MA 541 FINAL PROJECT

ANALYSIS OF ECONOMIC POLICY USING CONSUMER PRICE INDEX

GROUP – 01

Akanksha Wagh, Nihar Dugade, Varad Hattekar, Vidisha Parmar

Submitted to: Prof. Hong Do

Date: April 30, 2024

STEVENS INSTITUTE OF TECHNOLOGY

SPRING 2024

TABLE OF CONTENTS

INTRODUCTION

In the complex landscape of economic indicators, the Consumer Price Index (CPI) stands as a critical gauge of economic health and inflationary trends within an economy. Specifically, the CPI for All Urban Consumers (CPI-U) provides an essential snapshot of the economic pressures faced by urban households across the United States. Our exploration aims not only to chart the trajectory of consumer prices but also to decode the implications of these changes on the broader economic policy landscape.

The CPI is not just a number, but a multifaceted tool used by economists, policymakers, and businesses to gauge the price evolution of a meticulously selected basket of goods and services. This basket represents the typical consumption patterns of urban consumers and is pivotal in understanding the day-to-day economic reality faced by most of the population. By examining the CPI, we can deduce the inflationary pressures that influence both macroeconomic stability and individual financial wellbeing. Given its profound impact, the CPI is routinely used to adjust pensions, social security benefits, and other crucial financial instruments to shield against the erosive effects of inflation.

Our analysis in this report will delve into the CPI data for January 2024, focusing particularly on the city averages and detailed expenditure categories. We aim to identify trends and patterns that reveal the underlying dynamics of price changes, assess the effectiveness of current economic policies, and predict future economic conditions. In conclusion, this report is structured to illuminate the dynamics of the U.S. economy through the detailed lens of the Consumer Price Index. By dissecting changes in various expenditure categories, we will provide actionable insights that aim to enhance the understanding of economic health and guide informed decisions, thereby supporting the overarching goal of driving sustainable economic growth and stability.

DATA DESCRIPTION

1. Data Source:

The Organization for Economic Cooperation and Development (OECD) collects and validates the data, whereas the International Monetary Fund (IMF) takes care of the collection of data for all other countries. As per project requirement, we only extracted data of United States from October 2023 to January 2024.

2. Data Overview:

This project utilizes the Consumer Price Index (CPI) dataset to analyze economic trends and price movements across various consumer goods and services. The CPI is a vital metric that captures the average change over time in the prices paid by urban consumers for a basket of goods and services. These goods and services are compiled into a representative sample, reflecting typical consumption patterns of urban households, which is crucial for gauging inflationary pressures within the economy.

Our dataset comprises 266 entries, each corresponding to a major expenditure category within the CPI. These categories include essential sectors such as Food, Energy, and Other—encompassing broad utilities like housing, transportation, healthcare, and education. Each category in the dataset is described with specific attributes, including the relative importance as of December 2023 and both unadjusted and seasonally adjusted percentage price changes from November 2023 to January 2024.

The relative importance metric in our dataset provides a weight for each category, indicating its significance in the average consumer's budget. This weighing helps to underscore the impact of price changes in more significant areas of spending. Furthermore, our data distinguishes between unadjusted and seasonally adjusted percentage changes. Unadjusted changes offer a direct view of price fluctuations, whereas seasonally adjusted changes remove the effects of predictable seasonal variations, thus allowing for a clearer analysis of underlying economic trends and patterns.

Through this CPI dataset, we aim to discern inflationary or deflationary trends and assess their implications on both consumer behavior and broader economic policies. This detailed examination will enable policymakers, economists, and businesses to make well-informed decisions, fostering a more robust economic environment.

3. List of Variables:

Expenditure Category (String): Category includes consumer goods and services in Food, Energy, Education, Transportation sector and more.

Relative Importance of December 2023: The relative importance metric in our dataset provides a weight for each category, indicating its significance in the average consumer's budget. This weighing helps to underscore the impact of price changes in more significant areas of spending. Understanding the relative importance of different categories or items in the CPI helps analysts and policymakers interpret the index and assess the impact of price changes on overall inflation and consumer purchasing power.

Class (String): All the items from expenditure category are divided into 3 major categories- Food, Energy and Other.

Subclass and Subclass Division (String): The item is further divided into subcategories- Healthcare, Housing, Electricity, Transportation, Apparel, education and more.

Unadjusted percent change (Float): This metric captures the raw change in prices over a specified period, reflecting the actual fluctuations in costs without adjustments for external influences. It represents the direct price movements without modifications for seasonal variations, changes in the composition of the consumer goods basket, or other potential distortions. The data for the unadjusted percentage change was collected from January 2023 to January 2024, providing a straightforward view of price dynamics during this interval.

Adjusted percent change (Float): This measure accounts for various factors that can influence price changes, including seasonal variations, shifts in product quality, substitution effects (where consumers switch between products in response to price changes), and other relevant variables. The seasonally adjusted percentage change offers a clearer representation of underlying price trends by filtering out or mitigating the impact of these elements. Data for the seasonally adjusted percentage change was meticulously compiled from October 2023 to January 2024, ensuring an accurate depiction of the economic landscape during this period.

4. Outliers:

The series of box plots illustrate the distribution of percentage changes—both unadjusted and seasonally adjusted—from January 2023 to January 2024. These plots highlight the central tendency and variability within each period, with most data points clustering near the median, indicating stable changes over time. Notably, outliers are present in each period, suggesting occasional significant deviations from typical price changes. These outliers, especially prominent in the unadjusted data, underscore potential volatility in the

market or effects of external shocks on pricing, which the seasonal adjustments aim to mitigate, providing a clearer view of underlying economic trends.

