



# Nate's Investment Advisory Firm

---

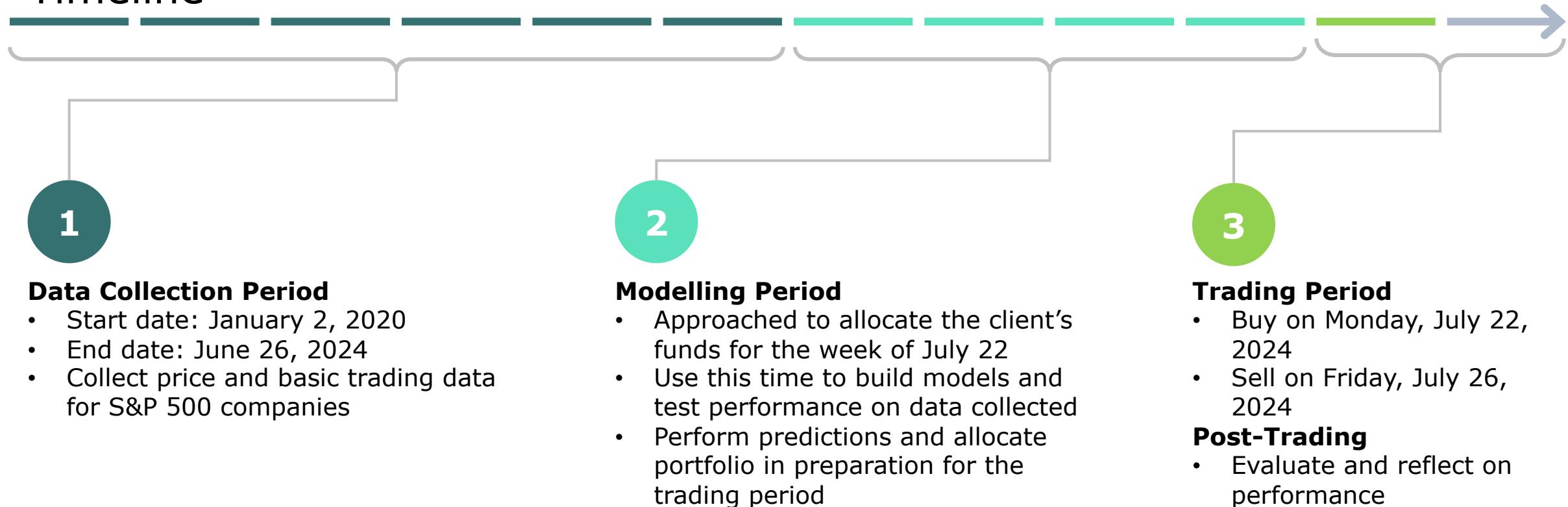
Stock Value Prediction

Callum Stevenson, Jake Vanderweyst, Joel Palmer,  
Nathan Farquharson, Tim Sankey

# Task Description

Key Areas	Description
Objective	The primary goal is to allocate \$1 million USD across six target companies: Salesforce, Uber, Chipotle, Kraft Heinz, Estée Lauder, and Expedia for the week of July 22 to July 26 to maximize returns.
Data Overview	We are analyzing the historic, and forward-looking, stock data of the companies that comprise the S&P 500 to understand the market and identify patterns/insights to inform our portfolio allocation.
Model Development	We will deploy predictive models/techniques learned in class, including machine-based regression and classification, to create the most accurate and informative model.
Deployment	The insights from our predictive models will be used to allocate the \$1 million USD between the six target companies that we believe will yield the greatest return over the selected period.

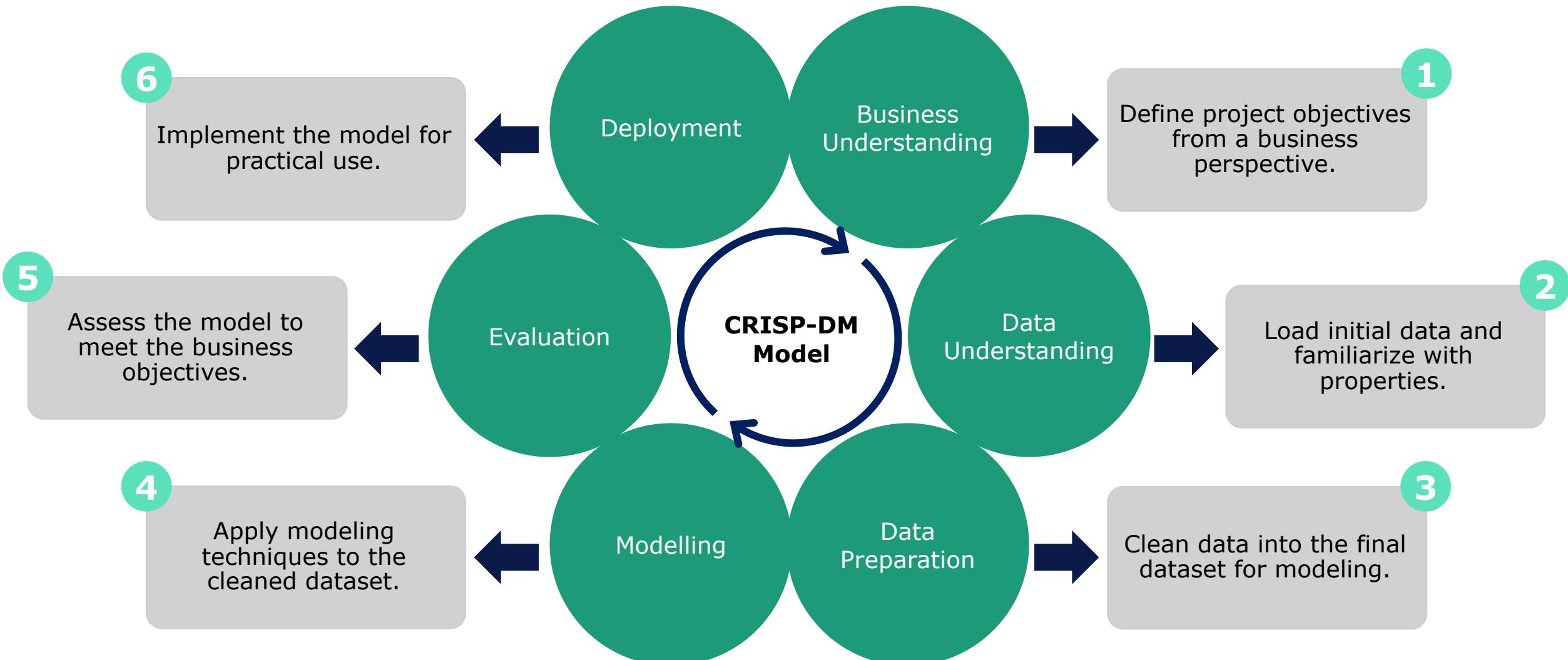
# Timeline



## Role

- We are financial advisors for a high-profile client with a very high-risk tolerance.
- The client recently received \$1 million USD from a friend who loaned them the money for a week, the client wants to generate a strong return with the money by investing in select S&P 500 stocks.

# Project Structure – CRISP-DM



# Business Understanding

CRISP-DM Model Part 1



# Six Target Stocks



Chipotle is a fast-casual restaurant chain known for its customizable burritos, bowls, and tacos made with fresh, high-quality ingredients.

**Opening Price: \$53.48**



Kraft Heinz is a global food industry giant renowned for its iconic brands and wide range of popular, high-quality products.

