**Capstone Project – AI-powered cash flow forecasting tool**

Group 4 members:

Lai Wing Man, Carmen; Law Tak Chun, Dirk; Keung Kwan Lam, Olivia;

Ho Chun Kit, Philip, Chan Tung Wai, Teddy; Tang Chi Fung, Jacob

1. **Project Background**

Our client project targeted to develop an AI-powered cash flow forecasting tool to enhance client financial management. The tool would gather and clean historical financial data, identify key cash flow indicators and trends, and build linear regression models to predict revenue and expenses. It would incorporate sensitivity analysis for scenario testing, provide interactive visualizations of cash flow projections, and integrate a large language model (LLM)-powered chatbot to answer client queries, to deliver a comprehensive and user-friendly solution for strategic financial planning.

Apple Inc. (Apple) was selected for analysis due to its stable financial performance, extensive historical data, and industry-leading position, making it an ideal case for developing an AI-powered cash flow forecasting tool. Its consistent revenue growth and diverse product portfolio provide rich data to model trends and test scenarios, ensuring the tool’s insights are both reliable and relevant for strategic financial planning.

1. **Dataset Description**

The dataset utilized for developing the AI-powered cash flow forecasting tool encompasses comprehensive financial data of Apple Inc. (US:AAPL) spanning over the past 20 years, from 1994 to the present. The data was sourced from EODHK via APIs, which provided access to various financial statements, including the balance sheet, cash flow statement, and income statement.

This dataset includes key financial metrics such as revenue, net income, total assets, cash flow, and relevant financial ratios. These metrics are critical for understanding the company's financial health and operational performance over time. Given the multi-source nature of the data, it contains some missing values and anomalies that require careful handling during preprocessing. The extensive time span allows for detailed trend analysis, capturing significant growth phases, seasonal patterns, and irregular fluctuations associated with market conditions or strategic events.

1. **Dataset Preprocessing**

The raw dataset underwent a series of preprocessing steps aimed at preparing it for accurate analysis and modeling. Initially, the 'date' column was converted into a datetime format to facilitate time series analysis and chronological filtering. To focus on recent trends and eliminate noise from earlier periods, the data was filtered to include only records from 1994 onwards. The 'date' column was then set as the dataset's index, enabling efficient time-based operations. For better interpretability and consistency, key financial metrics were summarized using descriptive statistics, with values converted to billions. This transformation not only improved readability but also ensured comparability across different metrics and time periods. Although not explicitly detailed here, typical preprocessing would also involve handling missing data through imputation techniques and detecting anomalies or outliers to maintain dataset integrity. These steps collectively ensured that the dataset was clean, consistent, and suitable for exploratory analysis and subsequent model development.

1. **Exploratory Data Analysis (EDA)**

The exploratory data analysis provided valuable insights into Apple’s financial trajectory and the characteristics of the dataset, supported by a variety of visualizations.

The initial analysis involved examining the long-term trends of key financial metrics through a series of time series plots. Figures 1.1 through 1.7 display these charts for revenue, net income, total assets, cash flow, and other relevant metrics. These visualizations collectively revealed a consistent upward trajectory across most metrics, indicating sustained growth over the years. Notably, these charts highlighted accelerated growth phases coinciding with major product launches, such as the iPhone and iPad, while also capturing periods of fluctuation corresponding to market conditions or strategic investments. Sudden dips or spikes observed in these figures often aligned with significant external events, including economic downturns or large-scale acquisitions.

Next, the distribution of financial metrics was analyzed using histograms and Kernel Density Estimate (KDE) plots, as shown in Figures 2.1 through 2.7. These charts illustrated that most metrics exhibited right-skewed distributions, with a concentration of years showing moderate to high performance, and a few years standing out with exceptionally high revenue or profits. Outliers identified through these distribution charts pointed to extraordinary years, either in terms of exceptional performance or financial difficulties, emphasizing the importance of considering these anomalies during modeling.

To further identify specific outliers, four boxplots are presented in Figures 3.1 through 3.4. These boxplots provided a clear visualization of years with unusually high or low values in key metrics, often corresponding to record-breaking financial results or periods of financial distress. Recognizing these outliers is vital, as they can significantly influence trend analysis and model accuracy if not appropriately managed.

The relationships among various financial variables were examined using a correlation matrix heatmap shown in Figure 4.1. This visualization revealed strong positive correlations between revenue and total assets, indicating that asset growth closely mirrors sales performance. Additionally, a significant correlation was detected between net income and gross profit, suggesting consistency in profitability. Moderate correlations between cash holdings and revenue suggested liquidity levels are partly dependent on operational performance, while weaker correlations between other pairs highlighted their relative independence. These insights are instrumental for feature selection in predictive modeling.

Further exploring specific variable relationships, two scatter plots are presented in Figures 5.1 and 5.2. Figure 5.1 depicts the positive linear relationship between revenue and net income, reinforcing the understanding that higher revenue generally translates into higher profits. Figure 5.2 illustrates the correlation between total assets and long-term debt, indicating that investments are often financed proportionally. Outliers observed in these scatter plots drew attention to particular years with anomalous relationships, possibly due to extraordinary expenses or revenue spikes, warranting deeper investigation.

Finally, the analysis of ratios over time is encapsulated in Figure 6.1, which plots profitability margins and working capital ratios as percentages of revenue. This visualization uncovered trends such as relatively stable or gradually improving profit margins, indicative of operational efficiencies. Conversely, fluctuations in working capital ratios reflected changing cash management strategies and operational needs. A general decline in these ratios over the period raised questions about rising operational costs or tightening liquidity, highlighting the importance of continuous monitoring for maintaining financial health.

Overall, these visualizations—spanning time series graphs, distribution analyses, outlier detection, correlation assessments, and ratio evaluations—provide a comprehensive understanding of Apple’s financial performance. They reveal key patterns, relationships, and anomalies that are essential for developing a robust and accurate cash flow forecasting model.

1. **Data Visualization**

Tableau was adopted to create effective visualization to convey complex data insights clearly and engagingly. Four dashboards were developed to provide a comprehensive analysis of Apple’s sales and financial metrics.

1. **Dashboard 1: Sales Overview 2007 - 2024 (Appendix A)**

Due to limited data availability, this dashboard visualized Apple’s sales across various products from 2007 to 2024 only.

Fig 7.1 - The top left featured a line chart displaying net sales over time, with a forecast projects extending to 2026. In 2024, for instance, net sales reached $391 billion, driven by strong iPhone performance and growth in Services. The forecast projects $423 billion in 2025 and $445 billion in 2026, based on historical growth rates and expected demand for new products.

Fig 7.2 - A stacked bar chart illustrated the product mix contributing to total sales. A consistent upward trend in net sales was shown over the years. By hovering over the chart, the year-over-year sales changes could be viewed as well. In 2024, the olive-green portion representing the iPhone contributed the largest share at $201 billion among other products like Mac and iPad. The orange bars, representing the iPod, indicated its discontinuation in 2015, which might primarily due to the rise of smartphones and streaming services.

Fig 7.3 - A sparkline chart (with motion) on the far right offered a compact visual of sales trends by product. All products, except the iPod, showed a persistent increase in sales.

Fig 7.4 - A summary table at the bottom left outlined the top-selling products and their total sales. Additionally, a brief summary of total sales of other products could be shown when hovering over the table. With global demand, the iPhone had remained the top-selling product since 2010. Over 50% of Apple’s total net sales was generated by the iPhone in 2024 as well.

Fig 7.5 - Regarding regional sales, a donut chart highlighted that the Americas contributed the most in 2024. A time filter was available to examine top contributors across different years.

Fig 7.6 - To support the dominance of the iPhone since 2010, a treemap in the bottom right visualized its sales breakdown by region over the last 5 years. In 2024, the total sales of iPhone in the Americas market reached $0.17 billion. While Europe, being the second, contributed $0.1 billion in sales.

