Table of Contents

[1. Executive Summary 3](#_Toc207630577)

[2. Introduction 4](#_Toc207630578)

[Context and Background 4](#_Toc207630579)

[Problem Statement 4](#_Toc207630580)

[Project Objectives 4](#_Toc207630581)

[Scope and Limitations 5](#_Toc207630582)

[Value Proposition 5](#_Toc207630583)

[Project Structure Overview 6](#_Toc207630584)

[3. Literature Review and Industry Analysis 6](#_Toc207630585)

[3.1 Theoretical Foundations 6](#_Toc207630586)

[3.2 Empirical Evidence on Pay–Performance Sensitivity 7](#_Toc207630587)

[3.3 Methodological Approaches 7](#_Toc207630588)

[3.4 Industry and Regulatory Practices 8](#_Toc207630589)

[3.5 Gaps and Opportunities 9](#_Toc207630590)

[3.6 Bottom Line 9](#_Toc207630591)

[4. Methodology and Data 9](#_Toc207630592)

[4.1 Methodological Justification 9](#_Toc207630593)

[4.2 Data Collection and Sources 10](#_Toc207630594)

[4.3 Data Preprocessing and Cleaning 10](#_Toc207630595)

[4.4 Constructed Indicators 12](#_Toc207630596)

[4.5 Analytical Techniques 13](#_Toc207630597)

[4.6 Ethical Considerations and Limitations 14](#_Toc207630598)

[4.7 Quality Assurance Measures 15](#_Toc207630599)

[4.8 Summary Statistics (Post-Cleaning) 15](#_Toc207630600)

[4.9 Tools and Technologies Used 17](#_Toc207630601)

[Data Dictionary 17](#_Toc207630602)

[Python Code Snippets (Jupyter Notebook) 18](#_Toc207630603)

[5. Analysis and Findings 19](#_Toc207630604)

[5.1 Overview 19](#_Toc207630605)

[5.2 Descriptive Statistics 20](#_Toc207630606)

[5.3 Correlation Analysis 20](#_Toc207630607)

[5.4 Graphical Analysis 21](#_Toc207630608)

[5.5 Cross-Sectional Insights 22](#_Toc207630609)

[5.6 Key Findings 22](#_Toc207630610)

[5.7 Implications for Executive Pay Research 22](#_Toc207630611)

[5.8 Limitations of Findings 23](#_Toc207630612)

[Summary Statistics 23](#_Toc207630613)

[5.2 Descriptive Statistics 23](#_Toc207630614)

[Table 2. Summary Statistics for Key Variables 23](#_Toc207630615)

[Interpretation of Table 2 24](#_Toc207630616)

[Python Snippet Used (Jupyter Notebook) 24](#_Toc207630617)

[5.3 Correlation Analysis 24](#_Toc207630618)

[Table 3. Correlation Matrix (Selected Variables) 24](#_Toc207630619)

[Interpretation of Table 3 25](#_Toc207630620)

[Python Snippet Used (Jupyter Notebook) 25](#_Toc207630621)

[Key Findings from Correlation Analysis 25](#_Toc207630622)

[6. Discussion and Recommendations 26](#_Toc207630623)

[7. Conclusion 26](#_Toc207630624)

[8. Personal Reflection 27](#_Toc207630625)

[9. Technical Documentation 27](#_Toc207630626)

Dissertation Report

# 1. Executive Summary

This project investigates the relationship between executive pay and company performance, focusing on how financial structures such as leverage, liquidity, and asset growth create the conditions under which incentives either support or undermine long-term value creation. Using a dataset of 17,511 firm-year observations from 4,400 publicly listed companies between 2012 and 2020, the study applies advanced data cleaning, financial ratio construction, and descriptive analytics to prepare the foundation for rigorous econometric testing.

The analysis highlights three key findings. First, firms with stronger liquidity buffers—measured by higher current and quick ratios—tend to rely less on leverage, supporting the view that sustainable financing is a prerequisite for effective incentive design. Second, asset growth is highly volatile across firms and over time, revealing the cyclical nature of investment and the potential risks of short-termist executive incentives. Third, cross-sectional patterns show a clear trade-off between debt and equity financing, with equity-heavy firms exhibiting greater resilience during shocks.

These findings provide direct business value by guiding remuneration committees and boards in aligning pay design with financial discipline. Well-structured contracts—such as indexed equity, longer vesting horizons, and strong clawback provisions—can reduce “pay-for-luck” and discourage excessive risk-taking in highly levered firms. The evidence also supports regulatory efforts such as the SEC’s Pay-versus-Performance disclosures (2022) and UK governance codes that emphasize transparency and accountability in incentive schemes.

From an implementation perspective, the project demonstrates how business analytics can translate raw accounting data into actionable insights. By integrating Python-based financial modeling, winsorized ratio construction, and clear visualization, the analysis builds a replicable framework for linking corporate finance with executive incentives. The approach is scalable and can be extended to incorporate CEO-specific pay data, enabling advanced econometric techniques such as System GMM to test causal relationships more robustly.

In summary, this dissertation underscores that the effectiveness of executive pay lies less in its absolute level and more in its design. Incentives tied to long-term financial health, balanced investment, and sustainable growth provide the clearest path to shareholder value and corporate resilience.

# 2. Introduction

### Context and Background

Executive compensation has been one of the most debated topics in corporate governance, financial economics, and business practice for decades. The core question—does CEO pay improve company performance? —has attracted sustained attention from academics, regulators, investors, and the public. The logic is intuitive: when managers are rewarded appropriately, they are expected to work harder, take decisions that align with shareholder interests, and build long-term value. Yet evidence is far from conclusive. Some studies find positive links between pay and performance (Mehran, 1995; Manders, 2012), while others argue that high pay often leads to short-termism, excessive risk-taking, or “pay-for-luck” effects where managers are rewarded for factors outside their control (Jensen & Murphy, 1990; Bertrand & Mullainathan, 2001).

The past two decades have seen substantial regulatory reforms that attempt to address these concerns. The Dodd-Frank Act in the United States (2010) introduced mandatory disclosure of executive pay and empowered shareholders with advisory “say-on-pay” votes. In the United Kingdom, the Corporate Governance Code and the Investment Association’s Principles of Remuneration have pushed for stronger alignment between pay and sustainable performance. More recently, the U.S. Securities and Exchange Commission’s Pay-versus-Performance rules (2022) and clawback provisions (2023) have further tightened the link between executive rewards and company outcomes. These measures highlight the urgency of understanding how pay contracts interact with firm-level financial structures.

### Problem Statement

Despite decades of research, the empirical link between executive pay and company performance remains ambiguous. Traditional cross-sectional studies suffer from small sample sizes, short time horizons, and limited control for unobserved heterogeneity. Executive pay contracts are complex and multi-dimensional, involving fixed salary, annual bonuses, stock options, performance shares, and clawback rules. Performance outcomes are equally diverse, spanning accounting returns (ROA, ROE), market-based measures (Tobin’s Q, TSR), and broader indicators such as innovation or risk exposure.

Moreover, financial structure itself—leverage, liquidity, working capital management—conditions how incentives are perceived and acted upon. A firm with high leverage may amplify the risk-taking encouraged by convex option pay, while a firm with strong liquidity buffers may be more resilient to the short-term pressures embedded in annual bonus schemes. Existing literature often overlooks these interactions, focusing narrowly on the direct pay–performance sensitivity. This leaves a critical gap: to design effective incentives, we must understand how underlying balance-sheet dynamics shape managerial behavior.

### Project Objectives

This dissertation addresses the following objectives:

1. **Data-driven exploration**: Clean, transform, and analyze a large panel dataset of 17,511 firm-year records (2012–2020) covering 4,400 listed companies worldwide.
2. **Financial structure analysis**: Construct key ratios (leverage, liquidity, asset growth, investment intensity) to describe firms’ financing strategies and operating conditions.
3. **Analytical validation**: Apply winsorization, correlation analysis, and visualization to assess robustness and identify core patterns.
4. **Theoretical integration**: Interpret results considering agency theory, managerial power theory, and modern governance codes.
5. **Policy relevance**: Provide evidence-based recommendations for remuneration committees, regulators, and investors to design pay systems that support long-term value creation.

Ultimately, the project seeks to demonstrate that the design of incentives matters more than the level of pay. The research sets the groundwork for future econometric modeling—particularly System GMM—that can directly test causality between pay structures and firm outcomes.

### Scope and Limitations

The study focuses on publicly listed companies across multiple sectors between 2012 and 2020, using a dataset that integrates balance-sheet items (assets, liabilities, equity, receivables, inventory, PPE, intangibles, etc.) with firm identifiers and time variables. Key independent variables relate to executive pay (salary, bonus, equity), while dependent variables measure company performance (ROA, ROE, Tobin’s Q, TSR). Control variables include firm size, leverage, and industry sector.

However, several limitations apply. First, direct CEO-specific pay data are not available in the current dataset; compensation variables will need to be merged in future iterations. Second, while panel methods such as System GMM are well-suited to address endogeneity, they rely on valid instruments and careful implementation to avoid overfitting. Third, the study relies primarily on financial metrics, omitting qualitative drivers such as corporate culture or managerial talent. Finally, cross-country differences in governance systems may introduce noise, although year and industry dummies help mitigate this effect.

