Name: Harish Satheesh
SRN: PES1UG21CS220

Title: Exploratory Data Analysis of Wholesale Price index of commodities in India

1. Introduction

1.1. About the Project

This project involves the exploration and analysis of the Wholesale Price Index (WPI) dataset, which includes data for approximately 800 commodities over the years 2013 to 2023. The purpose of this analysis is to gain insights into the trends, patterns, and variations in the wholesale prices of these commodities.

The Wholesale Price Index is the price of a representative basket of wholesale goods.

The index number for each year is a percentage and depicts the commodity price increase in percentage from the previous base year. For example, if the base year is 2012, and the index value for a particular commodity for the year 2013 is 109 it implies that the price of the commodity is 109% that of the base price, which means the commodity has seen a price increase of 9%.

Some countries use WPI changes as a central measure of inflation.

1.2. Purpose of the Analysis

The primary goals of this EDA are:

- To understand the overall trends in the WPI over the years.
- To identify the commodities with the highest and lowest average WPI.
- To determine the commodities with the least and most variation in WPI.
- To visualize the data using various plots to facilitate easy interpretation of results.

2. Dataset Overview

2.1. Description of the Dataset

The dataset consists of wholesale price index data for over 800 commodities in India from the years 2011 - 2023. The dataset consists of commodity name , serial number, followed by the index for each year.

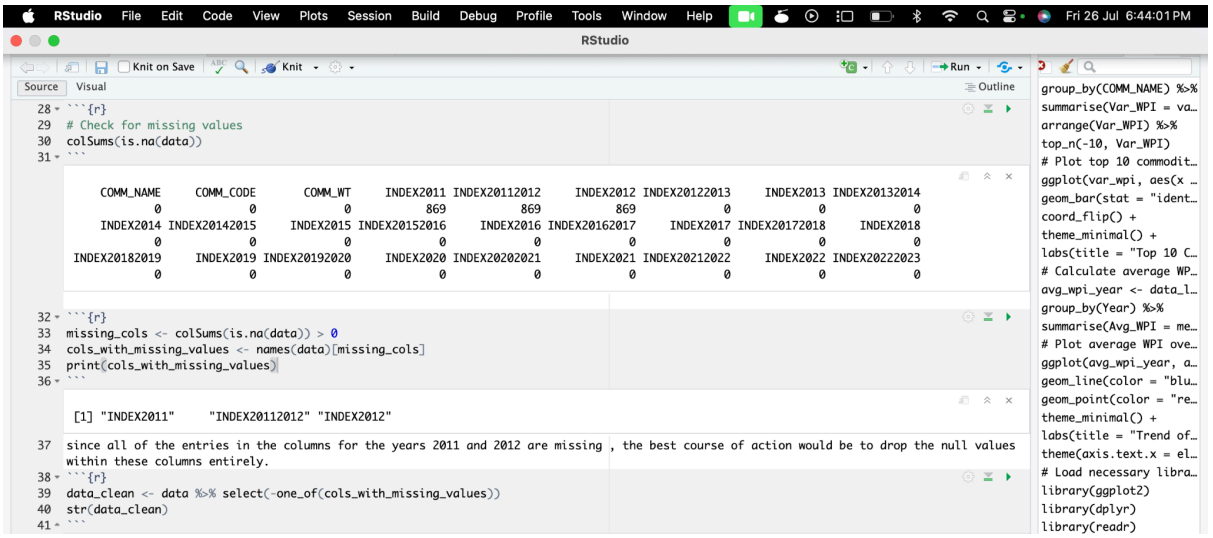
The dataset contains the following columns:

- `COMM_NAME`: Name of the commodity.
- `COMM_CODE`: Unique code for the commodity.
- `COMM_WT`: Weight of the commodity in the index calculation.
- `INDEX20122013` to `INDEX20222023`: WPI values for the respective years.

  Link: https://www.data.gov.in/resource/wholesale-price-index-base-year-2011-12-till-last-financial-year

2.2. Data Preprocessing

- Handling missing values.

