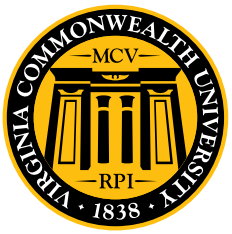


VIRGINIA COMMONWEALTH UNIVERSITY

Statistical analysis and modelling (SCMA 632)

A5: Visualization - Perceptual Mapping for Business

FERAH SHAN SHANAVAS RABIYA

V01101398

Date of Submission: 15-07-2024

CONTENTS

Introduction

The state of Kerala, located in the southern part of India, is known for its high literacy rates, advanced social development, and diverse cuisine. Understanding the consumption patterns of food across different districts of Kerala can provide valuable insights into the dietary habits, economic conditions, and social structures of the population. This report aims to analyze and visualize the food consumption data collected from the National Sample Survey Office (NSSO) 68th round survey. By examining the distribution and total consumption of food across various districts, we can identify trends, disparities, and potential areas for policy intervention.

The results from these visualizations will be analyzed to identify key patterns and trends in food consumption across Kerala. The discussion will focus on the implications of these findings for policy makers, local authorities, and other stakeholders involved in food security and public health in Kerala. By leveraging the NSSO68 dataset, this report aims to provide a detailed analysis of food consumption patterns across Kerala. The insights gained from this analysis will be valuable for developing targeted interventions to ensure equitable food distribution and improve nutritional outcomes across the state.

Objectives

- Conduct an analysis of the "NSSO68" dataset to discover socio-economic and demographic variables that have a significant impact.
- To present the distribution of total food consumption across different districts of Kerala using a histogram. This will help in understanding the overall spread and central tendencies of food consumption in the state.
- To create a barplot showcasing the total food consumption for each district, highlighting the differences in consumption levels across districts.
- To plot the total food consumption on the map of Kerala, providing a spatial understanding of the consumption patterns across the state.

Business Significance

Understanding food consumption patterns across different districts of Kerala is crucial for businesses in various sectors, including retail, agriculture, supply chain management, and health. This data can help identify high demand areas, optimize supply chain logistics, and develop customized retail strategies. By analyzing the total food consumption across districts, businesses can tailor marketing strategies and optimize inventory management to meet local demand efficiently.

The geographical distribution of food consumption allows businesses to plan efficient routes and allocate resources effectively, ensuring timely delivery and minimizing transportation costs. Knowing which districts have higher consumption rates reduces the risk of overstocking or stockouts. Retail expansion can be achieved by strategically selecting strategic locations and developing customized retail strategies based on consumption patterns.

Agricultural planning and support can also benefit from this data, as it allows farmers and agricultural businesses to plan crop production based on demand patterns in different districts, reducing waste and improving profitability. Support services, such as input suppliers and agritech companies, can tailor their offerings to meet the specific needs of farmers in high-consumption districts, providing targeted solutions to enhance productivity.

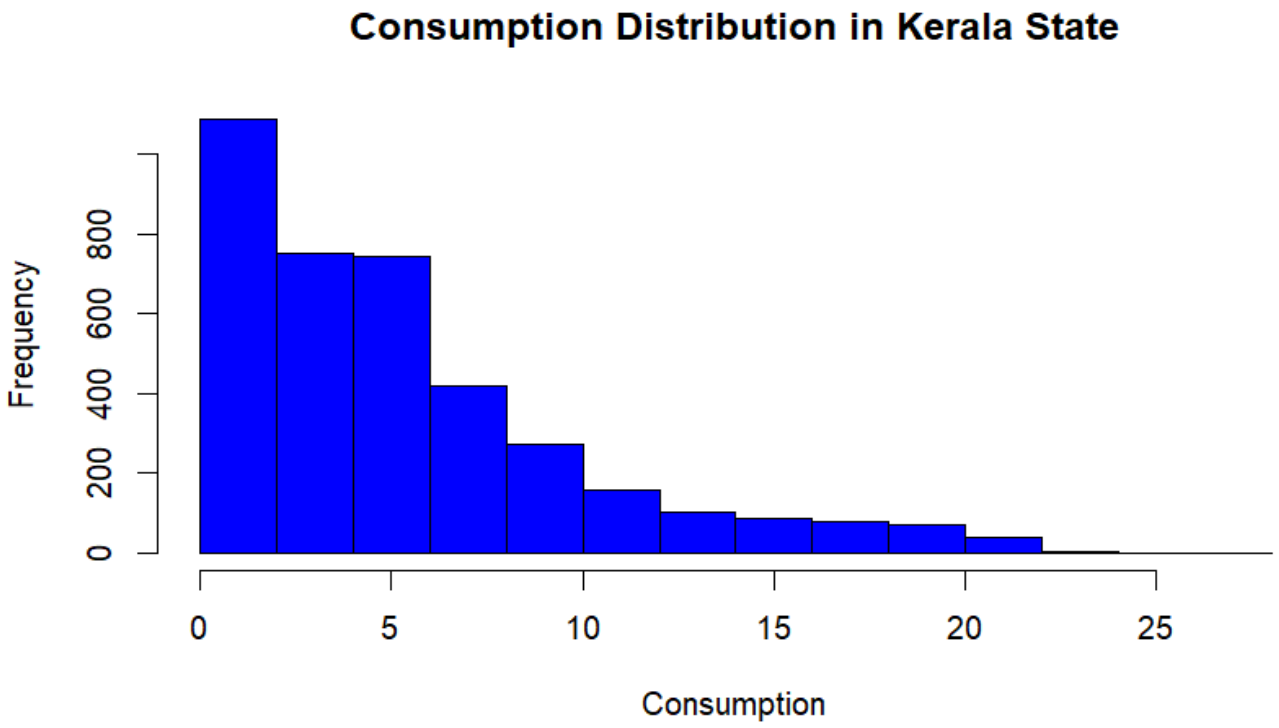
Policy and investment decisions can be made using the consumption data to identify areas with food security issues and design targeted interventions. Investors can identify high-potential areas for investing in food production, processing, and retail businesses, making informed decisions about where to allocate capital for maximum returns.