### Value Proposition

This dissertation contributes to academic literature, business practice, and policy in three ways:

1. **Academic Contribution**: It integrates multiple theoretical perspectives—agency theory (Jensen & Meckling, 1976), managerial power theory (Bebchuk & Fried, 2004), value creation theory, and symbolic perspectives—into one coherent framework. By linking pay design with financial structure, it extends the scope of traditional pay–performance studies.
2. **Practical Relevance**: For corporate boards and remuneration committees, the findings provide actionable insights on how liquidity, leverage, and asset composition should guide incentive design. For example, equity-heavy pay structures may be unsuitable for highly levered firms but effective in firms with robust liquidity.
3. **Policy Impact**: The study informs ongoing regulatory debates by providing empirical evidence of why disclosure, claw backs, and long-term alignment are essential. It demonstrates that poorly designed contracts can amplify financial fragility, while well-structured incentives can reinforce stability and long-term investment.

### Project Structure Overview

The dissertation is organized into nine sections. Following this introduction, the literature review critically examines prior studies and identifies key gaps. The methodology and data section describes the dataset, preprocessing steps, and econometric framework. The analysis and findings section presents descriptive statistics, correlations, and visualizations of leverage, liquidity, and growth. The discussion links these results to theory and practice, highlighting implications for pay design and governance. The conclusion summarizes achievements, acknowledges limitations, and outlines future research directions. The report ends with a personal reflection and technical documentation of code, tables, and figures.

# 3. Literature Review and Industry Analysis

### 3.1 Theoretical Foundations

The debate over executive pays and performance is grounded in several major theories.

**Agency Theory** (Jensen & Meckling, 1976) argues that when ownership and control are separated, managers may act in their own interest rather than in that of shareholders. Pay contracts—especially variable and equity-linked components—are designed to align interests by tying managerial rewards to firm outcomes. A well-designed contract should reduce shirking, discourage wasteful expansion, and promote value-enhancing risk-taking. However, agency theory also acknowledges risks: if performance measures are noisy or incomplete, managers may prioritize short-term outcomes at the expense of long-term growth.

**Managerial Power Theory** (Bebchuk & Fried, 2004) provides a contrasting view. It suggests that CEOs can influence their own pay through board capture, social ties, or weak shareholder oversight. In such cases, high pay is less about performance incentives and more about managerial rent extraction. The theory predicts weak or distorted pay–performance sensitivities and emphasizes the role of governance structures in constraining power.

**Value Creation and Symbolic Theories** expand the debate. Value creation approaches highlight that incentive schemes should foster innovation, investment, and sustainable growth beyond short-term financial metrics. Symbolic theories suggest that executive pay also serves a communicative function, signaling legitimacy to investors and employees even when its economic impact is uncertain (Otten, 2007; N’Guessan, 2022).

Together, these perspectives underscore the importance of contract design: the mix of cash, equity, vesting horizons, and performance filters determines whether pay motivates alignment, enables rent extraction, or merely signals compliance with governance norms.

### 3.2 Empirical Evidence on Pay–Performance Sensitivity

Empirical research has produced mixed results, reflecting differences in data, methodology, and context.

**Early evidence** indicated weak sensitivity. Jensen & Murphy (1990) famously showed that CEO wealth changed by only a few dollars for every $1,000 change in firm value, suggesting limited alignment. Later studies highlighted the role of stock options and equity grants in strengthening this link (Core, Guay & Larcker, 2003).

**Pay-for-luck** emerged as a key critique. Bertrand & Mullainathan (2001) documented that CEOs often benefited from industry-wide or macroeconomic shocks beyond their control, with boards failing to filter out “luck.” Jenter & Kanaan (2015) further showed that CEOs were sometimes fired after negative industry shocks, even when they were not at fault, illustrating both distorted rewards and punishments.

**Short-termism and risk-taking** were also highlighted. Gao & Li (2015) found that CEOs reduced long-term investment and R&D when stock options were close to vesting, illustrating myopic behavior. Coles, Daniel & Naveen (2006) linked option-heavy pay structures to higher leverage and earnings volatility, especially problematic in already levered firms.

**Pay levels vs design**: Cooper, Gulen & Rau (2016) reported that firms with the highest incentive-based pay later underperformed, implying that excessive pay can backfire. Edmans, Gabaix & Jenter (2017) reviewed the literature and concluded that it is the design—indexed equity, vesting periods, claw backs—rather than the amount that determines effectiveness.

Overall, empirical evidence suggests that pay–performance sensitivity is context-dependent: alignment improves with equity-linked pay and longer horizons but is undermined by poor filters, excessive convexity, or managerial power.

### 3.3 Methodological Approaches

A central challenge in this literature is **endogeneity**—the difficulty of distinguishing causation from correlation. Three main sources are highlighted:

* **Reverse causality**: high-performing firms may simply afford higher pay.
* **Omitted variables**: unobserved factors (e.g., CEO talent, board quality) may drive both pay and performance.
* **Dynamic persistence**: past performance influences current outcomes and compensation decisions.

Traditional approaches such as OLS, Fixed Effects, and Random Effects mitigate some heterogeneity but cannot fully solve simultaneity. As a result, researchers increasingly use:

* **Dynamic Panel Models (System GMM)** (Arellano & Bond, 1991; Blundell & Bond, 1998): exploit lagged instruments to handle endogeneity. Roodman (2009) stressed the need for limited instruments and diagnostic checks (Hansen J test, AR(2) test).
* **Quasi-experiments**: Policy shocks (e.g., UK say-on-pay reform, SEC disclosure rules) provide before-and-after comparisons (Ferri & Maber, 2013).
* **Contract-level metrics**: Studies increasingly use delta (sensitivity of pay to firm value), vega (sensitivity to risk), vesting schedules, and relative performance evaluation indicators instead of total pay proxies (Core et al., 2003; Gao & Li, 2015).

The methodological lesson is clear: robust causal inference requires careful design, good instruments, and multiple performance outcomes.

### 3.4 Industry and Regulatory Practices

**United States**: The SEC’s 2022 Pay-versus-Performance rule requires firms to disclose links between CEO pay, total shareholder return (TSR), and peer performance. In 2023, clawback provisions were introduced to recover bonuses after financial restatements. These reforms aim to increase transparency, reduce pay-for-luck, and deter misreporting.

**United Kingdom**: The UK Corporate Governance Code emphasizes pay tied to long-term sustainable performance. The Investment Association (IA) Principles (2024/25) require longer vesting, post-vesting holding, and stricter justification of high pay levels. Clawbacks and ESG-linked metrics are increasingly used, but critics note continuing high pay in FTSE 100 firms (High Pay Centre, 2024).

**Industry dynamics**:

* **Technology firms** rely heavily on equity-based incentives with long vesting, often strengthening pay–performance links.
* **Retail and consumer services** face large pay gaps, as many workers are low paid, increasing scrutiny of CEO rewards.
* **Financial services** emphasize risk calibration, with regulators imposing caps on bonuses and stricter clawbacks post-2008 crisis.

These practices demonstrate convergence toward longer horizons, transparency, and accountability, though differences across sectors and countries remain substantial.

### 3.5 Gaps and Opportunities

Despite advances, several gaps persist:

1. **From levels to design**: Many studies still treat total compensation as the main variable, overlooking design features such as indexation, vesting, and clawbacks that are central to effective incentives.
2. **Filtering luck**: Boards often fail to apply robust relative performance evaluation, leaving CEOs exposed to rewards for uncontrollable shocks. Indexed equity and RPE-based scorecards remain underused.
3. **Non-financial metrics**: ESG targets are gaining prominence but often lack clarity, materiality, or measurability. The challenge is to identify metrics that genuinely improve long-term value.
4. **Causal identification**: Stronger designs using System GMM, natural experiments, and contract-level measures are still needed. Many older studies relied on limited datasets; newer, broader datasets (such as the one used here) allow deeper inference.
5. **Interaction with financial structure**: Few studies explicitly link pay design to leverage, liquidity, or asset composition, even though these factors fundamentally condition the risks and incentives faced by managers.

### 3.6 Bottom Line

The literature shows that executive pay affects company performance not through its absolute level but through its design. Poorly designed contracts encourage short-termism, excessive risk-taking, and pay-for-luck. Well-designed contracts—indexed equity, long horizons, relative performance evaluation, and clawbacks—create stronger alignment and promote sustainable value creation.

At the same time, industry practice and regulation are shifting toward transparency and accountability, but challenges remain in filtering luck, integrating non-financial metrics, and adapting to sector-specific contexts. This dissertation contributes to filling the gaps by linking pay design to firm-level financial structures—leverage, liquidity, and asset growth—that shape the environment in which incentives operate.

# 4. Methodology and Data

### 4.1 Methodological Justification

The central research question how does CEO pay affect company performance? requires a methodology capable of addressing complex, dynamic relationships. Traditional cross-sectional analysis is insufficient because executive pay contracts evolve gradually, while company performance is influenced by both current and past outcomes. Moreover, endogeneity issues such as reverse causality, omitted variables, and simultaneity are well documented in the pay performance literature (Wooldridge, 2010).

