Disclaimer: This project is intended for educational and informational purposes only. The analysis, models, and methodologies used in this project are not intended to serve as financial or investment advice. The information presented should not be construed as a recommendation to buy, sell, or hold any securities or assets. Investing in the stock market carries inherent risks, and any decisions based on the content of this project are solely the responsibility of the individual. Always consult with a qualified financial advisor before making any investment decisions.

Table of Contents

Section 1: Stock Categories

Market Capitalization (Market Cap):

- Large-Cap: $10 billion or more (10,000+ million).
- Mid-Cap: $2 billion and $10 billion (2,000 million to 10,000 million).
- Small-Cap: $250 million and $2 billion (250 million to 2,000 million).

There are **11 primary sectors** recognized by the Global Industry Classification Standard (GICS):

- Communication Services
- Consumer Discretionary
- Consumer Staples
- Energy
- Financials
- Health Care
- Industrials
- Information Technology (IT)
- Materials
- Real Estate
- Utilities

Each sector includes multiple industries that group companies based on their products or services.

When comparing stocks, it's important to consider **market capitalization** and **industry** because these factors provide context for evaluating a company's performance relative to others in its sector.

Section 2: Stock Fundamentals

2.1. Price-to-Free Cash Flow (P/FCF)

Price-to-Free Cash Flow (P/FCF): A high P/FCF ratio indicates that the specific firm is trading at a high price but is not generating enough free cash flows to support the multiple (sometimes this is OK, depending on the firm, industry, and its specific operations). Smaller price ratios are generally preferred, as they may reveal a firm generating ample cash flows that are not yet properly considered in the current share price.

Undervalued stocks should have the following conditions:

- FCF must be positive, which means the company is making money.
- P/FCF that is lower than the industry's average P/FCF. Equivalently, the z-score of a stock's P/FCF is less than 0 (negative).

2.2. Return-on Equity & Price-to-Book

Return on Equity (ROE)	Price-to-Book (P/B)
This ratio is a gauge of a corporation's profitability and how efficiently it generates those profits. The higher the ROE, the better a company is at converting its equity financing into profits. A stock's ROE should be compared to its industry's ROE. "Good" ROE should be above the industry's average. ROE <10% should be considered "bad".	Book Value Per Share indicates a firm's net asset value (also known as NAV or total assets) minus total liabilities per share. When a stock is undervalued, it will have a higher BVPS than its stock price in the market. Price-to-book value (P/B) ratio measures the market's valuation of a company relative to its book value. Thus, P/B=1 means it is traded at exactly where it should be. Higher P/B means it might be overvalued. A stock's P/B is relative to a business and its industry. Typically under 1.0 is "good" but there are exceptions.

ROE & Price/Book should correlate.

- If a stock's ROE is high and P/B is low compared to its industry, it might be undervalued.
- If a stock's ROE is low and P/B is high compared to its industry, it might be overvalued.

		P/B	
		Low	**High**
ROE	**Low**	Fair	Overvalued
	High	Undervalued	Fair

Undervalued stocks should have the following conditions:

- Book value of a stock must be positive
- ROE must be larger than 10(%)
- P/B that is lower than the industry's average P/B. Equivalently, the z-score of a stock's P/B is less than 0 (negative).
- ROE that is higher than the industry's average ROE. Equivalently, the z-score of a stock's ROE is greater than 0 (positive)

2.3. Debt-to-equity (D/E) and Asset-to-Equity (A/E)

The D/E ratio measures how much debt (liability) a company has taken on relative to the value of its assets net of liabilities. Debt must be repaid or refinanced, imposes interest expense that typically can't be deferred, and could impair or destroy the value of equity in the event of a default. As a result, a high D/E ratio is often associated with high investment risk; it means that a company relies primarily on debt financing.

When using the D/E ratio, it is very important to consider the industry in which the company operates. Because different industries have different capital needs and growth rates, a D/E ratio value that's common in one industry might be a red flag in another. A D/E ratio below 1 would be seen as relatively safe, whereas values of 2 or higher might be considered risky. Companies in some industries, such as utilities, consumer staples, and banking, typically have relatively high D/E ratios.

Since $Assets = Liability + Equity$, we have

$$\frac{Assets}{Equity} = \frac{Liability + Equity}{Equity} = \frac{Liability}{Equity} + 1 \Leftrightarrow \frac{Liability}{Equity} = \frac{Assets}{Equity} - 1$$

$$\rightarrow \frac{D}{E} \leq 1 \Leftrightarrow \frac{Assets}{Equity} \leq 2$$

Undervalued stocks should have the following conditions:

- A/E must be greater than 1, which means D/E is greater than 0.
- A/E that is lower than the industry's average A/E. Equivalently, the z-score of a stock's A/E is less than 0 (negative)

2.4. Return-on-Asset (ROA)

Return on assets is a profitability ratio that shows how much profit a company generates from its assets. Return on assets (ROA) measures how effective a company's management is in generating profit from the total assets on its balance sheet.

