Financial Dashboard for NASDAQ: Real-Time Data Processing and Visualization

INTRODUCTION:

This project focuses on developing a financial dashboard that focuses on analysing and visualizing the current financial position for the companies, to provide actionable insights for investors and stakeholders in the financial market. The companies targeted in this project are Amazon, Apple, Microsoft, NVIDIA and Tesla which are basically in top 10 NASDAQ stock exchange. There is Real-time Data integration done using APIs and it is processed through Apache Spark for and then further visualized through an interactive dashboard using Streamlit. The API is specifically targeted to extract parameters related to the cash flow of the company.

BACKGROUND KNOWLEDGE OF DATASET:

The following is the general idea about the parameters involved in the dataset:

I) Cash Flow Components:

- Cash from Financing: Funds raised from debt or equity.

- Cash from Investing: Outflows or inflows due to investments in assets or securities.

- Cash from Operations: Core business-generated cash.

ii) Net Change in Cash: Summarizes overall cash movement.

iii) Free Cash Flow (FCF): The cash available after capital expenditures — critical for investors.

iv) Net Income: Profit after expenses, taxes, and interest.

v) Period Attributes: Month, Year, Date.

SYSTEM ARCHITECTURE:

The Architecture workflow includes APIs, where it extracts the data source of the financial metrics. Then the Backend for data extraction, transformation and loading is done through Spark. After that using Pandas data processing is done and finally inorder for front-end visualization and interactivity Streamlit is used. The technical choices that leverage this project are Apache Spark that is scalability, and the other is Real-time updates that are enabled by APIs.

TECHNOLOGY AND TOOLS:

The Data Sources for this project are APIs using RapidAPI from the following: link. In addition to the Data Processing Framework includes Apache Spark for efficient distributes data processing, where the key components are SparkSession, RDD and DataFrames. Furthermore, the Visualization Framework is Streamlit for interactive and real-time data representation. Notably the libraries used are Requests for API calls, Pandas for data manipulation and Plotly for interactive visualizations. Regrading the forecasting that is done in this project the modeling tools that are used are scikit-learn for linear regression.

DATA EXTRACTION AND PROCESSING

The API integration is done through RESTful API for cash flow data, where the key steps are HTTP GET requests using requests library. Furthermore, the parameters are symbol (company), period (quarterly) and language. Moreover, the status code checks and retries are there for API failures to cater with the Error handling.

Data Processing in Spark is achieved through converting JSON responses into Spark-compatible formats where we use RDD by SPARK.SPARKCONTEXT.PARALLELIZE(). Furthermore it is loaded in JSO data with **SPARK.READ.JSON()**. In addition to transformation and cleaning it is done by selecting the key metrics (data, net_income, cash_from_operations, cash_from_financing, cash_from_ investing, free_cash_flow. Moreover, the date fields are converted into datetime format to perform further data manipulation and after that the processed data is stored in the form of csv using Pandas.

DASHBOARD DEVELOPMENT

The interactive dashboard using Streamlit is created, where plotly is used to make visually appealing graphs. The features of the dashboard include a dynamic dataset selection by having a sidebar for choosing datasets by company, and the visualizations present on the dashboard include line charts, bar plots and correlation heatmaps. Furthermore, for predictive model linear regression is used to predict future trends.

The visualization details include Net income trends represented by line charts, Cash Flow Metrics represented by line charts, Aggregated Metrics represented by bar charts. Correlation Analysis by Heatmaps and Scatter plots.

Regarding the future prediction of net income, Linear Regression is used and for data scaling MinMaxScalar is used for consistent performance. Comparison of historical and predicted trends include historical peaks in net income and is carried forward with predicted values over the next 12 months visualized alongside historical data.

There is an Investor Insights section after each visualization, that provides actionable insights to help investors identify growth opportunities and assess risks. Promising predictions indicate potential growth trajectories, while slower forecasts may highlight operational inefficiencies or challenges. By comparing historical performance with future forecasts, investors can better evaluate trends and anticipate changes. For tailored recommendations, key metrics offer valuable signals: High or low Net Income reflects the company's profitability strength or concerns; Positive or negative Cash from Financing indicates capital-raising strategies or significant debt repayments; and Aggressive investments in assets or projects signal long-term growth strategies. Investors should also monitor external factors such as market trends and economic conditions to understand their influence on predictions and financial metrics. These insights equip stakeholders with the tools to make informed decisions by analyzing both historical and forecasted data effectively.