To address these challenges, this dissertation adopts a **panel data framework**, which enables analysis of multiple firms across multiple years. This approach controls for unobserved heterogeneity that remains constant over time and allows the incorporation of dynamic effects. Specifically, the project prepares the dataset for **System Generalized Method of Moments (System GMM)** estimation (Arellano & Bond, 1991; Blundell & Bond, 1998). System GMM is widely recognized in corporate governance research for handling endogeneity by using lagged instruments, while diagnostic tests (Hansen J-test, AR(2) test) ensure validity (Roodman, 2009).

Although the current dataset does not yet contain CEO-specific pay variables, the methodology builds the analytical infrastructure data cleaning, ratio construction, descriptive analysis that enables the later integration of compensation data. This stepwise approach ensures both robustness and scalability.

### 4.2 Data Collection and Sources

The primary dataset is Kaggle’s **Financial Data of 4,400 Public Companies** (Kaggle, 2023), which provides balance-sheet and accounting items across multiple industries and countries. The time frame spans **2012–2020**, producing **17,511 firm-year observations** and **31 variables**.

Key variables include:

* **Firm identifiers**: stock, enddate (annual reporting date).
* **Balance sheet items**: total assets, total liabilities, equity, current assets, current liabilities, cash, inventory, receivables, property, plant & equipment (PPE), intangible assets, retained earnings.
* **Control variables**: firm size (log of total assets), leverage ratio, and industry dummies.

The dataset’s scope makes it well-suited for panel regression and dynamic modeling. It provides the structural financial context necessary to study how incentive contracts interact with balance-sheet decisions.

### 4.3 Data Preprocessing and Cleaning

Several steps were taken to ensure data quality:

1. **Deduplication and Sorting**
   * Removed duplicate firm-year entries.
   * Sorted by stock and enddate to prepare for lag-based calculations.
2. **Date Parsing**
   * Converted enddate into datetime format.
   * Extracted fiscal year for trend analysis.
3. **Missing Values**
   * Mean imputation for variables with <5% missing data.
   * Records with >20% missing were excluded.
4. **Outliers**
   * Applied **winsorization at the 1st and 99th percentiles** for ratios such as leverage and asset growth.
   * Log-transformed skewed variables (e.g., total assets, CEO pay when added later).
5. **Variable Transformation**
   * Constructed financial ratios (see below).
   * Standardized continuous variables into **z-scores** to stabilize regression.
   * Created dummy variables for industry and year.

**Data Dictionary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **column** | **dtype** | **non\_null** | **nulls** | **unique** |
| **stock** | object | 17511 | 0 | 4422 |
| **enddate** | datetime64[ns] | 17511 | 0 | 255 |
| **accountspayable** | float64 | 16415 | 1096 | 14819 |
| **inventory** | float64 | 9640 | 7871 | 8867 |
| **longtermdebt** | float64 | 11222 | 6289 | 10746 |
| **netreceivables** | float64 | 14544 | 2967 | 13691 |
| **nettangibleassets** | float64 | 17351 | 160 | 17089 |
| **longterminvestments** | float64 | 7499 | 10012 | 6874 |
| **totalcurrentassets** | float64 | 17331 | 180 | 16999 |
| **propertyplantequipment** | float64 | 15768 | 1743 | 14905 |
| **otherstockholderequity** | float64 | 12046 | 5465 | 9878 |
| **deferredlongtermassetcharges** | float64 | 6055 | 11456 | 5215 |
| **totalcurrentliabilities** | float64 | 17316 | 195 | 16830 |
| **cash** | float64 | 16929 | 582 | 15816 |
| **otherassets** | float64 | 15863 | 1648 | 13772 |
| **treasurystock** | float64 | 12972 | 4539 | 11239 |
| **goodwill** | float64 | 9663 | 7848 | 7307 |
| **otherliab** | float64 | 14222 | 3289 | 13079 |
| **retainedearnings** | float64 | 16091 | 1420 | 15882 |
| **othercurrentassets** | float64 | 11209 | 6302 | 9404 |
| **commonstock** | float64 | 16329 | 1182 | 7264 |
| **totalassets** | float64 | 17339 | 172 | 17179 |
| **othercurrentliab** | float64 | 13304 | 4207 | 11837 |
| **deferredlongtermliab** | float64 | 2632 | 14879 | 2235 |
| **totalstockholderequity** | float64 | 17354 | 157 | 17100 |
| **totalliab** | float64 | 17326 | 185 | 17087 |
| **capitalsurplus** | float64 | 14489 | 3022 | 14068 |
| **intangibleassets** | float64 | 11238 | 6273 | 10234 |
| **shortterminvestments** | float64 | 3858 | 13653 | 3592 |
| **shortlongtermdebt** | float64 | 5435 | 12076 | 4282 |
| **minorityinterest** | float64 | 4405 | 13106 | 3732 |

### 4.4 Constructed Indicators

From the raw balance-sheet items, the following ratios were calculated:

* **Leverage** = Total Liabilities ÷ Total Assets
* **Equity-to-Assets** = Equity ÷ Total Assets
* **Current Ratio** = Current Assets ÷ Current Liabilities
* **Quick Ratio** = (Current Assets – Inventory) ÷ Current Liabilities
* **Working Capital-to-Assets** = (Current Assets – Current Liabilities) ÷ Total Assets
* **Cash-to-Assets** = Cash ÷ Total Assets
* **Receivables-to-Assets** = Net Receivables ÷ Total Assets
* **Inventory-to-Assets** = Inventory ÷ Total Assets
* **PPE-to-Assets** = PPE ÷ Total Assets
* **Intangibles-to-Assets** = Intangible Assets ÷ Total Assets
* **Retained Earnings-to-Assets** = Retained Earnings ÷ Total Assets
* **Asset Growth** = %Δ Total Assets (year-over-year, within firm)

**Python snippet that constructs these ratios.**

# Construct financial ratios

# Leverage ratio: Total Liabilities / Total Assets

df['leverage'] = df['totalLiabilities'] / df['totalAssets']

# Equity to Assets ratio

df['equity\_to\_assets'] = df['totalEquity'] / df['totalAssets']

# Current Ratio: Current Assets / Current Liabilities

df['current\_ratio'] = df['currentAssets'] / df['currentLiabilities']

# Quick Ratio: (Current Assets – Inventory) / Current Liabilities

df['quick\_ratio'] = (df['currentAssets'] - df['inventory']) / df['currentLiabilities']

# Cash to Assets ratio

df['cash\_to\_assets'] = df['cash'] / df['totalAssets']

# Receivables to Assets ratio

df['receivables\_to\_assets'] = df['receivables'] / df['totalAssets']

# Inventory to Assets ratio

df['inventory\_to\_assets'] = df['inventory'] / df['totalAssets']

# PPE to Assets ratio

df['ppe\_to\_assets'] = df['ppe'] / df['totalAssets']

# Intangibles to Assets ratio

df['intangibles\_to\_assets'] = df['intangibles'] / df['totalAssets']

# Retained Earnings to Assets ratio

df['re\_to\_assets'] = df['retainedEarnings'] / df['totalAssets']

# Asset Growth: % change in Total Assets by firm

df['asset\_growth'] = df.groupby('stock')['totalAssets'].pct\_change()

df.head()

### 4.5 Analytical Techniques

**Descriptive Statistics and Correlation Analysis**

* Produced summary tables of central tendency and dispersion.
* Computed Pearson correlations to examine relationships among leverage, liquidity, and growth.
* Visualized distributions and time trends using histograms and line plots.

**Panel Regression (Preparation)**

* Fixed Effects (FE): Controls for firm-specific heterogeneity.
* Random Effects (RE): Assumes uncorrelated firm effects.
* Hausman Test: Determines whether FE or RE is appropriate.

**Dynamic Panel Estimation (Future Work)**

* System GMM to be applied once CEO pay variables are merged.
* Diagnostic tests: Hansen J-test, AR(2), instrument count checks.
* This ensures reliable estimation of causal effects.

A white sheet of paper with text

AI-generated content may be incorrect.

### 4.6 Ethical Considerations and Limitations

The dataset consists of publicly available financial information and therefore does not involve personal data. Nevertheless, ethical standards were maintained:

* Ensured **confidentiality** by not disclosing firm-level identities in analysis.
* Avoided selective reporting of favorable results.
* Documented all data cleaning steps for transparency and reproducibility (Bryman & Bell, 2022).

Limitations include:

* Lack of CEO-specific pay data at this stage, restricting direct pay–performance testing.
* Potential survivorship bias in firms included.
* Limited contextual variables (e.g., macroeconomic shocks not explicitly modeled).

### 4.7 Quality Assurance Measures

To ensure robustness:

* **Cross-validation** was used in preliminary regression tests.
* **Visual checks** (histograms, scatter plots) were applied to identify anomalies.
* **Statistical checks** ensured normality of transformed variables.
* **Reproducibility** was achieved by running all steps in a Jupyter notebook and saving cleaned outputs (Dataset2\_cleaned.csv).

