Financial Health Data Report

This Data report summarizes the objectives, methodology, and key results of the financial data analysis project. The overall goal of the project is to develop a **data analysis and scoring system** to evaluate a company's financial health.

---

1. **Business Understanding**

Project Objectives and Scope

The project addresses the challenge that many financial professionals rely on intuition or outdated reports, which can lead to poor investment or lending decisions.

- **Main Objective:** To build a data analysis and scoring system that evaluates a company's financial health using real-world financial data.
- **Specific Goals:**
    1. Collect and preprocess financial data from the **Yahoo Finance API**.
    2. Analyze key financial metrics (e.g., revenue growth, debt-to-equity ratio).
    3. Build a **financial health scoring model**.
    4. Visualize financial insights for easier interpretation.
- **Success Criteria:** The system must produce **realistic health scores** based on financial fundamentals and output clear, explainable results for all users.

---

2. Data Understanding

A. Data Sources

- **Source:** Real financial datasets were fetched directly from the **Yahoo Finance API** using the yfinance library.
- **Ticker Universe:** The analysis focused on a final universe of **503 tickers**. This selection achieved **100.0% coverage** of the total market capitalization of the identified full universe of companies.
- **Data Types:** The extracted financial statements include Income Statements, Balance Sheets, and Cash Flow Statements.

3: Data Preparation

This phase focused on cleaning, standardizing, and merging the raw data into feature-ready tables.

- **Name Resolution:** A **fuzzy-matching mechanism** (resolve_item_names()) was implemented to consistently map desired financial line items to actual, often inconsistent, row names provided by Yahoo Finance.
- **Standardization:** All numeric values were **scaled (divided by 1e9)** to represent figures in **billions** for consistency.
- **Missing Data Handling:** All missing numeric values (NaN) in the final clean dataframes were replaced with **0** to ensure safe arithmetic during ratio calculations.
- **Integration:** Income, Balance Sheet, and Cash Flow master tables were concatenated and merged based on Ticker and Report Date.

4.Modeling

Financial Ratios

The analysis computed key financial ratios across four main categories (Profitability, Liquidity, Leverage, and Cash Flow):

- **Profitability Ratios:** Gross Margin, Operating Margin, Net Margin.
- **Liquidity Ratios:** Current Ratio, Quick Ratio (acid-test).
- **Leverage Ratios:** Debt-to-Equity Ratio, Total Debt/Total Assets.
- **Cash Flow Ratios:** Free Cash Flow (FCF), FCF Yield, and CapEx Ratio

Financial Risk Scoring (Altman Z-Score)

The core risk assessment was performed using the **Altman Z-Score**, which is designed to predict the probability of a company entering financial distress.

A. Z-Score Classification

The Z-Score results were categorized into three risk zones:

Classification	Z-Score Range	Interpretation
Safe	> 2.99	Low risk of financial distress
Grey	1.81 to 2.99	Moderate/Watch list risk

Classification	Z-Score Range	Interpretation
Distress	< 1.81	High probability of distress

B. Final Risk Distribution

Based on the analysis of **503 records**, the risk distribution across the universe is as follows:

Risk Category	Ticker Count
Distress	204
Grey	160
Safe	139