

Investgenics Financial Data Analysis Report

Background

Investgenics specializes in financial data analytics, providing investment advice to clients. A key business problem is determining the most stable and profitable investment among stocks, gold, and oil. Investors need to diversify their portfolios to maximize returns while minimizing risk. The objective of this analysis is to compare the performance of gold and oil using historical data to identify trends, volatility, and potential safe-haven assets during market instability, ultimately guiding clients towards informed investment decisions.

Analytical Approach

To address the business problem, a comprehensive analysis of historical gold and oil price data was conducted using Python within a Jupyter Notebook environment. The process began with importing the necessary libraries (pandas, seaborn, matplotlib) and loading the datasets (gold_stocks_price.csv, oil_price.csv).

Data Cleaning and Preparation:

- **Data Inspection:** Initial inspection involved checking the shape, data types, and the presence of missing values. The Volume column in the gold dataset was analyzed before subsetting to identify stocks with the highest and lowest trading volumes.
- **Subsetting:** To focus on relevant price metrics, subsets of both dataframes were created, retaining only the Date, Open, High, and Low columns.
- **Data Type Conversion:** The Date column was converted to datetime objects to facilitate time-series analysis and plotting.
- **Data Transformation:** New features were engineered to aid comparison. A user-defined function calculated the daily average price in USD, and a lambda function converted these prices to GBP (assuming an exchange rate of 1 USD = 0.8 GBP).

- **Merging:** The gold and oil datasets were merged on the Date column to create a single dataframe for side-by-side comparison of daily prices.

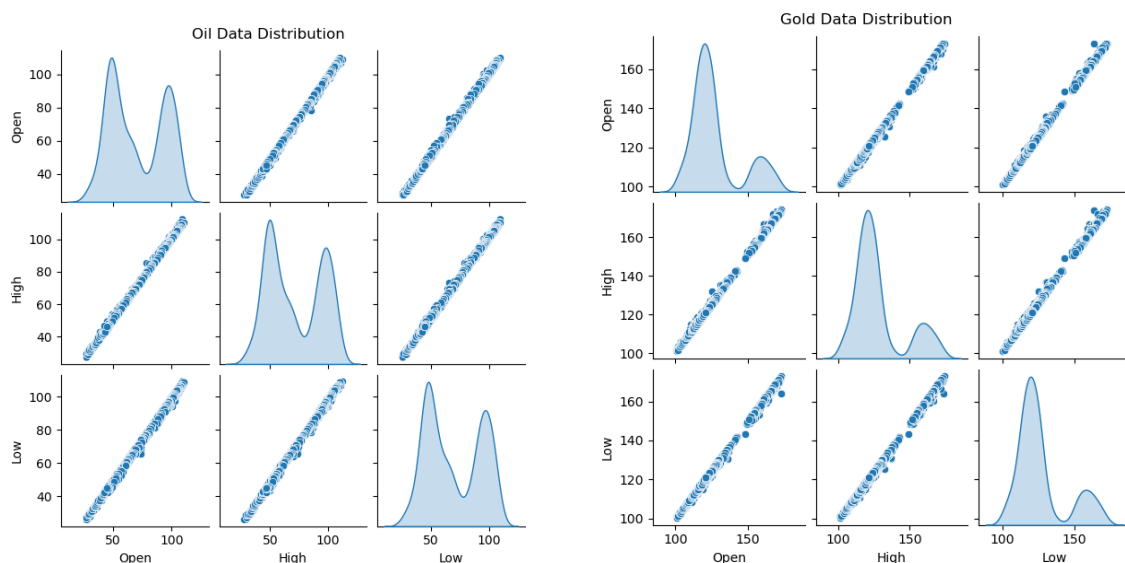
Analysis:

- **Descriptive Statistics:** Summary statistics (mean, standard deviation, min, max, percentiles) were generated using the describe() function to understand the central tendency and dispersion of the data.
- **Diagnostic Analytics:** Visualizations were employed to explore data distributions, identify outliers, and analyze specific market events. Pairplots revealed the distribution of prices, while boxplots highlighted outliers. Bar plots compared performance during specific periods (e.g., Dec 2015), and line plots tracked trends over time, particularly around significant events like the Brexit vote in June 2016.

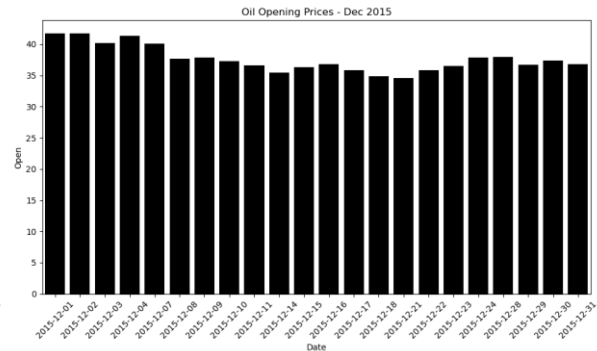
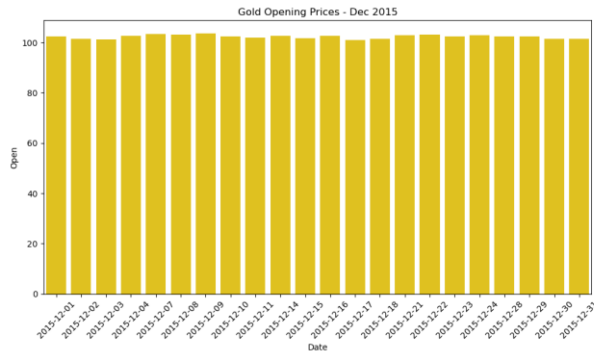
Visualization and Insights

The selection of visualizations was driven by the specific questions posed by the business scenario.

