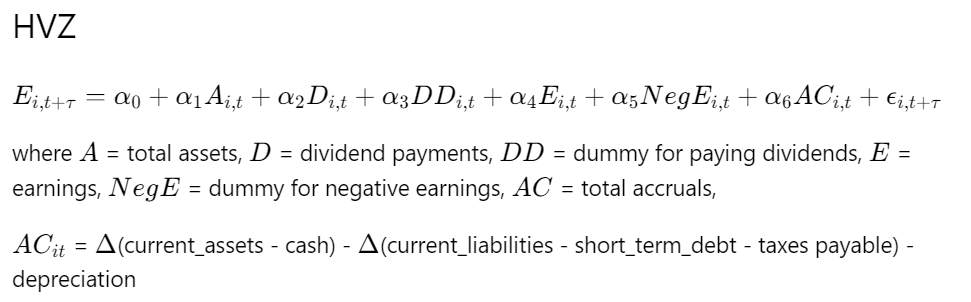
# Introduction

In this group project, I aim to investigate what factors can help predict future earnings. I started by training the three earnings prediction models discussed in class, including the HVZ model, the Earnings Persistence (“EP”) model and the Residual Income (“RI”) model using historical data between 2010 and 2019 to predict future earnings in 2020, 2021 and 2022. I evaluate the performance of my models using adjusted R squared, bias and accuracy. In the next step, I aim to improve the best of the three models by including additional variables, including the most recent analyst estimate, firm-level accounting variables, macro-economic variables and textual information from the MD&A section of 10-K filings. Furthermore, I handpicked a selection of 10 companies across 5 industries of small-cap and large-cap and compared my model’s bias and accuracy against analysts’ estimates. Finally, given my best model, I predict the top 10 companies which are most likely to experience the largest EPS growth in 2023, and select 3 companies to further investigate the potential investment return based on my model’s estimate.

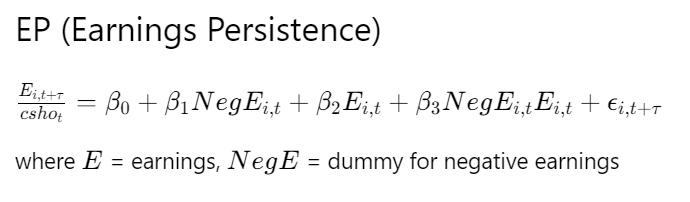
# Part 1

I trained the 3 earnings prediction models (HVZ, earnings persistence and residual income) with training data during 2009 to 2019. Their respective model specifications are as follows:

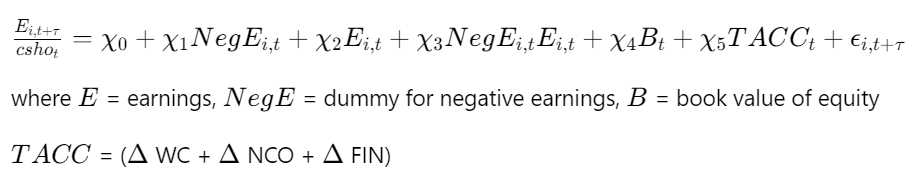
HVZ model



Earnings persistence (EP) model



Residual income (RI) model



Scaling variables with number of outstanding shares

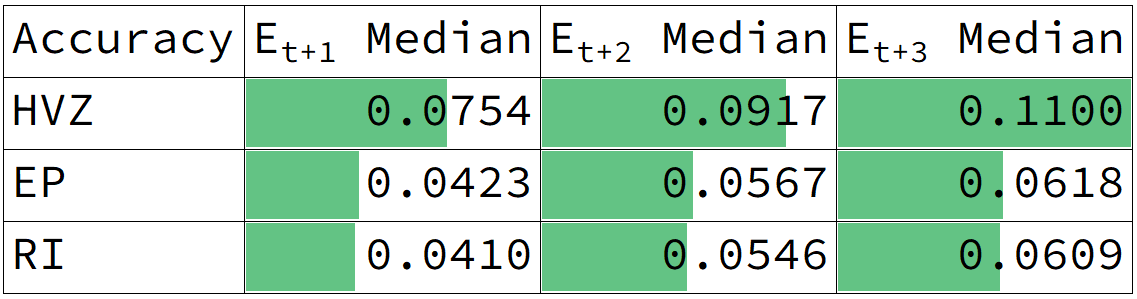
Both EP and RI models predict earnings deflated by the number of outstanding shares, which is the June-end number obtained from the “secm” table in the “comp” library. In addition, all the input variables in the EP and RI model are also deflated by the number of outstanding shares.

Out-of-sample performance comparison

I can compare the three models’ out-of-sample performance with bias and accuracy of each model’s prediction for T+1 (2020), T+2 (2021) and T+3 (2022) earnings. Bias is measured by actual earnings minus earnings forecasts, with the difference deflated by market capitalization for the HVZ model, and by share price for the EP and RI model in order to compare the three models fairly. Accuracy is measured by the absolute value of bias. I decided to only quote the median bias and accuracy as in my training dataset, the mean bias and accuracy are impacted by outliers even after winsorization due to the scaling of bias and accuracy by market cap or share price after fitting the models.

