Quantifying Market Sentiment and Its Impact on Stock Prices

An In-depth Analysis

MADE Project

Mohammad Ahmed

Agenda

01. Introduction

Exploring the relationship between market sentiments and stock prices.

02. Motivation

Influence of Social Media and News Sentiments

03. Methodology

Data Pipeline, Exploratory & Statistical analysis

04. Data Sources

Stock Price Data & News Datasets

05. In-dept Analysis

Financial Metrics and Technical Indicators

Introduction

- Exploring the relationship between market sentiments and stock prices.
- Seeks to determine the extent to which social media and news sentiments affect stock prices.

Motivation

- Recognizes the growing impact of sentiments on social media and news platforms in shaping stock market dynamics.
- Highlights the need to understand the correlation between market sentiments and stock prices for effective investment strategies in a digitally interconnected financial landscape.

Methodology – Pipeline

Kaggle Dataset Download

Utilizes the Kaggle API to download datasets related to stock exchange data and news articles.

Dataset Preprocessing

Renames columns, handles datetime formats, and filters data to focus on the relevant time period (e.g., 2021).

SQLite Migration

Creates an SQLite engine and migrates data to appropriate tables with specified rules

Overall Pipeline

Calls functions sequentially to ensure a streamlined process from dataset download to SQLite migration.

Methodology - Analysis

SQLite Data Retrieval

Executes SQL queries or utilizes Pandas to fetch relevant datasets for analysis.

Exploratory Data Analysis (EDA)

Creates visualizations and summary statistics to identify patterns and trends.

Statistical Analysis and Modeling

Applies modeling techniques to further explore correlations and relationships within the data.

Real-World Trading Scenarios

Explores the practical applicability of the insights gained in real-world

Data Sources

Stock Price Data

We sourced daily price data for global stock exchanges, including indexes from the United States, China, Canada, Germany, Japan, and more. This data was obtained from Kaggle, and it encompasses several decades of historical data from Yahoo Finance.

News Data

Two datasets from Kaggle were utilized for news sentiment analysis. The first dataset contains news headlines published over nineteen years by the Australian Broadcasting Corporation (ABC). The second dataset collects RSS feeds from BBC News using a self-updating mechanism.

- ABC News Data
- BBC News Data


SUCCESS
100.000.0
100,000.
PRODUCTION
PRODUCT
SUCCESS
RESULT
PROCCESS
TEAM
?
MEDIA
WEB
70000
community
GROUP
IDEA
WORKSTATION
IDEA

Analysis

Initial Configuration

- Generating a list of Stock Index
- Stock Index selection

	stock
0	NYA
1	IXIC
2	HSI
3	000001.SS
4	GSPTSE

```
Select a stock index for analysis (or type "exit" to quit): NYA
Data filtered for stock index: NYA
```

Derived Columns for Advanced Analysis

Daily Returns

- This column represents the percentage change in the stock's closing price from one day to the next.
- Formula:
  - Daily Return = $(\text{Close}_t - \text{Close}_{t-1}) / \text{Close}_{t-1}$

Log Return

- Logarithmic returns are often used in financial analysis and can be more appropriate for certain calculations.
- Formula:
  - Log Return = $\ln(\text{Close}_t / \text{Close}_{t-1})$

Moving Averages

- Moving averages smooth out price data to identify trends over a specified period.
- Common types include the Simple Moving Average (SMA) and Exponential Moving Aveo4 log returns.

Derived Columns for Advanced Analysis

Relative Strength Index (RSI)

- RSI is a momentum oscillator that measures the speed and change of price movements.
- It ranges from 0 to 100 and is used to identify overbought or ov5rsold conditions.