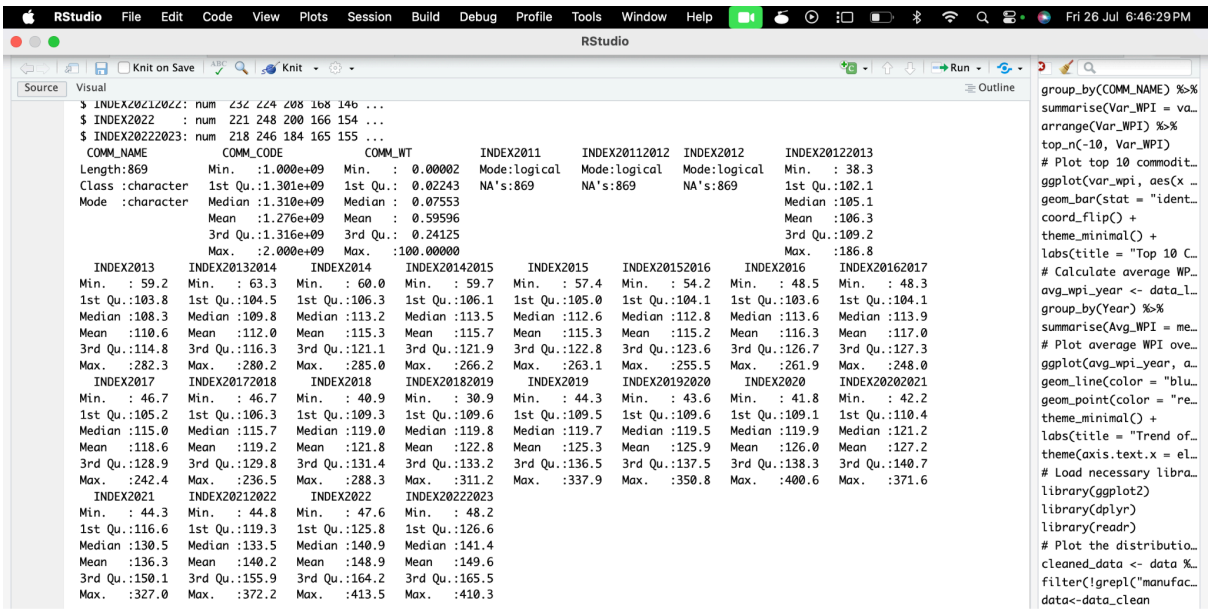


Here, the index values for the years 2011 and 2012 and missing across all commodities. Since all index values are relative to the baseline of these years, We can safely drop these two columns.

