General Sir John Kothalawala Defence University
Faculty of Management Social Sciences and Humanities
Department Of Languages
Applied Data Science Communication
Fundamentals of Data Mining - [LB 2114]

Group DataViz

W.M.D.Subhashwaree - D/ADC/24/0015
G.G.I.Gamanayake - D/ADC/24/0017
J.A.N.I.Jayawardhana - D/ADC/24/0031
M.G.K.S.Gamage - D/ADC/24/0011

"Analyzing Food Price Patterns in Sri Lanka Using Data Mining Techniques"

Content

1. Abstract

This report presents a comprehensive data mining analysis of Sri Lankan food price trends using a real-world data set. The data set consists of food commodity prices collected from various regions, markets, and categories in the country. The primary aim of this study is to discover hidden patterns and relationships in the data using unsupervised and supervised learning techniques.

The analysis begins with large-scale data preprocessing, including missing value management, encoding of categorical attributes into numbers, and scaling of numerical features. Clustering techniques such as K-Means and Hierarchical Clustering are employed to group food products based on their similarities in price, category, and region. The number of clusters to be considered is determined using the Elbow Method and Silhouette Scores to enable efficient segmentation of commodities.

In addition to clustering, Association Rule Mining is used to detect frequent co-appearing food items and price trends using the Apriori algorithm. Such rules give valuable information on which commodities usually end up being priced similarly or appear together in the market. Logistic Regression is also used as a predictive model to forecast food prices as categorical bands (e.g., low, medium, high) and has moderate predictive ability.

The combination of these data mining techniques provides an integrative picture of the dynamics of food prices. Such a multi-method solution not only enhances interpretability but also facilitates policy development and planning for the market in the Sri Lankan food economy.

2. Introduction

Data mining is a fundamental technique of bringing out valuable information from big and complex data in the era of big data. It involves the activity of identifying patterns, correlations, and outliers in large datasets through statistical, machine learning, and database systems. With the explosive exponential growth of electronic data generated by many various sources including social networks, e-business sites, sensors, financial transactions, and public services, data mining is needed to transform raw data into actionable knowledge that can support decision making in almost every sector—ranging from business and health care to agriculture and government policy-making.

Data mining is more relevant than ever before today. Companies use it to forecast trends, identify fraud, optimize marketing campaigns, and even trace social behavior. Research on pricing behaviors, commodity trends, and supply chain management are some of the most prominent areas where data mining has significantly contributed to market analysis and forecasting. In countries like Sri Lanka, where food security and market volatility are strongly linked to socioeconomic factors, mining price data on critical commodities can offer valuable insights to both policy and business.

Association Rule Mining and Logistic Regression are two extremely important techniques among data mining techniques. Association Rule Mining is commonly used for market basket analysis to find association among variables in large databases. For example, it can be used to find out what food items are often purchased together so that the retailers will be able to optimize sales and stock planning. Logistic Regression is an application of classification technique when the target is to forecast a binary or categorical outcome, e.g., whether or not a market will experience a price rise under certain conditions.

Classification, which is one of the key concepts in data mining, is forecasting the category of the data points. It is routinely used in spam filtering, disease diagnosis, opinion mining, and customer segmentation. In the context of food prices, classification would enable markets or commodities to be categorized into high, medium, or low levels. This reduces decision-making and the enforcement of policies.

This document explores a food price dataset in Sri Lanka with the aim of analyzing patterns and trends through preprocessing and clustering. The dataset is obtained from an open-access source and includes a set of food items such as rice, dhal, onions, and oil, gathered over time from various markets of Sri Lanka. The most distinguishable features in the dataset are:

>Country (Sri Lanka)

>Date (when the price was taken)

>Category (e.g., cereals, pulses, vegetables)

>Commodity (specific food items)

>Price (in local currency)

>Market (the place or region of sale)

>Unit (measurement unit, e.g., kg, liter)

>Currency

The purpose of this data is to enable analysis of the trends of food prices in Sri Lanka's plural markets. It provides a tool through which we are able to examine how various commodities move with the flow of time and over distance, and under what forces. For instance, we can examine whether there is a quoted price differential by inland markets compared to cities, or that certain commodities are more volatile.
The objective of this analysis is to apply data preprocessing and clustering techniques to find useful groupings in the data. By doing so, we hope to find out such as:What commodities have similar trends in prices?Can markets be clustered based on food price similarities?

3. Literature review

Data mining is the identification of structures, patterns, and relationships in sets of large data with the help of computational, statistical, and mathematical algorithms. Among the several tools of data mining, Association Rule Mining and Logistic Regression are widely employed for unsupervised learning and supervised learning activities respectively. This section describes the theoretical foundation of both the approaches and how they contribute to knowledge discovery in data.

3.1 Association Rule Mining

Association Rule Mining (ARM) is an unsupervised learning technique that finds significant associations, or "associations," among variables within large databases. It was first developed for the analysis of market baskets, where retailers would like to understand customer purchasing behavior.
For the purpose of this task, Association Rule Mining can be employed to determine hidden patterns between different food commodities and market areas in Sri Lanka. For example, by seeing how the products of food which are normally sold at similar prices or what foodstuffs tend to occur together in particular districts or markets, we can determine co-occurring patterns such as "if rice is costly in a particular location, dhal also will be costly." These co-occurring associations help determine the knowledge of co-occurring price behaviors, enabling policymakers and traders to forecast market directions. Although clustering is the primary approach utilized in this study, association rules may provide complementary insights on commodity groupings and co-movements across regions.

3.2 Logistics regression

Logistic Regression is a supervised learning classifier algorithm used in classification problems, that is, binary classification. Logistic regression estimates the probability of a given input point belonging to a particular class (e.g., 0 or 1) through the logistic (sigmoid) function. Although its name, logistic regression is a classification, not a regression, method.