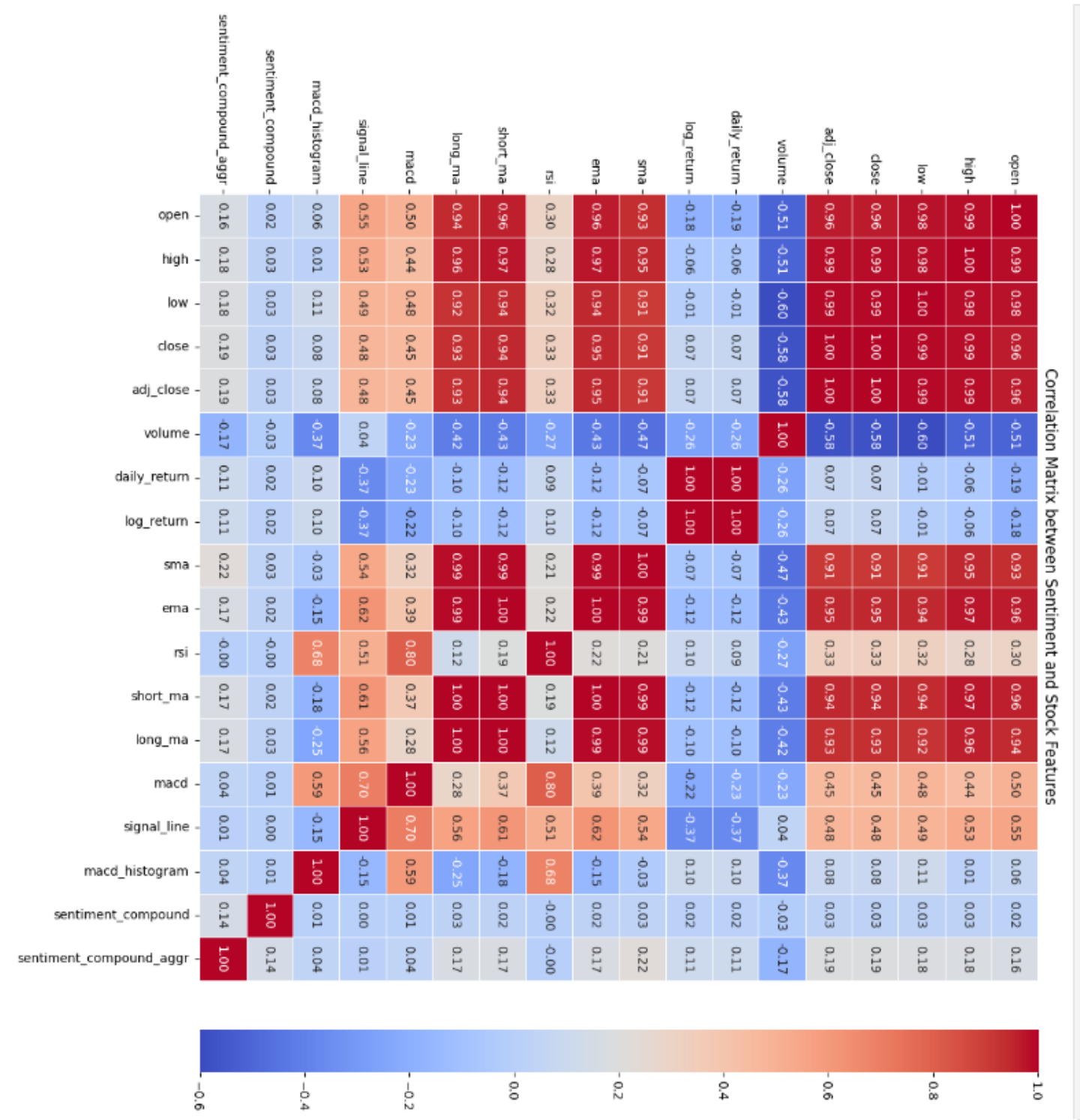
Moving Average Convergence Divergence (MACD)

- MACD is a trend-following momentum indicator that shows the relationship between two moving averages.
- It consists of the MACD line, sigcating overbought or oversold conditions.

