“Exploring Alternative Data in Systematic Investing for Retail Investors”

Author: Peter Zak

Contact Information: peter.zak@edu.escp.eu

Affiliation: ESCP Business School

Supervisor: Prof. Dr. Axel Buchner

Keywords: alternative data, systematic investing, retail investors, commodity data, consumer sentiment

JEL Class: G11, G14, C55, C58, Q41

Date: November 2025

Abstract

There are three main types of company analysis: fundamental, technical, and behavioral: each involving a different way of looking at and analyzing companies. The fourth type of analysis that has emerged in recent years involves alternative data. It relates to any kind of unofficially obtained data that is nonetheless publicly available and can give additional insight into companies' performance and operations. It can include sentiment analysis, web traffic, oil imports, satellite imagery, etc. and it has gained popularity tremendously among institutional investors in recent years. However, retail investors have largely been excluded from access to alternative data, due to the cost of obtaining and difficulty processing it, which is why this study explores various ways that they can leverage it. This is a largely unexplored research gap, as there are few papers that study how this type of data can be implemented and instead focus on the rate of adoption, they are rarely quantitative in nature and if they are – they don't explore data beyond sentiment analysis. That is why this paper will divide the research into the sector-wide shaping trends and company-specific factors sections and quantitatively explore different types of alternative data that can help predict stock prices, revenue levels, and cost of goods sold using multivariable linear regression and correlation analysis. For the sector-wide analysis, the study will focus on the iShares Global Clean Energy ETF (ICLN) and for the company-specific section on Tesla, Inc. The paper will showcase which specific types of data can help in predicting the above-mentioned factors, introduce various conceptual and academic evidence, and then explore these quantitively for the specific securities inspected. The thesis relies on secondary data sources and uses data from platforms like: Nasdaq Data Link, Yahoo Finance, Google Trends, Altindex, etc. Further, the study explains how these alternative data types can be included in the systematic investment process, especially if it involves the top-down approach and is longer-term oriented. The paper finds that including alternative data after identifying relevant data points increases the accuracy of the multivariable linear regression by increasing the adjusted R^2 and reducing the RMSE, compared with the benchmark models without these non-standard inputs. However, the study also finds that multivariable linear regression fails to capture nonlinear patterns, which are sometimes inherent to alternative data, even if the academic sources and correlation analysis suggest they should be relevant. Finally, the paper presents the most significant hindrances for retail investors to accessing alternative data, how they can overcome some of them, and include it in their investment analysis process. Despite the study being oriented to retail investors, the results may be used by investors of all sizes wishing to enrich their data-driven investment approach with alternative data.

Acknowledgements

I would like to express my gratitude to those who guided me throughout my research and thesis wiring process, in particular Prof. Dr. Axel Buchner and Research Associate & PhD Candidate - Mats Lützenkirchen. They supported me by guiding my exploratory process, suggesting on which statistical models and tools to use, how to present my findings, giving me feedback, and answering all my questions. They have both been tremendously helpful and without their insight and support, I would not have been able to make this thesis as insightful as it is. Once again, my profound gratitude and thanks to both of them for their support throughout this academic journey.

Table of Contents

List of Figures:

List of Tables:

List of Abbreviations:

AD – alternative data / non-standard data

SI – systematic investing

ETF – exchange traded fund

IV – independent variable

EV – electric vehicle

1. Introduction

Institutional investors have shown great ingenuity throughout the years in discovering non-conventional methods and leveraging non-financial data to gain additional insight into companies' performance in order to make more informed investment decisions. Traditional methods of evaluating a company can be divided into three main categories: fundamental, technical analysis, and behavioral analysis. The fourth type of analysis, which has gained traction among institutional investors, mainly due to the massive dissemination of data, involves the use of alternative data (AD) to enhance the investment selection process through the improvement of the predictive models. AD refers to unofficially obtained data that can provide insight into business operations and can be used to investigate financial markets instead of traditional data. (Wang & Ngai, 2024, p. 2). This non-financial type of data can include weather patterns, oil imports into a particular country, satellite imagery of manufacturing facilities, sentiment scores of companies' products, or any other type of information that can give additional insight into industry trends or companies' performance. Systematic investing (SI), on the other hand, pertains to an investment selection process that is primarily data-driven and has a predefined set of rules that guide the analysis process, making it replicable for a variety of other securities, while also combining some aspects of discretionary investing (Harvey, 2021). This paper will focus on how AD can be incorporated into the SI analysis process and show how retail investors, rather than institutions, can source data to train models and feed real-time information to enhance the predictive capability of statistical tools in order to make more informed investment decisions and gain an edge in analyzing companies.

Academic research on this topic is extremely scarce overall, mainly due to the high cost of collecting and processing AD. Many academic papers focus on the rates of adoption of non-traditional data itself but fail to prove its predictive capabilities using statistical tools (Hansen & Borch, 2022, p.12). Others generalize and try to logically prove cause-effect relationships, thus showing the added value of AD without quantitatively exploring the underlying correlations (In, Rook, & Monk, 2019, p.6). Furthermore, the few academic papers that try to quantitatively leverage non-financial data for improved model performance, such as (Frattini et al., 2022, p.1), only rely on sentiment analysis, which is relatively easy to obtain, considering the financial hindrances and processing complexity to accessing AD and focus mostly on the returns of the trading strategy, rather than the underlying model's accuracy and predictive capability itself. There are no research papers that also focus on the fundamentals of companies or entire sectors and show how these provide keener insight, while maintaining a general approach without delving too deeply into any specific industry. This study will quantitatively explore this research gap by focusing on non-traditional data pertaining to the actual performance of companies and sector shaping trends, rather than only the market participants perception of these. Moreover, this paper will concentrate on how incorporating this additional data increases the accuracy of the predictive models to assess continuous values, such as quarterly revenue or stock prices, instead of back testing and calculating returns from a specific trading strategy over a period of time.

The research questions this thesis will attempt to answer are the following:
RQ1: Which publicly accessible alternative data types have the potential to predict future stock prices or company earnings, and what is their predictive capability?

RQ2: How to incorporate them into a systematic investing analysis process?

RQ3: How can retail investors access and make use of alternative data to improve their investment choices and make more informed investment decisions?

This paper will attempt to answer these questions by dividing the research into two main parts: industry shaping factors and company-specific insight. The former will explore which non-traditional renewable energy industry shaping factors have the power to aid in predicting industry behavior as a whole, while the latter will delve into the company-specific factors that are beneficial to predicting company's stock price and quarterly performance. Analyses will be conducted across different time horizons and security types. Secondary data types have been collected from sources such as Nasdaq Data Link and Joint Organizations Data Imitative (JODI) or Altindex, organized as a time series to match securities financial metrics and analyzed using quantitative methodology. Specifically, multivariable linear regression models were compared with each other and the predictive value of adding AD was assessed. Additionally, the study relied on correlation analysis. For the literature review, the study relied on databases such as Google Scholar and Scopus. The researcher considered only studies in English and to guide his research, used keywords, such as: “Non-traditional Investment Data Sources”, “Rules Based Investment Strategies”, “Renewable Energy Transition”, or “Electric Vehicle Production Dynamics”. The coding strategy relied on iterative, inductive coding, as the author developed and narrowed down his research as he studied research articles, identified cause-effect relationships and correlations, which he later tested quantitatively.

The paper will contribute to the current research by illustrating how retail investors can leverage AD and streamline their decision-making process by making more accurate company performance predictions and later embedding this insight into a portfolio construction process. Non-financial data has been adopted by institutions, such as hedge funds, and insight obtained from it has been used to guide investment decisions. (Monk, Prins, & Rook, 2019, p. 16). However, retail investors, due to the high cost of obtaining this type of data, have been largely excluded from accessing AD (Ekster, Kolm, 2020, p.5) and this study will demonstrate which publicly available information can be collected, processed and incorporated to enhance the predictive models. Future researchers can delve even deeper into specific segments of my study to explore additional features that help assess the industry trends and individual companies’ performance.

A literature review following this introduction will clarify concepts explored in this paper and lay the foundation to conceptually understand the factors which are quantitively measured and proven in this paper. It will also be divided into the industry shaping trends and firm specific AD explored in this research. This is followed by the methodology section, which explains how data was collected, verified and analyzed. The actual findings and results are presented in the results section, following the discussion of the limitations of AD analyzed and obstacles for retail investors. The general conclusion, summarizes each part and provides a high-level answer to the research question reiterating what was already mentioned in more detail in each part.

The thesis ends with a list of references and an appendix with all the relevant datasets and model outputs attached.

2. Literature Review

The Literature Review aims to clarify concepts introduced in the thesis and established a context in which various parameters are analyzed, serving as a foundation, supporting the analysis process. First two subsections, will provide definitions of both AD and SI, which understanding of is imperative to comprehend the scope of the paper. Further subsections will explore on various aspects of both, industry shaping factors and company-specific metrics and explain how they are intertwined and influence each other using secondary literature.

For the sector-wide trends, the study focused on the renewable energy industry and the AD specific to that field. This orientation is substantiated by the available datasets, sourced from reliable sources, available to the majority of retail investors, toward whom this study is designated. The second, company-specific, part explores miscellaneous factors to gauge the performance of publicly traded firms, which are not specific to any enterprise and can be generalized to most publicly traded companies. The orientation towards retail investors is based on the extreme data availability and the difficulty in processing this data, so this study focuses on widely accessible datasets that the majority of above-mentioned investors can utilize.

2.1 Definition of Alternative Data

Alternative data is defined as unofficially obtained data that can be leveraged to provide additional insight into companies operations and can be used to investigate financial markets instead of or in combination with traditional data (Wang & Ngai, 2024). It can include any kind of information that can give investors a different perspective or keener view on business operations of a certain company or industry trends. It can include: web traffic data, social media sentiment analysis, satellite imagery of manufacturing facilities, oil and natural gas prices and imports into a particular economy, minerals prices that might influence a company's earnings,etc.

The dissemination of non-financial data for investment analysis purposes has been massive in recent years, with the expected revenue of AD providers to exceed that of traditional data providers by 2029 and grow at an exponential rate (Deloite, 2023). This growth showcases the added value of AD and has been primarily driven by asset managers and other institutional investors. This thesis will explore how retail investors, in contrast, can make use of non-traditional data to improve the decision-making processes.

Leveraging AD for improved portfolio selection and thus excess returns goes against the semi-strong and strong form of the Efficient Market Hypothesis, which respectively assume that prices already reflect all publicly available information and that they incorporate all information, both public and private (Fama, 1970). AD is information that is publicly available, but difficult to access, process and extract insight from. However, it adds value to the predictive models and allows investors to outperform benchmarks, for example in the study of Frattini, Bianchini, Garzonio, & Mercuri in 2022, where the researchers created "The Trend indicator", based on sentiment scores, and added it to the algorithmic trading model, which outperformed the conventional one without it.

2.2 Systematic Investing Analysis Process

Systematic Investing is a rules-based investment analysis, security selection and portfolio construction approach which is heavily data driven and where algorithms and predefined models are responsible for making trade decisions, rather than human discretion (Harvey, 2021). The models are usually back-tested and fine-tuned to predict securities and trends given specific market conditions. The data-oriented nature of SI allows for quantifiable information to be incorporated into the model (State Street Global Advisors, 2023), such as certain types of non-traditional data, which this study focuses on. If such AD improves the back-tested strategy returns and is deemed statistically significant, then this data type can be fed into the model in real time and used in a model with predefined coefficient or weight values.

SI differs from quantitative investing mostly by still allowing for a certain degree of human discretion to influence trading decisions. Analysts may translate their subjective opinion into a numeric figure, which can later be incorporated into the model (Harvey, 2021), which indirectly allows for discretionary judgment to steer the decision-making process.

The top-down portfolio selection is an approach that begins with analyzing macroeconomic trends, followed by sector evaluation and ending with individual security selection (Crescenzi, 2008). It goes in opposition to the bottom-up methodology, which focuses on individual security selection based on the micro-level data, as (Crescenzi, 2008) puts it, and de-emphasizes the importance of global trends and market cycles. It has great implications for the rule-based SI process, which can be applied for different time horizons, from intraday algorithmic trading to long-term investing and position holding. If an investor is long-term oriented the top-down approach is more favorable, as historically it has led to superior performance compared with the bottom-up, which can easily load large weights on noisy factors and result in underperformance compared to the top-down framework (Zurek & Heinrich, 2021).

That underscores the significance of exploring the underlying market factors that shape industry and market trends and that is exactly what the first research section will focus on.

2.3 Renewable Energy Industry Shaping Factors

For the industry-wide influencing trends section of this study, the renewable energy industry worldwide has been selected to explore how AD data can help predict the performance of the entire industry. The decision to focus on that specific sector can be substantiated with the fact that the non-standard data, which can aid in predicting its performance, can be relatively easily obtained compared with other industries, such as the pharmaceutical or the technology industry. This conclusion results from the research conducted by the author and the exploratory nature of this study.

In order to gauge the performance of the green energy industry, this paper focuses on Exchange Traded Funds (ETFs), which mirror the performance of the renewable energy companies. For that purpose, iShares Global Clean Energy (ICLN) has been selected, as it includes the world's biggest green energy focused manufacturers, such as Iberdrola, Vestas Wind Systems, SSE Plc, First Solar and it has been created already in June 2008 (BlackRock, 2025), which gave the researcher vast historical data to test it for. Moreover, it the largest renewable energy ETF with $1.6 B in Assets Under Management (AUM), next one being First Trust NASDAQ Clean Edge Green Energy Index Fund (QCLN) with only $366 M in AUM (Nasdaq, n.d.), thus additionally reinforcing the notion that ICLN reflects the entire industry best and is most suitable for this study.

The subsections below will explore the specific factors that are non-traditional financial data, yet may be used to better predict the performance of the green energy industry and thus by definition are considered AD.

2.3.1 Oil Price

According to Omri & Nguyen (2020), higher oil prices can entice companies to allocate a higher portion of their investment portfolio to green energy solutions, which should be cost effective in the long run. Such decisions impact the revenues, reduce costs per unit, due to economies of scale, and increase net earnings of renewable energy solutions manufacturers, which is reflected in their stock price. This means that oil prices have an indirect, negative and statistically significant impact on renewable energy consumption, indicating that oil and green energy products can act as substitutes, not only in the long run, but also in the short term (Omri & Nguyen, 2020).

Inversely, lower oil prices may propel companies to hold back on the renewable energy investments, mainly due to the high cost of implementing their systems (Kumar, Managi, & Matsuda, 2012). These price levels can be measured both globally and domestically and these changes are reflected in the prices of ETFs, which mirror the underlying companies performance and stock prices.

2.3.2 Gas and Oil Product Inflows & Exports

Historically, higher oil rents and increased primary and secondary oil products into an economy have resulted in decreased green energy adoption and increased cost of implementing renewable energy systems (Makki, Kaspard, Khalil, & Mawad, 2024). This follows the law of demand, which states that as the demand of a good decreases, the price of that good increases (OpenStax, 2022), which further reinforces the trend.

This finding, once again, points to the substituting effect between renewable energy solutions and oil products, demand for which can be gauged inspecting the import and export figures for an economy and comparing it with previous years trends and other benchmarks (Kumar et al., 2012).

2.4 Individual Company Alternative Data Dynamics

For the individual company analysis, the study focused on Tesla Inc. (TSLA) and explored various AD sources that might give additional insight into the performance of the company or its stock development, which is a reflection of how the investors and the market perceive the company.

Tesla went public in 2010 and, as any other publicly listed company in the United Sates, it has an obligation to disclose annual and quarterly earnings reports. Moreover, it has been the subject of numerous debates, and it has significantly gained on popularity, as indicated by search activity on Google Trends (Google Trends, n.d.), which means that more data providers, which this study utilized, such as Altindex, focused their efforts on providing information about this company with more detail, accuracy and granularity than for other enterprises.

2.4.1 General Interest in the Company

One of the most obvious and easily attainable type of non-standard data is the level of interest in the company and its products along with the respective sentiment level towards them, as it may translate to sales and stock price. While traditional investor sentiment survey data only produces lagged results of what has already been shown in the stock price, using Google Trends data and the number of mentions of financial terms on Twitter along with their overall sentiment is statistically significant for predicting returns for the respective entity (Mao, Counts, & Bollen, 2011). This aligns with the already conceptually explained relationship between these variables.

Moreover, the study conducted by Mao, Counts and Bollen (2011) proved that leveraging news media, web search and social media feeds can be useful in predicting the market returns the next day and references Philip E. Tetlock, who found that increased level of pessimism often precedes lower market returns the following day. This study borrows from their findings by trying to capture the “chatter” longer term and matching it with its respective sentiment to arrive at more accurate predictions, specifically for Tesla.

2.4.2 Cost of Manufacturing Electric Vehicles

The previously mentioned factors can significantly aid in the predicting the level of interest, sales and thus – revenue for a company, such as Tesla, but another tremendously important insight is how efficiently the enterprise manufactures their product or delivers their service and how much it costs them. This later translates to gross profit, net income and eventually – Earnings per Share (EPS), which is often one of the most significant factors that analysts look at.

Even though, TSLA provides also energy generating and storage products, the automotive sales of EVs accounted for 85.2% of their revenue in 2023, thus constituting their main revenue stream (Tesla, Inc., 2023, p. 54). The most expensive component of manufacturing EVs are lithium ion-batteries, which is associated with the high cost and nonfungibility of essential metals, such as lithium, cobalt and nickel. (Orangi, 2023) These elements are essential to building these batteries, which later power EVs, such as Tesla’s. Furthermore, Orangi (2023) concludes that many factors contribute to the adoption of the lithium-ion batteries and that even though technological advancements and beneficial market dynamics can lead to substantial cost reductions for EV manufacturers, elevated prices of these crucial metals may entirely counteract these benefits.

This underscores the importance of inspecting the prices of these metals and exploring for potential statistical significance when predicting the Cost of Goods Sold (COGS) for Tesla, since these metals are directly attributed to manufacturing EVs and fall into that category in their financial statements.

2.5 Retail Investors

According to the Financial Industry Regulatory Authority (FINRA), a retail investor is defined as "any person other than an institutional investor, regardless of whether the person has an account with a member" (FINRA, 2024: (a)(6)). Retail investors (also known as “individual investors”) are non-professional agents that buy and sell securities for their personal account, typically through traditional or online brokerage firms, and manage only their own funds.

The amount of assets held by retail investors depends on their level of education, with investors without a college degree holding a median of $28,332, those with an undergraduate degree - $73,044 and those who graduated their master’s holding $148,399 (Broadridge Financial Solutions, 2024). The study also delineates a clear trend of the rising popularity of retail investing with the collective share of online assets reaching 36% in 2023. This is mainly driven by millennials and gen Z entering the filed. The report also finds that more younger people in general adopt self-driven investment approaches and enter the filed to educate themselves about capital markets and investing. That young tech-savvy group retail investors interested in discovering new data-driven investment approaches is mainly who this study is directed towards.

3. Methodology

This study aims to explore the underlying relationships and correlations mainly using quantitative methodologies. Specifically, the paper attempts to identify the potentially significant factors that can serve as valuable AD by conceptually explaining the underlying forces and quantitatively proving these correlations by incorporating them into statistical models, testing their significance and comparing with benchmark models without the non-financial data inspected. Secondary literature and datasets have been used for all the research purposes.

The study is exploratory in nature, meaning that it tested for various cause-effect relationships or correlations over time to showcase how a specific data type can paint a different and deeper picture of a company's or industry's performance, and it presents both statistically significant and useful data and that that has been deemed as inconsequential. This approach is justified by the very nature of AD, as any investor who wishes to utilize it should expect to have to analyze data from different perspectives corresponding to different facets of an entity's functioning and explore for various, often deeply hidden, patterns from seemingly unrelated data types. The entangled nature of analyzing non-standard financial data to gain an edge in predicting performance has already been explained in the literature review section of this paper.

3.1 Coding Methodology of the Literature

For the coding methodology and literature review, the researcher relied on inductive coding, as the author developed his understanding of the underlying patterns and possible correlations or causations between data as he analyzed academic sources and data itself. This process was conducted in an iterative manner and the potential for predicting certain values representing companies performance was later tested quantitatively after obtaining the corresponding datasets. That is why the entire coding methodology can be described as iterative, inductive coding.

3.1 Data Collection

The data for this study was collected from a variety of reputable, sources (ex: government websites, financial data providers, industry reports, data gathering platforms, etc.) to ensure reliability for the analysis purposes. The vast majority of specialized AD providers charge astronomical fees for the access to their historical datasets and real time information, which will be explored in the discussion section. That is why this paper relies on a variety of official sources and ensembles the data points together index by timeframe. Sources used include:

- Nasdaq Data Link
- Joint Organization Data Initiative (JODI)
- Statista
- Our World in Data
- U.S. Energy Information Administration
- Google Trends
- Yahoo Finance
- Macrotrends
- Altindex

The majority of the data collected is time series in nature, which allowed for trend identification and correlation testing for data points dispersed in equal intervals over a period of time.

3.2 Data Verification Methods

Certain data sources, such as Nasdaq Data Link and the U.S. Energy Information Administration, provide data of absolute reliability, which does not need to be double-checked. However, other sources, such as Altindex and Yahoo Finance, although still reliable and academically acceptable, were deemed to require additional verification by the researcher.

Such double-checking was conducted using third party data providers, online articles and ChatGPT. Even though these sources are not independently reliable, their collective support for the figures presented by the primary data sources, adds additional legitimacy to data and ensures that these numbers are error-free for further analysis. These additional sources were not included in the bibliography, unless they contradicted the main data providers used in this study.

3.3 Data Manipulation & Cleaning

In order to ensure that datasets are comparable and suitable for analysis and useful for statistical tools, such as regression, the researched index data across using the timeframe as the reference variable. Moreover, many data types, such as Google Trends figures, are provided only on a monthly frequency so to make it comparable with financial data, the author averaged the values for each fiscal quarter by using arithmetic mean.

In case of missing values from the primary sources, the study looked for alternative ones to fill voids in data using methods such as the ones discussed in the Data Verification subsection. In order to additionally ensure that the inserted data points were coherent, the paper looked as seasonality trends and standard deviations to compare these scales with the ones of the newly inputted value.

3.4 Data Analysis & Visualization

The main objective of the majority of the models tested was to predict a continuous value related to a company's or industry's performance for the next fiscal quarter or year. That is the reason why the researcher manipulated the data by date as an index, so that all entries were aligned in time and suitable for regression or other model analysis. The purpose of these analyses, however, was not to arrive at more accurate predictions than the market expectations before earnings reports or expectations regarding the stock price, but to quantitatively demonstrate that the models perform better by incorporating AD as independent variables (IVs) to predict some value pertaining to companies' performance.

For the individual companies, these included the stock price, revenue, COGS and for the industry-wide trend, represented by an ETF performance, the researcher aimed at discovering the price of an ETF.

The statistical tools used mainly involved multivariable linear regression and correlation analysis in Excel. For requesting, merging and manipulating data the researcher used Python in Jupyter Notebook. In order to compare the additional value of incorporating non-standard data, the study compared the performance of the models, which only used typical financial metrics and benchmarked them with the ones that included AD. The metrics analyzed for this comparison include, but are not limited to, Root Mean Squared Error (RMSE), Adjusted R^2, standard error and p-value of the IVs.

4. Results

The following section will present the findings of the exploratory process, comment on the specific regression outputs and the specific IVs included in the regression to predict the corresponding dependent variables.