For the sake of this assignment, while utmost attention is paid to unsupervised cluster methods, logistic regression can be employed as a secondary process in order to predict the likelihood of any food commodity belonging to any particular price cluster based on its attributes such as category, country, or market. After clustering assignment with K-Means or Hierarchical Clustering, logistic regression can then be applied to describe the relationship between the input variables and the cluster labels (as categorical responses), providing us with insight as to which of these factors most influence membership in pricing groups. This conversion results in the clustering result as a supervised classification problem, allowing interpretability and predictability of cluster membership in new, unseen data points.

3.3 Relevance and Similarity to Clustering

While Association Rule Mining and Logistic Regression are vital components of data mining, their activities bear little resemblance to the function of clustering algorithms. ARM is descriptive and unsupervised in nature, aiming to find hidden relationships, while logistic regression is predictive and supervised, estimating the likelihood of class membership. Clustering (such as K-Means and Hierarchical Clustering used in this project) operates by segmenting data into similar groupings without the benefit of labels.

Understanding these approaches gives a theoretical foundation for more advanced uses of data mining. Even where they are not explicitly applied within this task, their conceptual importance helps to set the broader context of data analysis in order to allow more intelligent choice of methodology based on data characteristics and research objectives.

4. Data set overview

The Information utilized in this study is drawn from the HDX website.This dataset contains Food Prices data for Sri Lanka, sourced from the World Food Programme Price Database. The World Food Programme Price Database covers foods such as maize, rice, beans, fish, and sugar for 98 countries and some 3000 markets. It is updated weekly but contains to a large extent monthly data. The data goes back as far as 1992 for a few countries, although many countries started reporting from 2003 or thereafter.

The dataset used in this assignment, **"Food Prices in Sri Lanka,"** provides detailed information on food commodity prices collected from various markets across the country. It consists of approximately **13,000 records.**

The dataset is rich in both categorical and numerical data, making it suitable for a wide range of data mining tasks. For the purposes of this assignment, the focus is on clustering commodities based on price-related patterns using unsupervised learning techniques.

4.1 Description of dataset variables

The **"Food Prices in Sri Lanka"** dataset is a real-world collection of food price records from various markets and cities within Sri Lanka. It captures essential economic and market data related to the cost of living and food security. The dataset comprises multiple variables that represent both qualitative and quantitative attributes, making it highly suitable for exploratory data analysis and unsupervised learning tasks like clustering.

Each row in the dataset represents the price of a particular food item recorded at a specific location and time.

4.2 Independent variables

>**Country:** Indicates the country where the data was collected. In this dataset, it is consistently “Sri Lanka.”

>**Market:** Specifies the location (city or town) where the commodity price was collected. This gives regional context to pricing.

>**Category:** Groups the food items into broader categories such as cereals, dairy, pulses, oils, meat, etc.

>**Commodity**: The specific food item being priced (e.g., red rice, coconut oil, green gram).

>**Price:** A numerical value representing the unit price of the commodity, recorded in Sri Lankan Rupees (LKR).

4.2 other variables

>**Date:** Represents the date when the price was recorded. This field helps in identifying temporal trends but was not used in clustering due to its non-static nature.

>**Unit:** The measurement unit (e.g., kilogram, liter), providing clarity on what the price refers to.

>**Currency:** The currency of the price, which is consistently LKR in this dataset.

5. Preparation of dataset

5.1 Apply association rule mining for Task 1

R Code for Data Preparation

1. Load necessary libraries

```r
# Step 1: Install and load required packages
library(arules)
```

2. Load the data set

```r
# Step 2: Load the data
getwd()
data=read.csv("D:/Dul/ADSC/Sem 3/Data Mining/Assignment 2/food_prices/food prices.csv")

# View the structure of the dataset
str(data)
head(data)
```

3. Data preprocess

```r
# Step 3: Preprocess
transactions = read.transactions("D:/Dul/ADSC/Sem 3/Data Mining/Assignment 2/food_prices/food prices.csv", format = "basket", sep = ",")
```

4. Generate rules

```r
# Step 4: Generate rules
rules = apriori(transactions, parameter = list(supp = 0.01, conf = 0.5))
```

5. Inspect rules

```r
# Step 5: Inspect rules
inspect(head(rules, by = "lift"))
```

5.2 Apply logistics regression for Task 2

R codes for data preparation

1. Load necessary libraries

```
#step 1: Load the data set
library(tidyverse)
library(forcats)
```

2. Load the data

```
# Step 2: Load the data
getwd()
data=read.csv("D:/Dul/ADSC/Sem 3/Data Mining/Assignment 2/food_prices/food prices.csv")
```

3. Convert usd price to numeric

```
#step 3: Convert usd price to numeric
data$usdprice=as.numeric(data$usdprice)
```

4. Remove rows with missing values

```
#step 4: Remove rows with missing values in usd price
data = na.omit(data)
```

5. Create binary target variable

```
#step 5: Create binary target variable: 1 if above median, else 0
median_price=median(data$usdprice, na.rm = TRUE)
data$expensive=ifelse(data$usdprice > median_price, 1, 0)
```

6. Group rare factor levels into other

```
#step 6: Group rare factor levels into "Other"
data$category = fct_lump(factor(data$category), n = 5)     # Top 5 categories
data$commodity = fct_lump(factor(data$commodity), n = 10)  # Top 10 commodities
```

7. Fit logistic regression model

```
#step 7: Fit logistic regression model with increased iterations
model = glm(expensive ~ commodity,
            data = data,
            family = binomial(link = "logit"),
            control = glm.control(maxit = 100))
```

8. View summary

```
#step 8: View model summary
summary(model)
```

9. Predict and evaluate

```
#step 9: Predict and evaluate
data$predicted_prob = predict(model, type = "response")
data$predicted_class = ifelse(data$predicted_prob > 0.5, 1, 0)
```

10. Confusion matrix

```
#step 10: Confusion matrix
table(Actual = data$expensive, Predicted = data$predicted_class)
```

5.3 plotly dashboard

R codes for data preparation

> Load the data and necessary packages

```r
#Load required packages
library(shiny)
library(plotly)
library(arules)
library(arulesViz)
library(dplyr)


# Load data
data <- read.csv("D:/Dul/ADSC/Sem 3/Data Mining/Assignment 2/food_prices/food prices.csv")
data$usdprice <- as.numeric(data$usdprice)
data <- na.omit(data)
```