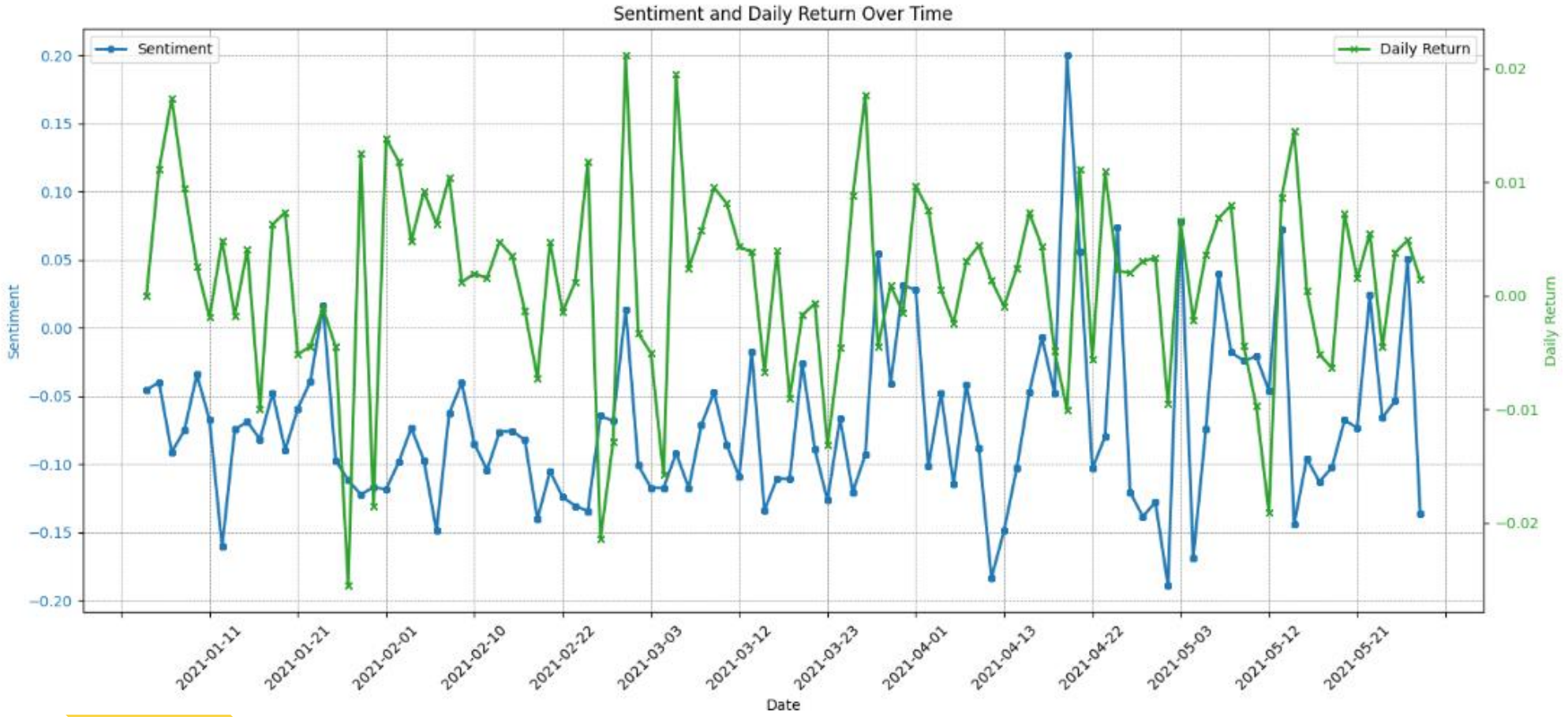
Aggregate Sentiment Scores

- Sentiment scores are aggregated by date to provide a holistic view of market sentiment over time.