The industry-wide shaping trends section will focus on the renewable energy industry, specifically on the ICLN ETF, and for the company-specific factors the paper will explore Tesla, Inc. (TSLA) for the reasons discussed in the Literature Review section.

4.1 Renewable Energy Industry-Wide Trends

The analysis below will only entail data from the ICLN ETF, as other major significant green energy following ETFs display very similar behavior over time, both in regard to their price and dividend yield, as presented in Fig. 1 and Fig. 2, where three biggest renewable energy ETFs by AUM and their price development are illustrated. Apart from ICLN in the graphs included are also: First Trust NASDAQ Clean Edge Green Energy Index Fund (QCLN), Invesco Solar ETF (TAN) and SPDR Kensho Clean Power ETF (CNRG), as these are the largest green energy ETFs.

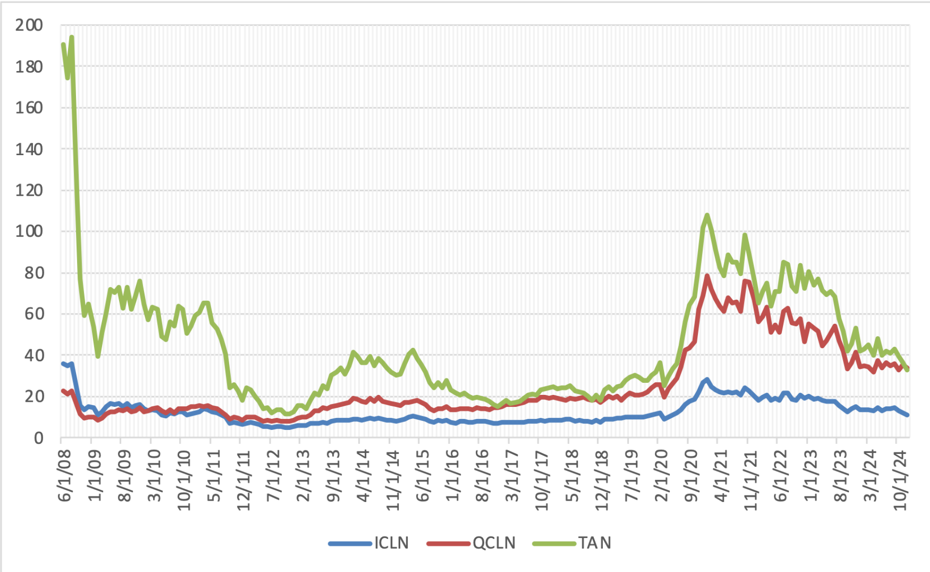


Figure 1: Green Energy ETFs Closing Prices (from 2008/06 to 2024/12 with monthly frequency)

Source: Yahoo Finance, Macrotrends

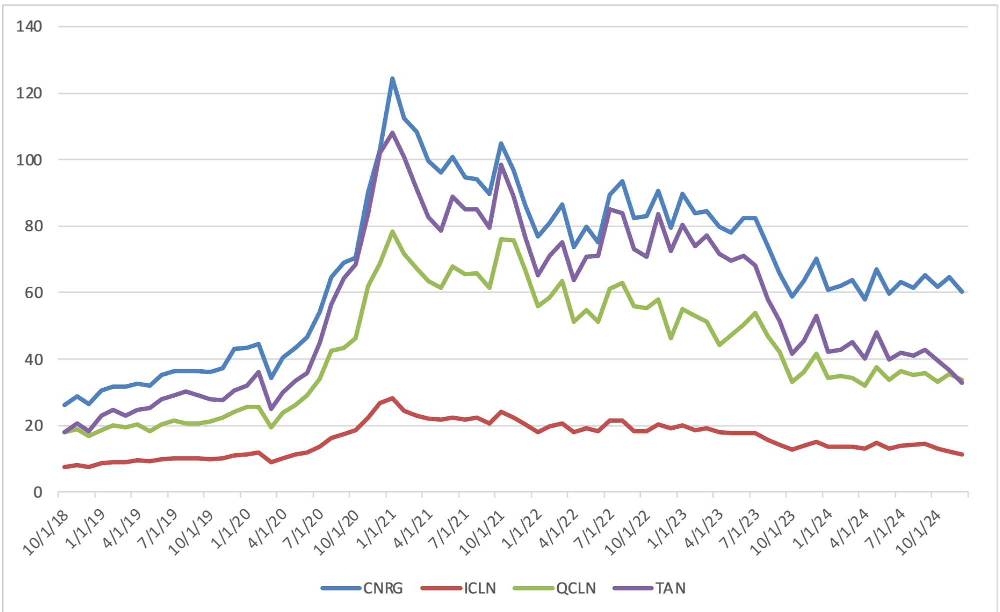


Figure 2: Green Energy ETFs Closing Prices (from 2018/10 to 2024/12 with monthly frequency)

Source: Yahoo Finance, Macrotrends

The graphs are included mostly for visualization purposes and to present the price development and variation in price of these financial products. It can be observed that all the above-mentioned ETFs move in a relatively commensurate fashion.

The AD included in this section will mostly relate to substitutes of renewable energy, such as oil prices and volume traded, imports into the United States and other factors described in the Literature Review section. The data has been collected from 2008 to 2024 with monthly frequency or daily and then averaged or summer per month where appropriate.

In order to be able to quantitatively compare the results of adding AD to multivariable linear regression to predict the closing price of ICLN next month, one should identify a point of reference to which the newly obtained results can be compared to. That point of reference will be a model, which uses exclusively conservative financial metrics as IVs to predict the closing price of ICLN. For this purpose, the researcher selected the: closing ICLN price, volume traded, total dividends distributed, dividend yield and total monthly net asset value (NAV) return, as DeFusco, Ivanov, & Karels (2011) proved that metrics such as NAV and trading volume can be used to predict short-term ETF behavior.

The study will explore 3 models:

Models Tested	*1st model*	*2nd model*	*3rd model*
Brief Description	*Conservative model for comparison purposes*	*1st model + including Brent crude oil price and volume traded*	*2nd model + including oil product inflows*
Regression Degrees of Freedom (df)	*5*	*7*	*9*
Features Included:	• *Close ICLN price* • *Volume ICLN traded* • *Total distribution*	• *Close ICLN price* • *Volume ICLN traded* • *Total distribution*	• *Close ICLN price* • *Volume ICLN traded* • *Total distribution* • *Dividend yield* • *Monthly total*

	• *Dividend yield* • *Monthly total (nav) return*	• *Dividend yield* • *Monthly total (nav) return* • *Brent crude oil close price* • *Brent crude oil volume traded*	*(nav) return* • *Brent crude oil close price* • *Brent crude oil volume traded* • *Oil product imports into the US* • *Primary oil product inflows into the us*

Table 1: Different models for predicting ICLN closing price
Source: Data compiled by author

Below presented are the regression output statistics for all models, along with the feature importance levels:

Regression Statistics	*1st model*	*2nd model*	*3rd model*
Multiple R	0.95769	0.96226	0.97435
R Square	0.91717	0.92594	0.94937
Adjusted R Square	0.91502	0.92322	0.94686
Standard Error	1.62379	1.54345	1.22734
RMSE	1.59912	1.51211	1.19495
Regression Significance F	2.58E-102	2.5824E-104	7.142E-113
Observations	199	199	192

Table 2: Different models regression outputs for predicting ICLN closing price
Source: Data compiled by author, Nasdaq Data Link, Yahoo Finance, Macrotrends

The regression statistics presented above demonstrate that both models outperformed the benchmark model, both in terms of explained variance, denoted by the adjusted R^2, standard error, and in terms of the RMSE. While the second model performs only marginally better, compared with the benchmark the third, most comprehensive, regression model reduces the RMSE by 34% compared with the first model.

The table below presents the regression coefficients and their respective statistical significance.

1st model	*Coefficients*	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*4.1771*	*0.5677*	*0.00005*	*3.0574*	*5.2968*
Close ICLN Price	*0.7229*	*0.0301*	*0.00008*	*0.6635*	*0.7824*
Volume ICLN traded	*9.892E-09*	*2.725E-09*	*3.62E-04*	*4.517E-09*	*1E-08*
Total Distribution / Div	*8.9846*	*4.7914*	*0.0623*	*-0.4657*	*18.4348*
Dividend Yield	*-180.9319*	*47.3451*	*0.0002*	*-274.3121*	*-87.55*
Monthly Total (NAV) Return	*-0.0041*	*0.0132*	*0.7569*	*-0.0301*	*0.0219*
2nd model	*Coefficients*	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*7.5602*	*0.8972*	*0.00003*	*5.7904*	*9.3300*
Close ICLN Price	*0.6820*	*0.0311*	*0.00009*	*0.6206*	*0.7435*
Volume ICLN traded	*1.812E-08*	*3.249E-09*	*8.25E-08*	*1.171E-08*	*2.45E-08*
Total Distribution / Div	*7.6639*	*4.5628*	*0.0947*	*-1.3361*	*16.663*
Dividend Yield	*-175.022*	*45.236*	*0.0001*	*-264.2492*	*-85.79*
Monthly Total (NAV) Return	*-0.0028*	*0.0125*	*0.8256*	*-0.0275*	*0.0220*
Brent Crude Oil Close Price	*-0.0264*	*0.0059*	*0.00002*	*-0.0382*	*-0.014*
Brent Crude Oil Volume Traded	*-2.199E-06*	*5.391E-07*	*6.61E-05*	*-3.26E-06*	*-1E-06*
3rd model	*Coefficients*	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*6.0388*	*1.5052*	*0.0001*	*3.0689*	*9.0087*
Close ICLN Price	*0.8734*	*0.0420*	*0.00008*	*0.7906*	*0.9561*
Volume ICLN traded	*0.0000*	*0.0000*	*0.3582*	*0.0000*	*0.0000*
Total Distribution / Div	*-3.9215*	*3.9738*	*0.3250*	*-11.7622*	*3.9193*
Dividend Yield	*-12.2715*	*41.1691*	*0.7660*	*-93.5016*	*68.958*

Monthly Total (NAV) Return	*-0.0065*	*0.0113*	*0.5632*	*-0.0288*	*0.0157*
Brent Crude Oil Close Price	*-0.0189*	*0.0052*	*0.0004*	*-0.0291*	*-0.008*
Brent Crude Oil Volume Traded	*0.0000*	*0.0000*	*0.0002*	*0.0000*	*0.0000*
Oil Product Imports into the US	*0.0003*	*0.0001*	*0.0092*	*0.0001*	*0.0006*
Primary Oil Product Inflows into the US	*-0.0001*	*0.0000*	*0.0083*	*-0.0002*	*0.0000*

Table 3: Different models regression coefficients table for predicting ICLN closing price
Source: Data compiled by author, Nasdaq Data Link, Yahoo Finance, Macrotrends

Analysis suggests that when adding the Brent Crude oil price and volume traded, their p-value is significantly below 0.05 and thus contributing to the model. The p-values of total monthly net asset value (NAV) returns is statistically insignificant, while the total dividend distribution becomes marginally significant. As for the third model: the "Oil Product Imports into the US" and "Primary Oil Product Inflows into the US" are both statistically significant with the p-value significantly below 0.05 and the 95% confidence interval not including 0. Furthermore, the correlation between these two variables is only 0.19, as presented in Table 6. The finding that the two IVs are weakly correlated additionally reinforces the notion that adding them to the regression improves its accuracy and explained variance in the dependent variable – the ICLN closing price next month.

These conclusions quantitatively confirm the findings of both, Omri & Nguyen (2020) and Kumar, Managi, & Matsuda (2012) that oil prices, volume traded, and consumption of oil, measured through various factors, directly influence the adoption of sustainable energy solutions, which is later reflected in green energy companies bottom line, their stock price, and by extension – the related ETF price.

In order to explore how adding additional features might additionally enhance the predictive capability of the model, the researcher added the following variables:

- Total Energy (related) CO_2 emissions (in the United States)
- Renewable Energy Consumption (in the US)
- Rotary Oil Rigs in Operation (in the US)
- Average Electricity Price to Ultimate Consumers (in the US)

Surprisingly, this resulted in marginally decreased model performance, as presented in Table 4:

Regression Statistics	
Multiple R	*0.974402426*
R Square	*0.949460087*
Adjusted R Square	*0.94576897*
Standard Error	*1.23990478*
RMSE	*1.193844384*
Regression Significance F	*7.0593E-108*
Observations	*192*

Table 4: Fourth model regression output for predicting ICLN closing price

Sources: Data compiled by author, Nasdaq Data Link, Yahoo Finance, Macrotrends, Our World in Data, U.S. Energy Information Administration

The level of explained variance declined infinitesimally from the third model explaining 94.69% of the variance to 94.57% and the RMSE also decreased by a extremely small amount. The analysis of the coefficients explains this result.

	Coefficients	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*5.689179*	*2.495169*	*0.02378*	*0.765260*	*10.613098*
Close ICLN Price	*0.873105*	*0.046155*	*0.00000*	*0.782024*	*0.964187*
Volume ICLN traded	*0.0000008*	*0.0000004*	*0.45769*	*-4.52E-09*	*9.99E-09*
Total Distribution / Div	*-4.492729*	*4.403837*	*0.30902*	*-13.183176*	*4.197719*
Dividend Yield	*-7.655117*	*44.856258*	*0.86468*	*-96.173599*	*80.863365*
Monthly Total (NAV) Return	*-0.006027*	*0.011580*	*0.60347*	*-0.028879*	*0.016826*

Brent Crude Oil Close Price	*-0.019124*	*0.007505*	*0.011672*	*-0.033933*	*-0.004314*
Brent Crude Oil Volume Traded	*-0.000002*	*0.000001*	*0.00105*	*-0.000004*	*-0.000001*
Oil Product Imports into the US	*0.000340*	*0.000130*	*0.00993*	*0.000083*	*0.000598*
Primary Oil Product Inflows into the US	*-0.000097*	*0.000040*	*0.01556*	*-0.000176*	*-0.000019*
Total Energy CO2 Emissions	*-0.000820*	*0.002878*	*0.77591*	*-0.006500*	*0.004859*
Renewable Energy Consumption	*-0.007136*	*0.088233*	*0.93563*	*-0.181252*	*0.166980*
Rotary Oil Rigs in Operation	*-0.000034*	*0.000468*	*0.94264*	*-0.000957*	*0.000890*
Average Electricity Price to Ultimate Consumers	*0.072062*	*0.244144*	*0.76821*	*-0.409728*	*0.553851*

Table 5: Fourth model regression coefficients table for predicting ICLN closing price
Sources: Data compiled by author, Nasdaq Data Link, Yahoo Finance, Macrotrends, Our World in Data, U.S. Energy Information Administration

The p-values of the of the newly added features are significantly higher than 0.05 and the 95% confidence intervals include 0, which makes them statistically insignificant. This finding is surprising, as the correlations between these variables are all below 0.52, with the exception of the correlation between “Renewable Energy Consumption” and “Average Electricity Price to Ultimate Consumers” and “Rotary Oil Rigs in Operation”, which stand respectively at 0.895 and -0.702. These findings are presented in the correlation matrix below, which only includes the non-standard data included in the regression as IVs, but it sources from the table including all variables listed in the coefficients table.

	Brent Crude Oil Close Price	Brent Crude Oil Volume Traded	Oil Product Imports into the US	Primary Oil Product Inflows into the US	Total Energy CO2 Emissions	Renewable Energy Consumption	Rotary Oil Rigs in Operation	Average Electricity Price to Ultimate Consumers	Next Month's Closing Price
Brent Crude Oil Close Price	1								
Brent Crude Oil Volume Traded	-0.569069742	1							
Oil Product Imports into the US	-0.208054175	0.047589991	1						
Primary Oil Product Inflows into the US	0.239307706	-0.698285628	0.19338315	1					
Total Energy CO2 Emissions	0.19975358	-0.404718935	0.11011586	0.52826567	1				
Renewable Energy Consumption	-0.177077833	0.72858253	-0.0108293	-0.7171537	-0.4858533	1			
Rotary Oil Rigs in Operation	0.646820078	-0.748501404	-0.21527	0.62993655	0.47687576	-0.70189089	1		
Average Electricity Price to Ultimate Consumers	0.013149597	0.515337388	-0.0676968	-0.6031988	-0.3965926	0.895826656	-0.5172938	1	
Next Month's Closing Price	-0.03989571	-0.00546435	0.43539709	-0.2983051	-0.1729161	0.199529631	-0.400566446	0.255512426	1

Table 6: Correlation matrix for selected AD inputs

Source: Data compiled by author, Nasdaq Data Link, Yahoo Finance, U.S. Energy Information Administration

However, the correlations between other IVs added, such as “Total Energy CO2 emission” and “Average Electricity Price to Ultimate Consumers” is rather low, at -0.39 and the correlations between the two abovementioned variables and the dependent variables are also low. However, these variables are statistically insignificant, even despite for the logical and factual evidence presented in the Literature Review. This might be explained by the limitation of linear regression in capturing complex, non-linear patterns, which are often the characteristic of the data types explored in this study.

4.2 Company-Specific Insight

For the company-specific chapter, the study focused specifically on Tesla, Inc. (TSLA) attempting to predict the company’s revenue, COGS and stock price for different time horizons. The reason why this company has been selected was already explored in the Literature Review section. That data has been collected from 2009 to 2024 with annual frequency for the revenue and COGS regression and for the stock price assessment, from August 2022 to March 2025 with daily frequency and then averaged or summed, where appropriate, per month in order to identify slightly longer-term trends and smoother out statistical noise.

4.2.1 Tesla's Stock Price

The first subsection will attempt to showcase the additive value of adding AD to predict the closing price next month of Tesla's stock. The data for this purpose was collected from August 2022 to March 2025 with daily frequency and then averaged or summed, where suitable, for every month. For the purpose of predicting the stock price for the next month, the AD was obtained from Altindex and included information regarding Tesla, Inc., such as: number of Twitter followers, Reddit mentions and sentiment, app download statistics for iOS and Android along with their respective rating, YouTube subscribers, number of comments on Glassdoor with the average score, etc. The logical rationale behind incorporating this information into a predictive model, such as multivariable linear regression is that they reflect the interest in a company along with the sentiment at any point in time that accompanies it. The models, which included these data points were compared with one that only relied on technical analysis related information, such as: open, high, low, close prices, volume traded, 20 and 50 period exponential moving averages.

4 multivariable linear regression models were compared against each other:

Model Tested	*model*	*2nd model*	*3rd model*	*4th model*
Regression Degrees of Freedom	*7*	*13*	*16*	*10*

Features Included	*Open TSLA* *High TSLA* *Low TSLA* *Close TSLA* *Volume TSLA* *EMA_20_TSLA* *EMA_50_TSLA*	*High TSLA* *Low TSLA* *Close TSLA* *Volume TSLA* *EMA_20 TSLA* *EMA_50_TSLA* *Twitter followers* *YouTube subscribers* *iOS downloads* *total iOS reviews* *iOS rating* *reddit mentions* *reddit sentiment*	*Open TSLA* *High TSLA* *Low TSLA* *Close TSLA* *Volume TSLA* *EMA_20 TSLA* *EMA_50 TSLA* *Android downloads* *total android reviews* *android rating* *app rating* *app downloads* *sentiment* *news mentions* *reddit mentions* *reddit sentiment*	*Low TSLA* *Close TSLA* *EMA_20_TSLA* *EMA_50_TSLA* *Reddit subscribers* *Instagram followers* *iOS rating* *patents* *app rating* *reddit sentiment*

Table 7: Different models description for predicting TSLA closing price
Source: Data compiled by author, Altindex, Yahoo Finance

The first model serves as a benchmark as it only includes the technical analysis related data, while the second one also includes AD, such as the number of Twitter and YouTube followers, Reddit sentiment, iOS downloads, reviews and ratings. The third one includes 16 IVs and the researcher experimented with the types of features included, but regardless of the selection the models yielded similar results in terms of accuracy. The fourth model was selected based on the lowest p-values across all models and included only 10 IVs.

The results of the regression are presented below in the Table comparing different models:

Regression Statistics	*1st model*	*2nd model*	*3rd model*	*4th model*
Multiple R	*0.86913204*	*0.93927115*	*0.95923448*	*0.9635141*
R Square	*0.7553905*	*0.8822303*	*0.92013079*	*0.92835941*

Adjusted R Square	*0.71557035*	*0.79717441*	*0.83493697*	*0.89424485*
Standard Error	*33.1667766*	*29.4768789*	*26.5916343*	*21.2848643*
RMSE	*30.45*	*22.1076592*	*18.21*	*17.2427108*
Regression Significance F	*2.8662E-11*	*7.5563E-06*	*1.6659E-05*	*8.53E-10*
Observations	*32*	*32*	*32*	*32*

Table 8 : Different models regression output for predicting TSLA closing price
Source: Data compiled by author, Altindex, Yahoo Finance

The results above prove that the 4th model, selected of only the most importance IV based on their respectively low p-values) performs best both in terms of accuracy, reflected in the Standard Error and RMSE, and in terms of explaining the highest amount of variance in the dependent variable (stock price).However, the biggest change can be observed between the first and second model in terms of reduced RMSE and improved adjusted R^2.

The reason why the number of observations for the models tested isn't higher like for the ICLN analysis is that the AD provided by Altindex date back only to August 2008 and the researcher had to average or sum the results for every month in order to make the results cogent. If the study were to consider data with daily frequency the variation and changes in the IVs would merely reflect statistical noise and any possible correlations or causations would have been coincidental. Averaged or summed per month, however, they reflect longer term trends and are more suitable for regression analysis.

4.2.2 Revenue

This subsection will attempt to predict the annual revenue of Tesla next year. The company discloses these figures by releasing their 10-K (annually) and 10-Q (quarterly) earnings reports. The AD added to the model here include:

- US Annual CO2 emissions
- California Estimated CO2 emissions
- "Tesla" Google Trends interest level
- EVSE Ports in California

- Station Locations in California

The study assumes that EVSE ports charging stations and CO2 emissions in California are most relevant for TSLA, as that state is one of the most significant profit centers for that company (Tesla, Inc., 2024).

For comparison purposes, the study included some traditional financial metrics commonly inspected by analysts. The regression result of that below is presented below in Table 9:

Regression Statistics	
Multiple R	*0.99481941*
R Square	*0.98966565*
Adjusted R Square	*0.98553192*
Standard Error	*4317.92549*
RMSE	*3,525.57*
Regression Significance F	*7.0115E-10*
Observations	*15*

Table 9 : Benchmark model regression output for predicting TSLA revenue
Source: Data compiled by author, Yahoo Finance, Statista, TSLA Annual Reports

The model explains 98.55% of the variance in the dependent variable with the RMSE being relatively high: $3.5 B.

	Coefficients	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*-988.295*	*3387.195*	*0.776422*	*-8535.437*	*6558.8*
Number of Employees	*0.99150*	*0.179469*	*0.00025*	*0.59162*	*1.391*
Change in Employee headcount	*2134.08*	*3956.66*	*0.601*	*-6681.90*	*10950*
R&D Spending (M)	*-11.931*	*5.7970*	*0.06658*	*-24.8482*	*0.9848*
Revenue (M)	*0.0914*	*0.21823*	*0.68416*	*-0.39485*	*0.5776*

Table 10 : Benchmark model regression coefficients table for predicting TSLA revenue
Source: Data compiled by author, Yahoo Finance, Statista, TSLA Annual Reports

These findings act as a point of reference. It is worth noting that even though, certain parameter, such as "Revenue (M)" have a p-value higher than 0.05, the model performs significantly worse without them.