3. Exploratory Data Analysis

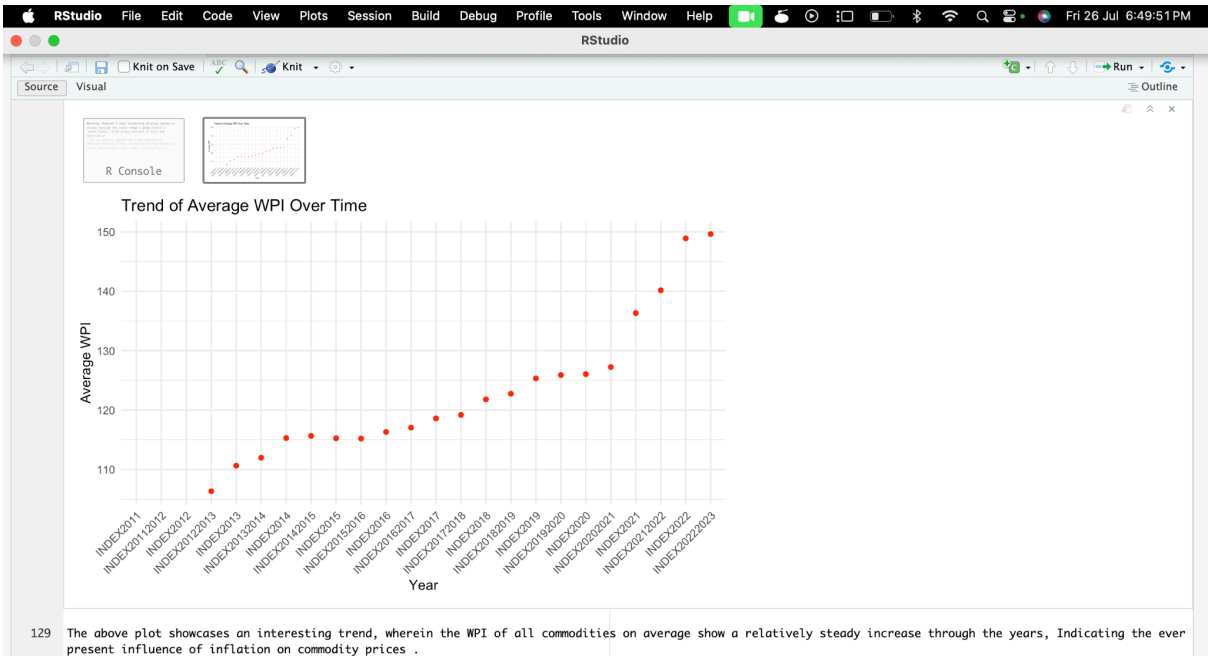
3.1. Summary Statistics

- Overview of summary statistics for the WPI values (mean, median, standard deviation, etc.).



3.2. Trends Over Time

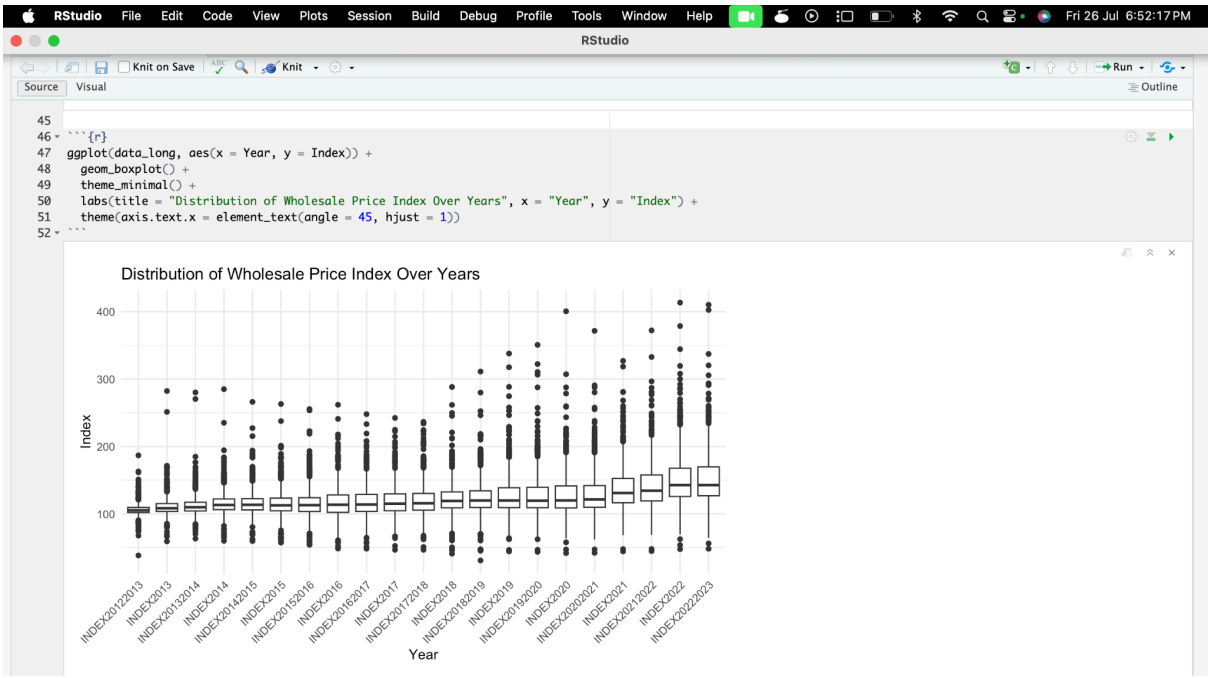
- **Plot**: Line plot showing the average WPI across all commodities over the years.



- **Interpretation**: The above plot showcases a steady increase in the average WPI of all commodities over the decade. This is an indication of inflation and how it affects commodity prices.