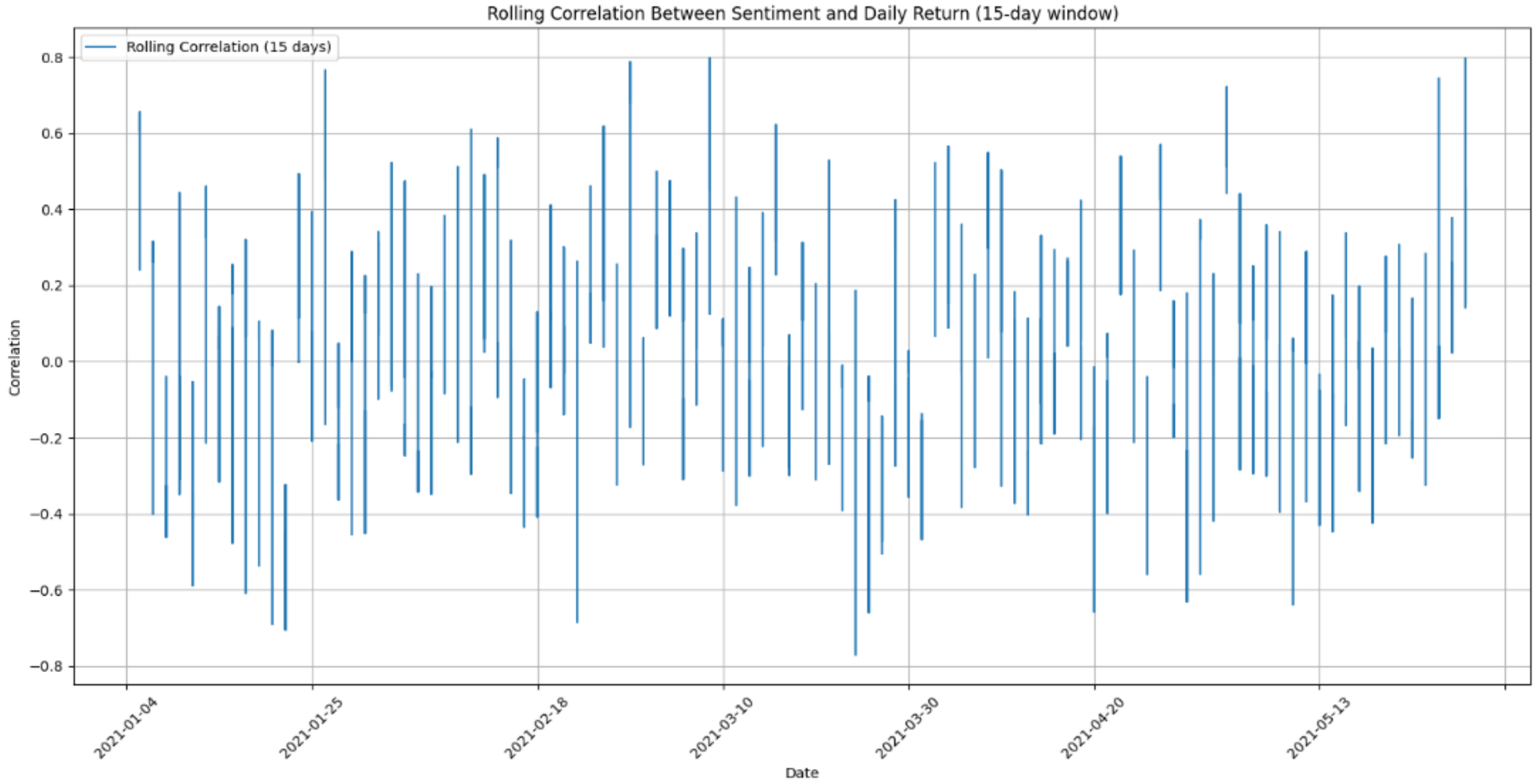
Correlation Matrix between Sentiment and Stock Features