**Opening Price: \$33.15**



Uber is a global technology company that revolutionizes transportation by connecting riders with drivers through a user-friendly mobile app.

**Opening Price: \$67.53**



ESTĒE LAUDER

Estée Lauder is a leading global beauty company recognized for its premium skincare, makeup, fragrance, and hair care products.

**Opening Price: \$98.89**



Expedia is a leading online travel company that offers comprehensive travel services, including booking flights, hotels, and vacations.

**Opening Price: \$134.83**



Salesforce is a global leader in cloud-based customer relationship management (CRM) software.

**Opening Price: \$248.75**

# Data Understanding

CRISP-DM Model Part 2

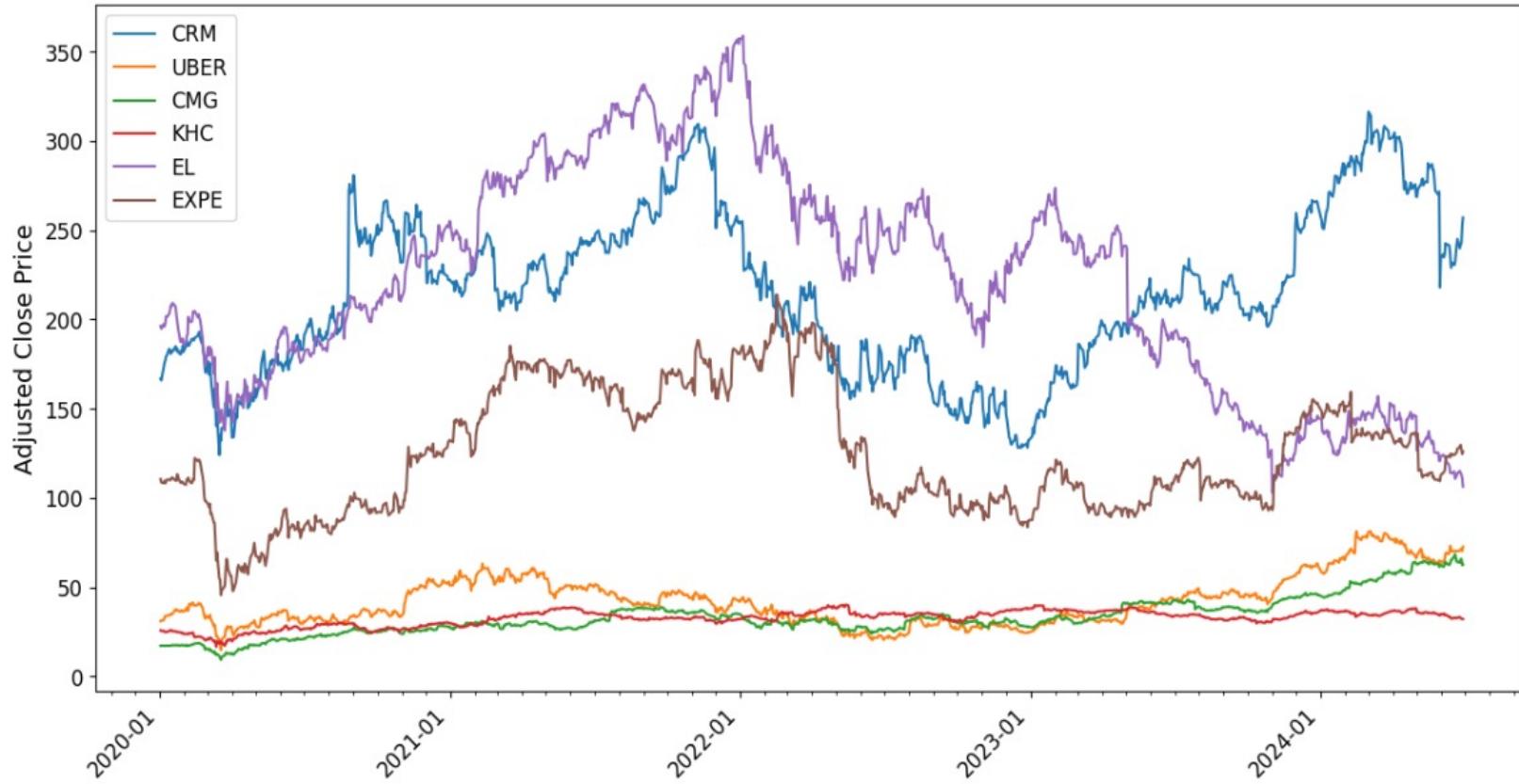


# Initial Variable Analysis

Variable	Description	Type
01 Date	Represents the specific trading day for which the stock data is recorded. It is typically formatted as YYYY-MM-DD.	Date String
02 Symbol	Unique identifier assigned to a publicly traded company on a stock exchange. For example, "CRM" for Salesforce.	Float
03 Adj Close	Represents the closing price after adjustments for all applicable splits and dividend distributions.	Float
04 Volume	Indicates the number of shares traded on a given trading day to reflect the level of activity.	Integer
05 Open, High, Low, Close	Variables represent the stock price at various points during a trading day	Float