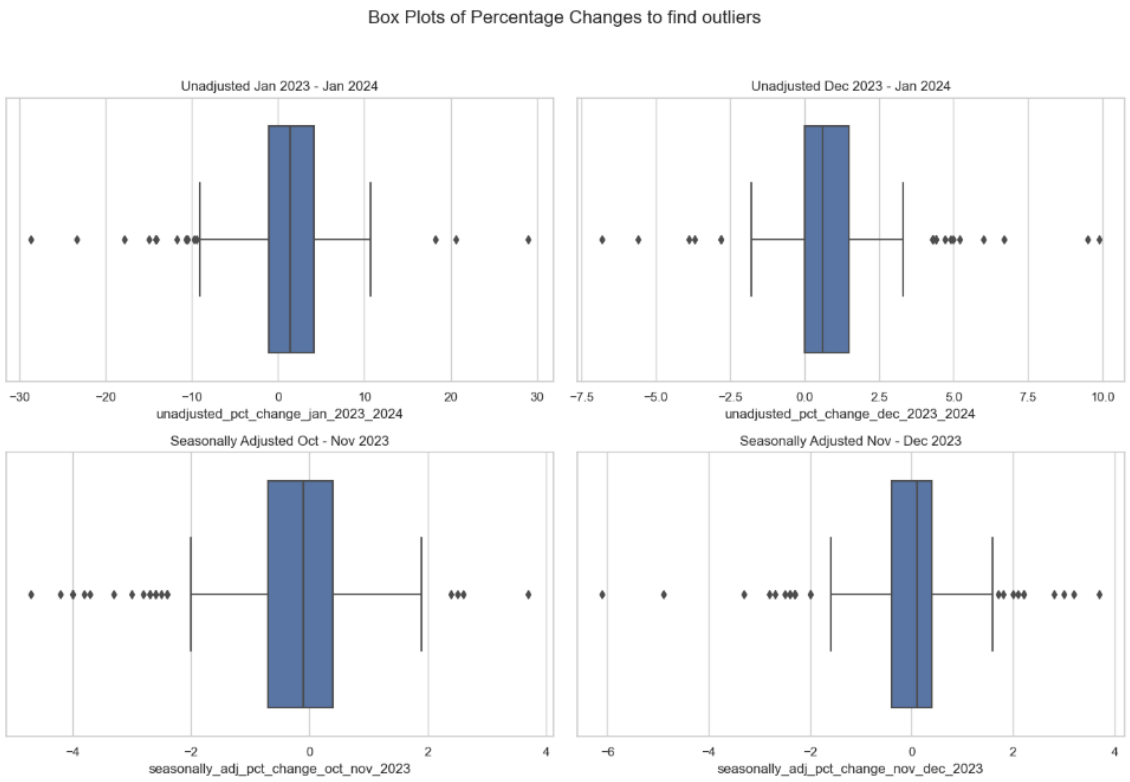

Figure 1. Analysis of Percentage Changes and Outliers

We will not eliminate these outliers, as we can see from Figure 1. there are not many points which lie outside of the clusters, and these outliers are dependent on each other since they might belong to the same consumer goods, commodities, or services.

SUMMARY STATISTICS

	Index	Relative importance Dec. 2023	Unadjusted percent change Jan. 2023- Jan. 2024	Unadjusted percent change Dec. 2023- Jan. 2024	Seasonally adjusted percent change Oct. 2023- Nov. 2023	Seasonally adjusted percent change Nov. 2023- Dec. 2023	Seasonally adjusted percent change Dec. 2023- Jan. 2024
count	265.0	265.000000	265.000000	263.000000	263.000000	263.000000	263.000000
mean	1.0	1.856894	1.076604	0.785551	-0.260076	-0.019392	0.268061
std	0.0	7.341149	5.646397	1.745818	1.126453	1.058176	1.384858
min	1.0	0.008000	-28.600000	-6.800000	-4.700000	-6.100000	-4.600000
25%	1.0	0.122000	-1.100000	0.000000	-0.700000	-0.400000	-0.300000
50%	1.0	0.294000	1.400000	0.600000	-0.100000	0.100000	0.400000
75%	1.0	0.854000	4.200000	1.500000	0.400000	0.400000	0.900000
max	1.0	79.790000	29.000000	9.900000	3.700000	3.700000	9.900000

Figure 2. Summary statistics for all attributes

The substantial fluctuations in unadjusted percent changes underscore the impact of external economic factors, such as supply chain disruptions or shifts in demand, that can significantly affect consumer prices. Meanwhile, the seasonally adjusted data typically present a more subdued variation, which may better reflect underlying economic trends by filtering out seasonal effects and other distortions.

These statistics not only highlight the dynamic nature of the consumer price index but also validate the effectiveness of seasonal adjustments in providing a clearer picture of the underlying economic conditions. As policymakers and analysts examine these trends, it is crucial to consider both unadjusted and adjusted data to gauge the real impact of economic policies and market conditions on consumer spending.

DATA VISUALIZATION

1. Analysis of Class-Wise Distribution of CPI Items

The pie chart provides a clear visual representation of the class-wise distribution of expenditure categories within the CPI as of December 2023. The largest portion of the chart, representing 'Other' items, constitutes 64.2% of the index, highlighting a diverse range of goods and services such as healthcare, housing, apparel, and education, underscoring their significant combined impact on the consumer budget. Food, vital for everyday sustenance, comprises 31.7% of the index, reflecting its crucial role in consumer spending.

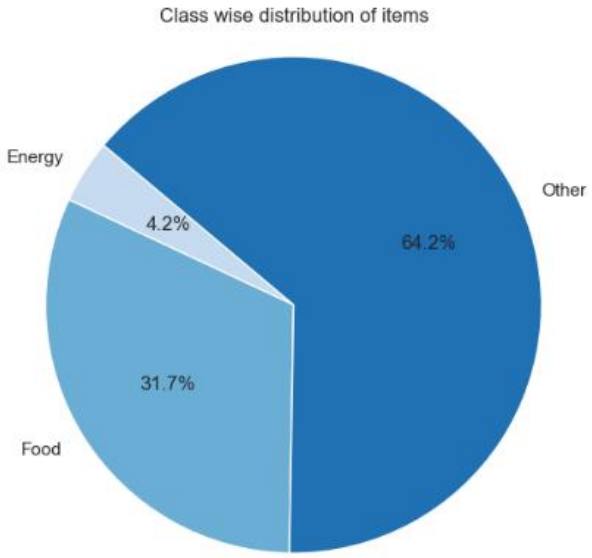

Figure 3. Analysis of Class-Wise Distribution of CPI Items

Energy, though essential, represents the smallest segment at 4.2%, indicating its relatively lower direct cost impact compared to other categories. This distribution aids in understanding how different sectors contribute to overall inflation and consumer spending, essential for economic analysis and policymaking.

2. Analysis of Distribution in the "Other" Expenditure Category

The bubble chart provides an insightful visual distribution of the "Other" category across various subclasses for December 2023. From Figure 2. we could see maximum data points are present in ‘Other’ category. Significant emphasis is placed on 'Shelter', 'Transportation services', and 'Recreational services', each represented by larger bubbles, indicating their substantial relative importance in consumer budgets. Subclasses such as 'Education and

communication', 'Medical care services', and 'Household furnishings and supplies' also show notable contributions but to a lesser extent.