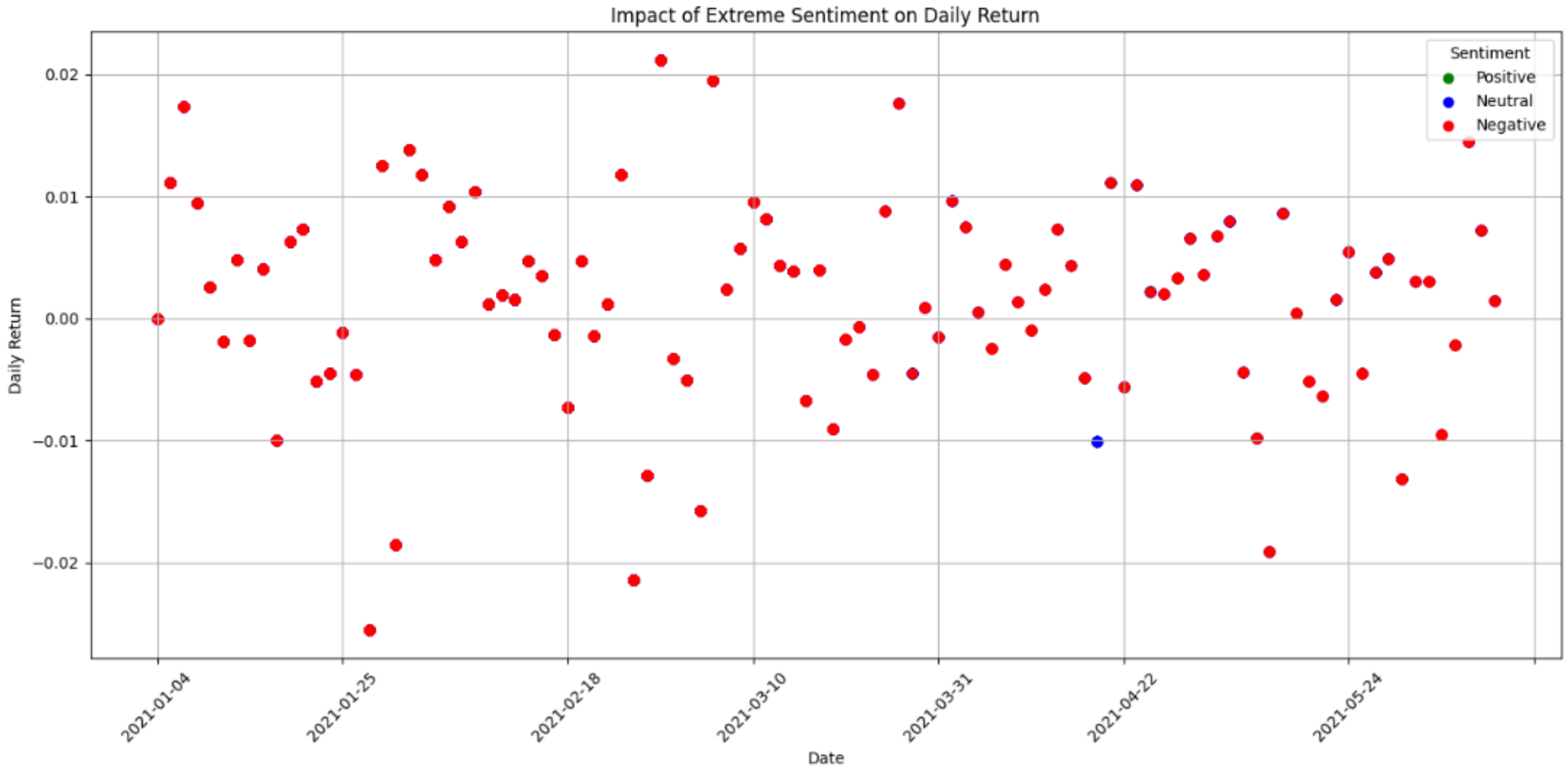
Sentiment and Daily Return Over Time