Health and nutrition programs can be developed using the data, with health organizations and NGOs developing nutritional programs aimed at improving dietary habits in districts with low consumption or poor dietary diversity. Companies producing health and wellness products can target marketing efforts towards districts with lower consumption of nutritious foods, promoting supplements and health foods to bridge nutritional gaps.

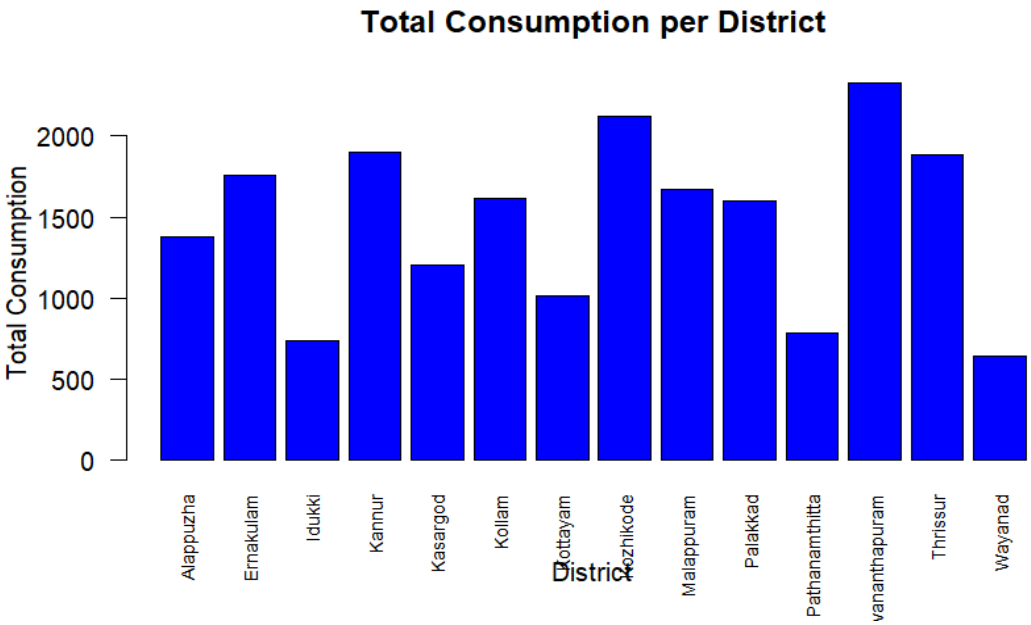
Results and Interpretation using R

- **Plotting a histogram depicting consumption distribution in Kerala and a bar graph showing total consumption per district.**

```
> hist(klnew$total_consumption, breaks = 10, col = 'blue', border = 'bl
ack',
+      xlab = "Consumption", ylab = "Frequency", main = "Consumption Di
stribution in Kerala State")
>
> klnew_consumption <- aggregate(total_consumption ~ District, data = k
lnew, sum)
> View(klnew_consumption)
```




```
> barplot(klnew_consumption$total_consumption,
+         names.arg = klnew_consumption$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)
```



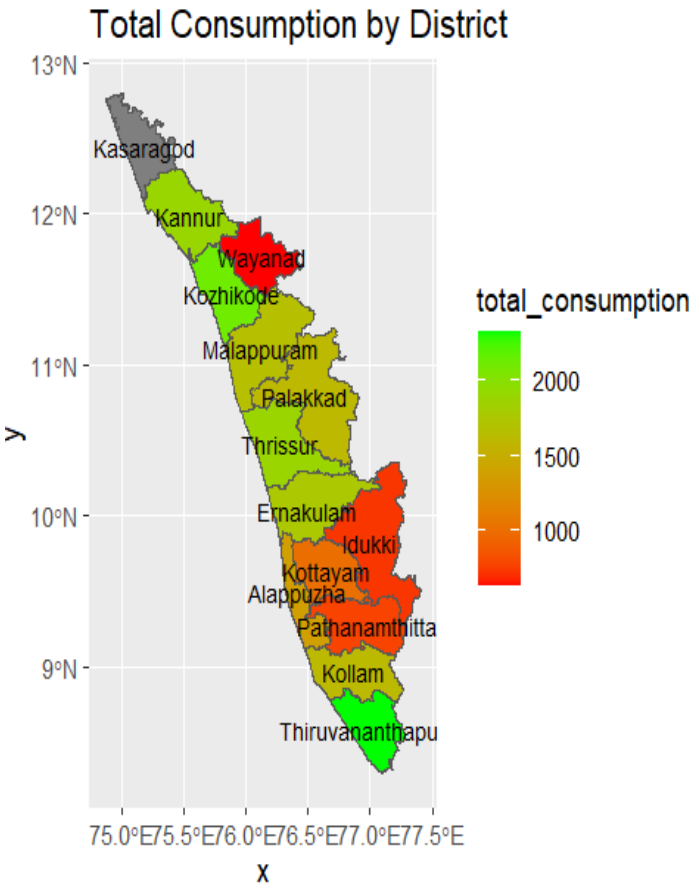
Interpretation:

The histogram in Kerala illustrates the distribution of total food consumption across different districts. It shows the central tendency of food consumption, indicating where most districts li e in terms of total consumption. The spread and skewness of the bars provide insights into the variability of food consumption across districts. If the histogram is skewed to the left or right, it suggests districts with significantly higher or lower consumption than the average.

