Predicting The Price of a Stock

An Interactive Qualifying Project
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfilment of the requirements for the
degree of Bachelor of Science

by
Samuel Kwok
Cole Manning
Mason Powell
Sam Rowe
Sebastian Villacorta

Date:
March 3, 2022

Report Submitted to:
Professor Mayer Humi

Abstract

The sharks of Wall Street have the leg up over small investors with their large financial backings and proprietary algorithms. The goal of this project was to apply historical stock data to mathematical models and algorithms to help investors make better guided decisions. Our models used a year's worth of trend data from the historical closing prices of stocks to make short-term predictions on their future performance.

Acknowledgements

We would like to thank Professor Mayer Humi for his guidance and assistance during this IQP. Professor Mayer Humi gave excellent instruction, feedback, and was always willing and available to answer our questions. Lastly, we want to thank him for encouraging us to think critically and outside the box.

Contents

1 Executive Summary **6**

2 Introduction **7**

3 Background **9**

3.1 The Stock Market and its Components 9

3.1.1 Stocks and the Market 9

3.1.2 Market Sectors 10

3.1.3 Market Indices 10

3.1.4 Volatility 11

3.1.5 Relative Strength Index 11

3.1.6 Bollinger Bands 12

3.1.7 Autocorrelation 13

3.1.8 Linear Regression (Least Squares) 13

3.1.9 Fourier Series 14

3.1.10 Savitzky-Golay Filter 14

4 Selected Stocks **15**

4.1 About Selected Stocks 15

4.1.1 Technology Sector 15

4.1.2 Materials Sector 16

4.1.3 Real Estate Sector 17

4.1.4 Consumer Discretionary Sector 19

4.1.5 Consumer Staples Sector 21

5 Stock Analysis and Prototype Fourier Model **22**

5.1 Autocorrelation 22

5.2 Trend Line and Prototype Fourier Series Model 23

5.3 Efficacy of Prototype Fourier Model Prediction 25

6 Fourier Model Accounting for Market Influence **31**

6.1 Technology Sector Index Results 32

6.2 Materials Sector Index Results 33

6.3 Real Estate Sector Index Results 34

6.4 Consumer Discretionary Sector Index Results 35

6.5 Consumer Staples Sector Index Results 36

7 RSI Signals Model 36

8 Bollinger Bands Model 37

9 Savitzky-Golay Model 39

10 Virtual Trading 41

10.1 Overall Performance 42
10.2 Technology Sector 42
10.3 Materials Sector 43
10.4 Real Estate Sector 43
10.5 Consumer Discretionary Sector 44
10.6 Consumer Staples Sector 44

11 Findings 45

11.1 Finding 1: Historical stock prices can be used to reasonably predict trends in the short term. 45
11.2 Finding 2: Stocks price trends and predictability are clearly affected by their sector. 47
11.3 Finding 3: Major real-world events have greater impact on stock prices than previous price trends. 48

12 Conclusions & Recommendations 48

13 Bibliography 50

14 Appendices 52

14.1 Appendix A: Matlab code 52

1 Executive Summary

To anticipate outcomes, predictive modeling uses statistics and data. Predictive modeling can be used to predict any form of unknown occurrence, regardless of when it occurred. When it comes to investing in the stock market, investors must make informed decisions. Brokers are aided by mathematical models that assist in making educated decisions. There exists models for short-term predictions as well as long-term predictions. Models can be used to get unbiased advice to make predictions or even be used in auto-trading applications.

Five sectors were chosen to be researched for this project. The first sector chosen was the Materials sector, which focuses on the production of raw materials and goods used in manufacturing. All of the stocks from the material sector come from the automotive or communications industry because they are expected to experience continuous and stable growth over a long period of time. The second sector chosen was the Real Estate sector, containing REITs (Real Estate Investment Trusts), mortgage companies, and property management companies. Stocks were selected from all three categories. This sector was chosen because of the volatility in the sector at the beginning of the project. The third sector was the Consumer Discretionary sector, the production of goods and services that are nonessential. This sector was chosen because it is expected to rise in value during the world's slow recovery from the COVID-19 pandemic. The fourth section chosen was the Consumer Staples sector, a sector that is composed of companies that produce and distribute essential everyday products. This sector was chosen because of its typically low volatility and steady growth due to it's non-cyclical nature. The fifth and last sector chosen was the Information Technology sector. This sector consists of both manufacturers of technological goods and the creators of software and virtual services.

This project starts off with the creation of a prototype model used to predict future stock prices. The Fourier Series acts as the foundation for the prototype model and was provided by Professor Humi. The Fourier Series is an expansion of a periodic function based on an infinite sum of sine and cosine functions. The prototype model takes in historical data and combined with the Fourier series would make predictions. The historical period initially used was from August 2020 to September 2021. The data is then autocorrelated to find the relevent peroid from the past year to use in the prototype model, the data was first reversed in order to have the most recent data be more impactful. This relevant period was used to simulate the behavior of each stock. One thing to note, is that this data comes from the peak of the COVID-19 pandemic. The efficiency of the prototype model was tested by the use of error bands, the area in which the prototype model was expected to deviate. The prototype model was able to predict prices up to eight days into the future. This differed slightly for each sector.

During this project, we simulated a trading exercise in which each member invested an initial amount of $100,000 into their chosen stocks for their sector. We then tracked the performance of our portfolios over this period, buying and selling greater shares as we saw fit. This activity allowed us to practically test our models while demonstrating that our methods could produce a greater profit than the general market.

In order to enhance the prototype model and assist in the virtual trading exercise, the RSI (Relative Strength Index) and Bollinger Bands were examined. The RSI is calculated using average price gains and losses over a given period of time. In this instance, the default time period is 14 days. The RSI produces a value from 0-100; a stock is considered overbought when the RSI is above 70 and oversold when it is below 30. The RSI can be used to pinpoint the general trend of a stock. The RSI assisted in making safer choices on which stocks to buy during the exercise than just using the prototype model alone. Bollinger Bands are calculated by a combination of moving averages of closing prices and standard deviation. Bollinger Bands are used to assess the volatility of a stock or market and when it is overbought or oversold. Bollinger bands helped enhance the prototype model by providing a more accurate indicator of when stocks were likely to deviate from its trend, this was used to buy and sell stocks to prevent profit loss.

2 Introduction

The stock market has been an ongoing source of commerce and business opportunity for generations. As far back as 1531, Belgium was the host of an exchange in which brokers would meet and deal with businesses, individuals, and individuals in debt, intending to sell bonds that would be paid back with interest (Beattie 2021). This later grew into outright investment through the East India companies. During the 1600s, the rapid spread of Eastern exploration led to the birth of dozens of different companies with their own associated expeditions. Prior to this point, ship owners would take investments in exchange for a partial return of the profits in order to balance the risk of a voyage gone afoul. Instead of a long-running partnership, these investments tended to last for a singular journey, thus requiring investors to fund multiple voyages to improve the odds of a return. The East India companies introduced the first version of a modern-day stock, which would allow individuals to invest in a company itself, paying for each ship's crew and allowing them to expect returns from all profits made under the company's authority (Beattie 2021). Consequently, this allowed said businesses to grow their fleets and demand more for each share, laying the groundwork for a relationship that defines the stock market.

In the 21st century, stock trading incentivizes individuals from all classes to invest through the allure of potential profit. Electronic trading has made this field increasingly accessible, with readily available services enabling anyone to put their money into a company nearly as easily as ordering any other product. The clear downside to this is that it takes a great deal of expertise and luck to accurately predict which venture will appreciate in value while avoiding those that stagnate or decline. In general, stocks are more well suited to long-term investments lasting multiple years, as stable stocks grow slowly but with little risk. Short-term investors deal with far greater risk but seek to make their profits over the span of weeks or days. Stock growth shows major variance between any given company and any give period of time, with profits only falling into the average returns window of 10% six years in the period from 1926 to 2014 (Royal 2021). Volatility is an inherent property of the stock market, and it is what investors want to take advantage of.

Trying to predict the health of the market as a whole has been an ongoing process, much like the market itself. This has taken the form of various indices, which collect data from multiple companies to track if prices have been increasing or decreasing. Using these as a measure is a major factor in determining an appropriate time to buy or sell stock for any given company. Many of these indices only monitor a smaller sector of the market, such as the technology-focused NASDAQ, in order to avoid losing accuracy by attempting to take data from several potentially unrelated fields (Caplinger 2021*b*). This allows people to grasp how specific segments of the market are performing, which may allow gains in a period of overall downturn. Following the trend of investing with an index as the only measure of the market has been a simple way for beginners to break into the field of stock trading with reasonable amounts of success, but there is still room for improvement.

More advanced investors employ methods that carefully examine the volatility and potential of individual stocks using their historical data, which allows them to control their risks. Through careful selection, they are able to build portfolios of company stocks that they can predict and profit from in the long term while investing reasonable amounts into stocks that might experience explosive, short-term growth. To find success in the latter, keeping inevitable losses in the 10-15% range so that they are always manageable (Seabury 2021).

The purpose of this project was to provide a way for small investors to compete against the sharks of Wall Street. The stock market is full of uncertainty, but big investors have inside trading tips and secrets that help them. The models derived from this project were designed to potentially aid a small investor in competing against these larger investors. Our team was introduced to this field of short-term investment while experimenting with potential methods for predicting a given stocks future value. Several computational models were fitted to analyze how chosen stock prices changed over time and how they compared to predictions made using said data. Emphasis was placed on the examination of multiple different sectors to obtain a better understanding of the stock market as a whole. Through this process, our team intended to provide an easier means for anyone to trade at the level of the most experienced investors.

3 Background

3.1 The Stock Market and its Components

3.1.1 Stocks and the Market

The stock market is where publicly owned companies and businesses are exchanged. These public companies are arranged in a way for portions of ownership to be sold away in the form of a stock. A stock is a form of security that represents ownership of a specific company (N/A 2021*e*). By buying a stock, or share, of a company, the buyer becomes owning a piece of that company. Investors can buy and maintain multiple stocks or sell the ones they already own to other interested investors. Originally, these transactions were done publicly in physical locations, but today, it is all done electronically (O'Shea 2021). Stocks can be bought or sold globally by investors through an exchange network, such as the New York Stock Exchange or brokers, such as Robinhood. The price of a stock depends on multiple factors but largely on the supply and demand of the market. If a company performs well, the stock tends to perform well too. As a result, the stock sees an increase in value. Investors buy stocks for numerous reasons including: for potential capital gains when stock prices rise, for any company payouts to investors (known as dividends), or to simply help support and influence a company towards a desired direction and goal. In return for a company selling its shares to investors, the company produces more money that it can use for its growth and development (Hayes 2021*b*).

3.1.2 Market Sectors

Stock market sectors are simply stocks organized into groups based on their companies' similarities. The market is broken up into sections and each section, or sector, contains stocks reflecting that of a shared or common industry. This form of stock-group classification is based on the Global Industry Classification Standard (GICS). There are currently 11 stock market sectors: Energy, Materials, Industrials, Utilities, Healthcare, Financials, Consumer Discretionary, Consumer Staples, Information Technology, Communication Services, and Real Estate (Caplinger 2021*a*). From there, each sector can be broken down even further into subcategories and specific trades. Market sectors help keep track of the performance of each part of the market, helping investors compare which companies and businesses are doing well in their respective divisions (O'Connell 2021).

3.1.3 Market Indices

A stock market index is a sample or collection of stocks for a specific sector, industry, or group that reflects how well it is currently performing in the market (N/A 2021*c*). Stock indices can help investors visualize how a particular section is performing and can give direction as to what stocks to buy, sell, observe, or ignore. They offer insight as to the "health" of the market (Staff 2021). As a result, there are multiple stock indices created for specific purposes that investors can use for any type of goal they might want to pursue. Stock indices vary in size, some containing as low as 30 companies while others exceeding over 2,000, all depending on the purpose of the index and what it is tracking (Tretina 2021). Some of the more popular used stock indices are: the S&P 500 Index, the Dow Jones Industrial Average, the Nasdaq, the NYSE Composite Index, and the Russell 2000 Index.