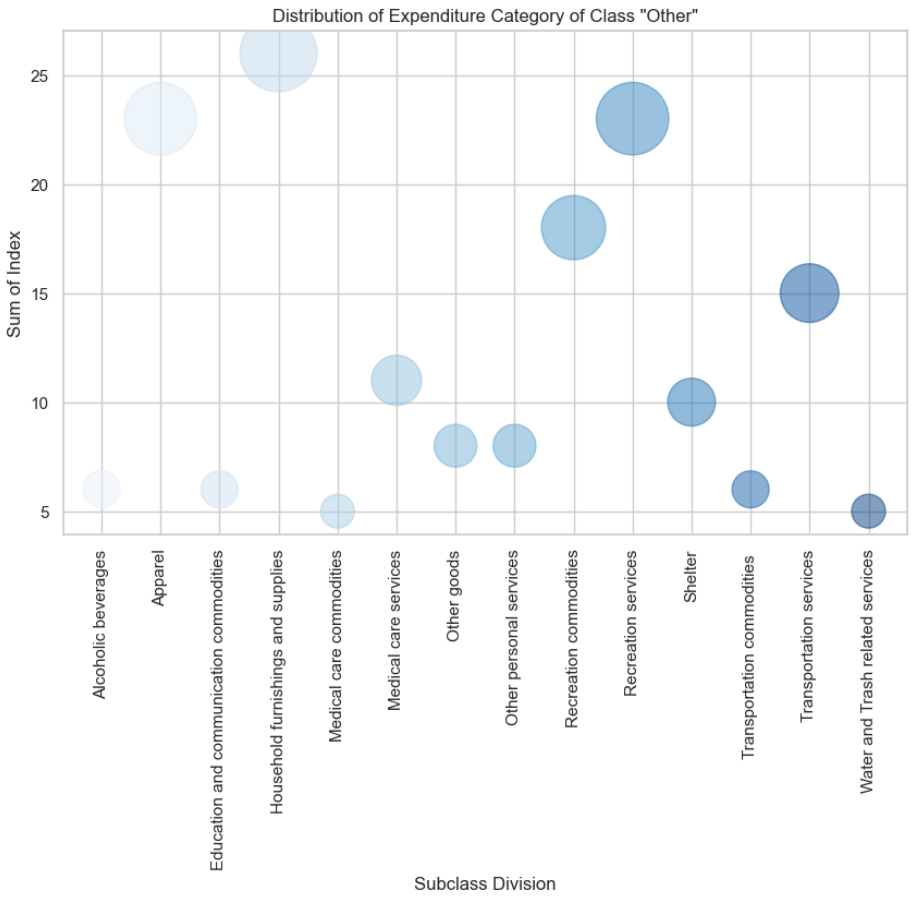

Figure 4. Analysis of Distribution in the "Other" Expenditure Category

The variation in bubble sizes across the chart highlights the diversity within the "Other" category, reflecting the broad spectrum of consumer expenditures that go beyond necessities like food and energy. This visualization effectively captures the varied impact of different services and goods on overall inflation and consumer spending, providing a detailed perspective on consumer priorities and economic conditions.

3. Analysis of Seasonally Adjusted Percentage Changes (October-November 2023)

The histogram overlaid with a density plot reveals a normal distribution of seasonally adjusted percentage changes in consumer prices, centered slightly above zero. This pattern indicates that most expenditure categories experienced minor price adjustments during this period, with the bulk of changes being modest. The range of these adjustments extends from approximately -4% to +4%, highlighting a few significant outliers on both ends.

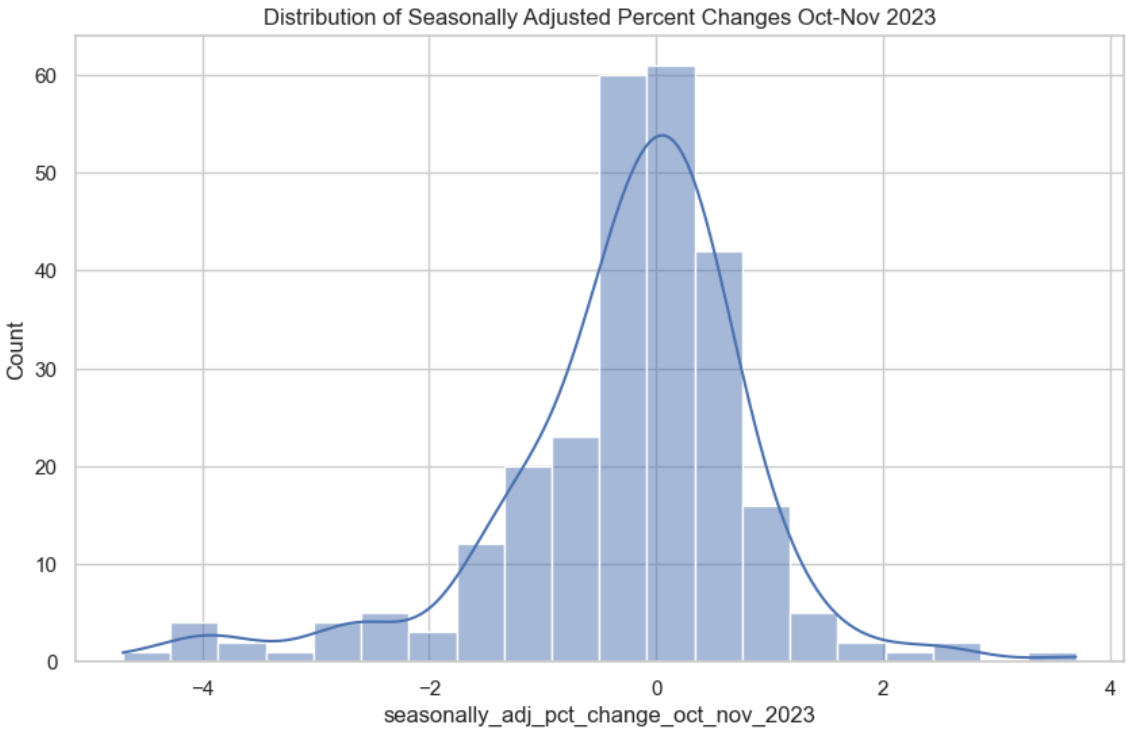

Figure 5. Analysis of Seasonally Adjusted Percentage Changes (October-November 2023)

A slight leftward skew suggests a prevalence of price reductions over increases, possibly due to seasonal economic influences. This graph provides key insights into the variability of price changes, crucial for economic analysis and decision-making related to consumer price sensitivity and inflation during the autumn of 2023.