The bar graph provides a detailed view of total food consumption in each district, listing the d istricts vertically. The Y-Axis shows the total amount of food consumed in each district, with each bar representing the total consumption for a specific district. High consumption districts have taller bars, possibly due to larger populations or higher per capita consumption, while lo w consumption districts have shorter bars, possibly due to smaller populations or lower per ca pita consumption. Here, Thiruvananthapuram is the highest consumer whereas Wayanad is th e lowest consumer.

- **Plotting the total consumption by districts on the map of Kerala.**

```
# Plot the data
ggplot(data_map_data) +
  geom_sf(aes(fill = total_consumption, geometry = geometry)) +
  scale_fill_gradient(low = "red", high = "green") +
  ggtitle("Total Consumption by District") +
  geom_sf_text(aes(label = District, geometry = geometry), size = 3, color
= "black")
```



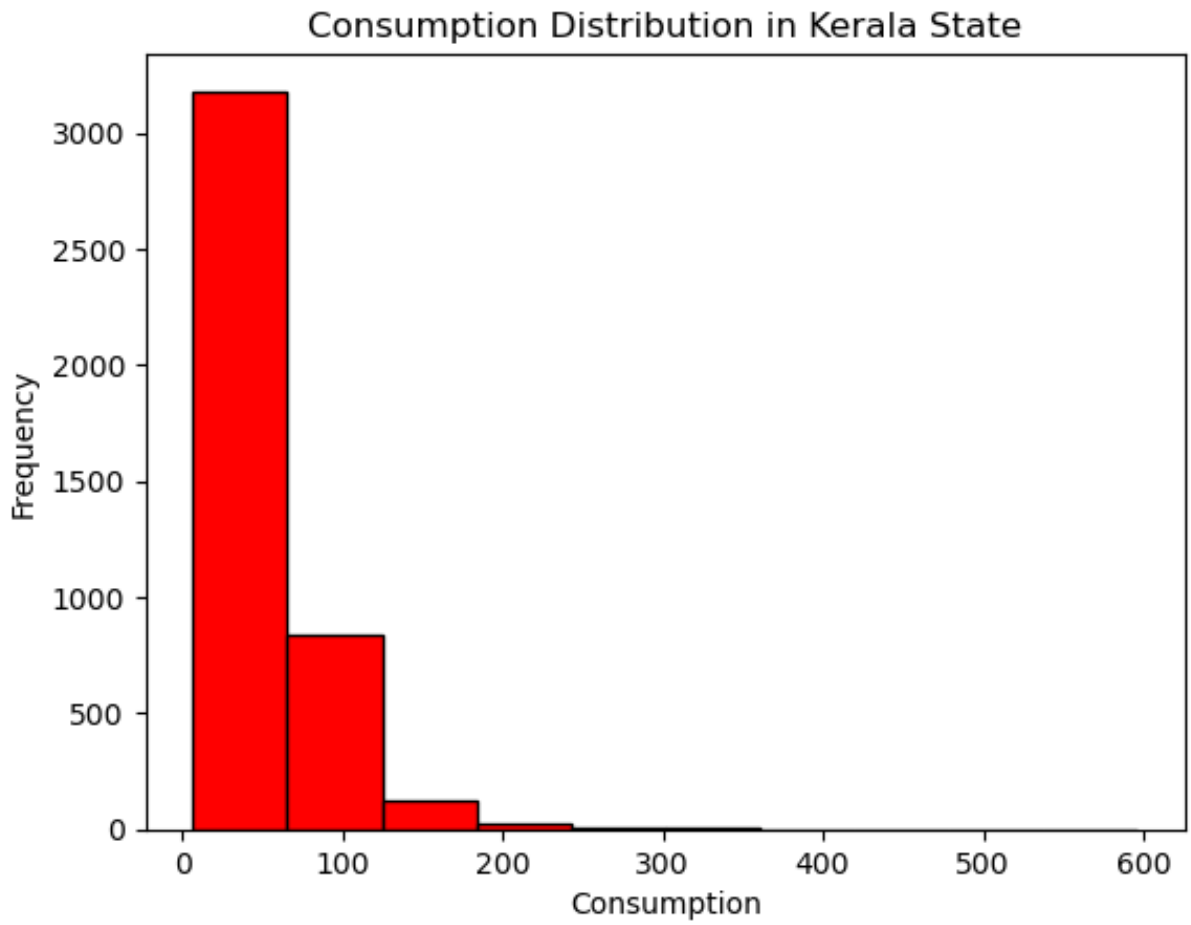
Interpretation:

The geospatial plot in Kerala shows the total food consumption across different districts, with each district labeled with its name. The color gradient indicates varying levels of consumption , providing a clear visual contrast between low and high consumption areas. High consumptio n areas are highlighted in green, indicating larger populations and higher per capita food intak e, while low consumption areas are highlighted in red, indicating smaller populations and low er per capita food intake. Here, Thiruvananthapuram shows the highest consumption wheres Wayanad and Idukki shows the least consumption. The plot also helps identify regional patter ns or clusters of high or low consumption, such as coastal districts consuming more food than inland districts or specific areas with notable consumption disparities. The plot provides a co mprehensive view of food consumption in Kerala.

