Queralt Zamora (254255)
Berta Noguera (254020)

Universitat Pompeu Fabra Barcelona

Visualizing trends and Predicting Gold prices

Problem Statement

Gold has always been considered a reliable store of value, particularly during times of inflation or economic uncertainty. However, the factors that influence gold prices remain complex, making it challenging for investors to anticipate trends and make informed decisions.
This project aims to address this problem by exploring historical gold prices from 2010 to 2024, identifying short and long-term trends, and analyzing correlations with key financial indices like the S&P 500 and CPI. Moreover, we aim to address the challenge of predicting gold prices, so that investors and market analysts can take advantage of it and provide meaningful insights.

Dataset Overview

Source

The dataset used is from Kaggle: https://www.kaggle.com/datasets/franciscogcc/financial-data/data

Exploratory Data Analysis

The EDA stage revealed strong correlations between historical gold prices and key financial and economic indicators. Gold prices are closely related to oil prices, reflecting similar sensitivity to global economic and geopolitical factors. Moderate correlations were also identified with NASDAQ, CPI and S&P 500 indicating the alignment that we expected with markets and inflation. In contrast, a weak negative correlation with platinum highlighted different market dynamics. These findings provide a foundation for developing visualization tools and models to predict gold price movements.
Overall, we have observed how gold prices increase during periods of uncertainty as for example, the European debt crisis, the COVID-19 pandemic, or the ongoing geopolitical tensions. On the other hand, declines during recovery periods, like the 2013–2015 taper tantrum, highlighted an inverse relationship with growing confidence in equity markets. Additionally, rising CPI aligned with gold's appeal during inflationary periods.

Business Questions and Objectives

Key Questions

1. What insights can be extracted to support informed investment decisions during economic uncertainty?
2. What are the key financial and economic indicators driving gold price movements, and how do they interact?
3. How accurately can short and long-term gold price fluctuations be predicted using a forecasting model?

Queralt Zamora (254255)
Berta Noguera (254020)

upf. Universitat Pompeu Fabra Barcelona

Objectives

- **Support investment decision-making:** Provide data-driven insights into gold price trends to help investors make informed decisions during periods of economic uncertainty.
- **Understand dynamics between gold and other financial indicators:** Analyze the relationships between gold prices and other financial variables to uncover key drivers of price movements.
- **Offer a forecasting tool for predicting gold price fluctuations:** Develop a reliable model for forecasting gold prices using influential factors, enabling better anticipation of market changes.
- **Visualize key trends:** Create visualizations to illustrate gold price dynamics and their connections to broader economic and financial trends.

Methodology

Data Preprocessing

The dataset was cleaned and prepared using Jupyter Notebooks. This involved handling missing values. For instance, data for variables like CPI or GDP was only available on the first day of each month or trimester. We used '.fillna(method='ffill)' to propagate the last available value forward in the column. This perspective considers the influence of CPI and GDP indicators in the subsequent month, assuming that once a new indicator is released, it affects gold price fluctuations. The remaining NaN values were handled by deleting the rows that contained at least one missing value, since we considered that it was the case for a few days and we had enough data. Finally, we had to adapt data and prepare it for the visual tool. So we implemented the separation of the commas 'data.to_csv('preprocessed_gold.csv', decimal=',')' to address issues with Tableau recognizing the values as doubles.

Exploratory Analysis and Storytelling

Tableau was used to create an interactive story composed of two dashboards. The first dashboard displays the historical trend of gold prices, marking key peaks and declines with relevant geopolitical or economic context. This same dashboard provides a comparison between gold and other metals and oil prices. It also includes a pie chart where the percentage of the volume sold per year is visualized. All the plots can be filtered, and the dashboard also has an explanation to guide the user to understand how to interact with it.

Similarly, the second dashboard contains the comparison of gold prices along with economic indicators. It also features a box plot of gold's closing prices categorized by NASDAQ, S&P 500, and CPI values. Observations reveal that as these indicator values increase, gold prices also rise. This relationship is illustrated through the color and shape of the plotted dots, indicating the direct link between these indicators and inflationary or global economic trends. Together, these dashboards explain the relationships between gold prices and other financial indicators, enabling users to explore correlations and trends interactively.

Queralt Zamora (254255)
Berta Noguera (254020)

upf. Universitat Pompeu Fabra Barcelona

Prediction and Visualization

The predictive model was developed with Jupyter Notebooks. A Streamlit page was created to serve as the primary interface for predicting gold prices and visualizing results. The app allows users to input exact variables and predict prices over short and long-term periods. The Streamlit page integrates the Tableau dashboards and the EDA insights and explanations consolidating all visual and interactive tools in one platform for seamless user experience.

Conclusions

1. **Influence of multiple factors:** While our analysis identified strong correlations between gold prices and variables such as oil prices, CPI, and other economic indices, it is evident that gold price movements are influenced by a multitude of factors beyond those captured in our dataset.
2. **Enhanced explainability:** Through our exploratory analysis and visualizations, we were able to explain a significant portion of the factors driving gold price dynamics.
3. **Oil as the most correlated variable:** Oil prices resulted to be the most correlated variable with gold prices. This relationship can be explained by their shared sensitivity to macroeconomic and geopolitical factors, such as inflation, supply chain disruptions, and global energy demand.
4. **Price predictions until 2030:** Using the developed predictive model, we successfully forecasted gold prices through 2030. However, as mentioned before, the analysis revealed that the most impactful driver of future gold prices remains the global and geopolitical context, which can cause sudden and unpredictable shifts.
5. **Comprehensive visualization and prediction tools:** The integration of Tableau dashboards in the Streamlit app enabled an interactive and user-friendly experience for understanding trends, exploring relationships, and forecasting gold prices, making the findings accessible and actionable for investors.

In conclusion, we have exceeded our goals and addressed the main business questions, providing valuable insights into gold price trends. However, predicting future movements will require continuous analysis of economic indicators and global events.