4. Analysis of the Relationship Between Relative Importance and Percent Change

The scatter plot presents the relationship between the relative importance of expenditure categories as of December 2023 and their corresponding unadjusted percent changes over the year. The plot reveals a concentration of data points with low relative importance, mostly fluctuating between significant decreases and increases in prices. As the relative importance increases, the variability in percent changes tends to decrease, albeit with a few exceptions at higher importance levels, suggesting a stability in prices of more significant categories. Notably, categories with very high relative importance show minimal fluctuations, indicating controlled price changes in essential consumer sectors. This visualization underscores the inverse correlation between the magnitude of price changes and the relative importance of categories, highlighting how crucial items are shielded from volatile price swings, thereby aiding in economic stability and predictability.

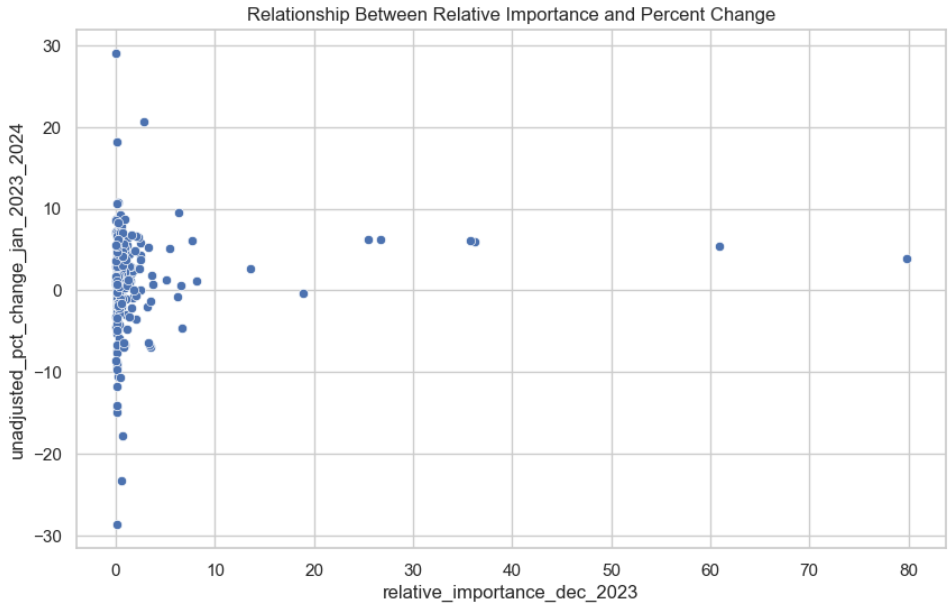

Figure 6. Analysis of the Relationship Between Relative Importance and Percent Change

5. Comparative Analysis of Seasonally Adjusted Percent Changes in Food and Energy Prices

The scatter plots display the seasonally adjusted percent changes for food and energy from October 2023 to January 2024, capturing monthly and yearly fluctuations. The data points are spread across three distinct periods: October-November 2023, November-December 2023, and a comprehensive span from January 2023 to December 2024.

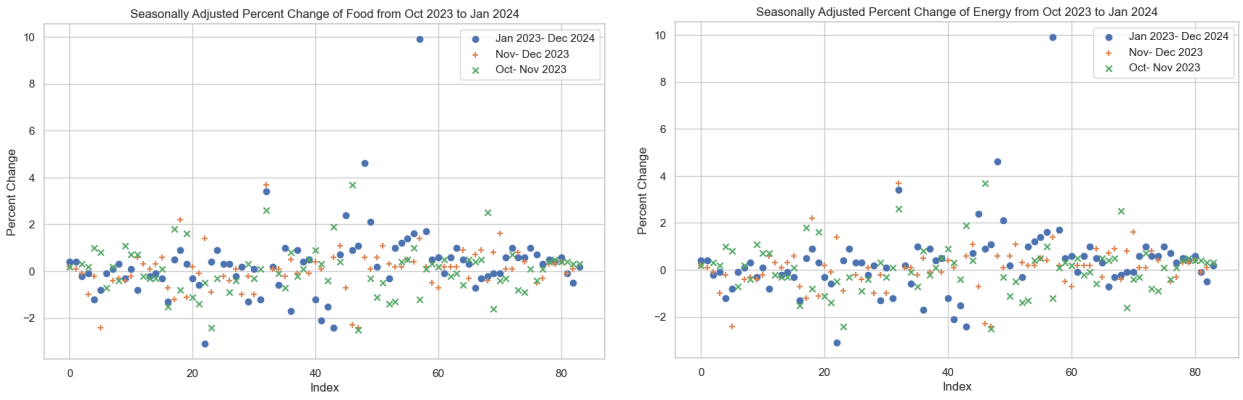

Figure 7. Comparative Analysis of Seasonally Adjusted Percent Changes in Food and Energy Prices

Both graphs reveal a general trend of stability in short-term monthly changes with most data points clustering near the zero mark, indicating minor fluctuations in both food and energy

prices. This suggests effective regulatory or market mechanisms that maintain price stability on a month-to-month basis. However, the yearly data, particularly for food, shows several outliers with substantial increases, indicating occasional spikes likely driven by seasonal trends, supply chain disruptions, or geopolitical events affecting food prices more significantly than energy.

In contrast, the energy sector appears to exhibit slightly higher volatility with a wider spread of data points, especially in the year-long overview. This could be attributed to the energy market's sensitivity to international political and economic conditions, which can result in more pronounced price fluctuations.

These visualizations not only underline the differences in volatility between food and energy sectors but also highlight the importance of considering seasonal adjustments and longer-term trends when analyzing consumer price indices. Such insights are crucial for policymakers and analysts focusing on inflation control and economic planning.

6. Comparative Analysis of Seasonally Adjusted and Unadjusted Percent Changes Across Expenditure Categories

The three-line graphs illustrate the dynamics of seasonally adjusted and unadjusted percent changes in food, energy, and other expenditure categories over a period from December 2023 to January 2024. Each graph presents two lines representing the adjusted and unadjusted percent changes, offering insights into how external factors and seasonal adjustments impact the reported price changes.

