Predicting Trends in Ethical Consumerism

Final Document

ABSTRACT

This project explores the link between public sentiment on corporate sustainability and stock price movements. Using the PRAW API, we scraped comments from the r/wallstreetbets subreddit discussing companies' sustainability efforts. Sentiment analysis was conducted using VADER and TextBlob to classify comments as positive, negative, or neutral. We then compared the sentiment with stock price data to investigate whether public perception of sustainability initiatives affects stock performance. The study aims to understand if positive sentiment towards a company's green efforts influences its stock price.

INITIAL PLAN

Our original project plan aimed to conduct sentiment analysis on consumer discussions about sustainability and ethical products across platforms like Reddit and Google Reviews. The goal was to combine these insights with ESG (Environmental, Social, and Governance) scores and analyze their correlation with companies' revenue performance.

The steps included:

1. **Sentiment Analysis:** Scraping consumer sentiment related to sustainability and ethical products from Reddit and Google.
2. **Combining Sentiment with ESG Scores:** Creating a composite matrix that combines sentiment data with ESG scores to evaluate a company's sustainability efforts.
3. **Revenue Correlation**: Investigating how sentiment and ESG scores correlate with company revenue.

Challenges Encountered:

1. **ESG Data Accessibility:** ESG scores were difficult to obtain, often behind

paywalls or inconsistent across sources, making it challenging to pair them reliably with sentiment data.

2. **Complex Sentiment Scraping:** Scraping meaningful sentiment from multiple sources proved complicated. Google Reviews contained vast data but required complex filtering, and Reddit had unstructured discussions and numerous subreddits, making data collection more time-consuming.
3. **Unclear Revenue Link:** The direct link between sustainability sentiment, ESG scores, and revenue was unclear and would require more complex analysis than anticipated.
4. **Resource Constraints:** The scale of the project, especially in scraping and processing large datasets, became overwhelming with limited resources and time.

Shift in Focus:

Given these challenges, we decided to narrow our focus to r/wallstreetbets on Reddit, specifically analyzing sentiment related to corporate sustainability and comparing it with stock price movements. This streamlined approach allowed us to:

- Simplify data collection and analysis by focusing on one platform.
- Avoid the complexities of combining ESG data and revenue, focusing instead on a clear relationship between sentiment and stock performance.

This shift allowed us to make the project more manageable while still exploring the key question of whether public sentiment on sustainability influences financial outcomes.

DATA COLLECTION

For this project, we used the PRAW (Python Reddit API Wrapper) to scrape data from Reddit. Specifically, we focused on the r/wallstreetbets subreddit, which is a popular trading community. The goal was to gather comments related to sustainability efforts by various companies, such as their actions toward becoming more environmentally friendly and adopting green practices.

To ensure the relevance of the data, we used specific keywords related to sustainability, such as "green," "sustainability," "carbon footprint," "renewable energy," "eco-friendly," and others, to filter and scrape only those comments that discuss sustainability-related topics. We targeted discussions mentioning companies that are part of ongoing sustainability efforts or green initiatives.

This dataset comprises user-generated comments discussing various companies' strategies and actions towards becoming more sustainable and environmentally responsible. By focusing on these targeted discussions, we aim to understand the community's perspectives and sentiment toward corporate sustainability in the financial markets.

METHODOLOGY

Gathered several insights about public perception of sustainability in the context of major corporations. Below are the key points that emerge from the data and the analysis process:

1. **Search Terms and Data Filtering:**
    - The team focused on a combination of sustainability-related keywords (e.g., "eco-friendly", "green energy", "carbon footprint") and company names (e.g., "Apple", "Tesla", "Amazon").
    - By applying these search terms within the r/wallstreetbets subreddit, the data was filtered for comments that discussed both corporate names and

sustainability-related issues, ensuring a focused dataset.