- A ROA that rises over time indicates that the company is doing well at increasing its profits with each investment dollar it spends.
- A falling ROA indicates that the company might have over-invested in assets that have failed to produce revenue growth. This is a sign the company may be in some trouble.
- A negative return on assets may indicate that the company is unable to utilize its total assets sufficiently enough to generate a profit.

A ROA of over 5% is generally considered good. Over 20% is excellent.

ROAs should always be compared among firms in the same industry.

Undervalued stocks should have the following conditions:

- ROA must be positive, which means the company is investing its resources so that it can generate profit.
- ROA that is greater than the industry's average ROA. Equivalently, the z-score of a stock's ROA is greater than 0 (positive)

2.5. Price-to-Earnings (P/E)

Price-to-Earnings (P/E): A high P/E ratio could mean that a company's stock is overvalued or that investors expect high growth rates. A low P/E can indicate that a company is undervalued or that a firm is doing exceptionally well relative to its past performance. A good price-to-equity (P/E) ratio is one that is between 20 and 25. The lower the P/E ratio, the better. When analyzing P/E ratios, it's important to do so in the context of the industry the business operates.

The P/E ratio has several limitations. It doesn't account for future earnings growth, can be influenced by accounting practices, and may not be comparable across different industries. It also doesn't consider other financial aspects such as debt levels, cash flow, or the quality of earnings.

Undervalued stocks should have the following conditions:

- P/E should be between 1 and 25.
- P/E that is lower than the industry's average P/E. Equivalently, the z-score of a stock's P/E is less than 0.

Section 3: Evaluation Criteria

The evaluation is based on a series of criteria. Each stock is only evaluated relative to the industry average within its own market cap category.

Preliminary criteria: All undervalued stocks must meet these criteria.

- P/FCF > 0
- P/FCF < Industry average P/FCF
- P/B > 0
- P/B < Industry average P/B
- ROE > 10
- ROA > 5
- A/E > 1
- A/E < Industry average A/E

Additional criteria: These criteria are used to compare between stocks that meet the preliminary criteria. For each criterion met, a stock gets 1 point. The highest score a stock can get is 5 points.

- P/B < (0.7 x Industry's average P/B)
- ROE > Industry average ROE
- ROA > Industry average ROA
- P/E < Industry average P/E
- 1 < P/E < 25

Z-score ranking: This method is applied to each fundamental indicator to compare a stock's performance against the industry average. The further a z-score deviates from 0, the higher the stock's ranking.

- For **P/FCF**, **P/B**, **A/E**, and **P/E**, a **negative z-score** indicates the stock is undervalued. For **ROE** and **ROA**, a **positive z-score** is preferred, as it signifies better performance relative to the industry average.
- A z-score further from 0 is typically seen as a sign of stronger performance relative to the industry, making stocks with higher absolute z-scores more attractive.
- However, if a z-score is greater than 3 or less than -3, the data point is likely an outlier and should be treated with caution.

Section 4: Calculating Mean & Removing Outliers

The mean of each fundamental metric (P/FCF, P/B, ROE, ROA, A/E and P/E) is calculated for stocks within each industry and market cap category (Large Cap, Mid Cap, Small Cap). The means are calculated without outliers to ensure that the results reflect the central tendency of the data.

An outlier is a data point that differs significantly from other observations. An outlier may be due to a variability in the measurement, an indication of novel data, or it may be the result of experimental error. There are different methods to identify and handle outliers. For this project, we will first identify the outliers then recalculate the mean without these outliers.

Outliers are identified using the first and third quartiles of the data:

- First quartile (Q1) is defined as the value in a data set where 25% of the data points are below it
- Third quartile (Q3) is defined as the value in a data set where 75% of the data points are below it
- The Interquartile Range (IQR) is given by $Q3 - Q1$

An outlier in a data set is any data point that is less than the **lower bound** or greater than the **upper bound** defined below:

$$Lower\ bound = Q1 - 1.5 \cdot (IQR), \qquad Upper\ bound = Q3 + 1.5 \cdot (IQR)$$

Once outliers are identified for a fundamental metric, we recalculate the mean and standard deviation excluding these outliers. Since standard deviation measures the spread of the data, excluding outliers generally results in a smaller standard deviation compared to the one that includes outliers.

Next, we use this new mean and standard deviation to calculate the z-scores of the data points (including outliers) using the following formula:

$$z = \frac{(data\ point) - mean}{\binom{standard}{deviation}}$$

Note: the z-score for an outlier calculated using the new mean and standard deviation is likely less than negative three (-3) or greater than positive three (3).