Food: The graph shows significant volatility in food prices with sharp peaks, particularly in the unadjusted data, which smoothens considerably in the adjusted line. This indicates substantial seasonal effects and external factors influencing food prices, which are mitigated in the seasonally adjusted data to provide a clearer view of underlying trends.

Energy: Energy prices exhibit less volatility compared to food, with both adjusted and unadjusted lines closely mirroring each other. However, there are noticeable adjustments that slightly smooth out peaks and troughs, reflecting minor seasonal influences.

Other Items: This category shows moderate fluctuations with occasional spikes in the unadjusted data, which are again smoothed in the adjusted view. The adjustments appear to reduce the extremity of changes, suggesting that these items are also subject to external variabilities, albeit less so than food.

Overall, these visualizations underscore the importance of seasonal adjustment in understanding economic indicators accurately. They reveal how raw data can overstate or understate trends due to external and seasonal factors, which the adjustments aim to

neutralize. This analysis is crucial for policymakers and analysts in assessing the true impact of economic policies and market conditions on consumer prices across different sectors.

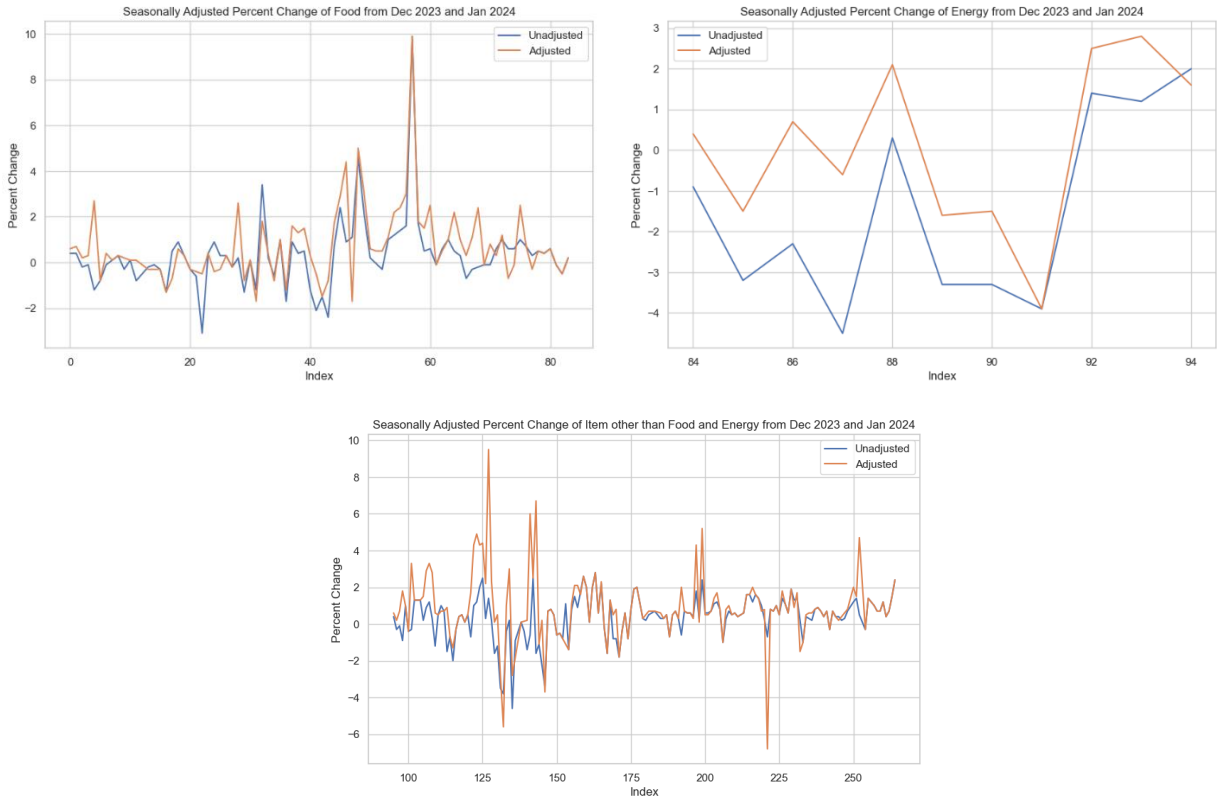

Figure 8. Comparative Analysis of Seasonally Adjusted and Unadjusted Percent Changes Across Expenditure Categories

CORRELATION ANALYSIS

The correlation heatmap presented in the analysis provides a detailed visual representation of the interrelationships between various expenditure categories, such as Energy, Food, and Utilities. This visualization is integral to understanding how price changes in one category can influence others, offering valuable insights for economic analysis and decision-making.