The regression output below represents the model including the above-mentioned non-standard data.

Regression Statistics	
Multiple R	*0.99944953*
R Square	*0.99889937*
Adjusted R Square	*0.99691824*
Standard Error	*1,992.82*
RMSE	*1,150.56*
Regression Significance F	*7.4648E-07*
Observations	*15*

Table 11 : Model with AD regression output for predicting TSLA revenue
Source: Data compiled by author, Google trends, Yahoo Finance, Statista

	Coefficients	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*41119.1*	*32439.298*	*0.2607*	*-42268.7*	*124506*
Number of Employees	*0.64755*	*0.225917*	*0.03514*	*0.06681*	*1.22829*
Change in Employee headcount	*4943.46*	*2506.775*	*0.10564*	*-1500.40*	*11387.3*
R&D Spending (M)	*-2.882*	*4.0679*	*0.510*	*-13.33*	*7.57445*
Revenue (M)	*-0.184*	*0.2113*	*0.42348*	*-0.727*	*0.3591*
US Annual CO2 emissions	*1.9E-05*	*6.45E-06*	*0.03018*	*2.7E-06*	*3.5E-05*

Estimated California CO2 emissions	*-4.6E-06*	*1.38E-06*	*0.020*	*-8.2E-06*	*-1.0E-06*
"Tesla" Google Trends	*-517.17*	*265.277*	*0.1087*	*-1199.08*	*164.746*
EVSE Ports in California	*8.1668*	*1.41622*	*0.002*	*4.52*	*11.807*
Station locations in California	*-4.043*	*2.82125*	*0.2112*	*-11.29*	*3.20908*

Table 12 : Model with AD regression coefficients table for predicting TSLA revenue
Source: Data compiled by author, Google trends, Yahoo Finance, Statista

The regression output analysis for both models suggests that even though adding the additional features improved the R Squared only marginally (from 98.55% to 99.69%), it significantly reduces the RMSE of the models from 3,525.57 to 1,150.56. Even though this model's mean error in predicting the revenue is still significantly too high to effectively leverage against the market's expectations for a given year, as these forecasts are usually very accurate, it illustrates the additive utility of the new features to enhance the accuracy of the model. It also means that investors can obtain more granular and accurate data, embed them into their prediction models and effectively leverage in their SI process.

4.2.3 Cost of Goods Sold (COGS)

In order to attempt to predict the COGS for Tesla, the researcher identified the most significant components of manufacturing EVs for that company (Tesla, Inc., 2024), as presented and discussed in the Literature Review. No comparison model will be analyzed here, as there are no standard financial indicators and metrics that can be used to gauge the efficiency with which the company will manufacture its products and their COGS the following year. That is why this regression includes only various mineral prices and industrial retail electricity prices as IVs with the COGS from previous year and gross profit margin for reference.

Regression Statistics	
Multiple R	*0.99191142*

R Square	*0.98388826*
Adjusted R Square	*0.96509123*
Standard Error	*5.42193885*
RMSE	*3.53*
Regression Significance F	*5.8577E-05*
Observations	*14*

Table 13: Regression output for predicting TSLA COGS

Source: Data compiled by author, Yahoo Finance, Statista

	Coefficients	*Standard Error*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	*231.214*	*95.1451*	*0.0511*	*-1.5974*	*464.026*
Lithium carbonate	*0.00092*	*0.000402*	*0.060*	*-5.61E-05*	*0.00191*
Cobalt	*-0.732*	*0.43295*	*0.141*	*-1.79202*	*0.3267*
Nickel	*-0.0015*	*0.00070*	*0.0746*	*-0.00322*	*0.00020*
Aluminum Ingots	*0.7055*	*0.2954*	*0.054*	*-0.0175*	*1.428*
Industrial Retail Electricity	*-41.53*	*15.954*	*0.040*	*-80.574*	*-2.495*
COGS (in B $)	*1.085*	*0.1490*	*0.00034*	*0.7211*	*1.450*
Gross Profit Margin	*66.664*	*33.414*	*0.0930*	*-15.097*	*148.426*

Table 14 : Regression coefficients table for predicting TSLA COGS

Source: Data compiled by author, Yahoo Finance, Statista

The regression summary output indicates that 98% of variance in the dependent variable (COGS of TSLA) is explained by the IVs and almost all of the features are statistically significant, with the exception of cobalt and the gross profit margin. The prices of lithium carbonate and nickel are marginally significant. These results align with the logical notion that by assessing the prices of the most financially significant components of manufacturing Tesla cars, one can predict the total cost manufacturing these vehicles and by extension – COGS.

Being able to predict the changes in COGS can be extremely significant for investors, as it has a tremendous impact on the company’s net income and Earnings per Share

(EPS), which influences that gapping up or down of Tesla's stock price after the company releases its financial statements annually or quarterly.

4.3 Summary and Implications of Findings for Systematic Investing

Even though certain regression output suggest models which explain almost all variation on the dependent variable, it is worth noting that the standard error and RMSE levels are too high to leverage these models for actual decision making and thus using them in SI for real investing. They serve for the purpose of illustrating how these non-standard data types can be used for more accurate predictions and to explain a larger degree of variance in the target variable, which is immensely important. It must be noted that the analysts expectation, available to all market participants prior to any 10-K or 10-Q based on which the company's performance is compared and which results in the gaping up or down of the stock, are more accurate than the models presented above and the relatively high standard errors render them not appropriate to compare on that scale.

However, this research proves that it is possible to consider AD that improve the performance of predictive models, by incorporating AD into the model and arriving at more accurate predictions. Furthermore, the p-values and the R^2 scores signify that investors can include these non-standard data types for whatever time horizon they are focused on, develop the models and use them by feeding the latest data. This also means that (retail) investors can obtain significantly more accurate predictions, which they can later use for the portfolio selection process or profiting from the gapping up or down right after a company releases their financial statements by having obtained more accurate picture of the companies performance for any period of time.

The SI process is heavily data driven and involves several steps, whether it is top-down or bottom-up or even if it pertains to high-speed algorithmic trading (Harvey, 2021). However, in longer term investing, which is the focus of the paper, in the top-down approach, these steps can involve: analyzing the nation's economy as a whole and determining its growth potential, sector-wide analysis and then industry selection. At each step an investor decide how much of their portfolio to allocate to risk free and risky assets, the exposure to different sectors and finally: to different companies (Crescenzi, 2008). The research presented above proves that AD can be included at every step of the way, whenever continuous data and quantitative assessment using statistical tools is involved, to arrive at more accurate predictions by explaining a larger portion of the variance in the dependent variable the investor attempts to predict. This insight is later included in the deciding how much of the portfolio to allocate to specific asset types, industries and enterprises.

5. Discussion

This section aims at identifying the limitations to accessing and leveraging AD, especially for retail investors, and the impact it has on the predictive models developed and thus on the portfolio selection process and expected returns. Moreover, in this section the researcher will explain the ethical considerations and legal restrictions that must be taken into account when processing AD for the purpose of analyzing securities and industry trends. Finally, the paper points to further research that can be done in this field to further understand the implications of non-standard data, especially considering the exploratory nature of this study and how other scholars and researchers can delve deeper into the specificities of AD in specific industries or for specific enterprises to further prove its predictive capabilities.

5.1 Limitations

There are two considerable limitations to accessing AD for retails investors: cost barriers and processing complexity. The former arises from the fact that AD has historically been reserved for institutional investors, and even though the data providers offer their services to anyone, they are targeted to clients whose budget constraints are not as significant as those of the majority of retail investors.

The processing complexity barrier relates to the very nature of AD, which this thesis explores, as it is significantly more difficult to analyze than typical financial statements, which have clear guidelines that investors can educate themselves about and follow in a relatively straightforward manner. In comparison, non-standard data can vary significantly and pertain to seemingly unrelated data points (Wang & Ngai, 2024). Furthermore, AD is very often unstructured and requires manipulation, cleaning and verification to make it useful for further analysis. These processes are often beyond the scope of even the more sophisticated retail investors and they require a lot of time, resources and sometimes even computing power for more voluminous datasets. They also do not guarantee significant added utility for the investment selection process, as discovering AD is often exploratory in nature. This difficulty multiplies for the purpose of SI, which required encoding the instruction process to make it repeatable across securities.

5.1.1 Accessing Alternative Data

The table below presents selected platforms and data providers, the information they provide, how this data can be used to gain insight into companies' performance, the prices they charge and how the researcher discovered their pricing or subscription models, as most of them don't directly list that information publicly.

This is the only part of the study for which the researcher partially obtained primary data by directly reaching out to companies' representatives and inquiring about the possibility to use their services for the purpose of this thesis.

Platform	**Type of Data provided**	**Pricing Model**	**Utility for Investors**	**Way of Discovering**
"Fastmarkets"	Mineral Prices (e.g. battery minerals, such as the prices of lithium hydroxide over time)	Annual models starting at £7,000/year	Possibility to inspect how the prices of elements and components of products manufactured by various companies might affect their	Reached out to a company representative (pricing models not

			COGS, profit margins and by extension EPS	disclosed online)
"SensorTower"	App statistics (e.g. downloads, time spent, uninstalls, mentions etc. with available historic data for comparison)	Annual models starting from $40,000 / year	Insight into the level of interest and sentiment towards companies' products or services by examining online behavior of (potential) customers, change in trends and interest. This can be used to predict revenue levels, net income and stock performance.	Reached out to a company represen tative (pricing models not disclosed online)
"Financial Times (FT) Professional – Developer Programme"	Financial (markets) related articles	Annual models starting from $17,000 / year	Possibility to collect hundreds of thousands or even millions of articles, papers, and financial data related studies, perform sentiment analysis and test for correlations and causations in order to be able to use real time articles for improved decision making.	Reached out to a company represen tative (pricing models not disclosed online)
"Planet Labs"	Satellite Imagery on demand or historical datasets	Cost per one image ranging from $100 - $500	Possibility to analyze (using ML methods) satellite imagery of companies	Available only after creating an account

	(optionally with machine learning enabled object detection or other analysis features)	Cost for a 3.7m resolution image: $\$0.1/km^2$ Minimum order thresholds for certain services (like PlanetScope Monitoring): $4,500	manufacturing facilities, parking lots, offices, or inspect how terrain is changing to adapt investment strategy and drive longer term trends.	and selecting a very specific image or service

Table 15 : Alternative data providers and platforms overview
Sources: Data compiled by author through active research and reaching out to the above-mentioned companies' representatives

The platforms listed above are sophisticated in nature and provide reliable and highly granular data, which is tremendously important for investors, especially considering different investment horizons of different retail investors. However, it also illustrates the financial hindrances in leveraging the sophisticated services of these data providers as the cost to accessing it is significantly too high for the vast majority of retail investors.

Considering the relatively limited scale at which retail investors operate, and that even those with a graduate degree held a median of $148,399 in 2024 in the United States (Broadridge Financial Solutions, 2024), the fees charged by AD providers make it untenable and unpractical for retail investors to engage with. This explains the widespread adoption of AD among institutional investors, who can easily afford such sophisticated solutions without compromise.

However, there are also numerous providers which provide information that might be considered AD for a significantly lower price. The examples include the platforms utilized in this paper, such as Altindex, that provide daily information related to web traffic, sentiment, job posts and social media followers all in one service, which charges only $100 per month (Altindex, 2025). Even though this type of data is not as granular and reliable as the data sources presented above, it should suit the needs of the majority of retail investors and allow them to enhance their predictive models. Additional platforms include Nasdaq Data Link, offering unique datasets ranging grom $49 to $500, or Similarweb, which provides shopping trends, more granular web traffic and other app intelligence solutions and datasets starting from $199 per month (Similarweb, 2025). These prices are within the range of the majority of retail investors (Broadridge Financial Solutions, 2024), and the data provided is of the perfect level of complexity for them to be able to analyze.

5.1.2 Impact on Predictive Models

This limitation to data accessibility is directly translated into the performance and accuracy of statistical predictive models, from multivariable linear regression to complicated deep neural networks, as they are only as good as the data they have been trained on. This means that if data of limited accuracy or granularity has been used to train them, they cannot produce predictions of superior accuracy.

That is why it is important for retail investors, especially, to be able to access only the data they require and are able to analyze, and for them to be able to assess whether that data adds predictive value and is statistically significant, considering the level of complexity of the models they want to develop and use and the future. That is especially important for the data driven nature of SI, which is very prone to biased or poorly suited data, especially when implemented in an iterative manner when the errors and skewed predictions of an erroneous model accumulate quickly, depending on how often the investors rebalances the portfolio.

5.2 Ethical & Legal Considerations

There are numerous legal and ethical considerations that need to be taken into account when collecting, storing, processing and analyzing AD. Web scarping is one of the most common ways of collecting textual content, numerical data, images, links and any other publicly available data type. It is relatively simple as even a beginner programmer can write a simple web scraping program. However, there are numerous regulation in many countries preventing certain aspects of that practice. For example, in the European Union (EU), General Data Protection Regulation (GDPR) prohibits web scraping of personal data (such as names, emails, IP addresses) without user's consent, even if the data is publicly available. Moreover, in the EU the Database Directive (Directive 96/9/EC) from 1996 established legal protection of various databases and prohibits extraction of substantial parts of data without proper authorization (European Parliament & Council, 1996). This also directly applies to web scraping, as this activity results in obtaining and analyzing the very data that these datasets store for the purpose of allowing users to access it through proper channels.

In the United States, the Computer Fraud and Abuse Act (CFFA) from 1986 restricts unauthorized access to computer systems, including databases and data available behind logins or payment systems, which means that unauthorized scraping of such data may be prosecuted. Additionally, may platforms prohibit scraping in their Terms of Service (TOS), such as LinkedIn in section 8, subsection 2 where the agreement stipulates that users agree not to develop any programs that would scrape or copy LinkedIn services, profiles and other data (LinkedIn, 2025).

Moreover, collecting AD raises privacy concerns and those of ethical nature, as the information analyzed was not meant for such inspection and many would argue that even if one were to find a loophole and collect such data through legal ways, it is unethical towards the profile creators and people who chose to share information, knowledge or insight to be collected in such as massive way without any regard to their individual contribution. In the future, it is likely that even more regulation will be implemented, especially in the EU, to additionally shield the internet users from their data, personal or not, being web scraped and collected. Investors should take these reservations into consideration when leveraging AD for their predictions and decision making.

5.3 Further Research

This study is exploratory in nature and addressed the gap in research by showcasing how seemingly unrelated in nature, publicly available, datasets can be used to make better predictions regarding companies' performance and facilitate the systematic investment analysis process, which leads to the construction of a portfolio, especially for retail investors with relatively limited resources, as opposed to institutional investors.

However, it paints a rather general picture of how AD can be leveraged and, while it supports its claims with statistical tools and quantitatively proven relationships, it also acts as a framework or point of reference for future scholars who can explore the field of non-standard data further. This means that there is relatively large space if of unexplored research, particularly regarding the specificities of individual industries or specific types of companies.

Other researchers can delve way deeper into the relationships between different AD and how it can used to predict specific companies performance with way greater granularity and accuracy than has been described and illustrated in this thesis. For example, scholars can attempt to feed AD to statistical tools or Machine Learning models and compare their predictions with market and analyst expectation, before a company releases their financial statements or month-to-month. If the difference between their assessments and the actual figures is smaller than that between market expectations, such findings would additionally reinforce the conclusions made in this study. Furthermore, researchers can attempt to embed the abovementioned models and back-test them on historical data that the model has not seen before and compare the returns of different strategies with the ones without non-standard data. Finally, for the purpose of this study the researcher focused on longer-term trends, such as time series with monthly or annual frequency, which meant that the data limitation was quite significant and linear regression was the best suited model for this analysis. However, other scholars can access data with higher granularity and apply more sophisticated ML models, such as Long Short-Term Memory (LSTM) or gradient boosting machines (GBMs) and test how various AD enhance these models on a shorter timeframe, for example for the purpose of high-speed algorithmic trading.

6. Conclusion

In conclusion, the research focused on exploring different AD types by incorporating them into multivariable regression models to assess how they can aid in predicting financial metrics, such as stock and ETF prices, revenue levels, and COGS. The study addressed the research gap resulting from the limited academic exploration of how these different non-standard data can help to gain a more accurate picture of the company's performance, especially longer term and for retail investors wishing to improve their portfolio construction and security selection process by employing top-down SI analysis. The research was divided into two main parts: the industry-wide assessment and company-specific factors. The former analyzed the ICLN ETF and attempted to predict its price, while the latter focused on Tesla, Inc. and demonstrated how different AD can help predict the company's stock price, revenue levels and COGS for different time horizons.

The study found that identifying relevant AD and incorporating them into predictive models and statistical tools, like multivariable linear regression improves the model accuracy by reducing the RMSE and increasing the percentage of explained variance in the dependent variable (adjusted R squared) investors try to predict, compared with the conservative models, which did not include these data points. However, the paper also finds that multivariable linear regression fails to capture nonlinear patterns, which are sometimes inherent to AD, even if the academic sources, logical arguments, and correlation analysis suggest they should be relevant.

The thesis focused on longer-term predictions in order to make the results more useful for SI-facilitated portfolio construction and longer term investing, rather than intraday trading, even though AD can be utilized in that area too. The paper explained how these findings can be incorporated into the SI analysis process, especially the top-down approach, which yields higher returns than bottom-up (Zurek & Heinrich, 2021), where investors can include the AD in their predictive models at every step and thus by extension: have more accurate predictions and construct better portfolios.

Moreover, the study found that AD can be collected with varying levels of complexity and it is crucial for retail investors to identify what degree of granularity and accuracy they require and are able to obtain and process to be able to maximize their utility of AD for investment analysis purposes. It is worth noting that retail investor's access to the majority of sophisticated AD providers and platforms is significantly restricted due to the high cost associated with these services. However, the paper also finds that there are certain platforms that retail investors can use, such as: Altindex, Similarweb, and Nasadaq Data Link. All of them provide relatively affordable intelligence data solutions and of a moderate level of complexity for retail investors to be able to process and analyze. This goes in addition to cost-free options, like Google Trends, JODI, and governmental websites.

There are also legal hindrances to collecting AD, such as GDPR, CFFA, and Terms of Service of individual platforms, which explicitly prohibit certain ways of obtaining their data, such as through web scraping. Furthermore, there are ethical considerations to be taken into account when collecting and analyzing non-standard data, as this type of information wasn't originally meant for scrupulous analysis, and investors should keep that in mind when accessing AD. Future researchers should delve more deeply into specific AD and test if they can train more sophisticated models on more granular data, like LSTM or GBMs, with more accurate predictions than market and analyst expectations prior to annual and quarterly earnings reports. Moreover, scholars can back-test various trading and investment strategies, which include non-standard data and compare them with the ones without.

References

Wang, Y., & Ngai, E. W. T. (2024). Alternative data in finance and business: Emerging applications and effects. Financial Innovation, 10, Article 2.
https://jfin-swufe.springeropen.com/articles/10.1186/s40854-024-00652-0

Monk, A., Prins, J., & Rook, D. (2019). Rethinking alternative data in institutional investment. Journal of Financial Data Science, 1(1), 16–31
https://caia.org/sites/default/files/014-031_monk_jfds.pdf

Financial Industry Regulatory Authority. (2024). FINRA Rule 2210: Communications with the Public
https://www.finra.org/rules-guidance/rulebooks/finra-rules/2210

Ekster, G., Kolm, P.N. (2020). Alternative data in investment management: Usage, challenges, and valuation, SSRN
https://smallake.kr/wp-content/uploads/2021/01/SSRN-id3715828.pdf

Frattini, A., Bianchini, I., Garzonio, A., & Mercuri, L. (2022). Financial technical indicator and algorithmic trading strategy based on machine learning and alternative data, MDPI, Risks 10(12), 1–24.
https://www.mdpi.com/2227-9091/10/12/225

Hansen, K. B., & Borch, C. (2022). Alternative data and sentiment analysis: Prospecting non-standard data in machine learning-driven finance. Big Data & Society, 9(1), 1–14.
https://journals.sagepub.com/doi/full/10.1177/20539517211070701

In, S. Y., Rook, D., & Monk, A. (2019). Integrating alternative data (also known as ESG data) in investment decision making. S&P Global, Global Economic Review, 48(3), 237–260.
https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3380835

Harvey, C. R. (2021). Why Is Systematic Investing Important? SSRN https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3785370

State Street Global Advisors. (2023). The Rise in Systematic Credit Investing. https://www.ssga.com/library-content/assets/pdf/global/systematic/2023/the-rise-in-systematic-credit-investing.pdf

Deloitte. (2023). Fueled by better information: Why investment management should embrace alternative data. Deloitte Insights https://www2.deloitte.com/content/dam/insights/articles/us176544_cfs_fsi-predictions_growth-of-alternative-data/DI_FSI-Predictions_Alternative-data%20postpub.pdf

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 25(2), 383–417 https://www.jstor.org/stable/2325486?origin=crossref&seq=1

Omri, A., & Nguyen, D. K. (2020). Do high oil prices obstruct the transition to renewable energy consumption? Sustainability, 12(11), 4689 https://www.mdpi.com/2071-1050/12/11/4689

Kumar, S., Managi, S., & Matsuda, A. (2012). Tail risk connectedness in clean energy and oil financial markets. Energy Economics, 34(1), 168–175 https://www.sciencedirect.com/science/article/abs/pii/S0140988311000727?via%3Dihub

Makki, M., Kaspard, J., Khalil, F., & Mawad, J. L. (2024). Renewable energy consumption determinants: Do they differ between oil-exporting countries and oil-importing ones? Sustainability, 16(17), 7295 https://www.mdpi.com/2071-1050/16/17/7295

OpenStax. (2022). Principles of Economics (3rd ed., Ch. 3) https://openstax.org/books/principles-economics-3e/pages/3-1-demand-supply-and-equilibrium-in-markets-for-goods-and-services

DeFusco, R. A., Ivanov, S. I., & Karels, G. V. (2011). The exchange traded funds' pricing deviation: analysis and forecasts. Journal of Economics and Finance, 35(2), 181–197
https://www.researchgate.net/publication/225157615_The_exchange_traded_funds'_pricing_deviation_Analysis_and_forecasts

Crescenzi, A. (2008). Investing from the top down: A macro approach to capital markets. McGraw-Hill
http://repo.darmajaya.ac.id/4058/1/Anthony%20Crescenzi%20-%20Investing%20from%20the%20top%20down_%20A%20macro%20approach%20to%20captial%20markets-McGraw-Hill%20%282008%29.pdf
Zurek, M., & Heinrich, L. (2021). Bottom-up versus top-down factor investing: An alpha forecasting perspective. Journal of Asset Management, 22(1), 11–29
https://link.springer.com/article/10.1057/s41260-020-00188-9

BlackRock. (2025, April 3). Supplement to the Summary Prospectus, Prospectus and Statement of Additional Information for the iShares Global Clean Energy ETF (ICLN). iShares Trust
https://www.blackrock.com/us/individual/resources/regulatory-documents/stream-document?stream=reg&product=IUS-ICLN-J&shareClass=NA&documentId=2022295%7E2280820%7E1709212%7E2317696%7E2251398%7E2253160%7E1896134%7E1871314&iframeUrlOverride=%2Fus%2Findividual%2Fliterature%2Fprospectus%2Fp-ishares-trust-global-and-us-broad-4-30.pdf