SNAPSHOTS OF THE DASHBOARD

Financial Data Dashboard

Selected Dataset: Amazon (AMZN)

Net Income Over Time
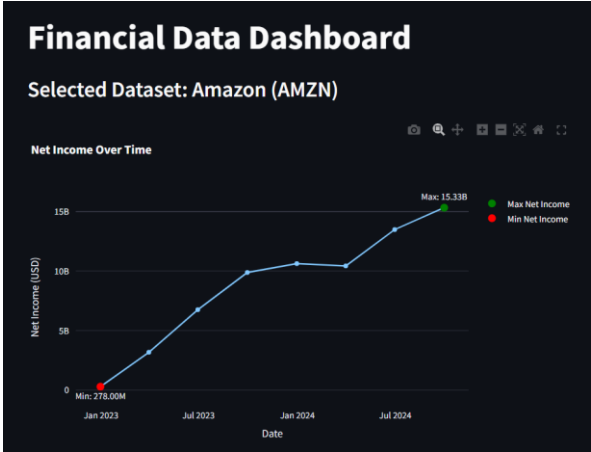
Investor Insights: Net Income

🔻 **Below Expectations**: Net Income peaked at only 15.33B. Investigate challenges affecting profitability.

📉 **Profitability Concern**: Net Income dropped to a low of 278.00M in Dec 2022. This may indicate operational inefficiencies or market-related challenges.

💡 **Investor Tip:**
- **Consistent Net Income growth** is a sign of strong profitability and operational excellence.
- **Sudden dips** may indicate challenges in cost management, revenue streams, or external factors.
- Evaluate trends over time to identify long-term sustainability and compare performance with peers in the same industry.

Aggregated Cash Flow Metrics by Year
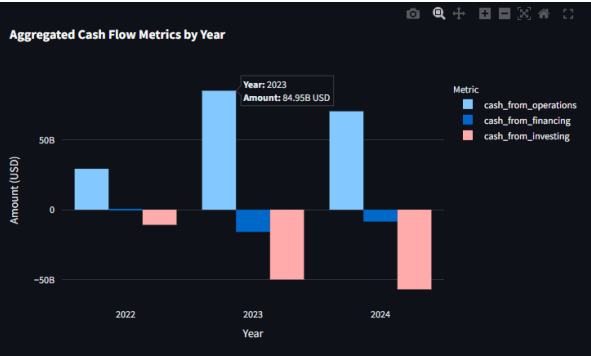
**Investor Insights: Aggregated Cash Flow Metrics**

✅ **Strong Operational Cash Flow**: Total Cash From Operations across years is 184.36B. This indicates strong core business performance and cash generation.

📉 **Debt Repayment Focus**: Total Cash From Financing is -24.30B, reflecting significant repayments or reduced external financing.

📈 **Aggressive Investments**: Total Cash From Investing is -117.55B. The company is investing heavily in assets, acquisitions, or R&D for future growth.

💡 **Investor Tip:**
- **Cash From Operations**: Consistently high values reflect strong core business performance.
- **Cash From Financing**: Positive values indicate capital raises; negative values show debt repayments or buybacks.
- **Cash From Investing**: Heavy investments may signal growth strategies but require monitoring ROI.
- Compare trends across years to evaluate financial stability and growth sustainability.

