

Task 3:

Presentation

Keniyah Chestnut

Analytics Programming - D598

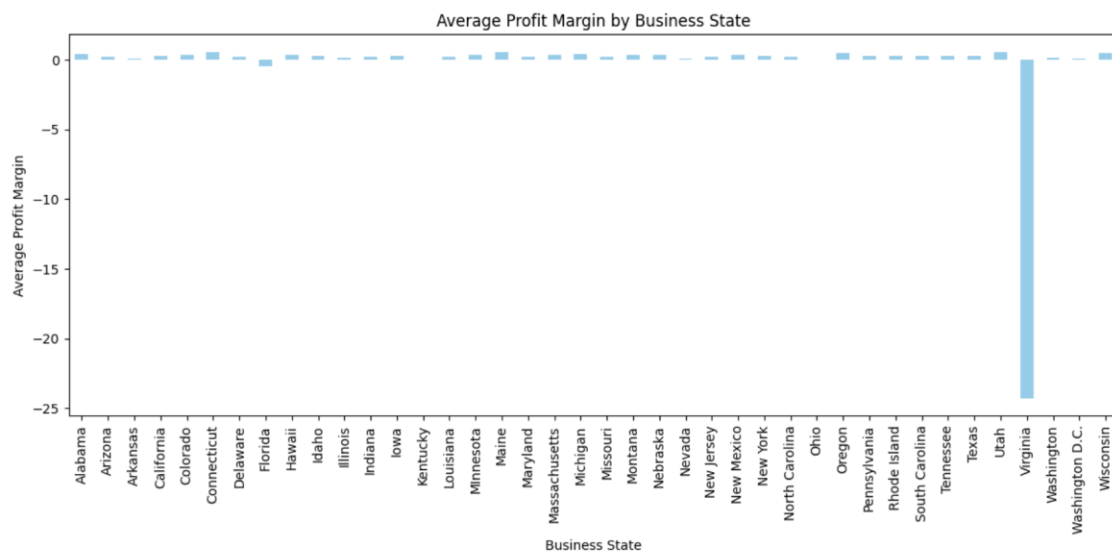
SID: 012601305

A. Code Explanation

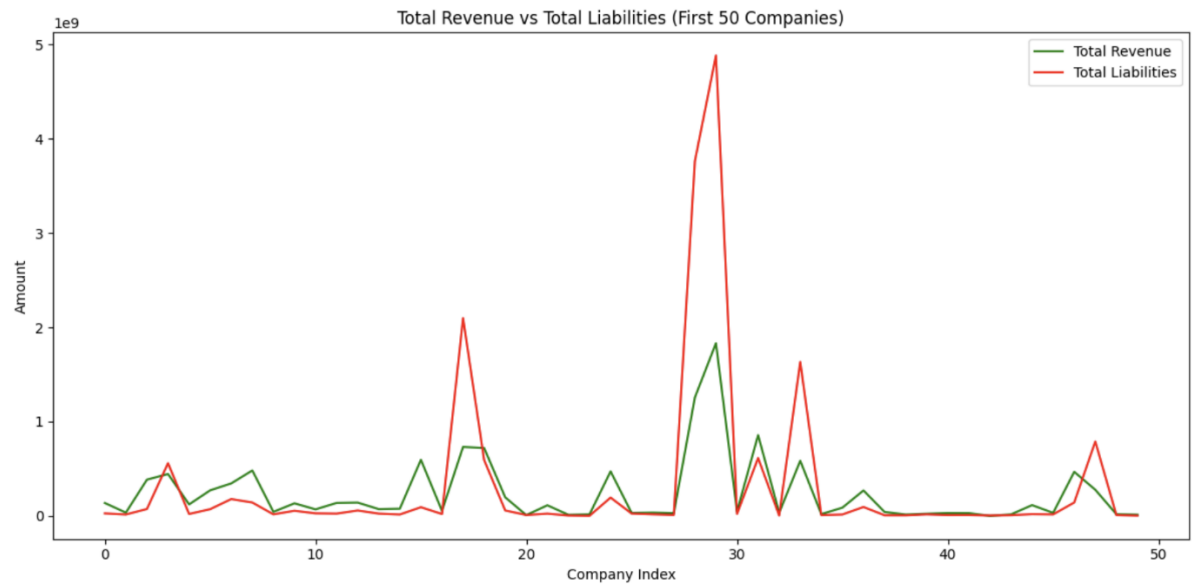
The program first imports necessary libraries (pandas and matplotlib) and loads the D598 dataset into a DataFrame, which aligns with Step 2 of the pseudocode. It identifies duplicate rows based on Step 3. Records are then grouped by the "Business State" column to calculate mean, median, minimum, and maximum statistics for major financial fields (Steps 4 and 5). These results are stored in a new DataFrame (Step 6). The program filters businesses with negative debt-to-equity ratios (Step 7) and builds a new debt-to-income DataFrame by safely handling division-by-zero cases (Step 8). Finally, the debt-to-income information is merged back into the original data to create a completed final dataset (Step 9). This logic follows the process outlined in the flowchart and pseudocode created in Task 1.