1. **Dashboard 2: Cash Flow 1994 - 2024 (Appendix B)**

This dashboard illustrated Apple’s cash flow trends in the last 30 years from 1994 to 2024.

Fig 8.1 - On the left, key indicators and ratios were shown, with a filter for previous data. In 2024, free cash flow stood at $109 billion, reflecting cash available after capital expenditures.

The operating cash flow ratio was 0.67, which indicated a moderate level.

The current ratio, at 0.867, measured Apple’s ability to cover short-term liquidity. Although a value below 1 suggested current liability exceed current assets, Apple had historically maintained a current ratio slightly above or below 1 (e.g. current ratio at 0.88 in 2022), demonstrating its efficiency in cash management and strong revenue streams.

Working capital in 2024 was negative $23 billion, which still reflected Apple’s operational efficiency through quick inventory turnover and strong cash flow from sales.

Fig 8.2 - A waterfall chart on the right compared annual cash flow changes. Each bar depicted beginning and ending cash flows, with green bars indicating increases and red bars showing decreases. There were notable growth and decline occurred in 2019 and 2022 respectively. In 2024, there was a slight decrease of $0.79 billion in the cash flow. Nevertheless, Apple’s cash flow demonstrated consistent growth over the past 30 years.

Fig 8.3 - At the bottom, a table detailed the cash flows from operating, investing and financing activities, along with year-over-year percentage changes.

1. **Dashboard 3: Profit and Loss 1994 - 2024 (Appendix C)**

This dashboard visualized Apple’s profit and loss data including incomes, profits and margins, with a scale at the top to show a particular year’s data from 1994 to 2024. While the plotted graph showed the overall trend.

Six text tables were created to directly show the value of Net Income, Operating Income, Gross Profit, Profit Margin, Operating Margin and Gross Margin. Whenever the number is negative, the number will be in red and in bracket.

Fig 9.1 – A combined chart of Total Revenue (bar chart) and Revenue Growth Rate (line chart). The revenue growth rate shows variability, indicating that Apple faced different market conditions, competitive pressures, or product cycles over the years. The rapid revenue growth from $8 billion to $395 billion demonstrated Apple's significant expansion and market capture, particularly driven by successful product launches (e.g., iPhone, services).

Fig 9.2 – A combined chart of Gross Profit (bar) and Gross Margin Percentage (line). Gross Profit followed a similar growth trajectory as revenue, highlighting effective cost management relative to sales. An increasing gross margin percentage signified improved efficiency in production and sales, possibly due to economies of scale and higher-margin products.

Fig 9.3 – Line charts showing Operating Income, Expense and Margin. In the early years, operating income and expenses were closely aligned, indicating a balanced approach to growth. The deviation starting in 2008 suggests that Apple began to optimize operations more effectively, allowing operating income to rise faster than expenses. The steady increase in operating margin to 25-30% reflects improved operational efficiency and cost control, contributing to stronger profitability.

Fig 9.4 – The line chart illustrating Net Income and Net Profit Margin shows there were fluctuation in the early years but then steadily increased and stabilized since 2011, indicating a strong competitive position and effective management of costs and pricing strategies.

Fig 9.5 – The chart showed the Return indicators on Equity, Assets and Investment. Improvements were observed for all three aspects, suggesting effective capital allocation and project management.

1. **Dashboard 4: Balance Sheet 1994 - 2024 (Appendix D)**

This dashboard provides an overall view on some key balance sheet data, and the data on a particular year can also be viewed and selected from the drop-down bar.

Fig 10.1 – A combined chart of Total Assets (bar) and Total Current Assets (line). It was clear that Apple has experienced a significant increase in total assets over the years, rising from $10 billion in the mid-2000s to over $350 billion by 2017. The trend in total current assets also shows an upward trajectory, although they represent about half of total assets, indicating a strong liquidity position.

Fig 10.2 – The liability balance chart illustrates the level of Current Assets and Liabilities. Both have shown similar growth trends since 2008, indicating that Apple is expanding its operational scale. However, the asset-to-liabilities ratio has decreased significantly from 3.3 in 2001 to a low of 0.8 in 2024, suggesting increased leverage and potential concerns about short-term liquidity.

Fig 10.3 – A line chart of Long Term Debt and Investments. Debt side has been on the rise since 2012, peaking at $97 billion in 2021 before declining. This suggests Apple has used debt strategically for growth during that period. Investment was always higher than Debt and demonstrated a similar pattern with larger fluctation. This reflects a fluctuating investment strategy.

Fig 10.4 – A combined chart of Asset Turnover Ratio (line) against total revenue (bar). As the revenue significantly increases, the asset turnover ratio has declined, reaching a low of 0.6 in 2017, indicating decreasing efficiency in generating revenue from assets during that period. Since then, the ratio has improved, reaching 1.1 in 2024, suggesting better asset utilization and operational efficiency.

Fig 10.5 – Three line charts showing Leverage Ratios. The debt-to-equity and debt-to-asset ratios have both exhibited a gradual increase throughout the 2010s, peaking in 2021-2022. This indicates a growing reliance on debt financing relative to equity. The debt repayment period showed significant fluctuations before 2003. Since then, excluding missing data from 2004-2012, it has remained relatively steady with an average of around 1 year over the past decade, suggesting manageable debt servicing.

1. **Liquidity Monitoring Dashboard**

This financial stress test evaluates Apple's cash flow resilience by simulating three economic scenarios—base, mild, and severe—to assess the company's liquidity under varying degrees of financial pressure. The model applies progressive shocks to key cash flow components, including operating cash flow reductions (up to 15%), capital expenditure increases (up to 20%), and financing activity adjustments (up to 25%), then quantifies their impact on free cash flow and cash reserves.

The analysis features intuitive waterfall charts that track cash movements across operating, investing, and financing activities, complemented by an interactive liquidity dashboard. This dashboard provides a comprehensive view of Apple's financial health through four key visualizations: (1) a cash balance trend line tracking historical positions, (2) a free cash flow composition pie chart, (3) color-coded warning indicators (red/yellow/green) based on cash runway calculations, and (4) critical metrics including monthly burn rate and working capital changes.

Together, these tools enable stakeholders to quickly identify potential liquidity risks, evaluate Apple's capacity to withstand financial stress, and make informed decisions about the company's financial stability across different economic conditions.

1. **LLM Application - AI Chatbot for Apple Annual Reports**

To enhance user interaction and support financial insight extraction, an AI-powered chatbot was developed using a large language model (LLM) architecture. The chatbot specializes in answering queries based on Apple official annual reports. It was implemented using a combination of modern AI and web technologies:

* **Python** served as the core development language.
* **LangChain** and **LangGraph** were adopted for orchestrating the document processing and conversational flow logic.
* **OpenAI’s GPT model** powered the chatbot’s language understanding and response generation.
* **Selenium** was used to automate web scraping, enabling the system to download all annual reports (in PDF format) from Apple’s official investor relations website.
* The PDF contents were parsed and embedded into a **vector database**, enabling semantic search and retrieval-based question answering (RAG).
* The front-end user interface was built using **Streamlit**, providing an intuitive and interactive experience.

Users can ask the chatbot questions such as “What was Apple’s revenue growth in 2020?” or “How did R&D expenses trend in recent years?”, and receive accurate answers grounded in the underlying PDF reports. This implementation not only demonstrates the integration of LLMs with financial documents but also highlights a practical application of generative AI in corporate analysis and client-facing tools.

1. **Challenges and Limitations**
2. **Challenges Faced During the Project**

The AI-powered cash flow forecasting tool, developed for this capstone project, advanced client cash flow management through predictive analytics. However, its approach and dataset had limitations that required acknowledgment for a balanced assessment.