Mao, H., Counts, S., & Bollen, J. (2011). Predicting financial markets: Comparing survey, news, Twitter, and search engine data. Proceedings of the Fifth International Conference on Weblogs and Social Media, 1–8.
https://arxiv.org/abs/1112.1051

Orangi, M. (2023). Trajectories for lithium-ion battery cost production: Can metal prices hamper the deployment of lithium-ion batteries? Batteries & Supercaps
https://chemistry-europe.onlinelibrary.wiley.com/doi/10.1002/batt.202300346

Tesla, Inc. (2024). Annual report on Form 10-K for the fiscal year ended December 31, 2023 (SEC Filing No. 001-34756). U.S. Securities and Exchange Commission https://www.sec.gov/Archives/edgar/data/1318605/000162828024002390/tsla-20231231.htm

Broadridge Financial Solutions. (2024). Landmark Broadridge Study of More than 40 Million U.S. Retail Investors Highlights the Dramatic Shifts in How Americans Invest https://www.broadridge.com/press-release/2024/broadridge-us-investor-study-highlights-growth

European Parliament & Council. (1996). Directive 96/9/EC of the European Parliament and of the Council of 11 March 1996 on the legal protection of databases. Official Journal of the European Communities, L77, 20–28
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31996L0009

United States Congress. (1986). Computer Fraud and Abuse Act of 1986, 18 U.S. Code § 1030
https://uscode.house.gov/view.xhtml?req=(title:18%20section:1030%20edition:prelim)

LinkedIn (2025). User Agreement
https://www.linkedin.com/legal/user-agreement

Nasdaq (2025). QCLN | First Trust NASDAQ Clean Edge Green Energy Index Fund. Retrieved April 29, 2025
https://www.nasdaq.com/market-activity/etf/qcln

Google Trends (2025). “Tesla”. Retrieved May 3, 2025
https://trends.google.com/trends/explore?date=all&q=%2Fm%2F0dr90d&hl=en

Altindex (2025). Pricing Plans & Features
https://altindex.com/pricing

Similarweb (2025). Pricing

https://account.similarweb.com/packages/marketing/

Additional Database References (not used as in-text citations, but relevant)

Nasdaq Data Link (2025). JODI Oil World Database. Retrieved 25 February 2025
https://data.nasdaq.com/databases/JODI

Yahoo Finance (2025). iShares Global Clean Energy ETF (ICLN). Retrieved 25 February 2025
https://finance.yahoo.com/quote/ICLN/

Statista (2025). Tesla's research and development expenses from FY 2010 to FY 2024. Retrieved 15 February 2025
https://www.statista.com/statistics/314863/research-and-development-expenses-of-tesla/

Statista (2025). Employees of Tesla from 2010 to 2024. Retrieved 15 February 2025
https://www.statista.com/statistics/314768/number-of-tesla-employees/

Statista (2025). Average cobalt spot price in the United States from 2010 to 2024. Retrieved 15 February 2025
https://www.statista.com/statistics/339743/average-spot-price-of-cobalt-in-the-us/

Statista (2025). Average lithium carbonate price from 2010 to 2024. Retrieved 15 February 2025
https://www.statista.com/statistics/606350/battery-grade-lithium-carbonate-price/

Statista (2025). Average nickel prices from 1980 to 2023. Retrieved 15 February 2025
https://www.statista.com/statistics/236578/average-nickel-prices/

Statista (2025). Average market spot price of aluminum ingots in the United States from 2011 to 2024. Retrieved 15 February 2025 https://www.statista.com/statistics/209336/price-of-aluminum-on-the-us-market/

Statista (2025). Average retail electricity price for industrial consumers in the United States from 1970 to 2023. Retrieved 15 February 2025 https://www.statista.com/statistics/190680/us-industrial-consumer-price-estimates-for-retail-electricity-since-1970/

Appendix

The data below presents various tables, datasets, regression outputs, correlation matrices and graphs. The majority of presented figures and tables do not include all the data, as the voluminous nature of the collected information would not fit into the Appendix without compromising the structure of the document and would not maintain readability. The regression summaries and outputs are not described, as the most relevant data is already presented there, including the names of the independent and dependent variables.

The sections and subsections are divided by the way they were presented in the paper.

Appendix 1: Industry-Wide Trends Related Research Material

1.1 Standard Financial Data for Comparison

Date	Date	Close ICLN Price	Volume ICLN traded	Total Distribution / Div	Dividend Yield	Monthly Total (NAV) Return	Next Month's Closing Price
2008-07-31T	2008-07-31	50.14	89,700.00	0.06	0.12%	-2.46	36.48
2008-08-31T	2008-08-31	36.48	215,600.00	0.06	0.16%	2.29	21.99
2008-09-30T	2008-09-30	21.99	323,300.00	0.06	0.27%	-27.83	19.17
2008-10-31T	2008-10-31	19.17	1,688,100.00	0.06	0.31%	-39.54	21.39
2008-11-30T	2008-11-30	21.39	398,200.00	0.06	0.28%	-13.03	20.27
2008-12-31T	2008-12-31	20.27	455,700.00	0.06	0.30%	9.33	15.23
2009-01-31T	2009-01-31	15.23	382,400.00	0.06	0.39%	-9.25	17.5
2009-02-28T	2009-02-28	17.5	455,800.00	0.06	0.34%	-19.74	21.05
2009-03-31T	2009-03-31	21.05	348,100.00	0.06	0.29%	15.03	23.67
2009-04-30T	2009-04-30	23.67	521,500.00	0.06	0.25%	20.06	22.8
2009-05-31T	2009-05-31	22.8	685,300.00	0.06	0.26%	14.20	23.35
2009-06-30T	2009-06-30	23.35	530,300.00	0.12	0.52%	-3.93	21.15
2009-07-31T	2009-07-31	21.15	545,000.00	0.12	0.58%	2.12	23.25
2009-08-31T	2009-08-31	23.25	502,700.00	0.12	0.52%	-8.86	20.43
2009-09-30T	2009-09-30	20.43	691,600.00	0.12	0.60%	9.96	21.59
2009-10-31T	2009-10-31	21.59	385,700.00	0.12	0.57%	-11.04	22.26
2009-11-30T	2009-11-30	22.26	417,600.00	0.12	0.55%	3.34	19.24
2009-12-31T	2009-12-31	19.24	701,500.00	0.17	0.88%	4.74	17.88
2010-01-31T	2010-01-31	17.88	469,600.00	0.17	0.94%	-11.96	18.84
2010-02-28T	2010-02-28	18.84	659,200.00	0.17	0.90%	-7.84	18.4
2010-03-31T	2010-03-31	18.4	902,300.00	0.17	0.92%	5.56	15.21
2010-04-30T	2010-04-30	15.21	695,800.00	0.17	1.11%	-1.69	14.6
2010-05-31T	2010-05-31	14.6	526,800.00	0.17	1.16%	-18.33	16.25
2010-06-30T	2010-06-30	16.25	296,600.00	0.15	0.95%	-2.63	15.53
2010-07-31T	2010-07-31	15.53	339,900.00	0.15	0.99%	10.61	17.27
2010-08-31T	2010-08-31	17.27	249,300.00	0.15	0.89%	-3.87	16.96
2010-09-30T	2010-09-30	16.96	358,000.00	0.15	0.91%	10.81	14.87
2010-10-31T	2010-10-31	14.87	295,700.00	0.15	1.03%	-1.73	15.84
2010-11-30T	2010-11-30	15.84	397,100.00	0.15	0.97%	-12.63	16.49
2010-12-31T	2010-12-31	16.49	432,500.00	0.11	0.66%	7.39	17.18
2011-01-31T	2011-01-31	17.18	464,800.00	0.04	0.22%	4.17	18.85
2011-02-28T	2011-02-28	18.85	1,366,200.00	0.04	0.20%	3.82	18.37
2011-03-31T	2011-03-31	18.37	965,900.00	0.04	0.20%	9.47	17.02
2011-04-30T	2011-04-30	17.02	592,700.00	0.04	0.22%	-2.03	16.2
2011-05-31T	2011-05-31	16.2	328,800.00	0.04	0.23%	-7.90	14.67
2011-06-30T	2011-06-30	14.67	568,500.00	0.17	1.18%	-3.09	13.22
2011-07-31T	2011-07-31	13.22	809,500.00	0.17	1.31%	-9.26	9.27
2011-08-31T	2011-08-31	9.27	753,900.00	0.17	1.86%	-9.52	9.8
2011-09-30T	2011-09-30	9.8	595,700.00	0.17	1.76%	-29.55	9.33
2011-10-31T	2011-10-31	9.33	392,700.00	0.17	1.85%	5.87	8.54
2011-11-30T	2011-11-30	8.54	1,067,100.00	0.17	2.02%	-6.55	9.28
2011-12-31T	2011-12-31	9.28	319,600.00	0.23	2.46%	-4.12	9.55
2012-01-31T	2012-01-31	9.55	279,200.00	0.23	2.39%	8.09	9.1
2012-02-29T	2012-02-29	9.1	398,900.00	0.23	2.51%	2.25	8.53

1.1.1 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.95768924
R Square	0.91716868
Adjusted R Square	0.91502279
Standard Error	1.62378703
Observations	199

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	5	5634.69048	1126.9381	427.407291	2.58E-102
Residual	193	508.8800704	2.6366843		
Total	198	6143.570551			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	4.17711203	0.567700666	7.35794809	5.225E-12	3.057418028	5.296806041	3.05741803	5.29680604
Close ICLN Price	0.72294943	0.030125733	23.9977378	7.779E-60	0.663531494	0.782367367	0.66353149	0.78236737
Volume ICLN traded	9.8918E-09	2.72494E-09	3.63011426	0.00036282	4.51735E-09	1.52663E-08	4.5174E-09	1.5266E-08
Total Distribution / Div	8.98457158	4.791409402	1.8751417	0.06228356	-0.465676951	18.43482012	-0.465677	18.4348201
Dividend Yield	-180.9319	47.34508585	-3.8215561	0.00017878	-274.312115	-87.55168997	-274.31211	-87.55169
Monthly Total (NAV) Return	-0.0040823	0.013168505	-0.3100073	0.75688991	-0.030054992	0.021890327	-0.030055	0.02189033

RESIDUAL OUTPUT | PROBABILITY OUTPUT

Observation	*ext Month's C*	*Residuals*	*ndard Residuals*	*Percentile*	*Next Month's Closing Price*	*squared residuals*	*Mean of squared residuals*	*RMSE*
1	40.7592886	-4.27928856	-2.6692942	0.251256281	6.31	18.3123106	2.55718628	1.59912047
2	30.7845803	-8.794580326	-5.4858003	0.753768844	6.51	77.3446431		
3	20.2369786	-1.066978603	-0.6655498	1.256281407	6.65	1.13844334		
4	18.1869437	3.203056303	1.99797222	1.75879397	6.72	10.2595697		
5	19.7296835	0.540316503	0.33703353	2.261306533	6.85	0.29194192		
6	18.8012253	-3.57122526	-2.2276252	2.763819095	7.05	12.7536499		
7	15.0554523	2.444547674	1.52483687	3.266331658	7.17	5.97581333		
8	16.8325573	4.217442651	2.63071657	3.768844221	7.22	17.7868225		
9	19.3606373	4.309362654	2.68805356	4.271356784	7.29	18.5706065		
10	21.2930304	1.506969649	0.9400033	4.773869347	7.32	2.27095752		
11	20.6721065	2.677893488	1.67039112	5.27638191	7.36	7.17111354		
12	21.2300443	-0.080044333	-0.0499293	5.778894472	7.88	0.0064071		

	Close ICLN Price	*Volume ICLN traded*	*Total Distribution / Div*	*Dividend Yield*	*Monthly Total (NAV) Return*
Close ICLN Price	1				
Volume ICLN traded	0.505395198	1			
Total Distribution / Div	-0.393883235	-0.315219473	1		
Dividend Yield	-0.680844795	-0.439114038	0.840721552	1	
Monthly Total (NAV) Return	0.051952647	0.022328168	-0.119175761	-0.114338546	1

1.2 Data with Additional AD

ICLN Financial Data							Brent Crude Oil Last Day Financ (BZ=F), frequency = monthly		US Total Oil Product Imports (Thousand Metric Tons) [code: TPIMKT]	Target Variable
Date	Date	Close ICLN Price	Volume ICLN traded	Total Distribution / Div	Dividend Yield	Monthly Total (NAV) Return	Brent Crude Oil Close Price	Brent Crude Oil Volume Traded	Oil Product Imports into the US	Next Month's Closing Price
2008-07-31T22:00:00.000Z	2008-07-31	50.14	89,700.00	0.06	0.12%	-2.46	114.05	23,653.00	7,644.00	36.48
2008-08-31T22:00:00.000Z	2008-08-31	36.48	215,600.00	0.06	0.16%	2.29	98.17	17,936.00	6,740.00	21.99
2008-09-30T22:00:00.000Z	2008-09-30	21.99	323,300.00	0.06	0.27%	-27.83	65.32	14,276.00	7,347.00	19.17
2008-10-31T23:00:00.000Z	2008-10-31	19.17	1,688,100.00	0.06	0.31%	-39.54	53.49	9,940.00	7,744.00	21.39
2008-11-30T23:00:00.000Z	2008-11-30	21.39	398,200.00	0.06	0.28%	-13.03	45.59	10,964.00	6,979.00	20.27
2008-12-31T23:00:00.000Z	2008-12-31	20.27	455,700.00	0.06	0.30%	9.33	45.88	4,510.00	8,015.00	15.23
2009-01-31T23:00:00.000Z	2009-01-31	15.23	382,400.00	0.06	0.39%	-9.25	46.35	1,928.00	8,993.00	17.5
2009-02-28T23:00:00.000Z	2009-02-28	17.5	455,800.00	0.06	0.34%	-19.74	49.23	2,175.00	7,228.00	21.05
2009-03-31T22:00:00.000Z	2009-03-31	21.05	348,100.00	0.06	0.29%	15.03	51.96	690	8,237.00	23.67
2009-04-30T22:00:00.000Z	2009-04-30	23.67	521,500.00	0.06	0.25%	20.06	65.52	1,609.00	6,723.00	22.8
2009-05-31T22:00:00.000Z	2009-05-31	22.8	685,300.00	0.06	0.26%	14.20	69.3	1,482.00	6,760.00	23.35
2009-06-30T22:00:00.000Z	2009-06-30	23.35	530,300.00	0.12	0.52%	-3.93	71.7	471	7,328.00	21.15
2009-07-31T22:00:00.000Z	2009-07-31	21.15	545,000.00	0.12	0.58%	2.12	69.65	198	6,952.00	23.25
2009-08-31T22:00:00.000Z	2009-08-31	23.25	502,700.00	0.12	0.52%	-8.86	69.07	441	6,407.00	20.43
2009-09-30T22:00:00.000Z	2009-09-30	20.43	691,600.00	0.12	0.60%	9.96	75.2	313	6,242.00	21.59
2009-10-31T23:00:00.000Z	2009-10-31	21.59	385,700.00	0.12	0.57%	-11.04	78.47	2,921.00	6,136.00	22.26
2009-11-30T23:00:00.000Z	2009-11-30	22.26	417,600.00	0.12	0.55%	3.34	77.93	396	5,305.00	19.24
2009-12-31T23:00:00.000Z	2009-12-31	19.24	701,500.00	0.17	0.88%	4.74	71.46	698	6,201.00	17.88
2010-01-31T23:00:00.000Z	2010-01-31	17.88	469,600.00	0.17	0.94%	-11.96	77.59	1,039.00	7,823.00	18.84
2010-02-28T23:00:00.000Z	2010-02-28	18.84	659,200.00	0.17	0.90%	-7.84	82.7	841	5,947.00	18.4
2010-03-31T22:00:00.000Z	2010-03-31	18.4	902,300.00	0.17	0.92%	5.56	87.44	851	6,412.00	15.21
2010-04-30T22:00:00.000Z	2010-04-30	15.21	695,800.00	0.17	1.11%	-1.69	74.02	1,683.00	7,800.00	14.6
2010-05-31T22:00:00.000Z	2010-05-31	14.6	526,800.00	0.17	1.16%	-18.33	75.01	3,920.00	7,099.00	16.25
2010-06-30T22:00:00.000Z	2010-06-30	16.25	296,600.00	0.15	0.95%	-2.63	78.18	5,460.00	6,805.00	15.53
2010-07-31T22:00:00.000Z	2010-07-31	15.53	339,900.00	0.15	0.99%	10.61	74.64	3,282.00	7,831.00	17.27
2010-08-31T22:00:00.000Z	2010-08-31	17.27	249,300.00	0.15	0.89%	-3.87	82.31	9,784.00	8,074.00	16.96
2010-09-30T22:00:00.000Z	2010-09-30	16.96	358,000.00	0.15	0.91%	10.81	83.15	7,714.00	7,199.00	14.87
2010-10-31T23:00:00.000Z	2010-10-31	14.87	295,700.00	0.15	1.03%	-1.73	85.92	12,944.00	6,773.00	15.84
2010-11-30T23:00:00.000Z	2010-11-30	15.84	397,100.00	0.15	0.97%	-12.63	94.75	8,990.00	6,501.00	16.49
2010-12-31T23:00:00.000Z	2010-12-31	16.49	432,500.00	0.11	0.66%	7.39	101.01	17,704.00	6,143.00	17.18
2011-01-31T23:00:00.000Z	2011-01-31	17.18	464,800.00	0.04	0.22%	4.17	111.8	27,851.00	7,864.00	18.85
2011-02-28T23:00:00.000Z	2011-02-28	18.85	1,366,200.00	0.04	0.20%	3.82	117.36	28,970.00	6,191.00	18.37
2011-03-31T22:00:00.000Z	2011-03-31	18.37	965,900.00	0.04	0.20%	9.47	125.89	18,795.00	6,997.00	17.02
2011-04-30T22:00:00.000Z	2011-04-30	17.02	592,700.00	0.04	0.22%	-2.03	116.73	4,612.00	7,907.00	16.2
2011-05-31T22:00:00.000Z	2011-05-31	16.2	328,800.00	0.04	0.23%	-7.90	112.48	19,525.00	7,340.00	14.67
2011-06-30T22:00:00.000Z	2011-06-30	14.67	568,500.00	0.17	1.18%	-3.09	116.74	12,092.00	6,688.00	13.22
2011-07-31T22:00:00.000Z	2011-07-31	13.22	809,500.00	0.17	1.31%	-9.26	114.85	26,360.00	6,122.00	9.27
2011-08-31T22:00:00.000Z	2011-08-31	9.27	753,900.00	0.17	1.86%	-9.52	102.76	17,587.00	5,427.00	9.8
2011-09-30T22:00:00.000Z	2011-09-30	9.8	595,700.00	0.17	1.76%	-29.55	109.56	23,438.00	5,324.00	9.33
2011-10-31T23:00:00.000Z	2011-10-31	9.33	392,700.00	0.17	1.85%	5.87	110.82	21,533.00	5,046.00	8.54
2011-11-30T23:00:00.000Z	2011-11-30	8.54	1,067,100.00	0.17	2.02%	-6.55	107.38	11,389.00	5,898.00	9.28
2011-12-31T23:00:00.000Z	2011-12-31	9.28	319,600.00	0.23	2.46%	-4.12	110.98	3,762.00	5,195.00	9.55
2012-01-31T23:00:00.000Z	2012-01-31	9.55	279,200.00	0.23	2.39%	8.09	122.66	19,706.00	6,021.00	9.1
2012-02-29T23:00:00.000Z	2012-02-29	9.1	398,900.00	0.23	2.51%	2.25	122.88	16,066.00	4,444.00	8.53
2012-03-31T22:00:00.000Z	2012-03-31	8.53	361,800.00	0.23	2.68%	-4.81	119.47	3,804.00	4,725.00	7.05
2012-04-30T22:00:00.000Z	2012-04-30	7.05	353,600.00	0.23	3.24%	-6.48	101.87	30,755.00	5,256.00	7.22
2012-05-31T22:00:00.000Z	2012-05-31	7.22	221,100.00	0.23	3.16%	-16.80	97.8	24,672.00	5,124.00	6.31

1.2.1 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.962256547
R Square	0.925937663
Adjusted R Square	0.923223337
Standard Error	1.543449579
Observations	199

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	7	5688.563359	812.6519085	341.129805	2.5824E-104
Residual	191	455.0071913	2.382236604		
Total	198	6143.570551			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	7.560198969	0.897241598	8.426045991	8.5057E-15	5.790424021	9.329973916	5.79042402	9.32997392
Close ICLN Price	0.682039835	0.031149778	21.89549561	5.6988E-54	0.620598081	0.743481589	0.62059808	0.74348159
Volume ICLN traded	1.81216E-08	3.24916E-09	5.577327391	8.2587E-08	1.17128E-08	2.45305E-08	1.1713E-08	2.453E-08
Total Distribution / Div	7.663910514	4.562812087	1.679646317	0.09466111	-1.336062893	16.66388392	-1.3360629	16.6638839
Dividend Yield	-175.0225287	45.23621127	-3.869080185	0.0001498	-264.249236	-85.7958214	-264.24924	-85.795821
Monthly Total (NAV) Return	-0.002766129	0.012537159	-0.220634455	0.82561274	-0.027495198	0.02196294	-0.0274952	0.02196294
Brent Crude Oil Close Price	-0.026420754	0.005949017	-4.441196347	1.5121E-05	-0.038154965	-0.014686544	-0.038155	-0.0146865
Brent Crude Oil Volume Traded	-2.1994E-06	5.39091E-07	-4.079818888	6.6154E-05	-3.26273E-06	-1.13606E-06	-3.263E-06	-1.136E-06

RESIDUAL OUTPUT					PROBABILITY OUTPUT					
Observation	*Predicted Next Month's Closing Price*	*Residuals*	*Standard Residuals*		*Percentile*	*Next Month's Closing Price*		*squared residuals*	*Mean of squared residuals*	*RMSE*
1	38.9511912	-2.471191196	-1.630158425		0.251256281	6.31		6.10678593	2.2864683	1.51210724
2	29.97737957	-7.987379568	-5.268995017		0.753768844	6.51		63.7982324		
3	20.8661762	-1.696176197	-1.118908129		1.256281407	6.65		2.87701369		
4	19.25179163	2.138208369	1.410501297		1.75879397	6.72		4.57193503		
5	20.93254123	-0.662541227	-0.437055281		2.261306533	6.85		0.43896088		
6	20.08725403	-4.857254026	-3.204160644		2.763819095	7.05		23.5929167		
7	16.52165678	0.978343215	0.64537881		3.266331658	7.17		0.95715545		
8	18.11303928	2.936960718	1.937410293		3.768844221	7.22		8.62573826		
9	20.4684888	3.201511197	2.111924993		4.271356784	7.29		10.2496739		
10	21.93959513	0.860404875	0.567579011		4.773869347	7.32		0.74029655		
11	21.24887818	2.101121821	1.386036598		5.27638191	7.36		4.41471291		
12	21.63143788	-0.481437881	-0.317587736		5.778894472	7.88		0.23178243		