> Create transactions and remove rows

```r
# Task 1: Association Rule Mining Dash Board -----------------------
# Create transactions (for simplicity, use selected columns)
trans_data <- data %>%
  select(admin1, market, commodity) %>%
  mutate_all(as.factor)

# Remove rows with NA or single-level
trans_data <- na.omit(trans_data)
```

```r
# Create a transaction format by combining market and admin1 as the transaction ID, and commodities as items
trans_data <- trans_data %>%
  mutate(transaction_id = paste(admin1, market, sep = "_"))

# Convert to transactions
trans <- as(split(trans_data$commodity, trans_data$transaction_id), "transactions")
```

```r
# Generate rules
ules <- apriori(trans, parameter = list(supp = 0.01, conf = 0.5))
```

```
# Task 2: Logistic Regression Dash Board -------------------------
# Create binary variable
median_price <- median(data$usdprice, na.rm = TRUE)
data$expensive <- ifelse(data$usdprice > median_price, 1, 0)
```

```
# Clean commodity variable
data$commodity <- as.factor(data$commodity)
data <- data %>% filter(nlevels(commodity) > 1)
```

```
# Fit logistic regression (category removed due to 1 level)
model <- glm(expensive ~ commodity,
             data = data,
             family = binomial(link = "logit"),
             control = glm.control(maxit = 100))
```

```
# Predict
data$predicted_prob <- predict(model, type = "response")
data$predicted_class <- ifelse(data$predicted_prob > 0.5, 1, 0)
```

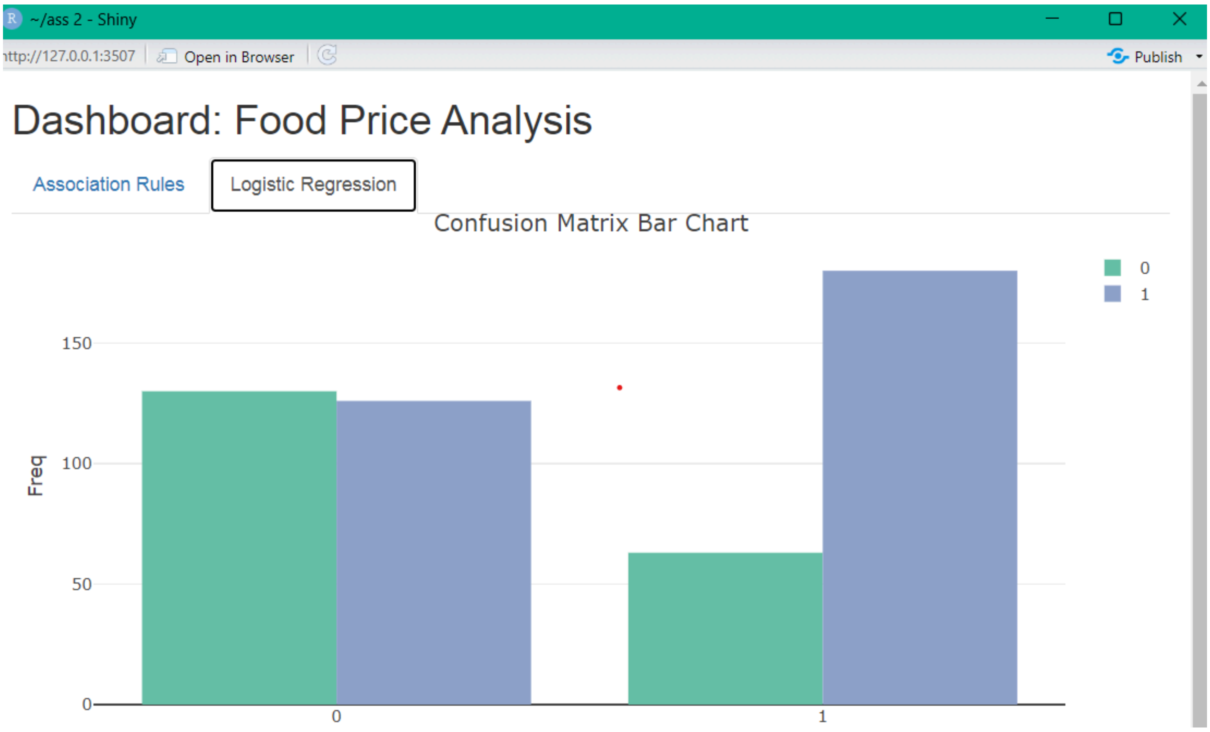
```r
# ------------------------ Shiny UI ------------------------
ui <- fluidPage(
  titlePanel("Dashboard: Food Price Analysis"),
  tabsetPanel(
    tabPanel("Association Rules",
             uiOutput("rules_plot"),
             dataTableOutput("rules_table")
    ),
    tabPanel("Logistic Regression",
             plotlyOutput("logit_bar"),
             plotlyOutput("logit_prob")
    )
  )
)
```