3.1.4 Volatility

Volatility has several different variations, but most generally it is a measure of a stock's price fluctuation over time. Mathematically, the volatility of an asset is measured by the variance and standard deviation of its prices over a given period, as given by these equations:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(u_i - \bar{u})^2}$$

$$\sigma_{an} = \sigma * \sqrt{h}$$

Where 'u' is the stock price and 'h' is the number of trading day intervals. These equations are used when estimating daily historical volatility and historical volatility, respectively. (*Compute Stock Volatility* 2011). If the price movement of a stock rises or falls quickly over a short period of time, it is considered to have high volatility due to the sudden fluctuation. Likewise, if the price movement of a stock rises or falls slowly over time, it is considered to have low volatility (Ashford 2021). This study was most concerned with actual current volatility, which is based on prices for a given period (initially a year for our stocks) that continues up until the most recent price. Alternatives include actual historical volatility, which does not end at the most recent price, and actual future volatility, which begins at the current price and ends at a time in the future. The volatility of an asset is important to investors because it determines a stock's reliability and therefore the amount that can be safely invested before risking major losses. However, high volatility is not strictly negative, as it can present opportunities to buy stocks at a low price before selling them for a much higher value. The amount of volatility that a stock experiences is due to many driving factors, such as those related to individual business decisions or practices, sectors being affected by specific industry-related events, and economic and political influence (N/A 2021*f*). Despite its nature, predicting volatile price changes through abnormal preceding movements is essential to generating and preserving profits made in the stock market.

3.1.5 Relative Strength Index

The Relative Strength Index, or RSI, is an indicator that tracks a stock's momentum. It is a useful tool that investors may use to help identify trading signals through the indication of whether a specific stock is being overbought or oversold (McBride 2021). It measures the change and speed of a stock's price over a period of time, which is commonly 14 days. It is calculated using the formula,

$$RSI = 100 - \frac{100}{1 + RS}$$

where RS is the Relative Strength value (N/A 2021d). The RS value is found by the summation of the average upward price change divided by the average downward price change, within the desired period. The values of the RS range from 0 to 100, which are then graphed to create the RSI. As most commonly accepted, an RSI value of 30 and below indicates that the stock is being oversold while an RSI value of 70 and above indicates it is being overbought. An oversold indication means that the stock has decreased rapidly but may still go back upward while an overbought indication means that the stock has increased rapidly but may go back downward. Investors use these values as entry and exit signals. When the RSI value rises above 30, many investors will interpret that as an entry, or to buy. Conversely, when the RSI value drops below 70, investors might interpret that as an exit, or to sell. While the RSI is a handy tool used by many investors, it is not always accurate since it does not take into account any external events or factors that may influence the price of the stock (McBride 2021).

3.1.6 Bollinger Bands

Bollinger bands are another analytical tool that investors can use to observe the amount of volatility for a stock. They help to indicate if the price of the stock is high or low on a relative basis (N/A 2021a). Bollinger bands are constructed by a series of three trendlines that form a "price envelope". These three lines each correspond to a lower, middle, and upper band. The middle band, or trendline, is the simple moving average of the price of the stock. The lower and upper bands are then the standard deviation of the simple moving average added and subtracted to it, forming the lower and upper bounds of the envelope. Typically, the standard deviation is multiplied by 2 and then both added and subtracted to the simple moving average for a 20-day period (N/A 2021b). The investor, however, could adjust the standard deviation factor as they see fit for any specified period of time, defining the overall amount of distance between the bands. Thus, the two parameters that are applied to the bands are the standard deviation size and the period. When the bands narrow and constrict together, known as a squeeze, it indicates a period with a low level of volatility. When the bands expand and widen apart, it indicates a period of high volatility, as the prices have more room to fluctuate and bounce around (Hayes 2021*a*). At times, the simple moving average can even break out of the envelope or hug one of the bands for an extended amount of time.

3.1.7 Autocorrelation

Autocorrelation is a statistical measurement of a data timeseries, in which the relationship of a current value in the series is measured to its previous values. This type is also known as a lagged correlation since it uses a specified number of lags (k) to shift the initial data values. Although this process uses just one time series, it is used twice: one that is shifted, or lagged, and the initial data time series (Smith 2021). To find the autocorrelation of the time series, it is lagged by a factor of k and by then finding its correlation between the lagged time series and the initial time series. The initial time series is lagged until the correlation between the two series is equal to zero. When the correlation reaches zero, the autocorrelation period is found, given by the number of lags (k) it took to get there. This autocorrelation period provides the most relevant data values to use within the time series. This is especially useful for determining the sample size needed to analyze each stock and for the modeling of short-term predictions. Below is the formula for the autocorrelation of a time series.

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t + \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2}$$

3.1.8 Linear Regression (Least Squares)

Linear regression is a simple mathematical technique used to help model the relationship between two variables: independent and dependent variables. This is done by calculating and inserting a line of best fit for the given data set to be used for interpretation and prediction. The dependent variable is the variable that is desired for its predicted value, while the independent variable is the variable used to predict it (Kenton 2022). The most common approach to achieve a linear regression line is by using least-squares. The least-squares is the smallest sum of squares of errors. With this technique, the vertical distance from each point to the regression line is minimized, thus having the least error (IBM 2022). The linear regression function is expressed simply with $y = mx + b$, where y is the dependent variable and x is the independent variable. The slope of the line is m and the intercept is b, and both are used to express the least-squares regression for the given data set. As the most basic modeling form, the least-squares linear regression can be used to help model the trend of a stock and which direction it can be headed. Thus, this can be a useful analysis tool that could help predict the behavior of a stock's price.

3.1.9 Fourier Series

The Fourier Series is the sum of a combination of both sine and cosine functions. It is used to delineate a periodic function using the summation of these waveforms up to an infinite amount for each wave, if desired. Due to it sharing similar time series characteristics such as infinite summations, the Fourier Series is often compared to that of the Taylor Series. The Fourier Series, however, is helpful when dealing with arbitrary periodic functions since it will split up these functions into a group of simpler terms to then be utilized, solved, and then substituted back into the original functions (Weisstein 2022). One consideration when using the Fourier Series is that the higher the k value or the larger the amount of sines and cosines being used, the less accurate it becomes. This is due to a major spike in magnitude within the time series, known as the Gibbs Phenomenon. Regardless, the Fourier Series could be a useful component in forecasting models, such as in the stock market.

3.1.10 Savitzky-Golay Filter

The Savitzky-Golay filter is a digital filter, used in signal processing to smooth out a set of data. The intention of using the Savitzky-Golay filter is to increase the precision of the data without distorting it. This is achieved via convolution, a mathematical operation on two functions that produces another function. This third function shows how the shape of the first function is modified by the second function. The Savitzky-Golay filter fits successive sub-sets of adjacent data points with a low degree polynomial by linear least sqaures. Moving average filters are commonly used in various industries with time series data to analyze noisy data, by removing short-term fluctuations and pinpointing longer-term trends. The Savitzky-Golay filter can be applied to the stock market to assist in analyzing price trends.

4 Selected Stocks

4.1 About Selected Stocks

4.1.1 Technology Sector

- Founded in 1937, **Canon Inc. (NYSE: CAJ)** is most well known as a camera manufacturer, specializing in producing many imaging products.
- **Calix, Inc (NYSE: CALX)** is a provider of cloud telecommunication solutions, such as VOIP servers. Calix has been active since 1999.
- Formerly part of Hewlett-Packard, **HP Inc. (NYSE: HPQ)** split from its predecessor in 2015, taking with it the personal computer and printer divisions. HP continues to produce new laptops, workstations, and printers at the consumer and enterprise level.
- **InterDigital, Inc. (NASDAQ: IDCC)** is an R&D company, focusing on mobile and video technologies, primarily in the wireless space.
- **Intel Corp. (NASDAQ: INTC)** is a leading manufacturer of CPUs and semiconductor chips, but also manufactures other small chipsets and computing devices.
- The American **MaxLinear, Inc. (NYSE: MXL)**, provides radio-frequency (RF), analog, digital, and mixed-signal semiconductor solutions to communications companies.
- **NCR Corp. (NYSE: NCR)** (previously National Cash Register) is a manufacturer of services and technology aimed at restaurants, retailers and banks. This includes full point-of-sale systems and services, along with ATM machines, and adjacent technology.
- Describing itself as operating in the fields of security, healthcare, and optoelectronics and manufacturing, **OSI Systems, Inc. (NASDAQ: OSIS)** is a marketer and developer of semiconductors and digital tools.
- **Raytheon Technologies Corp. (NYSE: RTX)** designs and manufactures advanced technology products in the aerospace and defense industries.

- The Israeli **Silicom Ltd. (NASDAQ: SILC)** specializes in the design, manufacturing, and marketing of networking equipment for servers. It is known for its line of high-end server networking cards.
- Silicon Valley's **Super Micro Computer, Inc. (NASDAQ: SMCI)**, more commonly referred to as simply Supermicro, is an IT infrastructure company, specializing in various technologies for servers and data centers.
- Staying true to its namesake, **Semtech Corporation (NASDAQ: SMTC)**, manufactures analog and mixed-signal semiconductors for consumer and industrial industries.
- **Sonos, Inc. (NASDAQ: SONO)** is a consumer audio company, well known for its multi-room audio solutions.

4.1.2 Materials Sector

- **Alcoa Corp. (NYSE: AA)** was founded in 1888 and primarily produces bauxite, alumina, and aluminum to sell to material processing companies.
- **Belden Inc. (NYSE: BDC)** operates using an Enterprise Solutions and Industrial Solutions branch. The former produces copper cable, fiber cable, and other connectivity solutions. The latter produces infrastructure components such as industrial Ethernet cables and I/O modules.
- The **Cieana Corp. (NYSE: CIEN)** produces network hardware and software for information services.
- **Cleveland-Cliffs Inc. (NYSE: CLF)** was founded in 1847 and operates as a flat-rolled steel producer supplying the automotive industry.
- **DuPont de Nemours, Inc. (NYSE: DD)** produces materials used in manufacturing photovoltaics, solar cells, and dielectrics for computer chips.
- **Freeport-McMoRan Inc. (NYSE: FCX)** is a mining company that produces copper, gold, silver, and other metals.
- The **Lithium Americas Corp. (NYSE: LAC)** focuses on excavating lithium deposits in the United States.

- **Lumentum Holdings Inc. (NASDAQ: LITE)** manufactures optical and photonic products such as fiber optic cables, transceivers, and transmitter modules.

- **Compagnie Générale des Établissements Michelin Société en commandite par actions. (OTC: MGDDY)** is a French company that primarily offers products related to the automotive industry, such as tires used in private use (cars, motorcycles) and professional use (trains, heavy machinery).

- **Plexus Corp. (NASDAQ: PLXS)** provides various electronic manufacturing and aftermarket services.

- **Green Plains Inc. (NASDAQ: GPRE)** was founded in 2004 and produces ethanol, distiller grains, and corn oil.

- **Teck Resources Limited. (NYSE: TECK)** mines for a variety of different minerals, with a focus on coal, copper, zinc, bitumen, and lead.

- The manufacturing segment of **Otter Tail Corp. (NASDAQ: OTTR)** engages in metal parts stamping, production of plastic horticultural containers, industrial packaging, and extruding raw material stock for various industries.

- **Ternium S.A. (NYSE: TX)** manufactures steel products via two segments, steel and mining. The former produces items such as slabs, reinforced bars, beams, tiles, and other similar products. The former sells iron ore and pellets.

- The **United States Steel Corporation (NYSE: X)** manufactures flat-rolled and tubular steel products as well as iron ore.

4.1.3 Real Estate Sector

- **Realty Income Corporation. (NYSE: O)** owns over 6,500 retail real estate properties with long-term leases with commercial clients. Realty Income Corporation is known as the Monthly Dividend Company because it provides its shareholders with a dependable monthly income from dividends.