1. **Dataset Limitations**

The dataset relied on 30 years of historical Apple financial data (1994–2024) from EODHK, including total revenue, operating expenses, and working capital. While robust for modeling Apple’s trends, its company-specific nature limited generalizability to other firms. Apple’s unique market position and focus on premium products did not mirror other industries or smaller businesses, restricting the tool’s applicability for Zenith’s diverse client base. Additionally, the absence of external factors—such as market trends or economic indicators like inflation and GDP—weakened predictions. For example, a recession could have reduced consumer spending on Apple products, an impact not captured by historical data alone, which potentially led to inaccurate forecasts in dynamic economic conditions.

1. **Model Limitations**

The tool used linear regression to predict revenue and expenses, assuming a linear relationship with features like year, lagged revenue/expenses, and working capital percentage. This assumption missed non-linear patterns, such as rapid growth from product launches (e.g., the 2007 iPhone) or supply chain disruptions. Standardized inputs and metrics like Mean Squared Error (MSE) and R-squared supported the model, but they could not address this limitation. Moreover, the 2025–2029 forecasts, based on historical data, failed to account for unforeseen events like geopolitical conflicts or regulatory changes, which could have altered Apple’s financial trajectory. This reliance on past trends limited the tool’s robustness for long-term planning, highlighting the need for further refinement.

1. **Future Work**

Despite its limitations, the AI cash flow forecasting tool provided a solid foundation for expansion through innovative enhancements. The following sections outlined potential improvements to boost functionality, accuracy, and user experience.

1. **Customizable Scenario Analysis**

A key enhancement is customizable scenario analysis, enabling users to adjust growth rates and variables via Streamlit dashboard sliders. Currently, the tool uses three scenarios—optimistic (10% revenue increase, 5% expense decrease), base (no change), and pessimistic (10% revenue decrease, 5% expense increase)—to evaluate cash flow. Allowing user-defined inputs, such as a 15% revenue boost from a new product or a 7% expense rise due to inflation, would let stakeholders tailor forecasts to specific strategies. This flexibility enhances decision-making but requires a dynamic forecasting pipeline, a complex task deferred due to time constraints in this capstone project.

1. **Email/SMS Notifications**

Integrating email or SMS notifications for cash shortfalls (below 5 billion USD) is another valuable addition. The current dashboard shows alerts visually, but automated notifications would enable real-time responses, such as emailing managers about a 2027 shortfall in the pessimistic scenario. This could use smtplib for email or Twilio for SMS, though it involves external service setup and authentication. Time limitations and the prototype’s focus on visual alerts led to its exclusion, though it’s a future priority for enterprise use.

1. **Sentiment Analysis**

Incorporating sentiment analysis from news or social media could improve forecast accuracy by reflecting public perceptions, like positive iPhone buzz or supply chain concerns. Using natural language processing (NLP) to extract sentiment scores from Twitter or financial news as model features is promising but requires external APIs, advanced libraries, and significant computational resources. This complexity, combined with the project’s emphasis on historical data, deferred its implementation for future resource availability.

1. **Integration with External Data Sources**

Enriching models with external data, such as GDP or unemployment rates via APIs (e.g., Alpha Vantage), would address the dataset’s lack of contextual factors. This could predict revenue drops during economic slowdowns or expense shifts from labor costs, with live updates enhancing responsiveness. However, API setup, data alignment, and cost considerations make this challenging. The project’s time-constrained scope favored static analysis, leaving this as a future goal to boost predictive power.

1. **Conclusion**

The AI-powered cash flow forecasting tool was successfully developed to enhance client financial management, using Apple’s historical data from 1994 to 2024 as a case study due to its stable financial performance and rich dataset. The tool gathered and cleaned financial data, identified key cash flow trends, and built linear regression models to predict revenue and expenses with high accuracy, as evidenced by low Mean Squared Error and strong R-squared values. It incorporated sensitivity analysis for optimistic, base, and pessimistic scenarios, delivering clear insights through interactive Tableau visualizations, including sales, cash flow, and profit trends. An LLM-powered chatbot was implemented to answer queries based on Apple’s annual reports, enhancing user interaction. Despite limitations, such as the dataset’s company-specific nature and the models’ reliance on linear assumptions, the tool provided a robust framework for strategic financial planning, laying a foundation for future enhancements like customizable scenarios and external data integration.

1. **Reference**

EODHD

<https://eodhd.com/>

Scikit-learn - machine learning in Python

<https://scikit-learn.org/stable/>

1. **Distribution of Work**

Dirk – Regression models & PowerPoint

Philip – Data Preprocessing & EDA

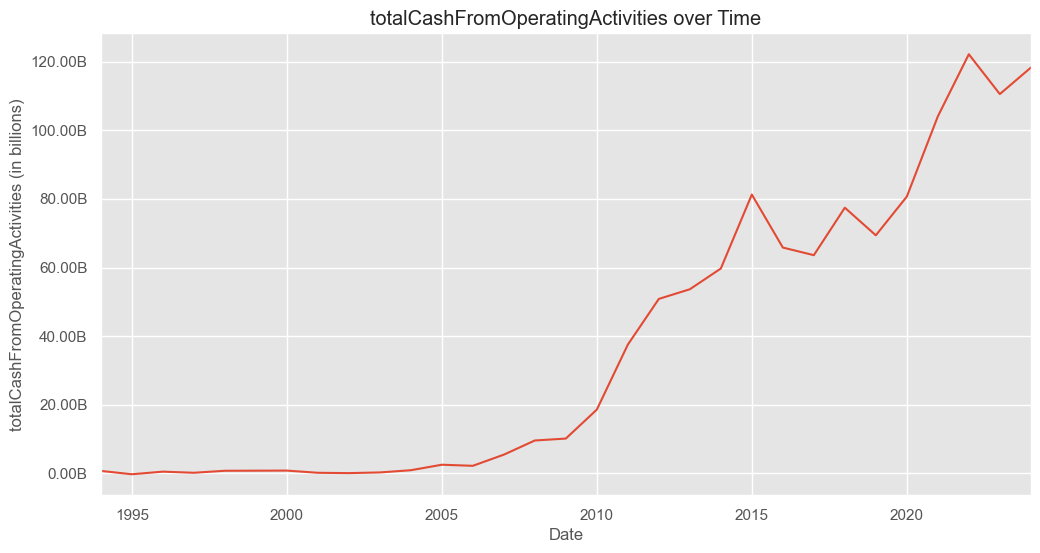
Olivia & Carmen – Visualization dashboard (Tableau)