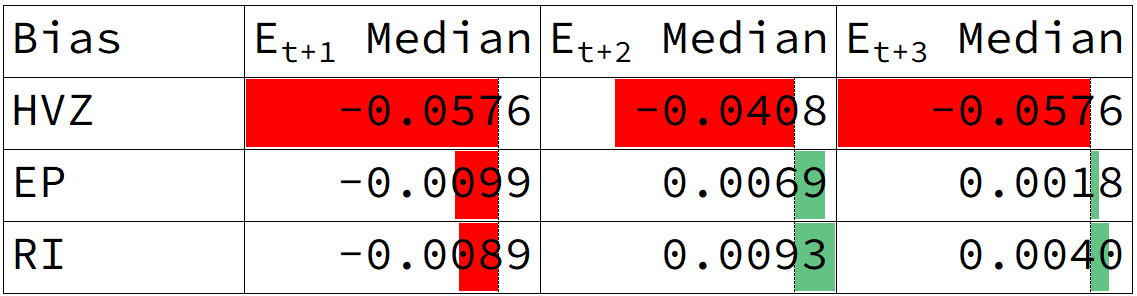
### Accuracy

In terms of accuracy, the RI model has the lowest median forecast error in 1-year, 2-year and 3-year forward earnings prediction. The EP model almost performs as well as the RI model with slightly higher median forecast error, but both of these models outperform the HVZ model by a margin. The performance boost in the EP and RI model likely comes from the interaction term between the indicator of negative earnings and earnings. The coefficients of the interaction term in the EP and RI model are -0.4692 and -0.2393, suggesting a persistent earnings per share with mean-reversal trend. As the current year’s earnings per share becomes more negative, the more positive the next year’s earnings per share will be.



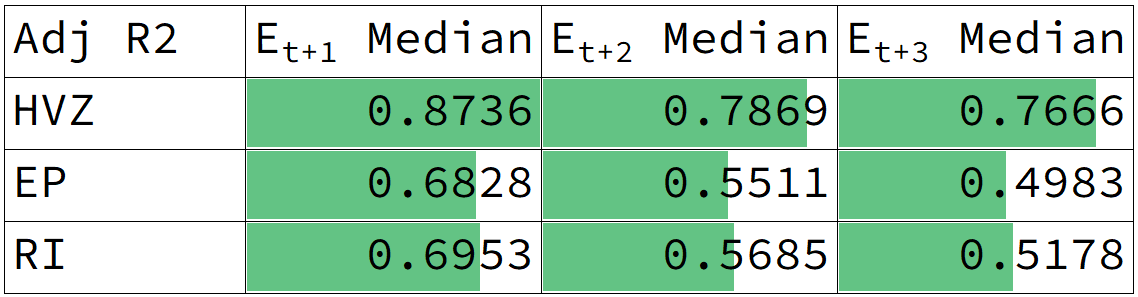
### Bias

In terms of bias, the RI model has the least negative (or most positive) bias, suggesting it provides the most conservation earnings prediction among the three models. On the other hand, the HVZ model has the most negative (or least positive) bias, suggesting it provides the most optimistic earnings prediction among the three models.



### Adjusted R squared

While the adjusted R squared for the HVZ model is the highest among the three models, it is because the HVZ model predicts at the earnings level, while the other two models predict at the earnings per share model. The inherent heteroskedasticity of the regression boosts adjusted R squared. my numbers are similar to the ones mentioned in Li and Mohanram (2014) paper, which mentioned that if the HVZ model is estimated at the per-share level, the adjusted R2 declines to 67.0%, 50.5% and 41.4% (untabulated) for t+1, t+2 and t+3 regressions, respectively. Therefore, the RI model still has the highest adjusted R squared for t+1, t+2 and t+3 regressions.



### Coefficients

The coefficients of each of the three models stay relatively consistent across T+1, T+2 and T+3 regression.

HVZ model

|  |  |  |
| --- | --- | --- |
| T+1 | T+2 | T+3 |
|  |  |  |

EP model

|  |  |  |
| --- | --- | --- |
| T+1 | T+2 | T+3 |
|  |  |  |

RI model

|  |  |  |
| --- | --- | --- |
| T+1 | T+2 | T+3 |
|  |  |  |

# Part 2

Based on the RI model, I consider the following additional input variables that may help explain the variation in 1-year, 2-year and 3-year forward earnings prediction.

### Variable: most recent analyst estimate from IBES library

The first variable I considered is the most recent analyst estimate of earnings per share from the IBES library. The reason why I include this variable is because research has shown analysts, with access to private communication with management and industry knowledge, tend to provide accurate estimates with an optimistic bias. For each company and each fiscal year end, I obtain the most recent analyst estimate based on the date the median analyst estimate is available and the date the financial results are released.

### Variable: firm-level accounting variables

In addition, I also considered firm-level accounting variables, including abnormal changes in production costs, percentage change in sales, gross profit margin, operating income margin, pretax income margin, asset turnover ratio and current ratio. Unlike the variables included in the RI model, which measures accrual-based earnings management, these additional variables measure real-based earning management. Abnormal changes in production costs measure the residual when I regress the production cost against total revenue and the change in total revenue. Unexplained variation in production costs could indicate the management tries to push up or down production costs in order to manage earnings in a way that is not inline with the change in total revenue. From DuPont analysis, return on equity can be broken into profitability, efficiency and leverage. In terms of profitability, I also considered various measures of margins as a percentage of total revenue to indicate the profitability of the core business. In terms of efficiency, I considered asset turnover, which is total revenue divided by total asset. In terms of leverage, I considered the current ratio, which is current assets divided by current liabilities. I made reference to the 2021 paper ("Earning management estimation and prediction using machine learning: A systematic review of processing methods and synthesis for future research." written by Almaqtari, Faozi A., et al.).