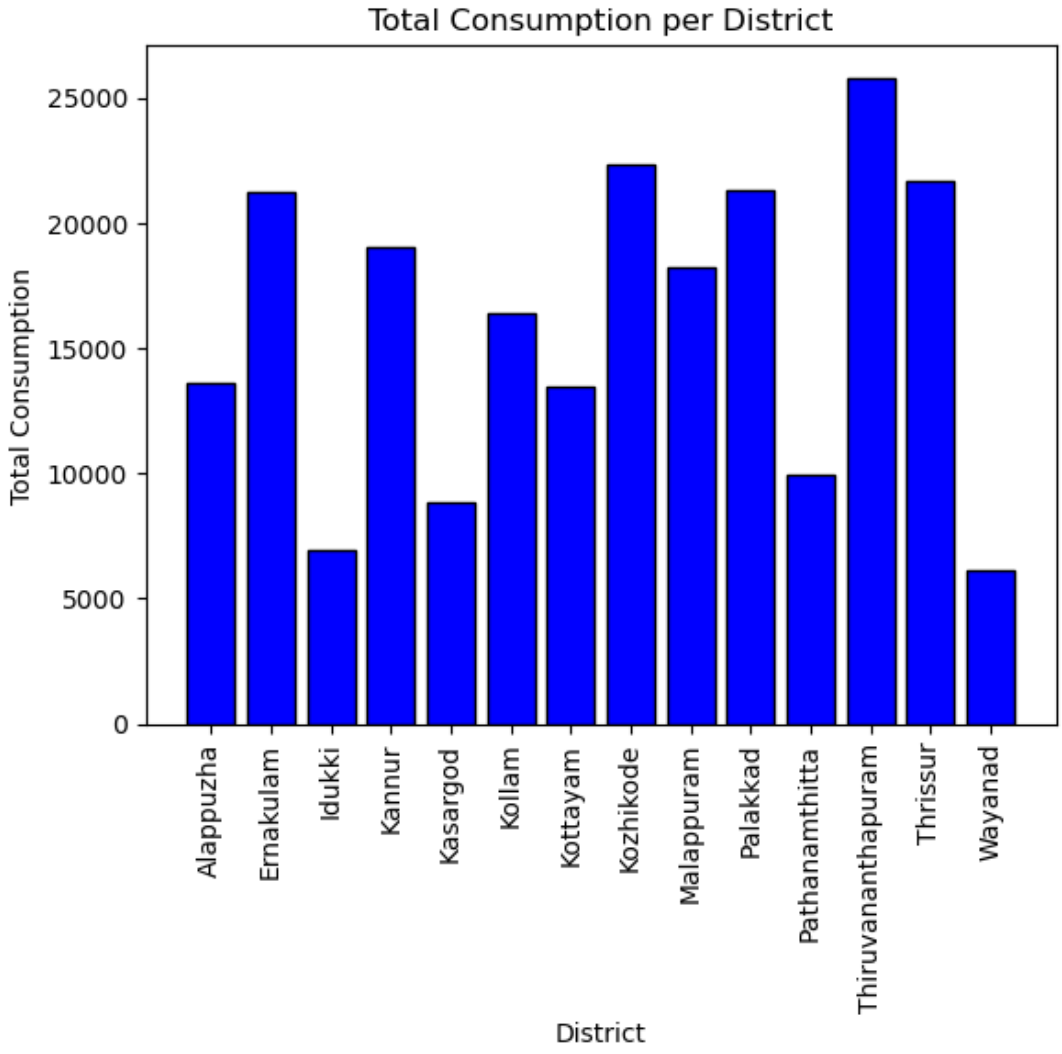
Results and Interpretation using Python

- **Plotting a histogram depicting consumption distribution in Kerala and a bar graph showing total consumption per district.**

```
plt.hist(KE_clean['total_consumption'], bins=10, color='red',
edgecolor='black')
plt.xlabel("Consumption")
plt.ylabel("Frequency")
plt.title("Consumption Distribution in Kerala State")
plt.show()
```




```
plt.bar(KE_clean_consumption['District'],
KE_clean_consumption['total_consumption'], color='blue', edgecolor='black')
plt.xlabel("District")
plt.ylabel("Total Consumption")
plt.title("Total Consumption per District")
plt.xticks(rotation=90)  # Rotate district names for better visibility
plt.show()
```



Interpretation:

The histogram in Kerala illustrates the distribution of total food consumption across different districts. It shows the central tendency of food consumption, indicating where most districts li e in terms of total consumption. The spread and skewness of the bars provide insights into the variability of food consumption across districts. If the histogram is skewed to the left or right, it suggests districts with significantly higher or lower consumption than the average. The bar g raph provides a detailed view of total food consumption in each district, listing the districts ve rtically. The Y-Axis shows the total amount of food consumed in each district, with each bar r epresenting the total consumption for a specific district. High consumption districts have talle r bars, possibly due to larger populations or higher per capita consumption, while low consum ption districts have shorter bars, possibly due to smaller populations or lower per capita consu mption. Here, Thiruvananthapuram is the highest consumer whereas Wayanad is the lowest c onsumer.