# Trends of Target Companies

Six Selected Stock Prices Over Time



**KraftHeinz**

Upward trend since beginning



Positive trend since 2022



Large negative trend since 2022



Negative since 2022 but bouncing back

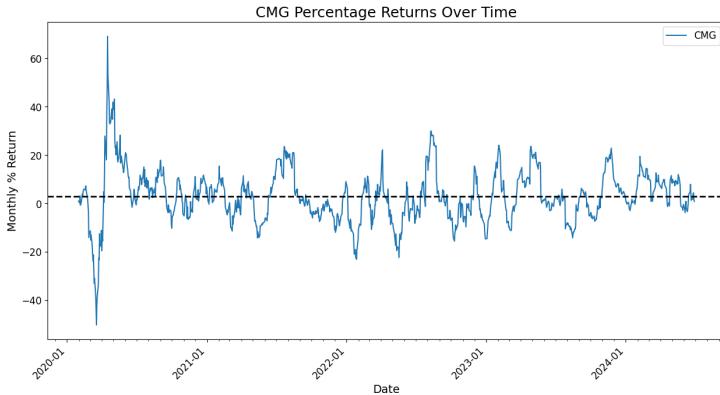


Negative since 2022 but bouncing back

# Average Monthly Return for Target Companies

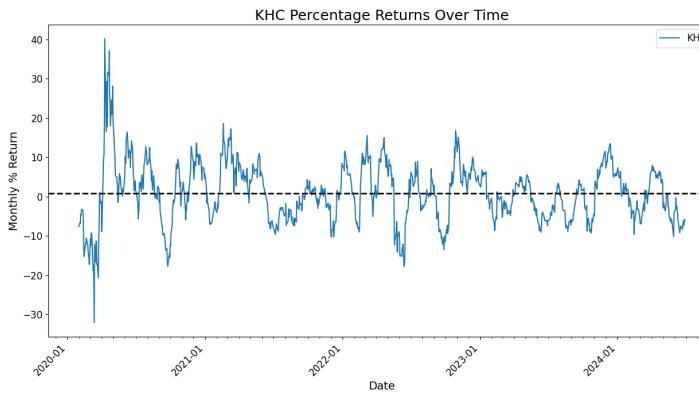
## Chipotle

**Avg Return = 2.97%**



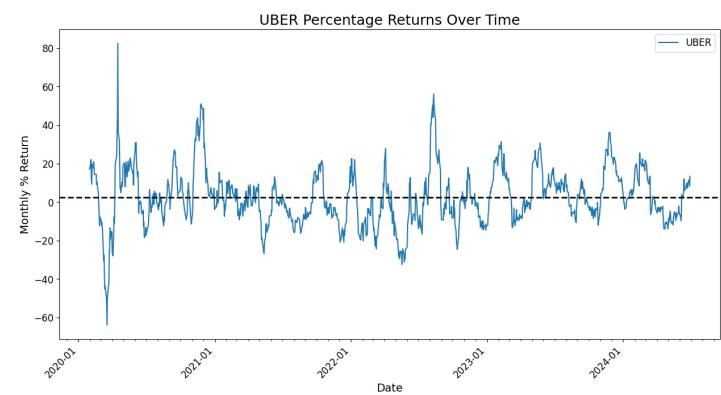
## KraftHeinz

**Avg Return = 0.79%**



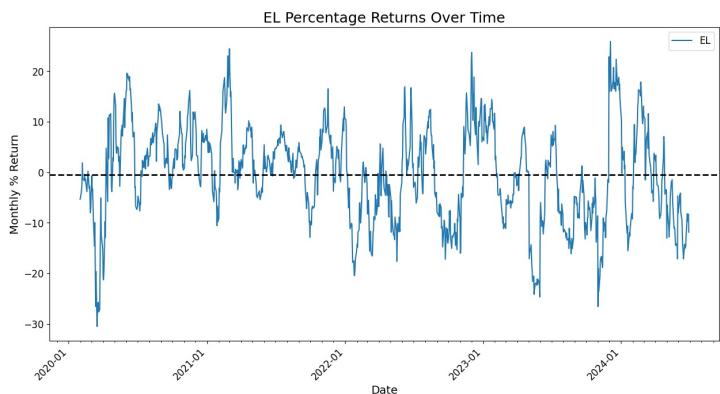
## Uber

**Avg Return = 2.34%**



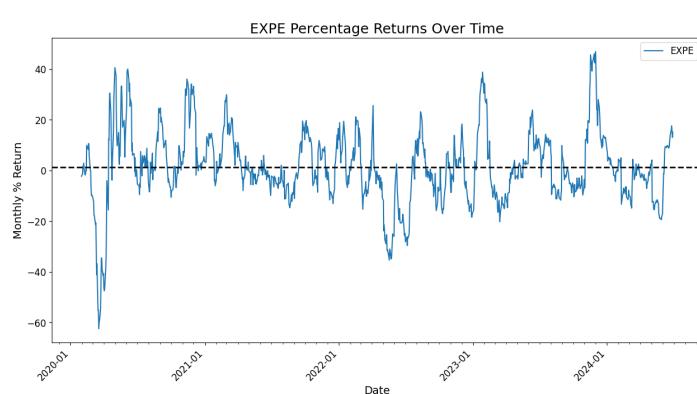
## Estee Lauder

**Avg Return = -0.50%**



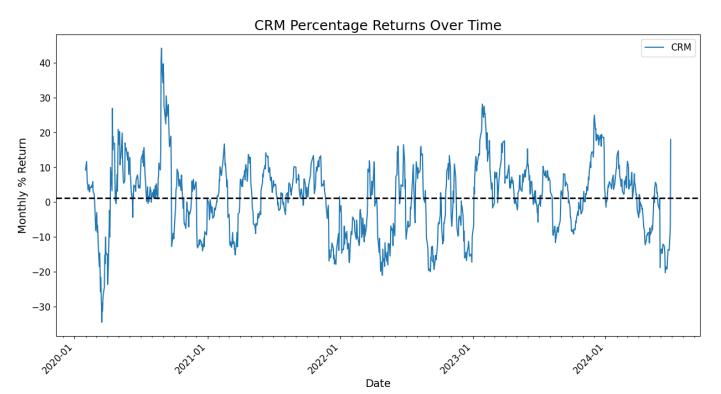
## Expedia

**Avg Return = 1.33%**



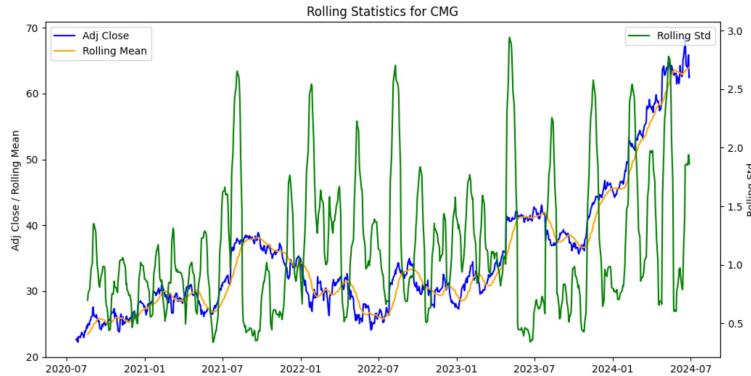
## Salesforce

**Avg Return = 1.06%**



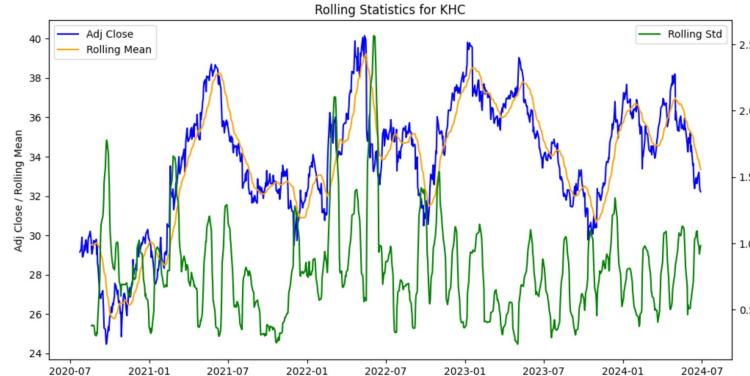
# Rolling Statistics for Target Companies

## Chipotle



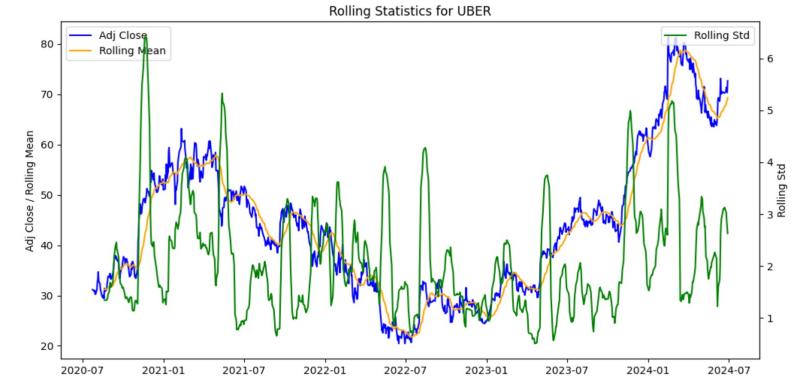
- Very high volatility
- Strong upward trend in 2024

## KraftHeinz



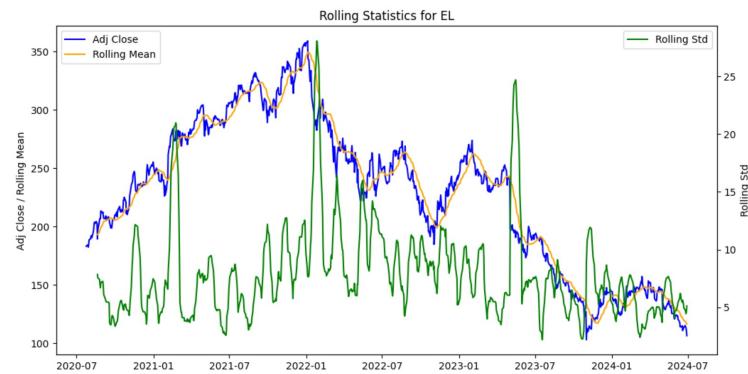
- Very seasonal returns
- Price oscillates along ~\$34