- **Iron Mountain Incorporated. (NYSE: IRM)** is a global leader in storage and information management services. They provide a service to store and protect billions of valuable assets ranging from business information to historical artifacts. These services include record storage, information management, secure destruction, and online cloud storage.
- **Invitation Homes Inc. (NYSE: INVH)** is primarily a single-home leasing company. They provide homes that are in close proximity to jobs and schools for people.
- **FRP Holdings, Inc. (NASDAQ: FRPH)** is a real estate investment and development company in the United States. FRP Holdings is broken into four segments: Asset Management, Mining Royalty Lands, Development, and Stabilized Joint Venture.
- **National Health Investors, Inc. (NYSE: NHI)** is a real estate investment trust that specializes in sale-leaseback, joint-venture, mortgage, and mezzanine financing of need-driven and discretionary senior housing and medical investments.
- **Universal Health Realty Income Trust (NYSE: UHT)** a real estate investment trust, that focuses on healthcare and human service-related facilities. This includes hospitals, medical and office buildings, rehabilitation centers, free-standing emergency departments, and childcare centers.
- **RE/MAX Holdings, Inc. (NYSE: RMAX)** is a franchisor of real estate and mortgage brokerage services, primarily in the United States and Canada. The company is divided into three segments: Real Estate, Mortgage, and Marketing Funds.
- **PotlatchDeltic Corporation (NASDAQ:PCH)** is a real estate investment trust that owns over 1.8 million acres of timberlands, primarily in Alabama, Arkansas, Idaho, Louisiana, Minnesota, and Mississippi. They also operate sawmills and run rural timberland sales programs, and a residential and commercial real estate development business.
- **Rayonier Inc. (NYSE: RYN)** is a timberland real estate investment trust with 2.7 million acres in the United States and New Zealand.
- **First Industrial Realty Trust, Inc. (NYSE: FR)** own, operate, and have under development millions of square feet of industrial real estate.

- **NexPoint Residential Trust, Inc. (NYSE: NXRT)** is primarily focused on acquiring, owning, and operating middle-income multifamily properties in large cities and suburbs, mostly in the Southeastern and Southwestern United States.
- **American Assets Trust, Inc. (NYSE: AAT)** is focused on acquiring, improving, developing, and managing office, retail, and residential properties in high-barrier-to-entry markets, mostly in California, Oregon, Washington, Texas, and Hawaii.
- **Blackstone Mortgage Trust, Inc. (NYSE: BXMT)** is a real estate finance company that issues senior loans collateralized by commercial real estate.
- **American Homes 4 Rent (NYSE: AMH)** operates in the single-family home rental industry, focused on acquiring, developing, renovating, leasing, and operating single-family homes.
- **American Campus Communities, Inc. (NYSE: ACC)** is the largest owner, manager, and developer of high-quality student housing communities in the United States.

4.1.4 Consumer Discretionary Sector

- **Bed Bath & Beyond Inc (NASDAQ: BBBY)** is an American chain of domestic merchandise retail stores.
- **DISH Network Corp (NASDAQ: DISH)** is an American television provider and the owner of the direct-broadcast satellite provider Dish.
- **eBay Inc. (NASDAQ: EBAY)** is an American multinational e-commerce corporation that facilitates consumer-to-consumer and business-to-consumer sales through its website.
- **The Fox Corporation (NASDAQ: FOXA)** is an American mass media company primarily in the business of television broadcasting stations.
- **The Formula One Group (NASDAQ: FWONK)** is a group of companies responsible for the promotion of the FIA Formula One World Championship, and the exercising of the sport's commercial rights.

- **The Gap, Inc (NYSE: GPS)** is an American worldwide clothing and accessories retailer.
- **The Huazhu Hotels Group Ltd (NASDAQ: HTHT)** is a hotel management company in China.
- **Leggett & Platt (NYSE: LEG)**, based in Carthage, Missouri, is a diversified manufacturer that designs and produces various engineered components and products that can be found in homes and automobiles.
- **The Las Vegas Sands Corporation (NYSE: LVS)** is an American casino and resort company based in Nevada. Its resorts feature accommodations, gambling and entertainment, convention and exhibition facilities, restaurants and clubs.
- **Macy's, Inc (NYSE: M)** is an American department store company founded by Xavier Warren in 1929. It sells products from clothing to fragrances and house appliances.
- **The Rogers Communications Inc (NYSE: RCI)** is a Canadian communications and media company operating primarily in the fields of wireless communications, cable television, telephony, and Internet.
- **Skechers USA, Inc. (NYSE: SKX)** is a footwear company that is the third largest athletic footwear brand in the United States. Its headquarters is located in California.
- **The TJX Companies, Inc. (NYSE: TJX)** is an American multinational off-price department store corporation, headquartered in Framingham, Massachusetts.
- **Toyota Motor Corporation (OTCMKTS: TOYOF)** is a Japanese multinational automotive manufacturer headquartered in Toyota City, Aichi, Japan. It is one of the largest automobile producers in the world.
- **The VF Corporation (NYSE: VFC)** is an American worldwide apparel and footwear company founded in 1899 and headquartered in Denver, Colorado. It consists of 13 total brands including: The North Face, Vans, and Supreme.

4.1.5 Consumer Staples Sector

- Established in 1937, **British American Tobacco PLC (NYSE: BTI)** is a British multinational company that sells cigarettes, tobacco, and other nicotine products. As of 2019, it is the largest tobacco company in the world based on net sales.
- **Conagra Brands, Inc. (NYSE: CAG)** is an American consumer packaged goods holding company headquartered in Chicago, Illinois.
- **Church & Dwight Co., Inc. (NYSE: CHD)** is a major American manufacturer of household products that is headquartered in Ewing, New Jersey. It is best known for its Arm & Hammer line which includes baking soda and a variety of products made with it, including laundry detergent.
- **Colgate-Palmolive (NYSE: CL)** is a company that sells various household products, including toothpaste, soap, pet food, cleaning products.
- **Campbell Soup Company (NYSE: CPB)** is an American processed food and snack company, best known for its canned soup products.
- **General Mills, Inc. (NYSE: GIS)**, is an American multinational manufacturer and marketer of branded consumer foods sold through retail stores.
- Founded by Mark R. Hughes in 1980, **Herbalife Nutrition Ltd (NYSE: HLF)** is a global multi-level marketing corporation that develops and sells dietary supplements.
- **Ingredion Incorporated (NYSE: INGR)** is an American multinational ingredient provider based in Westchester, Illinois producing mainly starch, modified starches and starch sugars.
- **The Kellogg Company (NYSE: K)** is an American multinational food manufacturing company headquartered in Battle Creek, Michigan.
- **The Kraft Heinz Company (NASDAQ: KHC)** is an American food company formed by the merger of Kraft Foods and Heinz.
- **Lamb Weston Holdings Inc (NYSE: LW)** is an American food processing company and one of the world's largest producers and processors of various frozen potato products including frozen french fries, waffle fries, and others.

- **Monster Beverage Corp (NASDAQ: MNST)** is an American beverage company that manufactures energy drinks including Monster Energy, Relentless and Burn.
- **New York Times Co (NYSE: NYT)** is an American mass media company that publishes The New York Times newspaper.
- Headquartered in Houston, Texas, **Service Corporation International (NYSE: SCI)** is an American provider of funeral goods and services as well as cemetery property and services.
- **Tyson Foods, Inc. (NYSE: TSN)** is an American multinational corporation based in Springdale, Arkansas, that operates in the food industry. It is the second largest processor and marketer of chicken, beef, and pork in the world.

5 Stock Analysis and Prototype Fourier Model

With the chosen stocks downloaded into their respective CSV filess, we could begin work on the prototype model. This process was completed using MathWorks' MATLAB programming language.

5.1 Autocorrelation

The first step was to get the **correlated period** for each stock's daily closing price over time. The correlated period is the range of data for each stock that is used when generating prediction models. Autocorrelation investigates the degree of similarity between a time series and a lagged version of the series across subsequent intervals. This will measure the link between a variable's current and past values. MATLAB's sample autocorrelation function was used for this. The autocorrelation or serial correlation, finds the correlation between a time-series of data, y_t. The autocorrelation for a lag k is given by;

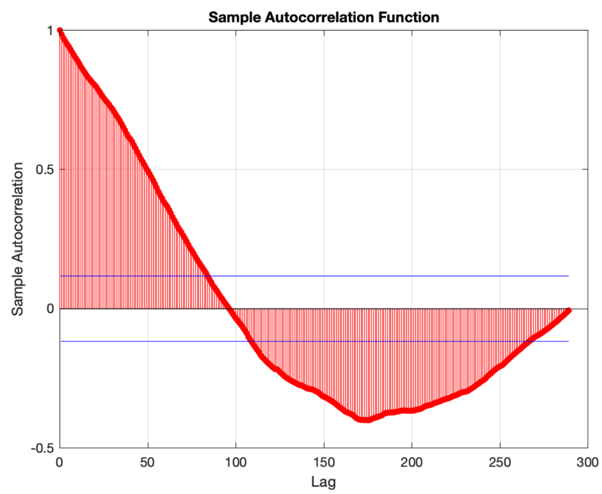
$$r_k = \frac{c_k}{c_0}$$

where,

$$c_k = \frac{1}{T}\sum_{t=1}^{T-k}(y_t - \bar{y})(y_{t+k} - \bar{y})$$

c_0 is the sample variance of the time series. The data was first reversed - so that the *most recent* data would be parsed first - and fed into the `autocorr()` function. The usable portion of the data was determined as being from the beginning (newest closing price), to the first time the function was equal to zero (deemed uncorrelated). This process effectively trimmed off uncorrelated closing prices, allowing us to only process relevant data.

Figure 1: Autocorrelation Function Graph for Canon Inc. (CAJ)



5.2 Trend Line and Prototype Fourier Series Model

The correlated data was then graphed over time, along with its line of best fit (MATLAB polyfit with $N = 1$). This trend line linearly maps to the overarching pattern of our stock prices with the equation:

$$y = mx + b$$

Next, the difference between the trend line and the actual closing price was graphed. Typically, these difference graphs have a resemblance to a sinusoidal function, oscillating up and down. A Fourier Series is an extension of a periodic function with regard to the infinite sum of sines and cosines. The sin and cosine functions have an orthogonality relationship that the Fourier Series uses for its expansions. A Fourier Series with three terms was then fit to the difference graph. If the period of relevant data included at least 50 data points, then we used a Fourier 3 model. Otherwise, we opted to use a Fourier 2 model. The equation for the Fourier series is given as:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi kx}{T}\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi kx}{T}\right)$$

for a periodic function f(x) where,

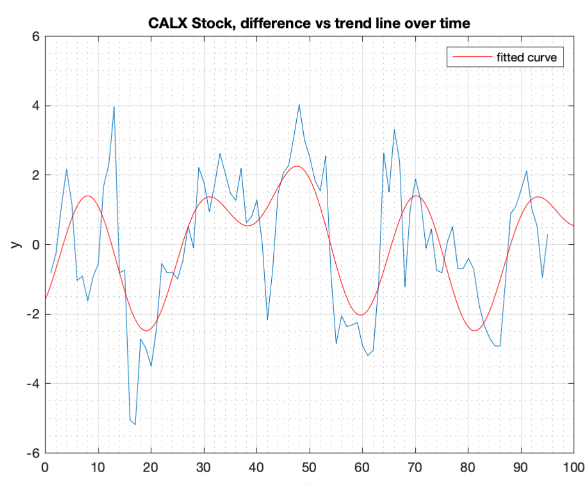
$$a_0 = \frac{2}{T}\int_0^T f(x)\,dx$$

$$a_k = \frac{2}{T}\int_0^T f(x)\cos\left(\frac{2\pi kx}{T}\right)\,dx$$

$$b_k = \frac{2}{T}\int_0^T f(x)\sin\left(\frac{2\pi kx}{T}\right)\,dx$$

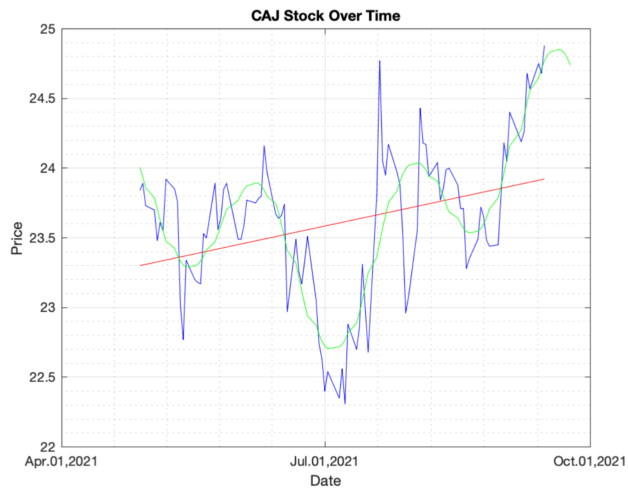
and T is the period length. Using these equations, our group was able to obtain an approximation of where a given stock's closing price would be in the near future. If the period of relevant data included at least 50 data points, we used a Fourier 3 model. Otherwise, we opted to use a Fourier 2 model, with said number corresponding to the amount of sin and cos pairs used in the series.