### Variable: macro-enconomic variables

Along with firm-level features, I can also incorporate macroeconomic variables that affect the global economy and different societal factors. These variables can provide insight into the impact of time trend and international business exposure on a firm’s performance measured by earnings per share.

The American Federal Reserve interest rate has long been considered a key indicator of the health of the US economy. By influencing the cost of borrowing and lending money, the interest rate can affect investment decisions and spending patterns, ultimately impacting the performance of firms in the US and the world.

Exchange rates between the United States dollar and various currencies, such as the Chinese yuan, G10 currency average, and other major currencies, can also impact a firm’s performance in a given year. Fluctuations in exchange rates can impact the demand for a firm’s products and services, as well as its costs of production and supply chain. These factors can ultimately influence the firm’s profitability and financial performance.

Global GDP, which measures the total value of goods and services produced by all countries in the world, is another key macroeconomic variable that can affect firm performance. Changes in global GDP can indicate shifts in overall economic growth and demand for goods and services, which can in turn affect a firm’s earnings.

Similarly, global inflation is a key factor to consider in predicting a firm’s future earnings. As prices of goods and services increase, firms may face higher input costs or demand for their products and services may decline. This can lead to lower profitability and revenue for the firm. Understanding and incorporating global inflation rates into a predictive model can help identify potential risks and opportunities for a given firm.

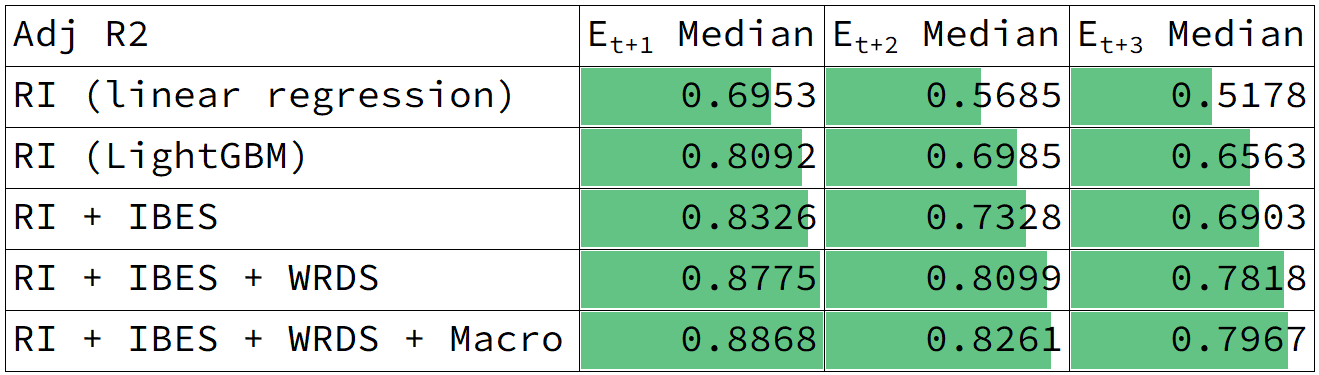
Global consumer spending is also an important macroeconomic factor to consider. Changes in consumer spending patterns and preferences can impact the demand for a firm’s products and services, and ultimately affect its financial performance. Understanding global consumer trends and incorporating them into a predictive model can help identify potential opportunities and risks for a given firm.

In addition to global factors, the United States GDP, inflation, and consumer spending can also impact a firm’s performance. As one of the largest economies in the world, the United States has a significant impact on global economic conditions. Understanding trends in the United States GDP, inflation, and consumer spending can help predict how a given firm may perform in the future, as these factors can influence demand for the firm’s products and services, as well as its costs and access to financing.

### Performance evaluation

Instead of linear regression, I have picked LightGBM regressor as this tree-based model is more flexible than linear regression and can be trained with missing data. Here are the four model specifications to compare based on adjusted R squared:

1. The RI model
2. The RI model + IBES analyst forecast
3. The RI model + IBES analyst forecast + firm-level accounting variables (“WRDS”)
4. The RI model + IBES analyst forecast + firm-level accounting variables (“WRDS”) + macro-economic variables



Given the same set of input variables from the RI model, I was able to boost the adjusted R squared significantly by training a LightGBM regressor instead of linear regression from 69.53% to 80.92% for T+1 regression, from 56.85% to 69.85% for T+2 regression, and from 51.78% to 65.63% for T+3 regression. As I added IBES analyst estimates, firm-level accounting ratios, and macro-economic variables, I was able to continue to increase the adjusted R squared to 88.68%, 82.61% and 79.67% for T+1, T+2 and T+3 regression respectively.

Unlike R squared, which will always go up as I increase the number of input variables regardless of their explanatory power, the adjusted R squared penalizes with the number of input variables. Therefore, the additional input variables should have enough explanatory power in earnings prediction in order to boost the adjusted R squared. By adding the latest analyst estimates, I consider the private communication between analysts and management into my model. By adding firm-level accounting variables, I consider the real-based earnings management that might be performed by the management. By adding macro-economic variables, I consider the impact of time trends and the firm’s international exposure on its future earnings.

