish/prawn, goat meat, beef, pork, chicken, and other birds. The analysis follows these key steps:

1. Removing Outliers: The `remove-outliers` function is applied to specified columns to eliminate extreme values that could skew the analysis. This helps in obtaining a more accurate representation of the data.

2. Summarizing Consumption: Total non-vegetarian consumption per district and region is calculated and summarized. This is achieved by summing up the consumption of various non-vegetarian items and grouping the data by district and region. The top-consuming districts and regions are then identified.

3. Mapping Districts and Sectors: A mapping dictionary is used to convert numerical district and sector codes to their respective names, making the data more readable and interpretable.

4. Geospatial Analysis: The code merges the consumption data with geographic data (a Geo JSON file of Karnataka districts). This merged dataset is then used to create a map that visually represents the total non-vegetarian consumption by district in Karnataka.

5. Visualization: Finally, the code plots a map with the consumption data, displaying the intensity of non-vegetarian consumption across different districts. The map provides a visual summary of the distribution of non-vegetarian food consumption, highlighting areas with higher or lower consumption levels.

This analysis helps in understanding the geographical distribution of non-vegetarian food consumption in Karnataka, identifying key districts and regions with significant consumption patterns.

Interpretation

The map of Karnataka illustrates varying levels of meat consumption ('Non-total_v') across its districts based on the NSSO68 dataset. Districts such as Bangalore, Hassan, and Kolar appear as regions with higher meat consumption, indicated by deeper shades on the map. These districts likely reflect urban centers and regions with higher population densities and economic activities, contributing to increased meat consumption. Conversely, districts like Dakshina Kannada and Kodagu exhibit lighter shades, indicating lower levels of meat consumption.