Figure 2: Calix Inc., Difference vs Trend Line, with Fitted Fourier Curve



In order to match the Fourier Model with the closing prices of the stock, and not the differences, the Fourier equation was added to the equation of the trend line for each point: $adjusted(x) = fourier(x) + trend(x)$. The resulting equation is visible as the green line on Figure 3

Figure 3: Canon Stock Over Time with Trend Line and Adjusted Fourier



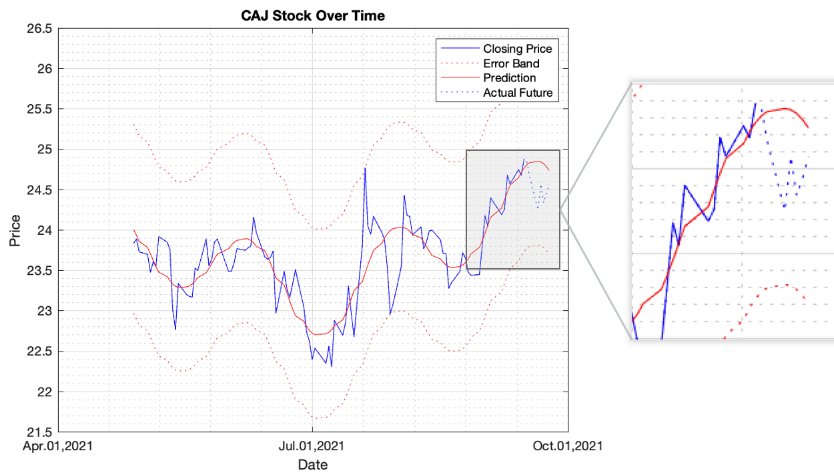
5.3 Efficacy of Prototype Fourier Model Prediction

The final step for this portion of the project was the analysis of the efficacy of the generated Fourier Model. This was done by first generating "error bands," to represent the area in which the Fourier prediction is expected to deviate. The only data needed to make the error bands are the differences between the model's predicted prices and the actual closing prices over the correlated period. Each stock's error band is then formed by adding and subtracting the absolute value of the maximum difference found to every closing price.

The future[i] closing prices of each stock were then overlain onto a graph with the error bands, Fourier prediction, and known closing prices, as seen in Figure 4.

[i]Relative to original downloaded data the prediction model was made against. 'Future' data ran until 2021-09-15.

Figure 4: Final Canon Stock Analysis Graph



The graphs were programmatically analyzed to determine if they exited the error band and if they ever reentered. The below tables display this data for each sector, with a maximum value of 10 days within the bands.

Table 1: Prediction Analysis (Technology Sector)

Stock	Days Till Out of Error Band	Error Bound	Total Days in Band
CAJ	Never	1.1698	8
CALX	6	3.6842	6
HPQ	4	1.6863	4
IDCC	3	4.0112	3
INTC	2	1.5276	2
MXL	3	3.6737	3
NCR	6	2.4316	6
OSIS	Never	2.8078	8
RTX	5	3.3216	4
SILC	5	1.8739	5
SMCI	Never	1.9001	8
SMTC	5	3.3856	5
SONO	Never	2.8274	8
T	1	0.2656	6
VSAT	2	3.256	2
Mean	3.8182	2.5215	5.2
STD	1.6414	1.0461	2.1039

Table 2: Prediction Analysis (Materials Sector)

Stock	Days Till Out of Error Band	Error Bound	Total Days in Band
AA	2	5.3430	1
BDC	1	3.0413	0
CIEN	1	1.4793	0
CLF	1	2.9304	0
DD	5	3.7202	7
FCX	1	3.2913	0
GPRE	1	3.5469	0
LAC	4	3.2262	3
LITE	1	2.2499	0
MGDDY	3	0.91685	2
OTTR	1	1.4979	0
PLXS	1	3.6301	0
TECK	1	2.2894	0
TX	3	3.0230	2
X	3	3.0334	2
Mean	1.9333	2.8813	1.1333
STD	1.2893	1.0524	1.8571

Table 3: Prediction Analysis (Real Estate Sector)

Stock	Days Till Out of Error Band	Error Bound	Total Days in Band
O	3	2.1880	2
IRM	1	1.9225	2
INVH	3	1.2166	3
FRPH	Never	4.0297	8
NHI	Never	1.6441	8
UHT	3	0.7909	2
RMAX	1	0.9684	0
PCH	2	3.0615	1
RYN	Never	1.4454	8
FR	3	1.5008	2
NXRT	5	2.2242	7
AAT	Never	1.8664	8
BXMT	Never	1.4778	8
AMH	1	1.3203	0
ACC	1	1.3417	1
Mean	2.3	1.7999	4
STD	1.2689	0.80608	3.2249

Table 4: Prediction Analysis (Consumer Discretionary Sector)

Stock	Days Till Out of Error Band	Error Bound	Total Days in Band
BBBY	Never	7.7212	10
DISH	Never	4.4789	10
EBAY	Never	7.1774	10
FOXA	Never	1.9322	10
FWONK	Never	3.4736	10
GPS	Never	4.0602	10
HTHT	Never	5.3982	10
LEG	Never	2.4921	10
LVS	Never	3.9212	10
M	9	3.5102	9
RCI	Never	2.2414	10
SKX	Never	5.6491	10
TJX	Never	4.6604	10
TOYOF	4	5.6871	6
VFC	Never	4.9392	10
Mean	6.5	4.4895	9.6667
STD	2.5	1.6245	1.0111

Table 5: Prediction Analysis (Consumer Staples Sector)

Stock	Days Till Out of Error Band	Error Bound	Total Days in Error Band
BTI	Never	1.6853	10
CAG	2	1.0017	1
CHD	5	2.3662	4
CL	1	1.6436	0
CPB	3	3.1661	2
GIS	2	2.3252	1
HLF	1	8.5662	0
INGR	Never	5.9866	10
K	8	2.1455	7
KHC	Never	2.0355	10
LW	6	8.2758	5
MNST	4	7.2544	3
NYT	1	2.6887	0
SCI	2	2.7397	1
TSN	Never	7.1458	10
Mean	3.1818	3.9351	4.2667
STD	2.2082	2.5843	3.9407

6 Fourier Model Accounting for Market Influence

To improve the Fourier model further, we took into account the stock market's performance as a whole via multiple indices. Each sector made separate models using the DOW, NASDAQ, and a third index tailored to the given sector.

The index first had to be scaled down (I_s) by dividing each day's value by that of the first value in our relevant period. This same process was repeated for the closing prices of each stock so that the two values could be accurately compared. To get this result, we found the linear correlation between the scaled versions of both the stock and each index (α). We then utilized MATLAB's corr function to get each index's correlation coefficient. Using those values, we weighted the indices to the Fourier model with the following equations:

$$I_s * \alpha + (1 - \alpha) * F_s$$

$$I_s * (1 - \alpha) + \alpha * F_s$$

By combining the two predicted price values in this way, we could account for the behavior of the market as a whole while still granting the appropriate amount of influence to our own model as determined by the α value.

6.1 Technology Sector Index Results

Table 6: Prediction Analysis with Indices (Technology Sector)

Stock	NDXT	DOW	NASDAQ	Best Index
CAJ	8	8	8	TIE
CALX	6	6	6	TIE
HPQ	5	5	5	TIE
IDCC	4	5	4	DJI
INTC	7	7	7	TIE
MXL	3	3	3	TIE
NCR	1	1	1	TIE
OSIS	1	1	1	TIE
RTX	6	6	6	TIE
SILC	8	8	8	TIE
SMCI	8	8	8	TIE
SMTC	1	1	1	TIE
SONO	2	0	0	NDXT
T	8	8	8	TIE
VSAT	8	8	8	TIE
Mean	5.0667	4	4.9333	NDXT

For the technology sector, the NASDAQ-100 Technology Sector (NDXT) was chosen along with our standard choices of the DOW and NASDAQ as the three incorporated indices. When compared to DJI & NASDAQ, the NDXT showed a slight gain in performance for our limited testing. It is expected that it would better represent the data for any given tech stock, as this index contains leaders in the tech industry.

6.2 Materials Sector Index Results

Table 7: Prediction Analysis with Indices (Materials Sector)

Stock	DOW Material	DOW	NASDAQ	Best Index
AA	9	9	8	TIE
BDC	8	5	9	NASDAQ
CIEN	4	4	4	TIE
CLF	4	4	4	TIE
DD	1	1	1	TIE
FCX	9	9	8	TIE
GPRE	6	6	6	TIE
LAC	1	1	1	TIE
LITE	5	4	3	DJUSBM
MGDDY	8	8	5	TIE
OTTR	7	1	9	NASDAQ
PLXS	3	3	3	TIE
TECK	6	6	6	TIE
TX	5	5	5	TIE
X	9	4	4	DJUSBM
Mean	5.6667	4.6667	5.0667	TIE
STD	2.7168	2.6367	2.6040	TIE

The DOW Jones U.S. Basic Materials Index (DJUSBM) was chosen as the third measurement because it specializes in the given sector. The inclusion of the indices universally improved the model's performance, with the NASDAQ and DOW Material index being equally consistent over the measured period.

6.3 Real Estate Sector Index Results

Table 8: Prediction Analysis with Indices (Real Estate Sector)

Stock	S&P REIT	DOW	NASDAQ	Best Index
AAT	8	8	8	TIE
ACC	8	8	8	TIE
AMH 3	2	4	NASDAQ	
BXMT	2	2	3	NASDAQ
FR	8	8	8	TIE
FRPH	8	8	8	TIE
INVH	8	8	8	TIE
IRM	4	4	3	S&P/DOW
NHI	8	8	8	TIE
NXRT	8	8	8	TIE
O	8	8	8	TIE
PCH	8	8	8	TIE
RMAX	1	1	1	TIE
RYN	8	8	8	TIE
UHT	8	8	8	TIE
Mean	6.5333	6.4666	6.6	NASDAQ
STD	2.5875	2.6956	2.4727	NASDAQ

The third index chosen for the real estate sector was the S&P United States REIT index. This was chosen to gauge the sector's health because it measured how well the S&P real estate investment trusts were performing. By incorporating the indices, the model's predictions in the real estate sector saw improvements overall. The NASDAQ ended up slightly beating the DOW and S&P REIT indices.

6.4 Consumer Discretionary Sector Index Results

Table 9: Prediction Analysis with Indices (Consumer Discretionary Sector)

Stock	S&P 500	DOW	NASDAQ	Best Index
BBBY	10	10	10	TIE
DISH	10	10	10	TIE
EBAY	5	10	4	DOW
FOXA	10	10	10	TIE
FWONK	5	6	9	NASDAQ
GPS	10	10	10	TIE
HTHT	10	10	10	TIE
LEG	10	10	10	TIE
LVS	10	10	10	TIE
M	10	10	10	TIE
RCI	10	10	10	TIE
SKX	10	10	10	TIE
TJX	5	5	5	TIE
TOYOF	9	9	9	TIE
VFC	10	10	10	TIE
Mean	8.9300	9.3333	9.1333	DOW
STD	1.9821	1.5345	1.8571	DOW

The third chosen index for the consumer discretionary sector was the S&P 500 index. This index was chosen since it contains several large companies from this sector. However, both the DOW and NASDAQ beat out the chosen index, with the DOW becoming the best index overall from the three.