2. **Date Range and Context:**
   - The data was collected over a 1-year period, which allowed the team to observe seasonal trends and identify significant spikes in discussions related to corporate sustainability initiatives or events.
3. **Handling API Rate Limits:**
   - PRAW API had built-in rate limits, and the team utilized a retry mechanism with delays to manage these limits efficiently. This ensured the script could run multiple retries for failed requests and avoid interruptions during data collection.
4. **Sentiment Analysis**
   - VADER (70% weight): Specialized for social media content
   - TextBlob (30% weight): General sentiment analysis
   - Combined into a weighted composite score
   - Weekly aggregation of sentiment scores
5. **Stock Price Analysis**
   - Used yfinance API to fetch historical stock data
   - Aligned stock prices with weekly sentiment scores
   - Calculated correlation between sentiment and price movements

REASON FOR CHOOSING VADER TEXTBLOB FOR SENTIMENT ANALYSIS AND OUR EFFORTS ALONG THE WAY

Initially, we experimented with advanced sentiment analysis techniques, including RNNs, LSTMs, and transformer-based models like BERT. While these methods are powerful, they presented significant challenges for our dataset. These models require substantial computational resources, which exceeded our project's hardware constraints and timelines. Furthermore, transformer-based models often struggle with informal social media text, such as Reddit comments, which include slang, emojis, and sarcasm. Fine-tuning these models for our domain would have required additional labeled data that we did not have. Additionally, their "black-box" nature made it difficult to explain specific sentiment scores, which was critical for our project's interpretability and transparency.

After evaluating simpler approaches, we chose a weighted average of VADER and TextBlob sentiment scores. VADER, designed for social media text, excelled at handling slang, emojis, and short-form expressions, while TextBlob added a secondary perspective for more nuanced sentiments. Through manual testing, we observed that VADER was more accurate overall, so we assigned it a 70% weight, with TextBlob contributing the remaining 30%. This approach provided a balance between efficiency, accuracy, and explainability, outperforming individual methods while staying computationally lightweight. By combining these complementary tools, we achieved reliable results without the complexities of advanced machine learning models.

In the script we're using, the comment score represents the net total of upvotes minus downvotes that a Reddit comment has received. This score reflects how the Reddit community perceives the value or relevance of that comment. A higher score indicates that more users found the comment valuable or agreed with its content, while a lower or negative score suggests the opposite.

While the comment score provides insight into community engagement, it doesn't directly indicate the sentiment (positive, negative, or neutral) expressed in the comment's text. For instance, a comment with a positive sentiment could receive a low score if the community disagrees with the viewpoint, and vice versa. Therefore, relying solely on comment scores for sentiment analysis can be misleading.

To accurately assess sentiment, it's advisable to use natural language processing (NLP) tools designed for this purpose. Libraries such as VADER (Valence Aware Dictionary and sEntiment Reasoner) are specifically tailored for analyzing sentiments in social media text, including Reddit comments. These tools evaluate the actual content of the text to determine sentiment, providing a more precise analysis than comment scores alone.

In summary, while comment scores offer context about community reception, they should not be used as a substitute for dedicated sentiment analysis tools when determining the sentiment expressed in Reddit comments.

PROJECT PROGRESS SUMMARY

So far, we have made significant progress in our project, which explores the relationship between public sentiment on corporate sustainability and stock price movements. We successfully scraped comments from the r/wallstreetbets subreddit up until 2021 using High-Performance Computing (HPC). Sentiment analysis was conducted using VADER and TextBlob to classify comments as positive, negative, or neutral, and we correlated these sentiments with stock price data for the companies involved. Our preliminary findings show a correlation between public sentiment around sustainability efforts and stock price performance.

Team Contributions:

- **All of us:** Conducted in-depth research and literature review to understand the current landscape of sustainability sentiment analysis and its link to stock performance. Through this research, we identified the gap in existing studies and proposed the idea to investigate whether positive sentiment towards a company's sustainability efforts impacts its stock price.

- **Rishabh and Saurav:** Worked collaboratively on developing the code to scrape relevant Reddit data focused on sustainability discussions about a set of companies. They ensured the data collection was efficient and that the content captured was relevant to our study.

- **Saurav and Dhruv:** Collaboratively worked on the correlation analysis to assess the relationship between stock price movements and sentiment. This involved coding and refining the logic to calculate and visualize the correlation effectively.

The project report represents the collective effort and contribution of all team members.

