Financial Visualization Written Report

David Pogrebitskiy, Jacob Kulik
Khoury College of Computer Sciences, Northeastern University

Abstract

Through the investigation of balance sheet accounts across companies, a domain task that we believe the end user in our use case will perform will be to explore how firms in different industries are made up of assets, liabilities, and equity over the course of time. It is important to support these tasks as it will help newcomers to accounting and finance better understand trends in different industries and the customers of these firms better understand where their money is being utilized.

1 Introduction

A balance sheet is a financial statement that reports a company's assets, liabilities, and equity at a specific point in time. Understanding how a balance sheet works is essential for anyone interested in finance or accounting, as it provides insights into a company's financial health, solvency, and liquidity. By analysing a company's balance sheet, stakeholders can make informed decisions about investing, lending, or doing business with that company, which can have significant implications for their financial performance and profitability.

This project seeks to analyse the distribution of resources and sources of capital across 16 different companies over a 12-year period and offer insightful information about how they operate. For finance and accounting students who wish to better understand the basics behind a company's structure, the analysis will also assist in identifying any patterns or modifications in a firm's asset strategy over time, offering a more complete picture of its sector. Utilizing data from 10-K forms, which are required to be filed yearly for publicly traded companies, the visualization will make it easier to digest the information on these forms.

2 Related Work

Financial statements provide essential data points for the evaluation of a company's financial health in areas like productivity, leverage, liquidity, profitability, and value. Understanding these health measures can aid in making informed decisions about investing, lending, and doing business with companies. Priester and Wang (2010) discuss the numerous financial ratios and indicators that can be utilized to establish quantitative measures of health, all of which can be derived from a business's financial statements [1]. Overall, they offer valuable insight into the importance of financial statements analysis, which can be relevant for researchers, practitioners, and students studying business and accounting.

Visualization is often considered a crucial method for analysing these financial statements and data, especially with the increasing amounts of data available today. Tekusova and Kohlhammer (2007) found that systems that utilize animation, focus, zoom, details on demand, and time period selection were positively received by both users within and outside finance. The interactivity and dynamically changing visualizations allowed users to personalize their experience with the system, which was particularly useful for finance professionals who need to analyse data quickly and efficiently [2]. This supports the relevance of visualization in financial analysis. Utilizing data from 10-K forms can, similarly, help users gain an understanding of resource and capital distribution across different companies over time.

3 Use Case

We propose a balance sheet visualizer tool that provides users with an in-depth view of the financial health and stability of various companies across a 12-year period. The tool is designed for individuals with limited accounting or financial knowledge who want to gain a better understanding of how companies allocate their resources and manage their finances over time. This intended user would benefit from a visual display of both industry and firm values.

In this scenario, a user would use the visualization tool to compare the balance sheets of different companies in various sectors. The tool would allow the user to highlight different components of the balance sheet, including the factors that make up assets, liabilities, and equity. By presenting the data in an interactive and visually engaging format, the tool would enable the user to quickly identify patterns and trends in the financial performance of different companies.

4 Data

Bank data was queried directly from the Compustat North American Fundamentals Yearly database through the Wharton Research Data Services (WRDS). WRDS is a well-known data provider of business-related data run by the University of Pennsylvania. Although WRDS requires an account to query from their databases, Northeastern students are easily able to apply for an account that is sponsored by the university. The dataset that resulted from our query on WRDS is here.

There are several biases and ethical considerations that may be present in yearly financial reports. People argue that the requirement for reports pushes companies to prioritize short-term decisions to make reports look better while sacrificing future success. These pushes for short-term results often negatively affect work environments and employees because the performance expectations are kept high. On the opposing side, people argue that these reports are necessary for consumers to make wise financial decisions. Yearly reports don't always expose the true performance of a company but utilizing many of them in collaboration may help reveal the truth. However, these reports are audited by an independent company, so they are trustworthy, even with confirmation bias.

The original dataset contained quarterly asset information across a range of 10 years for Apple, Verizon, Coca-Cola, Target, Exxon Mobil, Intel, IBM, Johnson & Johnson, Merck, 3M, Pfizer,

Stryker, Union Pacific, Microsoft, Accenture, Netflix from January 2010 to January 2023. After extracting the data from the database, we noticed there were several columns full of missing values. Because this dataset was rich in columns, we found no issues in just deleting them. The removed columns are as follows: Assets Held for Sale, Provision of Other Assets, Aggregate Miscellaneous Assets, and Deferred Tax Assets

5 Design Process

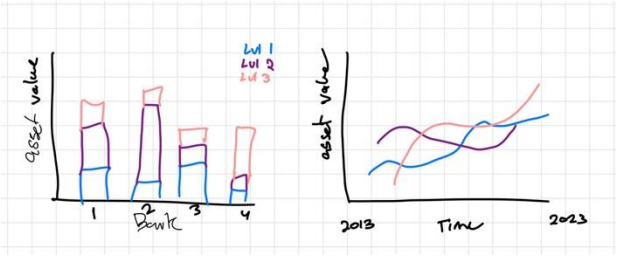

Figure 1: Sector and firm view. Stacked chart and time series line chart that shows proportions in both a vertical cross section and time series format.