Jacob - Alert systems and cash flow analysis

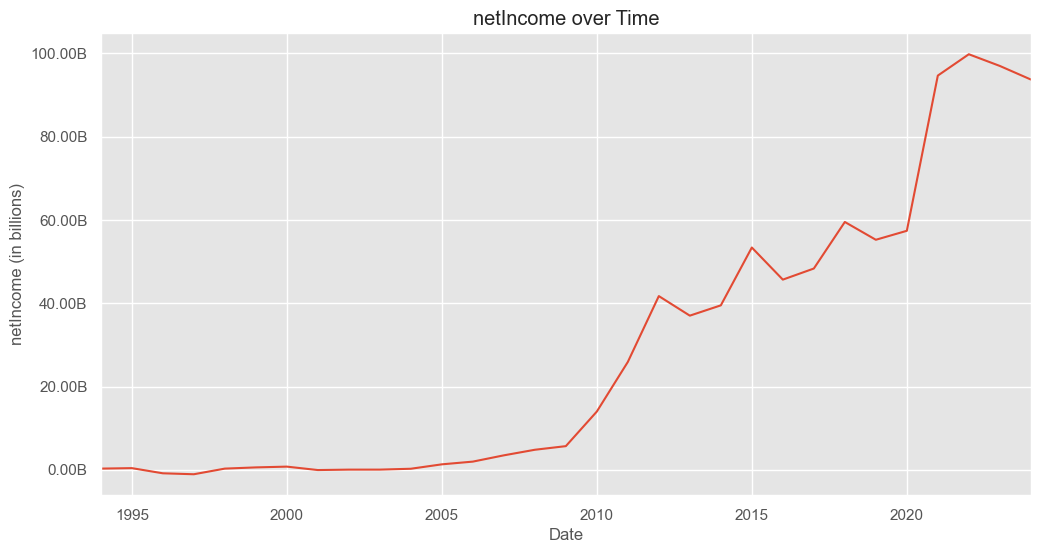
Teddy - LLM Application - AI Chatbot

1. **Appendices**

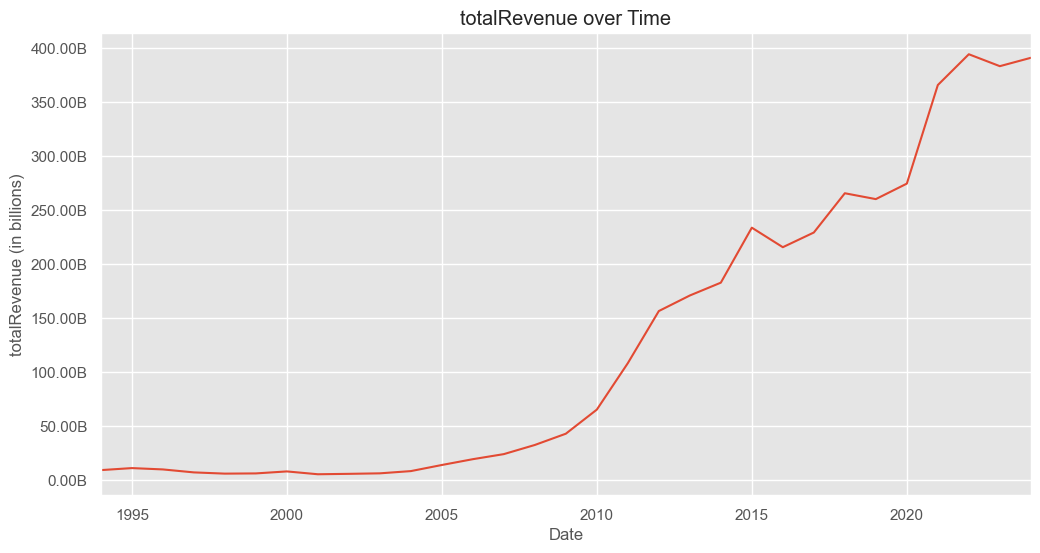
Figures 1.1



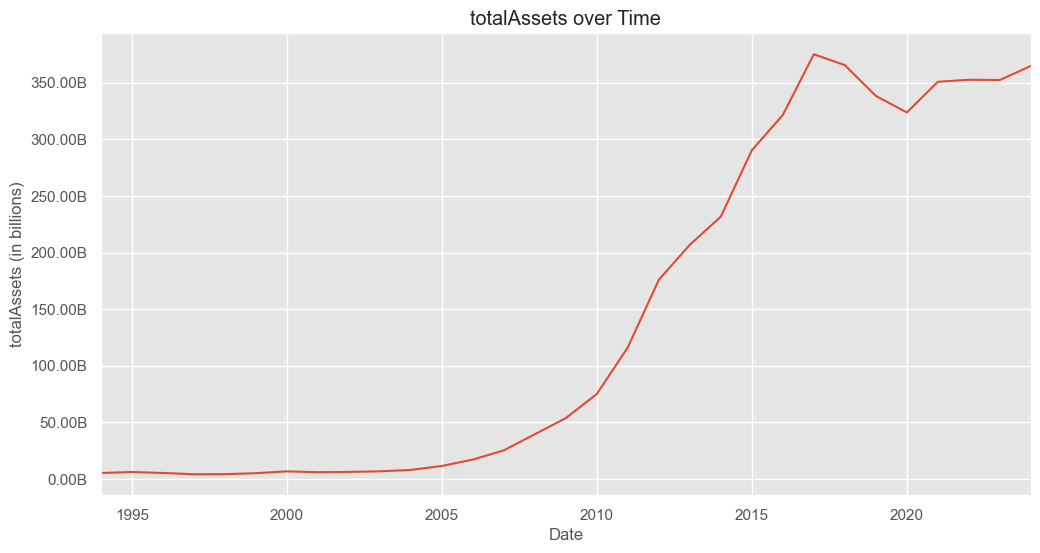
Figures 1.2



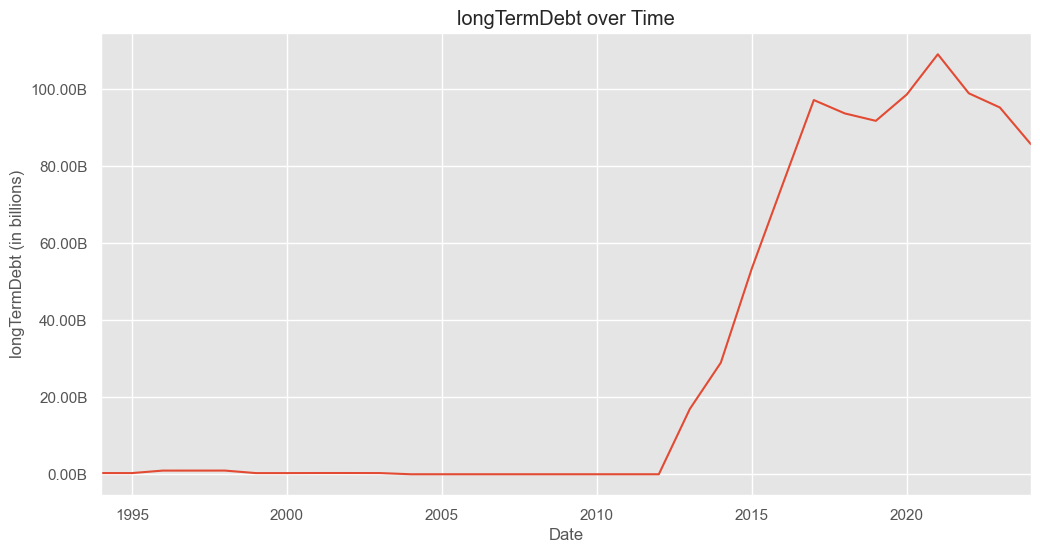
Figures 1.3