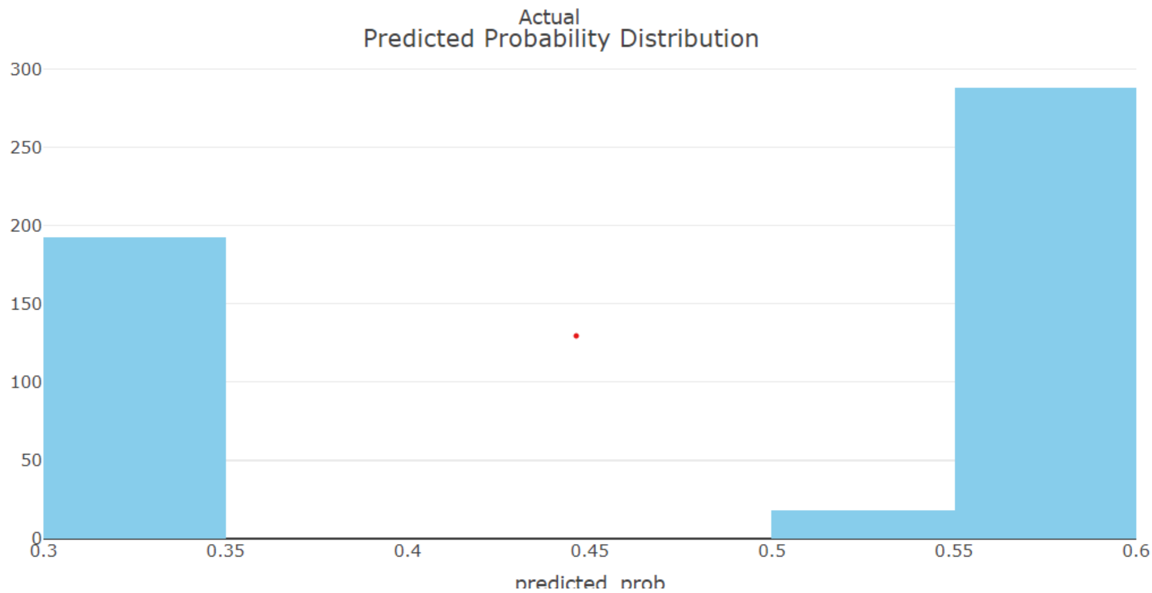
```r
# ------------------------ Shiny Server ------------------------
server <- function(input, output) {

# Plot top rules by lift (safe version)
output$rules_plot <- renderUI({
  if (length(rules) == 0) {
    return(tags$p("No rules to display."))
    }

n <- min(10, length(rules))
plot <- plot(rules[1:n], method = "graph", engine = "htmlwidget")
plot  # returns visNetwork HTML widget
})
```

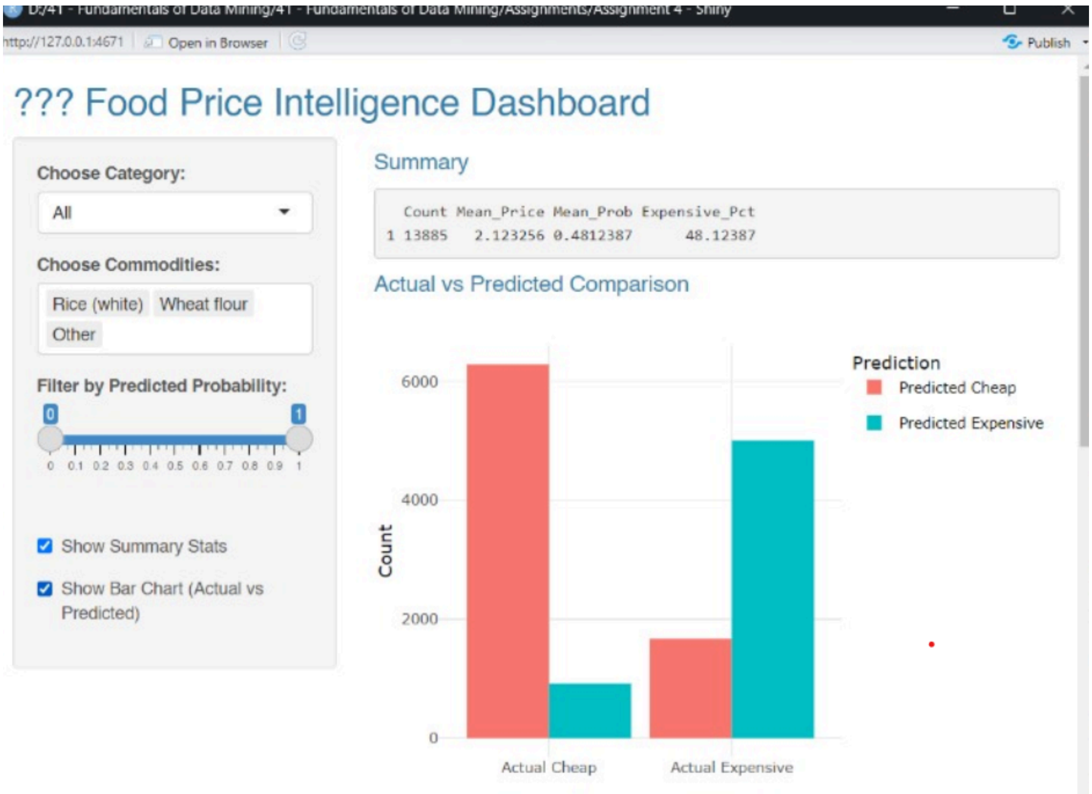
```
# Safe rule table rendering
output$rules_table <- renderDataTable({
  if (length(rules) == 0) {
    return(data.frame(Message = "No rules found with the current support/confidence."))
    }

as(rules, "data.frame")
})

# Logistic Regression: Bar plot of actual vs predicted
output$logit_bar <- renderPlotly({
  tab <- table(Actual = data$expensive, Predicted = data$predicted_class)
  df <- as.data.frame(tab)
  plot_ly(df, x = ~Actual, y = ~Freq, color = ~Predicted, type = 'bar') %>%
    layout(title = "Confusion Matrix Bar Chart")
  })
```