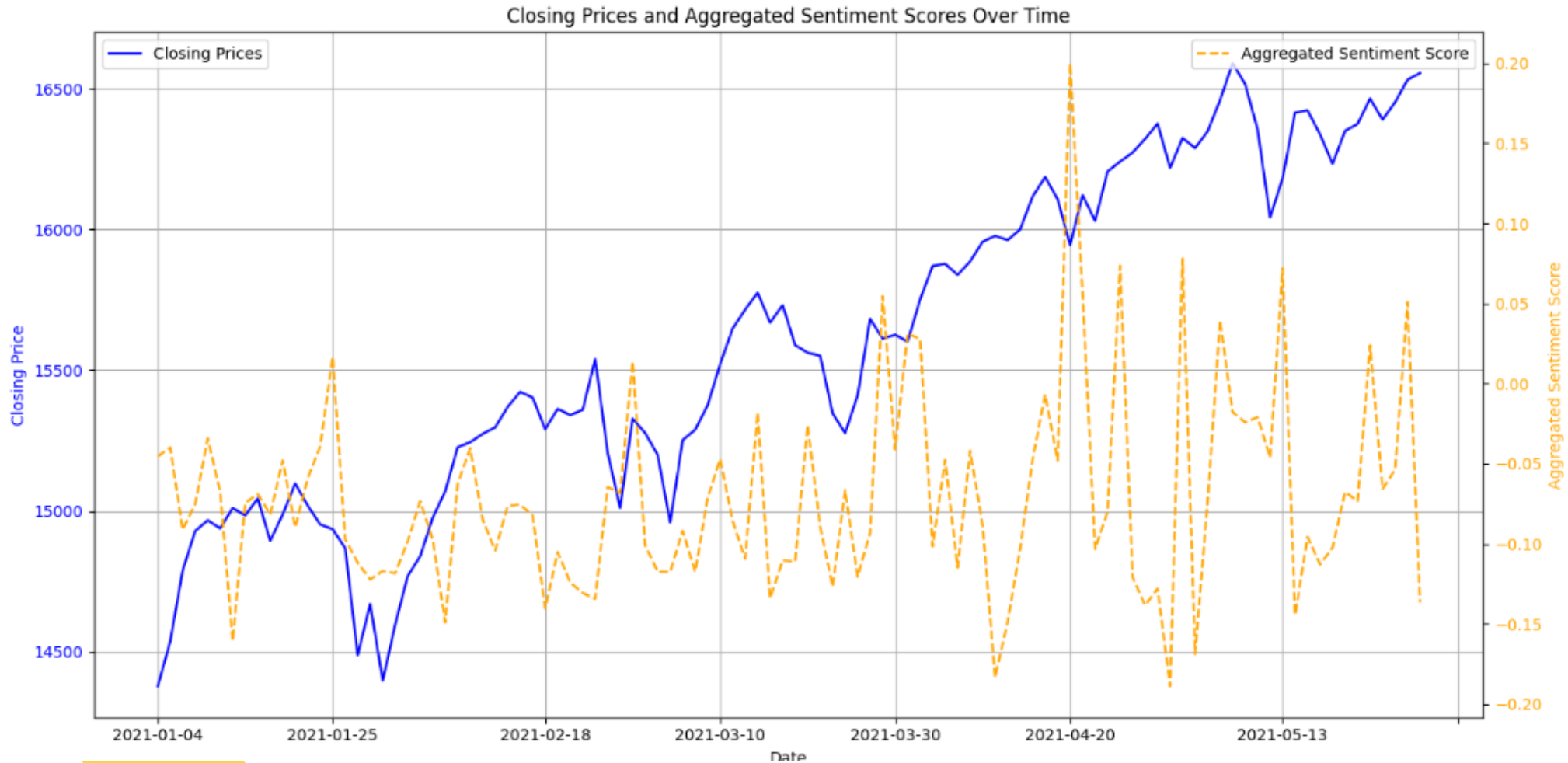
Rolling Correlation Between Sentiment and Daily Return