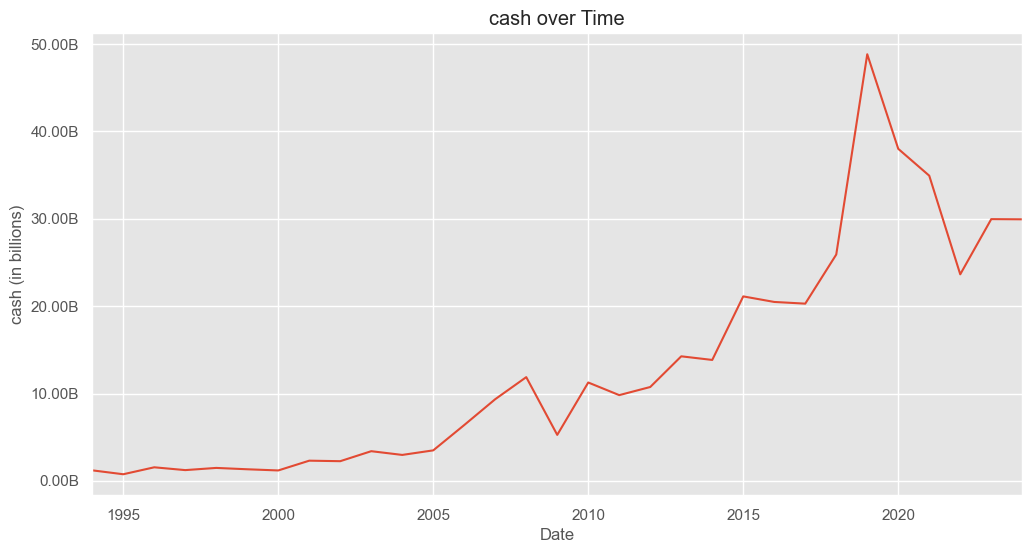
Figures 1.4



Figures 1.5



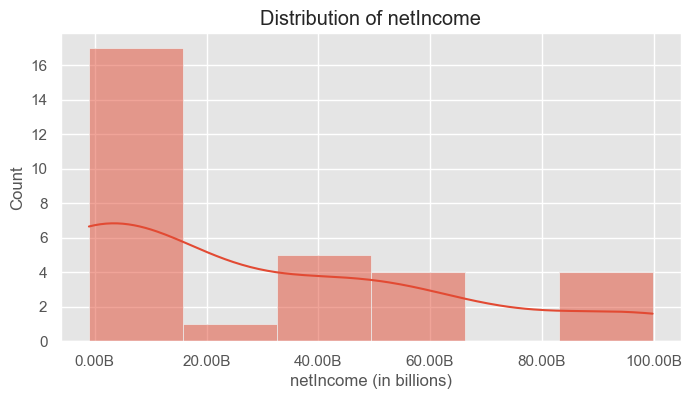
Figures 1.6



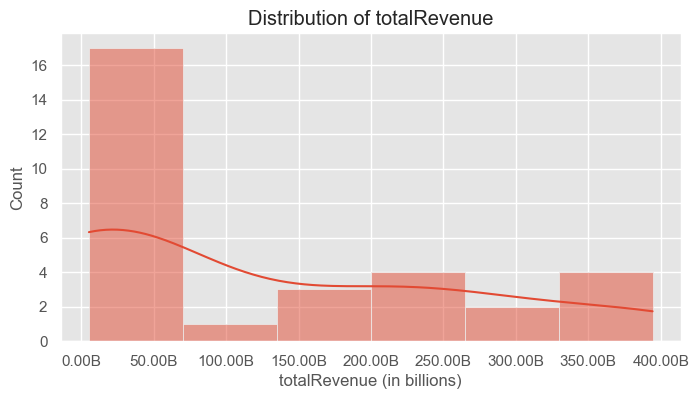
Figures 1.7



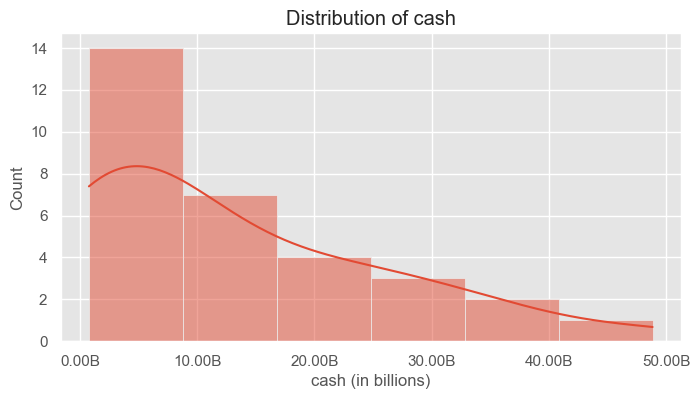
Figures 2.1



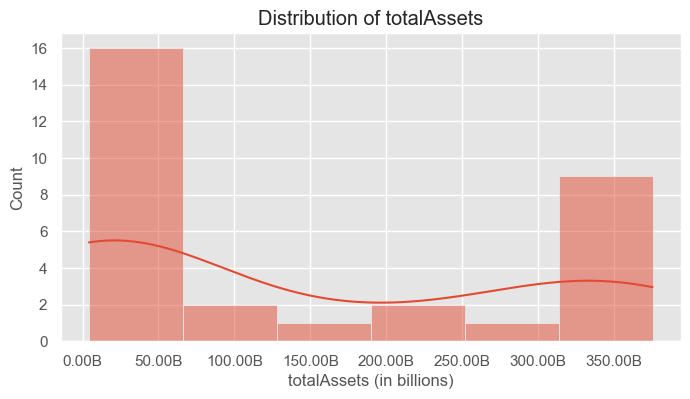
Figures 2.2



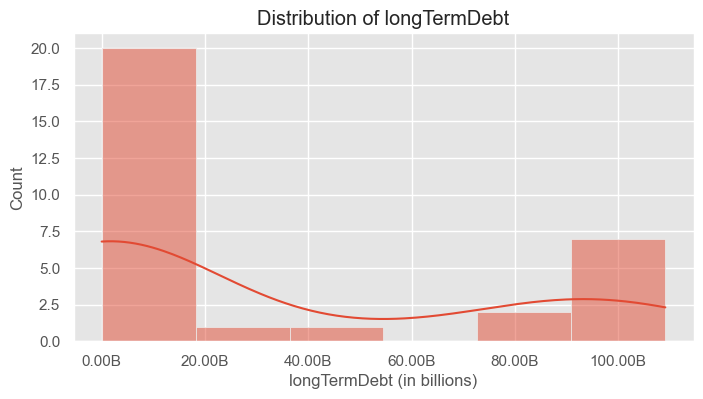