3.3. Distribution of WPI

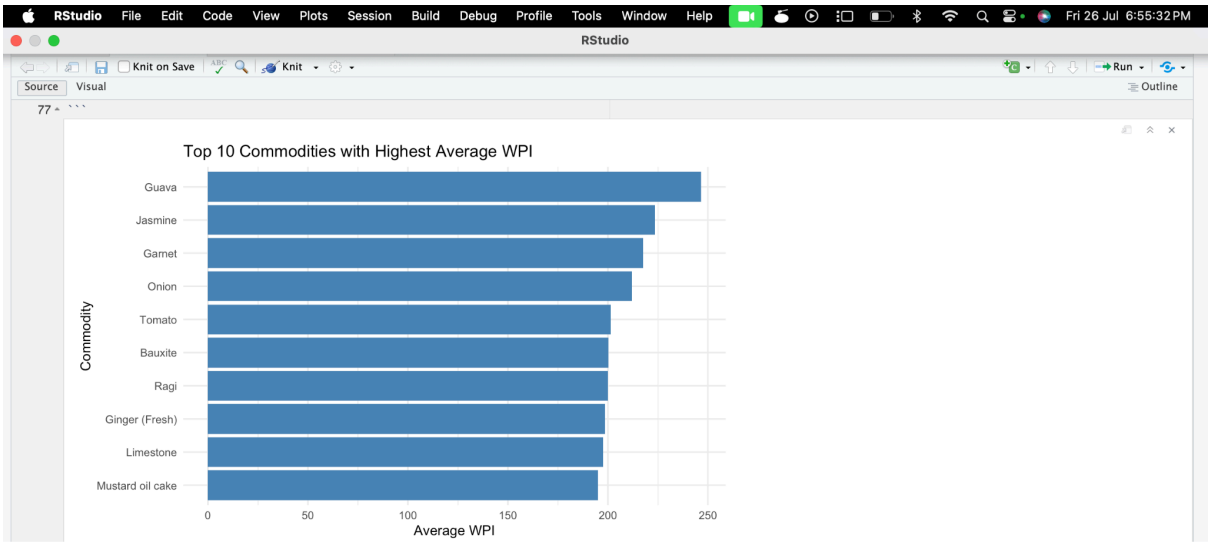
- **Plot**: Boxplot of WPI values for each year.



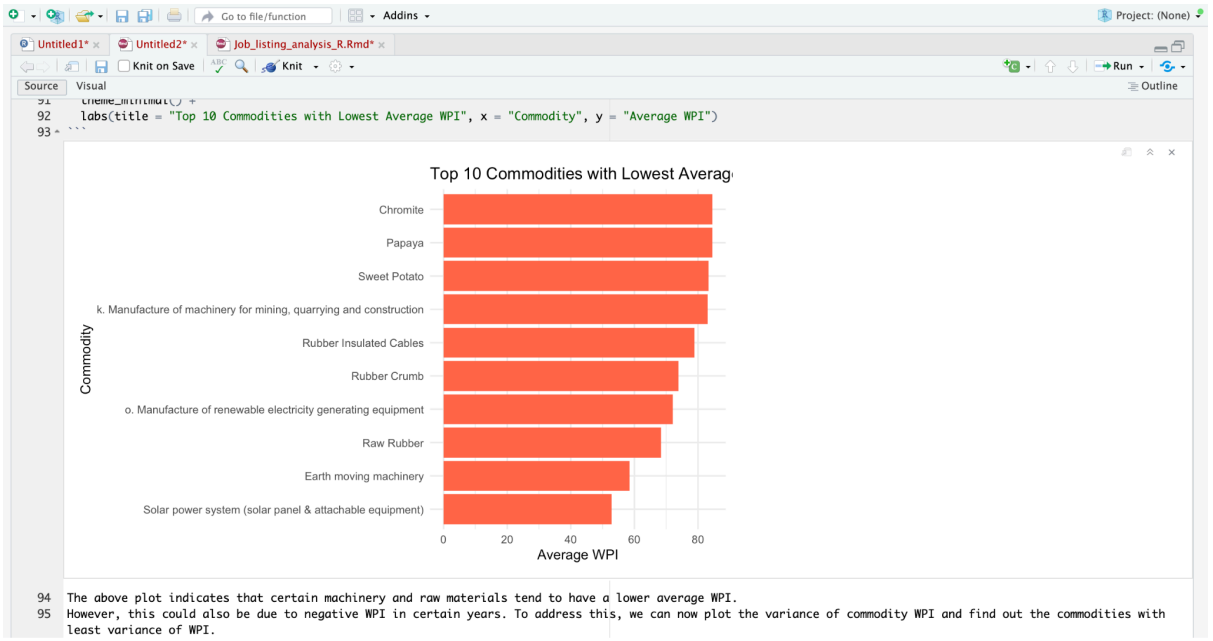
- **Interpretation**: This box plot gives us an understanding of the distribution and variability of WPI values annually, wherein we can observe that the median WPI remains fairly stable while exhibiting a slight increase over the years. The general volatility of the WPI along with the outliers seems to have increased over the decade.