6.5 Consumer Staples Sector Index Results

Table 10: Prediction Analysis with Indices (Consumer Staples Sector)

Stock	DOW	NASDAQ	S&P 500 Consumer Staples	Best Index
BTI	10	10	10	TIE
CAG	9	8	9	TIE
CHD	3	10	4	NASDAQ
CL	9	9	9	TIE
CPB	10	4	10	TIE
GIS	6	7	5	NASDAQ
HLF	5	5	5	TIE
INGR	10	8	10	TIE
K	9	9	9	TIE
KHC	10	9	10	TIE
LW	10	10	10	TIE
MNST	4	10	4	NASDAQ
NYT	10	7	7	DOW
SCI	3	10	3	NASDAQ
TSN	10	10	10	TIE
Mean	7.8667	8.4	7.6667	NASDAQ
STD	2.7047	1.8547	2.5991	NASDAQ

The S&P 500 Consumer Staples index was chosen as the third index for the Consumer Staples index. This was chosen because it seemed appropriate, as this index was specifically related to the sector in question. It was still, however, beat by NASDAQ by all metrics for these stocks in this time period.

7 RSI Signals Model

Our next step after moving into the application period was the implementation of the relative strength index (RSI) to determine when to invest more or less into any given stock. The RSI is given by dividing the average of the percent gains by the average of the percent losses within a specified period.

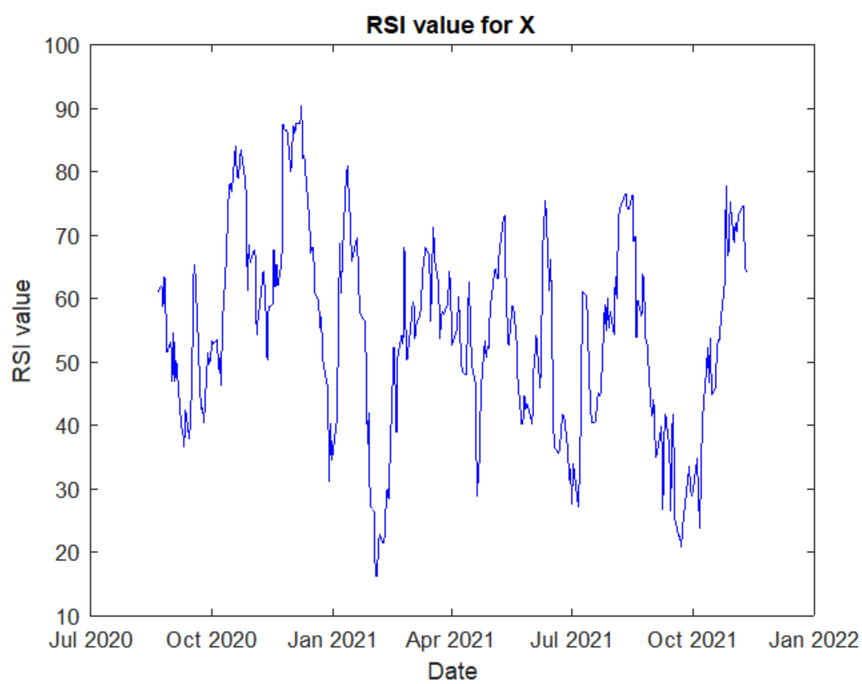
$$RSI = 100 - \frac{100}{1 + RS}$$

where,

$$RS = \frac{average\ gains}{average\ losses}$$

At its core, the RSI is a momentum indicator that predicts if a stock or some other asset is overbought or oversold. In general, RSI values above 70 indicate the former, while those below 30 indicate the latter. This indicator was a major factor in our decisions to redistribute our investments to avoid imminent dips in value. The RSI combined with the prototype model allowed for safer investments to be made than with just the model.

Figure 5: RSI Function Graph for the United States Steel Corporation (X)



8 Bollinger Bands Model

To expand on our methods of handling potentially problematic stocks, we formed Bollinger Bands using our data. The two bands are created one standard deviation above and below the moving average of a given stock. The Bollinger

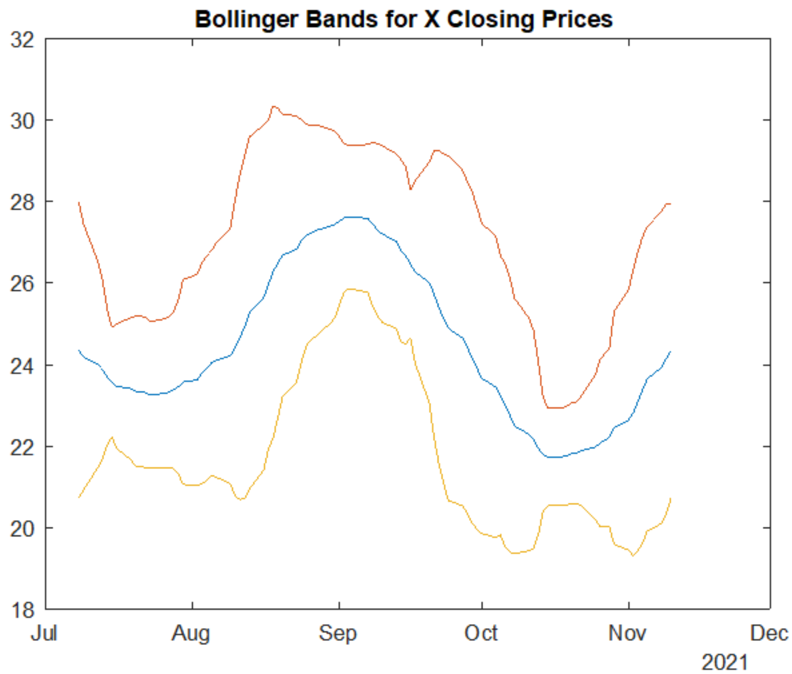
Band formulas for standard deviation and the moving average are given as:

$$\sigma = \sqrt{\frac{\sum_{j=1}^{N}(X_j - \bar{X})^2}{N}}$$

$$\bar{X} = \frac{\sum_{j=1}^{N} X_j}{N}$$

The upper band, will be calculated as $\bar{X} + 2\sigma$, the middle as $\bar{X}$, and lastly the lower band as $\bar{X} - 2\sigma$. These bands scale the data and can be used to detect high or low values relative to the rest of the stock's prices, with proximity to the upper band indicating that a stock is overvalued and vice versa. Through this method, we can determine if a stock is more likely to experience a major change in value while using other indicators to confirm our results. For example, the bands can also be applied to RSI, with the behavior of the latter being interpreted similarly to that of a stock itself. More specifically, if the RSI fails to touch the upper Bollinger band during its extreme highs, that could be interpreted as a sell signal due to the stock being perceived as overvalued. The bands returned by this function tighten during periods of low volatility and vice versa, with especially strong trends resulting in the price leaving the bands and generating buy or sell signals. The usage of Bollinger bands helped our group to determine when our stocks were likely to experience major changes in value, and thus were a major factor in deciding when to buy and sell during our initial virtual trading period. The bands were able to provide a better insight into determining the optimal time to sell or buy than the prototype alone.

Figure 6: Bollinger Band Graph for the United States Steel Corporation (X)



9 Savitzky-Golay Model

To further increase the accuracy of our market predictions further into the future, Savitzky-Golay filtering was incorporated into the basis of our model. The initial Fourier model used a linear trendline with the equation:

$$y = mx + b$$

The Savitzky-Golay filter modifies the trendline equation to incorporate the use of an exponent between the chosen range of .9 to 1.1, resulting in the equation:

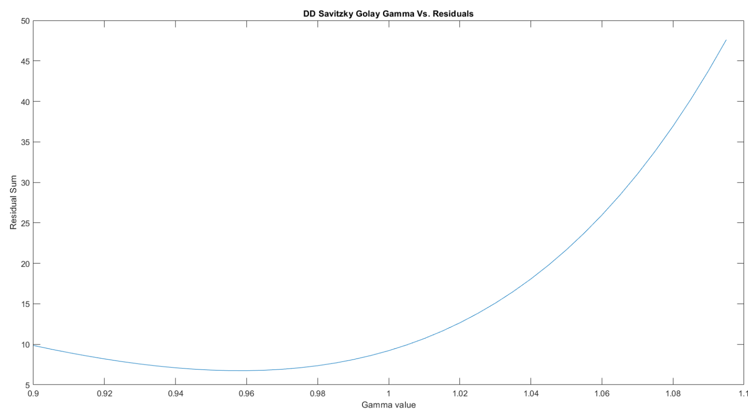
$$y = mx^{\gamma} + b$$

This gamma value smooths out the curves of data by scaling each value downwards, consequently reducing the standard deviation and impact from outliers in the data set. This allowed us to better map to general trends rather than the potentially more extreme data itself. We determined the optimal gamma value manually by calculating the sum of the residuals for different gamma values within our range, with the minimum denoting the gamma value that was most accurate to actual historical prices. The sum of the residuals can be calculated by the formula:

$$\sum_{i=1}^{i}(p_i - mi^{\gamma} + b)$$

Where "p" refers to the actual price of a stock on a given day. The resulting plot for each gamma for a given stock produces the following:

Figure 7: Savitzky-Golay Comparison for DuPont de Nemours Inc (DD)



Although it is possible for optimal gamma values to vary between each stock, we found that stocks within a given sector were often receptive to similar values. In the case of the Real Estate Sector, .95 worked the best for all of the stocks. This smoothed out some of the noise in the data and extended the model's predictions slightly. The figure below showcases the effectiveness of the Savitzky-Golay filtering on the Real Estate Sector predictions.

Table 11: Savitzky-Golay Filtering (Real Estate Sector)

Stock	Without Filtering	With Filtering
O	2	4
IRM	2	8
INVH	2	8
FRPH	8	9
NHI	8	7
UHT	2	2
RMAX	0	0
PCH	1	2
RYN	8	9
FR	2	8
NXRT	4	8
AAT	8	9
BXMT	8	9
AMH	0	8
ACC	1	8
Mean	3.7333	6.6000

All other sectors demonstrated similar improvements to the maximum number of days inside each error band when using Savitzky-Golay values around .95-.985.

10 Virtual Trading

With our initial measurements of each stock taken, it was now appropriate to begin employing practical experimentation using our predictions. Each member of our group began with $100,000 of virtual currency and selected a subset of our sector's stocks that we would use to track how our measurements performed against the present-day stock market. Given the time that had passed between the start of the project and the start of this process, we moved our "present" date to November 10th to allow our models to use the data up until that point. From there, we managed our stocks as though they were genuine investments, buying and selling as we felt appropriate.

10.1 Overall Performance

Table 12: Net Change Over Initial Period

Sector	Net Change	Percent Change
Technology	$4,500	4.5%
Material	-$2,830.87	-2.83%
Real Estate	-$2119.42	-2.12%
Consumer Discretionary	-$4,307.89	-4.31%
Consumer Staples	-$2,507.8	-2.51%

Table 13: Index Net Change Over Initial Period

Index	Net Change	Percent Change
DJI	-$852.91	-2.36%
NASDAQ	-$397.56	-2.55%
NDXT	-$-299.06	-3.13%
DJUSBM	-$-20.49	-3.64%
S&P REIT	$1.08	-.27%
S&P 500	-$-55.04	-1.18%

The initial period began November 10th and ran until December 6th. In general, our portfolios roughly followed the performance of the overall market in this short period of time, with nearly every chosen sector index showing a decrease around 1-3%. This result was expected due to the very brief period of collection. While it is uncertain, our team believes the sudden drop in the stock market was caused by the announcement of the Omicron variant of SARS-CoV-2, which occurred on November 24th.

10.2 Technology Sector

The stocks chosen for this endeavor were CALX, INTC, SONO, RIVN, VSAT, OSIS, MXL, BELL, & HPQ, in order of highest to least dollar amount invested.