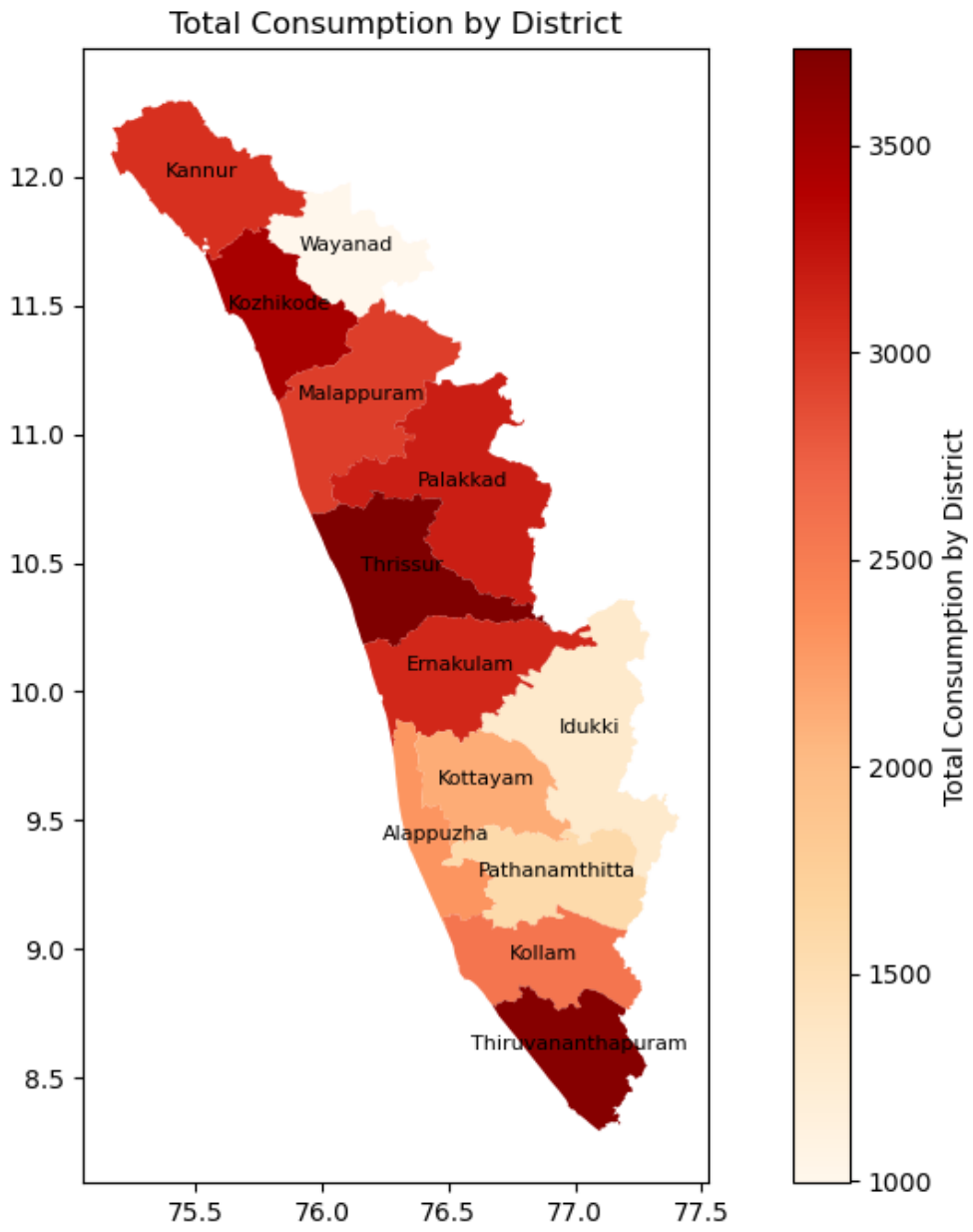
- **Plotting the total consumption by districts on the map of Kerala.**

```
# Load and plot Kerala state map
data_map = gpd.read_file("C:\\Users\\Ferah
Shan\\Downloads\\KERALA_DISTRICTS.geojson")

data_map = data_map.rename(columns={'dtname': 'District'})

# Merging kl_consumption and data_map tables
data_map_data = data_map.merge(kl_consumption, on='District')
print(data_map_data)
```

```
              District  stname stcode11 dtcode11 year_stat   Shape_Length  \
0            Kasaragod  KERALA       32      588    2011_c  382258.965224
1              Wayanad  KERALA       32      590    2011_c  292261.797789
2            Kozhikode  KERALA       32      591    2011_c  298776.567326
3           Malappuram  KERALA       32      592    2011_c  409487.628759
4              Palakkad  KERALA       32      593    2011_c  551293.272056
5              Thrissur  KERALA       32      594    2011_c  397475.610728
6                Idukki  KERALA       32      596    2011_c  570539.472994
7              Ernakulam  KERALA       32      595    2011_c  412906.029329
8              Alappuzha  KERALA       32      598    2011_c  299423.689877
9              Kottayam  KERALA       32      597    2011_c  321896.971631
10      Pathanamthitta  KERALA       32      599    2011_c  406371.586244
11              Kollam  KERALA       32      600    2011_c  349910.199376
12  Thiruvananthapuram  KERALA       32      601    2011_c  286375.498107
13              Kannur  KERALA       32      589    2011_c  346881.951323

      Shape_Area  OBJECTID  test  Dist_LGD  State_LGD  \
0   2.078638e+09        90     0       558         32
1   2.240878e+09        96     1       567         32
2   2.451416e+09        99     0       561         32
3   3.710887e+09       101     0       562         32
4   4.671520e+09       108     0       563         32
5   3.147926e+09       112     0       566         32
6   4.518944e+09       114     0       556         32
7   3.170929e+09       115     0       555         32
...
10  POLYGON ((76.67578 9.48423, 76.67573 9.4843, 7...
11  POLYGON ((76.93026 9.17391, 76.92989 9.17427, ...
12  POLYGON ((76.7828 8.86194, 76.78261 8.86203, 7...
13  POLYGON ((75.39207 12.2902, 75.39165 12.29051,...
```