### Textual analysis

In addition, I also explored textual analysis using the Management Discussion and Analysis section of 10-K files to see if the textual information would contain additional information to predict future earnings. Due to disk space limit, I was only able to perform the analysis for a selected list of firms, but not for all the 10,000+ firms across 10 years of historical data. Therefore, I decided to only use the textual information to support my analysis in Question 3 (comparison vs analyst estimates) and Question 5 (stock pitches). Here are the definition of the features I created from the textual information:

**Number of words:** Longer MD&A sections can be associated with more thorough analysis and understanding of the firm's finances and macro-economic conditions - this could help improve the accuracy of the model. On the other hand, shorter sections may indicate a simpler business model.

**Proportion of positive/negative/uncertain words:** These features can be used to determine the tone of the company. A higher proportion of positive words could indicate a relatively favorable outlook; while a high proportion of negative words may indicate risks that could have a negative impact on future earnings. Further, an uncertain tone may indicate a lack of confidence and variability in future earnings.

**Fog index:** High fog index indicates complex language which may make the report harder to understand, leading to misinterpretation and errors in earnings predictions.

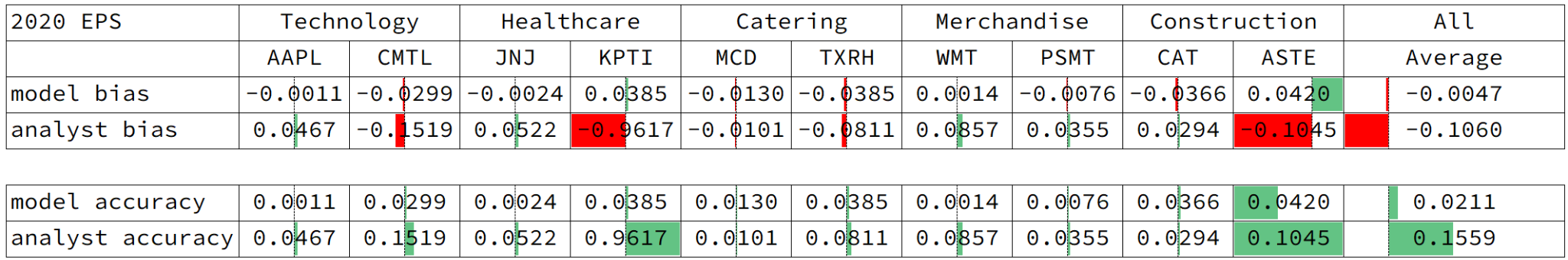
**Polarity:** Positive polarity may indicate a better financial performance leading to higher future predicted earnings, while negative polarity may indicate a decline.

**Subjectivity:** High subjectivity may indicate the report is based on personal opinions rather than facts, making it less reliable as a source of financial data and negatively impact the accuracy of earnings prediction models.

# Part 3

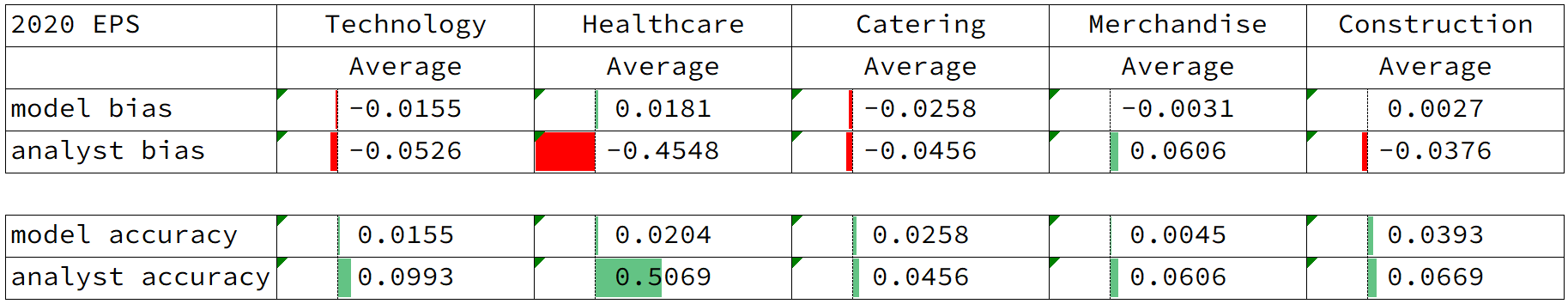
I selected 10 firms across 5 industries in Dow Jones and Russell 2000 index for my out-of-sample 2020 earnings per share prediction, such that I have a broad representation across industry and firm size. The detailed methodology to pick the 10 firms is presented after the results.