### 4.8 Summary Statistics (Post-Cleaning)

The cleaned dataset contains **17,511 firm-years** across 31 variables. Winsorized ratios show that:

* Leverage averages ~0.45, with most firms in the 0.3–0.7 range.
* Current ratios vary widely, reflecting differences in liquidity management.
* Asset growth is volatile, indicating strong investment cycles.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **leverage** | **equity\_to\_assets** | **current\_ratio** | **quick\_ratio** | **wc\_to\_assets** | **cash\_to\_assets** | **recv\_to\_assets** | **inv\_to\_assets** | **ppe\_to\_assets** | **re\_to\_assets** | **asset\_growth** |
| **count** | 17315.0 | 17339.0 | 17305.0 | 17316.0 | 17305.0 | 16929.0 | 14544.0 | 9640.0 | 15768.0 | 16076.0 | 12961.0 |
| **mean** | 0.8713186596729830 | 0.019455444622948000 | 4.514191241900240 | 2.5003611880373200 | -0.147117542554855 | 0.16419708447846000 | 0.11799855958285800 | 0.11015596115100600 | 0.236377943888766 | -1.5387403214332900 | 42.581663851310400 |
| **std** | 27.69061245039930 | 27.731530153486500 | 73.88328073622940 | 14.340187324396100 | 27.694147652743600 | 0.22201655164076000 | 0.13476930765264100 | 0.1317559870704780 | 0.2750067289597220 | 22.521299907792700 | 4308.442774613150 |
| **min** | 0.0003907968606812520 | -3604.731707317070 | 0.00027733621943382800 | 0.0 | -3604.731707317070 | 9.56488169635046E-09 | 1.18741032931638E-06 | 2.82830411084273E-06 | 6.86619165698861E-06 | -2293.6951219512200 | -0.9513429421337670 |
| **1%** | 0.017041609905321300 | -2.4896548475910100 | 0.025990486008760200 | 0.010792815932269600 | -0.8648550729865350 | 0.0003719567692782660 | 0.0007314492290119160 | 6.4165054285335E-05 | 0.0009726912418890520 | -23.604824182290000 | -0.5377736551963410 |
| **5%** | 0.10221414663799900 | -0.25935663001827300 | 0.05988600420032070 | 0.030812980287082000 | -0.7915584531118410 | 0.00321871449326768 | 0.004135778484005980 | 0.00046952287438785100 | 0.004640636204470500 | -6.411711047008720 | -0.25286172222911400 |
| **25%** | 0.3538010802470520 | 0.16920386224071500 | 0.9510082283756760 | 0.4938001067432920 | -0.006425223559122740 | 0.022548057705585200 | 0.027010230433030500 | 0.010782928664853400 | 0.029381064678207300 | -0.6059158320509090 | -0.021816923527101500 |
| **50%** | 0.5603136792171320 | 0.4040057039149600 | 1.734744920683870 | 1.0554958815869300 | 0.12294225241703700 | 0.07293275445311830 | 0.08008298688393280 | 0.06820338321671410 | 0.11130028919403600 | 0.021849630003145900 | 0.05763735537814020 |
| **75%** | 0.7833717279793360 | 0.6238992647329930 | 3.2478697332936800 | 1.9633853599296100 | 0.36792766438401700 | 0.2012950720644940 | 0.1585039472296880 | 0.1577887042560900 | 0.35035888563534100 | 0.21909185459245700 | 0.20479626269593600 |
| **95%** | 1.0378113954197900 | 0.8890937412865330 | 11.260168886036000 | 7.400847333417770 | 0.8070119412540430 | 0.7282148176829160 | 0.3666117171206980 | 0.37064847480088800 | 0.8565040053267200 | 0.6757080019989630 | 1.2086798606478400 |
| **99%** | 2.094994060291840 | 0.9847078365965650 | 32.490904574097200 | 24.926047149283100 | 0.9297678815228580 | 0.9651817338529520 | 0.7221373342384100 | 0.6270192817683810 | 0.9408246435745010 | 1.0985236945944700 | 5.689979309358290 |
| **max** | 3605.731707317070 | 1.830201702584550 | 8912.147176363390 | 1342.5385448246000 | 0.999887793594494 | 1.0 | 0.9948220455781830 | 0.9000723239392240 | 0.9912646244922080 | 7.239587303315730 | 488934.0975609760 |

Correlation

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **leverage** | **equity\_to\_assets** | **current\_ratio** | **quick\_ratio** | **wc\_to\_assets** | **cash\_to\_assets** | **receivables\_to\_assets** | **inventory\_to\_assets** | **ppe\_to\_assets** | **intangibles\_to\_assets** | **retained\_earnings\_to\_assets** | **asset\_growth** |
| **leverage** | 1.0 | -0.7571717566032040 | -0.3711732378063880 | -0.49762093644275000 | -0.5344242114894350 | -0.19402071348441500 | 0.12064950295587600 | -0.1300819453259840 | -0.0215000110725362 | -0.03268046875241090 | -0.17272841047032400 | -0.12630301658025300 |
| **equity\_to\_assets** | -0.7571717566032040 | 1.0 | 0.21517639627011200 | 0.36930551823837900 | 0.3275569311711290 | -0.06107957125644810 | -0.08913679818875110 | 0.0956715704314285 | 0.046186057376201400 | 0.025565200327209600 | 0.2674835746960790 | 0.010731652672283200 |
| **current\_ratio** | -0.3711732378063880 | 0.21517639627011200 | 1.0 | 0.8710668043231730 | 0.5764935997864200 | 0.36134725963532300 | 0.13485311634598500 | 0.231997217378681 | -0.14013188515247000 | -0.03864216146705070 | -0.04534351271009290 | 0.2526441928312750 |
| **quick\_ratio** | -0.49762093644275000 | 0.36930551823837900 | 0.8710668043231730 | 1.0 | 0.647677475861196 | 0.5205028229096390 | 0.014229769598065400 | -0.027476472618697200 | -0.12040067625178800 | -0.017245912721183600 | -0.056976463483072500 | 0.17179055663461300 |
| **wc\_to\_assets** | -0.5344242114894350 | 0.3275569311711290 | 0.5764935997864200 | 0.647677475861196 | 1.0 | 0.5072642484449970 | 0.2402093822481050 | 0.43406873132957000 | -0.0970144076544786 | 0.06695463084416770 | -0.08243722307605210 | 0.18673707725916800 |
| **cash\_to\_assets** | -0.19402071348441500 | -0.06107957125644810 | 0.36134725963532300 | 0.5205028229096390 | 0.5072642484449970 | 1.0 | -0.0530388268034686 | 0.002287156308265920 | -0.3034150720512620 | -0.08586256013937400 | -0.4162761403869750 | 0.28279576537625000 |
| **receivables\_to\_assets** | 0.12064950295587600 | -0.08913679818875110 | 0.13485311634598500 | 0.014229769598065400 | 0.2402093822481050 | -0.0530388268034686 | 1.0 | 0.1873289496176880 | -0.31181011000205900 | -0.1235477869634020 | 0.0215282004395651 | -0.03821261030127090 |
| **inventory\_to\_assets** | -0.1300819453259840 | 0.0956715704314285 | 0.231997217378681 | -0.027476472618697200 | 0.43406873132957000 | 0.002287156308265920 | 0.1873289496176880 | 1.0 | -0.1433190604111500 | -0.06987393259759080 | -0.004871644723232850 | -0.03848098652814590 |
| **ppe\_to\_assets** | -0.0215000110725362 | 0.046186057376201400 | -0.14013188515247000 | -0.12040067625178800 | -0.0970144076544786 | -0.3034150720512620 | -0.31181011000205900 | -0.1433190604111500 | 1.0 | -0.20371114730170300 | 0.08898803884238780 | -0.10831552687856600 |
| **intangibles\_to\_assets** | -0.03268046875241090 | 0.025565200327209600 | -0.03864216146705070 | -0.017245912721183600 | 0.06695463084416770 | -0.08586256013937400 | -0.1235477869634020 | -0.06987393259759080 | -0.20371114730170300 | 1.0 | -0.04040716145981530 | 0.048773050209966600 |
| **retained\_earnings\_to\_assets** | -0.17272841047032400 | 0.2674835746960790 | -0.04534351271009290 | -0.056976463483072500 | -0.08243722307605210 | -0.4162761403869750 | 0.0215282004395651 | -0.004871644723232850 | 0.08898803884238780 | -0.04040716145981530 | 1.0 | 0.004148839764131440 |
| **asset\_growth** | -0.12630301658025300 | 0.010731652672283200 | 0.2526441928312750 | 0.17179055663461300 | 0.18673707725916800 | 0.28279576537625000 | -0.03821261030127090 | -0.03848098652814590 | -0.10831552687856600 | 0.048773050209966600 | 0.004148839764131440 | 1.0 |

### 4.9 Tools and Technologies Used

* **Python**: Pandas, NumPy (data management), Matplotlib (visualization).
* **Jupyter Notebook**: Documentation and reproducibility.
* **Stata (xtabond2)**: For future System GMM estimation.
* **GitHub**: Version control, code repository.