1.2.2 Correlation Matrix

	Close ICLN Price	*Volume ICLN traded*	*Total Distribution / Div*	*Dividend Yield*	*Monthly Total (NAV) Return*	*Brent Crude Oil Close Price*	*Brent Crude Oil Volume Traded*	*Oil Product Imports into the US*
Close ICLN Price	1							
Volume ICLN traded	0.506509598	1						
Total Distribution / Div	-0.393887197	-0.315441963	1					
Dividend Yield	-0.681024554	-0.439361314	0.84071456	1				
Monthly Total (NAV) Return	0.051965135	0.022343883	-0.119175114	-0.114337722	1			
Brent Crude Oil Close Price	0.042655592	0.037797452	0.072214664	0.068385745	-0.044908677	1		
Brent Crude Oil Volume Traded	-0.101299387	0.386692628	-0.16417833	-0.104781404	0.05817196	-0.529976288	1	
Oil Product Imports into the US	0.449428746	0.02472423	-0.162026996	-0.324638016	-0.060746393	-0.213930466	-0.013468957	1

1.3 Data including Additional AD (2)

ICLN Financial Data							Brent Crude Oil Last Day Financ (BZ=F), frequency = monthly		US Total Oil Product Imports (Thousand Metric Tons) [code: TPIMKT]	US Total Primary Oil Product Imports (Thousand Metric Tons) [code: TCIMKT]	Total Energy CO2 Emissions (Million Metric Tons of Carbon Dioxide)	Renewable Energy Consumption: Residential and Commercial Sectors (Trillion Btu)	Crude Oil and Natural Gas Rotary Rigs in Operation, Total (Number of Rigs)	Average Price of Electricity to Ultimate Customers, Total (Cents per Kilowatthour, Including Taxes)	Target Variable
Date	Date	Close ICLN Price	Volume ICLN traded	Total Distribution / Div	Dividend Yield	Monthly Total (NAV) Return	Brent Crude Oil Close Price	Brent Crude Oil Volume Traded	Oil Product Imports into the US	Primary Oil Product Inflows into the US	Total Energy CO2 Emissions	Renewable Energy Consumption	Rotary Oil Rigs in Operation	Average Electricity Price to Ultimate Consumers	Next Month's Closing Price
2009-01-31T23:00:00.000Z	2009-01-31	15.23	382,400.00	0.06	0.39%	-9.25	46.35	1,928.00	8,993.00	44,973.00	533.309	10.878	1553	9.68	17.5
2009-02-28T23:00:00.000Z	2009-02-28	17.5	455,800.00	0.06	0.34%	-19.74	49.23	2,175.00	7,228.00	38,055.00	456.571	9.794	1320	9.78	21.05
2009-03-31T22:00:00.000Z	2009-03-31	21.05	348,100.00	0.06	0.29%	15.03	51.96	690	8,237.00	43,002.00	460.966	11.151	1105	9.64	23.67
2009-04-30T22:00:00.000Z	2009-04-30	23.67	521,500.00	0.06	0.25%	20.06	65.52	1,609.00	6,723.00	41,052.00	418.649	10.856	1004	9.57	22.8
2009-05-31T22:00:00.000Z	2009-05-31	22.8	685,300.00	0.06	0.26%	14.20	69.3	1,482.00	6,760.00	41,105.00	412.306	11.331	918	9.77	23.35
2009-06-30T22:00:00.000Z	2009-06-30	23.35	530,300.00	0.12	0.52%	-3.93	71.7	471	7,328.00	40,285.00	428.582	11.026	895	10.13	21.15
2009-07-31T22:00:00.000Z	2009-07-31	21.15	545,000.00	0.12	0.58%	2.12	69.65	198	6,952.00	42,963.00	447.494	11.259	931	10.28	23.25
2009-08-31T22:00:00.000Z	2009-08-31	23.25	502,700.00	0.12	0.52%	-8.86	69.07	441	6,407.00	40,084.00	458.837	11.359	980	10.27	20.43
2009-09-30T22:00:00.000Z	2009-09-30	20.43	691,600.00	0.12	0.60%	9.96	75.2	313	6,242.00	40,635.00	417.568	10.72	1009	10.09	21.59
2009-10-31T23:00:00.000Z	2009-10-31	21.59	385,700.00	0.12	0.57%	-11.04	78.47	2,921.00	6,136.00	38,700.00	432.272	11.003	1044	9.7	22.26
2009-11-30T23:00:00.000Z	2009-11-30	22.26	417,600.00	0.12	0.55%	3.34	77.93	396	5,305.00	38,831.00	427.341	10.798	1107	9.34	19.24
2009-12-31T23:00:00.000Z	2009-12-31	19.24	701,500.00	0.17	0.88%	4.74	71.46	698	6,201.00	36,826.00	510.075	11.064	1172	9.4	17.88
2010-01-31T23:00:00.000Z	2010-01-31	17.88	469,600.00	0.17	0.94%	-11.96	77.59	1,039.00	7,823.00	38,369.00	524.209	11.114	1267	9.28	18.84
2010-02-28T23:00:00.000Z	2010-02-28	18.84	659,200.00	0.17	0.90%	-7.84	82.7	841	5,947.00	35,445.00	471.433	9.969	1350	9.47	18.4
2010-03-31T22:00:00.000Z	2010-03-31	18.4	902,300.00	0.17	0.92%	5.56	87.44	851	6,412.00	41,551.00	469.037	11.29	1419	9.48	15.21
2010-04-30T22:00:00.000Z	2010-04-30	15.21	695,800.00	0.17	1.11%	-1.69	74.02	1,683.00	7,800.00	42,307.00	420.441	11.276	1479	9.53	14.6
2010-05-31T22:00:00.000Z	2010-05-31	14.6	526,800.00	0.17	1.16%	-18.33	75.01	3,920.00	7,099.00	43,011.00	435.608	11.953	1513	9.72	16.25
2010-06-30T22:00:00.000Z	2010-06-30	16.25	296,600.00	0.15	0.95%	-2.63	78.18	5,460.00	6,805.00	43,577.00	457.383	11.294	1531	10.18	15.53
2010-07-31T22:00:00.000Z	2010-07-31	15.53	339,900.00	0.15	0.99%	10.61	74.64	3,282.00	7,831.00	44,290.00	484.454	11.468	1573	10.46	17.27
2010-08-31T22:00:00.000Z	2010-08-31	17.27	249,300.00	0.15	0.89%	-3.87	82.31	9,784.00	8,074.00	42,903.00	490.663	11.579	1638	10.4	16.96
2010-09-30T22:00:00.000Z	2010-09-30	16.96	358,000.00	0.15	0.91%	10.81	83.15	7,714.00	7,199.00	40,112.00	441.784	10.998	1655	10.17	14.87
2010-10-31T23:00:00.000Z	2010-10-31	14.87	295,700.00	0.15	1.03%	-1.73	85.92	12,944.00	6,773.00	39,114.00	427.192	10.978	1668	9.81	15.84
2010-11-30T23:00:00.000Z	2010-11-30	15.84	397,100.00	0.15	0.97%	-12.63	94.75	8,990.00	6,501.00	37,810.00	444.028	10.639	1683	9.55	16.49
2010-12-31T23:00:00.000Z	2010-12-31	16.49	432,500.00	0.11	0.66%	7.39	101.01	17,704.00	6,143.00	39,514.00	527.582	11.028	1711	9.52	17.18
2011-01-31T23:00:00.000Z	2011-01-31	17.18	464,800.00	0.04	0.22%	4.17	111.8	27,851.00	7,864.00	41,433.00	529.291	11.496	1711	9.48	18.85
2011-02-28T23:00:00.000Z	2011-02-28	18.85	1,366,200.00	0.04	0.20%	3.82	117.36	28,970.00	6,191.00	32,865.00	454.482	10.566	1718	9.56	18.37
2011-03-31T22:00:00.000Z	2011-03-31	18.37	965,900.00	0.04	0.20%	9.47	125.89	18,795.00	6,997.00	40,879.00	464.574	11.739	1720	9.55	17.02
2011-04-30T22:00:00.000Z	2011-04-30	17.02	592,700.00	0.04	0.22%	-2.03	116.73	4,612.00	7,907.00	38,424.00	418.515	11.474	1790	9.54	16.2
2011-05-31T22:00:00.000Z	2011-05-31	16.2	328,800.00	0.04	0.23%	-7.90	112.48	19,525.00	7,340.00	40,760.00	425.525	12.088	1836	9.78	14.67
2011-06-30T22:00:00.000Z	2011-06-30	14.67	568,500.00	0.17	1.18%	-3.09	116.74	12,092.00	6,688.00	40,498.00	445.152	11.952	1863	10.26	13.22
2011-07-31T22:00:00.000Z	2011-07-31	13.22	809,500.00	0.17	1.31%	-9.26	114.85	26,360.00	6,122.00	42,189.00	476.216	12.244	1900	10.47	9.27
2011-08-31T22:00:00.000Z	2011-08-31	9.27	753,900.00	0.17	1.86%	-9.52	102.76	17,587.00	5,427.00	40,668.00	480.293	12.357	1957	10.49	9.8
2011-09-30T22:00:00.000Z	2011-09-30	9.8	595,700.00	0.17	1.76%	-29.55	109.56	23,438.00	5,324.00	39,568.00	428.329	11.722	1978	10.29	9.33
2011-10-31T23:00:00.000Z	2011-10-31	9.33	392,700.00	0.17	1.85%	5.87	110.82	21,533.00	5,046.00	40,496.00	424.911	11.854	2017	9.83	8.54
2011-11-30T23:00:00.000Z	2011-11-30	8.54	1,067,100.00	0.17	2.02%	-6.55	107.38	11,389.00	5,898.00	38,756.00	432.51	11.71	2011	9.58	9.28
2011-12-31T23:00:00.000Z	2011-12-31	9.28	319,600.00	0.23	2.46%	-4.12	110.98	3,762.00	5,195.00	39,999.00	475.055	11.941	2003	9.53	9.55
2012-01-31T23:00:00.000Z	2012-01-31	9.55	279,200.00	0.23	2.39%	8.09	122.66	19,706.00	6,021.00	39,070.00	477.264	11.392	2003	9.61	9.1
2012-02-29T23:00:00.000Z	2012-02-29	9.1	398,900.00	0.23	2.51%	2.25	122.88	16,066.00	4,444.00	35,974.00	440.142	10.846	1990	9.58	8.53
2012-03-31T22:00:00.000Z	2012-03-31	8.53	361,800.00	0.23	2.68%	-4.81	119.47	3,804.00	4,725.00	39,148.00	417.405	11.597	1979	9.52	7.05
2012-04-30T22:00:00.000Z	2012-04-30	7.05	353,600.00	0.23	3.24%	-6.48	101.87	30,755.00	5,256.00	37,272.00	390.729	11.464	1962	9.47	7.22
2012-05-31T22:00:00.000Z	2012-05-31	7.22	221,100.00	0.23	3.16%	-16.80	97.8	24,672.00	5,124.00	40,717.00	414.721	11.932	1977	9.64	6.31
2012-06-30T22:00:00.000Z	2012-06-30	6.31	143,200.00	0.15	2.40%	4.09	103.84	3,014.00	5,392.00	40,143.00	427.533	11.562	1972	10.13	6.72

1.3.1 Correlation Matrix

	Close ICLN Price	*Volume ICLN traded*	*Total Distribution / Div*	*Dividend Yield*	*Monthly Total (NAV) Return*	*Brent Crude Oil Close Price*	*Brent Crude Oil Volume Traded*	*Oil Product Imports into the US*	*Primary Oil Product Inflows into the US*	*Total Energy CO2 Emissions*	*Renewable Energy Consumptio n*	*Rotary Oil Rigs in Operation*	*Average Electricity Price to Ultimate Consumers*	*Next Month's Closing Price*
Close ICLN Price	1													
Volume ICLN traded	0.644670291	1												
Total Distribution / Div	-0.383599365	-0.339481939	1											
Dividend Yield	-0.708475304	-0.47146187	0.83535877	1										
Monthly Total (NAV) Return	0.129079688	-0.000221165	-0.17855	-0.1778592	1									
Brent Crude Oil Close Price	0.00535406	0.033541083	0.06510189	0.06380884	-0.0877118	1								
Brent Crude Oil Volume Traded	-0.011086218	0.376021778	-0.2127031	-0.1582167	0.00545511	-0.569069742	1							
Oil Product Imports into the US	0.424144967	0.051862242	-0.1235513	-0.2895382	0.00050868	-0.208054175	0.047589991	1						
Primary Oil Product Inflows into the US	-0.285144326	-0.653010694	0.42179399	0.41289052	-0.0945589	0.239307706	-0.698285628	0.193383149	1					
Total Energy CO2 Emissions	-0.162249965	-0.366099191	0.19456683	0.17046727	-0.0322813	0.19975358	-0.404718935	0.110115858	0.528265673	1				
Renewable Energy Consumption	0.205676873	0.600440899	-0.2539616	-0.3066332	0.00119178	-0.177077833	0.72858253	-0.01082926	-0.717153679	-0.485853264	1			
Rotary Oil Rigs in Operation	-0.388684001	-0.524502001	0.32114742	0.39864893	-0.1105592	0.646820078	-0.748501404	-0.215270026	0.629936551	0.476875762	-0.7018909	1		
Average Electricity Price to Ultimate Consumers	0.273443073	0.611113338	-0.1489796	-0.2905769	-0.071705	0.013149597	0.515337388	-0.067696761	-0.603198786	-0.396592622	0.89582666	-0.5172938	1	
Next Month's Closing Price	0.969910493	0.624153051	-0.4168073	-0.7174033	0.12735855	-0.03989571	-0.00546435	0.435397087	-0.298305126	-0.172916052	0.19952963	-0.4005664	0.255512426	1

1.3.2 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.97435439
R Square	0.94936647
Adjusted R Square	0.94686261
Standard Error	1.22733891
Observations	192

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	9	5140.39002	571.1544466	379.161781	7.142E-113
Residual	182	274.1576672	1.506360809		
Total	191	5414.547687			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	6.03884046	1.505210706	4.011956884	8.7849E-05	3.06893318	9.008747734	3.06893318	9.00874773
Close ICLN Price	0.8733709	0.041952969	20.81785687	5.02E-50	0.79059416	0.95614763	0.79059416	0.95614763
Volume ICLN traded	3.1944E-09	3.46798E-09	0.921106821	0.358214	-3.648E-09	1.0037E-08	-3.648E-09	1.0037E-08
Total Distribution / Div	-3.9214708	3.973847033	-0.986819776	0.32504108	-11.762205	3.919263467	-11.762205	3.91926347
Dividend Yield	-12.271546	41.16907119	-0.298076819	0.76598454	-93.501585	68.95849307	-93.501585	68.9584931
Monthly Total (NAV) Return	-0.0065384	0.011289321	-0.579163305	0.56319512	-0.0288131	0.015736419	-0.0288131	0.01573642
Brent Crude Oil Close Price	-0.0188592	0.005197844	-3.628275817	0.0003705	-0.029115	-0.008603429	-0.029115	-0.0086034
Brent Crude Oil Volume Traded	-2.329E-06	6.08364E-07	-3.827933604	0.00017757	-3.529E-06	-1.12842E-06	-3.529E-06	-1.128E-06
Oil Product Imports into the US	0.0003324	0.000126309	2.63164269	0.00922656	8.3182E-05	0.00058162	8.3182E-05	0.00058162
Primary Oil Product Inflows into the US	-0.0001014	3.79601E-05	-2.670752754	0.00825381	-0.0001763	-2.64836E-05	-0.0001763	-2.648E-05

RESIDUAL OUTPUT					PROBABILITY OUTPUT				
Observation	*ext Month's C*	*Residuals*	*Standard Residuals*		*Percentile*	*Next Month's Closing Price*	*squared residuals*	*Mean of squared residuals*	*RMSE*
1	16.6695576	0.830442411	0.693148051		0.26041667	6.31	0.6896346	1.42790452	1.19494959
2	18.7869864	2.26301362	1.888876891		0.78125	6.51	5.12123064		
3	21.4526937	2.217306293	1.850726209		1.30208333	6.65	4.9164472		
4	23.1490321	-0.349032125	-0.291327771		1.82291667	6.72	0.12182342		
5	22.362784	0.987215967	0.824002742		2.34375	6.85	0.97459537		
6	22.9152715	-1.765271535	-1.473424897		2.86458333	7.05	3.11618359		
7	20.5904894	2.659510619	2.219822322		3.38541667	7.17	7.07299673		
8	22.6237108	-2.193710775	-1.831031661		3.90625	7.22	4.81236696		
9	19.8034517	1.786548298	1.491184041		4.42708333	7.29	3.19175482		
10	21.0500243	1.209975679	1.00993431		4.94791667	7.32	1.46404114		
11	21.2699081	-2.029908115	-1.694309966		5.46875	7.36	4.12052695		
12	19.0225408	-1.142540811	-0.953648231		5.98958333	7.88	1.3053995		
13	18.2013376	0.638662398	0.533074407		6.51041667	8.04	0.40788966		

1.3.3 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.92079124
R Square	0.84785651
Adjusted R Square	0.83765694
Standard Error	2.14526658
Observations	192

ANOVA

	df	SS	MS	F	Significance F
Regression	12	4590.75949	382.563291	83.1267424	1.6176E-66
Residual	179	823.788195	4.60216869		
Total	191	5414.54769			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	23.5374968	3.99655984	5.88943936	1.8775E-08	15.6510636	31.4239301	15.6510636	31.4239301
Volume ICLN traded	4.229E-08	5.2374E-09	8.07459205	9.6065E-14	3.1955E-08	5.2625E-08	3.1955E-08	5.2625E-08
Total Distribution / Div	36.7315062	6.621143	5.54760805	1.0278E-07	23.6659687	49.7970437	23.6659687	49.7970437
Dividend Yield	-556.3368	59.2015267	-9.3973388	2.6495E-17	-673.1595	-439.51411	-673.1595	-439.51411
Monthly Total (NAV) Return	0.01627567	0.01993173	0.81657069	0.41525921	-0.0230557	0.05560706	-0.0230557	0.05560706
Brent Crude Oil Close Price	-0.0020705	0.01289045	-0.1606193	0.87257435	-0.0275073	0.02336634	-0.0275073	0.02336634
Brent Crude Oil Volume Traded	-7.095E-06	1.1785E-06	-6.0209146	9.5997E-09	-9.421E-06	-4.77E-06	-9.421E-06	-4.77E-06
Oil Product Imports into the US	0.00140169	0.00020382	6.87696137	9.8524E-11	0.00099948	0.00180389	0.00099948	0.00180389
Primary Oil Product Inflows into the US	-0.0002113	6.8267E-05	-3.0953058	0.00228232	-0.000346	-7.66E-05	-0.000346	-7.66E-05
Total Energy CO2 Emissions	-0.0051009	0.00496415	-1.0275443	0.30555075	-0.0148967	0.0046949	-0.0148967	0.0046949
Renewable Energy Consumption	-0.4216426	0.14787622	-2.8513214	0.00486563	-0.7134476	-0.1298377	-0.7134476	-0.1298377
Rotary Oil Rigs in Operation	-0.0030972	0.00075944	-4.0783287	6.8179E-05	-0.0045958	-0.0015986	-0.0045958	-0.0015986
Average Electricity Price to Ultimate Consumers	0.41419448	0.42125438	0.98324076	0.32681607	-0.4170691	1.24545804	-0.4170691	1.24545804

RESIDUAL OUTPUT

Observation	ext Month's C	Residuals	ndard Residuals
1	18.3002375	-0.8002375	-0.3853256
2	19.0086168	2.04138316	0.98295453
3	20.2788248	3.39117523	1.63289829
4	19.4229181	3.37708191	1.62611216
5	19.4956721	3.8543279	1.85591277
6	21.2661881	-0.1161881	-0.0559462
7	19.7321078	3.51789215	1.69391425
8	19.4294458	1.00055415	0.48178081
9	19.2968688	2.29313116	1.10417471
10	18.8913975	3.36860247	1.62202918
11	17.8153154	1.42468456	0.68600553
12	18.7175418	-0.8375418	-0.4032881
13	19.5599219	-0.7199219	-0.3466524
14	18.4553676	-0.0553676	-0.0266603
15	17.1618588	-1.9518588	-0.9398473
16	17.8610237	-3.2610237	-1.5702285
17	15.7863902	0.4636098	0.2232346
18	16.4005017	-0.8705017	-0.4191587
19	17.4582339	-0.1882339	-0.0906372
20	18.0412361	-1.0812361	-0.5206303
21	17.9167576	-3.0467576	-1.4670564
22	16.4648854	-0.6248854	-0.3008911
23	16.4513382	0.03866178	0.01861619
24	15.2311383	1.94886172	0.93840416
25	16.7064588	2.14354124	1.03214507

PROBABILITY OUTPUT

Percentile	Next Month's Closing Price
0.26041667	6.31
0.78125	6.51
1.30208333	6.65
1.82291667	6.72
2.34375	6.85
2.86458333	7.05
3.38541667	7.17
3.90625	7.22
4.42708333	7.29
4.94791667	7.32
5.46875	7.36
5.98958333	7.88
6.51041667	8.04
7.03125	8.11
7.55208333	8.25
8.07291667	8.26
8.59375	8.41
9.11458333	8.48
9.63541667	8.5
10.15625	8.53
10.6770833	8.54
11.1979167	8.56
11.71875	8.56
12.2395833	8.6
12.7604167	8.61

1.3.4 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.974402426
R Square	0.949460087
Adjusted R Square	0.94576897
Standard Error	1.23990478
Observations	192

ANOVA

	df	SS	MS	F	Significance F
Regression	13	5140.896919	395.4536092	257.228376	7.059E-108
Residual	178	273.6507675	1.537363863		
Total	191	5414.547687			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	5.689178611	2.495169077	2.280077396	0.02378912	0.76525964	10.6130976	0.76525964	10.6130976
Close ICLN Price	0.873105456	0.046155073	18.91678214	1.843E-44	0.78202392	0.96418699	0.78202392	0.96418699
Volume ICLN traded	2.73816E-09	3.67897E-09	0.744274566	0.45769181	-4.522E-09	1.00E-08	-4.522E-09	9.9982E-09
Total Distribution / Div	-4.492728508	4.40383688	-1.020185041	0.30902528	-13.183176	4.19771903	-13.183176	4.19771903
Dividend Yield	-7.655116702	44.85625778	-0.170658835	0.86468586	-96.173599	80.8633651	-96.173599	80.8633651
Monthly Total (NAV) Return	-0.006026551	0.01158016	-0.520420311	0.60341736	-0.0288786	0.01682552	-0.0288786	0.01682552
Brent Crude Oil Close Price	-0.01912354	0.007504665	-2.548220291	0.01167246	-0.0339331	-0.004314	-0.0339331	-0.004314
Brent Crude Oil Volume Traded	-2.42327E-06	7.24514E-07	-3.344684063	0.00100455	-3.853E-06	-9.935E-07	-3.853E-06	-9.935E-07
Oil Product Imports into the US	0.000340026	0.00013049	2.605766174	0.00994301	8.252E-05	0.00059753	8.252E-05	0.00059753
Primary Oil Product Inflows into the US	-9.749E-05	3.99127E-05	-2.44257797	0.01555968	-0.0001763	-1.873E-05	-0.0001763	-1.873E-05
Total Energy CO2 Emissions	-0.000820478	0.00287805	-0.285081204	0.77591336	-0.0065	0.00485901	-0.0065	0.00485901
Renewable Energy Consumption	-0.007135971	0.08823255	-0.080876856	0.93563075	-0.1812524	0.16698046	-0.1812524	0.16698046
Rotary Oil Rigs in Operation	-3.37073E-05	0.000467857	-0.072046173	0.94264605	-0.000957	0.00088955	-0.000957	0.00088955
Average Electricity Price to Ultimate Consumers	0.072061893	0.244144222	0.295161166	0.76821476	-0.4097276	0.55385143	-0.4097276	0.55385143

RESIDUAL OUTPUT

Observation	Predicted Next Month's Closing Price	Residuals	Standard Residuals		squared residuals	Mean of squared residuals	RMSE
1	16.65605382	0.843946176	0.705071399		0.71224515	1.42526441	1.19384438
2	18.80970707	2.240292928	1.871643611		5.0189124		
3	21.49988329	2.170116706	1.813015173		4.70940652		
4	23.2089451	-0.408945096	-0.341651517		0.16723609		
5	22.43892769	0.911072306	0.761151651		0.83005275		
6	22.97475623	-1.824756229	-1.524485167		3.32973529		
7	20.65669657	2.593303426	2.166564796		6.72522266		
8	22.65373168	-2.223731682	-1.857807586		4.94498259		
9	19.87086051	1.719139487	1.43624809		2.95544058		
10	21.05221766	1.207782339	1.009036841		1.45873818		
11	21.25048117	-2.010481172	-1.679648315		4.04203454		
12	18.92602945	-1.046029453	-0.873901051		1.09417762		
13	18.09274459	0.747255413	0.624291494		0.55839065		
14	18.52259464	-0.122594636	-0.10242119		0.01502944		
15	17.5198005	-2.309800495	-1.929713426		5.33517833		
16	15.45742022	-0.857420218	-0.716328232		0.73516943		
17	14.68502136	1.564978642	1.307455039		2.44915815		
18	15.91323874	-0.383238735	-0.320175242		0.14687193		
19	15.54920911	1.720790889	1.437627746		2.96112128		
20	17.20608238	-0.246082378	-0.205588521		0.06055654		

Appendix 2: Company-Specific Related Research Material

2.1 Stock Price Related

The two image pairs below depict two table in Excel, but have been presented this way for improved visibility.