Figures 2.3



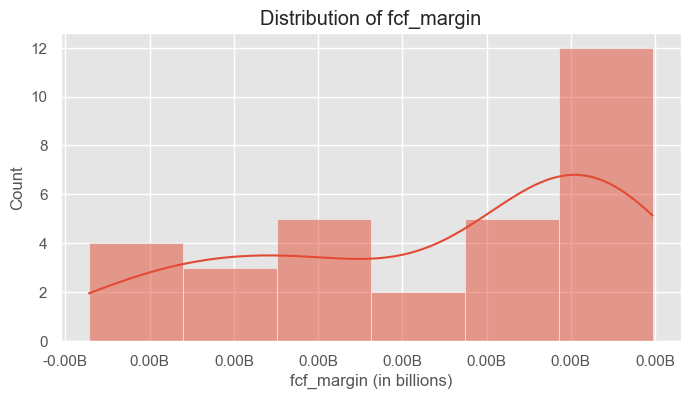
Figures 2.4



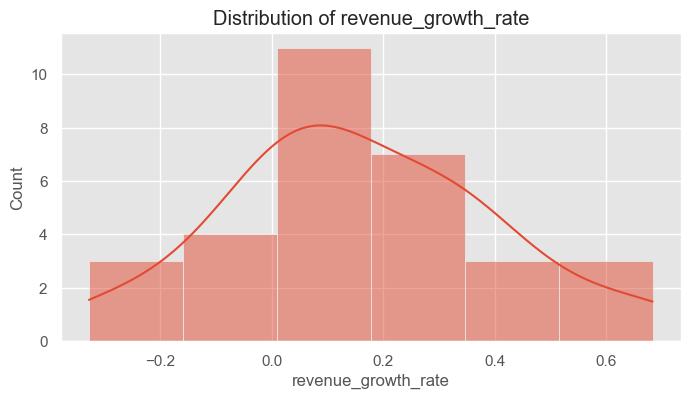
Figures 2.5



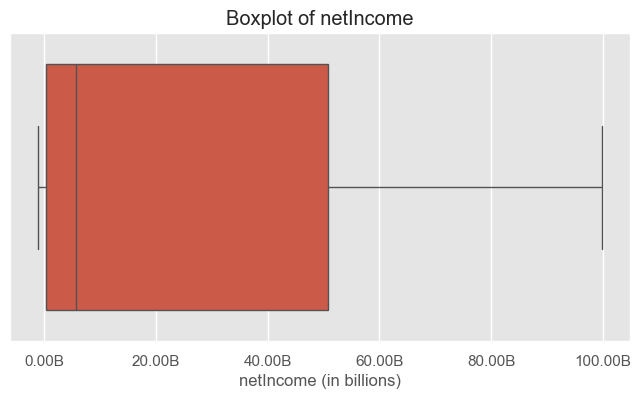
Figures 2.6



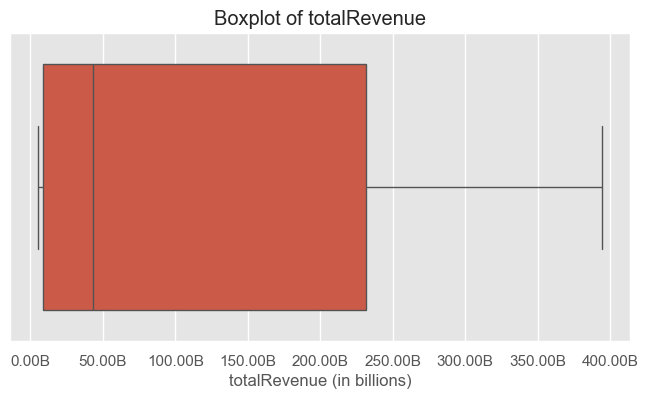
Figures 2.7



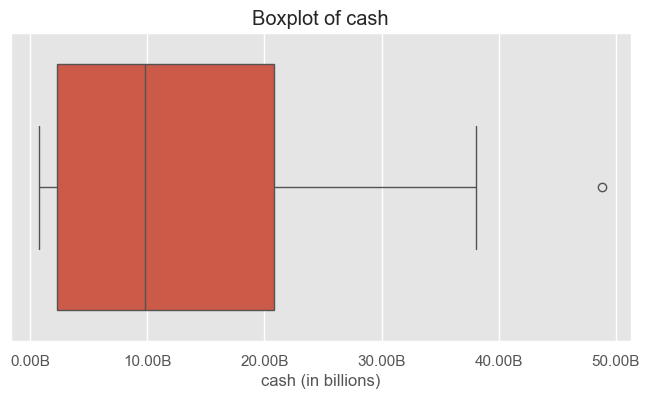
Figures 3.1