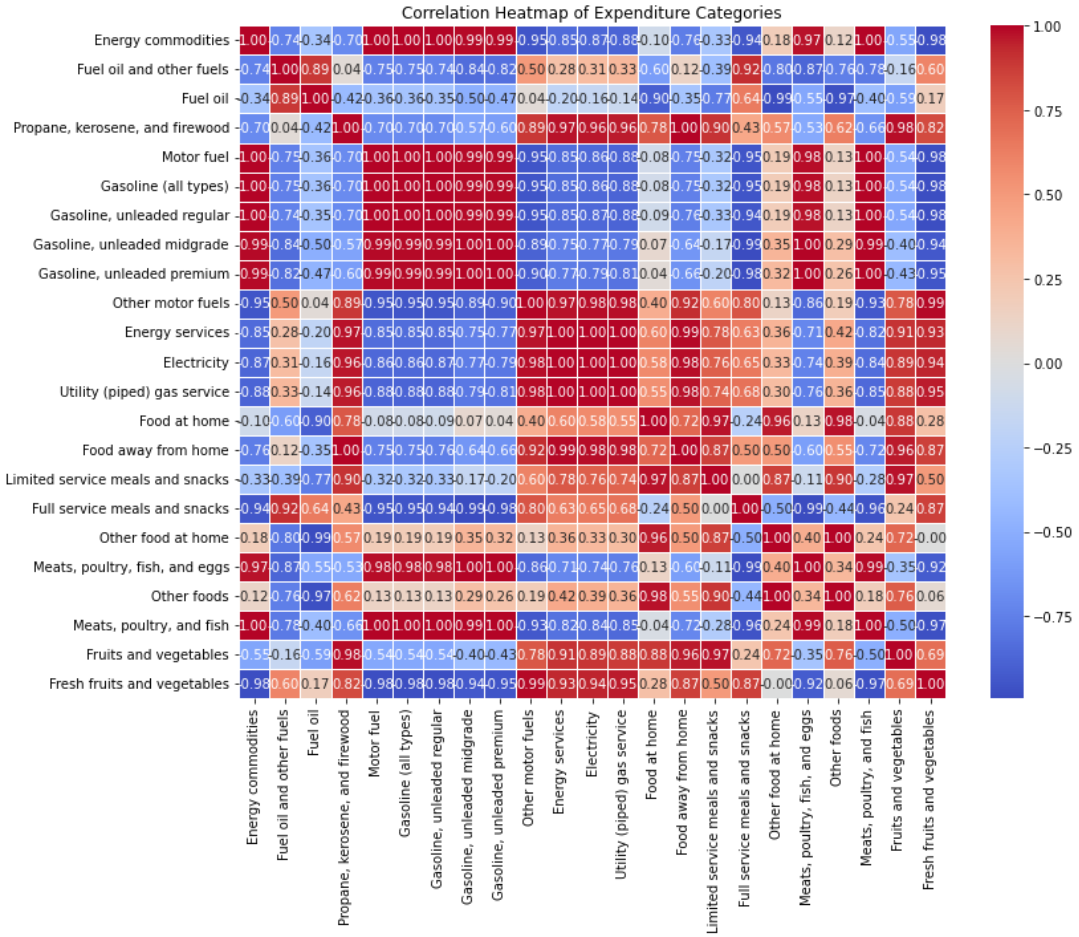

Figure 9. Correlation between expenditure categories, such as Energy, Food, and Utilities

Key Observations:

1. Energy Sector Correlations:

The heatmap shows a very highly positive correlation among different types of motor fuels and gasoline categories (regular, midgrade, premium). This suggests that price fluctuations in these subcategories are tightly linked, likely due to common underlying factors such as global oil prices, taxation, and refining costs. For instance, correlations close to 1.00

between regular and premium gasoline types indicate that any change in the price of one is almost perfectly mirrored by the other.

2. Negative Correlations Between Energy and Food:

Notably, there are significant negative correlations between motor fuels and food categories like fresh fruits and vegetables. This pattern suggests that increases in fuel prices might lead to higher transportation costs, which could negatively affect the prices of perishable goods, making them more expensive and possibly reducing their consumption.

3. Utility Services:

Electricity and piped gas services exhibit strong positive correlations with each other but show negative correlations with motor fuels. This indicates divergent pricing factors affecting these services compared to those affecting fuel prices, possibly due to different regulatory and market dynamics.

HYPOTHESIS TESTING

1. ANOVA for Expenditure Category: Food

In the analysis of seasonal adjustments in food expenditure categories over three consecutive time periods (October-November 2023, November-December 2023, and December 2023-January 2024), an ANOVA test was conducted to determine the statistical significance of the differences observed across these periods. The null hypothesis posited no significant variation in seasonally adjusted percent changes among these intervals.

The ANOVA results yielded an F-statistic of 2.344 and a p-value of 0.105. The F-statistic indicates the ratio of the variance between the groups to the variance within the groups, and in this context, suggests a moderate distinction between the mean adjustments across different months. However, the p-value exceeds the common alpha level of 0.05, which implies that the observed differences are not statistically significant at the conventional 95% confidence level.

Given this outcome, the null hypothesis that there is no significant difference in the seasonally adjusted percent changes between expenditure categories or time periods cannot be rejected. This suggests that while there are variations in the seasonally adjusted percent changes across the different periods, these differences are not sufficient to assert statistical significance.

Overall, the consistency in seasonal adjustments suggests a stable inflationary environment for food-related expenditures during this period, allowing for predictable consumer spending patterns and facilitating more straightforward economic planning and forecasting.

2. ANOVA for Expenditure Category: Energy

ANOVA analysis for the Energy sector expenditure categories over three consecutive time periods—October to November 2023, November to December 2023, and December 2023 to January 2024—yields significant insights. The test was aimed at detecting whether there are statistically significant differences in the seasonally adjusted percent changes across these periods.

The results of the ANOVA test revealed an F-statistic of 0.0346 and a p-value of 0.966. The very low F-statistic indicates that the variance among the group means is not significantly greater than the variance within the groups. Correspondingly, the high p-value, far exceeding the typical alpha level of 0.05, suggests that we fail to reject the null hypothesis. This implies

that there are no statistically significant differences in the seasonally adjusted percent changes across the examined time periods for the Energy sector.

This outcome suggests that the price changes within the energy sector are relatively stable over the studied periods, despite the inherent volatility associated with energy commodities. For policymakers, this stability implies that external factors during this period did not significantly disrupt energy prices, allowing for predictable budgeting and planning. For analysts, these findings underscore the importance of looking beyond short-term fluctuations to understand the broader trends in energy pricing. This analysis contributes to a more nuanced understanding of the economic landscape, particularly in how energy prices react to seasonal adjustments and potential external disturbances.

3. Chi Square Test for Food and Energy:

In the analysis of the association between expenditure categories and the months using the Chi-square test of independence, the results indicated no significant relationship. The Chi-square statistic was notably low at 0.0346, and the corresponding p-value was exceedingly high at 0.966. These statistics led to the failure to reject the null hypothesis, suggesting that there are no significant associations between the different expenditure categories and the months considered in the study.

This outcome implies that variations in expenditure across different categories do not significantly depend on the month. This stability is critical for economic analysis and forecasting, as it suggests that consumer spending in these categories is not subject to substantial changes over the months included in the study. This analysis confirms the consistency in consumer behavior across different months, underscoring the reliability of using such data for ongoing economic assessments and planning. This finding supports the utility of the Consumer Price Index (CPI) data in gauging economic stability and consumer confidence without the need for adjustments for temporal variations in the specific expenditure categories studied.

Since we could not come up with any significant conclusions from Chi-Square Test, we decided to implement T-test instead.

4. T-test Analysis

In our analysis of price changes across different months using paired t-tests, the statistical findings reveal varying levels of significance which are crucial for understanding seasonal price dynamics. Notably, the comparison between October-November and November-December periods demonstrates an insignificant difference (t-statistic: 0.174, p-value: 0.864), indicating stability in price changes across these months. This suggests that

consumer price variations were minimal and did not significantly deviate during the early quarter of the observed period.