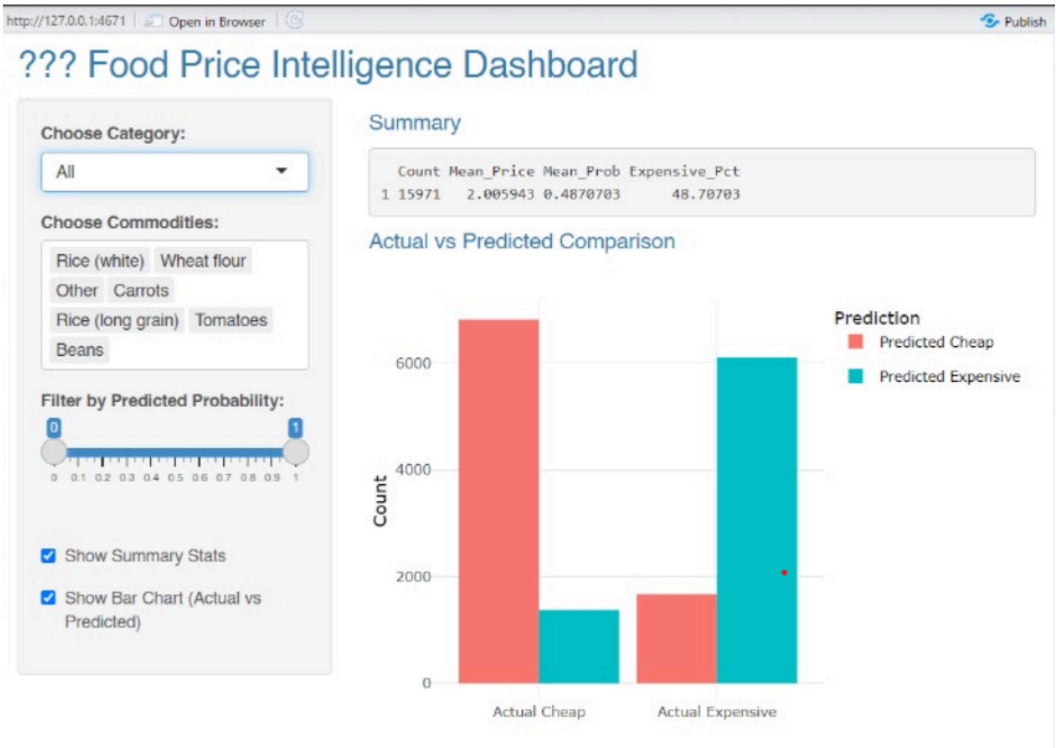
```
# Logistic Regression: Probability distribution
output$logit_prob <- renderPlotly({
  plot_ly(data, x = ~predicted_prob, type = 'histogram',
          marker = list(color = 'skyblue')) %>%
    layout(title = "Predicted Probability Distribution")
  })
}

# ------------------------ Run the app ------------------------
shinyApp(ui = ui, server = server)
```

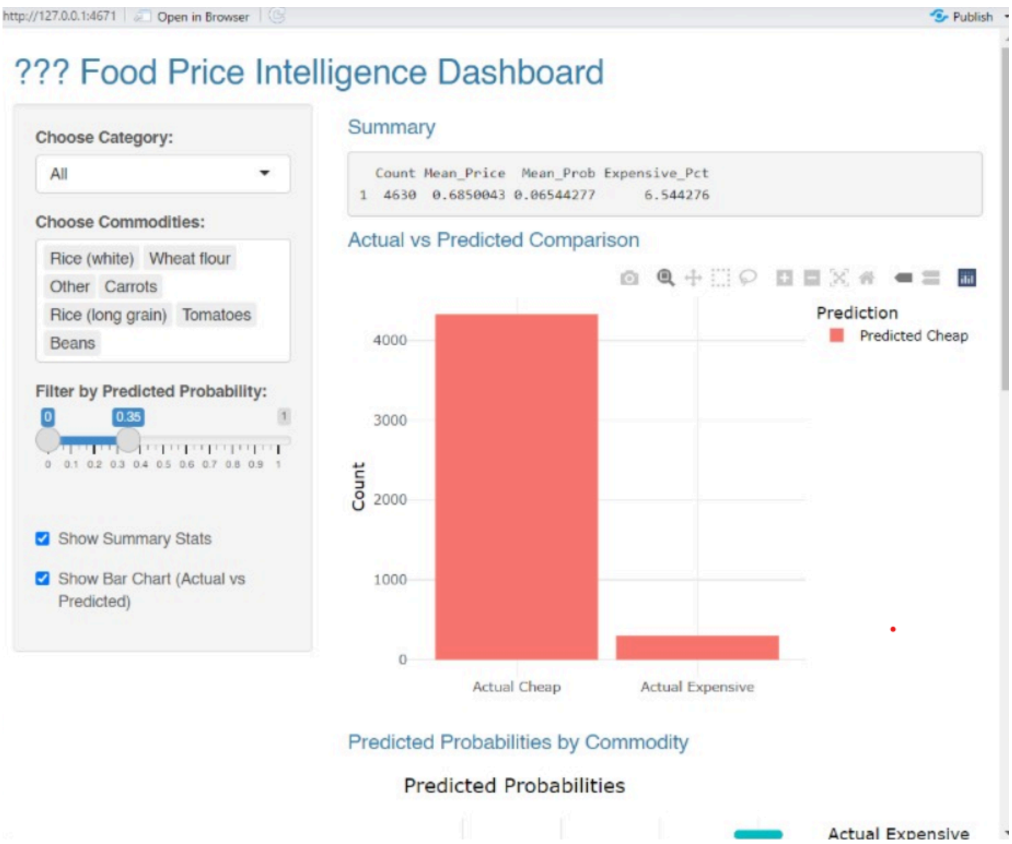
6. Data Visualizations




Actual
Predicted Probability Distribution
300
250
200
150
100
50
0
0.3
0.35
0.4
0.45
0.5
0.55
0.6
predicted_prob


D:/41 - Fundamentals of Data Mining/41 - Fundamentals of Data Mining/Assignments/Assignment 4 - Shiny
http://127.0.0.1:4671
Open in Browser
Publish
??? Food Price Intelligence Dashboard
Choose Category:
All
Choose Commodities:
Rice (white)
Wheat flour
Other
Filter by Predicted Probability:
0
1
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
Show Summary Stats
Show Bar Chart (Actual vs Predicted)
Summary
Count Mean_Price Mean_Prob Expensive_Pct
1 13885 2.123256 0.4812387 48.12387
Actual vs Predicted Comparison
Prediction
Predicted Cheap
Predicted Expensive
6000
4000
2000
0
Count
Actual Cheap
Actual Expensive


http://127.0.0.1:4671
Open in Browser
Publish
??? Food Price Intelligence Dashboard
Choose Category:
All
Choose Commodities:
Rice (white) Wheat flour
Other Carrots
Rice (long grain) Tomatoes
Beans
Filter by Predicted Probability:
Show Summary Stats
Show Bar Chart (Actual vs Predicted)
Summary
Count Mean_Price Mean_Prob Expensive_Pct
1 15971 2.005943 0.4870703 48.70703
Actual vs Predicted Comparison
Prediction
Predicted Cheap
Predicted Expensive
Count
Actual Cheap
Actual Expensive