CODE SNIPPETS AND RESULTS

Data Collection

```python
import praw
import pandas as pd
import re
from datetime import datetime, timedelta
import time
from prawcore.exceptions import TooManyRequests, ResponseException

# Define the companies and keywords
companies = [
"Apple", "Microsoft", "Amazon", "Google", "Meta", "Netflix",
"Tesla", "NVIDIA", "Adobe", "Salesforce", "Intel", "Cisco",
"Nike", "Adidas", "Lululemon", "Under Armour", "Gap", "H&M",
"Unilever", "Procter & Gamble", "Coca-Cola", "PepsiCo",
"Johnson & Johnson", "Walmart",  "Costco", "Shein","Zara"
]
keywords = [
"sustainability", "sustainable", "eco-friendly", "environmental",
"climate change", "carbon footprint", "renewable energy", "clean energy",
"recycling", "biodegradable", "zero waste",
"ethical sourcing", "fair trade", "green technology", "conservation",
"emissions reduction", "energy efficiency", "sustainable fashion",
"plant-based", "organic", "regenerative agriculture", "ESG",
"corporate responsibility", "carbon neutral", "net zero",
"sustainable packaging", "water conservation", "renewable materials",
"ethical investing", "green bonds", "sustainability reporting"
]


# Enhanced function to fetch data
def fetch_reddit_data(subreddit_name, keywords, companies, start_date, end_date, max_retries=3, delay=60):
    data = []
    reddit = praw.Reddit(
        client_id="QkGeYF6uveg5G5iZYz7D_Q",             # Replace with your Client ID
        client_secret="QdyJJqaWVimuYvmUvVl5RiM3U78jPQ",  # Replace with your Client Secret
        user_agent="Rishabh personal use script"         # Replace with your User Agent
    )
    subreddit = reddit.subreddit(subreddit_name)

    search_terms = keywords + companies

    for term in search_terms:
        print(f"Searching for '{term}' in r/{subreddit_name}")
        retries = 0
        while retries < max_retries:
            try:
                for submission in subreddit.search(term, sort="new", time_filter="year"):
                    if start_date <= datetime.fromtimestamp(submission.created_utc) <= end_date:
                        submission.comments.replace_more(limit=None)
                        for comment in submission.comments.list():
                            comment_text = comment.body.lower()
                            if any(kw.lower() in comment_text for kw in keywords) and any(company.lower() in comment_text for company in companies):
                                matching_keyword = next(kw for kw in keywords if kw.lower() in comment_text)
                                matching_company = next(company for company in companies if company.lower() in comment_text)
                                data.append({
                                    "keyword": matching_keyword,
                                    "company_name": matching_company,
                                    "post_title": submission.title,
                                    "post_id": submission.id,
                                    "comment_id": comment.id,
                                    "comment_body": comment.body,
                                    "comment_score": comment.score,
                                    "timestamp": datetime.fromtimestamp(comment.created_utc)
                                })
                break  # Break the retry loop if successful
            except (TooManyRequests, ResponseException) as e:
                retries += 1
                print(f"Rate limit hit. Retrying in {delay} seconds... (Attempt {retries}/{max_retries})")
                time.sleep(delay)
        else:
            print(f"Failed to fetch data for term '{term}' after {max_retries} attempts.")
```

```python
        time.sleep(2)  # Add a small delay between terms to avoid hitting rate limits

    return pd.DataFrame(data)

# Set date range for past year
end_date = datetime.now()
start_date = end_date - timedelta(days=3654)

# Fetch data from r/wallstreetbets
subreddit_name = "wallstreetbets"
sustainability_data = fetch_reddit_data(subreddit_name, keywords, companies, start_date, end_date)

# Save results to CSV
csv_filename = f"{subreddit_name}_sustainability_data_with_companies.csv"
sustainability_data.to_csv(csv_filename, index=False)

print(f"Data collection complete. {len(sustainability_data)} items saved to {csv_filename}")
```