Conversely, the November-December and December-January comparison yields a significant t-statistic of -2.986, with a corresponding p-value of 0.007, pointing to a statistically significant difference in price changes. This finding highlights a notable increase in price fluctuations as the year progresses, which may be attributed to factors such as holiday season effects, year-end sales activities, or other economic variables affecting consumer prices.

The comparison between October-November and December-January presents a borderline significance with a t-statistic of -1.834 and a p-value of 0.080. Although not definitively significant, this result suggests a potential trend in increasing price variability towards the end of the period studied.

These insights from the t-test analysis provide a nuanced understanding of the temporal stability and variability in consumer prices, indicating specific periods of significant change that could influence economic decisions and policy considerations. Such analyses are essential for crafting targeted economic policies that accommodate seasonal fluctuations and ensure economic stability.

PREDICTION MODEL BUILDING

In the comparative analysis of different regression models—Support Vector Machine (SVM), Linear Regression, and Lasso Regression—on the seasonally adjusted percentage changes from December 2023 to January 2024, several key insights emerge regarding their predictive accuracies and coefficient behaviors.

Model	MSE
SVR	0.850702
Linear Regression	0.976038
Lasso Regression	0.913847

Figure 10. Model MSE Comparison

The SVM model shows a robust fit, as evidenced by the lowest Mean Squared Error (MSE) of 0.850702 among the models tested. This suggests that the SVM model, with its regularization and kernel tricks, is most effective at capturing complex patterns in the data, leading to more accurate predictions. The coefficient values for SVM highlight the significant role of specific predictors, with a major coefficient of 1.310117, suggesting a strong influence on the model's output.

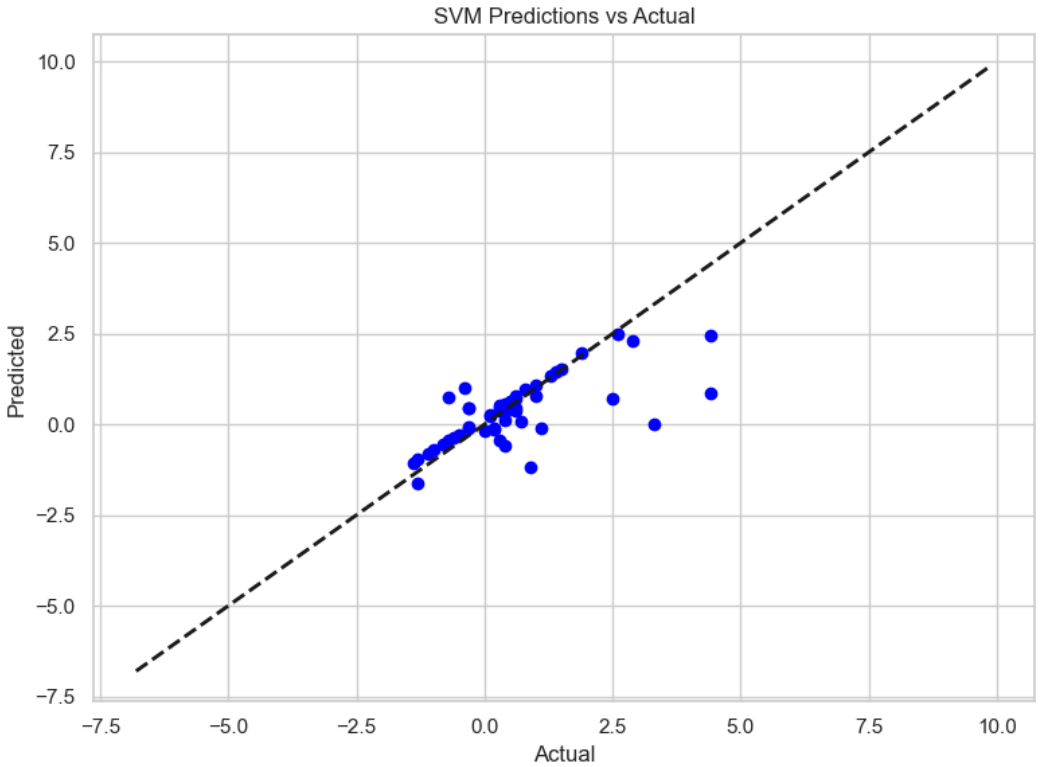

Figure 11. SVM Prediction vs Actual

On the other hand, the Linear Regression model, often the baseline for predictive modeling, displayed a slightly higher MSE of 0.976038 with R^2 Score of 0.21. Its coefficients, notably lower than those in SVM, indicate a relatively simpler model which, while easy to interpret, lacks the predictive power seen in SVM due to the absence of regularization, which helps in reducing overfitting.

