Data Visualization
Kevin Zeng

Project #4: Interactive Visualization using Python

The Consequences of Hyperinflation

Introduction

Over the past decade, due to several severe issues across the globe, such as the COVID-19 pandemic lockdown, Russian-Ukraine war fears, shortage of labor force, hyperinflation becomes a big concern for people around the world because all these issues impact people's daily lives. Recently, the Russian-Ukraine war causes crude oil price increases rapidly since people are afraid there is a shortage of crude oil and natural due to policy changes and sanctions. There is a correlation between oil price and inflation, which means high oil price directly causes high inflation (Lioudis, 2022).

In this project, we will investigate the consequences and trends of high crude oil prices and gasoline prices from 1995 to the present. There are three data visualizations that will be generated to observe the patterns of trend and correlation between them, and they are the bar chart for gasoline prices, the line plot for crude oil prices, and the scatter plot for inflation rates.

Datasets

The dataset of crude oil price includes the dates, crude oil price, percent change, and change from 1995 to the present by month. The dates mean the actual dates on which the price is captured. The price is the crude oil price in USD/BBL. The percent change is the change of percentage to the previous date. The change is the absolute change value to the previous date.

Crude oil price:
File name: crude.csv
Source link: https://www.kaggle.com/datasets/sc231997/crude-oil-price

The dataset of gasoline prices includes dates, average prices of all grades, price of regular gasoline, price of midgrade gasoline, and price of premium gasoline in USD/gallon from 1995 to the present by month. The dates are the actual dates on which the price of gasoline is captured. The average price of all grades is the average of regular gasoline, midgrade gasoline, and premium gasoline.

Gasoline price:
File name: gas_price.csv
Source link:
https://www.eia.gov/dnav/pet/hist/LeafHandler.ashx?n=PET&s=EMM_EPM0_PTE_NUS_DPG&f=M

The dataset of inflation rate includes years, amount, and inflation rate from 1701 to the present by year. However, only the data from 1995 to the present is the focus of this project. The years are the actual years in which the inflation is captured. The amount is the dollar amount. The inflation rate is the inflation rate for each year, for example, 0.01 means 1%.

Inflation rate:
File name: inflation.csv
Source link: https://www.kaggle.com/datasets/prasertk/300-years-of-inflation-rate-in-us

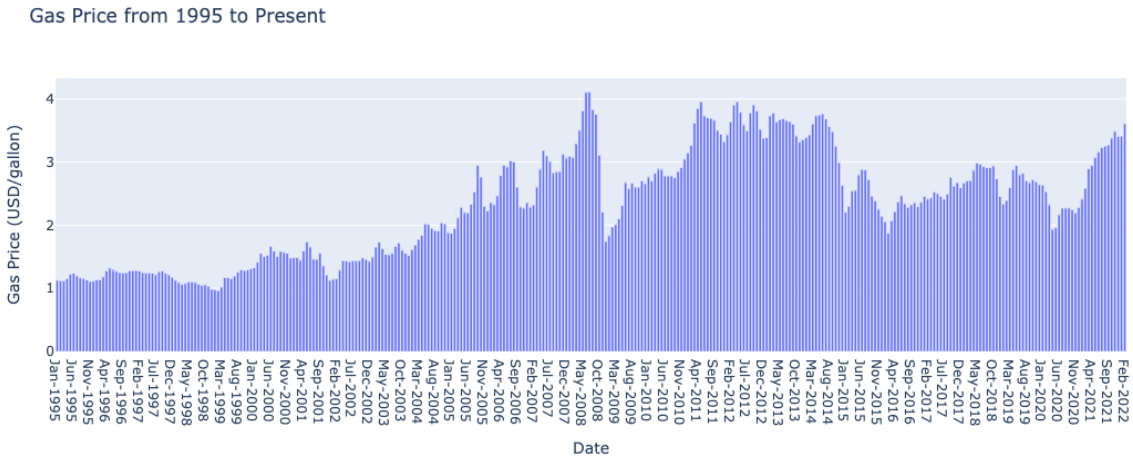
Approach

In order to generate three data visualizations for crude oil prices, gasoline prices, and inflation rates. The programming language that is used for this project is Python. First, the data on gasoline prices is imported into the data frame using the Python library, Pandas. Then, the data visualization library, Plot.ly, is used to generate a bar chart. The title, x-axis, and y-axis are modified. The text font size is modified to the size of 12. The color of bars is set to be blue. The title is labeled for the relationship between gasoline price and date from 1995 to the present. The x-axis is labeled for dates. The y-axis is labeled for gasoline price. This bar chart also includes a feature of tooltip to allow users to explore more detailed information for each data point.