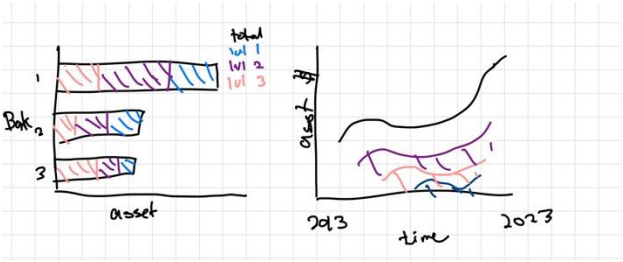

Figure 2: Sector and firm view. Horizontal stacked chart and time series area chart that shows proportions in both a horizontal cross section and time series format.

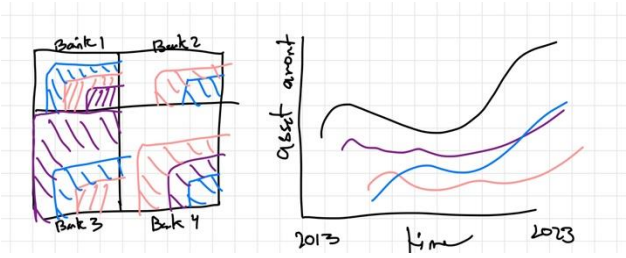

Figure 3: Sector and firm view. Area plot and time series line chart that shows proportions in with parts of a whole and time series format.

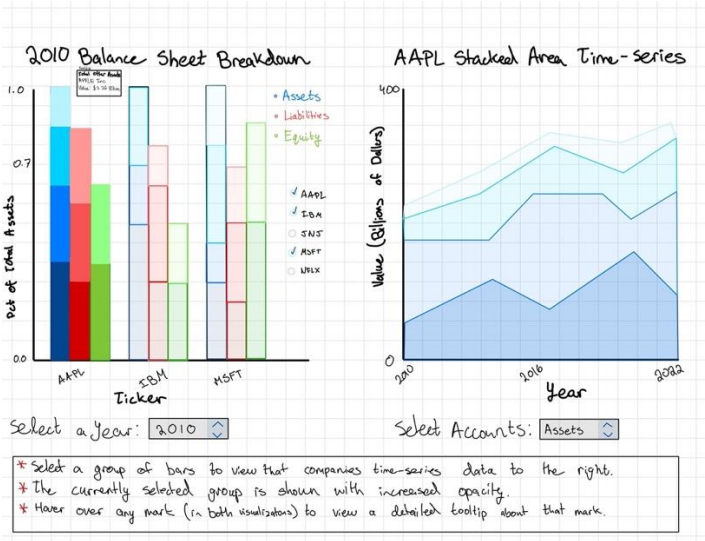

Figure 4: Final design. Stacked chart and time series area chart that shows proportions in both a vertical cross section and time series format.

The final sketch is a culmination of the rough sketches. After narrowing down to a sector-wide and firm-specific dual view, the specific method of displaying both is what was tweaked. In order to better display cross section balance sheet information, we narrowed down to a stacked bar chart method as there are different values that make up total assets, liabilities, and equity. The vertical stacked bars from the first sketch were utilized for this fact. Second, the variation in time series visualization was best with a focus around highlighting a trend over time, so a stacked area was utilized. Here, it shows a trend over time while also showing how different values make up parts of the whole.

The left visualization focuses on a line mark, with color, size, and both positions as channels. The size of the line marks represents different proportions, while the color helps with categorizing the items. Position also helps refer to the value and company. The right visualization focuses on an area mark, with area, color, and both positions as channels. The areas represent proportions, with color helping categorize and position referring to year and value.

The visualization is designed to be as interactive as possible, with hover tooltips and feature filtering on both. Year can be specified on the left visualization, as well as which companies to inspect. The right visualization is updated to reflect the company whose bar was clicked on the left visualization. The views are meant to be used together, as the left gives a very broad understanding of trends for a given year, while the right dives deeper into a specific company. With this functionality, we hope it is as intuitive and beginner friendly as possible.

During the implementation process, the final visualization design evolved to fit our need to make the data as easy for a beginner to use as possible. This meant adding interactive elements, as well as descriptions and a concise legend. The goal was to utilize hover functionality to make it as intuitive as possible for someone to pick up and explore. We chose to focus solely on assets/liabilities/equity in order not to confuse the user with the liabilities and equity stacked graph. We also decreased the number of default tickers in order to stop the cramming of bars. We conducted 4 main tasks for our usability testing; how

many total current assets does Verizon (VZ) have in 2010, filter down the visualization to solely MMM, calculate the difference (in billions of dollars) between Target's PPE and IBM's PPE in 2015, and describe the trend over time of Apple's stockholders' equity. These tasks were simple for our users, with a minimal number of clicks and a short period of time required per task. It was tricky for certain users to figure out how to look at the stockholders' equity trend, but they were able to spend more time on it to figure it out.