## Uber



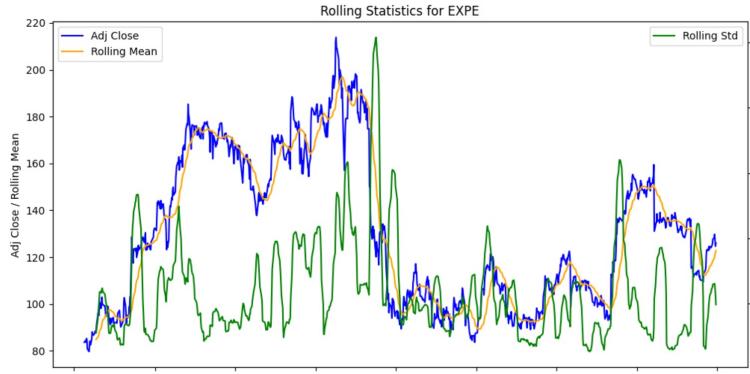
- High volatility through period
- Trending upward since 2022

## Estee Lauder



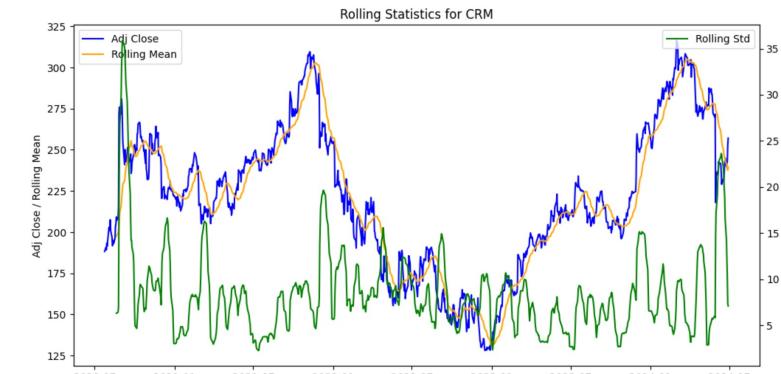
- Bursts of volatility
- Trending down since 2022

## Expedia



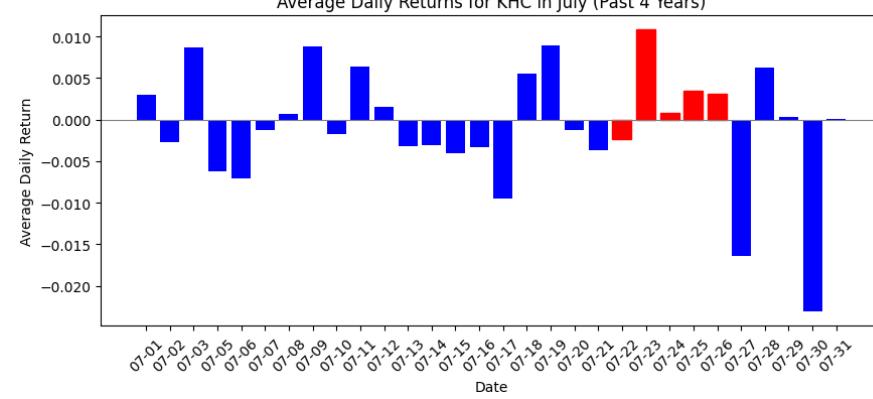
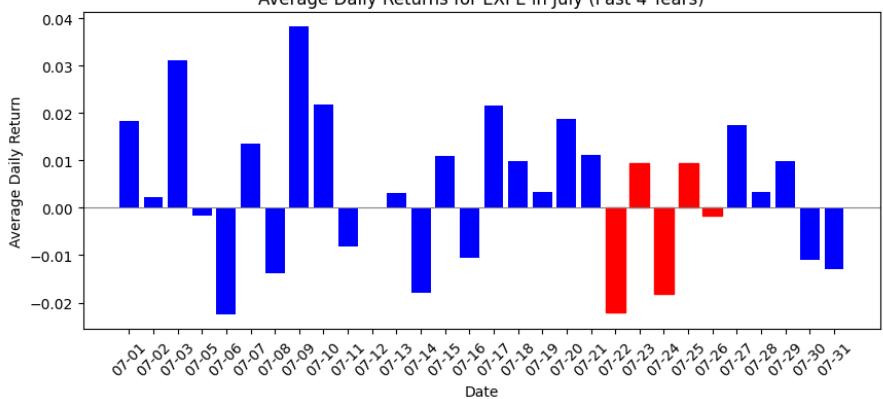
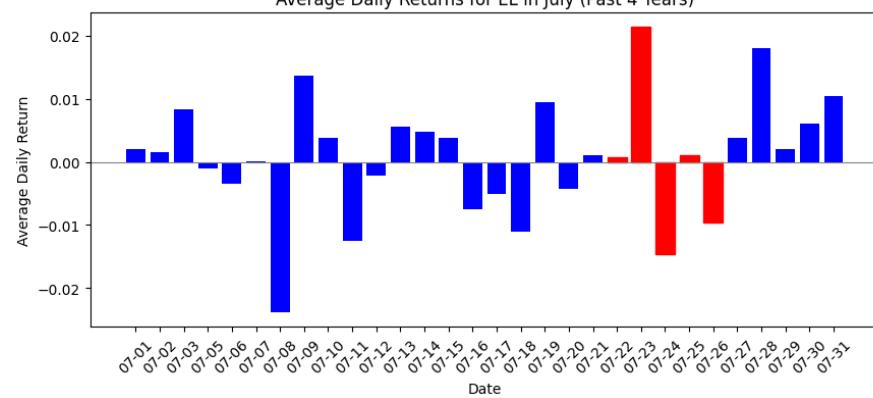
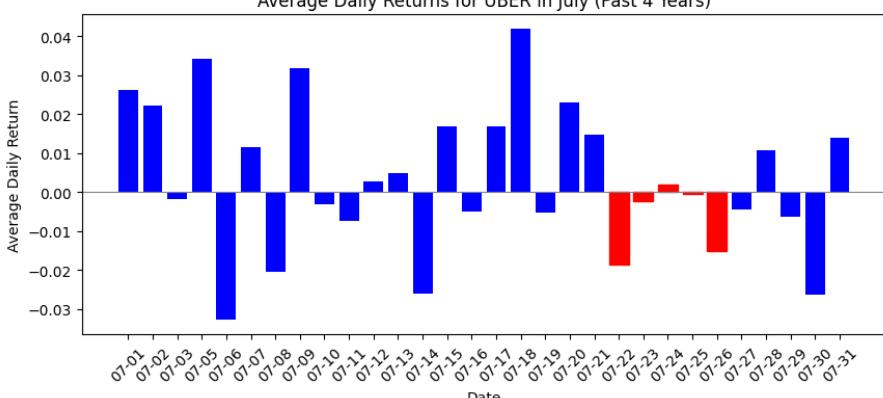
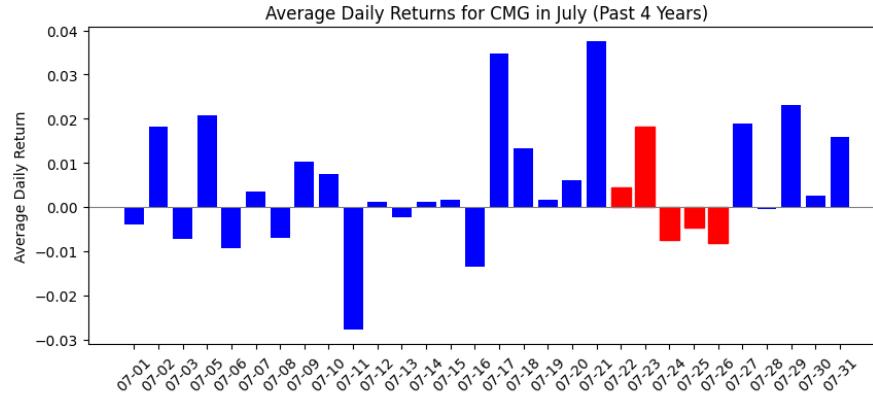
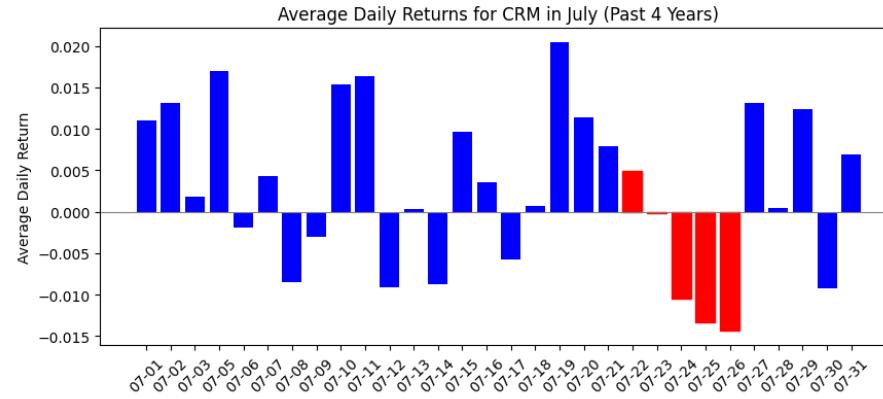
- Bursts of volatility
- Dramatic fall in 2022