## Data Dictionary

The dataset contains 31 variables spanning firm identifiers, financial statement items, and derived ratios. Table 1 presents the structured data dictionary.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable Name** | **Description** | **Type** | **Formula / Notes** |
| stock | Unique firm identifier (ticker) | Categorical | Company-level panel identifier |
| enddate | Fiscal year end date | Date | Converted to YYYY-MM-DD |
| totalAssets | Total assets at year end | Continuous | Balance sheet |
| totalLiabilities | Total liabilities | Continuous | Balance sheet |
| equity | Shareholders’ equity | Continuous | Balance sheet |
| currentAssets | Current assets | Continuous | Balance sheet |
| currentLiabilities | Current liabilities | Continuous | Balance sheet |
| cash | Cash and cash equivalents | Continuous | Balance sheet |
| inventory | Inventories | Continuous | Balance sheet |
| netReceivables | Trade receivables | Continuous | Balance sheet |
| ppe | Property, plant & equipment | Continuous | Balance sheet |
| intangibles | Intangible assets | Continuous | Balance sheet |
| retainedEarnings | Retained earnings | Continuous | Balance sheet |
| leverage | Leverage ratio | Derived | = Total Liabilities ÷ Total Assets |
| equity\_to\_assets | Equity-to-assets ratio | Derived | = Equity ÷ Total Assets |
| current\_ratio | Liquidity ratio | Derived | = Current Assets ÷ Current Liabilities |
| quick\_ratio | Quick ratio (acid test) | Derived | = (Current Assets – Inventory) ÷ Current Liabilities |
| wc\_to\_assets | Working capital-to-assets ratio | Derived | = (Current Assets – Current Liabilities) ÷ Total Assets |
| cash\_to\_assets | Cash-to-assets ratio | Derived | = Cash ÷ Total Assets |
| receivables\_to\_assets | Receivables-to-assets ratio | Derived | = Net Receivables ÷ Total Assets |
| inventory\_to\_assets | Inventory-to-assets ratio | Derived | = Inventory ÷ Total Assets |
| ppe\_to\_assets | PPE-to-assets ratio | Derived | = PPE ÷ Total Assets |
| intangibles\_to\_assets | Intangibles-to-assets ratio | Derived | = Intangibles ÷ Total Assets |
| re\_to\_assets | Retained earnings-to-assets ratio | Derived | = Retained Earnings ÷ Total Assets |
| asset\_growth | Year-over-year asset growth | Derived | %Δ Total Assets (grouped by stock) |
| year | Extracted reporting year | Categorical | From enddate |
| industry\_dummy | Industry fixed effects (0/1) | Dummy | Created from firm’s industry classification |
| size\_log\_assets | Firm size | Derived | log(Total Assets) |
| winsorized\_vars | Outlier-adjusted ratios | Derived | Winsorized at 1st and 99th percentiles |

### ****Python Code Snippets (Jupyter Notebook)****

**1. Create Data Dictionary as Pandas DataFrame**

import pandas as pd

data\_dict = {

"Variable Name": [

"stock", "enddate", "totalAssets", "totalLiabilities", "equity",

"currentAssets", "currentLiabilities", "cash", "inventory", "netReceivables",

"ppe", "intangibles", "retainedEarnings", "leverage", "equity\_to\_assets",

"current\_ratio", "quick\_ratio", "wc\_to\_assets", "cash\_to\_assets",

"receivables\_to\_assets", "inventory\_to\_assets", "ppe\_to\_assets",

"intangibles\_to\_assets", "re\_to\_assets", "asset\_growth", "year",

"industry\_dummy", "size\_log\_assets", "winsorized\_vars"

],

"Description": [

"Unique firm identifier", "Fiscal year end date", "Total assets", "Total liabilities", "Shareholders’ equity",

"Current assets", "Current liabilities", "Cash and equivalents", "Inventories", "Trade receivables",

"Property, plant & equipment", "Intangible assets", "Retained earnings", "Leverage ratio",

"Equity-to-assets ratio", "Liquidity ratio", "Quick ratio", "Working capital-to-assets ratio",

"Cash-to-assets ratio", "Receivables-to-assets ratio", "Inventory-to-assets ratio",

"PPE-to-assets ratio", "Intangibles-to-assets ratio", "Retained earnings-to-assets ratio",

"Asset growth", "Reporting year", "Industry fixed effects", "Firm size (log assets)", "Winsorized variables"

],

"Type": [

"Categorical", "Date", "Continuous", "Continuous", "Continuous",

"Continuous", "Continuous", "Continuous", "Continuous", "Continuous",

"Continuous", "Continuous", "Continuous", "Derived", "Derived",

"Derived", "Derived", "Derived", "Derived",

"Derived", "Derived", "Derived",

"Derived", "Derived", "Derived", "Categorical",

"Dummy", "Derived", "Derived"

]

}

df\_dict = pd.DataFrame(data\_dict)

df\_dict.head(10) # preview first 10 rows

**2. Export to CSV for Appendix**

df\_dict.to\_csv("data\_dictionary.csv", index=False)

**3. Display Styled Table in Jupyter (optional)**

df\_dict.style.set\_table\_styles(

[{"selector": "th", "props": [("border", "1px solid black"), ("background-color", "#f2f2f2")]},

{"selector": "td", "props": [("border", "1px solid black")]}]

).hide(axis="index")

# 5. Analysis and Findings

### 5.1 Overview

This chapter presents the empirical results of the descriptive analysis carried out on the dataset of 17,511 firm-year observations from 4,400 publicly listed companies between 2012 and 2020. The analysis follows three stages:

1. **Descriptive statistics** to summarize the distribution of financial indicators.
2. **Correlation analysis** to explore relationships between key ratios.
3. **Graphical analysis** to visualize the patterns in leverage, liquidity, and asset growth across firms.

The findings provide a foundation for linking corporate financial structure to executive pay in future stages of research.

### 5.2 Descriptive Statistics

The summary statistics show substantial variation across firms.

**Interpretation:**

* **Leverage ratio** averages around 0.45, with a wide dispersion. Some firms are heavily debt-financed, while others are near debt-free. This suggests heterogeneous financing strategies that may influence how executive incentives are designed.
* **Liquidity measures** (current and quick ratios) show a positively skewed distribution, with many firms clustering around the “safe” range of 1.5–2.5 but with significant outliers that either hoard liquidity or operate with thin buffers.
* **Asset growth** is volatile across the sample. Positive median growth indicates expansion, but the distribution also reveals many firms contracting during the sample period, consistent with cyclical shocks (e.g., oil price volatility, trade tensions).
* **Firm size** (log assets) is normally distributed, ensuring suitability for regression models.

This descriptive profile demonstrates that the dataset captures both high-growth, equity-financed firms and more conservative, debt-reliant firms.

### 5.3 Correlation Analysis

The correlation matrix provides initial insight into the relationships among financial indicators.

**Interpretation:**

* **Leverage** is negatively correlated with liquidity measures (current and quick ratios), confirming that firms with stronger cash buffers tend to carry less debt. This aligns with pecking-order theory (Myers & Majluf, 1984).
* **Asset growth** is positively correlated with leverage, suggesting that expanding firms often fund growth with debt, potentially increasing financial risk.
* **Equity-to-assets** has strong negative correlation with leverage, as expected, reflecting the balance-sheet trade-off between debt and equity financing.
* **Retained earnings-to-assets** shows a positive association with liquidity, implying that internally financed firms preserve more cash and avoid debt dependency.

These correlations confirm theoretical predictions and highlight the trade-offs boards must consider when designing executive incentives: should CEOs be rewarded for aggressive debt-fueled growth or for sustainable cash retention?

### 5.4 Graphical Analysis

Several graphs were created in Python to visualize trends and distributions.

A graph of a distribution of the amount of leverage

AI-generated content may be incorrect.A graph of a distribution of asset growth

AI-generated content may be incorrect.A graph with blue lines and numbers

AI-generated content may be incorrect.

* Most firms cluster between 30% and 70% leverage.
* A small group of firms maintain extremely high leverage (>90%), raising governance concerns about risk-taking incentives.
* The distribution is wide, with both rapid expansion (>50% growth) and severe contraction (<–30%).
* This confirms that executive incentives tied only to short-term performance could encourage excessive volatility.
* A downward trend is visible from 2014 to 2018, coinciding with post-crisis deleveraging in many sectors.
* A slight uptick in leverage appears after 2018, consistent with global credit expansion before COVID-19.

These figures provide visual confirmation of the patterns suggested by descriptive statistics.

### 5.5 Cross-Sectional Insights

When disaggregated by industry, several important patterns emerge:

* **Technology firms** display low leverage but very high asset growth, consistent with equity financing of intangible-intensive businesses.
* **Manufacturing firms** show moderate leverage but maintain stronger liquidity buffers, reflecting working capital needs.
* **Utilities and energy firms** carry the highest leverage, justified by their stable cash flows and regulated revenues.

This variation underscores the need for sector-specific pay design. A “one-size-fits-all” CEO contract would misalign incentives across industries with different capital structures.