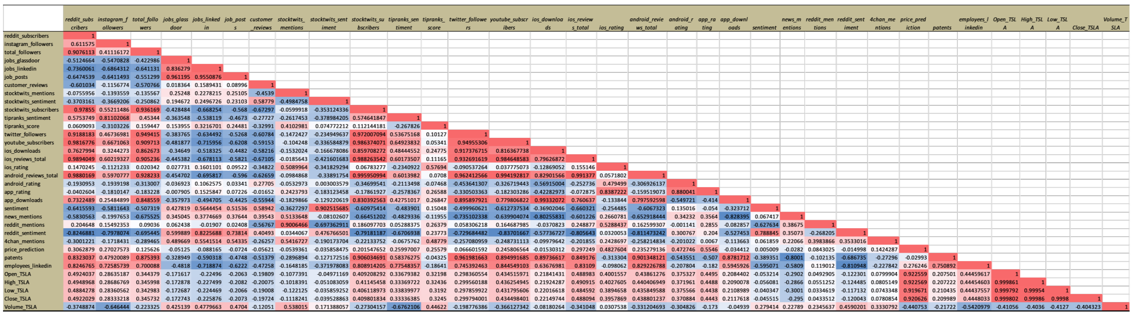
date	reddit_subscribers	stocktwits_mentions	stocktwits_sentiment	stocktwits_subscribers	instagram_followers	youtube_subscribers	total_followers	jobs_glassdoor	jobs_linkedin	job_posts			financials	news_mentions	reddit_mentions		tipranks_followers_total	tipranks_sentiment	tipranks_score
2025-04-25 00:00:00	3440446	27372	42		9570967	2820000	39736819	3865	3000	3433	59	0	{'price': 28	53	2794	65	178304	69	8
2025-04-24 00:00:00	3440487	26472	40	1025853	9572229	2820000	40761725	3871	3000	3436	53	0	{'price': 25	87	1055	49	178304	69	8
2025-04-23 00:00:00	3440613	23879	36	1025339	9573176	2820000	40762117	3855	3000	3428	56	0	{'price': 25	211	2338	64	178304	69	8
2025-04-22 00:00:00	3440777	14214	31	1024544	9574856	2820000	40758444	3867	3000	3434	51	0	{'price': 23	58	3924	52	178304	69	8
2025-04-21 00:00:00	3440896	12222	35	1024232	9576256	2820000	40755730	3878	4000	3939	51	0	{'price': 22	67	1624	48	178304	69	8
2025-04-17 00:00:00	3441414	16604	38	1023859	9584752	2820000	40751397	3943	3000	3472	53	0	{'price': 24	43	708	52	178304	69	8
2025-04-16 00:00:00	3441566	18862	45	1023710	9586444	2820000	40750002	3964	4000	3982	53	0	{'price': 24	35	770	46	178304	69	8
2025-04-15 00:00:00	3441672	21640	42	1023570	9589236	2820000	40750698	3930	3990	3960	54	0	{'price': 25	35	510	50	178304	69	8
2025-04-14 00:00:00	3441762	24129	41	1023405	9590324	2820000	40749337	3969	4000	3985	54	0	{'price': 25	26	578	52	178304	69	8
2025-04-11 00:00:00	3442386	29397	34	1023098	9594364	2820000	40742908	3912	4000	3956	48	0	{'price': 25	73	794	40	178304	69	8
2025-04-10 00:00:00	3442496	30405	33	1022820	9595626	2820000	40738697	3922	4000	3961	50	0	{'price': 25	68	695	48	178304	69	8
2025-04-09 00:00:00	3442570	35452	32	1022436	9600906	2820000	40731496	3959	3000	3480	51	0	{'price': 27	52	853	53	178304	69	8
2025-04-08 00:00:00	3442605	34892	34	1022138	9608010	2820000	40736666	3965	3000	3483	50	0	{'price': 22	58	782	46	178304	69	8
2025-04-07 00:00:00	3442181	35076	34	1021686	9610915	2820000	40742707	3905	4000	3953	50	0	{'price': 23	56	888	48	178304	69	8
2025-04-04 00:00:00	3437473	32782	37		9614448	2820000	39709550	3947	4000	3974	52	0	{'price': 23	71	360	49	178304	69	8
2025-04-03 00:00:00	3436158	33736	40	1020559	9615694	2820000	40726957	3996	3000	3498	55	0	{'price': 26	117	1035	55	178304	69	8
2025-04-02 00:00:00	3435579	30150	39	1019935	9617091	2820000	40723318	3969	3000	3485	52	0	{'price': 28	95	2500	49	178304	69	8
2025-04-01 00:00:00	3435092	31348	41	1019662	9618519	2820000	40720146	3967	4106	4037	52	0	{'price': 26	71	1362	47	178304	69	8
2025-03-31 00:00:00	3434706	33886	46		9619517	2820000	39696736	3968	3000	3484	56	0	{'price': 25	56	1670	53	178304	69	8
2025-03-28 00:00:00	3434022	36301	49	1018887	9622021	2820000	40705683	3921	4000	3961	59	0	{'price': 26	65	1263	58	178304	69	8
2025-03-27 00:00:00	3433742	37646	46	1018523	9622696	2820000	40702339	3914	3000	3457	60	0	{'price': 27	70	1771	64	178304	69	8
2025-03-26 00:00:00	3433402	37390	47	1018096	9624019	2820000	40701836	3897	4000	3949	58	0	{'price': 27	58	2538	57	178304	69	8
2025-03-25 00:00:00	3433030	40708	42	1017542	9622452	2820000	40694075	3947	4000	3974	56	0	{'price': 28	81	3217	58	178304	69	8
2025-03-24 00:00:00	3432505	39906	36	1016859	9625647	2820000	40692900	3978	4000	3989	57	0	{'price': 27	59	3095	67	178304	69	8
2025-03-21 00:00:00	3431544	37866	33	1015805	9629678	2820000	40679132	3965	3000	3483	54	0	{'price': 24	99	4349	61	178304	69	8

twitter_mentions	twitter_followers	ios_downloads	ios_reviews_total	ios_rating	android_downloads	android_reviews_total	android_rating	app_rating	app_downloads	4chan_mentions	employees	patents	price_prediction	ai_scores	employees_linkedin	google_trend	bitrank	index	Close	High	Low	Open	Volume
17561	23905406	17516	20349	3.9	4280	42646	3.6	3.75	10898	12	140473	468	266.319	{'score': 39, 'audience': 33, 'employment': 29, 'customer': 50, 'fundamental': 50}				830	285	287	260	262	167560700
12157	23903156	17807	20348	3.9	4295	42621	3.6	3.75	11051	0	140473	468	242.542	{'score': 39, 'audience': 33, 'employment': 29, 'customer': 50, 'fundamental': 50}				829	260	260	249	251	94464200
18003	23902989	18018	20348	3.9	4288	42600	3.6	3.75	11153	0	140473	468	234.859	{'score': 39, 'audience': 33, 'employment': 29, 'customer': 50, 'fundamental': 52}				828	251	259	244	255	150381900
21731	23898267	18395	20341	3.9	4352	42588	3.6	3.75	11374	0	140473	468	221.495	{'score': 38, 'audience': 33, 'employment': 29, 'customer': 50, 'fundamental': 46}				827	238	243	230	231	120858500
12321	23894346	18067	20328	3.9	4211	42570	3.6	3.75	11139	0	140473	468	214.588	{'score': 41, 'audience': 33, 'employment': 38, 'customer': 50, 'fundamental': 50}				826	228	232	223	230	97768000
12410	23881372	18921	20307	3.9	4188	42485	3.6	3.75	11555	0	140473	468	230.889	{'score': 43, 'audience': 33, 'employment': 29, 'customer': 58, 'fundamental': 50}				825	241	244	238	243	83404800
11138	23878282	19091	20303	3.9	4203	42461	3.7	3.8	11647	0	140473	468	235.674	{'score': 46, 'audience': 33, 'employment': 38, 'customer': 60, 'fundamental': 50}				824	242	252	234	248	112378700
12834	23876220	18928	20298	3.9	4121	42440	3.7	3.8	11525	0	140473	468	248.832	{'score': 47, 'audience': 33, 'employment': 38, 'customer': 58, 'fundamental': 58}	71833			823	254	259	248	250	79594300
13008	23873846	18928	20286	3.9	4121	42431	3.7	3.8	11525	42	140473	468	246.826	{'score': 46, 'audience': 33, 'employment': 38, 'customer': 58, 'fundamental': 61}				822	252	262	246	258	100135200
11410	23863060	19743	20258	3.9	4130	42377	3.7	3.8	11937	41	140473	468	237.394	{'score': 40, 'audience': 33, 'employment': 46, 'customer': 39, 'fundamental': 43}				821	252	258	241	252	128948100
14892	23857755	19743	20248	3.9	4130	42358	3.7	3.8	11937	65	140473	468	241.933	{'score': 43, 'audience': 33, 'employment': 46, 'customer': 50, 'fundamental': 43}				820	252	262	239	260	181722600
18141	23845584	19803	20244	3.9	4126	42326	3.7	3.8	11965	96	140473	468	260.534	{'score': 43, 'audience': 33, 'employment': 36, 'customer': 56, 'fundamental': 45}				819	272	275	224	225	219433400
14596	23843913	20022	20228	3.9	4173	42307	3.7	3.8	12098	107	140473	468	208.548	{'score': 40, 'audience': 33, 'employment': 36, 'customer': 50, 'fundamental': 35}				818	222	250	218	245	171603500
16048	23847925	19556	20204	3.9	4069	42290	3.7	3.8	11813	114	140473	468	224.027	{'score': 43, 'audience': 33, 'employment': 46, 'customer': 56, 'fundamental': 35}				817	233	252	214	224	183453800
12372	23837629	20121	20134	3.9	4004	42160	3.7	3.8	12063	81	140473	468	229.923	{'score': 43, 'audience': 33, 'employment': 46, 'customer': 56, 'fundamental': 35}				816	239	261	236	255	181229400
11391	23834546	20121	20130	3.9	4004	42131	3.7	3.8	12063	102	140473	468	255.825	{'score': 43, 'audience': 33, 'employment': 36, 'customer': 56, 'fundamental': 45}				815	267	276	262	265	136174300
20224	23830713	19543	20089	3.9	3776	42092	3.7	3.8	11660	80	140473	468	268.591	{'score': 39, 'audience': 33, 'employment': 36, 'customer': 47, 'fundamental': 45}				814	283	285	251	255	212787800
16214	23826873	20175	20077	3.9	3904	42053	3.7	3.8	12040	11	140473	468	260.169	{'score': 45, 'audience': 33, 'employment': 46, 'customer': 56, 'fundamental': 45}				813	268	277	259	264	146486900
14027	23822513	20363	20018	3.9	3915	42039	3.7	3.8	12139	62	140473	468	259.16	{'score': 50, 'audience': 33, 'employment': 42, 'customer': 67, 'fundamental': 45}				812	259	261	243	249	134008900
14925	23810753	19705	19955	3.9	3723	41977	3.7	3.8	11714	54	140473	468	258.208	{'score': 47, 'audience': 33, 'employment': 61, 'customer': 50, 'fundamental': 35}				811	264	276	261	276	123809400
16291	23807378	19993	19930	3.9	3824	41962	3.7	3.8	11909	80	140473	468	261.75	{'score': 43, 'audience': 33, 'employment': 42, 'customer': 50, 'fundamental': 35}				810	273	292	272	272	162572100
15159	23806319	19874	19911	3.9	3840	41918	3.7	3.8	11857	95	140473	468	265.442	{'score': 46, 'audience': 33, 'employment': 57, 'customer': 50, 'fundamental': 35}				809	272	285	267	283	153629800

month	Date	reddit_subscribers	instagram_followers	total_followers	jobs_glassdoor	jobs_linkedin	job_posts	customer_reviews	stocktwits_mentions	stocktwits_sentiment	stocktwits_subscribers	tipranks_sentiment	tipranks_score	twitter_followers	youtube_subscribers	ios_downloads	ios_reviews_total
2025-03	2025-03-31 00:00:00	3428386.194	9642454.355	40004258.87	3986.8	3813.74194	3900.74	0	37318.03226	34.19354839	1017509.909	69	8	23752366.42	2820000	19032.03226	19790.83871
2025-02	2025-02-28 00:00:00	3403376.25	9697162.25	39351839.61	3799.64	3496.32143	3641.07	32.64285714	19687.75	44.57142857	1000088.316	69	8	23603139	2812500	18984	19384.46429
2025-01	2025-01-31 00:00:00	3350170.645	9715453.935	40277140.55	3332.3	3017.74194	3169.9	38	12429.67742	62.41935484	993318.7742	69	8.096774194	23418519.77	2799677.419	18296.41935	18288.6129
2024-12	2024-12-31 00:00:00	3293718.516	9711194.613	40002786.32	3001.483871	3118.74194	3060.35	38	14755.87097	67.90322581	986090.3226	69.14285714	9.095238095	23224363.52	2787419.355	18585.70968	17406.83871
2024-11	2024-11-30 00:00:00	3240095.267	9669016.6	39668176.27	2175.133333	3234.93333	2705.2	40	13635	64.9	978885.4333	67	6	23011178.97	2769000	15747	16664.03333
2024-10	2024-10-31 00:00:00	3187736.774	9640367.645	39284806.87	2440.387097	3000	2720.48	39.35483871	12879.12903	68.61290323	971121.8065	67	6.482758621	22761709.68	2723870.968	15648	16216.03226
2024-09	2024-09-30 00:00:00	3132708.655	9625728.833	38849299.7	2114.133333	2304.16667	2209.37	38.66666667	6362.166667	67.5	966860.9	67	4.4	22547091.6	2681333.333	15763.96667	15717.43333
2024-08	2024-08-31 00:00:00	3066514.677	9656602.581	38790509.13	1419.354839	1258.06452	1338.71	40.51612903	7499.612903	60.96774194	964161.871	67.22580645	3.451612903	22430004.19	2673225.806	17316.29032	15346.41935
2024-07	2024-07-31 00:00:00	2977970.871	9679854.742	36178876.97	959.5806452	982.225806	971.065	40.77419355	16422.77419	62.22580645	959810.7097	68.61290323	4.483870968	22303498.71	2663333.333	17249.48387	15015.64516
2024-06	2024-06-30 00:00:00	2894415.2	9689375.733	38969907.23	404.7333333	521.3	463.267	44.33333333	8694.066667	61.4	953955.5333	69.26315789	3.631578947	22780827.43	2651333.333	17939.73333	14666.73333
2024-05	2024-05-31 00:00:00	2807516.548	9702623.097	38996423.81	119.3225806	136.741935	128.323	44.12903226	11073.25806	62.12903226	950014.1613	69	5.290322581	22894979.68	2641290.323	18274.96774	14295.87097
2024-04	2024-04-30 00:00:00	2718020.567	9706322.667	38119893.27	2689.933333	2606.06667	2648.07	41.33333333	17530.63333	58.66666667	944953.1667	69.4	5.166666667	22119596.87	2631000	18193.53333	13950.83333
2024-03	2024-03-31 00:00:00	2620247.419	9707621.806	36939824.26	5987.096774	5612.90323	5800	40	12324.83871	62.35483871	939543	69.5483871	5.677419355	21761159.1	2613225.806	17072.87097	13599
2024-02	2024-02-29 00:00:00	2575031.577	9708990.69	33436547.03	6165.517241	5965.51724	6065.52	40	12631.51724	68.13793103	934247.2414	69.37931034	4.965517241	21615630.96	2595862.069	15676.03448	13286.31034
2024-01	2024-01-31 00:00:00	2507576.296	9702641.032	36892300.87	5977.419355	6000	5988.71	40	14563.77419	60.25806452	927932	69.35483871	5.580645161	21502871.06	2574838.71	15848.45161	13004.29032
2023-12	2023-12-31 00:00:00	2499171.194	9618255.806	35589615.94	5829.032258	5548.3871	5688.71	40	11137.48387	70.4516129	923982.6452	67.48387097	4.870967742	21402220.52	2526774.194	16118.6129	12654.93548
2023-11	2023-11-30 00:00:00	2468257.133	9497292.8	36586591.13	5976.7	5700	5838.37	40	11043.13333	68.36666667	919644	64.26666667	4.733333333	21241397.2	2460000	14699.13333	12325.6
2023-10	2023-10-31 00:00:00	2439171.29	9490078.871	36391074.29	6529.032258	6064.51613	6296.77	40	12246.70968	58.51612903	914709.7097	67.29032258	5.096774194	21096791.84	2450322.581	14894.6129	12020.45161
2023-09	2023-09-30 00:00:00	2412953	9499288.1	36183867.53	6700	6166.66667	6433.33	40	11654.56667	66.76666667	911645.7667	67.63333333	6.233333333	20927980.67	2432000	15833.7	11719.26667
2023-08	2023-08-31 00:00:00	2372800.774	9499618	35957401.74	7109.677419	6612.90323	6861.29	40	11932.19355	63.90322581	908301.9032	66.4516129	5.806451613	20747648.81	2429032.258	16090.54839	11388.90323
2023-07	2023-07-31 00:00:00	2293365.129	9478815.71	35590165.74	7345.16129	6903.22581	7124.19	40	17924.64516	75.06451613	903865.129	66.29032258	8.35483871	20494119.77	2420000	16487.16129	11066.45161
2023-06	2023-06-30 00:00:00	2248780.133	9486199.533	35348782.6	7823.333333	7333.33333	7578.33	40.4	16700.53333	79.13333333	899254.6	65.23333333	8.8	20305881.67	2408666.667	14078.7	10753.46667
2023-05	2023-05-31 00:00:00	2221082.29	9500185.968	35132149.48	6073.806452	7935.48387	7004.74	41.58333333	9918.419355	70.38709677	895096.129	65.83870968	8.032258065	20128365.74	2387419.355	15186.19355	10502.45161
2023-04	2023-04-30 00:00:00	2197587.5	9513307.8	34882249.83	2273.933333	7000	4637.2	42	14878.53333	63.3	891716.1333	66	7.966666667	19905638.4	2374000	15952.53333	10194.8
2023-03	2023-03-31 00:00:00	2172253.387	9514362.806	34590192.19	3306.806452	6451.6129	4879.52	42	15312.41935	63.58064516	888326.4194	64.29032258	8.548387097	19650088.29	2365161.29	14886.16129	9945.290323
2023-02	2023-02-28 00:00:00	2148445.893	9497145.321	34231808.43	4996.464286	6000	5498.5	44.08	23933.75	69.67857143	883982.1071	62.64285714	8.428571429	19369020.82	2333214.286	14833.96429	9737.821429
2023-01	2023-01-31 00:00:00	2120713.484	9521507.226	33920635.84	6397	6258.06452	6327.77	44.5	29192.45161	64.22580645	876758.4194	63.38709677	7.193548387	19084559.94	2317096.774	11962.83871	9557.193548
2022-12	2022-12-31 00:00:00	2087272.387	9559929.323	33500551.1	7309.935484	7483.87097	7396.94	46.07142857	28374.80645	56.58064516	866525.5484	64.74193548	5.935483871	18678759.32	2308064.516	9172.83871	9274.806452
2022-11	2022-11-30 00:00:00	2056384.1	9592160.667	31585229.2	7642	6900	7271	48	15466.66667	61.43333333	860735.3	66.46666667	3.466666667	17984945.5	2290000	9373.633333	8924.033333
2022-10	2022-10-31 00:00:00	2026671.484	9624446.742	32008089.19	6967.741935	6129.03226	6548.39	49.06666667	18330.06452	60.16129032	854694.2258	66.90322581	4.290322581	17216147.71	2286129.032	9489.451613	8649.225806
2022-09	2022-09-30 00:00:00	1985887.967	9650593.6	31656934.2	2650.933333	5966.66667	4309.03	50.96551724	11064.53333	60.53333333	846450.7667	68.1	5.1	16920335.2	2253666.667	9065.966667	8381.433333
2022-08	2022-08-31 00:00:00	1945595.154	9619124.074	29967435.56	1714.111111	4300.55556	3007.56	52.58333333	16636.18519	66.07407407	835812.12	66.72	5.44	16686945.32	2250000	8167.28	8189.6