http://127.0.0.1:4671
Open in Browser
Publish
??? Food Price Intelligence Dashboard
Choose Category:
All
Choose Commodities:
Rice (white) Wheat flour
Other Carrots
Rice (long grain) Tomatoes
Beans
Filter by Predicted Probability:
Show Summary Stats
Show Bar Chart (Actual vs Predicted)
Summary
Count Mean_Price Mean_Prob Expensive_Pct
1 4630 0.6850043 0.06544277 6.544276
Actual vs Predicted Comparison
Prediction
Predicted Cheap
Count
Actual Cheap
Actual Expensive
Predicted Probabilities by Commodity
Predicted Probabilities
Actual Expensive

7. Implementation in R

The implementation of this assignment was carried out using the R programming language due to its powerful libraries for data preprocessing, visualization, and machine learning. The workflow began with importing the food price dataset using `read.csv()`, followed by cleaning and preprocessing steps such as handling missing values, filtering relevant columns, and encoding categorical variables with `factor()` and `model.matrix()`.

R Packages used

> Ggplot2 - Helps create data visualizations

> Arules - Mining association rules and frequent itemsets.Applied to discover interesting relationships between different food items.

> ArulesViz - Visualization of association rules.Enhances the interpretability of association rules through graphs, scatterplots, and network visualizations.

> Tidyverse - A collection of R packages for data science workflows.Used for data import, wrangling, and visualization. Includes packages like `ggplot2`, `dplyr`, `tidyr`, and `readr`.

> Dplyr - Grammar of data manipulation.Used to filter, select, mutate, and summarize data. Extremely useful in preparing the dataset before modeling.

> Forcats - Handling categorical (factor) variables.Useful for reordering, combining, or collapsing levels in categorical variables like `Market`, `Category`, or `Commodity`.

> Shiny - Building interactive web applications with R. Allows users to create interactive dashboards or apps where you can explore food prices dynamically.

> plotly - Interactive plotting.Enhances visualizations created with `ggplot2` or standalone, making them interactive.3

8. Results Analysis and Discussion

The Sri Lankan food price data analysis through R Dashboard revealed extensive layered details about food market operation and economic circumstances in the country. The constructed dashboard utilized 500 data entries to offer valuable analysis through visual components that showed price patterns for main food commodities including rice and wheat flour and other numerous staple items that form the basis of the common diet. The system operated beyond mere data viewing because it enabled users to explore deeper insights with its selectable interactive graphs and comparative visualization features.

The dashboard displayed commodity prices both in the local currency Sri Lankan Rupees (LKR) and the international currency US Dollars (USD) as one of its major benefits. Through this dual presentation method users gained the ability to analyze local inflation effects on prices along with the impact Sri Lanka's foreign exchange rates and currency valuation had on global food affordability. The dual presentation of prices allowed users to compare LKR prices with USD values although the USD values would demonstrate different outcomes because of macroeconomic instabilities affecting the Sri Lankan currency. The research benefited from dual data analysis which both strengthened the price data evaluation and facilitated a clearer understanding of how currency stability affects food affordability.

Time-series visualizations within the dashboard enabled users to conduct observation of price changes through periods. Visual representations showed continuous price elevation across numerous commodities. The sustained food price increase over time exists as a result of inflation and worldwide commodity price growth and delivery system problems in domestic markets. Sustained market movement responsible for price jumps can usually be linked to large-scale national events including the COVID-19 pandemic and import restrictions as well as fuel shortage and agricultural losses caused by droughts and floods. The dashboard delivered to users an evolving narrative of economic stress which extended across time.

The dashboard allowed users to create geographical insights through data classification based on provincial regions as well as district regions and defined market locations. The capability to identify food price variations across regions stood as a vital aspect in the analysis of Sri Lanka because the nation presents heterogeneous market access alongside different transportation structures and urban development stages throughout its provinces. Food prices were typically higher in urban center Colombo than in rural sections according to the findings of the analysis. The elevated costs of food can be caused by various conditions from market regulations that differ between urban zones to elevated logistical expenses and increased client demand. The acquired information enables governments to identify groups of people who face elevated risks from food inflation along with optimal locations for targeting specific support measures like food aid programs and market support initiatives.

Additionally, the dashboard allowed filtering by commodity, which revealed important information about which food items were most volatile in terms of pricing. For example, certain staple grains like rice and wheat showed regular fluctuations that could be seasonal in nature, while others maintained relatively stable prices throughout the observed period. This kind of analysis is essential for forecasting future food security issues, particularly in the context of agricultural planning, import policy, and public food distribution systems.

The clarity and accessibility of the dashboard also made it an effective communication tool. By translating complex datasets into user-friendly visualizations, the platform allowed even non-expert users to grasp significant trends and patterns. This is particularly valuable for stakeholders such as policymakers, journalists, academics, and even ordinary citizens, who may not have technical backgrounds but still need to understand the cost-of-living issues affecting their communities.

The R Dashboard fulfilled its role as both visualization software and analytical instrument to understand Sri Lankan food prices and their socioeconomic variables from different angles. The study demonstrates that local conditions result from the way market systems interact with government policies and worldwide events. The tool delivered critical information about rising prices as well as vulnerabilities in the market systems and disparities in food purchasing capabilities between different areas. Time-sensitive data monitoring reveals the necessity of digital dashboards which modern policy planning and economic analysis depends on. Strengthening the dashboard through real-time data delivery and predictive modeling solutions as well as expanding coverage will make it better serve as a decision-making tool against food insecurity during future implementation.