Predictive Model: Future Trend of Net Income

Net Income: Historical and Predicted Trend
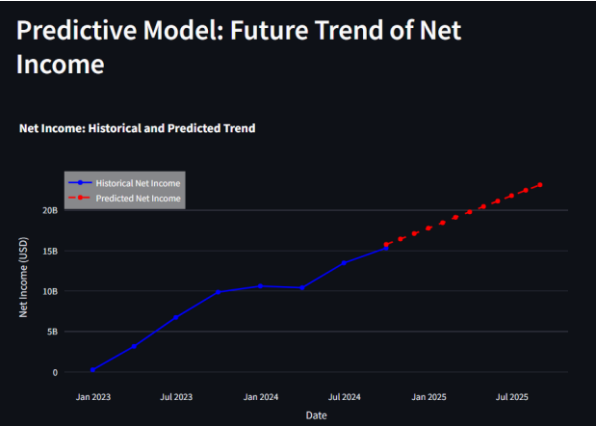
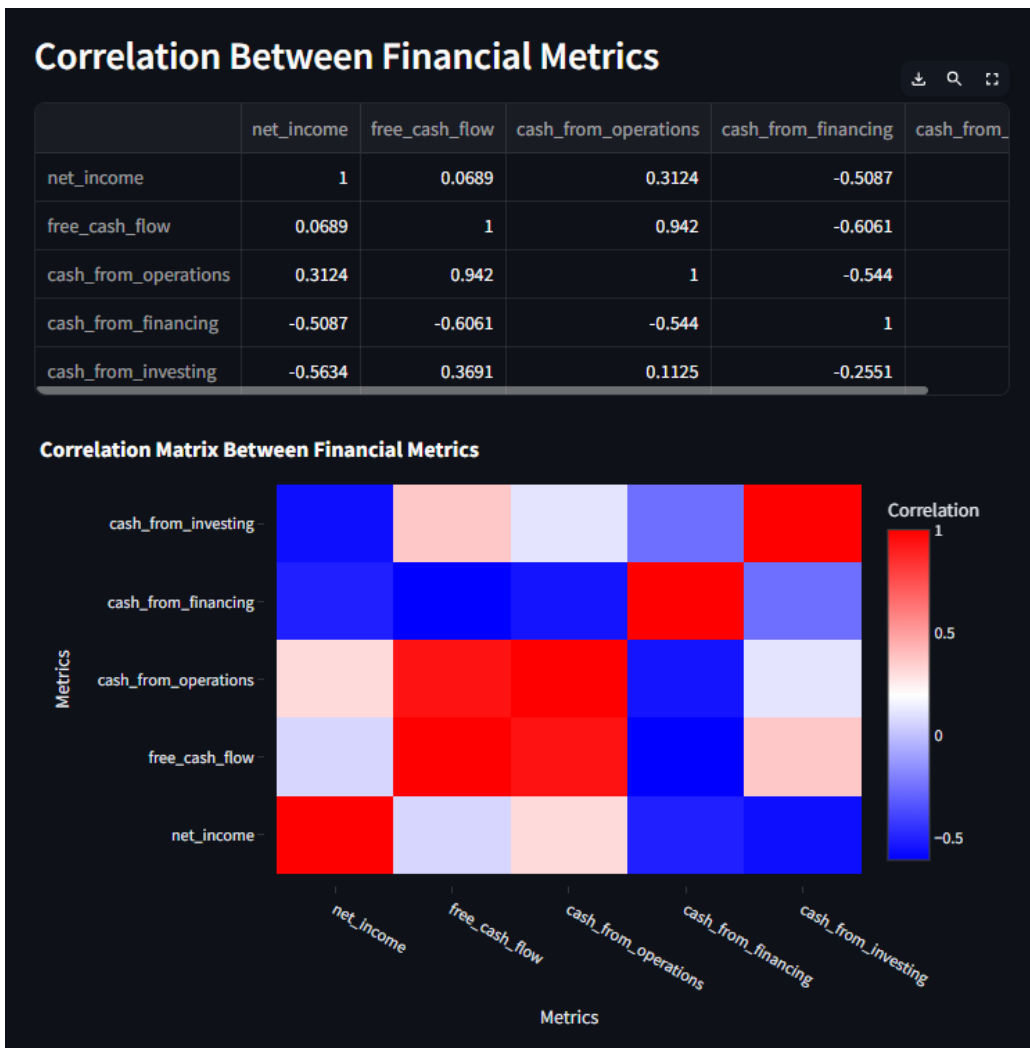**Investor Insights: Predictive Model for Net Income**

⚠️ **Moderate Historical Performance**: The peak Net Income was 15.33B USD. There may be room for improvement in profitability.

📉 **Downtrend Concern**: Historical Net Income dropped as low as 278.00M USD. Evaluate past challenges in profitability or market pressures.

📊 **Promising Future**: The model predicts Net Income could rise to 23.14B USD, surpassing historical highs. This suggests potential growth opportunities.

💡 **Investor Tip:**
- Compare **historical performance** with the predicted trend to evaluate the company's growth trajectory.
- A rising predicted trend suggests opportunities for expansion, while a flat or declining prediction warrants caution.
- Investors should consider external factors (market trends, competition) that could influence the company's ability to meet or exceed these predictions.

Correlation Between Financial Metrics

	net_income	free_cash_flow	cash_from_operations	cash_from_financing	cash_from_
net_income	1	0.0689	0.3124	-0.5087	
free_cash_flow	0.0689	1	0.942	-0.6061	
cash_from_operations	0.3124	0.942	1	-0.544	
cash_from_financing	-0.5087	-0.6061	-0.544	1	
cash_from_investing	-0.5634	0.3691	0.1125	-0.2551	

Correlation Matrix Between Financial Metrics

Conclusion

In conclusion this project delivers a valuable tool for investors and decision-makers to analyze, forecast, and act upon critical financial insights into the financial performance of the leading NASDAQ-listed companies.

Code appendix

Link