B. Visualization

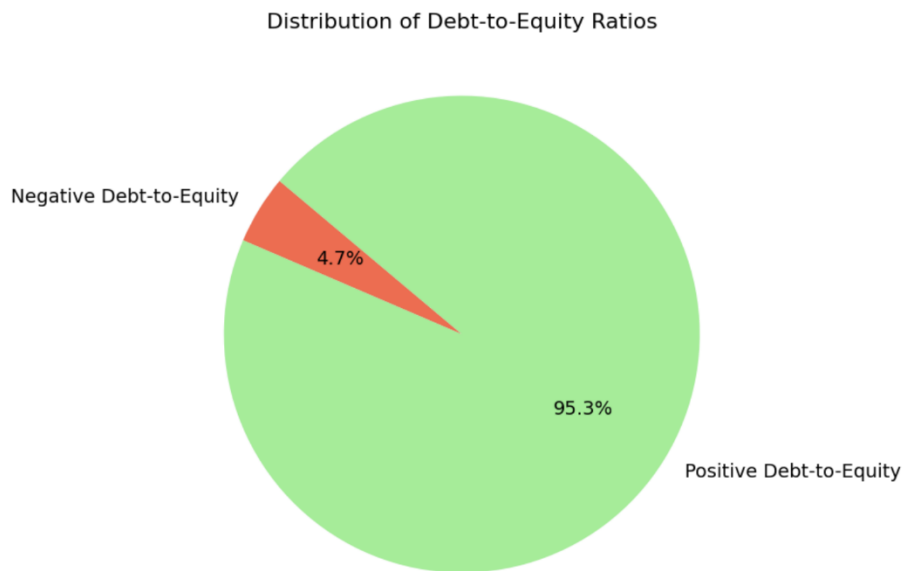
Visualization 1: Average Profit Margin by Business State



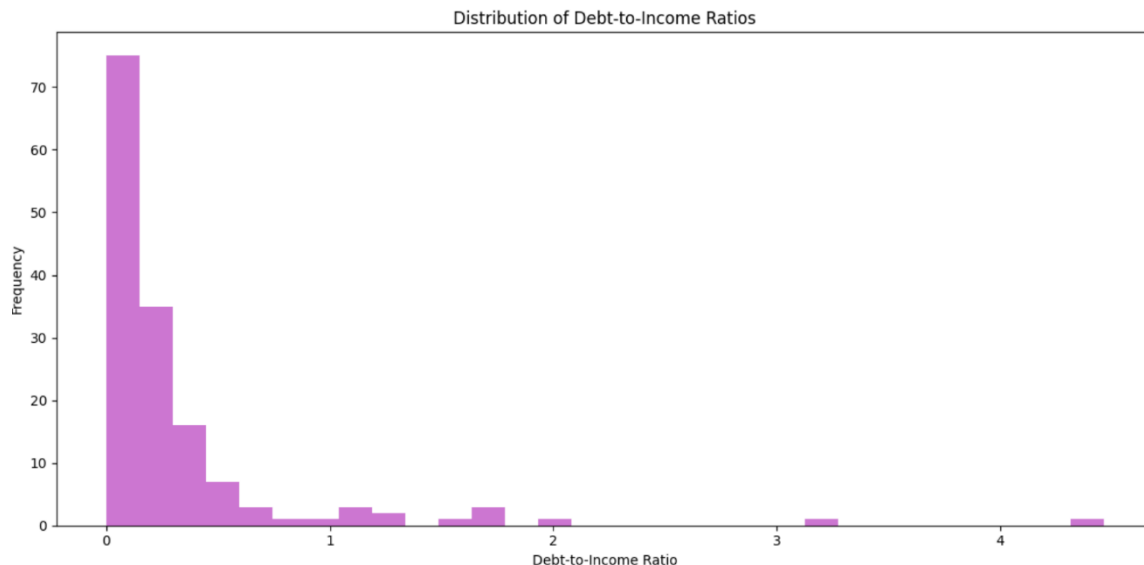
Visualization 2: Total Revenue vs Total Liabilities (First 50 companies)



Visualization 3: Distribution of Debt-to-Equity Ratios



Visualization 4: Distribution of Debt-to-income Ratio



C. Explanation of Customized Visualizations

Visualization 1: Average Profit Margin by Business State (Bar Chart)

This chart was created using the grouped data from `df_grouped`. I used the "Profit_Margin_Mean" column and plotted it as a bar chart. I customized the bar color to "skyblue" to make the chart more visually appealing and easier to read. A bar chart was a good choice because it makes it simple to compare the average profit margins across different states and spot where businesses are performing better or worse.

Visualization 2: Total Revenue vs Total Liabilities (Line Graph)

This graph was built using the `df_final` DataFrame. It plots the "Total Revenue" and "Total Liabilities" values for the first 50 companies. I customized the lines by using green for revenue and red for liabilities to make the comparison clearer. A line graph was the right choice because it shows trends and differences between the two financial measures over the companies, making it easy to spot patterns or gaps between revenue and liabilities.

Visualization 3: Distribution of Debt-to-Equity Ratios (Pie Chart)

This pie chart shows the breakdown of companies with negative versus positive debt-to-equity ratios, using counts from the `df_final` DataFrame. I customized the slice colors to "tomato" and

"lightgreen" to make the chart easier to read. I chose a pie chart because it quickly shows the proportion of companies in each category, giving a clear view of how many businesses are carrying more debt than equity.

Visualization 4: Distribution of Debt-to-Income Ratios (Histogram)

This histogram was made by plotting the "Debt to Income Ratio" column from the `df_final` DataFrame. I customized the bars by setting the color to "orchid" to make it visually stand out. I used a histogram because it shows the distribution of debt-to-income ratios across all companies, helping stakeholders understand how the ratios are spread out and whether most businesses have high or low debt compared to their income.

References

No outside sources were used in the creation of this submission. All information was based on WGU course materials and provided resources.