ios_rating	android_reviews_total	android_rating	app_rating	app_downloads	sentiment	news_mentions	reddit_mentions	reddit_sentiment	4chan_mentions	price_prediction	patents	employees_linkedin	Open_TSLA	High_TSLA	Low_TSLA	Close_TSLA	Volume_TSLA	EMA_20_TSLA	EMA_50_TSLA	Next Month's Closing Price
3.9	41787.6129	3.7	3.8	11474.90323	52.83871	56.90322581	1961.806452	55.19354839	24.43333333	249.5085484	468	69761	256.7123813	263.5028578	247.0380947	255.8080967	2718957500	267.4843959	254.8351609	255.8975014
3.9	41341.35714	3.775	3.8375	11559.39286	56.5	49.03571429	916.5714286	55.78571429	12.89285714	358.5751429	468	69278	347.4263161	353.7784231	336.8668438	343.3394727	1581998600	268.71348	254.7954492	255.8080967
3.83871	40707.6129	3.738709677	3.78871	11066.41935	63.83871	51.93548387	604.9354839	60.25806452	17.22580645	413.9943226	468	66168	405.8294998	415.3145004	394.9695038	405.0189972	1510656500	260.8581123	251.1814075	343.3394727
3.8	40073.03226	3.703225806	3.751613	11382.83871	65.80645	33.74193548	690.0967742	60.83870968	16.48387097	453.2519355	468	66134	415.4738087	426.2109491	404.0423831	415.9614287	1903365900	245.6832823	244.9023222	405.0189972
3.8	39449.73333	3.8	3.8	9833.766667	63.83333	37.03333333	585.5666667	61.83333333	6.633333333	333.5264667	468	65258	318.2600014	326.4454987	309.8530006	318.2195	2091913900	227.7592669	237.9203178	415.9614287
3.8	38926.22581	3.8	3.8	9843.322581	63.83871	35.90322581	573.1612903	56.38709677	3.451612903	239.844	468	64958	239.4834794	243.0213053	234.3791285	238.780436	1910807800	218.2371371	234.6428002	318.2195
3.8	38392.66667	3.793333333	3.796667	9934.333333	64	21.16666667	240.5666667	56.37037037	2.185185185	229.8409	468	64549	234.3015015	239.4784996	230.0955009	235.1599991	1611273300	216.0746846	234.4739171	238.780436
3.73871	37707.93548	3.7	3.719355	10832.80645	61.64516	28.06451613	191.3548387	56.61290323	2.774193548	188.7361935	468	65271	209.3709113	214.086819	204.0949998	209.0109107	1618373100	214.0657041	234.4459138	235.1599991
3.70323	37098.83871	3.7	3.701613	10755.19355	62.96774	37.61290323	714.7741935	58.06451613	6.483870968	216.3344839	468	64873	239.7295442	247.027273	234.3622721	240.6213636	2946114300	214.5977877	235.4840771	209.0109107
3.7	36583.26667	3.7	3.7	11009.06667	61.83333	32.93333333	425	58.76666667	9.566666667	169.9206667	468	64816	181.9142103	185.6947367	178.8110528	182.2199996	1407037300	211.8584639	235.274392	240.6213636
3.7	36090	3.690322581	3.695161	11190.12903	62.3871	34.7	333.3870968	55.6	3.4	177.4816452	468	63861	177.5040907	180.4959086	174.1959083	177.2318192	1668180900	214.9783022	237.4398774	182.2199996
3.7	35599.73333	3.62	3.66	11058	61.1	58.76666667	871.1333333	54.22222222	11.66666667	162.9887333	468	67525	165.4113645	169.2040912	162.1540888	165.8722721	2478540800	218.9516162	239.8973491	177.2318192
3.7	35151.32258	3.6	3.65	10383.48387	63.3871	41.87096774	425.0322581	58.35483871	18.32258065	172.2309355	468	66557	177.0000015	180.1024994	173.1084999	176.1630005	1898904200	224.5389156	242.9187809	165.8722721
3.7	34720.06897	3.6	3.65	9471.172414	65.89655	44.82758621	355.6896552	60.34482759	19	191.4514138	468	64443	191.7359985	195.6045006	188.311499	192.7930008	2019907700	229.6311172	245.6435066	176.1630005
3.7	34269	3.683870968	3.691935	9666.806452	65.32258	53.61290323	482.5806452	66.03225806	22.58064516	227.3555806	466.74	59849	217.4571446	220.4428558	213.2419041	216.1304765	2343784600	233.5088137	247.8006701	192.7930008
3.7	33577.45161	3.7	3.7	9940.870968	69.45161	41.48387097	218.2580645	70.51612903	8.161290323	255.0163548	460.5	61441	246.9670021	251.2225014	243.2389977	247.1379982	2294598400	235.3381123	249.0933311	216.1304765
3.7	32994.3	3.7	3.7	9281.8	67.36667	43.13333333	265.4333333	69.63333333	11.8	216.6975	449.47	61394	229.0119026	233.1723807	223.7419041	229.4119052	2650798400	234.0960191	249.1731406	247.1379982
3.7	32574.87097	3.719354839	3.709677	9564	64.90323	35.5483871	394.9032258	69.09677419	13.22580645	214.3630645	438.97	57589	237.3695464	242.3645456	232.3668178	236.9077266	2590570100	234.5890837	249.9797216	229.4119052
3.70333	32105.76667	3.78	3.741667	10070.4	68.3	43.73333333	332.3	70.13333333	15	241.0885667	431.13	62591	257.3050018	262.094001	252.2130013	256.968	2439306100	234.345016	250.5132724	236.9077266
3.7871	31644.25806	3.8	3.793548	10100.96774	67	38.16129032	255.0645161	70.64516129	12.77419355	242.9157742	420.61	59634	242.0373907	246.7547826	237.6600003	242.333043	2501580900	231.9636492	250.2498141	256.968
3.8	31156.83871	3.761290323	3.780645	10298.41935	71.03226	52.87096774	521.7419355	71.80645161	19.83870968	280.9277742	411.9	56119	274.6194992	278.525	269.1740005	273.5055	2392089000	230.8721341	250.5729477	242.333043
3.8	30665.23333	3.7	3.75	8916	72.13333	46.96666667	485.5	72.3	15.13333333	255.9003333	405.1	57485	244.0590501	250.3104771	239.9628565	246.0904759	3440477900	226.3844114	249.6369251	273.5055
3.8	30276.70968	3.741935484	3.770968	9409.258065	69.48387	46.74193548	229.3548387	72.41935484	9.612903226	173.9773226	401.06	49534	174.7731816	178.6186371	171.9763621	175.7990896	2681994800	224.3100888	249.7816781	246.0904759
3.8	29818.06667	3.7	3.75	10405.5	66.73333	55.9	413.2333333	70.8	15.06666667	196.8249333	393.8	55855.36594	177.734212	180.4736842	173.5278947	176.955262	2505176300	229.4165098	252.8013756	175.7990896
3.8	29371.19355	3.7	3.75	9799.451613	66.22581	55.58064516	365.2258065	70.70967742	18.51612903	196.8563226	388	51585	188.425218	192.462608	184.2426081	188.5226089	3311619900	234.9387464	255.8971354	176.955262
3.8	28980.78571	3.7	3.75	9524.75	67.85714	286.6785714	657.1785714	71.28571429	22.27272727	203.5200714	376.5	51980	198.1784202	204.2347372	193.3321051	199.7752637	3624845100	239.8246556	258.647116	188.5226089
3.8	28561.80645	3.709677419	3.754839	7160.548387	65.32258	465.6451613	940.483871	68.25806452	14.39207831	143.5333226	361.52	53143.27588	132.3159996	137.2429993	128.4305	134.1959991	3897499400	244.0403811	261.0500488	199.7752637
3.8	27983.70968	3.764516129	3.782258	5853.709677	63.12903	458.1666667	1361.193548	68.09677419	14.43302167	154.1395484	350.26	50380	155.8971427	158.9214278	150.1614293	152.9704764	2944247700	255.6029476	266.2277651	134.1959991
3.8	27413.63333	3.8	3.8	5764.833333	65.2	450.0666667	656.9	67.83333333	14.47396502	317.203	341.57	51697.5	193.479049	197.3523807	186.9152382	191.2466663	1885275300	266.4063656	270.8505116	152.9704764
3.8	26819.96774	3.8	3.8	5896.096774	65.03226	366.0322581	745.0645161	68.19354839	14.51490838	310.27	334.23	59322	224.1261902	229.7885735	217.4819053	223.8104771	1735263100	274.3179129	274.0996481	191.2466663
3.8	26301.4	3.886666667	3.843333	5783.4	65.86667	210.1	518.5666667	68.93333333	14.55585174	308.11	326	60070	287.8857102	294.0271403	282.8238089	288.734763	1299271000	279.6344851	276.1522673	223.8104771
3.8	25801.48	3.876	3.838	5265.12	68.03704	280.04	477.1538462	70.92307692	14.5967951	325.5047778	314.04	57920	297.2230424	301.8520282	290.2385506	294.8698571	1695263200	278.6765611	275.6386961	288.734763

2.1.1 Correlation Matrix



2.1.2 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.86913204
R Square	0.7553905
Adjusted R Sq	0.71557035
Standard Erro	33.1667766
Observations	32

ANOVA

	df	SS	MS	F	Significance F
Regression	7	146074.09	20867.7272	18.9700563	2.8662E-11
Residual	24	47301.508	1100.03507		
Total	31	193375.598			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	128.614442	147.182791	0.87384157	0.38705885	-168.20794	425.436828	-168.20794	425.436828
Open_TSLA	-10.791811	7.53271588	-1.4326587	0.15918582	-25.982981	4.39935817	-25.982981	4.39935817
High_TSLA	-0.750991	8.02300949	-0.0936046	0.92585802	-16.930932	15.4289497	-16.930932	15.4289497
Low_TSLA	-4.9811131	6.5091305	-0.7652501	0.44830256	-18.108026	8.14579961	-18.108026	8.14579961
Close_TSLA	17.2898534	7.79075327	2.21927878	0.03179411	1.57830208	33.0014048	1.57830208	33.0014048
Volume_TSLA	-1.251E-08	9.9179E-09	-1.2618475	0.21380559	-3.252E-08	7.4865E-09	-3.252E-08	7.4865E-09
EMA_20_TSLA	-0.650892	1.01711099	-0.6399419	0.5256056	-2.7020918	1.40030781	-2.7020918	1.40030781
EMA_50_TSLA	0.37202855	1.39665873	0.26637041	0.79122714	-2.4446022	3.18865932	-2.4446022	3.18865932

RESIDUAL OUTPUT / PROBABILITY OUTPUT

Observation	ext Month's Cl	Residuals	ndard Residuals	Percentile	onth's Closing Price	squared residuals	Mean of squared residuals	RMSE
1	239.369503	16.5279986	0.53736302	0.98039216	134.195999	273.174738	927.480548	30.45
2	271.977354	-16.169257	-0.5256995	2.94117647	152.970476	261.444872		
3	377.16351	-33.824037	-1.0996968	4.90196078	165.872272	1144.06548		
4	411.527276	-6.5082791	-0.211599	6.8627451	175.79909	42.357697		
5	321.497061	94.4643681	3.07125257	8.82352941	176.163	8923.51684		
6	243.988416	74.2310835	2.41342223	10.7843137	176.955262	5510.25376		
7	266.405703	-27.625267	-0.8981606	12.745098	177.231819	763.155374		
8	233.126736	2.03326282	0.06610602	14.7058824	182.22	4.13415768		
9	259.962875	-50.951965	-1.6565649	16.6666667	188.522609	2596.1027		

2.1.3 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.95776302
R Square	0.91731
Adjusted R Square	0.82910734
Standard Error	27.0571363
Observations	32

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	16	121819.85	7613.74064	10.4000259	2.1213E-05
Residual	15	10981.3294	732.088626		
Total	31	132801.18			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	4781.18548	1436.59143	3.32814562	0.0045864	1719.16333	7843.20764	1719.16333	7843.20764
app_downloads	-0.0093594	0.00917167	-1.0204655	0.32368515	-0.0289083	0.01018958	-0.0289083	0.01018958
sentiment	-2.7990159	6.80163156	-0.4115212	0.68651022	-17.29635	11.6983186	-17.29635	11.6983186
news_mentions	0.01346751	0.11833011	0.11381304	0.91089537	-0.2387471	0.26568217	-0.2387471	0.26568217
reddit_mentions	-0.0023966	0.03566422	-0.0672	0.94731005	-0.0784131	0.07361984	-0.0784131	0.07361984
reddit_sentiment	9.26893512	3.3656418	2.75398741	0.01476727	2.09523944	16.4426308	2.09523944	16.4426308
4chan_mentions	1.11783986	1.69845936	0.65814931	0.52041619	-2.5023406	4.73802029	-2.5023406	4.73802029
price_prediction	0.12482782	0.2946022	0.42371653	0.67778587	-0.5031019	0.75275755	-0.5031019	0.75275755
patents	-1.907737	0.75969582	-2.5111853	0.02397044	-3.5269903	-0.2884837	-3.5269903	-0.2884837
employees_linkedin	0.0015575	0.00260478	0.59794021	0.55880029	-0.0039945	0.00710946	-0.0039945	0.00710946
Open_TSLA	9.05947901	14.1097627	0.64207168	0.5305177	-21.014768	39.1337264	-21.014768	39.1337264
High_TSLA	-18.893946	14.9033978	-1.2677609	0.22420995	-50.659786	12.8718948	-50.659786	12.8718948
Low_TSLA	-12.186677	13.1265628	-0.9283982	0.36790051	-40.165284	15.791929	-40.165284	15.791929
Close_TSLA	22.3924802	13.8132422	1.62108792	0.12582551	-7.0497487	51.8347091	-7.0497487	51.8347091
Volume_TSLA	6.6198E-09	1.924E-08	0.34406061	0.73557647	-3.439E-08	4.7629E-08	-3.439E-08	4.7629E-08
EMA_20_TSLA	8.69076746	4.93750475	1.76015374	0.09875195	-1.8332748	19.2148097	-1.8332748	19.2148097
EMA_50_TSLA	-25.300843	9.8780046	-2.5613313	0.02170463	-46.355311	-4.2463742	-46.355311	-4.2463742

RESIDUAL OUTPUT					PROBABILITY OUTPUT					
Observation	*ext Month's Cl*	*Residuals*	*ndard Residuals*		*Percentile*	*onth's Closing Price*		*squared residuals*	*Mean of squared residuals*	*RMSE*
1	267.606686	-11.709185	-0.6221284		1.5625	134.195999		137.105006	343.166543	18.52
2	250.067579	5.74051727	0.30500321		4.6875	152.970476		32.9535386		
3	345.516891	-2.1774178	-0.1156898		7.8125	165.872272		4.74114845		
4	391.538943	13.4800542	0.71621766		10.9375	175.79909		181.71186		
5	379.30652	36.6549089	1.94753617		14.0625	176.163		1343.58235		
6	315.464096	2.75540364	0.14639917		17.1875	176.955262		7.5922492		
7	285.923936	-47.1435	-2.5048124		20.3125	177.231819		2222.50958		
8	251.720379	-16.56038	-0.8798805		23.4375	182.22		274.246197		
9	246.830282	-37.819371	-2.009406		26.5625	188.522609		1430.30481		

2.1.4 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.89038099
R Square	0.7927783
Adjusted R Square	0.59850795
Standard Error	41.472345
Observations	32

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	15	105281.893	7018.79289	4.08079936	0.00406605
Residual	16	27519.2863	1719.9554		
Total	31	132801.18			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	8135.97033	4446.41766	1.82978095	0.08597791	-1290.014	17561.9547	-1290.014	17561.9547
reddit_subscribers	0.00025009	0.00040042	0.62457052	0.54105775	-0.0005988	0.00109896	-0.0005988	0.00109896
instagram_followers	-0.0007573	0.00048107	-1.5742937	0.13498347	-0.0017772	0.00026248	-0.0017772	0.00026248
total_followers	9.4943E-06	1.2309E-05	0.77131214	0.45176335	-1.66E-05	3.5589E-05	-1.66E-05	3.5589E-05
jobs_glassdoor	-0.6358375	4.90739503	-0.1295672	0.89852371	-11.03905	9.76737526	-11.03905	9.76737526
jobs_linkedin	-0.6402167	4.91847268	-0.1301657	0.89805774	-11.066913	9.78647964	-11.066913	9.78647964
job_posts	1.27490831	9.82612723	0.12974677	0.89838391	-19.555551	22.1053675	-19.555551	22.1053675
customer_reviews	-1.6701348	4.74036259	-0.3523222	0.72919461	-11.719255	8.37898494	-11.719255	8.37898494
stocktwits_mentions	-0.0011951	0.00365051	-0.3273784	0.747626	-0.0089338	0.00654364	-0.0089338	0.00654364
stocktwits_sentiment	2.44936236	3.12047824	0.78493172	0.44396136	-4.165756	9.06448071	-4.165756	9.06448071
stocktwits_subscribers	-0.0055466	0.00701132	-0.7910931	0.44045965	-0.0204099	0.00931673	-0.0204099	0.00931673
tipranks_sentiment	9.84988333	11.908187	0.82715222	0.42031643	-15.394345	35.094112	-15.394345	35.094112
tipranks_score	11.0766753	14.7131523	0.75284175	0.4624789	-20.113814	42.2671647	-20.113814	42.2671647
twitter_followers	-3.318E-05	5.6158E-05	-0.5907779	0.56292222	-0.0001522	8.5873E-05	-0.0001522	8.5873E-05
youtube_subscribers	0.00137943	0.00094053	1.46665295	0.16184961	-0.0006144	0.00337325	-0.0006144	0.00337325
ios_downloads	-0.0044722	0.01581031	-0.2828684	0.7809046	-0.0379886	0.02904412	-0.0379886	0.02904412

RESIDUAL OUTPUT					PROBABILITY OUTPUT					
Observation	*ext Month's Cl*	*Residuals*	*ndard Residuals*		*Percentile*	*lonth's Closing Price*		*squared residuals*	*Mean of squared residuals*	*RMSE*
1	247.555291	8.34221042	0.27999057		1.5625	134.195999		69.5924746	859.977698	29.33
2	268.260516	-12.452419	-0.417942		4.6875	152.970476		155.062748		
3	326.311943	17.0275301	0.57149695		7.8125	165.872272		289.936781		
4	370.284647	34.7343503	1.165793		10.9375	175.79909		1206.47509		
5	353.461575	62.4998534	2.09768978		14.0625	176.163		3906.23168		
6	365.611352	-47.391852	-1.5906182		17.1875	176.955262		2245.98761		
7	314.801485	-76.021049	-2.5515032		20.3125	177.231819		5779.19991		

2.1.5 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.95923448
R Square	0.92013079
Adjusted R Sq	0.83493697
Standard Erro	26.5916343
Observations	32

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	16	122194.454	7637.1534	10.8004402	1.6659E-05
Residual	15	10606.7252	707.115014		
Total	31	132801.18			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	6032.31479	4147.03711	1.45460834	0.16638326	-2806.8856	14871.5152	-2806.8856	14871.5152
android_dow	-0.0435184	0.02665832	-1.6324497	0.1233953	-0.1003392	0.0133025	-0.1003392	0.0133025
android_revie	-0.0415933	0.02960363	-1.4050077	0.18038976	-0.104692	0.02150531	-0.104692	0.02150531
android_ratin	-545.81778	475.2808	-1.1484112	0.26878386	-1558.8548	467.219262	-1558.8548	467.219262
app_rating	993.906	544.84225	1.82420875	0.08810348	-167.39777	2155.20977	-167.39777	2155.20977
app_downloa	0.00277016	0.01478092	0.18741448	0.85384722	-0.0287346	0.03427494	-0.0287346	0.03427494
sentiment	-11.08693	8.32327204	-1.3320398	0.20273325	-28.827564	6.65370474	-28.827564	6.65370474
news_mentio	0.01484316	0.14552178	0.10199956	0.92010806	-0.2953292	0.32501549	-0.2953292	0.32501549
reddit_menti	-0.0027699	0.03358617	-0.0824709	0.93536271	-0.0743571	0.06881734	-0.0743571	0.06881734
reddit_sentim	8.37453192	3.76741311	2.22288655	0.04201677	0.34448096	16.4045829	0.34448096	16.4045829
Open_TSLA	20.8137186	14.8329895	1.40320457	0.18091662	-10.80205	52.4294873	-10.80205	52.4294873
High_TSLA	-23.82918	15.731501	-1.5147429	0.15062339	-57.360081	9.70172089	-57.360081	9.70172089
Low_TSLA	-26.021252	13.0715436	-1.9906794	0.06505741	-53.882588	1.84008341	-53.882588	1.84008341
Close_TSLA	29.8887463	13.206551	2.26317578	0.0388903	1.73964925	58.0378434	1.73964925	58.0378434
Volume_TSLA	6.2642E-09	1.8659E-08	0.33571595	0.74173476	-3.351E-08	4.6036E-08	-3.351E-08	4.6036E-08
EMA_20_TSLA	11.3343633	12.1521469	0.9327046	0.36574362	-14.567325	37.2360513	-14.567325	37.2360513
EMA_50_TSLA	-34.744679	26.2072852	-1.3257641	0.20475471	-90.604185	21.1148273	-90.604185	21.1148273

RESIDUAL OUTPUT					PROBABILITY OUTPUT					
Observation	*ext Month's Cl*	*Residuals*	*ndard Residuals*		*Percentile*	*lonth's Closing Price*		*squared residuals*	*Mean of squared residuals*	*RMSE*
1	260.922169	-5.0246676	-0.2716424		1.5625	134.195999		25.247284	331.460163	18.21
2	257.590423	-1.7823265	-0.0963557		4.6875	152.970476		3.17668789		
3	340.823182	2.51629048	0.13603509		7.8125	165.872272		6.3317178		
4	398.859838	6.15915937	0.332975		10.9375	175.79909		37.9352442		

2.1.6 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.93927115
R Square	0.8822303
Adjusted R Square	0.79717441
Standard Error	29.4768789
Observations	32

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	13	117161.225	9012.40189	10.3723593	7.5563E-06
Residual	18	15639.9551	868.886392		
Total	31	132801.18			

	Coefficients	*tandard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	2972.8678	2957.34499	1.0052489	0.32810021	-3240.2835	9186.01908	-3240.2835	9186.01908
twitter_followers	1.2033E-05	3.0938E-05	0.38894638	0.7018785	-5.296E-05	7.7031E-05	-5.296E-05	7.7031E-05
youtube_subscribers	0.00015015	0.00029481	0.5093026	0.616727	-0.0004692	0.00076952	-0.0004692	0.00076952
ios_downloads	-0.0034918	0.00805477	-0.4335052	0.66979847	-0.0204142	0.01343066	-0.0204142	0.01343066
ios_reviews_total	-0.0603633	0.05417138	-1.1143029	0.27980743	-0.1741732	0.05344652	-0.1741732	0.05344652
ios_rating	560.206446	318.587563	1.75840651	0.09566861	-109.12119	1229.53408	-109.12119	1229.53408
reddit_mentions	0.02863451	0.0336821	0.85014029	0.40641365	-0.042129	0.09939797	-0.042129	0.09939797
reddit_sentiment	6.58218128	3.11728038	2.11151403	0.04897205	0.03301821	13.1313443	0.03301821	13.1313443
High_TSLA	-3.291695	7.77563318	-0.4233347	0.67706618	-19.627694	13.0443041	-19.627694	13.0443041
Low_TSLA	-12.108462	5.54708391	-2.1828518	0.0425356	-23.762453	-0.4544714	-23.762453	-0.4544714
Close_TSLA	15.4585248	9.34749985	1.65376037	0.11550962	-4.1798437	35.0968932	-4.1798437	35.0968932
Volume_TSLA	-5.615E-09	2.071E-08	-0.2711284	0.78937707	-4.912E-08	3.7894E-08	-4.912E-08	3.7894E-08
EMA_20_TSLA	14.8910733	12.5144399	1.18991129	0.24953983	-11.400789	41.1829359	-11.400789	41.1829359
EMA_50_TSLA	-34.656009	25.2467308	-1.3726929	0.18671022	-87.697422	18.3854047	-87.697422	18.3854047

RESIDUAL OUTPUT / PROBABILITY OUTPUT

Observation	*ext Month's Cl*	*Residuals*	*ndard Residuals*	*Percentile*	*Month's Closing Price*	*squared residuals*	*Mean of squared residuals*	*RMSE*
1	257.363886	-1.4663848	-0.0652846	1.5625	134.195999	2.15028446	488.748596	22.11
2	247.428784	8.37931307	0.37305384	4.6875	152.970476	70.2128875		
3	354.067694	-10.728221	-0.4776291	7.8125	165.872272	115.094733		
4	399.543526	5.475471	0.24377243	10.9375	175.79909	29.9807827		
5	384.458492	31.5029366	1.40253638	14.0625	176.163	992.435012		