Second, the data of crude oil prices is imported into the data frame using the Python library, Pandas. Then, the data visualization library, Plot.ly, is used to generate a line plot. The title, x-axis, and y-axis are modified. The markers are set to be green. The title is labeled for the relationship between crude oil prices and dates. The x-axis is labeled for dates. The y-axis is labeled for crude oil prices. This line plot also includes a feature of tooltip to allow users to explore more detailed information for each data point.

Finally, the dates of inflation rates are imported into the data frame using the Python library, Pandas. Then, the data visualization library, Plot.ly, is used to generate a scatter plot. The title, x-axis, and y-axis are modified. The color of the markers is modified to be red with a size of 12. The border color of markers is set to be dark slate grey with a width of 2. The title is labeled for the relationship between inflation rate and date. The x-axis is labeled for the date. The y-label is labeled for inflation rates. This scatter plot also includes a feature of tooltip to allow users to explore more detailed information for each data point.

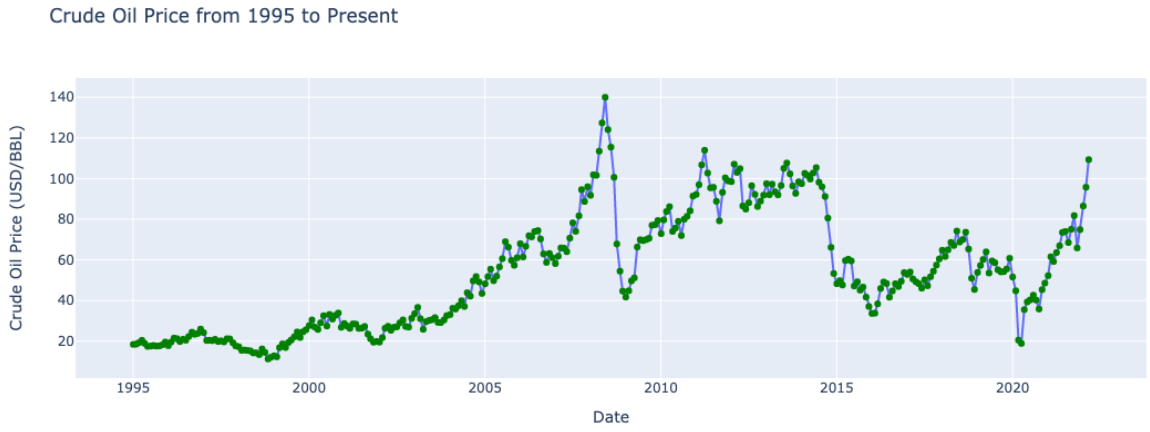
Visualization #1



After running the Python code for the gasoline dataset, the bar chart of gasoline prices from 1995 to the present is generated. Each bar shows the gasoline price for each specific month and year. By hovering over each bar, the visualization provides more detailed information for specific dates and average gas prices for those dates. By observing this visualization, one can notice the trend that the peaks are around July 2008 with the average gasoline price of 4.114 and Feb 2022 with the price of 3.611. Also, the high gasoline prices stay high throughout the years from 2011 to 2014.

The visualization of the bar chart is chosen because it clarifies the trend of gasoline prices, which shows the highest and lowest gas prices in this date range. The use of a bar chart provides an overview of key values. The use of blue bars can provide better visualization of the trend to users. Using the tooltip feature allows users to see more detailed information for both the x-axis and y-axis. It is easy for users to understand without a detailed description of the visualization because it is widely used in both the industry and academia. Furthermore, this bar chart helps summarize a huge dataset. The dark-ink ratio is applicable for this visualization because none of the information is redundant, and it avoids confusion for users. The use of bars also avoids confusion for users. The use of 2D data visualization helps to visualize the data more clearly.