### Results

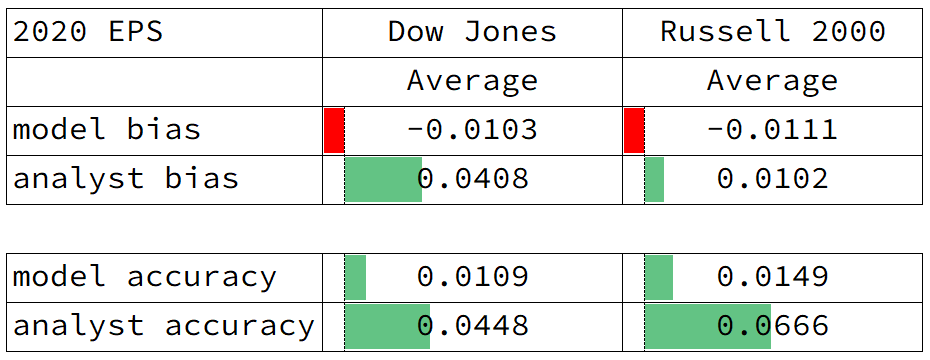


The above results show the bias and accuracy of my model’s prediction for the 10 firms compared with the bias and accuracy of the most recent median analyst estimate based on the “statpers” field from the IBES library.

Based on my final model using the RI model with the most recent analyst estimate, firm-level accounting variables and macro-economic variables, my model is more accurate than the analyst estimate on average, with a mean absolute forecast error of 0.0211. my model is also less optimistic than the analyst forecasts, with a mean forecast error of -0.0047.



If I group by industry as above, it shows that my model’s accuracy ranges from 0.0045 to 0.0393 across five industries, while analysts’ accuracy has a higher variance ranging from 0.0456 to 0.5069. In particular, analysts' estimates are the most optimistic for the healthcare industry, with a mean bias of -0.4568, suggesting they tend to overestimate.



If I group by the index the company is listed in as a proxy of company size, it shows that my model’s bias and accuracy is rather stable across small-cap and large-cap firms. On the other hand, while analysts’ accuracy is also rather stable across small-cap and large-cap firms, they are particularly optimistic for small-cap firms that are listed in the Russell 2000 index, suggesting they tend to overestimate earnings per share.

### Methodology

The 10 firms are selected based on their market capitalization and industry classification. The firms are chosen from both the Dow Jones and Russell 2000 indices. Two firms are selected from the same industry for comparison of the performance of my model in different situations. The list of firms covers a range of sectors including technology, healthcare, catering, general merchandise, and services. Refer to the table below for the results:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **gvkey** | **tic** | **company name** | **industry (sic)** | **Index** |
| 001690 | AAPL | APPLE INC | technology (3663) | DJIA |
| 003358 | CMTL | COMTECH TELECOMMUN | technology (3663) | Russell |
| 006266 | JNJ | [Johnson & Johnson](https://markets.ft.com/data/equities/tearsheet/summary?s=JNJ:NYQ) | healthcare (2834) | DJIA |
| 018930 | KPTI | KARYOPHARM THERAPEUTICS INC | healthcare (2834) | Russell |
| 007154 | MCD | MCDONALD'S CORP | catering (5812) | DJIA |
| 160376 | TXRH | TEXAS ROADHOUSE INC | catering (5812) | Russell |
| 011259 | WMT | WALMART INC | general merchandise (5331) | DJIA |
| 065343 | PSMT | PRICESMART INC | general merchandise (5331) | Russell |
| 002817 | CAT | CATERPILLAR INC | construction (3531) | DJIA |
| 012262 | ASTE | ASTEC INDUSTRIES INC | construction (3531) | Russell |

**Manufacturing - Electronic Equipment Sector -** SIC: 3663

APPLE INC

Business: Apple Inc. is a multinational technology company that designs, develops, and sells consumer electronics, computer software, and online services. The company's products include iPhones, iPads, Mac computers, Apple Watch, Apple TV, AirPods, and a range of services such as Apple Music, iCloud, and Apple Pay.

Industry: Apple operates in the consumer electronics industry, which includes companies that design, manufacture, and sell products such as smartphones, tablets, computers, and other digital devices. The company is a major player in the industry and competes with other large technology companies such as Samsung, Google, and Microsoft.

COMTECH TELECOMMUN

Business: Comtech Telecommunications Corp. is a leading provider of communication solutions for mission-critical applications. The company designs, develops, and markets advanced communication technologies and systems that enable wireless communication, satellite communication, and location-based services for government and commercial customers.

Industry: Comtech operates in the telecommunications industry, which includes companies that provide services and equipment for voice and data communication, such as telephone, internet, and wireless communication. The company's products and services compete with other telecommunications providers, as well as other technology companies that offer similar solutions.

**Manufacturing - Chemical and Allied Products Sector -** sic 2834

[Johnson & Johnson](https://markets.ft.com/data/equities/tearsheet/summary?s=JNJ:NYQ)

Business: Johnson & Johnson is a multinational healthcare company that develops, manufactures, and sells a wide range of products and services in the pharmaceutical, medical devices, and consumer health sectors. The company's products include prescription drugs, over-the-counter medications, medical devices, surgical instruments, and consumer healthcare products such as baby care, skin care, and oral care products.

Industry: Johnson & Johnson operates in the healthcare industry, which includes companies that provide medical products, devices, and services to consumers and healthcare professionals. The company's products and services compete with other healthcare providers and pharmaceutical companies, as well as generic drug manufacturers and biotechnology companies.

KARYOPHARM THERAPEUTICS INC

Business: Karyopharm Therapeutics, Inc. is a commercial-stage pharmaceutical company, which engages in the discovery, development, and commercialization of drugs directed against nuclear export for the treatment of cancer and other diseases

Industry: Karyopharm Therapeutics operates in the pharmaceutical preparations industry, which involves the research, development, manufacturing, and commercialization of various pharmaceutical products and therapies used to diagnose, prevent, treat, or manage diseases and medical conditions.

