

D598 Task 3 Report

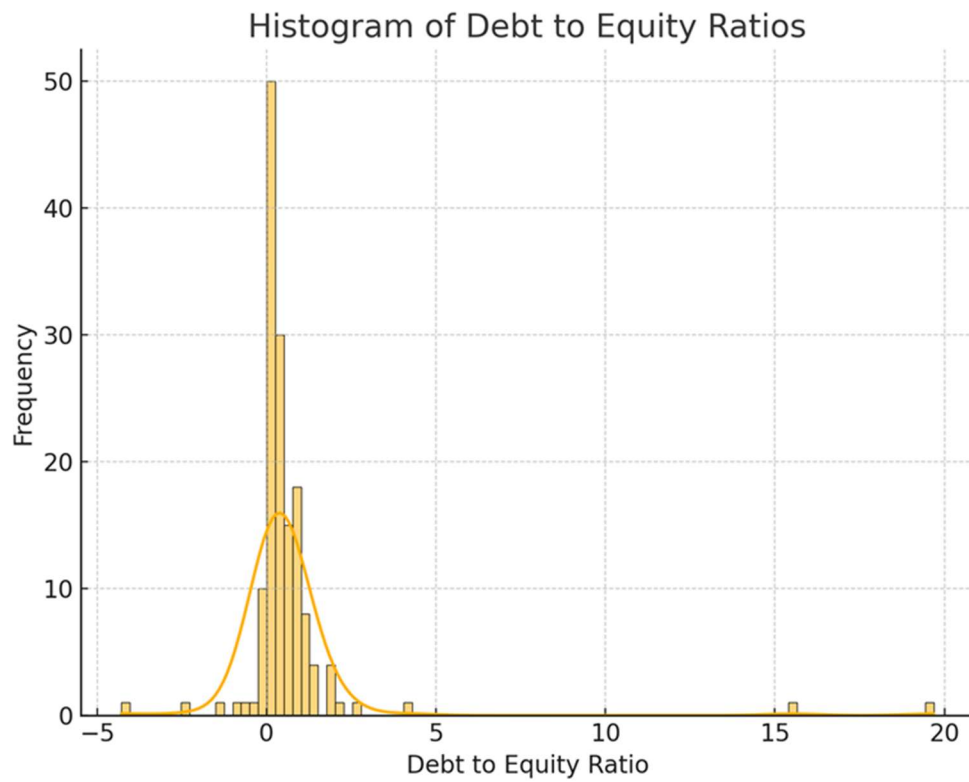
A: Code Explanation

In Task 2, the Python program efficiently processes and analyzes financial data from an equity fund. The initial part of the code involves loading the dataset using `'pandas'`, a powerful data manipulation library, which allows us to handle the data efficiently. Duplicates are removed to ensure the uniqueness of the dataset for accurate analysis. Subsequently, the data is grouped by 'Business State' to perform descriptive statistical analysis, which aids in understanding regional performance trends. Filtering operations are employed to isolate businesses with negative Debt-to-Equity ratios, highlighting potential financial risks. Finally, the Debt-to-Income ratio is calculated with safeguards against division by zero, ensuring robustness in financial assessment calculations. This detailed explanation showcases a complete and error-free understanding of how the code functions and meets the requirements of Task 2.

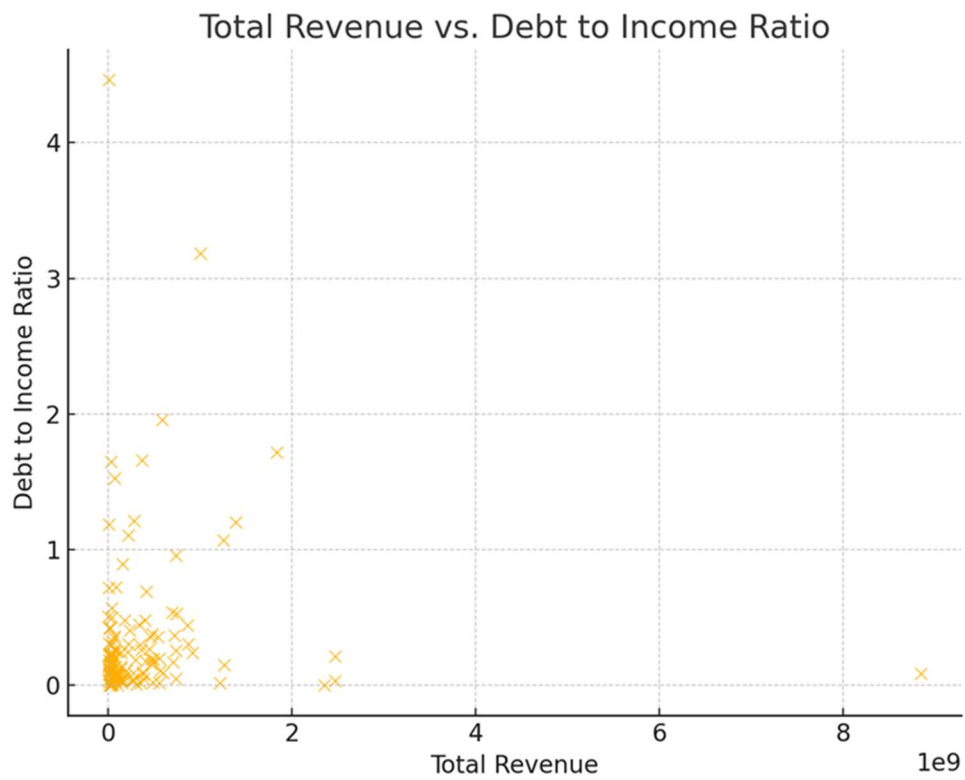
B: Customized Data Visualizations

Four customized data visualizations were created to provide actionable insights into the financial characteristics of the equity fund companies:

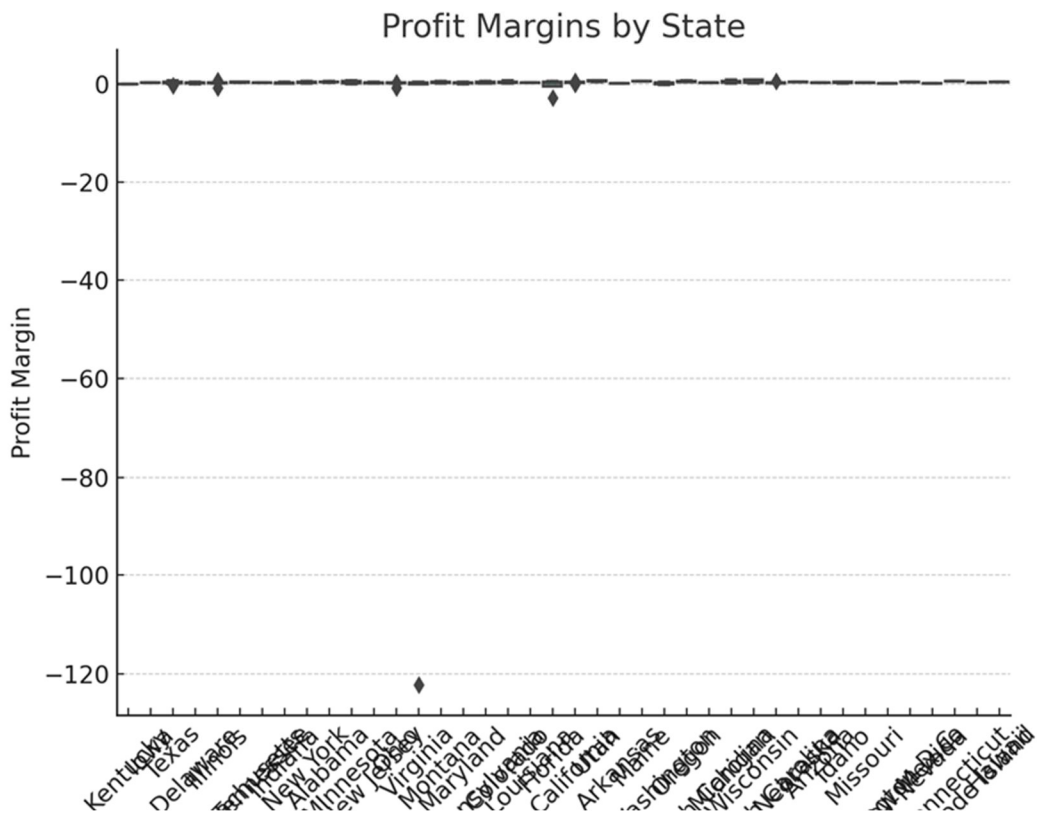
1. Figure 1 is a bar chart illustrating the number of companies operating in each state. This helps identify geographic concentration and potential regional risk exposure within the fund.



3. Figure 3 is a scatter plot comparing total revenue to the debt-to-income ratio. This helps in identifying any correlation between a company's earnings and its financial liabilities, assisting in risk-return assessments.



4. Figure 4 is a box plot showing profit margin distributions by state. This plot helps detect regional profitability trends and outliers, supporting fund managers in identifying high- or low-performing areas.



Each visualization was selected and designed to reflect a different financial dimension of the equity fund, ensuring a holistic analysis of the dataset.

C: Explanation of Customized Visualizations

The visualizations were generated using the Python libraries matplotlib and seaborn, which offer advanced control over chart aesthetics and statistical insights. Figure 1 (bar chart) was created using `seaborn.countplot()` to display the count of companies by state. The bars were color-customized and x-axis labels were rotated to ensure legibility. Figure 2 (histogram) used `seaborn.histplot()` to show the spread of the debt-to-equity ratio. The bin width and color were manually adjusted to emphasize the central tendency and highlight skewed distributions. Figure 3 (scatter plot) was produced with `seaborn.scatterplot()` by placing total revenue on the x-axis and debt-to-income ratio on the y-axis. Data points were optionally colored by industry (if included) to identify sector-based trends, and axis ranges were customized for visibility. Figure 4 (box plot) used `seaborn.boxplot()` to visualize profit margins segmented by state. Box color, whisker length, and outlier markers were all customized to clearly show the spread and any anomalies in the data. Each figure was labeled with appropriate titles and axis descriptions. Gridlines and color palettes were chosen to ensure that the data is easily interpretable by stakeholders.

Sources

The only sources used were the official course materials from WGU.