## Salesforce



- Recovered from 2022
- Trending down in 2024

# Historic Performance Week of July 22



These graphs show the historical performance on average for each stock from July 22 - July 26

Weaker Performing



Stronger Performing



KraftHeinz

# Data Preparation

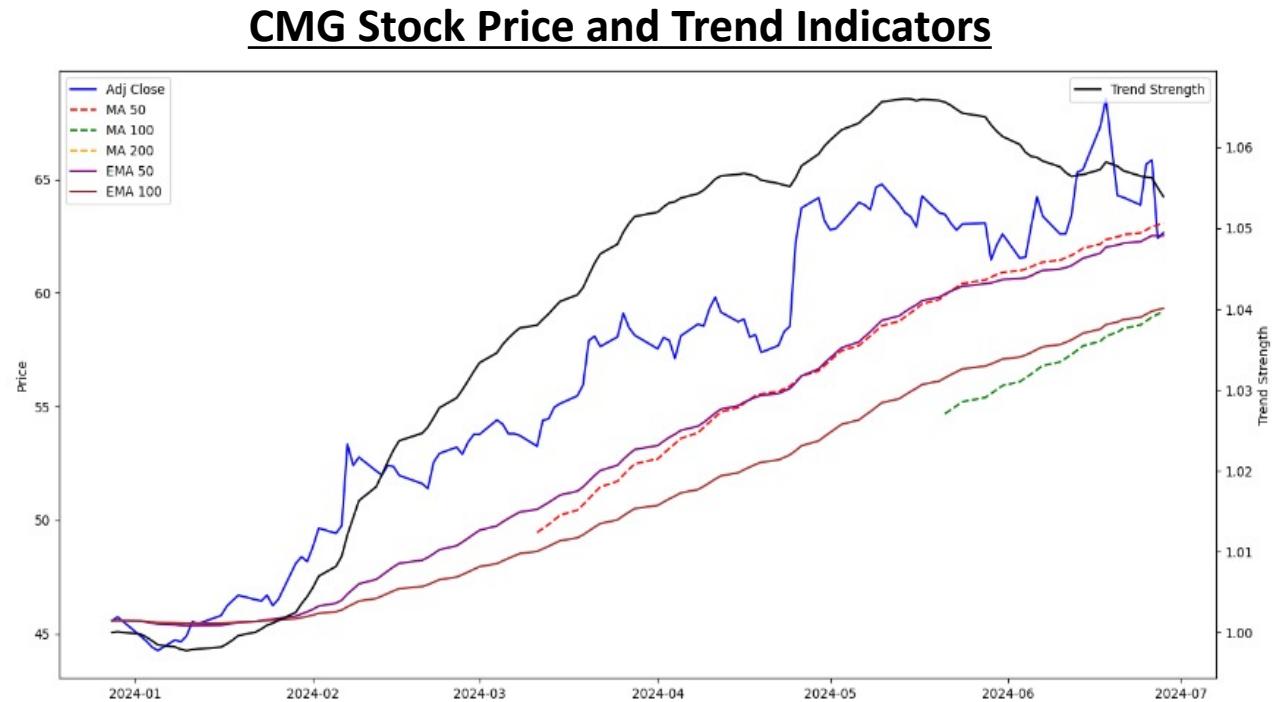
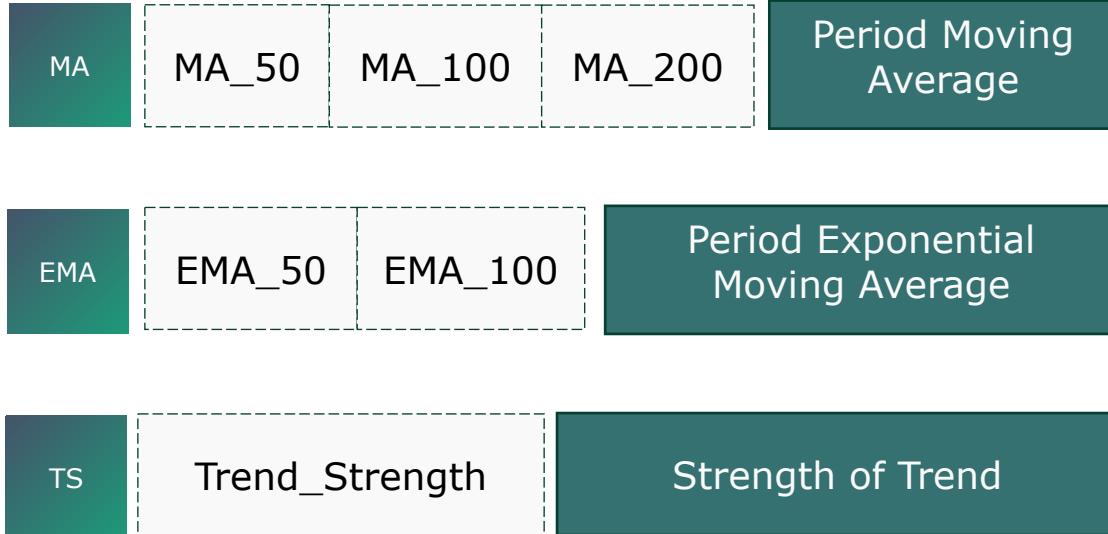
CRISP-DM Model Part 3



# Categories of Variables: Trend Indicators



**Trend Indicators:** Mathematical calculations based on historical data that identify the strength and direction of a trend.