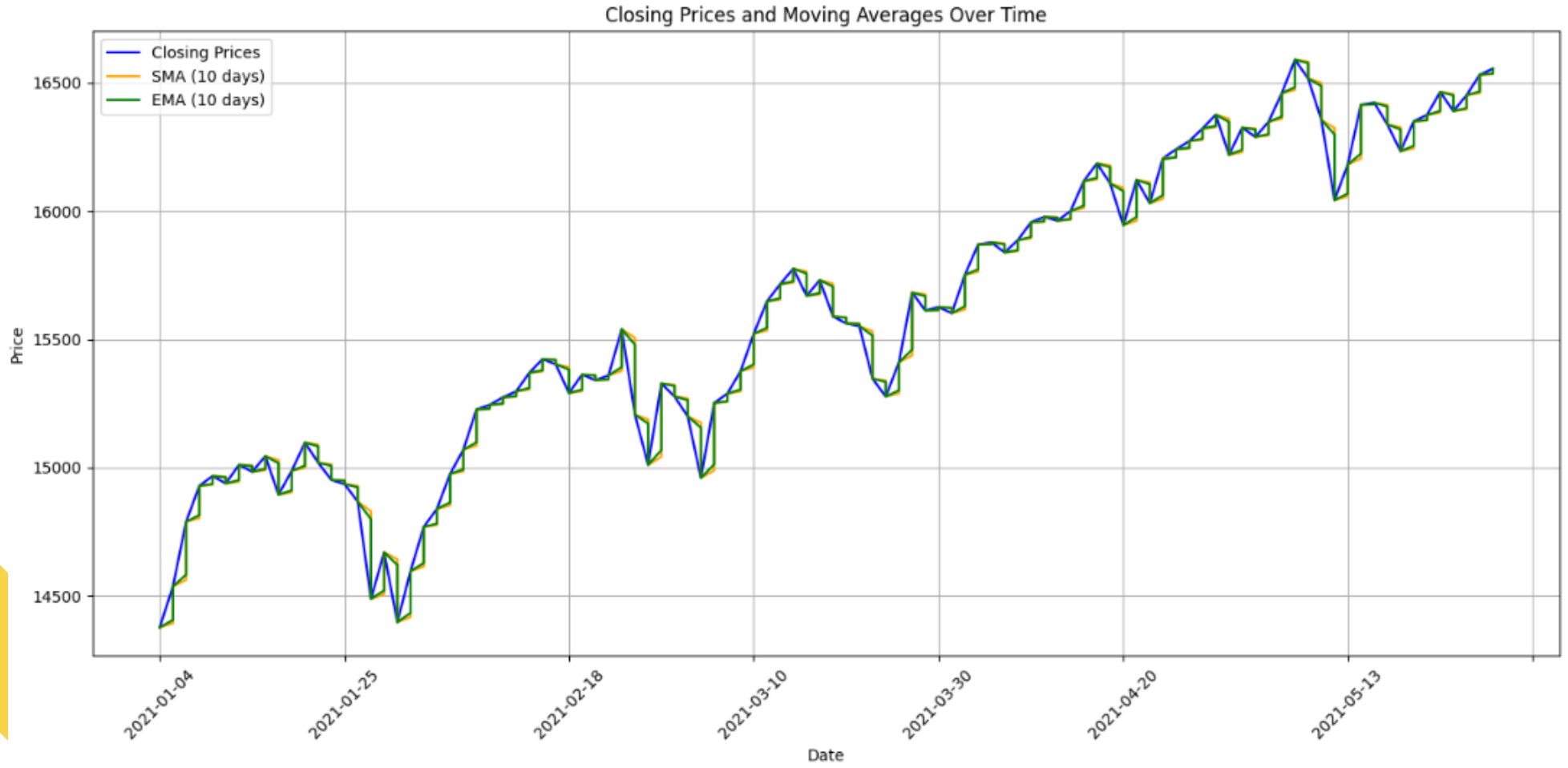
Impact of Extreme Sentiment on Daily Return