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...

```
# Plot with labeled district names
fig, ax = plt.subplots(1, 1, figsize=(12, 8))
data_map_data.plot(column='total_consumption', cmap='OrRd', legend=True, ax=ax
, legend_kwds={'label': "Total Consumption by District"})
data_map_data.apply(lambda x: ax.annotate(text=x['District'], xy=x.geometry.ce
ntroid.coords[0], ha='center', fontsize=8, color='black'), axis=1)
plt.title('Total Consumption by District')
plt.show()
```



Interpretation:

The geospatial plot in Kerala shows the total food consumption across different districts, with each district labeled with its name. The color gradient indicates varying levels of consumption , providing a clear visual contrast between low and high consumption areas. High consumptio n areas are highlighted in green, indicating larger populations and higher per capita food intak e, while low consumption areas are highlighted in red, indicating smaller populations and low er per capita food intake. Here, Thiruvananthapuram shows the highest consumption wheres Wayanad and Idukki shows the least consumption. The plot also helps identify regional patter ns or clusters of high or low consumption, such as coastal districts consuming more food than inland districts or specific areas with notable consumption disparities. The plot provides a co mprehensive view of food consumption in Kerala.

Recommendation

The analysis of food consumption patterns across Kerala's districts has identified several recommendations for policymakers, businesses, NGOs, health organizations, and researchers. These recommendations include targeted nutritional programs, education and awareness campaigns, resource allocation, support local agriculture, market linkages, market expansion and targeting, tailored products, supply chain optimization, product innovation, community support programs, health and nutrition education, and school programs.

Policymakers should implement targeted nutritional programs and food subsidies in low-consumption districts to address potential food security and malnutrition issues. They should also conduct awareness campaigns to educate communities about balanced diets and the importance of nutrition, especially in low-consumption areas. Resource allocation should focus on improving food storage and distribution infrastructure in high-consumption districts to prevent food wastage and ensure efficient supply chains.

Businesses should focus on market expansion and targeting, developing and marketing products that cater to specific dietary preferences and needs of different districts. Supply chain optimization should be optimized to reduce costs and improve availability of food products in both high and low-consumption districts. Local partnerships with suppliers and distributors can enhance supply chain efficiency and reduce transportation times.

NGOs and health organizations should develop community support programs, such as food assistance programs targeting vulnerable populations in low-consumption districts, community kitchens, health and nutrition education workshops, and school programs to instill healthy eating habits in children from an early age. Researchers and academics should conduct further research to understand the socio-economic and cultural factors driving consumption patterns in different districts, study the impact of government policies and business initiatives on food consumption patterns and nutritional outcomes, and advocate for more granular and frequent data collection on food consumption.

In conclusion, adopting a multi-faceted approach that involves policymakers, businesses, NGOs, health organizations, and researchers can enhance food security, improve nutritional outcomes, and foster equitable economic development across Kerala.