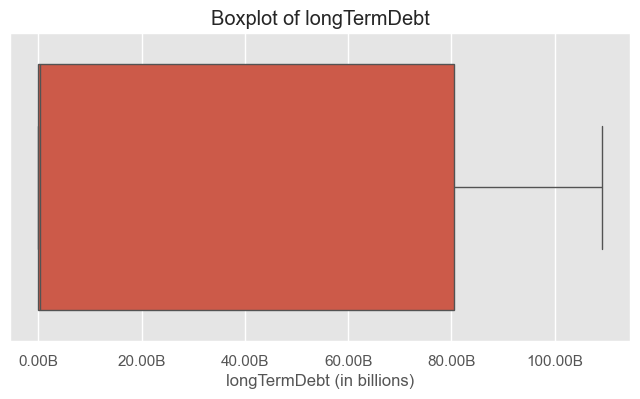
Figures 3.2



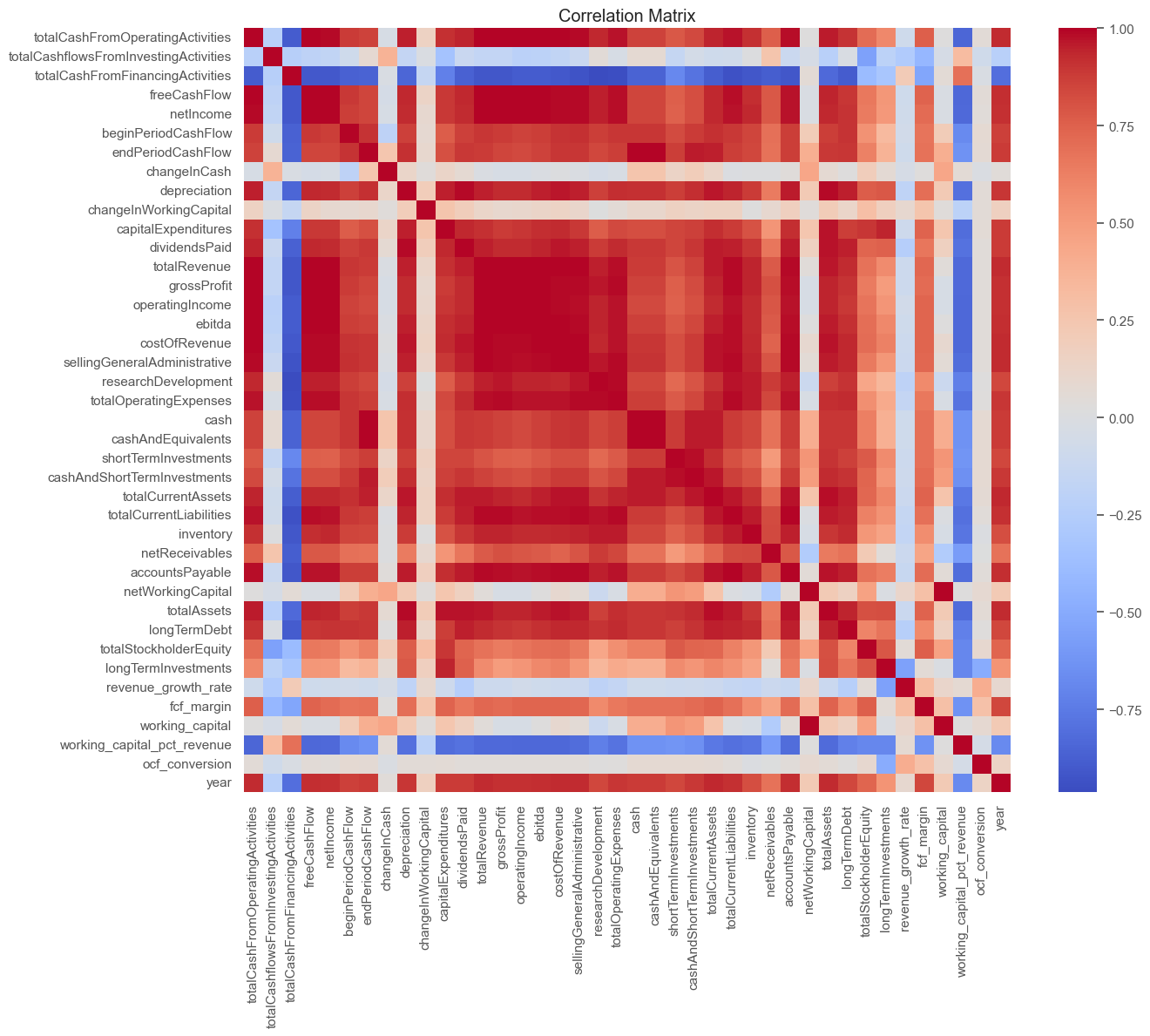
Figures 3.3



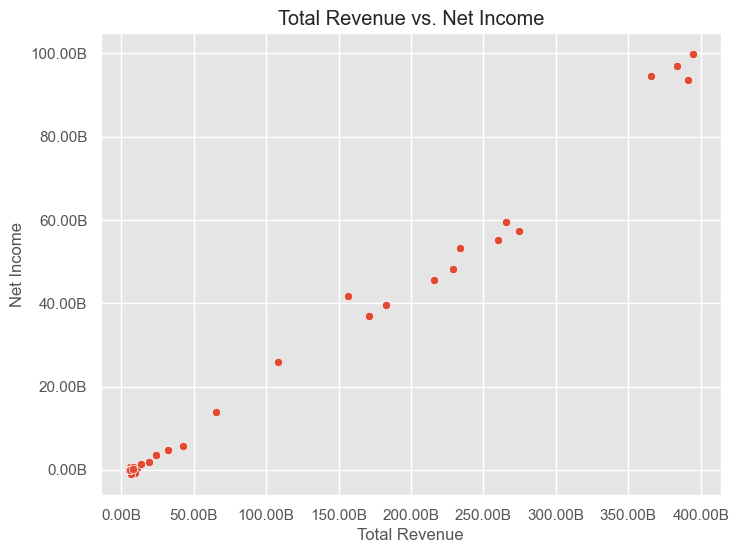
Figures 3.4



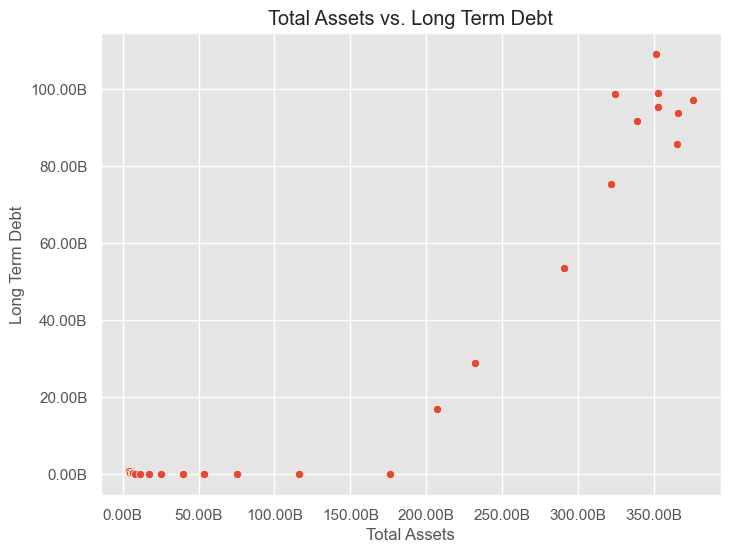
Figures 4.1



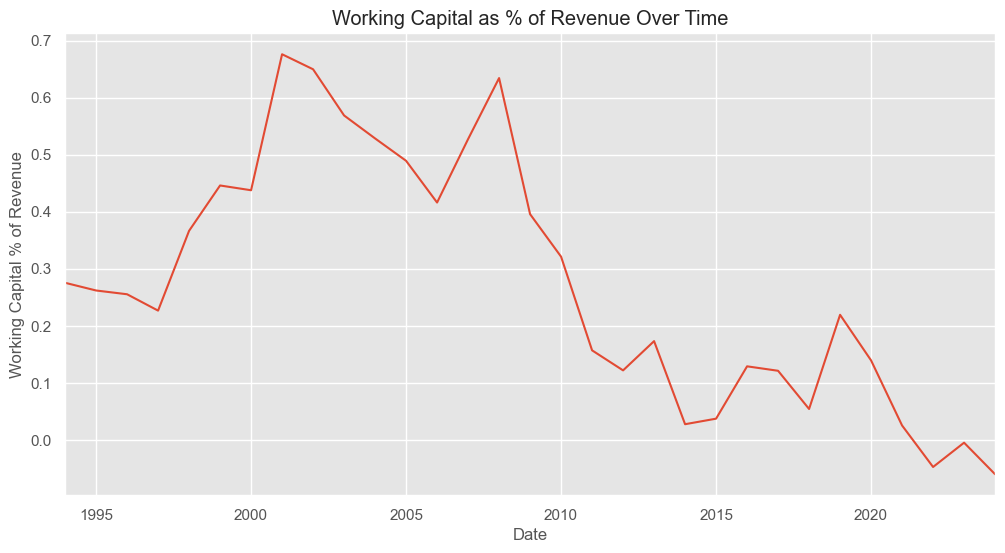
Figures 5.1



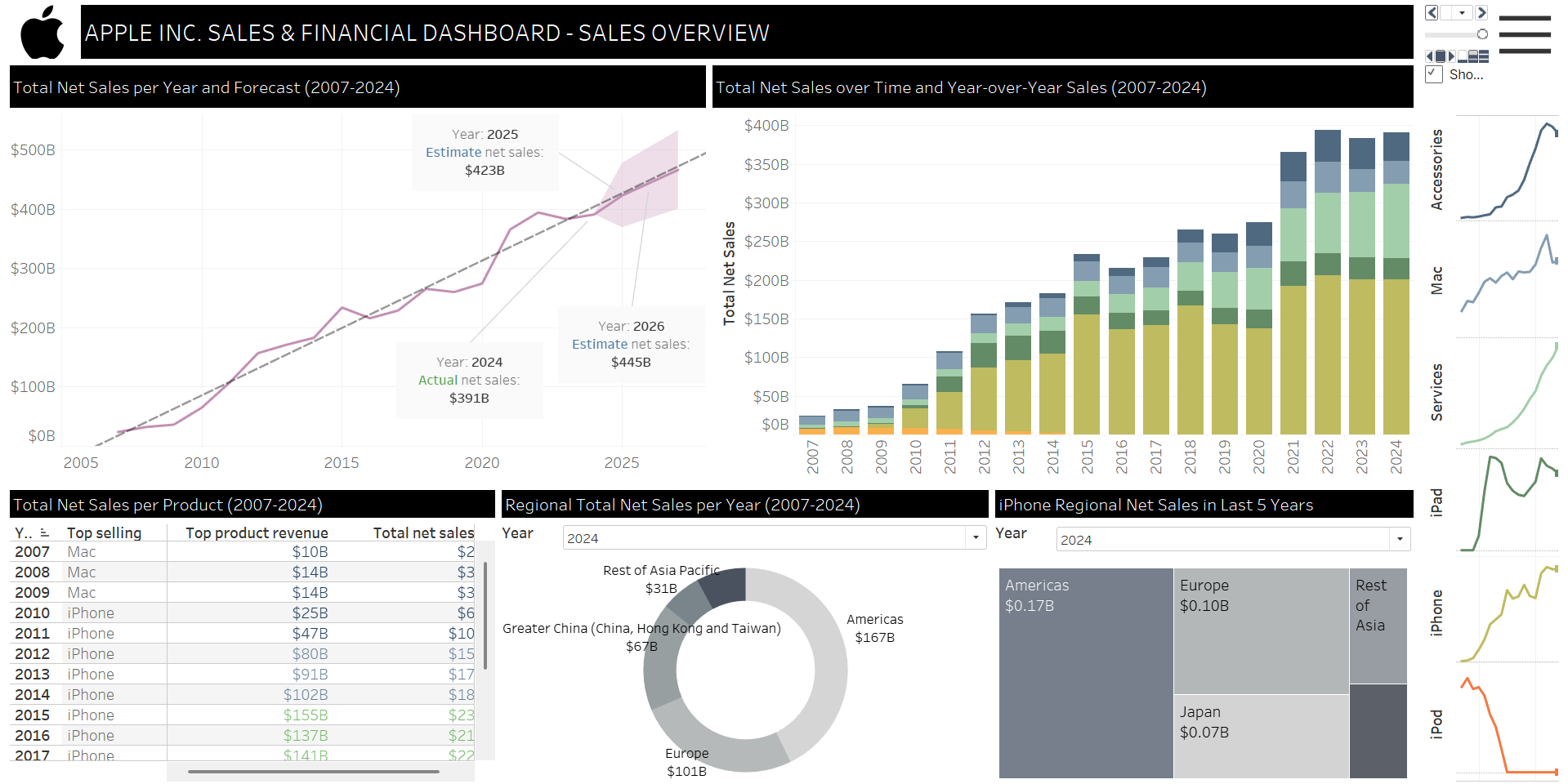
Figures 5.2