6 Final Design

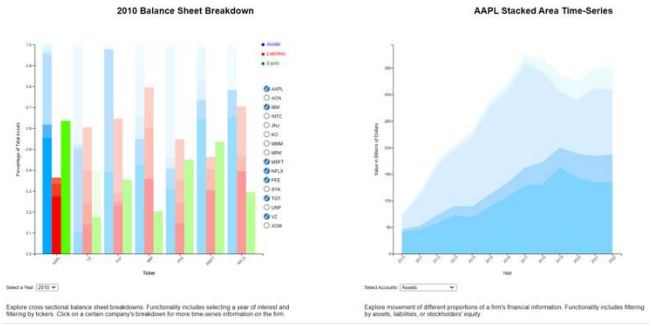

Figure 5: Final Design Implementation

The final design aims to focus on the combination of an industry and firm-wide analysis. The left plot is a stacked bar chart covering cross-sectional data at a certain point in time, where the length of different stacked bars represents the percentage of the whole. Here, length and color are channels that represent different categories of assets. Both horizontal and vertical positions are used to categorize the information to a specific firm. Through linking, a certain firm is highlighted, and the time series information is conveyed with a stacked area chart, as it highlights the sum and movement of values over time. Here, position is a channel as it shows direction and a general trend over time. Color is also used as a channel to categorize the different values making up the account. The year can be selected by the user to delve deeper into a specific point in time. Clicking on a company's values updates the firm-specific visualization. The legend will be utilized to show the aspects of a balance sheet with the given color channel.

If our target user, an individual with limited accounting or financial knowledge, wanted to solve our domain problem, or the different financial parts that make up a company and how they would change over time, they would be able to use our visualization. Firstly, they would observe the different breakdowns of assets, liabilities, and equity for the default companies using the tooltip. Next, they would add or remove companies from the graph. After this, they would click on a company of interest and use the second view to delve deeper. Here, they would change the specific account and use the second tooltip to observe the time-series changes.

7 Discussion

We hope the analysis of the breakdown assets, liabilities, and equity across companies over a twelve-year period will provide valuable insights to new students in the field of finance. The goal of our analysis was to reveal trends over time banks and highlight how different types of companies allocate their funds differently, and this goal was achieved.

The tool fully addresses the domain problem we set out to solve, which was to explore how firms in different industries are made up of assets, liabilities, and equity over the course of time. Certain limitations, however, are that the visualization only encompasses twelve companies and does not allow for different ticker queries. Due to inconsistencies with the data associated with certain companies on WRDS, not all companies can be used. Further, some companies have negative equity, leading to plotting errors. However, incorporating the WRDS API in the future would be an interesting way to expand on the project.

In the future, improvements can be made to the stacked bar visualization include options to stack liabilities and equity for better clarity. Further, more datapoints, such as total revenue, can be used and visualized for the user to explore. Since 10-K forms contain vast amounts of information, there are always extensions to the views shown and the data that gets included.

8 Conclusion

In conclusion, we believe that through a sector and firm approach to visualizing asset, liability, and equity breakdowns of different firms, we can help students better understand the basics behind a balance sheet. Supporting these tasks is crucial as it can aid individuals new to the fields of accounting and finance in comprehending industry trends, while also applying what they've learned in class. Contributions throughout this project were split up equally. David covered the stacked bar charts and general plot generation, while Jacob created tooltips, linked views, and made the stacked area chart. The final report and demo video were equally split.

References

[1] C. Priester, J. Wang. Financial statements analysis. In *C. Priester, J. Wang (Eds.), Financial Strategies for the Manager* (pp. 19-39). Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-540-70966-4_2.

[2] T. Tekusova and J. Kohlhammer, "Applying Animation to the Visual Analysis of Financial Time-Dependent Data," *2007 11th International Conference Information Visualization (IV '07)*, Zurich, Switzerland, 2007, pp. 101-108, doi: 10.1109/IV.2007.28.

Appendix

[1] Data Abstraction

a. Each row in the data represents the financial information for a given firm during a given year from 2010-2022

b. Categorical Attributes: gvkey, tic (Ticker), conm (Company Name), indfmt, consol, popsrc, datafmt, curcd, costat

c. Sequential Attributes: fyear, datadate

d. Continuous Attributes: act (Total Current Assets), ppent (Total Property, Plant, and Equipment), intan (Total Intangible Assets), ivaeq (Investments and Advances - Equity Method), ivao (Investments and Advances - Other), ao (Total Other Assets), lct (Total Current Liabilities), txditc (Deferred Taxes and Investment Tax Credit), lo (Other Liabilities), dltt (Total Long-Term Debt), ceq (Total

Common/Ordinary Equity), pstk (Total Preferred/Preference Stock (Capital)), mibn (Nonredeemable Noncontrolling Interest), at (Total Assets), lt (Total Liabilities), teq (Total Equity)

[2] Task Abstraction

a. High level -> Consume -> Discover
    i. The visualization being both sector-wide and firm-specific views
b. Targets of trends and distribution
    i. Stacked bar charts highlight distribution and stacked area time series data conveys trends