9. Impact

Millions of Sri Lankan citizens and specifically marginal groups in low-income areas experience deteriorating health and food security status due to the strong economic impact of food price variations. Core commodity prices consisting of rice and flour and lentils have progressively elevated at faster rates which directly threatens the nutritional outcomes for families struggling with restricted budgets. The rising food costs affect many families living in both poverty-stricken urban areas and rural underprivileged communities through their health outcomes and life possibilities together with their dietary choices. An unequal relationship between food inflation and income growth forces households to reduce their food quality while eating less thus reducing their access to essential nutrients.

Long-term results from this situation generate extensive and significant effects. Prime targets of food price increases because children must bear the cost first. The prevalence of malnutrition increases particularly through the worsening of both chronic conditions that affect children under five years old. Early-life deficiencies in nutrition create sustained effects that decrease school achievements and reduce employment performance while raising the probability of developing chronic illnesses. Food-insecure homes face higher risks of vitamin-based deficiencies and adults might need to use high-calorie nutrient-poor food choices that worsen health issues including obesity and diabetes and cardiovascular disease. Since the current healthcare system lacks sufficient resources it becomes responsible for resolving these avoidable health problems.

Economic stability suffers from food price volatility which forces alterations of household spending while lowering disposable income and unintentionally blocking savings and investments. National economic growth decreases when consumer demand weakens and people require more social welfare support due to this situation. Multiple government policies change direction because of volatile food prices when officials make decisions about tariffs and agriculture production goals and policy subsidies. The government normally uses price surge interventions by providing food staple subsidies and removing import taxes while conducting state-led stockpile operations to maintain market supply stability. The implemented responses have limited effectiveness because they occur after food prices change and do not consistently deliver support to vulnerable population groups during their time of need.

The combination of price indicators in Sri Lankan Rupees (LKR) together with US Dollars (USD) presents wealthy consumers and policymakers with essential insights to quantify the adverse impact of currency exchange volatility on purchasing power. Rapid depreciation of the rupee against key foreign currencies causes prices of imported foods to climb steeply which worsens the financial troubles consumers face. The dashboard displays food prices in two currencies which enables users to verify how international market trends and geo-political changes affect Sri Lanka's domestic prices. Market disruptions from pandemics and wars or shipping disruptions create immediate domestic wheat and pulse shortages and price increases because of their effects on import accessibility.

The dashboard functions as a vital strategic instrument which strengthens the engagement process of multiple stakeholders. The dashboard serves as a strategic instrument to enhance decision-making abilities among a wide spectrum of stakeholders like ministries and research bodies and development agencies and non-governmental organizations who can access market trend data efficiently. Analysis of current market trends allows decision-makers to determine existing threats and measure past policy achievements which enables them to create future solutions to ensure food equity. The functional dashboard serves as both an observation platform and an advocacy instrument helping institutions address food security problems through active engagement.

The widespread impact of rising food prices determines how society will be affected across all aspects of health and education and working levels and social harmony. The R Dashboard functions as an integral part of national resilience frameworks because this tool provides valuable academic tools that advance resilience strategies. These tools create fundamental connections between statistical information and societal choices which enable evidence-based development that includes everyone. Concerted action based on dashboard insights should first serve the food needs of vulnerable groups while building a self-sufficient food system that protects everyone.

10. Key Findings

The analysis of R Dashboard revealed multiple essential discoveries that uncover vital aspects of Sri Lanka's food pricing system and its characteristics and weaknesses. Rice together with wheat flour and other grains displayed a constant upward price trajectory as the main discovery during the analysis phase. Staples such as rice and wheat flour together with other grains deliver dietary requirements to most people particularly among individuals in low- and middle-income living situations. Sri Lankan staple food costs have risen steadily and sometimes dramatically due to both internal inflationary forces alongside agricultural production limitations and farm cost increases together with food transport problems. The patterns emerged across different periods and markets verifying this problem represents a national-scale challenge for the country.

Research data revealed that some parts of Sri Lanka including urban Colombo continuously maintained higher food prices than other rural marketplaces. Various interacting elements help to clarify these price differences. The increased market demand density throughout urban regions leads to higher price levels. The transportation along with storage expenses charged to commodities brought from agricultural regions create additional prices which urban customers must bear. The urban residents who depend on commercial retail face increased financial pressures to pay higher prices than rural consumers since they lack alternative subsistence options. Rural markets close to agricultural regions can obtain lower product prices because these locations possess shorter supply routes enabling direct sales of local produce. Economic and welfare policies need to be specifically designed based on the comprehension of area-based price variations.

The research revealed significant results through the exchange rate effect. The depreciation of Sri Lankan Rupees (LKR) against United States Dollars (USD) seriously diminished food affordability because it increased prices of imported products and globally priced items. The declining national currency value made it necessary for consumers to allocate more finances to acquire equivalent food portions even when the LKR nominal prices remained stable. The food system's susceptibility to worldwide economic changes illustrates why Sri Lanka needs stable macroeconomic policies which include both monetary and fiscal measures to protect its food security.

The dashboard detected recurring price changes that corresponded with normal agricultural cycles together with national festivities while demonstrating possible climate-induced effects. The prices increase during significant national festivals along with post-harvest periods because demand rises simultaneously with restricted supplies. Predictable cycles help build early warning systems as well as develop inventory management strategies that assist the government and private sector to prepare in advance for short-term price disturbances. Seasonal analysis leads households to modify their purchasing behavior before such expected price shifts occur.

The volatility levels between various food items appeared as an important discovery within the research project. The prices of staple food items like rice grew steadily but slowly while flour and lentil costs demonstrated inconsistent patterns because of their reliance on imports and storage difficulty factors and processing requirements. Stakeholders can focus their price interventions and stockpile plans on sensitive food commodities through this information which enables them to improve supply chains as well. Commodities that maintain stable price stability provide chances for businesses to create sustainable market plans together with export business development prospects.

The dashboard presented subdued indications about how economic impacts are distributed among different social classes suggesting uneven food price burden distribution. People with lower incomes face greater food budget distribution because they allocate more money to dietary expenses so even slight price hikes result in significant strain. The dashboard lacked specific data regarding distributional impacts even though its identified patterns suggest rising food accessibility challenges would persist unless governments implement food subsidy programs or community-based support initiatives.