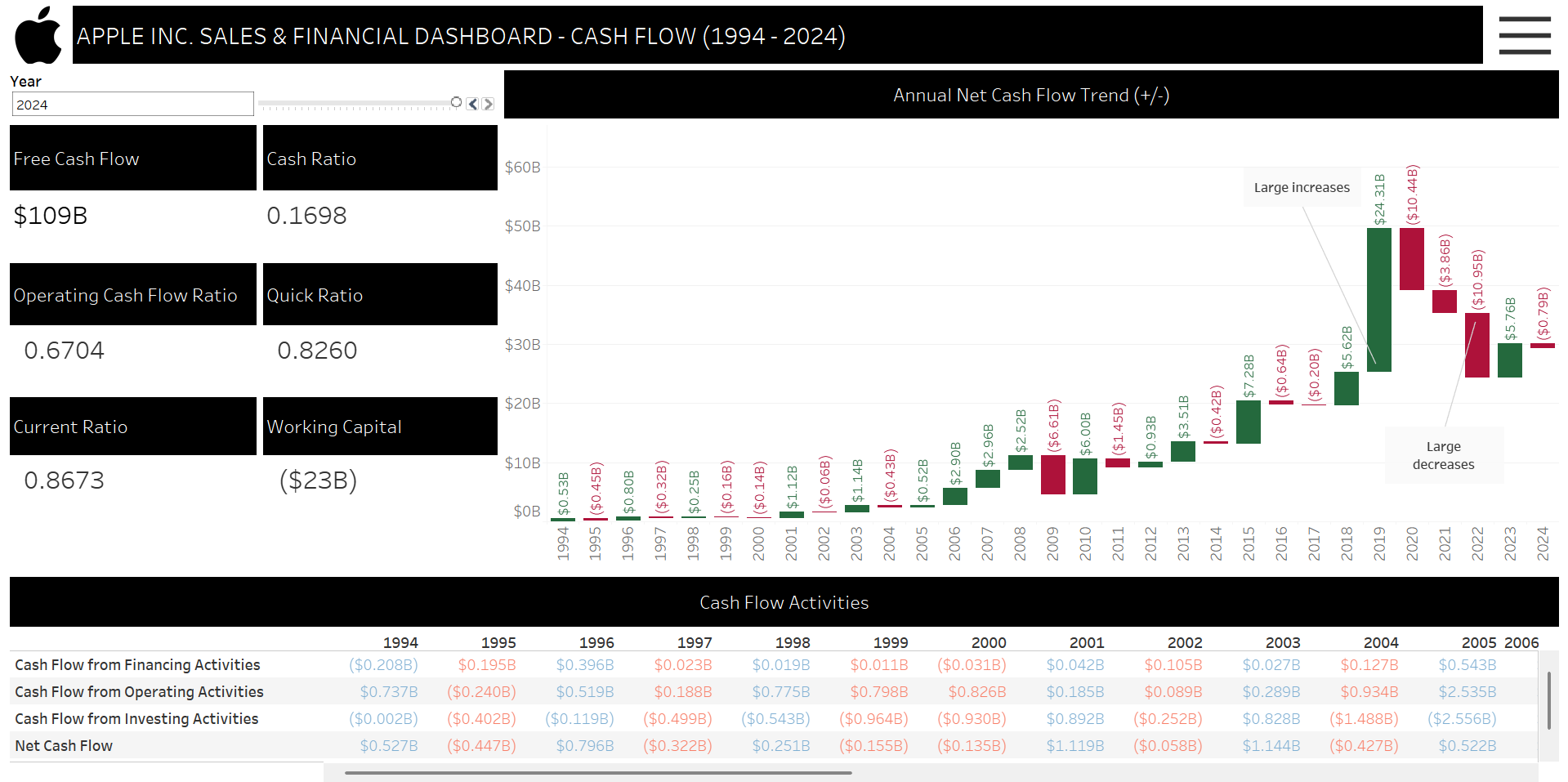
Figures 6.1



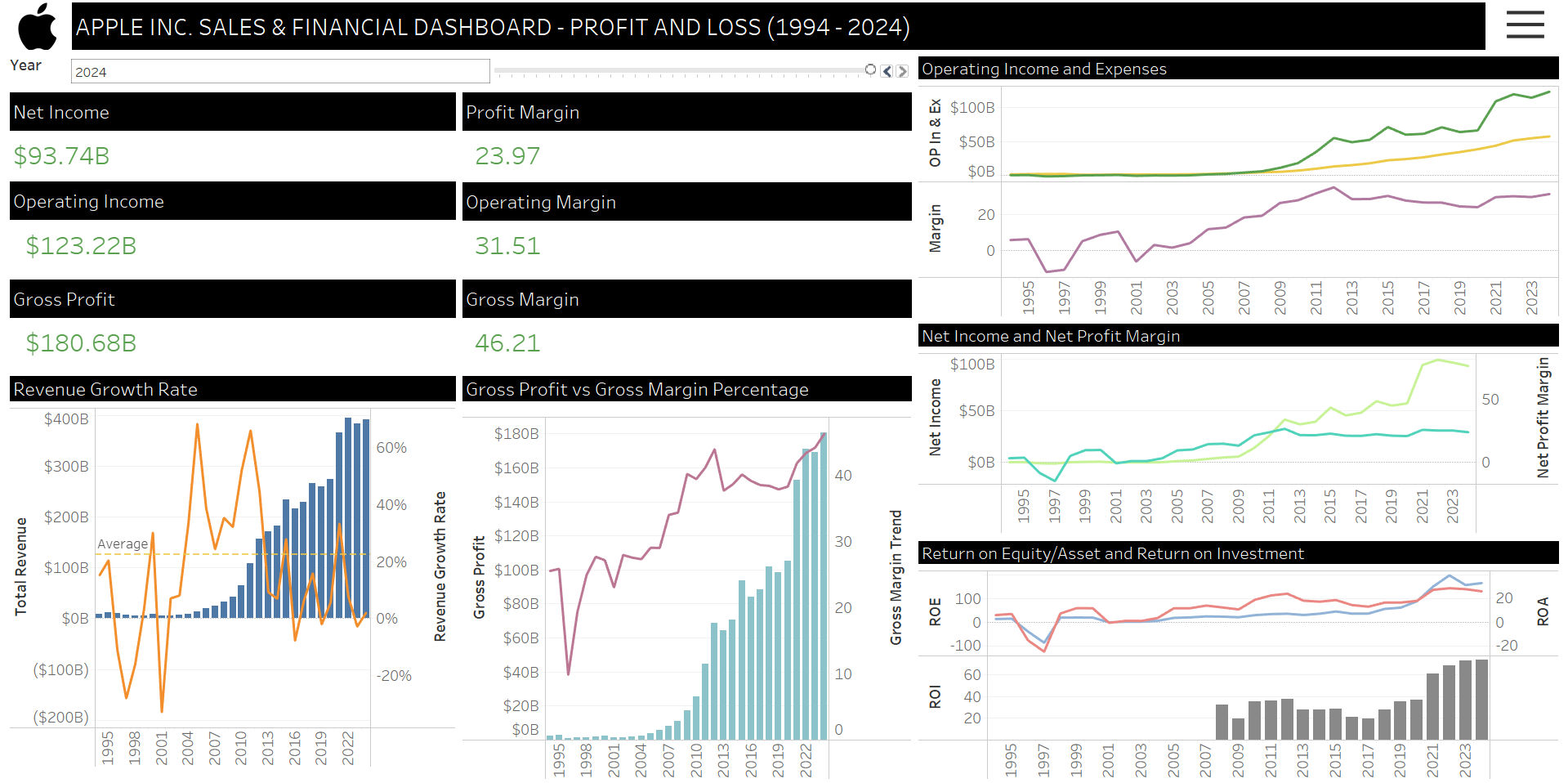
Appendix A - Dashboard 1: Sales Overview 2007 - 2024



Appendix B - Dashboard 2: Cash Flow 1994 - 2024



Appendix C - Dashboard 3: Profit and Loss 1994 - 2024



Appendix D - Dashboard 4: Balance Sheet 1994 - 2024