**Retail Trade - Eating and Drinking Sector -** sic 5812

MCDONALD'S CORP

Business: McDonald's Corporation is a multinational fast-food restaurant chain that operates and franchises a wide range of quick-service restaurants. The company offers a menu of burgers, sandwiches, wraps, salads, and other fast food items, as well as breakfast items and desserts. McDonald's also sells various beverages, including soft drinks, coffee, and juices.

Industry: McDonald's operates in the fast food and quick-service restaurant industry, which includes companies that provide a range of food and beverage products in a convenient, fast-paced environment. The company's products and services compete with other fast food chains and quick-service restaurants, as well as other companies that offer similar products.

TEXAS ROADHOUSE INC

Business: Texas Roadhouse, Inc. is a full-service, casual dining restaurant chain, which offers assorted seasoned and aged steaks hand-cut daily on the premises and cooked to order over open gas-fired grills.

Industry: The Texas Roadhouse operates in the casual dining and restaurant industry, which includes companies that provide a range of food and beverage products in a sit-down restaurant environment. The company's products and services compete with other casual dining and restaurant chains, as well as other companies that offer similar products.

**Retail Trade - General Merchandise Sector -** sic 5331

WALMART INC

Business: Walmart Inc. is a multinational retail corporation that operates a chain of discount department stores, grocery stores, and hypermarkets. The company offers a wide range of merchandise at affordable prices, including groceries, electronics, home goods, clothing, and other consumer goods.

Industry: Walmart operates in the retail industry, which includes companies that provide a range of consumer products in various formats such as brick-and-mortar stores, e-commerce, and other channels. The company's products and services compete with other retailers, both online and offline, as well as other companies that offer similar products.

PRICESMART INC

Business: PriceSmart Inc. is a membership-based warehouse club that operates in Latin America and the Caribbean. The company offers a range of products at low prices to its members, including groceries, electronics, home goods, and other consumer goods. PriceSmart also operates its own private label brand of products.

Industry: PriceSmart operates in the retail industry, specifically in the warehouse club sector, which includes companies that offer a range of consumer products at discounted prices to members. The company's products and services compete with other warehouse clubs and retailers, both online and offline, as well as other companies that offer similar products.

**Construction Sector -** sic 3531

CATERPILLAR INC

Business: Caterpillar Inc. is a Fortune 500 company that designs, manufactures and markets construction and mining equipment, diesel and natural gas engines, diesel-electric locomotives and industrial gas turbines. It is the world's largest manufacturer of mining and construction equipment and one of the world's largest manufacturers of diesel and natural gas engines.

Industry: Caterpillar Inc. operates in the construction machinery industry, which involves the design, manufacturing, and sale of equipment and machinery used in construction, infrastructure development, and related sectors. The industry serves a wide range of customers, including contractors, construction companies, governments, and other entities involved in building and maintaining infrastructure.

ASTEC INDUSTRIES

Business: Astec Industries, Inc. engages in the design, engineer, manufacture, and market of equipment and components used in road building and construction activities. It operates through the following segments: Infrastructure Solutions, Material Solutions, and Corporate and Other.

Industry: Astec Industries operates in the construction machinery industry, which involves the design, manufacturing, and sale of equipment and machinery used in construction, infrastructure development, and related sectors. The industry serves a wide range of customers, including contractors, construction companies, governments, and other entities involved in building and maintaining infrastructure.

# Part 4

Based on my final model using the RI model with the most recent analyst estimate, firm-level accounting variables and macro-economic variables, I have picked the top 10 firms that are most likely to exhibit the largest earnings growth in 2023 in descending order of EPS growth, defined as (2023 predicted earnings - 2022 realized earnings) / absolute value of 2022 realized earnings - 1.

The “earnings\_lead1” column refers to the predicted earnings per share in 2023. The “earnings” column refers to the actual earnings per share in 2022. The “eps\_growth” column refers to the growth in predicted earnings per share between 2022 and 2023.



# Part 5 - 3 selected stocks

Within the top 10 firms, I have selected the following 3 companies using qualitative information.

### Stock 1 - QuoteMedia Inc.

QuoteMedia Inc. is a leading provider of customizable financial market data, news, and research solutions to the investment industry. The company's services include real-time stock quotes, market news and analysis, historical data, financial content syndication, mobile solutions, and web-based financial applications. QuoteMedia's data feeds cover global equity, option, and futures exchanges, as well as over-the-counter markets and mutual funds. Its clients include banks, brokerage firms, financial planners, investment websites, and mobile app developers.

**Market Opportunities:**

The financial market data and research industry has experienced significant growth over the past few years and is projected to continue growing in the foreseeable future. The increasing demand for real-time and historical financial data and analysis, as well as the growing adoption of digital platforms by investors, are the key drivers of this growth. QuoteMedia is well-positioned to capitalize on this trend, with its comprehensive suite of financial data and research solutions that cater to the needs of various types of investors, including individuals, institutional investors, and financial advisors.