The dashboard provides a comprehensive view which examines various elements that affect food prices in Sri Lanka. A comprehensive analysis shows that the Sri Lankan food market operates under a combination of forces which drive continuous price changes and display regional inequality with currency adaptations while showing seasonal patterns with specific volatile commodities in the mix. The research findings provide crucial knowledge for statistical and academic needs as well as operating power to real-world policy decisions about food safety and poverty elimination and agricultural expansion and trade policies. Sri Lanka can establish a better feeding structure by exploring patterns early and investigating their essential elements.

11. Recommendations

The analytical findings from the dashboard provide the basis for multiple practical and policy-level proposals to overcome Sri Lanka's escalating food prices issues. The government together with its relevant stakeholders should make the development of domestic agricultural production systems their top priority. Modern farming methods combined with irrigation interventions and proper agricultural instruction will enable Sri Lanka to decrease its food imports while improving national food security. National food price monitoring systems need to expand their real-time tracking capabilities throughout Sri Lanka. The systems link up with dashboards including the analyzed one to give alerts about anomalous price increases so authorities can respond promptly. When inflation reaches high levels the program should provide food subsidies and cash payments to poor families to prevent hunger among the population. Data analytics should get increased support for its implementation in policy development. National planning systems need to integrate these dashboards for stakeholders to make decisions based on evidence. The efficiency of supply chains increases by enhancing both warehouse facilities and transportation networks along with market regulatory frameworks which restricts price changes effectively and minimizes food losses. Government institutions along with non-profits and international agencies should collaborate to acquire necessary resources for continuous food security monitoring while developing long-term food security plans.

12. Challenges

The implementation of the R Dashboard for food price data presentation encountered various obstacles throughout the development phase along with the analytical tasks. The dataset presented the initial major problem because of its poor quality standards. Data analysis became impossible until the CSV file underwent cleaning to remove header rows embedded in its data body. The entries all arrived as strings during initial import and required extra preprocessing before conversion into numerical pricing data and time-series date objects were possible. The dataset presented a drawback because it consisted of only 500 records. The dataset limitations prevented performing extensive temporal or seasonal analyses because of its minimal size. The available amount of data did not extend into long-term measurements or provide extensive information across all regions of Sri Lanka. Data updates happened manually which created an additional difficulty for the team. The dashboard constitutes a static tool that needs periodic manual updating through API integration because it lacks an automated data pipeline which hampers real-time monitoring capabilities. The limitations of R Shiny technology and its graphic rendering capabilities probably restricted the potential visualization complexity. A design challenge persisted as the team worked to build an accessible user interface which would suit users spanning from policymakers to academics and the overall public. The implementation challenges emphasize the necessity for strengthened data resources as well as improved data network management systems and thoughtful design approaches that prioritize end-user needs.

13. Conclusion

The Sri Lankan Food Prices R Dashboard along with its detailed evaluation positions it as a vital tool for comprehending food affordability patterns and consumer and marketplace behaviors in the nation. The dashboard transformed fundamental pricing information into visually appealing user-friendly displays and provided an analytical system which revealed essential patterns and discrepancies in food costs throughout time and spatial boundaries and different commodities. This dashboard provided an inclusive perspective of the combined influence that inflation distribution along with economic macro factors such as monetary exchange rates and global logistics disruptions have on Sri Lankans' daily food purchasing abilities.

The dashboard contained visual presentation features which served as interpretation tools for various groups such as experts and decision-makers along with scholars and development specialists. The dashboard's ability to break information down between regional locations and commodity categories and price types LKR and USD exposed both increasing food price trends across time and regional disparities that affected food purchasing ability. The system generated useful information about urban-rural pricing differences and commodity market volatilities together with seasonal trends which policy-makers can utilize for precise action. Analysis demonstrated that food price inflation exists beyond statistical figures because it produces tangible effects on household wellbeing and national nutritional results together with national development progress.

The research investigation confirmed that contemporary food security goes beyond agricultural production for achieving sustained food availability. Modern food security demands highly developed systems for data gathering and current monitoring together with advanced technological instruments for risk prediction and outcome simulation that allow proactive measure development. Through the dashboard digital platforms successfully interpolated complicated information into actionable evidence that allowed swift interventions. An increase in core commodity prices detected by the system would prompt authorities to investigate supply chain issues and long-term inflation patterns could steer decisions about payment incentives and trade patterns as well as public budget modifications.

The tool provided many beneficial accessible insights but unfortunately faced certain shortcomings in its design. Data inconsistencies and the requirement to preprocess data and the static nature of the collected dataset reveal opportunities for betterment. The dashboard falls short as an ongoing decision-support system because it lacks real-time connectivity between data sources while its 500-entry dataset makes it difficult to achieve statistically solid and long-lasting conclusions. The dashboard presents a successful proof of concept because it developed an operational model which can be developed further to establish national food monitoring systems and academic research programs.

Advanced dashboard potency will result from expanding datasets and adding API-driven real-time data and improving user interface capabilities in future development work. The dashboard's effectiveness in multiple sectors can be improved by adding forecasting models and integration with demographic or nutritional data and mobile accessibility features. The usefulness of the dashboard will be maximized through training development workers and policymakers to read and apply its outputs thereby integrating data-driven solutions into existing decision-making practices.

Data science and digital tools have proven effective in addressing critical problems which developing nations currently face according to this project example. The challenges of food security together with economic resilience and resource access need intelligent adaptive systems which match political determination and investment needs so the system can react to changing circumstances. This Sri Lankan Food Prices R Dashboard shows how proper systems must be constructed for public good purposes using well-designed approaches and defined goals. The role of analytical tools including this dashboard will rise in importance for sustainable development among Sri Lanka and other nations as they face economic insecurities and international food crises.

14. References

https://www.worldbank.org/en/programs/icp/brief/foodpricesfornutrition

https://tradingeconomics.com/sri-lanka/food-inflation

https://www.wfp.org/countries/sri-lanka

https://www.wfp.org/publications/hunger-hotspots-fao-wfp-early-warnings-acute-food-insecurity-november-2024-may-2025

Dataset - https://data.humdata.org/group/lka