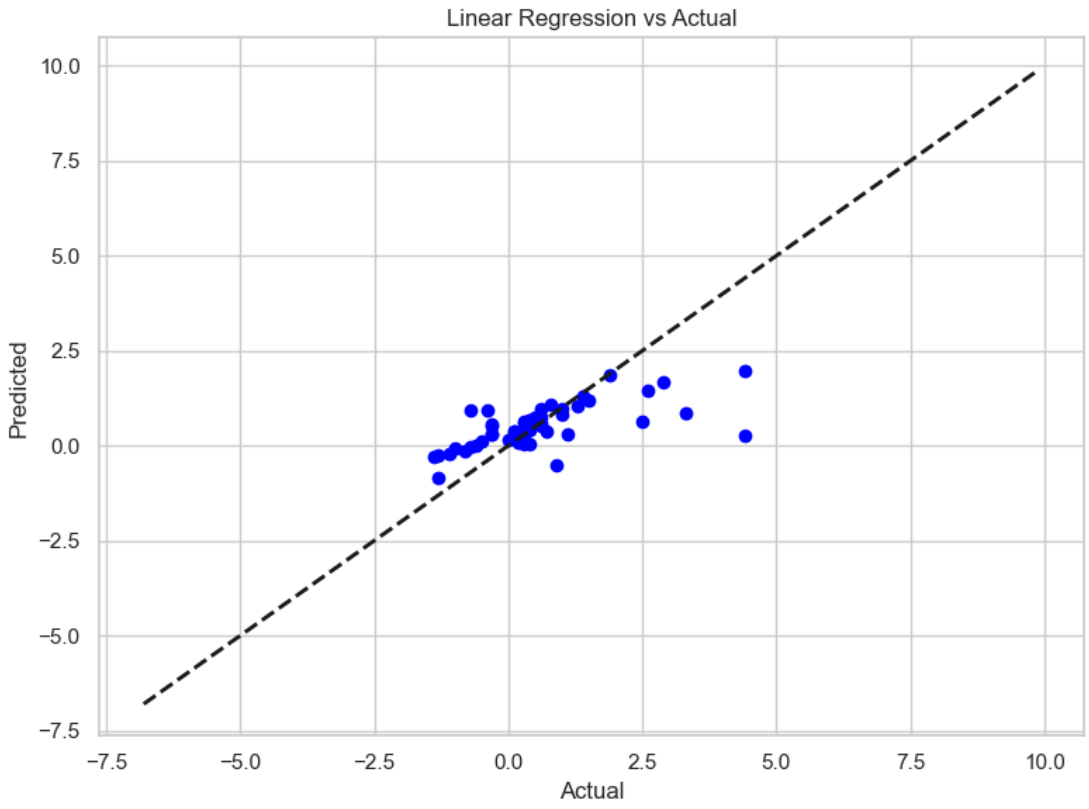

Figure 12. Linear Regression Prediction vs Actual

Lasso Regression, known for its ability to perform feature selection by shrinking coefficients, resulted in an MSE of 0.913847. The Lasso model penalizes less significant features more heavily, resulting in some coefficients being driven to zero, as seen with a significant drop in coefficient magnitude compared to Linear Regression. This leads to a simpler model that retains only the most significant predictors.

The scatter plots comparing predicted values against actual values for all three models closely follow the ideal line (45-degree line), indicating good model performances. However, the concentration and distribution of data points around this line vary slightly among the models, with SVM demonstrating the tightest clustering, thereby underscoring its superior predictive ability.

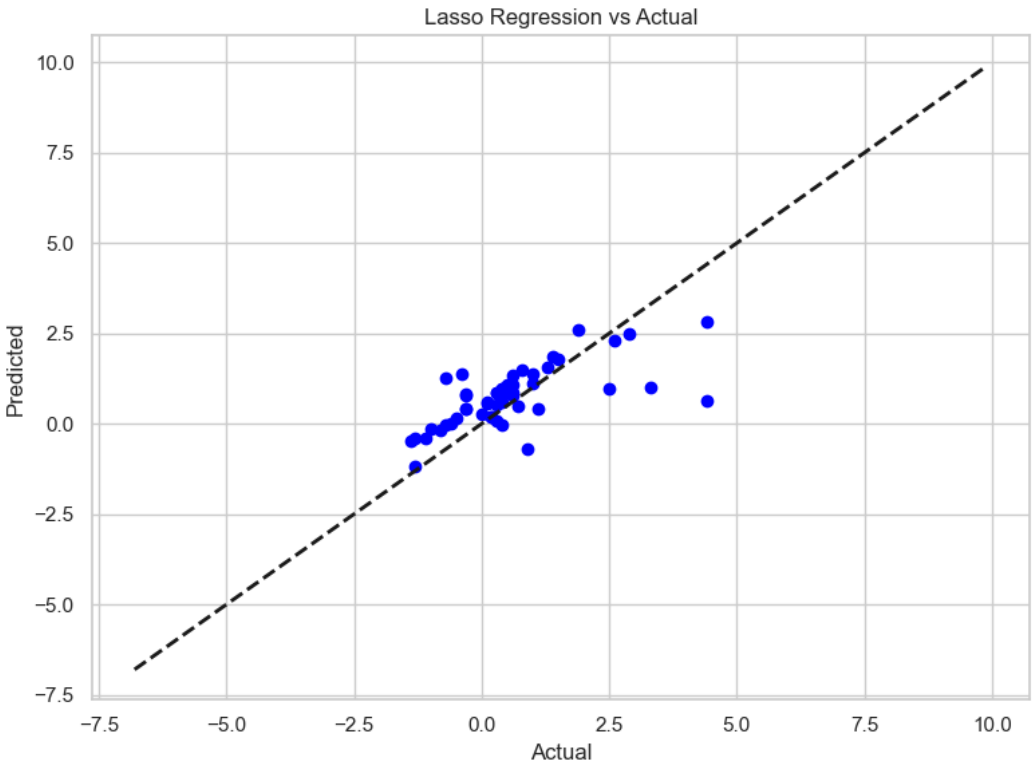

Figure 13. Lasso Regression Prediction vs Actual.

In summary, while all models provide valuable insights, the SVM stands out for its accuracy and robustness in handling complex datasets, making it particularly suitable for datasets with intricate relationships and higher dimensionality.

CONCLUSION

In the case of United States, energy sector always acted as a game changer and this data analysis proved the same. The government and policy makers can control fuel prices and regulate them according to the revenue needs of the country and spending capacity of the average consumer. It can be said that percentage changes in prices of energy sector commodities can lead to overall inflation/deflation in the country.

Higher energy prices led to increased transportation expenses, which were passed on to consumers in the form of higher food prices. Additionally, businesses in other sectors reliant on transportation, such as retail and manufacturing, also faced higher operational costs.

The production of food is energy-intensive, requiring fuel for machinery, irrigation, and fertilizers. Fluctuations in energy prices can directly affect the cost of agricultural inputs, including fuel and electricity.

The percentage changes in all the sectors were affected heavily due to percentage changes in the energy sector. The energy sector served as an underlying reason for positive and negative percentage changes in prices of the various goods.

From the data of the past 3 months, it can be concluded that the economy is slowly trying to recover from volatile fluctuations in the price changes of the commodities.

Policymakers can leverage these insights to design comprehensive strategies that mitigate the impact of fuel price volatility on other critical sectors, particularly food. Implementing subsidies or adjusting taxes on fuel during periods of high volatility could help stabilize food prices, especially for perishables.

Businesses, especially those in logistics and retail, can use this data to anticipate cost changes and adjust pricing or supply chain strategies accordingly. For example, a logistics company might plan for alternative routes or consolidate shipments to manage fuel costs more effectively.

Consumers benefit from understanding these correlations for better budgeting and spending strategies, especially in anticipating changes in food and utility costs based on trends in fuel prices.

APPENDIX

Used Equations:

1. T-test for single sample

t-statistic = $(\bar{x} - \mu) / (s/\sqrt{n})$

2. χ2 test

χ2 – statistic = $[(n-1)* s2] / \sigma 2$

Dataset:

IMF Website: https://data.imf.org/?sk=388DFA60-1D26-4ADE-B505-A05A558D9A42&sId=1479329132316