The investment made into the Technology Sector is net positive, producing a profit of $4,500. However, this is only due to the large gains seen in the early IPO period of Rivian Automotive (RIVN), Amazon's new electric car company.

RIVN boasted a +48.3% change at its best, at which point the majority of holdings were sold for a large profit.

The largest investments were into CALX and INTC, with about 900 shares purchased between the two. These two stocks showed the highest slope in the fourier prediction models in the short term. The latter has risen by about +4.7%, and the former has fallen by 2.8%.

The majority of the smaller stocks were quickly sold, due to a lack of trust in predictions during the long term. New stocks are to be chosen to reinvest the money.

10.3 Materials Sector

The materials sector's traded stocks included BDC, DD, LAC, LITE, OTTR, and X, and they saw an overall negative performance during the first four weeks of tracking. They were chosen because they mapped fairly well to previous predictors and thus would likely be easier to manage. LITE and X received the largest investments at $30,000 each due to demonstrating reasonable growth while mapping well to the Fourier model. BDC, LAC, and OTTR received roughly the same amounts, while DD received only $5,000 due to being visibly volatile even prior to the introduction of the Bollinger Bands. Another $5,000 was moved from OTTR to X due to the former showing extremely high values from the RSI. While the sector experienced overall growth for three of the four weeks measured, the last week of November showed near-universal losses, resulting in a net loss of $2,830.87 from the beginning $100,000.

10.4 Real Estate Sector

The real estate sector had six stocks initially purchased. The six chosen stocks were: O, RDFN, PCH, FRPH, RYN, and AAT. O had the most money invested into it because it is known to be a reliable and low volatility stock, so it provided security at the beginning of this simulation. The second largest investment was RDFN because the stock was declining and was predicted to start climbing back up. The remaining funds went into PCH, FRPH, RYN, and AAT because the prototype model was able to predict these best out of the selected stocks. O was later sold in week two because of its low volatility and replaced by two stocks. The two stocks RMAX and SAFE were chosen to replace O because of their low RSI values. The Bollinger Bands did not show any of the chosen stocks had high volatility. The stocks were slightly profitable for the first two weeks but dropped slightly in week three and took a larger drop in the final week. This sector ended with a net loss of $2119.42 and was most likely due to the Real Estate sector losing a bit of volatility and the overall drop in the market during the fourth week.

10.5 Consumer Discretionary Sector

Five different companies from the original list of consumer discretionary stocks were chosen to compile this portfolio. The five chosen stocks were: Ebay (EBAY), Liberty Media Formula One Series (FWONK), Skechers (SKX), Macy's (M), and VF Corporation (VFC). All of these stocks were chosen for their relative low volatility levels. Most of the virtual amount of money went towards EBAY. This was due to the anticipated levels of online shopping done for the upcoming holiday season through this online merchant store, and thus expected its stock price to rise. The rest of the virtual money was split among the remaining four stocks by the manner of their descending volatility. At the beginning of the tracking period for these stocks, they performed generally well, exceeding past $3,000 in capital gains and earnings within the first two weeks. Throughout the remainder of the term, however, the stocks performed poorly, in that most of their prices began to decrease. By the end of the term, the amount that this portfolio lost was just over $4,000 of the original $100,000 that it started with. None of the stocks' shares were sold during this negative downtrend, due to the hope that they may rise in value and self-stabilize once again towards the end of the year.

10.6 Consumer Staples Sector

The stocks chosen for Consumer Staples included GIS, TSN, KO, HLF, BTI, and WMT. The initial selection was GIS, TSN, KO, HLF, and BTI. These stocks were generally not very volatile. The most money was allocated to Coca-Cola (KO), which ended up being a net loss in the end. I chose not to sell a good portion of these intial stocks for some time thinking they'd recover, and some did to an extent. Like in the other sectors, all stocks dropped a significant amount during the period around Thanksgiving, resulting in a net loss by the end, although there was some recovery. This was partially mitigated by choosing to short-sell WMT, and the stocks that remained continued to rise after the end of the term.

11 Findings

Through the evaluation of the data we gathered, we found that it is entirely possible to estimate the relative price of a given stock using readily available data. We will first examine our data obtained from our Fourier models, identifying which sectors were most compatible with our methods. Furthermore, we found that results vary by sector, with some exhibiting far more symptoms of volatility. After we had gathered those results, we found that major changes in trends are possible at any time and hypothesized a possible cause. Even so, our findings also demonstrate the potential for a sector's price trends to run counter to those affecting the rest of the market. These findings would likely serve as extremely helpful tools for those looking for an introduction to stock market trading.

11.1 Finding 1: Historical stock prices can be used to reasonably predict trends in the short term.

Our models and virtual stock market trading experiment confirmed that stock prices follow measurable, short-term patterns that can be used to improve the odds of making a profit. Our initial Fourier series models all demonstrated efficacy in generating error bands that reasonably gauged the range in which a given stock's closing price would be over a 10-day period.

Table 14: Mean Days within Fourier Model's Error Band

Sector	Mean Days Inside
Technology	5.20
Material	1.13
Real Estate	4.00
Consumer Discretionary	9.67
Consumer Staples	4.27
Mean	4.85
STD	2.77

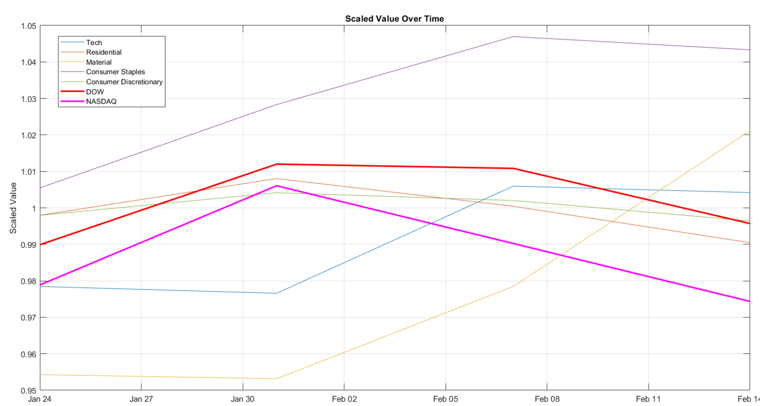
As shown by the results above, most sectors demonstrated accurate results for about 5 days. Our results improved further with the introduction of overall market influence to our models. The Fourier models combined with our chosen index resulted in a much higher average in prediction accuracy for all of our chosen sectors, showing slight drawbacks for the better performing ones in exchange for major improvements to the sectors that were previously poorly fit. This was especially notable with the Real Estate Sector, which had performed the worst when its data was fed into the original model.

Table 15: Mean Days within Best Performing Fourier/Index Model's Error Band

Sector	Mean Days Inside
Technology	5.07
Material	5.67
Real Estate	6.60
Consumer Discretionary	9.33
Consumer Staples	8.40
Mean	7.01
STD	1.62

The results of our new models proved to be very effective in choosing which stocks to choose and when to redistribute our assets during our second virtual trading period, which ran from January 24th to February 14th. Their total values were scaled using their initial amount to compare them to the theoretical returns one would have gotten by investing alongside the DOW or NASDAQ.

Figure 8: Scaled Stock and Index Values



As shown above, our portfolios generally performed better than the market because we were able to determine which stocks would appreciate or depreciate in value with reasonable accuracy. By the end of the period, four of our five chosen sectors had outperformed the DOW and NASDAQ. The clearest limitation we face in confirming this assessment is the uncertain nature of the stock market itself, as it is also entirely possible that our portfolios profited out of coincidence rather than intention. However, ultimately our objective was to earn any amount of profit over our trading period, and that goal appears to have been consistently met through the methods we employed.

11.2 Finding 2: Stocks price trends and predictability are clearly affected by their sector.

While historical prices can help predict short-term trends and certain events will have effects across all markets, each sector tends to perform differently. While the most pronounced events have effects seen across different sectors, the differences between individual sectors' performances are still quite pronounced. We can use the S&P 500 sectors' historical data to demonstrate this difference fairly clearly.

Table 16: One Year Return by Sector based on S&P 500

Sector	Percent Change
Technology	17.35
Material	14.89
Real Estate	24.33
Consumer Discretionary	6.04
Consumer Staples	18.18
Mean	16.16
STD	5.42

These differences can also be seen based on our overall return from the initial experiment towards the end of 2021, which resulted in a net loss in all sectors except Technology. Naturally, inherent differences between sectors also play into this result. A sector such as Consumer Staples will generally have less volatile stocks and therefore be overall lower risk and perform more consistently, while the Technology sector would have potential for greater gains and be generally more volatile.

11.3 Finding 3: Major real-world events have greater impact on stock prices than previous price trends.

Major events – either within a sector, or for a specific stock – will *consistently* overrule the prediction of any generated model that does not take these market events into consideration. Market trends alone can only carry an investor so far.

Any semi-random selection of stocks (ones not chosen for large events unique to them) will trend the same as the market. Overall, we found that when the market went down significantly, our stocks would trend down as well, with the vise-versa holding true as well. The aforementioned exception to this being the selection of stocks with major events occurring, or being predicted to occur (the latter of which we did not pursue in our research). We saw a solid example of this within our initial pseudo-investment round. In this round, the stock market as a whole was down by about 19% in January alone, and this trend was reflected heavily in four out of our five different investment sectors. This downward trend was due, at least in part, by the new pandemic announcement of the Omicron variant (Telford 2022). The fifth industry selection, information technology, however, was by-and-large increasing, seeing at most a positive 7% delta week-to-week.

The reason behind this was the investment into a stock based off of current events – the IPO of Amazon's new electric car company, Rivian Automotive (RIVN). The fresh IPO of this ticker was able to outshine the overall negative trend of the stock market so much as to seemingly ignore it.

Part of any effective investment scheme should be an attentive eye (or algorithm) on the happenings of both the market as a whole (along on an industry level), and the major events concerning the stocks being invested in. An investor lacking this element, is bond to get wrapped up in factors beyond what can be seen by the raw data.

12 Conclusions & Recommendations

Predicting the price of a stock is difficult due to the many underlining factors that affect a stock's behavior. With the assistance of mathematics and technology, people have developed efficient ways to predict the price of stocks, but no method is perfect. Models are limited in scope because they typically require assumptions that simplify the number of variables involved in a given real-world scenario. Improvements to accuracy are possible by identifying which variables are the most important while trimming away irrelevant or inconsequential data. This was the foundation for this project, in which we created and improved several mathematical models to assist in short-term predictions regarding the future price of stocks. Our base model combines the Fourier model with a year of processed historical data in order to make predictions. With just the Fourier model alone, we were able to accurately predict up to ten days into the future. These predictions overall showed more success as additional variables were added to the prototype model. When the Fourier model was combined with indices as a way to incorporate the market influence, we saw greater success seen in our means and standard deviations. The RSI incorporated with our model allowed for safer investments as we were able to predict some dips that occurred in the market and redistribute the funds somewhere else. When combining Bollinger bands with the model, we were able to isolate volatile stocks. This allowed us to predict when stocks were going to experience major changes in value, allowing us to determine the safest time for stocks to be bought or sold. The additional components added to the model combined with our analysis of each chosen stock allowed us to consistently demonstrate profit during our virtual trading experiment, confirming its real-word efficacy. Some sectors saw better success during our trading periods, but there was no sector where the model did not contribute to profits made. This model should not be used as a standalone investment tool but rather in conjunction with other sources to help with short-term decisions in the stock market. One of the major limitations affecting our model was the COVID-19 pandemic and its effect on the stock market. We found no feasible way to predict certain fluctuations in the market due to COVID-19. However, that was not seen as solely a limitation, as this project gave insight on how our methods performed during periods of extreme volatility. Should another similar situation arise in the future, our data would act as a useful reference. Another limitation of this project was the amount of time given. There was not a sufficient amount of time to test the models practically through our virtual trading experiments, as they only lasted for about six weeks each. With the allocation of more time, we believe that more favorable results could be seen from more testing and refinement. The true measure of this project's success is the amount of knowledge it has created in the areas of model creation, in addition to acting as a proof of concept in allowing amateur, individual investors to produce short-term profits in the stock market. In regard to future IQP groups, we recommend the use of Python, as others in the industry are starting to use the programming language more and more. The functionality of each model we created in MATLAB can ported over or recreated through the language's analogous stock libraries. Prior programming knowledge in Python or a similar coding language is greatly recommended if that course of action is taken. Furthermore, Python is well-suited for machine learning, which may prove to be a valuable asset in this experiment as it is fundamentally a regression analysis problem. Beyond that, some method of incorporating relevant events in news via a web crawler that identifies significant keywords could provide an early warning to potential shifts in stock behavior.

