Reason for companies to go bankrupt:

1. Failure of Business Strategy:
   1. Lack of governance
      1. Failure to understand key metrics can lead to bankruptcy.
         1. **A1: Net Profit / Total Assets:** Measures overall profitability relative to assets.
         2. **A9: Net Profit / Sales:** Indicates how effectively the company converts sales into profit.
         3. **A76: Return on Equity:** Measures profitability relative to shareholders' equity.
   2. Market trends
      1. Failing to see what competitors are doing
      2. Market sentiment shifts and not adjusting
         1. **~~A21: Sales (n) / Sales (n-1):~~** ~~Measures the growth rate of sales compared to the previous period.~~
2. Inefficient use of assets
   1. Management fails to generate income from assets
      1. **A7: Operating Income to Total Asset Ratio:** Assesses the efficiency of generating operating income from total assets.
   2. Holding inventory too long
      1. **A47: (Inventory \* 365) / Cost of Products Sold**: Calculates the average number of days inventory is held before being sold.
   3. Negative Working Capital
      1. **A3: Working Capital to Total Assets Ratio**: Assesses a company’s ability to cover its short-term obligations with its current assets.
      2. **A77: working capital Indicator**: Measures the short-term financial health by calculating the difference between current assets and current liabilities. 1= positive value, 0 = negative
3. Too much Debt
   1. Debt used to pay for Working capital
      1. **A5: Defensive Interval Ratio**: Measures how long a company can operate with its existing liquid assets without additional cash inflows.
      2. **~~A71: Debt Coverage Ratio~~**~~: Measures a company's ability to service its debt with its earnings.~~
   2. Debt used to pay for equipment, plant
      1. **A10: Equity to Total Assets Ratio**: Indicates the proportion of assets financed by shareholders’ equity.
      2. **A2: Total Liabilities / Total Assets**: Shows the proportion of assets financed by liabilities.
      3. **A75: Financial Leverage:** Assesses the use of debt to finance the company's assets.
   3. Operating profits can’t cover interest
      1. **~~A70: Financial Expense to Operating Income:~~** ~~Measures the proportion of operating income used to cover financial expenses.~~
      2. **~~A68: Cash to Interest:~~** ~~Measures the ability to cover interest expenses with available cash.~~
      3. **A65: Interest to Sales:** Measures the burden of interest expenses relative to sales.
      4. **~~A74: Cash to Financial Expense:~~** ~~Measures the ability to cover financial expenses with cash on hand.~~
4. Not enough Cash flow
   1. Not enough cash on hand
      1. **A66: Cash to Assets**: Measures the proportion of assets held in cash.
   2. Not collecting from account receivables
      1. **A44: (Receivables \* 365) / Sales**: Highlights the average number of days it takes to collect receivables.
   3. Sales are not enough to cover expenses
      1. **A58: Total Costs / Total Sales**: Compares total costs to total sales, highlighting cost management efficiency.
   4. Cash conversation cycle is too long
      1. **A73: Cash Conversion Cycle**: Measures the time it takes for a company to convert its investments in inventory and other resources into cash flows from sales.
   5. Cash can’t cover interest
      1. **A68: Cash to Interest**: Measures the ability to cover interest expenses with available cash.
   6. Days payable is too high
      1. **A52: (Short-term Liabilities \* 365) / Cost of Products Sold**: Calculates the number of days it would take to pay off short-term liabilities with the cost of goods sold.

Initial analysis of the data determined their were quite a few null values. These values were investigated to understand why they were missing. In most of the case, there was no data available. The other cases were errors from dividing by zero. This research used imputation methods to feel in the missing dataset. The imputation methods that were reviewed were the “Mean/Median/Mode” imputation; KNN imputation and Random Forest Imputation. From the three methods, random forest imputation was selected as these feature are fairly complex. It was determined that for null values exceeding 10% of the data the feature was removed from the dataset.