2.2 Revenue Related

	2009 - 514 employees, 2008 - 363		(for R&D in 2009 considering Daimler contirbution)					Volume of private electric vehicle charging stations and chargers in the United States		
Year	Number of Emploees	Change in Employee headcount	R&D Spending (M)	Revenue (M)	US Annual CO2 emissions	Worl Annual CO2 emissions	"Tesla" Google Trends	EVSE Ports	Station locations	Revenue Next Year
2009	514	41.60%	43.20	111.90	5,480,725,500	31,541,922,000	5.94	161	40	117.00
2010	899	74.90%	93.00	117.00	5,679,715,300	33,355,903,000	5.75	359	89	204.00
2011	1,417	57.62%	208.98	204.00	5,546,116,000	34,504,540,000	5.92	1,759	738	413.00
2012	2,964	109.17%	273.98	413.00	5,344,086,000	35,000,020,000	7.75	3,025	1,491	2,014.00
2013	5,859	97.67%	231.98	2,014.00	5,480,156,700	35,301,655,000	17.17	3,196	1,624	3,198.00
2014	10,161	73.43%	464.70	3,198.00	5,528,681,000	35,450,094,000	18.83	3,697	1,867	4,046.00
2015	13,058	28.51%	717.90	4,046.00	5,376,473,000	35,404,900,000	21.00	4,365	2,010	7,000.00
2016	17,782	36.18%	834.41	7,000.00	5,252,932,000	35,416,658,000	26.50	5,811	2,628	11,759.00
2017	37,543	111.13%	1,378.07	11,759.00	5,212,162,000	35,989,897,000	31.83	6,232	2,651	21,461.00
2018	48,817	30.03%	1,460.37	21,461.00	5,377,797,000	36,730,425,000	38.58	6,863	2,786	24,578.00
2019	48,016	-1.64%	1,343.00	24,578.00	5,262,145,000	37,104,275,000	49.75	9,631	3,080	31,536.00
2020	70,757	47.36%	1,491.00	31,536.00	4,714,628,000	35,126,526,000	62.25	11,040	2,965	53,823.00
2021	99,290	40.33%	2,593.00	53,823.00	5,032,213,000	36,991,734,000	70.67	14,189	3,614	81,462.00
2022	127,855	28.77%	3,075.00	81,462.00	5,078,871,000	37,293,834,000	71.83	14,966	3,789	96,773.00
2023	140,473	9.87%	3,969.00	96,773.00	4,911,391,000	37,791,570,000	74.42	15,780	3,834	97,690.00
2024	125,655	-10.55%	4,540.00	97,690.00	4,890,981,355	38,072,757,867	76.33	16,088	4,333	UNAVAILABLE DA

2.2.1 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99481941
R Square	0.98966565
Adjusted R Square	0.98553192
Standard Error	4317.92549
Observations	15

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	4	1.7855E+10	4463708518	239.411793	7.0115E-10
Residual	10	186444806	18644480.6		
Total	14	1.8041E+10			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-988.29553	3387.19563	-0.291774	0.77642255	-8535.4377	6558.84664	-8535.4377	6558.84664
Number of Emploees	0.99150774	0.17946903	5.52467329	0.00025289	0.59162583	1.39138966	0.59162583	1.39138966
Change in Employee headcount	2134.08891	3956.66384	0.53936574	0.60143631	-6681.9075	10950.0853	-6681.9075	10950.0853
R&D Spending (M)	-11.931675	5.79700405	-2.0582485	0.06658227	-24.848205	0.98485527	-24.848205	0.98485527
Revenue (M)	0.09141225	0.21823756	0.41886579	0.68416878	-0.3948513	0.57767583	-0.3948513	0.57767583

RESIDUAL OUTPUT | PROBABILITY OUTPUT

Observation	*d Revenue Ne*	*Residuals*	*ndard Residuals*	*Percentile*	*venue Next Year*	*squared residuals*	*Mean of squared residuals*	*RMSE*
1	-96.145915	213.145915	0.05840713	3.33333333	117	45431.1809	12429653.7	3,525.57
2	402.610136	-198.61014	-0.054424	10	204	39445.986		
3	-828.50933	1241.50933	0.34020356	16.6666667	413	1541345.41		
4	1049.12333	964.876668	0.26439953	23.3333333	2014	930986.985		
5	4321.55141	-1123.5514	-0.3078802	30	3198	1262367.78		
6	5401.06723	-1355.0672	-0.3713212	36.6666667	4046	1836207.2		
7	4371.36681	2628.63319	0.72030903	43.3333333	7000	6909712.44		
8	8098.72278	3660.27722	1.00300443	50	11759	13397629.3		
9	23239.71	-1778.71	-0.4874095	56.6666667	21461	3163809.27		
10	32592.1341	-8014.1341	-2.1960665	63.3333333	24578	64226345.8		
11	32807.4148	-1271.4148	-0.3483984	70	31536	1616495.65		
12	55271.1998	-1448.1998	-0.3968417	76.6666667	53823	2097282.52		
13	72300.3359	9161.66412	2.51051742	83.3333333	81462	83936089.5		
14	97151.6135	-378.61352	-0.1037493	90	96773	143348.2		
15	99991.8054	-2301.8054	-0.6307503	96.6666667	97690	5298308.26		

2.2.2 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99944953
R Square	0.99889937
Adjusted R Square	0.99691824
Standard Error	1992.82499
Observations	15

ANOVA

	df	SS	MS	F	Significance F
Regression	9	18021422118	2002380235	504.206254	7.4648E-07
Residual	5	19856757.22	3971351.44		
Total	14	18041278876			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	41119.116	32439.29812	1.26757108	0.26077505	-42268.754	124506.9865	-42268.754	124506.9865
Number of Emploees	0.64755739	0.225917634	2.86634282	0.03514462	0.06681762	1.228297157	0.06681762	1.228297157
Change in Employee headcount	4943.46679	2506.775022	1.97204246	0.10564766	-1500.4036	11387.33712	-1500.4036	11387.33712
R&D Spending (M)	-2.8825122	4.06793759	-0.708593	0.5102368	-13.339479	7.574454235	-13.339479	7.574454235
Revenue (M)	-0.1841499	0.211371374	-0.871215	0.42348648	-0.7274973	0.359197512	-0.7274973	0.359197512
US Annual CO2 emissions	1.9354E-05	6.45693E-06	2.99739223	0.03018959	2.7559E-06	3.5952E-05	2.7559E-06	3.5952E-05
Worl Annual CO2 emissions	-4.639E-06	1.38591E-06	-3.3471129	0.02039312	-8.201E-06	-1.0762E-06	-8.201E-06	-1.0762E-06
"Tesla" Google Trends	-517.17132	265.2775217	-1.9495482	0.10873304	-1199.0889	164.746263	-1199.0889	164.746263
EVSE Ports	8.16682443	1.416221951	5.76662749	0.00220316	4.52631001	11.80733885	4.52631001	11.80733885
Station locations	-4.0431957	2.821258876	-1.4331176	0.21126643	-11.295473	3.209081133	-11.295473	3.209081133

RESIDUAL OUTPUT / PROBABILITY OUTPUT

Observation	ed Revenue N	Residuals	ndard Residuals	Percentile	evenue Next Year	squared residuals	Mean of squared residuals	RMSE
1	1201.97887	-1084.97887	-0.9110268	3.33333333	117	1177179.148	1323783.81	1,150.56
2	-93.026487	297.0264873	0.24940493	10	204	88224.73414		
3	-152.94251	565.9425139	0.47520628	16.6666667	413	320290.929		
4	3309.59237	-1295.59237	-1.0878731	23.3333333	2014	1678559.59		
5	1664.92559	1533.074411	1.2872802	30	3198	2350317.151		
6	4861.12599	-815.125988	-0.6844388	36.6666667	4046	664430.3771		
7	4651.12272	2348.877283	1.97228732	43.3333333	7000	5517224.491		
8	11229.8742	529.1258216	0.44429232	50	11759	279974.1351		
9	22426.9887	-965.988727	-0.811114	56.6666667	21461	933134.2205		
10	24581.6363	-3.63627606	-0.0030533	63.3333333	24578	13.22250355		
11	33930.9382	-2394.93824	-2.0109634	70	31536	5735729.186		
12	53456.6231	366.3768753	0.30763653	76.6666667	53823	134232.0147		
13	80539.3798	922.6201859	0.77469866	83.3333333	81462	851228.0075		
14	96523.1784	249.8216425	0.20976833	90	96773	62410.85305		
15	97942.6047	-252.604747	-0.2121052	96.6666667	97690	63809.15797		

2.3 COGS Related

	MATERIALS PRICES										TSLA FINANCIALS	
Year	Lithium carbonate	Cobalt	Nickel	Aluminum Ingnots	Industrial Retial Electricity	COGS (in B $)	Gross Profit Margin	COGS (in B $) Next Year			COGS (billions)	GROSS PROFIT MARGIN
2010	5,180	20.85	21,809	104.00	6.77	0.09	26.50%	0.09			0.09	26.50%
2011	5,180	17.99	22,910	116.10	6.82	0.09	29.90%	0.38			0.09	29.90%
2012	6,060	14.07	17,548	101	6.67	0.38	7.26%	1.56			0.38	7.26%
2013	6,800	12.89	15,032	94.20	6.89	1.56	22.65%	2.32			1.56	22.65%
2014	6,690	14.48	16,893	104.50	7.10	2.32	27.55%	3.12			2.32	27.55%
2015	6,500	13.44	11,863	88.20	6.91	3.12	22.81%	5.4			3.12	22.81%
2016	8,650	12.01	9,595	80.40	6.76	5.4	22.84%	9.54			5.4	22.84%
2017	15,000	26.97	10,410	98.30	6.88	9.54	18.90%	17.42			9.54	18.90%
2018	16,000	37.43	13,144	114.70	6.92	17.42	18.84%	20.51			17.42	18.84%
2019	12,100	16.95	13,914	99.50	6.81	20.51	16.55%	24.91			20.51	16.55%
2020	8,600	15.70	13,787	89.70	6.67	24.91	21.02%	40.22			24.91	21.02%
2021	12,600	24.21	18,465	138.50	7.18	40.22	25.28%	60.61			40.22	25.28%
2022	68,100	30.78	25,834	150.60	8.32	60.61	25.60%	79.11			60.61	25.60%
2023	46,000	17	21,521	130	8.06	79.11	18.25%	80.24			79.11	18.25%

2.3.1 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.911304527
R Square	0.83047594
Adjusted R Square	0.724523403
Standard Error	13.02456889
Observations	14

ANOVA

	df	SS	MS	F	Significance F
Regression	5	6648.32761	1329.66552	7.83818832	0.0059913
Residual	8	1357.11516	169.639395		
Total	13	8005.44277			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	-7.801476829	201.838275	-0.0386521	0.9701149	-473.24137	457.638419	-473.24137	457.638419
Lithium carbonate	0.001021104	0.00078492	1.30090653	0.22950783	-0.0007889	0.00283113	-0.0007889	0.00283113
Cobalt	-1.25996411	0.89024886	-1.4152943	0.19470953	-3.3128817	0.79295345	-3.3128817	0.79295345
Nickel	-0.00236647	0.00138747	-1.7056062	0.12647482	-0.005566	0.00083303	-0.005566	0.00083303
Aluminum Ingnots	1.057090796	0.58655065	1.8022157	0.1091826	-0.2954974	2.40967903	-0.2954974	2.40967903
Industrial Retial Electricity	-5.59566647	33.1984171	-0.1685522	0.8703324	-82.151354	70.9600207	-82.151354	70.9600207

RESIDUAL OUTPUT | PROBABILITY OUTPUT

Observation	Predicted COGS (billions)	Residuals	Standard Residuals	Percentile	COGS (billions)	squared residuals	Mean of squared residuals	RMSE
1	-8.337980887	8.42798089	0.82487258	3.57142857	0.09	71.0308618	96.936797	9.85
2	5.17104796	-5.081048	-0.4972979	10.7142857	0.09	25.8170484		
3	8.574971404	-8.1949714	-0.8020672	17.8571429	0.38	67.1575563		
4	8.352121158	-6.7921212	-0.6647659	25	1.56	46.1329098		
5	11.54540079	-9.2254008	-0.9029185	32.1428571	2.32	85.1080198		
6	8.397696054	-5.2776961	-0.5165445	39.2857143	3.12	27.8540756		
7	10.35601441	-4.9560144	-0.4850605	46.4285714	5.4	24.5620788		
8	14.31273264	-4.7727326	-0.4671221	53.5714286	9.54	22.7789769		
9	12.79714445	4.62285555	0.45245318	60.7142857	17.42	21.3707934		
10	17.34446555	3.16553445	0.30982066	67.8571429	20.51	10.0206084		
11	6.070002493	18.8399975	1.84392888	75	24.91	354.945506		
12	37.09401608	3.12598392	0.30594972	82.1428571	40.22	9.77177547		
13	74.46053396	-13.850534	-1.3555946	89.2857143	60.61	191.837291		
14	59.14183394	19.9681661	1.95434623	96.4285714	79.11	398.727656		

2.3.2 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.99191142
R Square	0.98388826
Adjusted R Square	0.96509123
Standard Error	5.42193885
Observations	14

ANOVA

	df	SS	MS	F	Significance F
Regression	7	10771.193	1538.74186	52.3427502	5.85771E-05
Residual	6	176.3845253	29.3974209		
Total	13	10947.57752			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	231.214321	95.14514549	2.4301221	0.05115394	-1.597463091	464.0261051	-1.5974631	464.026105
Lithium carbonate	0.00092869	0.000402491	2.30736498	0.06048682	-5.61664E-05	0.001913555	-5.617E-05	0.00191356
Cobalt	-0.7326247	0.43295262	-1.692159	0.14156354	-1.79202155	0.326772244	-1.7920216	0.32677224
Nickel	-0.0015109	0.000701197	-2.1547058	0.07462004	-0.00322664	0.000204894	-0.0032266	0.00020489
Aluminum Ingnots	0.70552127	0.295498221	2.38756522	0.0542059	-0.017536825	1.428579374	-0.0175368	1.42857937
Industrial Retial Electricity	-41.534899	15.95450946	-2.6033328	0.04048082	-80.57417684	-2.495620297	-80.574177	-2.4956203
COGS (in B $)	1.08586906	0.149045287	7.28549741	0.00034064	0.721168376	1.450569736	0.72116838	1.45056974
Gross Profit Margin	66.664374	33.41443266	1.99507724	0.09305896	-15.09779728	148.4265453	-15.097797	148.426545

RESIDUAL OUTPUT / PROBABILITY OUTPUT

Observation	Predicted COGS (in B $) Next Year	Residuals	ndard Residuals	Percentile	COGS (in B $) Next Year	squared residuals	Mean of squared residuals	RMSE
1	-2.254154809	2.344154809	0.63639642	3.571428571	0.09	5.49506177	12.492079	3.53
2	6.90433191	-6.52433191	-1.7712403	10.71428571	0.38	42.5669069		
3	-0.506276828	2.066276828	0.56095748	17.85714286	1.56	4.26949993		
4	2.452560549	-0.132560549	-0.0359878	25	2.32	0.0175723		
5	1.010151834	2.109848166	0.57278632	32.14285714	3.12	4.45145928		
6	3.295757786	2.104242214	0.57126441	39.28571429	5.4	4.4278353		
7	12.9897134	-3.449713402	-0.9365359	46.42857143	9.54	11.9005226		
8	16.20906178	1.21093822	0.32874823	53.57142857	17.42	1.46637137		
9	23.76957854	-3.259578545	-0.8849177	60.71428571	20.51	10.6248523		
10	29.66208753	-4.752087525	-1.2901074	67.85714286	24.91	22.5823358		
11	34.17781715	6.04218285	1.64034538	75	40.22	36.5079736		
12	57.30129366	3.308706338	0.89825503	82.14285714	60.61	10.9475376		
13	76.43809101	2.671908989	0.72537586	89.28571429	79.11	7.13909765		
14	83.97998648	-3.739986483	-1.0153399	96.42857143	80.24	13.9874989		

2.4 Other company-specific metrics used to validate results

Date	Averaged Stock Price	Number of Employees	Revenue (10-Q) (M USD)	COGS (10-Q) (M USD)	EBITDA (10-Q) (M USD)	Net Income (10-Q) (M USD)	EPS (10-Q)	R&D Spending (based on last 10-Q)	Next Quarter's EPS
30-Sep-2024	235.16	134,000	25182	20185	4,065	2167	0.62	1,039	0.66
30-Jun-2024	197.88	138,000	25500	20922	2,883	1478	0.42	1,074	0.62
31-Mar-2024	187.99	137,000	21301	17605	2,417	1171	0.34	1,151	0.42
31-Dec-2023	240.80	140,473	25167	20729	3,296	7925	2.26	1,094	0.34
30-Sep-2023	252.11	140,000	23350	19172	2,999	1853	0.53	1,161	2.26
30-Jun-2023	244.39	140,000	24927	20394	3,553	2703	0.78	943	0.53
31-Mar-2023	195.33	140,000	23329	18818	3,710	2518	0.73	771	0.78
31-Dec-2022	151.94	127,855	24318	18541	4,890	3719	1.07	810	0.73
30-Sep-2022	277.00	120,000	21454	16072	4,644	3292	0.95	733	1.07
30-Jun-2022	274.00	115,000	16934	12700	3,386	2259	0.65	667	0.95
31-Mar-2022	289.00	110,000	18756	13296	4,483	3313	0.95	865	0.65
31-Dec-2021	352.26	99,290	17719	12872	3,461	2326	0.68	740	0.95
30-Sep-2021	265.00	94,000	13757	10097	2,765	1618	0.48	611	0.68

Date	Averaged Stock Price	Revenue (10-Q) (M USD)	COGS (10-Q) (M USD)	EBITDA (10-Q) (M USD)	Net Income (10-Q) (M USD)	EPS (10-Q)	Averaged Stock Price Next Quarter
30-Sep-2024	235.16	25182	20185	4,065	2167	0.62	333.62
30-Jun-2024	197.88	25500	20922	2,883	1478	0.42	235.16
31-Mar-2024	187.99	21301	17605	2,417	1171	0.34	197.88
31-Dec-2023	240.80	25167	20729	3,296	7925	2.26	187.99
30-Sep-2023	252.11	23350	19172	2,999	1853	0.53	240.80
30-Jun-2023	244.39	24927	20394	3,553	2703	0.78	252.11
31-Mar-2023	195.33	23329	18818	3,710	2518	0.73	244.39
31-Dec-2022	151.94	24318	18541	4,890	3719	1.07	195.33
30-Sep-2022	277.00	21454	16072	4,644	3292	0.95	151.94
30-Jun-2022	274.00	16934	12700	3,386	2259	0.65	277.00

2.4.1 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.969456726
R Square	0.939846343
Adjusted R Square	0.933284126
Standard Error	28.41384345
Observations	62

ANOVA

	df	SS	MS	F	Significance F
Regression	6	693773.131	115628.8551	143.220854	1.00325E-31
Residual	55	44404.0575	807.3464998		
Total	61	738177.188			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	3.508942824	6.13333829	0.572109781	0.56957873	-8.782541777	15.80042742	-8.7825418	15.8004274
Averaged Stock Price	1.003684691	0.0939933	10.67825738	5.0462E-15	0.8153179	1.192051481	0.8153179	1.19205148
Revenue (10-Q) (M USD)	-0.034508551	0.02612999	-1.320649234	0.19208744	-0.086874222	0.017857119	-0.0868742	0.01785712
COGS (10-Q) (M USD)	0.039995797	0.02795587	1.430675909	0.15818037	-0.016029029	0.096020622	-0.016029	0.09602062
EBITDA (10-Q) (M USD)	0.032551201	0.02364039	1.376931673	0.17411232	-0.014825198	0.079927601	-0.0148252	0.0799276
Net Income (10-Q) (M USD)	-0.022449903	0.01365414	-1.644182345	0.10584326	-0.049813419	0.004913614	-0.0498134	0.00491361
EPS (10-Q)	18.53914137	44.7593427	0.414196015	0.68034149	-71.16058591	108.2388686	-71.160586	108.238869

RESIDUAL OUTPUT / PROBABILITY OUTPUT

Observation	Predicted Averaged Stock Price Next Quarter	Residuals	Standard Residuals	Percentile	Averaged Stock Price Next Quarter	squared residuals	Mean of squared residuals	RMSE
1	273.022216	60.597784	2.246002999	0.806451613	1	3672.09142	716.194476	26.76
2	227.3926682	7.7673318	0.287889249	2.419354839	1.1	60.3314433		
3	219.9417046	-22.061705	-0.817697471	4.032258065	1.2	486.718809		
4	177.0641179	10.9258821	0.40495811	5.64516129	1.3	119.3749		
5	283.4197614	-42.619761	-1.579663572	7.258064516	1.4	1816.44406		
6	273.7119275	-21.601928	-0.800656242	8.870967742	1.5	466.643272		

2.4.2 Regression

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.848159994
R Square	0.719375375
Adjusted R Square	0.677016941
Standard Error	0.247851336
Observations	62

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	8	8.34618846	1.04327356	16.9830494	3.55336E-12
Residual	53	3.25580509	0.06143028		
Total	61	11.6019935			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	-0.0744807	0.05738557	-1.2978994	0.19994216	-0.189581585	0.040620186	-0.189581585	0.04062019
Averaged Stock Price	0.002856759	0.00090984	3.13985162	0.00276264	0.001031853	0.004681665	0.001031853	0.00468166
Number of Employees	-1.01938E-05	5.4146E-06	-1.8826606	0.06524042	-2.1054E-05	6.66454E-07	-2.1054E-05	6.6645E-07
Revenue (10-Q) (M USD)	-0.000624754	0.00030199	-2.0687873	0.04345903	-0.001230471	-1.90379E-05	-0.001230471	-1.904E-05
COGS (10-Q) (M USD)	0.000679985	0.000303	2.24420557	0.02901905	7.22523E-05	0.001287719	7.22523E-05	0.00128772
EBITDA (10-Q) (M USD)	0.000741873	0.00029408	2.52270322	0.01468287	0.000152026	0.001331721	0.000152026	0.00133172
Net Income (10-Q) (M USD)	-0.000144225	0.00012013	-1.200552	0.2352635	-0.00038518	9.67299E-05	-0.00038518	9.673E-05
EPS (10-Q)	0.157982961	0.39058571	0.40447707	0.68748967	-0.625432755	0.941398677	-0.625432755	0.94139868
R&D Spending (based on last 10-Q)	0.001132229	0.00061175	1.85079346	0.06977565	-9.47927E-05	0.002359251	-9.47927E-05	0.00235925

RESIDUAL OUTPUT | PROBABILITY OUTPUT

Observation	*Predicted Next Quarter's EPS*	*Residuals*	*ndard Residuals*		*Percentile*	*Next Quarter's EPS*		*squared residuals*	*Mean of squared residuals*	*RMSE*
1	1.201807901	-0.5418079	-2.3452063		0.806451613	-0.34		0.2935558	0.05251299	0.23
2	0.587518486	0.03248151	0.14059568		2.419354839	-0.28		0.00105505		
3	0.710397314	-0.2903973	-1.2569798		4.032258065	-0.28		0.0843306		
4	0.451634012	-0.111634	-0.4832059		5.64516129	-0.27		0.01246215		
5	1.023154892	1.23684511	5.35366305		7.258064516	-0.27		1.52978582		
6	0.927881736	-0.3978817	-1.7222243		8.870967742	-0.27		0.15830988		