13 Bibliography

References

Ashford, K. (2021), 'What is market volatility—and how should you manage it?'.
URL: *https://www.forbes.com/advisor/investing/what-is-volatility/*

Beattie, A. (2021), 'The birth of stock exchanges'.
URL: *https://www.investopedia.com/articles/07/stock-exchange-history.asp*

Caplinger, D. (2021*a*), 'Stock market sectors'.
URL: *https://www.fool.com/investing/stock-market/market-sectors/*

Caplinger, D. (2021*b*), 'What is a stock market index? defined, listed & monitored'.
URL: *https://www.fool.com/investing/stock-market/indexes/*

Compute Stock Volatility (2011).

Hayes, A. (2021*a*), 'Bollinger band definition'.
URL: *https://www.investopedia.com/terms/b/bollingerbands.asp*

Hayes, A. (2021*b*), 'How does the stock market work?'.
URL: *https://www.investopedia.com/articles/investing/082614/how-stock-market-works.asp*

IBM (2022), 'Linear regression'.
URL: *https://www.ibm.com/topics/linear-regression*

Kenton, W. (2022), 'Least squares method'.
URL: *https://www.investopedia.com/terms/l/least-squares.asp*

McBride, G. (2021), 'What is rsi in stocks?'.
URL: *https://www.bankrate.com/investing/what-is-rsi-in-stocks/*

N/A (2021*a*), 'Bollinger bands?'.
URL: *https://corporatefinanceinstitute.com/resources/knowledge/trading-investing/bollinger-bands/*

N/A (2021*b*), 'Bollinger bands'.
URL: *https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/bollinger-bands*

N/A (2021*c*), 'Market indices'.
URL: *https://www.sec.gov/fast-answers/answersindiceshtm.html*

N/A (2021*d*), 'Relative strength index (rsi)'.
URL: *https://www.fidelity.com/learning-center/trading-investing/technical-analysis/technical-indicator-guide/RSI*

N/A (2021*e*), 'Stocks'.
URL: *https://www.investor.gov/introduction-investing/investing-basics/investment-products/stocks*

N/A (2021*f*), 'Understanding stock market volatility and how it could help you'.
URL: *https://www.fidelity.com.hk/en/start-investing/learn-about-investing/your-guide-to-stock-investing/understanding-stock-market-volatility-and-how-it-could-help-you*

O'Connell, B. (2021), 'Stock market sectors help you understand the market'.
URL: *https://www.forbes.com/advisor/investing/stock-market-sectors/*

O'Shea, A. (2021), 'What is the stock market and how does it work?'.
URL: *https://www.nerdwallet.com/article/investing/what-is-the-stock-market*

Royal, J. F. (2021), 'What is the average stock market return?'.
URL: *https://www.nerdwallet.com/article/investing/average-stock-market-return*

Seabury, C. (2021), 'Mastering short-term trading'.
URL: *https://www.investopedia.com/articles/trading/09/short-term-trading.asp*

Smith, T. (2021), 'Autocorrelation'.
URL: *https://www.investopedia.com/terms/a/autocorrelation.asp*

Staff, T. (2021), 'What is a stock market index? definition, examples & faq'.
URL: *https://www.thestreet.com/dictionary/s/stock-market-index/*

Telford, T. (2022), 'January has been a nightmare for investors. here's what wall street is watching'.
URL: *https://www.washingtonpost.com/business/2022/01/25/january-stock-trading/*

Tretina, K. (2021), 'Investing basics: What is a market index?'.
URL: *https://www.forbes.com/advisor/investing/stock-market-index/*

Weisstein, E. (2022), 'Fourier series'.
URL: *https://mathworld.wolfram.com/FourierSeries.html*