Trend Strength Calculated as the Ratio of the 50-day EMA to the 100-day EMA.

Helps Analyst and Traders to Identify direction and strength of a price movement in the market.

# Categories of Variables: Seasonal Components



**Time Series Decomposition:** Technique used to break down a time series into constituent components, typically including trend, seasonal, and residual components.

## Trend

- Long-term movement, displays overall direction of the data over time
- General direction of their data

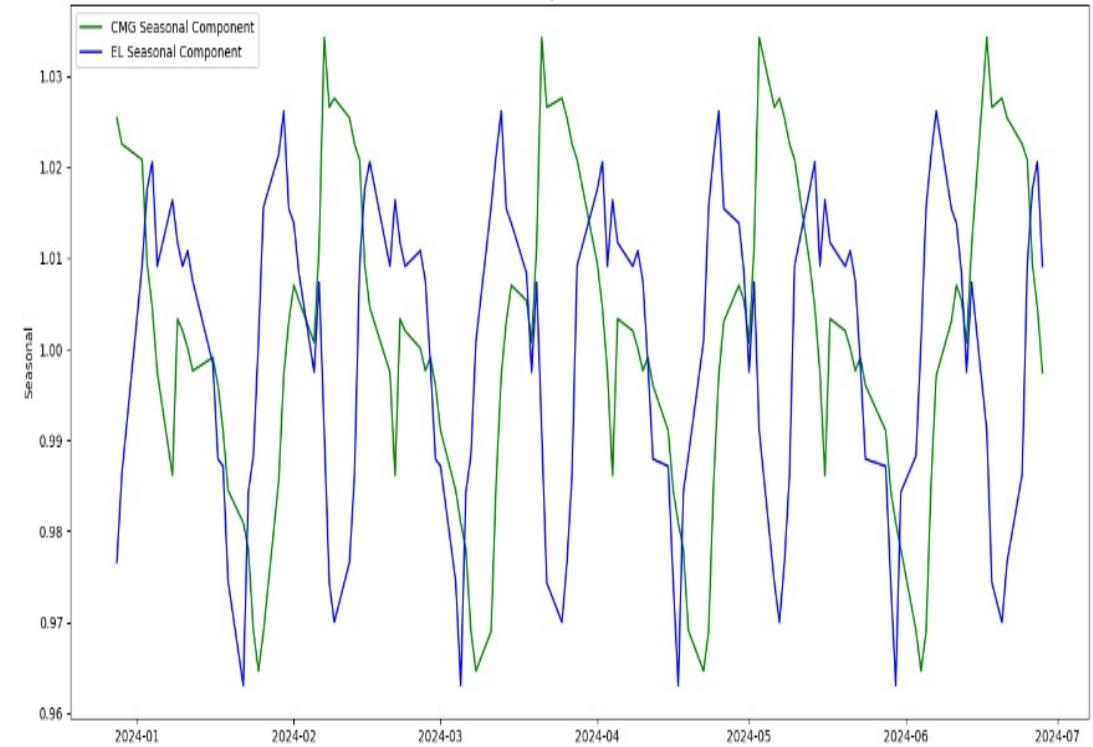
## Seasonal

- Repeating short-term cycles in the time series at regular intervals
- Seasonal analysis is useful for businesses with cyclical patterns

## Residual

- The random noise or irregular fluctuations
- Identify unusual or unexpected variations in the data

## Seasonal Components of CMG and EL

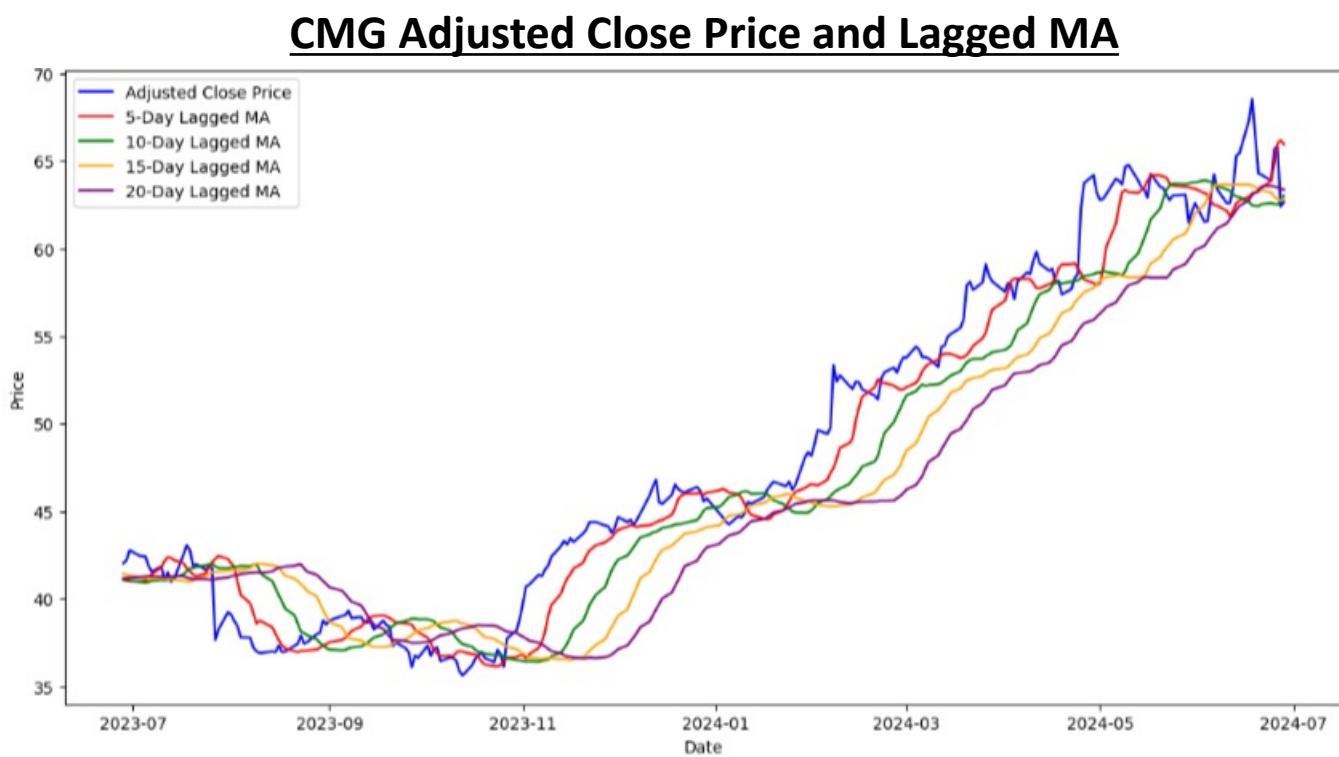


# Categories of Variables: Lagged Moving Averages



**Lagged Moving Averages:** Incorporate past data points to smooth out short-term fluctuations and highlight longer-term trends

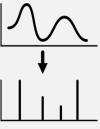
01	Lag_MA_5	5 Period Moving Average
02	Lag_MA_10	10 Period Moving Average
03	Lag_MA_15	15 Period Moving Average
04	Lag_MA_20	20 Period Moving Average



Calculated by Shifting Time Series Data by # of Days and Applying a Moving Average

Helps Analyst and Traders to Identify Trends and Potential Reversals in the Market