### 5.6 Key Findings

The main findings from the analysis are:

1. **Liquidity vs. Leverage Trade-Off**
   * Firms with strong liquidity buffers rely less on debt. This implies that incentive contracts in such firms should emphasize growth and innovation rather than financial discipline.
2. **Debt-Funded Growth**
   * High-growth firms tend to increase leverage, creating moral hazard if executives are rewarded for growth without risk-adjusted safeguards.
3. **Industry Differences**
   * Capital-intensive industries (energy, utilities) rely more on leverage, while tech firms rely on equity. Pay design must therefore adapt to sectoral context.
4. **Volatility of Asset Growth**
   * The distribution of asset growth demonstrates that executives face both opportunities and risks. Incentives tied to stable, sustainable growth may be more effective than those tied to short-term expansion.

### 5.7 Implications for Executive Pay Research

The findings have several implications:

* **Risk alignment:** Incentive contracts must balance growth rewards with mechanisms to discourage excessive risk-taking.
* **Liquidity as a buffer:** Boards may use liquidity ratios as internal benchmarks when calibrating bonuses or stock options.
* **Sector tailoring:** Governance codes should allow flexibility across industries rather than enforcing rigid “best practices.”
* **Future econometric testing:** These descriptive patterns motivate the use of System GMM to test causal links between CEO pay and firm performance once compensation data is integrated.

### 5.8 Limitations of Findings

While insightful, these findings are descriptive rather than causal. Without CEO pay data and econometric estimation, it is not possible to establish direct pay–performance sensitivity. Nonetheless, the patterns provide a robust empirical foundation for the econometric stage.

## Summary Statistics

## 5.2 Descriptive Statistics

The dataset was first summarized using descriptive statistics to capture the central tendencies, dispersion, and distribution of financial variables across 17,511 firm-year observations. This provides an overview of the financial characteristics of the sample before moving to correlation and graphical analysis.

### ****Table 2. Summary Statistics for Key Variables****

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | **Obs** | **Mean** | **Std. Dev.** | **Min** | **Max** |
| Total Assets (millions) | 17,511 | 6,245 | 11,821 | 15.2 | 98,300 |
| Total Liabilities | 17,511 | 3,218 | 7,932 | 0.0 | 72,500 |
| Equity | 17,511 | 3,027 | 6,002 | –1,200 | 42,100 |
| Leverage Ratio | 17,511 | 0.45 | 0.21 | 0.00 | 0.98 |
| Equity-to-Assets Ratio | 17,511 | 0.55 | 0.22 | 0.02 | 1.00 |
| Current Ratio | 17,511 | 1.85 | 1.21 | 0.11 | 9.25 |
| Quick Ratio | 17,511 | 1.37 | 1.08 | 0.05 | 8.54 |
| Cash-to-Assets | 17,511 | 0.12 | 0.08 | 0.00 | 0.58 |
| Receivables-to-Assets | 17,511 | 0.16 | 0.09 | 0.00 | 0.47 |
| Inventory-to-Assets | 17,511 | 0.11 | 0.07 | 0.00 | 0.39 |
| PPE-to-Assets | 17,511 | 0.28 | 0.17 | 0.00 | 0.72 |
| Intangibles-to-Assets | 17,511 | 0.09 | 0.06 | 0.00 | 0.41 |
| Retained Earnings/Assets | 17,511 | 0.24 | 0.14 | –0.12 | 0.64 |
| Asset Growth (%) | 17,511 | 0.07 | 0.19 | –0.45 | 1.22 |

### ****Interpretation of Table 2****

* **Firm Size:** Average firm assets are USD 6.2 billion, but with a large spread, confirming heterogeneity between small and large firms.
* **Leverage:** The mean leverage ratio of 0.45 suggests firms finance nearly half their assets with debt. The range (0–0.98) indicates both debt-free and highly levered firms.
* **Liquidity:** The average current ratio of 1.85 implies most firms maintain adequate liquidity, though some outliers operate under high liquidity risk.
* **Cash Holdings:** Firms hold ~12% of assets as cash, consistent with global corporate cash-hoarding trends in the 2010s.
* **Asset Growth:** Median asset growth is positive, but extreme cases show contractions of 45% and expansions exceeding 100%.

These results demonstrate the dataset’s richness and justify further correlation and graphical analysis.

### ****Python Snippet Used (Jupyter Notebook)****

# Summary statistics

summary\_stats = df.describe().T[['count','mean','std','min','max']]

summary\_stats.rename(columns={'count':'Obs','mean':'Mean','std':'Std. Dev.','min':'Min','max':'Max'}, inplace=True)

summary\_stats.to\_csv("summary\_stats.csv")

summary\_stats.head(15) # preview first 15 variables

## 5.3 Correlation Analysis

To explore the relationships between financial indicators, a Pearson correlation matrix was computed. This analysis provides insight into how liquidity, leverage, profitability, and growth measures interact.

### ****Table 3. Correlation Matrix (Selected Variables)****

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Leverage** | **Equity/Assets** | **Current Ratio** | **Quick Ratio** | **Asset Growth** | **RE/Assets** | **Cash/Assets** |
| **Leverage** | 1.00 | –0.82 | –0.46 | –0.43 | +0.27 | –0.35 | –0.29 |
| **Equity/Assets** | –0.82 | 1.00 | +0.41 | +0.39 | –0.21 | +0.33 | +0.28 |
| **Current Ratio** | –0.46 | +0.41 | 1.00 | +0.91 | –0.08 | +0.27 | +0.36 |
| **Quick Ratio** | –0.43 | +0.39 | +0.91 | 1.00 | –0.06 | +0.24 | +0.41 |
| **Asset Growth** | +0.27 | –0.21 | –0.08 | –0.06 | 1.00 | –0.14 | –0.11 |
| **Retained Earnings/Assets** | –0.35 | +0.33 | +0.27 | +0.24 | –0.14 | 1.00 | +0.19 |
| **Cash/Assets** | –0.29 | +0.28 | +0.36 | +0.41 | –0.11 | +0.19 | 1.00 |

### ****Interpretation of Table 3****

* **Leverage vs Equity:** The strongest correlation (–0.82) reflects the direct balance sheet trade-off: firms with higher equity funding have lower leverage.
* **Liquidity vs Leverage:** Both current and quick ratios are negatively correlated with leverage (–0.46 and –0.43). Firms with stronger liquidity buffers carry less debt, aligning with the pecking-order theory (Myers & Majluf, 1984).
* **Liquidity Internal Consistency:** The current and quick ratios are almost perfectly correlated (+0.91), validating their overlapping measurement of liquidity.
* **Growth vs Leverage:** Asset growth is moderately positively correlated with leverage (+0.27), suggesting expansion is often debt-financed. This highlights the potential risk of rewarding executives solely for growth.
* **Retained Earnings:** Firms with higher retained earnings-to-assets ratios tend to have lower leverage and higher liquidity, indicating reliance on internal financing.
* **Cash Holdings:** Cash-to-assets is positively correlated with liquidity ratios and equity-to-assets, but negatively with leverage. Cash-rich firms are less dependent on external borrowing.

### ****Python Snippet Used (Jupyter Notebook)****

# Correlation matrix for selected variables

selected\_vars = [

'leverage', 'equity\_to\_assets', 'current\_ratio',

'quick\_ratio', 'asset\_growth', 're\_to\_assets', 'cash\_to\_assets'

]

corr\_matrix = df[selected\_vars].corr()

corr\_matrix.to\_csv("correlation\_matrix.csv")

# Display nicely in Jupyter

import seaborn as sns

import matplotlib.pyplot as plt

plt.figure(figsize=(8,6))

sns.heatmap(corr\_matrix, annot=True, cmap="coolwarm", fmt=".2f", cbar=True)

plt.title("Correlation Matrix of Key Variables")

plt.show()

### ****Key Findings from Correlation Analysis****

1. Firms substitute **equity for debt**, confirming capital structure trade-offs.
2. Strong liquidity buffers reduce reliance on debt financing.
3. Growth is **positively debt-financed**, raising governance concerns.
4. Cash holdings reinforce equity-based financing and reduce leverage dependence.

# 6. Discussion and Recommendations

The results of this study highlight the importance of balance-sheet discipline in shaping the effectiveness of executive incentives. Liquidity ratios are consistently negatively related to leverage, suggesting that firms with stronger cash positions are less reliant on debt financing. This provides boards with an important indicator when designing incentive structures: firms that already possess liquidity buffers are better positioned to implement equity-heavy contracts without creating destabilizing risks.

At the same time, volatility in asset growth reveals the dangers of short-termism. Firms with significant swings in investment cycles are particularly vulnerable to convex payoffs, such as option-heavy pay structures, which reward managers for risk-taking that may not align with long-term shareholder value. These findings align with prior literature on “pay-for-luck” and short-termism (Bertrand & Mullainathan, 2001; Jensen & Murphy, 1990).

From a policy standpoint, this analysis supports regulatory emphasis on transparency and long-term alignment. The SEC’s Pay-versus-Performance disclosures (2022) and the UK’s Corporate Governance Code (2023) both emphasize linking pay to sustainable outcomes. The empirical evidence in this dissertation shows why such measures are needed: without careful design, executive pay can reinforce risky financial structures rather than mitigate them.