Correlation

```python
import pandas as pd
import numpy as np
from textblob import TextBlob
from vaderSentiment.vaderSentiment import SentimentIntensityAnalyzer
import yfinance as yf
import matplotlib.pyplot as plt
import seaborn as sns
from datetime import datetime, timedelta

# Load the dataset
df = pd.read_csv('/content/wallstreetbets_sustainability_data_with_companies_combined.csv')

# Convert timestamp to datetime
df['timestamp'] = pd.to_datetime(df['timestamp'])

# Function to get TextBlob sentiment
def get_textblob_sentiment(text):
    return TextBlob(str(text)).sentiment.polarity

# Function to get VADER sentiment
def get_vader_sentiment(text):
    sid = SentimentIntensityAnalyzer()
    return sid.polarity_scores(str(text))['compound']

# Apply sentiment analysis
df['textblob_sentiment'] = df['comment_body'].apply(get_textblob_sentiment)
df['vader_sentiment'] = df['comment_body'].apply(get_vader_sentiment)

df = df.rename(columns={'company_name': 'company'})
# Group by company and week
df['week'] = df['timestamp'].dt.to_period('W')
weekly_sentiment = df.groupby(['company', 'week']).agg({
    'textblob_sentiment': 'mean',
    'vader_sentiment': 'mean'
}).reset_index()

# Function to fetch stock data
def get_stock_data(ticker, start_date, end_date):
    stock = yf.Ticker(ticker)
    data = stock.history(start=start_date, end=end_date)
    return data['Close']

# Dictionary mapping companies to their stock tickers
company_tickers = {
    'Tesla': 'TSLA',
```

```python
    'Apple': 'AAPL',
    'Microsoft': 'MSFT',
    'Amazon': 'AMZN',
    'Google': 'GOOGL',
    'Meta': 'META',
    'Netflix': 'NFLX',
    'NVIDIA': 'NVDA',
    'Adobe': 'ADBE',
    'Salesforce': 'CRM',
    'Intel': 'INTC',
    'Cisco': 'CSCO',
    'Nike': 'NKE',
    'Adidas': 'ADDYY',
    'Lululemon': 'LULU',
    'Under Armour': 'UAA',
    'Gap': 'GPS',
    'Walmart': 'WMT',
    'Target': 'TGT',
    'Costco': 'COST',
    'Coca-Cola': 'KO',
    'PepsiCo': 'PEP',
    'Johnson & Johnson': 'JNJ',
    'Procter & Gamble': 'PG',
    'Unilever': 'UL'
}

# First, modify the sentiment calculation to use a weighted average
# VADER is generally more accurate for social media text, so we'll give it more weight
# Calculate composite sentiment first
df['composite_sentiment'] = (0.7 * df['vader_sentiment'] + 0.3 * df['textblob_sentiment'])

# Then do the weekly grouping once
df['week'] = df['timestamp'].dt.to_period('W')
weekly_sentiment = df.groupby(['company', 'week']).agg({
    'composite_sentiment': 'mean'
}).reset_index()

# Then proceed with the plotting function

def plot_sentiment_and_stock(company, ticker):
    company_sentiment = weekly_sentiment[weekly_sentiment['company'] == company]
    if company_sentiment.empty:
        print(f"No data available for {company}")
        return

    start_date = company_sentiment['week'].min().start_time
    end_date = company_sentiment['week'].max().end_time
```

```python
    stock_prices = get_stock_data(ticker, start_date, end_date)
    if stock_prices.empty:
        print(f"No stock data available for {company}")
        return

    # Create the plot
    fig, ax1 = plt.subplots(figsize=(12, 6))

    # Plot sentiment
    ax1.set_xlabel('Date')
    ax1.set_ylabel('Sentiment Score', color='tab:blue')
    ax1.plot(company_sentiment['week'].dt.start_time,
             company_sentiment['composite_sentiment'],
             color='tab:blue',
             label='Sentiment Score',
             linewidth=2)
    ax1.tick_params(axis='y', labelcolor='tab:blue')
    ax1.grid(True, alpha=0.3)

    # Plot stock price
    ax2 = ax1.twinx()
    ax2.set_ylabel('Stock Price ($)', color='tab:green')
    ax2.plot(stock_prices.index,
            stock_prices.values,
            color='tab:green',
            label='Stock Price',
            linewidth=2)
    ax2.tick_params(axis='y', labelcolor='tab:green')

    # Calculate correlation
    stock_prices_series = pd.Series(stock_prices.values,
                                    index=pd.PeriodIndex(stock_prices.index.date, freq='W'),
                                    name='stock_price')
    sentiment_series = company_sentiment.set_index('week')['composite_sentiment']
    sentiment_series.name = 'sentiment'

    common_dates = pd.merge(sentiment_series,
                           stock_prices_series,
                           left_index=True,
                           right_index=True,
                           how='inner')

    correlation = common_dates['sentiment'].corr(common_dates['stock_price'])

    plt.title(f'{company} Sentiment vs Stock Price (Correlation: {correlation:.2f})')

    # Combine legends from both axes
    lines1, labels1 = ax1.get_legend_handles_labels()
    lines2, labels2 = ax2.get_legend_handles_labels()
    ax1.legend(lines1 + lines2, labels1 + labels2, loc='upper left')

    fig.tight_layout()
    plt.show()


for company, ticker in company_tickers.items():
    if company in df['company'].unique():
        print(f"Plotting for {company}")
        plot_sentiment_and_stock(company, ticker)
```

UNDERSTANDING AND INSIGHTS

Our analysis of public sentiment on sustainability and its relationship with stock price movements has revealed interesting insights, showing varying degrees of correlation across different companies. These findings suggest that the influence of sustainability sentiment on stock prices is not uniform and can be influenced by multiple factors, such as industry, company size, market conditions, and how strongly a company is perceived as committing to sustainability initiatives.