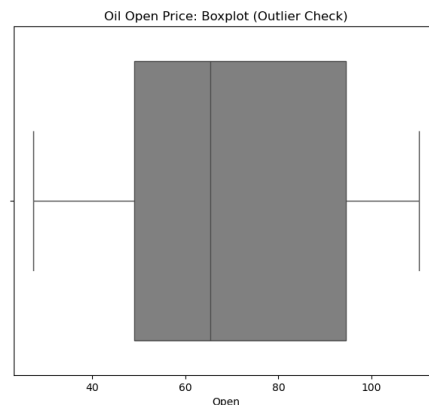
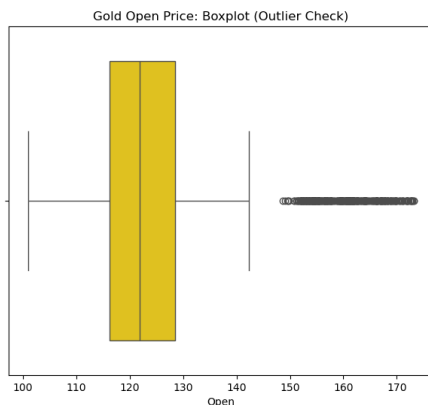
- **Distribution Analysis (Pairplots):** Pairplots with Kernel Density Estimation (KDE) were used to visualize the distribution of Open, High, and Low prices.
 - **Insight:** The distributions for both gold and oil were non-normal, often showing multi-modality or skew. This indicates that prices tend to cluster around certain support and resistance levels rather than following a smooth bell curve, suggesting that standard risk models assuming normality might be insufficient.



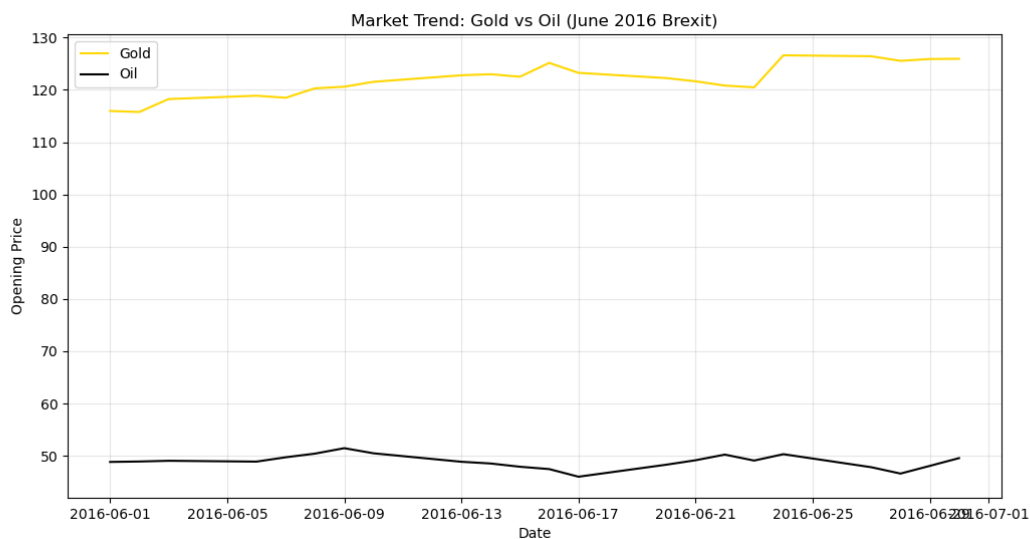
- December 2015 Performance (Barplots): To compare performance during a specific downturn, bar plots of Open prices for December 2015 were generated.
 - Insight: The visualization clearly showed a downward trend for Oil prices throughout the month, reflecting the global supply glut. In contrast, Gold prices remained relatively stable, demonstrating its resilience during periods of commodity market stress.



- Outlier Detection (Boxplots): Boxplots were chosen to identify outliers in the Open prices.
 - Insight: Both assets displayed outliers. However, Gold's outliers were often on the upper end (positive price spikes), while Oil showed outliers on the lower end (price crashes). This suggests that extreme volatility in Gold can be profitable, whereas in Oil, it often signifies a sharp loss in value.



- June 2016 Market Event (Lineplots & Scatterplots): Line plots were used to visualize the price trends during June 2016 to assess the impact of the Brexit referendum.
 - Insight: A distinct spike in Gold prices was observed around June 24, 2016, correlating with the Brexit vote result. This visual evidence confirms Gold's status as a "safe-haven" asset that investors flock to during geopolitical uncertainty, while Oil prices did not exhibit such a protective surge.



Patterns and Predictions

The analysis revealed several key patterns:

1. **Volatility:** Gold exhibited higher standard deviation in the initial subset, indicating greater price fluctuation. However, its "upward" volatility (positive outliers) often acts as a benefit during market panic.
2. **Safe-Haven Status:** The sharp rise in Gold prices during the Brexit vote (June 2016) and its stability during the Oil crash (Dec 2015) confirm its negative correlation with riskier assets. It serves as an effective hedge against market downturns.
3. **Trend Divergence:** Gold and Oil prices do not move in tandem. While Oil is sensitive to supply-demand dynamics (like the 2015 oversupply), Gold is more sensitive to market sentiment and geopolitical risk.

Recommendations:

- **Diversification:** Investors should include Gold in their portfolios to hedge against stock market and currency volatility.
- **Further Exploration:** Future analysis should investigate the correlation between these commodities and major stock indices (like the S&P 500) to quantify the diversification benefit precisely. Additionally, modelling the impact of interest rate changes on Gold prices would provide deeper insights for forecasting.