R Codes

```
# Set the working directory and verify it
setwd("E:\\VCU\\Summer 2024\\Statistical Analysis & Modeling")
getwd()

# Load necessary libraries
library(tidyr)
library(ggplot2)
library(dplyr)
library(BSDA)
library(readr)
library(glue)

# Read the file into R
data <- read.csv("NSSO68.csv")

# Filter for Kerala
kl_df <- data %>% filter(state_1 == "KE")

# Display dataset info
cat("Dataset Information:\n")
print(names(kl_df))
print(head(kl_df))
print(dim(kl_df))

# Finding missing values
missing_info <- colSums(is.na(kl_df))
cat("Missing Values Information:\n")
print(missing_info)

any(is.na(kl_df))
sum(is.na(kl_df))
```

```
# Sub-setting the data
klnew <- kl_df %>%
  select(state_1, District, Region, Sector, State_Region, Meals_At_Home, ricepds_v,
Wheatpds_q, chicken_q, pulsep_q, wheatos_q, No_of_Meals_per_day)

# Check for missing values in the subset
cat("Missing Values in Subset:\n")
print(colSums(is.na(klnew)))

# Impute missing values with mean for specific columns
impute_with_mean <- function(column) {
  if (any(is.na(column))) {
    column[is.na(column)] <- mean(column, na.rm = TRUE)
  }
  return(column)
}
klnew$Meals_At_Home <- impute_with_mean(klnew$Meals_At_Home)

# Check for missing values after imputation
cat("Missing Values After Imputation:\n")
print(colSums(is.na(klnew)))

# Find outliers and removing them
boxplot(klnew$ricepds_v)
remove_outliers <- function(kl_df, column_name) {
  Q1 <- quantile(kl_df[[column_name]], 0.25)
  Q3 <- quantile(kl_df[[column_name]], 0.75)
  IQR <- Q3 - Q1
  lower_threshold <- Q1 - (1.5 * IQR)
  upper_threshold <- Q3 + (1.5 * IQR)
  kl_df <- subset(kl_df, kl_df[[column_name]] >= lower_threshold & kl_df[[column_
name]] <= upper_threshold)
  return(kl_df)
}
```

```
outlier_columns <- c("ricepds_v", "chicken_q")
for (col in outlier_columns) {
  klnew <- remove_outliers(klnew, col)
}

# Renaming districts and sectors
district_mapping <- c( "1" = "Kasargod", "2" = "Kannur", "3" = "Wayanad", "4" = "K
ozhikode", "5" = "Malappuram", "6" = "Palakkad", "7" = "Thrissur", "8" = "Ernakula
m", "9" = "Idukki", "10" = "Kottayam", "11" = "Alappuzha", "12" = "Pathanamthitta",
"13" = "Kollam", "14" = "Thiruvananthapuram")
sector_mapping <- c("1" = "Rural", "2" = "Urban")

klnew$District <- as.character(klnew$District)
klnew$Sector <- as.character(klnew$Sector)
klnew$District <- ifelse(klnew$District %in% names(district_mapping), district_map
ping[klnew$District], klnew$District)
klnew$Sector <- ifelse(klnew$Sector %in% names(sector_mapping), sector_mapping
[klnew$Sector], klnew$Sector)

# Summarize consumption
klnew$total_consumption <- rowSums(klnew[, c("ricepds_v", "Wheatpds_q", "chicke
n_q", "pulsep_q", "wheatos_q")], na.rm = TRUE)

district_summary <- summarize_consumption("District")
region_summary <- summarize_consumption("Region")

cat("Top 3 Consuming Districts:\n")
print(head(district_summary, 3))
cat("Bottom 3 Consuming Districts:\n")
print(tail(district_summary, 3))

cat("Region Consumption Summary:\n")
```

```
print(region_summary)

View(klnew)

hist(klnew$total_consumption, breaks = 10, col = 'blue', border = 'black',
    xlab = "Consumption", ylab = "Frequency", main = "Consumption Distribution in Kerala State")

klnew_consumption <- aggregate(total_consumption ~ District, data = klnew, sum)
View(klnew_consumption)

barplot(klnew_consumption$total_consumption,
        names.arg = klnew_consumption$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)

#(b) Plot on the Kerala state map using NSSO68.csv data

library(ggplot2)
library(sf) # mapping
library(dplyr)
Sys.setenv("SHAPE_RESTORE_SHX" = "YES")

data_map <- st_read("C:\\Users\\Ferah Shan\\Downloads\\KERALA_DISTRICTS.geojson")
View(data_map)

data_map <- data_map %>%
  rename(District = dtname)
```

```
colnames(data_map)

# Merge klnew_consumption with data_map
data_map_data <- data_map %>%
  left_join(klnew_consumption, by = c("District" = "District"))

library(sf)

# Reproject data_map to UTM zone 45N
data_map_utm <- st_transform(data_map, crs = 32645)

# Merge klnew_consumption with data_map_utm
data_map_data_utm <- data_map_utm %>%
  left_join(klnew_consumption, by = c("District" = "District"))

# Plot the data
ggplot(data_map_data) +
  geom_sf(aes(fill = total_consumption, geometry = geometry)) +
  scale_fill_gradient(low = "red", high = "green") +
  ggtitle("Total Consumption by District") +
  geom_sf_text(aes(label = District, geometry = geometry), size = 3, color = "black")
```