**Financial Performance:**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Market Cap | Earning | EPS growth | Accruals | Gross Profit/Revenue | Sales Change | Asset Turnover |
| 19.11M | 0.005M | 44.36 | -2.55 | 0.61 | 0.15 | 2.67 |

QuoteMedia appears to be in good financial health with large growth potential. With a positive earning, the EPS of QuoteMedia is expected to grow about 44 times. This growth could be explained by the increasing sales, which indicates that the company is experiencing growth in revenue. Additionally, QuoteMedia has reported negative accruals in its financial reports, which suggests that its earnings are of high quality and reliable. QuoteMedia has a relatively high gross profit/revenue ratio, which suggests that the company is generating a high level of gross profit relative to its revenue. This indicates that the company is operating efficiently and is able to generate profits from its core business activities to reach long-term success. Lastly, QuoteMedia has a high assets turnover rate, which suggests that it is using its assets efficiently to generate sales and profits. This is a good sign for its financial health and indicates that the company is making the most of its resources.

**Competitive Advantage:**

QuoteMedia has several competitive advantages that have helped it establish a leading position in the financial data industry. These advantages include comprehensive data coverage, a robust technology platform, customizable solutions, an experienced management team, and strong partnerships with major financial institutions and news providers.

**Company News & Management Expectation:**

In 2022, QuoteMedia signed and launched major multi-year agreements with two of Canada's largest banking institutions and several other multinational financial firms. The company is currently in negotiations with several large firms and expects its revenue growth in fiscal 2023 to match or exceed the annual revenue growth achieved in 2022, with a significant improvement in net income figure. In 2022, QuoteMedia achieved SOC2 Type II certification, which provides independent assurance that the company maintains a high level of information security, data integrity, and business resiliency. The company's growth in revenue and market share is fueled by its development of exciting new data applications and products, as well as the expansion of its global market coverage, which is expected to continue throughout 2023 and beyond.

### Stock 2 - International Isotopes Inc.

International Isotopes specializes in the supply of radioisotopes used in medical, industrial and research applications. It also provides nuclear engineering services. The company has a unique expertise in the development, production, and distribution of a variety of isotopes and their derivative products. The company's products and services are used in a wide range of applications, including cancer treatment, cardiac imaging, and radiopharmaceutical research.

**Financial Performance:**

International Isotopes Inc (INIS) had the highest indicated earnings growth potential, as indicated by its earnings per share (EPS) growth of 219 times in just the next year. This suggests that the company is expected to experience significant growth soon. INIS has a relatively higher market cap than my previous recommendation (Quote Media) – of around $35M, with significant potential for further growth.

The relatively low value of Beta indicates its volatility is on the low end, and this could be a safe investment to hedge against the more volatile stocks (see Stock 3 - Mammoth Energy). A current ratio of 1.63 suggests that it has sufficient liquidity to meet any of its short-term obligations. It has seen steady growth in operating profit margins over the last decade, and will likely continue this trend.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Market Cap | Revenue | Total Accruals (RI Model) | Book Value | Current Ratio | Beta |
| 35.1M | 11.2M | -0.000571 | 8578 | 1.63 | 0.49 |

Assuming the same price-to-earnings ratio for the next year, I can forecast the share price to the potential increase.

Current Share Price=$0.07; Earnings (2021)=$0.000598; Predicted Earnings (2022)=$0.132

Potential Future Price=0.07/0.000598\*0.131993= $15.4

This is a massive predicted increase, indicating that it is likely to be undervalued, to accommodate for risks and statistical nature of my predictions model, I can set a floor value of the expected future share price to be around $1-$2.

Based on text analysis and reading through the MD&A - I can infer that the company serves many markets and has a wide array of product lines. This can help mitigate risks associated with a single product line, industry or market segment.

**Textual Analysis:**

Based on comparison of the language used in the MD&A in 2010 and 2022, I see some subtle differences. In terms of the positive words used – they seem to be more hopeful in 2011, using words ‘improve’, stabilizes’, and ‘enhance’. Negative words like ‘loss’, ‘limited’ and ‘depletion’ could indicate poor financial performance and funds.

In comparison to 2022, words like ‘profitable’, ‘highest’, ‘excellent’, and ‘opportunities’ seem to indicate better performance and seem somewhat indicative of future growth. The most negative words seem to be associated with fraud and legal issues, rather than financial performance. See word clouds below for reference.

### 

Word Cloud from MD&A (2011):

Positive Words: 1% | Negative Words: 0.7%

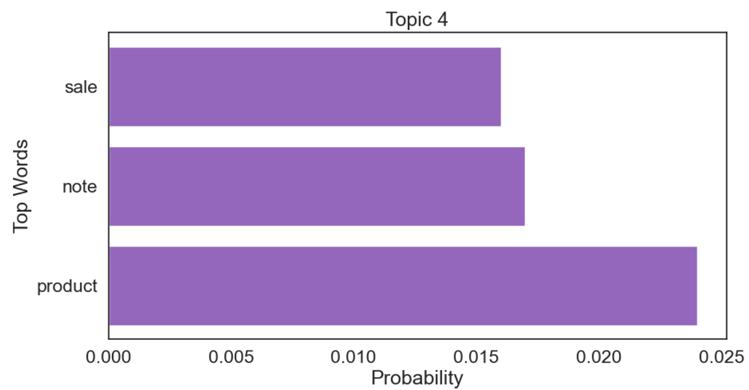
### 

Word Cloud generated from MD&A (2022):

Positive Words: 0.79% | Negative Words: 0.81%

## **Topic Modeling:**

I used LDA modeling technique to look for words and topics that stood out in the filing. Based on the most important topics, I see a trend of words involving product and sales. Having checked their website and going through the MD&A, I see a strong emphasis on product development and their operations. INIS has a diversified business and product portfolio, including its production of isotopes and radiological services. This diversification could help mitigate risks associated with a single product line or industry. Most recently, they also completed the sale of certain unused assets for $4.0 million in cash, resulting in a net gain of $1.8 million, which can also be seen in the topic visualized below.