3.4. Commodity-Specific Analysis

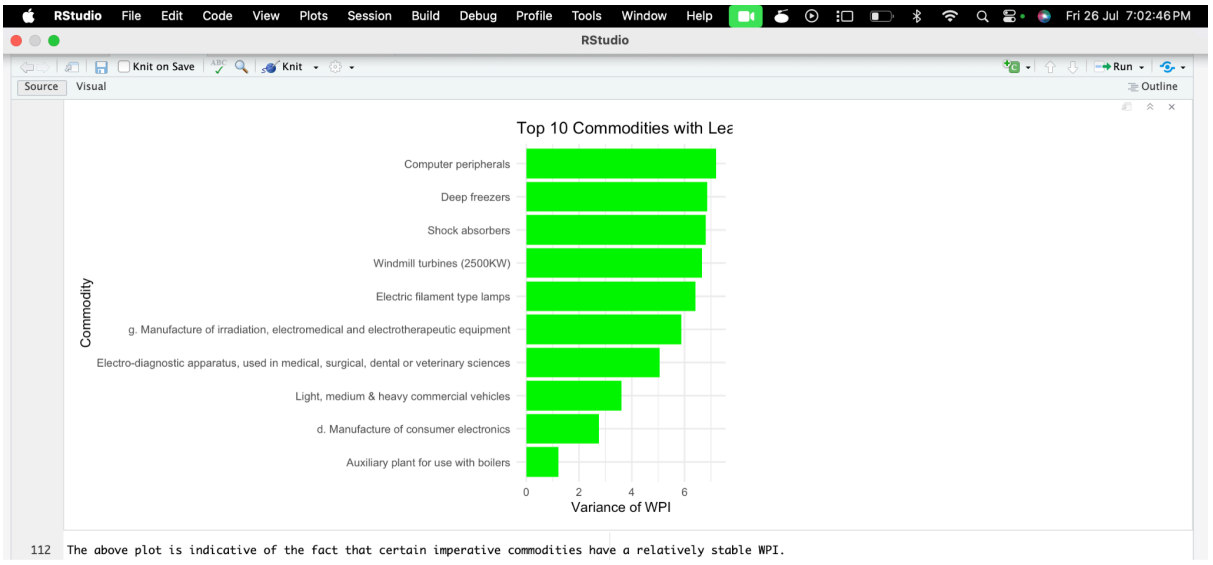
- **Plot**: Bar plots for the top 10 commodities with the highest average WPI.



- **Plot**: Bar plots for the top 10 commodities with the lowest average WPI.



Plot: Bar plots for the top 10 commodities with the least variance in WPI.



4. Insights and Observations

4.1. Key Findings

- There is a dominant presence of agricultural products in the top 10 commodities with the highest average WPI, which are subject to seasonal variations. This potentially contributes to a higher WPI. The higher increase in WPI for these commodities also suggests that they are contributing to overall inflation, since rises in commodities like onions and tomatoes can have a direct impact on cost of living.

- A steady increase in the average WPI indicates a general trend of rising prices for wholesale commodities, reflecting overall inflationary pressures in the economy. This suggests that the cost of goods is increasing consistently over time, as is seen in the markets around us.

- Economic Growth : A steadily increasing WPI may also be associated with economic growth. As the economy expands, there is generally more consumption and production, which can lead to higher prices due to increased demand and greater cost pressures.

- The commodities which have the lowest average WPI over the years indicate factors such as market saturation, wherein there is an abundance of supply and intense competition, and prices tend to remain low. This could be the case for commodities like rubber cables and raw rubber. Perishable goods: For commodities like papaya and sweet potato, their perishable nature might lead to more stable prices as producers aim to sell their produce quickly to avoid spoilage, often accepting lower price increases.