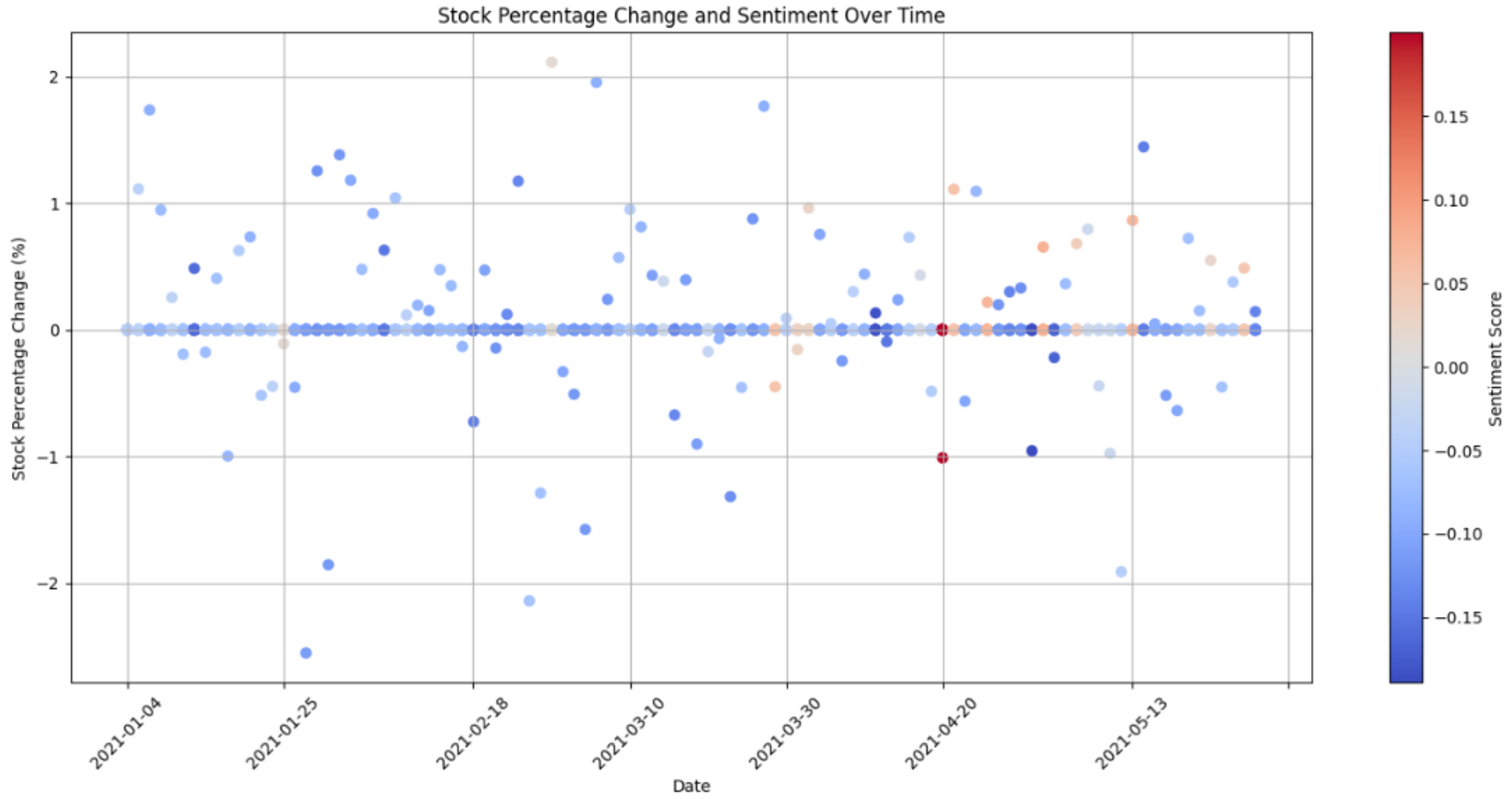
Closing Prices and Aggregated Sentiment Scores Over Time



Closing Prices and Moving Averages Over Time



Stock Percentage Change and Sentiment Over Time



OLS Regression Results

Dep. Variable:	daily_return	R-squared:	0.011
Model:	OLS	Adj. R-squared:	0.011
Method:	Least Squares	F-statistic:	95.61
Date:	Wed, 10 Jan 2024	Prob (F-statistic):	1.85e-22
Time:	08:54:39	Log-Likelihood:	27508.
No. Observations:	8391	AIC:	-5.501e+04
Df Residuals:	8389	BIC:	-5.500e+04
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
const	0.0028	0.000	14.882	0.000	0.002	0.003
sentiment_compound_aggr	0.0195	0.002	9.778	0.000	0.016	0.023

Omnibus:	474.314	Durbin-Watson:	0.022
Prob(Omnibus):	0.000	Jarque-Bera (JB):	573.846
Skew:	-0.580	Prob(JB):	2.46e-125
Kurtosis:	3.543	Cond. No.	20.2

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

Conclusion

News Sentiment Analysis

Examined the influence of news sentiments on stock prices using 2021 data.

Key Findings

High sentiment scores positively affect market dynamics; low scores correlate with downturns.

Statistical Insight

OLS regression shows a modest but significant link between sentiment scores and stock returns.

Thank you

Mohammad Ahmed

mohammad.ahmed@fau.de