# Categories of Variables: Fourier Transform



**Fourier Transform:** Mathematical technique that transforms time series into its constituent frequencies

## Magnitude Spectrum

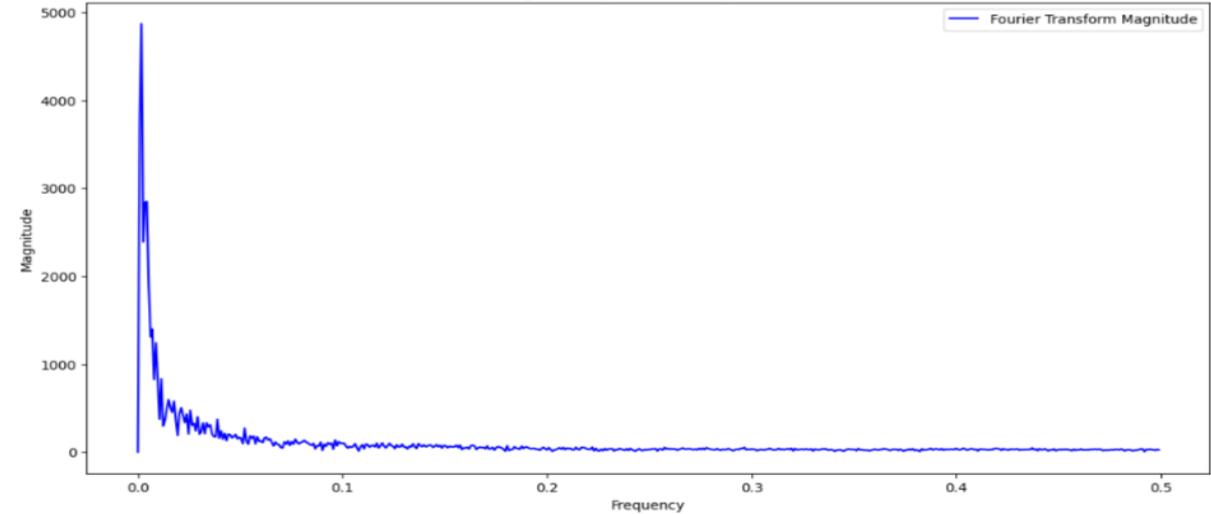
Low Frequency → Longer Term Trend

High Frequency → Short-Term Cycles

## Financial Analysis Implications

- Both Short and Long Investment Horizon
- Anticipate Potential Fluctuations and Manage Risk
- Optimize Entry and Exit Points

## CMG Fourier Transform Magnitude Spectrum



Period	Frequency	Trend/Cycle
565.60 Days	0.0018	Economic Cycle and Company Performance
113.60 Days	0.0088	Quarterly Reports

# Categories of Variables: Technical Indicators



**Technical Indicators:** Calculations based on price, volume, or open interest that traders use to analyze securities and predict movements

RSI

Measures the Speed and Change of Price Movements

Bollinger Bands

Middle Band (SMA) and Two Outer Bands (Standard Deviation)

MACD

Indicates the Strength and Direction of Trend

EMA

Exponential Moving Averages (12 and 26 Day)

ADX

Measures Trend Strength

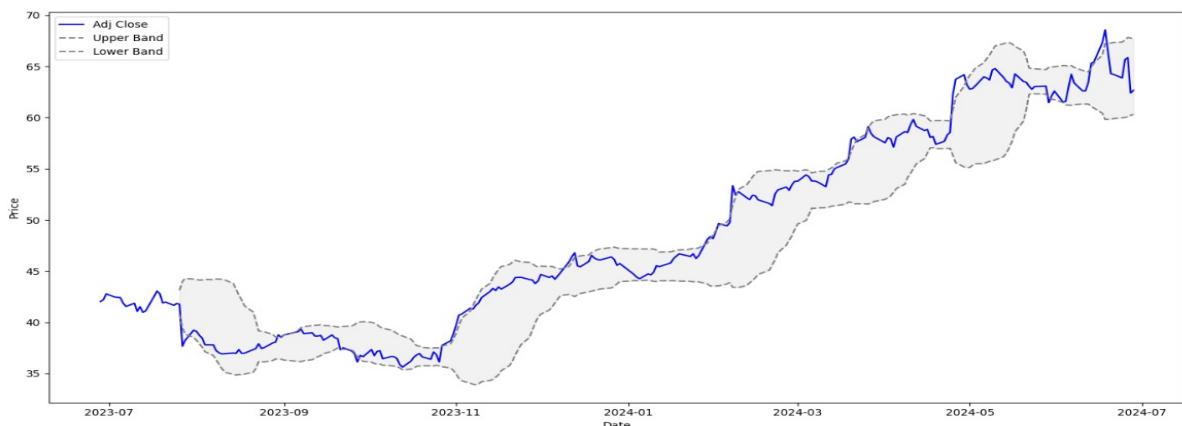
SMA

Simple Moving Averages (50 and 200 Day)

CMG: Relative Strength Index (Index)



CMG Adjusted Closing Price with Bollinger Bands



# Categories of Variables: Interaction Features



**Interaction Features:** Created by Combining Two or More Existing Variables to Capture the Combined Effect

RSI \* Adj Close

Combines Momentum with Actual Price Levels

Can Highlight Oversold or Overbought Conditions

EMA 12 \* Volume

Combines Trend Aspects with Trading Volume

High Values Volume suggest Strong Investor Interest and Confidence

Stoch Osc \* ADX

Combines Momentum Aspects with Trend Strength

Identify Changes in Trend Strength and Momentum

## CMG Performance Relative to Interaction Features



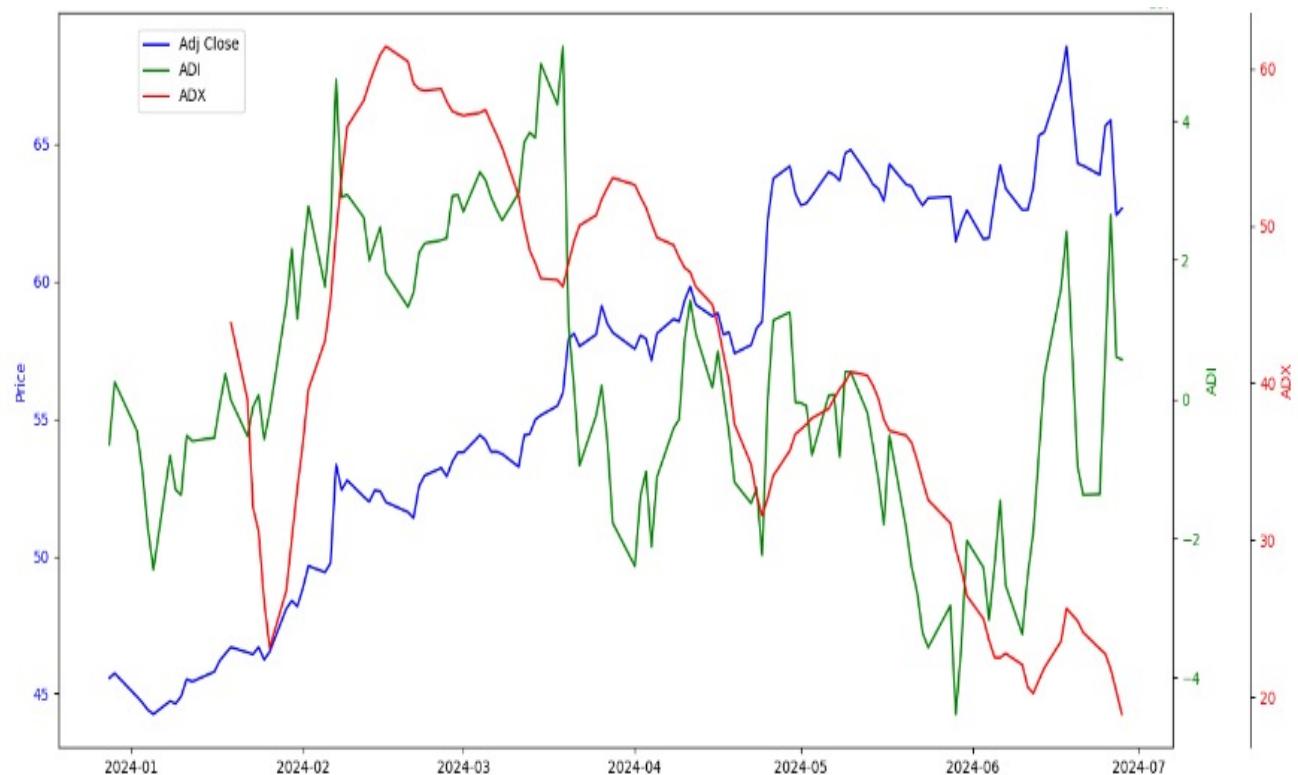
# Categories of Variables: Volume-Based Features