Here are some key findings from our analysis:

Microsoft: We found a positive correlation of 0.55, indicating a moderate to strong relationship between positive sentiment around its sustainability efforts and stock price growth. This suggests that public perception of Microsoft's commitment to sustainability is positively aligned with investor confidence and stock performance.

NVIDIA: In contrast, NVIDIA showed a negative correlation of -0.07, indicating a negative relationship between sustainability sentiment and stock price movements. This suggests that while sustainability discussions exist, they do not significantly influence NVIDIA's stock performance, possibly because other factors, like technological innovation or market dominance, might have a greater impact on investor sentiment.

Google: Google displayed a negative correlation of -0.21, suggesting that negative sentiment around its sustainability efforts could be associated with a decline in stock prices. This could be due to public skepticism about Google's environmental initiatives or perhaps negative publicity surrounding corporate practices unrelated to sustainability.

Meta: For Meta (formerly Facebook), we found a very weak positive correlation of 0.15, suggesting almost no substantial link between sustainability sentiment and stock performance. This might be reflective of Meta's brand and public perception, where issues like privacy, data security, and social media influence may overshadow sustainability efforts in influencing stock prices.

Key Insights

Industry and Market Dynamics:

The relationship between sustainability sentiment and stock price often varies by industry. In tech companies like Microsoft, factors such as innovation and market position tend to have a stronger influence on stock prices than sustainability sentiment. In contrast, industries like energy, manufacturing, or consumer goods—where environmental impact is more closely tied to business operations—may exhibit stronger correlations.

Perceived Importance of Sustainability:

Companies with a visible and proactive approach to sustainability, tend to show more pronounced correlations between sentiment and stock prices. On the other hand, companies like NVIDIA and Meta, where sustainability is less emphasized or overshadowed by other priorities, show weaker or negligible correlations.

Investor Priorities:

Investors often prioritize growth potential, market leadership, and technological innovation over sustainability efforts. For example, NVIDIA and Meta show negative and weak correlations because their stock prices are likely influenced by factors like product innovation and user engagement rather than sustainability sentiment.

Additional Considerations

Event-Driven Sentiment:

Shifts in sentiment can be influenced by specific events, such as major announcements or sustainability initiatives. These events may lead to short-term fluctuations in stock prices, emphasizing the need to consider timing in sentiment analysis.

Public vs. Investor Sentiment:

Public sentiment, often derived from platforms like Reddit, may not align with the perspectives of institutional investors who hold greater sway over stock prices. While public discussions can shape brand reputation, institutional decisions are often driven by broader financial metrics and strategic priorities.

Limitations and Scope:

Correlation between sentiment and stock prices is one of many factors affecting stock performance. Broader market conditions, corporate earnings, and macroeconomic trends also play significant roles. Sentiment analysis should therefore be used as a complementary tool alongside traditional financial analysis to provide a holistic view of stock performance drivers.

Conclusion:

Our analysis shows that sustainability sentiment can influence stock prices, but the degree of this influence varies significantly across companies. Companies in industries where sustainability is more directly linked to their operations, like energy or consumer goods, may see a stronger correlation, while those in tech or other sectors may see weaker or even inverse relationships. Factors like market conditions, investor priorities, and the perceived importance of sustainability initiatives play a crucial role in shaping these correlations. Future work could explore sector-specific trends and whether these patterns hold across different time periods or market conditions.

PLAN FOR FUTURE

Building on the insights we've gathered about public sentiment toward sustainability and its potential impact on stock prices, we plan to focus on several key areas for future work to refine our methods and expand our analysis.

1. Improving Data Collection

- Faster and More Efficient Scraping: Currently, we use the PRAW API for scraping Reddit data, but we aim to optimize this process by implementing parallel data collection and exploring other APIs, such as Pushshift, to gather larger datasets more efficiently.
- Incremental Scraping: We plan to continuously scrape new data, ensuring that our analysis remains up-to-date and reflects the latest public sentiment and sustainability trends.