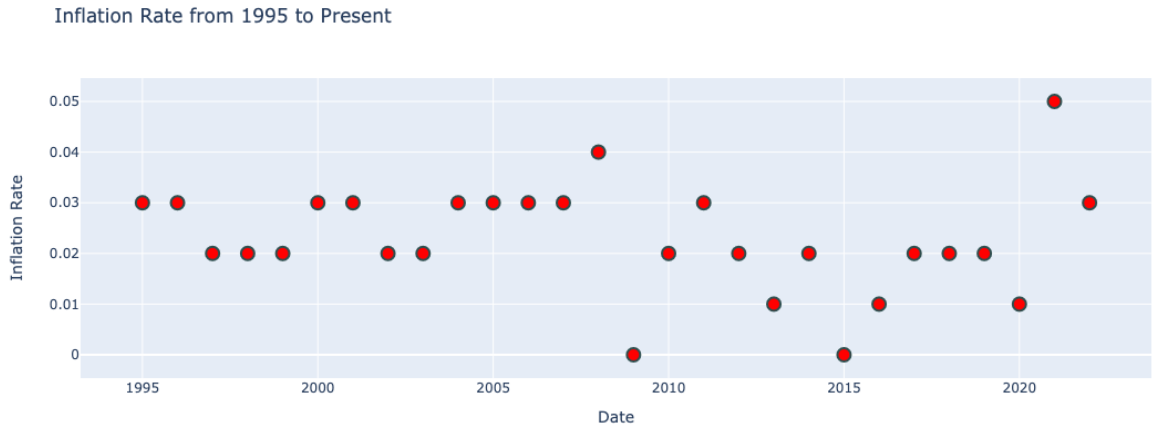
Visualization #2



After running the Python code for the crude oil price dataset, the line plot of crude oil prices from 1995 to the present is generated. Each data point shows crude oil prices for each specific month and year, and they are connected by a blue line to illustrate the trend. By hovering over each green data point, more detailed information for date and price is shown through the tooltip feature. After observing this visualization, one can notice the trend that the peaks are around Jun 2008 and Mar 2022. The high crude oil prices also stay throughout the years from 2011 to 2014.

The visualization of the line plot is chosen for the crude oil price trend because the line plot is useful in showing relationships between the x-axis and y-axis datasets, and it can show the trend of data effectively. The use of green data points with a blue line provides a better illustration of the trend to users. Also, the use of a line plot shows the highest and lowest crude oil prices in this specific date range. Using the tooltip feature is helpful to see more detailed information regarding the crude oil prices for specific dates. It is easy for users to understand the data without further descriptions of the visualization. the 2D visualization of the line plot helps to visualize the data more clearly. The use of data points also avoids confusion for users. Furthermore, the dark-ink ratio is applicable for this visualization because none of the information on this visualization is redundant, so it can avoid confusion for users.

Visualization #3



After running the Python code for the inflation rate dataset, the scatter plot of inflation rates from 1995 to the present is generated. Each red circle shows the inflation rates for specific years. By hovering over each red circle, the scatter plot provides more detailed information for specific years and inflation rates for that specific year. By observing this visualization, one can notice the years 2008 and 2021are relatively high.

The visualization of the scatter plot is chosen because it effectively shows the relationships between the x-axis and y-axis. The use of a scatter plot is good at showing non-linear patterns, and the maximum and minimum data points can be observed easily. The trend is easy to see without further description of the visualization. The use of the color red provides a vivid illustration for users to understand the data points. The use of a scatter plot provides key values of the datasets. Using the tooltip feature of the visualization provides more detailed information regarding the inflation rate and the years for those inflation rates. The use of larger circles with borders for markers of the visualization avoids confusion for users. Furthermore, the 2D data visualization helps to visualize the dataset more clearly. The dark-ink ratio is applicable for this visualization because none of the information is redundant, and it can avoid confusion for users.

Conclusion

By comparing the bar chart of gasoline prices, the line plot of crude oil prices, and the scatter plot of inflation rates from 1995 to the present, the correlation between oil price and inflation is confirmed. All three data visualizations provide good patterns for users to see the trends easily. Since there are the peaks of high crude oil prices and gasoline prices in 2008 and 2022, and there are peaks of high inflation rates in 2008, the end of 2021, and 2022, all three data visualizations show similar patterns, which also means all three visualizations of crude oil price, gasoline price, and inflation rate have a strong correlation. The root cause of high oil prices will

possibly increase the transportation of merchandizes to stores. Meanwhile, there is a chance of high transportation fees of raw materials to manufacturers, which also causes an increase in the price of merchandizes. The control of oil prices might be critical in order to control inflation.

References

Lioudis, N. (2022, March 10). What is the relationship between oil prices and inflation? Investopedia. Retrieved March 29, 2022, from https://www.investopedia.com/ask/answers/06/oilpricesinflation.asp