14 Appendices

14.1 Appendix A: Matlab code

```
import matplotlib.*

%Toggle Pausing
pause('off');
%pause('on');

%Get filenames to read in data
AA='D:\SchoolFiles\IQP\Stocks\AA.csv';

BDC='D:\SchoolFiles\IQP\Stocks\BDC.csv';

CIEN='D:\SchoolFiles\IQP\Stocks\CIEN.csv';
```

```
CLF='D:\SchoolFiles\IQP\Stocks\CLF.csv';

DD='D:\SchoolFiles\IQP\Stocks\DD.csv';

FCX='D:\SchoolFiles\IQP\Stocks\FCX.csv';

GPRE='D:\SchoolFiles\IQP\Stocks\GPRE.csv';

LAC='D:\SchoolFiles\IQP\Stocks\LAC.csv';

LITE='D:\SchoolFiles\IQP\Stocks\LITE.csv';

MGDDY='D:\SchoolFiles\IQP\Stocks\MGDDY.csv';

OTTR='D:\SchoolFiles\IQP\Stocks\OTTR.csv';

PLXS='D:\SchoolFiles\IQP\Stocks\PLXS.csv';

TECK='D:\SchoolFiles\IQP\Stocks\TECK.csv';

TX="D:\SchoolFiles\IQP\Stocks\TX.csv";

X="D:\SchoolFiles\IQP\Stocks\X.csv";

%Change file path to corresponding index
Dow="D:\SchoolFiles\IQP\Stocks\^DMI.xlsx";

DowIndustrial="D:\SchoolFiles\IQP\Stocks\^DJI.csv";

NASDAQ="D:\SchoolFiles\IQP\Stocks\^IXIC.csv";

%Put filepaths into String array
Stocks = [AA, BDC, CIEN, CLF, DD, FCX, GPRE, LAC, LITE,
   MGDDY, OTTR, PLXS,  TECK,  TX, X];
StockNames = ["AA", "BDC", "CIEN", "CLF", "DD", "FCX", "
```

```
    GPRE", "LAC", "LITE", "MGDDY", "OTTR", "PLXS", "TECK",
     "TX", "X"];

%Comment out above lines with stock file paths and use
    below for folder-based file read-in
%{
%Change to directory holding stock csv files
path_directory='D:\SchoolFiles\IQP\input-data\';
original_files=dir([path_directory '/*.csv']);

Stocks = strings([1,15]);
StockNames = strings([1,15]);

 for k=1:length(original_files)
    Stocks(k) = [path_directory '/' original_files(k).
        name];
    StockNames(k) = original_files(k).name;
 end

%}


stdPrices = zeros(1, length(Stocks));
days = zeros(1,length(Stocks));
errBnd = zeros(1,length(Stocks));
daysInBand = zeros(1,length(Stocks));

dowDaysAccurate = zeros(1, length(Stocks));
dowMaterialDaysAccurate = zeros(1,length(Stocks));
nasdaqDaysAccurate = zeros(1,length(Stocks));



%Read in data from Dow Material data sheet
dowIndex = readtable(Dow, 'PreserveVariableNames', true);
dowX = dowIndex.Date;
```

```
dowY = dowIndex.Close;

%Set date to act as baseline
startDate = datetime(2022,1,20,0,0,0);

%Scale Dow
startIndex = find(dowX==startDate);

scDowX = dowX(startIndex:end);
scDowY = dowY(startIndex:end);

for i = 1 : length(scDowX)
    scDowY(i) = (scDowY(i) / dowY(startIndex));
end


%Read in data from Dow Industrial data sheet
dowIndIndex = readtable(DowIndustrial, '
   PreserveVariableNames', true);
dowIndX = dowIndIndex.Date;
dowIndY = dowIndIndex.Close;

%Scale
startIndexDowInd = find(dowIndX==startDate);

scDowIndX = dowIndX(startIndexDowInd:end);
scDowIndY = dowIndY(startIndexDowInd:end);

for i = 1 : length(scDowIndX)
    scDowIndY(i) = (scDowIndY(i) / dowIndY(
       startIndexDowInd));
end

%Read in data from Nasdaq sheet
nasdaqIndex = readtable(NASDAQ, 'PreserveVariableNames',
   true);
```

```
nasdaqX = nasdaqIndex.Date;
nasdaqY = nasdaqIndex.Close;

%Scale
startIndexNASDAQ = find(nasdaqX==startDate);

scNASDAQX = nasdaqX(startIndexNASDAQ:end);
scNASDAQY = nasdaqY(startIndexNASDAQ:end);

for i = 1 : length(scNASDAQX)
    scNASDAQY(i) = (scNASDAQY(i) / nasdaqY(
        startIndexNASDAQ));
end

%Primary loop executed on each stock
%Produces Fourier model, Fourier+index models, RSI,
   Bollinger Bands, and
%PCA
for i = 1 : length(Stocks)

    %Read in data
    curPath = Stocks(i);
    curStock = readtable(curPath, 'PreserveVariableNames'
        , true);
    presentDate = find(curStock.Date==startDate);

    curX = curStock.Date(1:presentDate);
    curY = curStock.Close(1:presentDate);


    %Plot data, set labels
    plot(curX,curY);
    title(strcat(StockNames(i),' Closing Price Vs. Date')
        );
    xlabel('Date');
    ylabel('Price');
```

```
pause

%Add data to standard deviation array
stdPrices(i) = std(curY);

%Find relevant period of data
reverse = curY(end:-1:1);
[m,z]=size(curX);
[acf, lags, bounds] = autocorr(reverse, m-1);

intersect = 0;
for j = 1 : length(acf)
    if acf(j) < 0
        %Subtract 2 to account for acf mismatch
        intersect = j - 2;
        break
    end
end

%Get starting point of relevant data
yy=linspace(1,intersect,intersect);

%Make linspace for future prediction
presentDate = find(curStock.Date==startDate);
%extPoint is the starting point of relevant data +
   the number of data
%points past the set present date
extPoint = intersect + length(curStock.Date) - length
   (curStock.Date(1:presentDate));
yyExt=linspace(1,extPoint, extPoint);

%Add 1 to ignore day of first negative correlation
period = length(curX) - intersect + 1;

relX = curX(period:length(curX));
relY = curY(period:length(curY));
```

```
plot(relX,relY);
title(strcat(StockNames(i),' Closing Price Vs. Date
   for Relevant Period'));
xlabel('Date');
ylabel('Price');
pause

%Get trend line
curP = polyfit(datenum(relX), relY, 1);
curF = polyval(curP, datenum(relX));
p=polyfit(yy,transpose(relY),1);
trend = zeros(1,intersect);


%Get Savitzky-Golay with lowest residuals
SGResult = zeros(1,40);
SGX = zeros(1,40);
SGSum = 0;

%Start at .9, add until we reach 1.1
gamma = .9;
checkY = curY(length(curY) - extPoint:end);
for j=1:40
    for k=1:extPoint
        trend(k)=p(1)*(k^gamma)+p(2); % Trendline
           with each gamma
        SGSum = SGSum + (checkY(k) - trend(k))^2;
    end
    SGResult(j) = SGSum / extPoint;
    SGSum = 0;
    SGX(j) = gamma;
    gamma = gamma + .005;
end

%Plot data, set labels
```

```
plot(SGX,SGResult);
title('DD Savitzky Golay Gamma Vs. Residuals');
xlabel('Gamma value');
ylabel('Residual Sum');

%disp("Optimal Gamma = " + SGX(find(SGResult == min(
    SGResult))));

bestGamma = SGX(SGResult == min(SGResult));


%Combine slopes of dow trendline and price trendline
for j=1:extPoint
    %trend(j)=p(1)*j+p(2); %Linear trendline
    trend(j)=p(1)*(j^bestGamma)+p(2); % Savitzky-
        Golay trendline
end





%Plot data, set labels
plot(relX,relY,'b',relX,curF,'r');
title(strcat(StockNames(i),' Closing Price Vs. Date
    for Relevant Period with Trendline'));
xlabel('Date');
ylabel('Price');

pause

%Display difference between trend and closing price
diff = transpose(relY) - trend(1:intersect);
plot(relX, diff)
title('Difference between trend and closing prices
    for ' + StockNames(i))
```

```
pause

%Make Fourier function
%Uses Fourier3 if more than 50 relevant data points
if intersect > 50
    fourierFunc = fit(yy',diff','fourier3');

else
    fourierFunc = fit(yy',diff','fourier2');

end

display(fourierFunc)

%Combine trend with fourier adjustment and Dow
modFunc = zeros(1,extPoint);

for j=1:extPoint
    modFunc(j) = trend(j) + fourierFunc(j);
end


plot(relX,modFunc(1:intersect),'r')
hold
plot(relX,relY,'b')
xlabel('Date')
ylabel('Closing Price')
title('Closing Stock Prices for ' + StockNames(i) + '
     with Fourier Model')
pause
close

%Find difference between historical data and trend +
   fourier
tsDiff = transpose(relY)- modFunc(1:intersect);
```

```
plot(yy, tsDiff, 'b')
xlabel('Date')
ylabel('Difference')
title('Difference b/w Fourier Model and Closing Price
    for ' + StockNames(i))

pause

%Get absolute max of difference, compare to recent
   historical data
[maxDiff, maxIndex] = max(abs(tsDiff));

recX = curStock.Date(presentDate:end);
recY = curStock.Close(presentDate:end);

boundUp = zeros(1,extPoint);
boundDown = zeros(1,extPoint);

for j=1:extPoint
    boundUp(j) = modFunc(j) + maxDiff;
    boundDown(j) = modFunc(j) - maxDiff;
end

%Display closing prices compared to Fourier band
plot(yyExt, boundUp, 'r')
hold
plot(yyExt, boundDown, 'r')
plot(yy, relY, 'b')
xlabel('Date')
ylabel('Closing Price')
title('Closing Price and Fourier Band for ' +
   StockNames(i))
hold off

pause
```

```
compareInterval = length(modFunc) - length(recX) + 1;

%Compare recent closing prices to extended Fourier
    band
plot(yyExt(compareInterval:end), boundUp(
    compareInterval:end), 'r')
hold
plot(yyExt(compareInterval:end), boundDown(
    compareInterval:end), 'r')
plot(yyExt(compareInterval:end), recY, 'b')
xlabel('Date')
ylabel('Closing Price')
title('Recent Closing Price and Fourier Band for ' +
    StockNames(i))
hold off

%Sample scaled to data from the 15th
breakPoint = length(recY);
errorBound = 0;
inside = true;
beenOut = false;
totalIn = 0;
%See if true data left the band
for j=1:length(recY)
    %Triggers when recY is outside of band and has
        not yet exited it
    %if(and(not(beenOut) ,(or((recY(j) > boundUp(j +
        length(relX))),(recY(j) < boundDown(j + length
        (relX)))))))
    if(and(not(beenOut) ,(or((recY(j) > boundUp(j)),(
        recY(j) < boundDown(j))))))
         breakPoint = j;
         errorBound = abs(abs(recY(j)) - (abs(recY(j))
              + abs(maxDiff)));
         beenOut = true;
         inside = false;
```

```
    end

    %Triggers when recY is inside band and was
       outside of it
    %if(and(not(inside) ,not((or((recY(j) > boundUp(j
        + length(relX))),(recY(j) < boundDown(j +
       length(relX))))))))
    if(and(not(inside) ,not((or((recY(j) > boundUp(j)
       ),(recY(j) < boundDown(j)))))))
        inside = true;
    end

    %Counts number of days within band
    if(inside)
        totalIn = totalIn + 1;
    end
end

%Store Values
days(i) = breakPoint;
errBnd(i) = errorBound;
daysInBand(i) = totalIn;
pause

%Scale recent period
basePrice = recY(1);
scaleRecY = recY;
for j = 1:length(recY)
    scaleRecY(j) = scaleRecY(j) / basePrice;
end


%Graph scaled recent period and scaled Dow

plot(recX,scaleRecY,'b')
hold
```

```
plot(scDowX(1:length(recX)),scDowY(1:length(recX)),'r
    ')
xlabel('Date')
ylabel('Scaled Value')
title('Scaled Values for ' + StockNames(i) + ' and
    Material Dow Jones')
pause
close

%Get correlation b/w scaled period and Dow
corDowMaterial = corr(scaleRecY,scDowY(1:length(
    scaleRecY)));

%Combine predicted stock prices with Dow
scRecDow = zeros(1,length(recY));

futureModFunc = modFunc(length(relX):length(modFunc))
    ;


%Scales futureModFunc vals to the modFunc val on
    start date
for j = 1:length(futureModFunc)
    futureModFunc(j) = futureModFunc(j) / modFunc(
        length(relX));
end

%Combines dow material and fourier models
for j = 1:length(recY)
    scRecDow(j) = ((futureModFunc(j) * corDowMaterial
        ) + (scDowY(j) * (1 - corDowMaterial)));
    scRecDow(j) = scRecDow(j) * modFunc(length(relX)+
         j - 1);
end

%Plot prices with Dow
```

```
plot(recX,scRecDow,'b')
hold
plot(recX,recY,'r')
xlabel('Date')
ylabel('Closing Price')
title('Predicted Price Using Dow (b) Vs. Actual Price
    (r) For ' + StockNames(i))
pause
close

%Plot difference between Dow function and data
dowDiff = transpose(recY)- scRecDow;
plot(recX, dowDiff, 'b')
xlabel('Date')
ylabel('Difference')
title('Difference b/w Dow/Fourier Model and Closing
   Price for ' + StockNames(i))
pause



%Repeat for Dow Industrial and NASDAQ

%Dow Industrial
corDow = corr(scaleRecY,scDowIndY(1:length(scaleRecY)
   ));

%Combine predicted stock prices with Dow
scRecDowInd = zeros(1,length(recY));

%Combines fourier and dow industrial models
for j = 1:length(recY)
    scRecDowInd(j) = ((futureModFunc(j) * corDow) + (
       scDowIndY(j) * (1 - corDow)));
    scRecDowInd(j) = scRecDowInd(j) * modFunc(length(
       relX)+ j - 1);
```

```
end

dowIndDiff = transpose(recY) - scRecDowInd;

%NASDAQ
corNASDAQ = corr(scaleRecY,scNASDAQY(1:length(
    scaleRecY)));

%Combine predicted stock prices with Dow
scRecNASDAQ = zeros(1,length(recY));

%Combines fourier and NASDAQ models
for j = 1:length(recY)
    scRecNASDAQ(j) = ((futureModFunc(j) * corNASDAQ)
        + (scNASDAQY(j) * (1 - corNASDAQ)));
    scRecNASDAQ(j) = scRecNASDAQ(j) * modFunc(length(
        relX)+ j - 1);
end

nasdaqDiff = transpose(recY) - scRecNASDAQ;

%Record number of days each index mapped to function
dowMaterialCutoff = (max(relY) - min(relY)) / 2;
dowCutoff = (max(relY) - min(relY)) / 2;
nasdaqCutoff = (max(relY) - min(relY)) / 2;

dowMaterialIsMapping = true;
dowIsMapping = true;
nasdaqIsMapping = true;

for j = 1:length(dowDiff)
    if(and(dowMaterialIsMapping, abs(dowDiff(j)) >
        dowMaterialCutoff))
         dowMaterialDaysAccurate(i) = j;
         dowMaterialIsMapping = false;
    end
```

```
    if(and(dowIsMapping, abs(dowIndDiff(j)) >
       dowCutoff))
        dowDaysAccurate(i) = j;
        dowIsMapping = false;
    end

    if(and(nasdaqIsMapping, abs(nasdaqDiff(j)) >
       nasdaqCutoff))
        nasdaqDaysAccurate(i) = j;
        nasdaqIsMapping = false;
    end
end

%Set value to length of time if was accurate the
   whole time
if(dowMaterialIsMapping)
    dowMaterialDaysAccurate(i) = length(dowDiff);
end

if(dowMaterialIsMapping)
    dowDaysAccurate(i) = length(dowDiff);
end

if(dowMaterialIsMapping)
    nasdaqDaysAccurate(i) = length(dowDiff);
end
close

%Get RSI values of current index
rsi = rmmissing(rsindex(curY));

plot(curX((length(curX) - length(rsi) + 1):end), rsi,
    'b')
xlabel('Date')
ylabel('RSI value')
```

```
    title('RSI value for ' + StockNames(i))
    pause


    [middle,upper,lower] = bollinger(relY);
    CloseBolling = [middle, upper,lower];
    plot(relX,CloseBolling)
    title('Bollinger Bands for ' + StockNames(i) + '
       Closing Prices')
    pause
    close


    %PCA
    varTable = [curY dowY(1:startIndex) dowIndY(1:
       startIndexDowInd) nasdaqY(1:startIndexNASDAQ)];
    [coeff,score,latent,tsquared,explained] = pca(zscore(
       varTable));
    scatter(score(:,1),score(:,2), 'filled')
    title('PCA 1 vs 2')
    xlabel('PC 1 (94.8%)');
    ylabel('PC 2 (3.5%)');
    pause
    close

end

%Combine values
finalResultErrorBands = table(transpose(StockNames),
   transpose(days), transpose(errBnd), transpose(
   daysInBand));

%Find best index
bestIndexes = strings([1,length(Stocks)]);
for i = 1:length(Stocks)
    if(and(dowMaterialDaysAccurate(i) > dowDaysAccurate(i
```

```
        ), dowMaterialDaysAccurate(i) > nasdaqDaysAccurate
        (i)))
         bestIndexes(i) = "Dow Material";
    elseif(and(dowDaysAccurate(i) >
        dowMaterialDaysAccurate(i), dowDaysAccurate(i) >
        nasdaqDaysAccurate(i)))
         bestIndexes(i) = "Dow";
    elseif(and(nasdaqDaysAccurate(i) >
        dowMaterialDaysAccurate(i), nasdaqDaysAccurate(i)
        > dowDaysAccurate(i)))
         bestIndexes(i) = "NASDAQ";
    else
         bestIndexes(i) = "TIE";
    end
end

finalResultIndexes = table(transpose(StockNames),
   transpose(dowMaterialDaysAccurate), transpose(
   dowDaysAccurate), transpose(nasdaqDaysAccurate),
   transpose(bestIndexes));
finalResultIndexes.Properties.VariableNames = {'Stock' '
   Dow Material' 'Dow' 'NASDAQ' 'Best Index'};

%Get mean and standard deviation for index accuracies
meanDaysDowMaterial = mean(dowMaterialDaysAccurate);
meanDaysDow = mean(dowDaysAccurate);
meanDaysNASDAQ = mean(nasdaqDaysAccurate);
stdDaysDowMaterial = std(dowMaterialDaysAccurate);
stdDaysDow = std(dowDaysAccurate);
stdDaysNASDAQ = std(nasdaqDaysAccurate);

meanRow = {'Mean', meanDaysDowMaterial, meanDaysDow,
   meanDaysNASDAQ, ''};
stdRow = {'STD', stdDaysDowMaterial, stdDaysDow,
   stdDaysNASDAQ, ''};
```

```
finalResultIndexes = [finalResultIndexes;meanRow];
finalResultIndexes = [finalResultIndexes;stdRow];

disp("Mean Days Dow Material = " + meanDaysDowMaterial);
disp("Mean Days Dow = " + meanDaysDow);
disp("Mean Days NASDAQ = " + meanDaysNASDAQ);

%Print mean and std values for days in the error bands.
%{
meanDays = mean(days);
stdDays = std(days);
meanErrBnd = mean(errBnd);
stdErrBnd = std(errBnd);

disp("Mean Days = " + meanDays);
disp("Day STD = " + stdDays);
disp("Mean Error Bound = " + meanErrBnd);
disp("Error Bound STD = " + stdErrBnd);
%}
```