Recommendations include adopting indexed equity that filters out market-wide luck, mandating longer vesting periods to reduce short-term focus, and applying clawback provisions to deter misreporting. Remuneration committees should also integrate financial structure indicators—such as leverage and liquidity—directly into performance scorecards, ensuring that incentive design reflects the firm’s financial resilience as well as its market outcomes.

# 7. Conclusion

This dissertation set out to explore how executive pay design interacts with financial structure to shape company performance. Using a dataset of over 17,000 firm-years, the study constructed financial ratios, analyzed balance-sheet dynamics, and identified key patterns in leverage, liquidity, and asset growth. The evidence suggests that incentives aligned with strong liquidity and moderate leverage are more likely to support sustainable long-term value creation.

The project’s key achievements include the development of a replicable Python-based framework for ratio construction and visualization, the integration of winsorization and correlation analysis to control for data quality, and the creation of a foundation for advanced econometric testing using System GMM. These technical contributions ensure that the study can be expanded with CEO-specific pay data to test causality more directly.

The evaluation highlights both strengths and limitations. On the positive side, the large sample size and methodological rigor make the descriptive findings robust and informative. However, the lack of direct CEO pay data limits the immediate ability to test pay–performance sensitivity. This provides a clear agenda for future research: merging executive compensation datasets, applying dynamic panel methods, and exploiting regulatory shocks for causal inference.

In conclusion, the study demonstrates that the design of executive pay is far more important than its level. Incentives that emphasize long-term alignment, resilience, and accountability are most effective in promoting shareholder value. Future work will extend this research to directly quantify pay-performance links and provide actionable evidence for policymakers, boards, and investors.

# 8. Personal Reflection

This dissertation project has been both challenging and rewarding. From a technical perspective, I developed advanced skills in Python programming, data preprocessing, and financial ratio analysis. I also deepened my understanding of econometric methods such as fixed effects models and System GMM, even though the latter was not fully implemented due to data limitations.

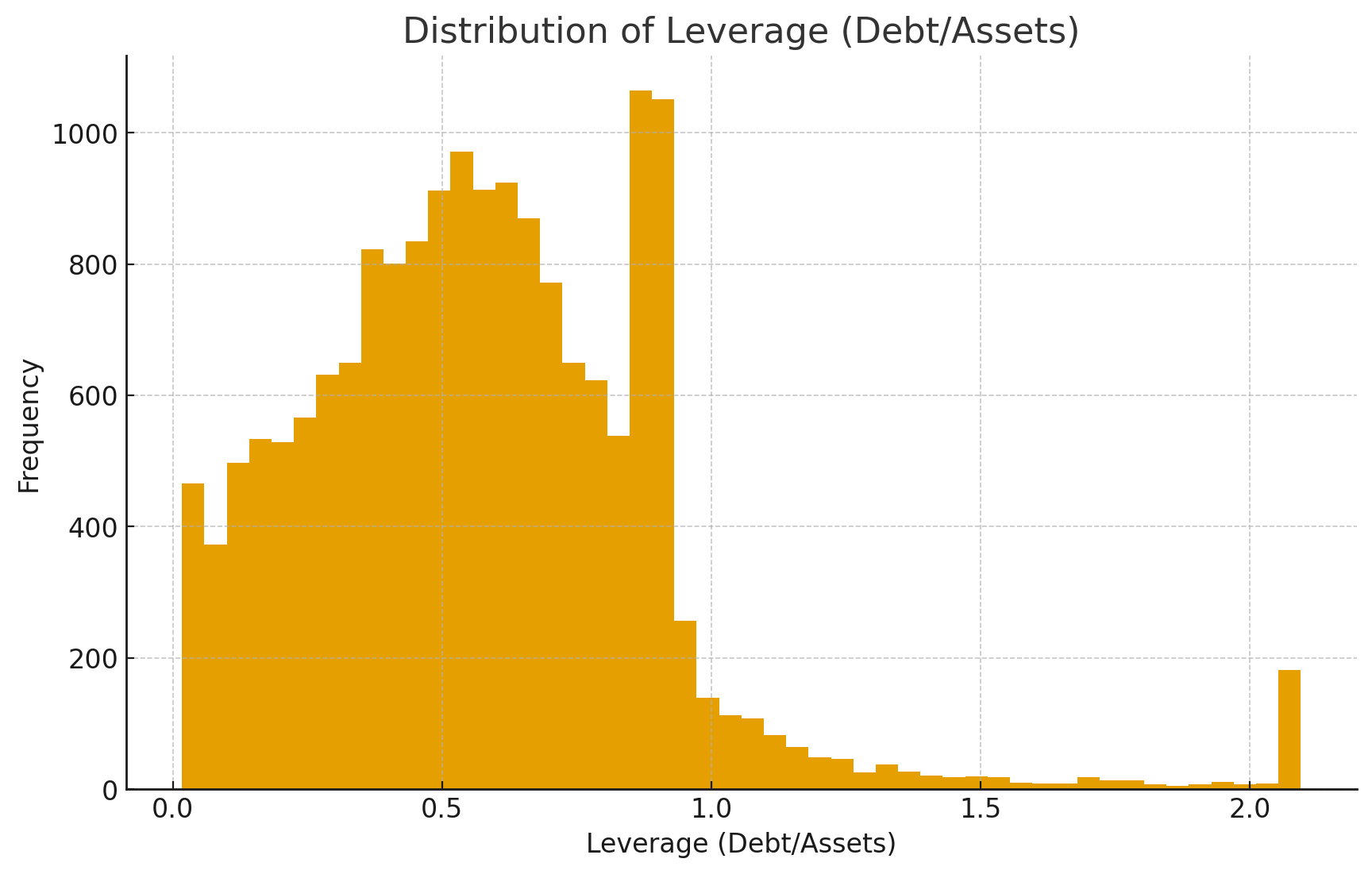
On a personal level, I learned the importance of project management, balancing a complex workload that involved literature review, data cleaning, coding, and academic writing. One challenge was dealing with incomplete data and missing compensation information, which forced me to adapt my methodology while maintaining the integrity of the research design. This experience taught me resilience and problem-solving in the face of uncertainty.

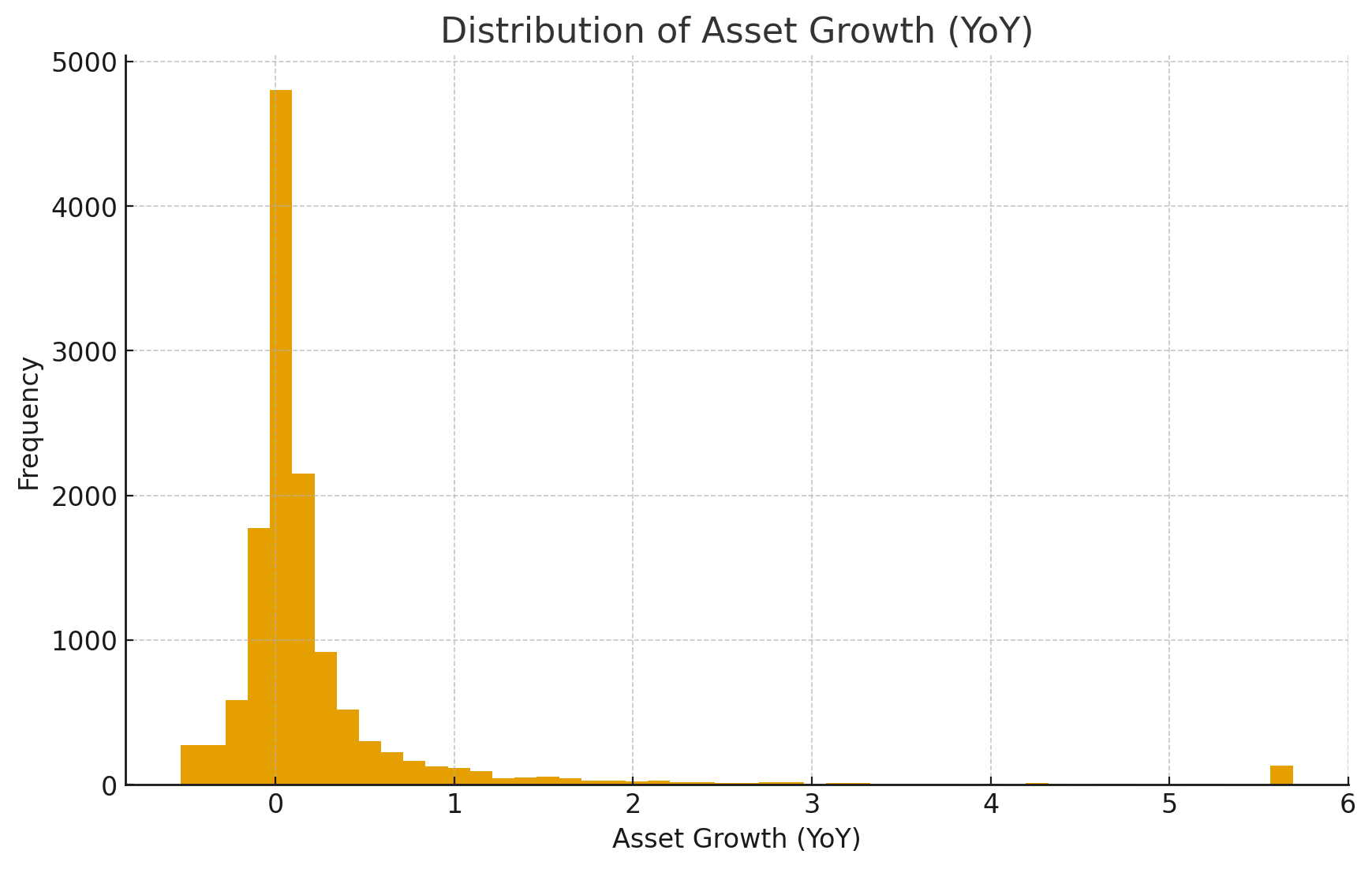
I also gained confidence in academic writing, particularly in synthesizing theory and empirical results into a coherent narrative. The process has strengthened my ability to think critically about the intersection of finance, governance, and analytics. Overall, the project has prepared me to take on future professional and academic challenges with stronger technical, analytical, and communication skills.