Words associations (LDA Topic Model)

### Stock 3 - Mammoth Energy Services

#### **Introduction**

Mammoth Energy Services is a US-based oil and gas exploration and production services provider founded in 2014 and headquartered in Oklahoma. The company operates through three segments: Infrastructure Services, Pressure Pumping Services, and Natural Sand Proppant Services. The Infrastructure Services segment provides construction, upgrade, maintenance, and repair services to the electric power delivery systems, and telecommunications and renewable energy industries. The Pressure Pumping Services segment offers hydraulic fracturing services in oil and natural gas basins across North America. The Natural Sand Proppant Services segment mines and sells silica sand used in hydraulic fracturing operations.

#### **Broad Statistics**

### 

Earnings growth of top 10 analyzed companies

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Market Cap | Revenue | Cash Flow | Adj EBIDTA | EPS | Beta |
| 182.7M | 416.11M | 20.89M | 86.1M | -0.01 | 1.76 |

#### **Current Status**

Mammoth Energy Services has been growing its industrial businesses continuously, while its equipment manufacturing operations have provided the company with the ability to repair much of its existing equipment in-house and manufacture certain new equipment it may need in the future. Additionally, the demand for well completion, infrastructure, and natural sand proppant services has remained strong in 2022, with adjusted EBITDA of $86.1 million. The company has doubled its active frac fleet count and converted pumping fleets to dual-fuel spreads, which has resulted in increased upfront costs but has saved the company in long-term environmental and financial costs.

Moreover, Mammoth Energy Services has exhibited strong performance in its well completion services division during 2022, fueled by the increase in demand in the pressure pumping industry. Similarly, operational improvements and an increase in crew count have driven enhanced results for its infrastructure services, with the crew count rising from 82 in December 2021 to approximately 91 in December 2022.

The COVID-19 pandemic and resulting economic conditions have not had a significant impact on demand or pricing for Mammoth Energy Services' infrastructure services. The company has continued to add crew capacity, while funding for projects in the infrastructure space remains strong with added opportunities expected from the Infrastructure Investment and Jobs Act, which was signed into law on November 15, 2021.

#### **Looking Ahead**

Looking ahead, Mammoth Energy Services will continue to monitor the industry and market conditions resulting from the COVID-19 pandemic and take mitigating steps in an effort to preserve liquidity, reduce costs, and lower capital expenditures. These actions may include reducing headcount, adjusting pay, and limiting spending.

Furthermore, improvements in the oilfield services industry and in both pricing and utilization of the company's well completion and drilling services are expected to continue in 2023. Mammoth Energy Services plans to continue adding crew capacity and focusing on operational improvements to drive enhanced results. In addition, the company expects increased opportunities in the infrastructure space due to the Infrastructure Investment and Jobs Act, which is expected to drive further growth in this segment.

#### **Overall View**

Mammoth Energy Services has been expanding and growing its operations in the energy sector, with its income and assets primarily attributable to its four reportable segments. The company has exhibited strong performance in its well completion and infrastructure services divisions, and the demand for its services remains strong in 2022. The COVID-19 pandemic and resulting economic conditions have not had a significant impact on demand or pricing for the company's infrastructure services. Looking ahead, the company will continue to monitor market conditions and take mitigating steps to preserve liquidity and reduce costs, while focusing on driving enhanced results through operational improvements and crew capacity additions. Overall, the company is poised for growth and continued success in the energy and infrastructure sectors.

# Conclusion

In this group project, I aim to investigate what factors can help predict future earnings. I started by showing the RI model is the best model among the three earnings prediction models discussed in class in terms of out-of-sample bias and accuracy. The boost in performance compared with the HVZ model is likely due to the additional interaction term between the indication of negative earnings and current earnings per share, suggesting a persistent mean-reversal pattern in earnings per share.

Next, I improved the RI model by changing my choice of model from linear regression to a more flexible LightGBM regressor, boosting my adjusted R squared from 69.53% to 80.92% in T+1 regression, and similarly boosting the performance in T+2 and T+3 regression. I also saw a continuous boost in performance as I gradually include the most recent analyst estimate, firm-level accounting variables, and macro-economic variables.

To compare my model’s prediction with analysts’ forecast, I handpicked a selection of 10 companies across 5 industries of small-cap and large-cap and showed that my model is both more accurate in terms of accuracy and less optimistic in terms of bias. In particular, I observed that analysts’ forecasts are the most optimistic for healthcare companies and small-cap companies, while my model’s performance has been consistent across different industries and varying firm sizes.

Finally, given my best model, I predict the top 10 companies which are most likely to experience the largest EPS growth in 2023, and further investigate the potential investment return in 3 companies, including International Isotopes (INIS), QuoteMedia Inc. (QMCI) and Mammoth Energy Services (TUSK) based on my model’s estimate of earnings per share in 2023. I supported my analysis with both qualitative information and qualitative information such as textual information from the Management Discussion and Analysis section in their 10-K filings.