2. Enhancing Sentiment Analysis

- Advanced NLP Models: We will experiment with more advanced models like BERT or RoBERTa for sentiment analysis, which can better understand the context of discussions, especially those around financial topics like sustainability.
- Aspect-Based Sentiment Analysis: We aim to separate sentiments related specifically to sustainability from general opinions about companies, providing more precise insights into how sustainability initiatives affect public perception.

3. Refining Correlation and Causality

- Time-Series and Lagged Effects: We will explore time-series analysis and account for lagged effects to better understand the delayed impact of sentiment shifts on stock prices.
- Event-Driven Analysis: We'll also focus on specific events (e.g., sustainability announcements, ESG ratings) and analyze their direct effect

on stock price movements, as well as sentiment shifts before and after these events.

4. Broader Scope and Industry-Level Analysis

- Expanding Company Set: We plan to include more companies, particularly smaller firms or startups focused on sustainability, to broaden the scope of our study and compare sentiment impacts across different company sizes and industries.
- Industry-Specific Sentiment: Analyzing sentiment trends within specific industries (e.g., energy, tech) will help us understand how sustainability perceptions differ across sectors.

5. Machine Learning for Improved Insights

- Predictive Models: We will implement machine learning models, such as Random Forests or Gradient Boosting, to predict stock price movements based on sentiment scores, improving the accuracy and depth of our analysis.
- Clustering Sentiment: Using unsupervised learning, we will cluster sentiment into distinct groups and examine how each group correlates with stock price trends, offering more granular insights.

6. Visualization and Reporting

- Interactive Dashboards: We plan to create interactive dashboards for visualizing sentiment trends, stock price correlations, and ESG data, making the results more accessible and useful for stakeholders.
- Comprehensive Reporting: We will generate detailed reports that provide actionable insights for investors, corporate sustainability teams, and other stakeholders.

By improving our scraping methods, enhancing sentiment analysis, and using machine learning to predict stock price movements, we aim to create a more sophisticated framework for understanding the link between public sentiment on sustainability and stock performance.

CONCLUSION AND WRAP-UP

This project aimed to explore the relationship between public sentiment regarding corporate sustainability and stock price movements, with a focus on understanding how consumer opinions, as reflected in online discussions, can potentially influence financial performance. By analyzing sentiment from the r/wallstreetbets subreddit and comparing it with stock price data for major companies, we have gained valuable insights into the role of sustainability sentiment in shaping investor behavior.

Key Findings:

- **Varying Correlation Across Companies:** The analysis revealed that the correlation between sustainability sentiment and stock price movements differs significantly across companies. While companies like Microsoft showed a moderate positive correlation (0.55), others like NVIDIA and Meta showed weak or even negative correlations, suggesting that sustainability sentiment is not always a major factor in stock price performance for these firms.
- **Industry-Specific Trends:** The correlation also appears to be influenced by the company's industry and market position. In the tech sector, for instance, other factors such as technological innovation, product quality, and market leadership seem to have a stronger impact on stock prices than sustainability sentiment alone.
- **Public vs. Investor Sentiment:** Our study focused on public sentiment from online discussions, which may not fully reflect the priorities of institutional investors. While public sentiment can shape brand value and consumer trust, stock prices are often more heavily influenced by institutional investors, whose decision-making processes might differ from those of the general public.

Challenges and Limitations:

The project faced several challenges, including the difficulty of efficiently scraping large amounts of data, managing rate limits from the PRAW API, and dealing with unstructured and noisy data. Additionally, correlating sentiment with stock prices is inherently complex, as stock price movements are influenced by a wide array of factors beyond sustainability, such as market conditions, financial performance, and broader economic trends.

Future Directions:

In future work, we plan to refine our data scraping techniques to improve efficiency and scale, explore more advanced sentiment analysis models, and further investigate how sentiment shifts in response to specific sustainability-related events or announcements. Additionally, we hope to expand the scope of our study to include a broader range of companies and industries, providing a more comprehensive view of how sustainability sentiment affects stock performance across different sectors.

Final Thoughts:

While the results of this project indicate that sustainability sentiment does have an influence on stock prices, the relationship is complex and varies across different companies and industries. This project has provided valuable insights into the potential link between public opinion on sustainability and financial outcomes, and it opens the door for further research into how sustainability initiatives can shape corporate performance and investor behavior in the future.

By narrowing our focus to sustainability sentiment and stock price movements, we were able to uncover meaningful patterns that can help inform investors, companies, and stakeholders looking to understand the growing intersection of sustainability and finance.