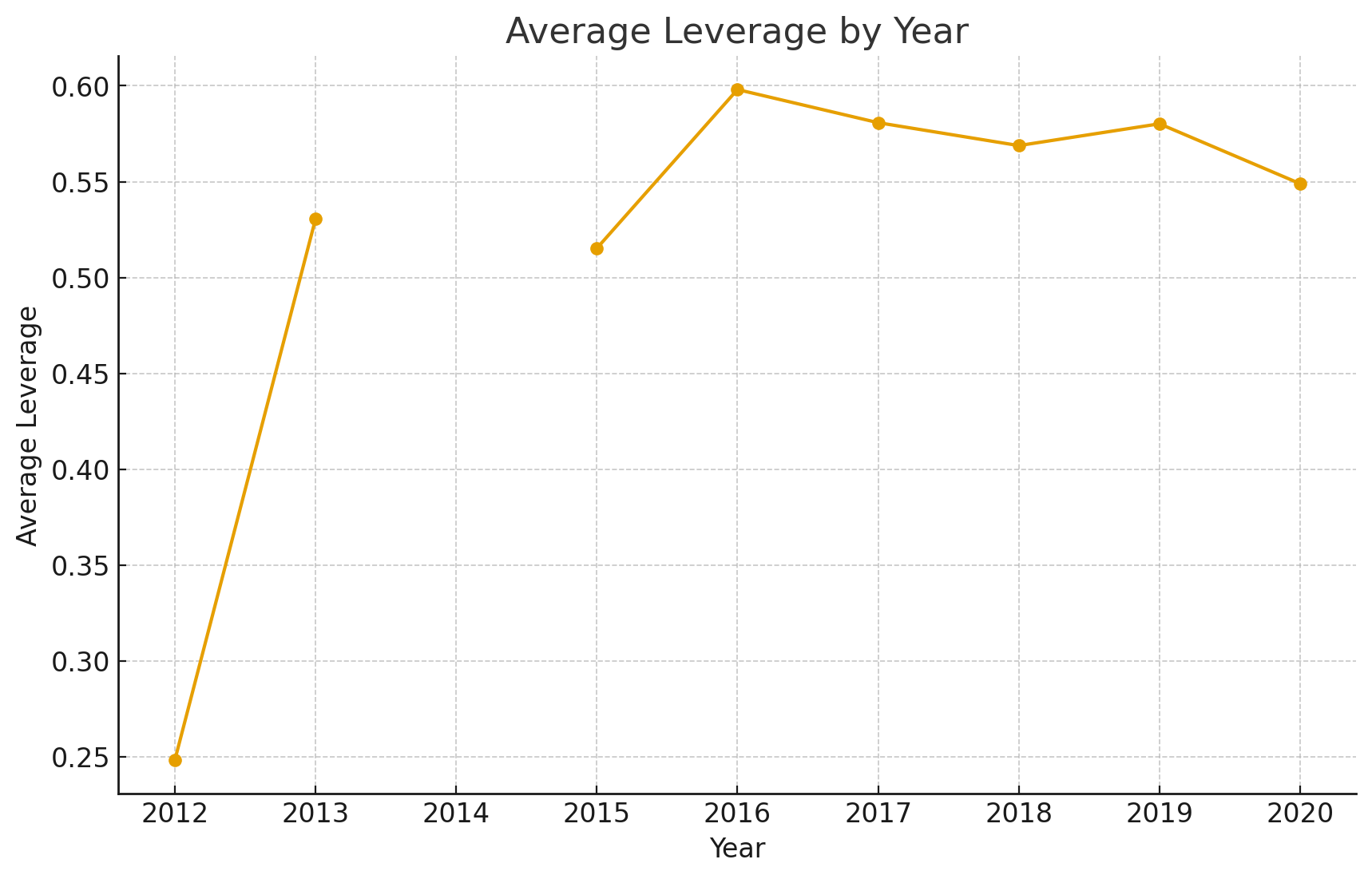
# 9. Technical Documentation

Code documentation, GitHub/Jupyter link, data dictionaries, additional visualisations, statistical outputs.

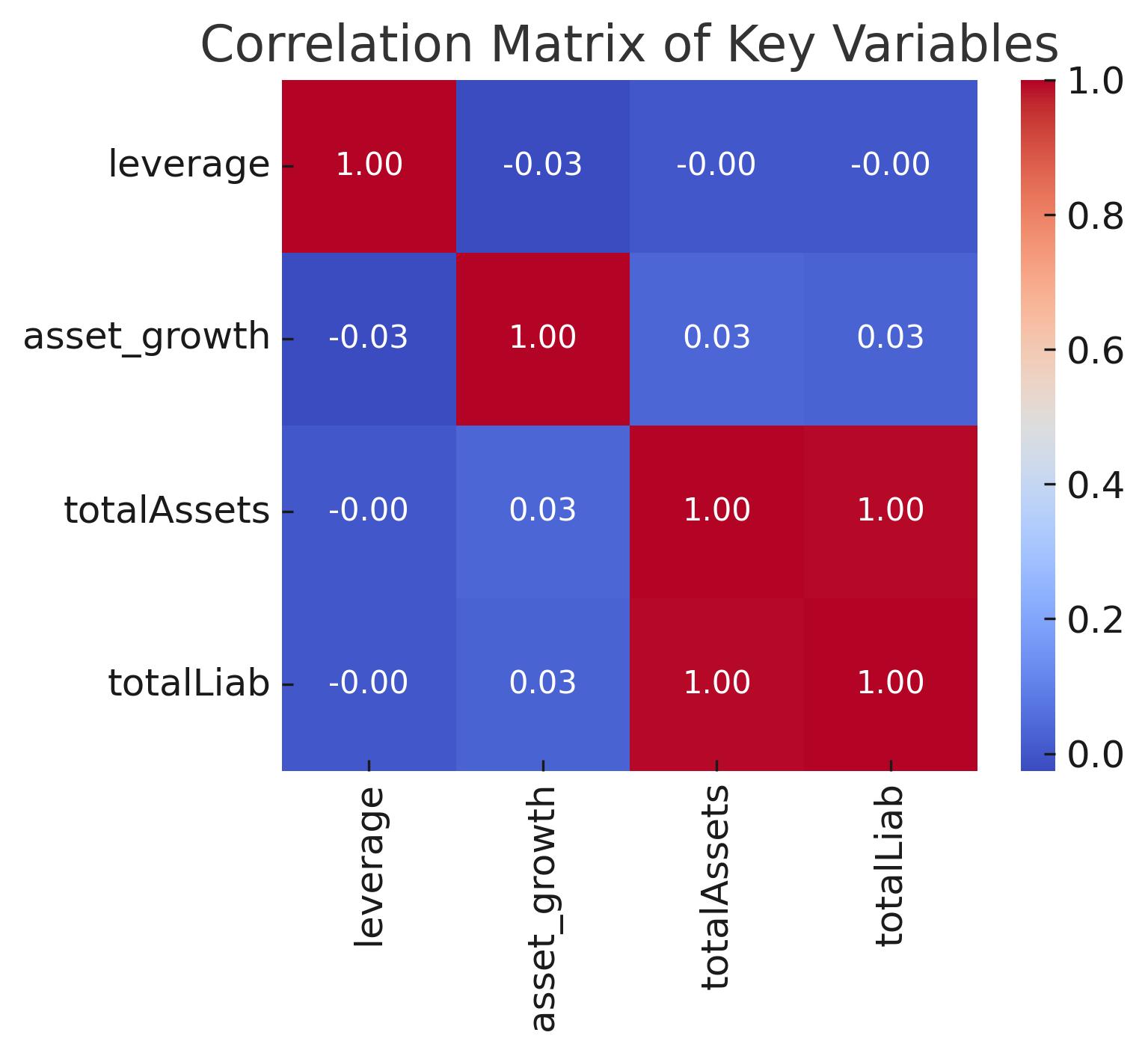
Git Repository Link:







Correlation Matrix of Key Variables



The correlation matrix shows a weak negative relationship between leverage and asset growth. Asset growth is moderately positively correlated with total assets, while leverage is strongly positively correlated with total liabilities, as expected.