Python Codes

```
pip install geopandas

import pandas as pd
import numpy as np
from scipy import stats
import matplotlib.pyplot as plt
import seaborn as sns
import geopandas as gpd
```

```
df = pd.read_csv("E:\\VCU\\Summer 2024\\Statistical Analysis & Modeling\\NSSO68.csv", low_memory=False)
display(df)

KE = df[df['state_1']=="KE"]

KE.isnull().sum().sort_values(ascending = False)

df.columns

KE_new = KE[['state_1', 'District', 'Sector','Region','State_Region','ricetotal_q','wheattotal_q','moong_q','Milktotal_q','chicken_q','bread_q','foodtotal_q','Beveragestotal_v','Meals_At_Home']]

KE_new.isnull().sum().sort_values(ascending = False)

KE_clean = KE_new.copy()
KE_clean.loc[:, 'Meals_At_Home'] = KE_clean['Meals_At_Home'].fillna(KE_new['Meals_At_Home'].mean())
KE_clean.isnull().any()

def remove_outliers(df, column_name):
    Q1 = df[column_name].quantile(0.25)
    Q3 = df[column_name].quantile(0.75)
    IQR = Q3 - Q1
    lower_threshold = Q1 - (1.5 * IQR)
    upper_threshold = Q3 + (1.5 * IQR)
    df = df[(df[column_name] >= lower_threshold) & (df[column_name] <= upper_threshold)]
    return df

outlier_columns = ['ricetotal_q']
for col in outlier_columns:
    KE_clean = remove_outliers(KE_clean, col)
```

```
import matplotlib.pyplot as plt
plt.figure(figsize=(8, 8))
plt.boxplot(KE_clean['ricetotal_q'])
plt.xlabel('ricetotal_q')
plt.ylabel('Values')
plt.title('Boxplot of ricetotal_q')
plt.show()

rice1 = KE_clean['ricetotal_q'].quantile(0.25)
rice2 = KE_clean['ricetotal_q'].quantile(0.75)
iqr_rice = rice2-rice1
up_limit = rice2 + 1.5*iqr_rice
low_limit = rice1 - 1.5*iqr_rice

KE_clean=KE_new[(KE_new['ricetotal_q']<=up_limit)&(KE_new['ricetotal_q']>=low_limit)]

plt.boxplot(KE_clean['ricetotal_q'])

KE_clean['District'].unique()

KE_clean.loc[:,'Sector'] = KE_clean['Sector'].replace([1, 2], ['URBAN', 'RURAL'])

KE_clean.columns

KE_clean.loc[:, 'total_consumption'] = KE_clean[['ricetotal_q', 'wheattotal_q', 'moong_q', 'Milktotal_q', 'chicken_q', 'bread_q', 'foodtotal_q', 'Beveragestotal_v']].sum(axis=1)
KE_clean.head()

KE_clean.loc[:,"District"] = KE_clean.loc[:,"District"].replace({1: "Kasargod", 2: "Kannur", 3: "Wayanad", 4: "Kozhikode", 5: "Malappuram", 6: "Palakkad", 7: "Thrissur", 8: "Ernakulam", 9: "Idukki", 10: "Kottayam", 11: "Alappuzha", 12: "Pathanamthitta", 13: "Kollam", 14: "Thiruvananthapuram"})
```

```
KE_clean.loc[:,"Region"] = KE_clean.loc[:,"Region"].replace({1: "Rural", 2: "Urban"
})

KE_clean.groupby('Region').agg({'total_consumption':['std','mean','max','min']})
KE_clean.groupby('District').agg({'total_consumption':['std','mean','max','min']})

total_consumption_by_districtname=KE_clean.groupby('District')['total_consumption
'].sum()
total_consumption_by_districtname.sort_values(ascending=False).head(3)
total_consumption_by_districtname.sort_values(ascending=True).head(3)

KE_clean['District'] = KE_clean['District'].astype(str)
KE_clean['Sector'] = KE_clean['Sector'].astype(str)
print(KE_clean.head())

plt.hist(KE_clean['total_consumption'], bins=10, color='red', edgecolor='black')
plt.xlabel("Consumption")
plt.ylabel("Frequency")
plt.title("Consumption Distribution in Kerala State")
plt.show()

KE_clean_consumption = KE_clean.groupby('District')['total_consumption'].sum().re
set_index()
print(KE_clean_consumption.head())

plt.bar(KE_clean_consumption['District'], KE_clean_consumption['total_consumptio
n'], color='blue', edgecolor='black')
plt.xlabel("District")
plt.ylabel("Total Consumption")
plt.title("Total Consumption per District")
plt.xticks(rotation=90)  # Rotate district names for better visibility
plt.show()
```

```
# b) Plotting total consumption on the Kerala state map

# Filtering for Kerala
df_kl = data[data['state_1'] == "KE"]

# Sub-setting the data
kl_new = df_kl[['state_1', 'District', 'Region', 'Sector', 'State_Region', 'Meals_At_Hom
e', 'ricetotal_q', 'Wheatpds_q', 'chicken_q', 'pulsep_q', 'wheatos_q', 'No_of_Meals_per
_day']]

# Check for missing values in the subset
print("Missing Values in Subset:")
print(kl_new.isnull().sum())

# Impute missing values with mean for specific columns
kl_new['Meals_At_Home'].fillna(kl_new['Meals_At_Home'].mean(), inplace=True)

# Check for missing values after imputation
print("Missing Values After Imputation:")
print(kl_new.isnull().sum())

# Remove outliers
for col in outlier_columns:
    kl_new = remove_outliers(kl_new, col)

# Summarize consumption
kl_new['total_consumption'] = kl_new[['ricetotal_q', 'Wheatpds_q', 'chicken_q', 'pulse
p_q', 'wheatos_q']].sum(axis=1)

district_summary = kl_new.groupby('District')['total_consumption'].sum().reset_index
().sort_values(by='total_consumption', ascending=False)
print("District Consumption Summary:")
print(district_summary)
```

```
# Mapping districts so that merging of the tables will be easier
district_mapping = {
    "1": "Kasargod", "2": "Kannur", "3": "Wayanad", "4": "Kozhikode", "5": "Malappu
ram", "6": "Palakkad", "7": "Thrissur", "8": "Ernakulam", "9": "Idukki", "10": "Kottay
am", "11": "Alappuzha", "12": "Pathanamthitta", "13": "Kollam", "14": "Thiruvananth
apuram"
}

kl_new['District'] = kl_new['District'].astype(str).map(district_mapping).fillna(kl_new
['District'])
print(kl_new)

# kl_consumption stores aggregate of total consumption district-wise
kl_consumption = kl_new.groupby('District')['total_consumption'].sum().reset_index(
)
print(kl_consumption)

# Load and plot Kerala state map
data_map = gpd.read_file("C:\\Users\\Ferah Shan\\Downloads\\KERALA_DISTRICT
S.geojson")

data_map = data_map.rename(columns={'dtname': 'District'})
print(data_map)

# Merging kl_consumption and data_map tables
data_map_data = data_map.merge(kl_consumption, on='District')
print(data_map_data)

import geopandas as gpd
import pandas as pd
import matplotlib.pyplot as plt

# Plot with labeled district names
fig, ax = plt.subplots(1, 1, figsize=(12, 8))
```

```
data_map_data.plot(column='total_consumption', cmap='OrRd', legend=True, ax=ax,
legend_kwds={'label': "Total Consumption by District"})
data_map_data.apply(lambda x: ax.annotate(text=x['District'], xy=x.geometry.centroi
d.coords[0], ha='center', fontsize=8, color='black'), axis=1)
plt.title('Total Consumption by District')
plt.show()
```

References

1. www.github.com
2. www.geeksforgeeks.com
3. www.datacamp.com
4. www.icssrdataservice.in
5. www.medium.com