**Volume-Based Features:** Analyzing trading volume data in conjunction with price movements

VPT	Volume Price Trend	Combines price and volume to indicate the strength of price movements
OBV	On-Balance Volume	Measures buying and selling pressures
ADI	Accumulation/Distribution Index	Combines price and volume to show supply and demand
CMF	Chaikin Money Flow	Measures the money flow volume over a specific period
FI	Force Index	Combines price and volume to measure the strength of price movements

## CMG Stock Price, ADI, and ADX Trends



# Modelling

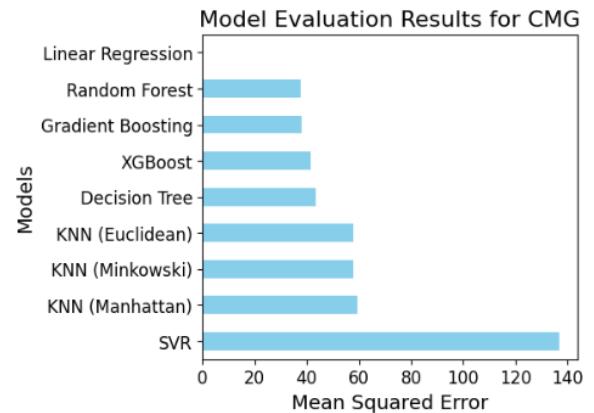
CRISP-DM Model Part 4



# Historical Model Building

## 1. Linear Regression

We used our created variables as 'X' and Adj Close as 'Y' and evaluated the MSE of each model we've covered. Linear Regression was consistently the strongest



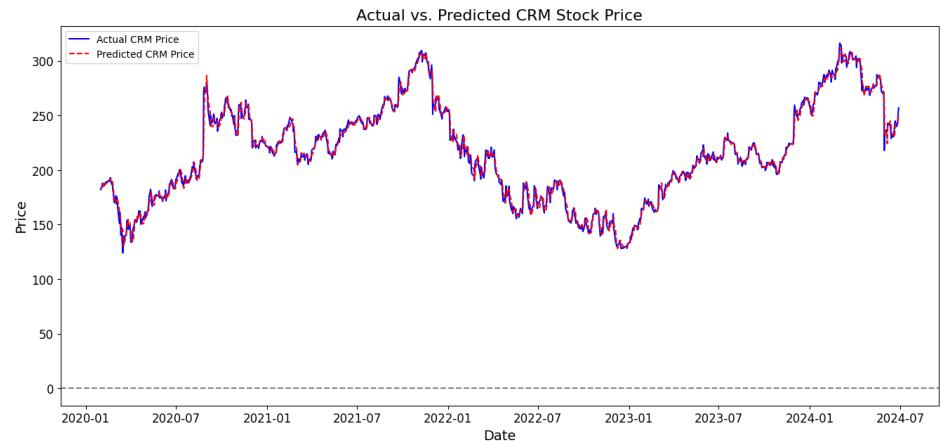
## 2. lag1, MA\_5, MA\_100, EMA\_12, lag5

We tested the R squared of each individual variable on Adj Close to determine the 5 strongest in explaining variance. These variables predominately focus on historical moving averages

Feature evaluation results (sorted by R-squared):	
	R-squared
lag1	0.99
MA_5	0.99
EMA_12	0.97
MA_10	0.97
lag5	0.94
EMA_26	0.93
MA_20	0.92
Bollinger_Low	0.90
lag10	0.88
Bollinger_High	0.86
lag20	0.77
VPT	0.72

## 3. Create the Model

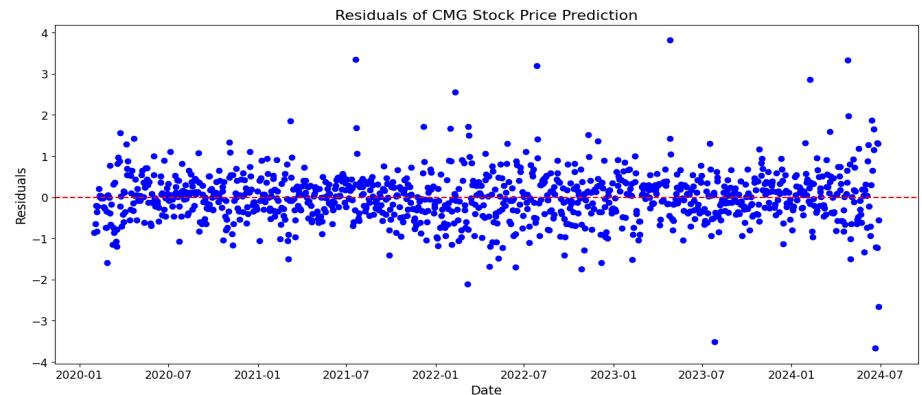
We ran the model for each stock overlaying it with historical prices



# Historical Model Building

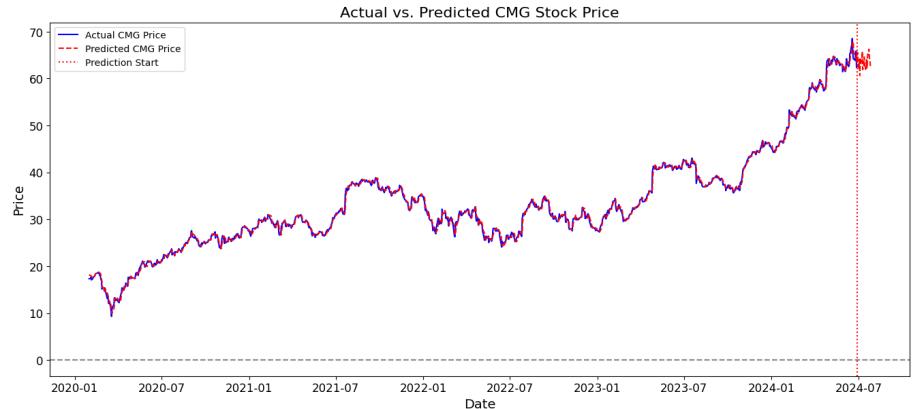
## 4. Evaluate the Model

Each model had an R square value >0.98, a reasonable MSE score and cross val score on R squared >0.96 showing strong predictability, unbiased and strong fitting



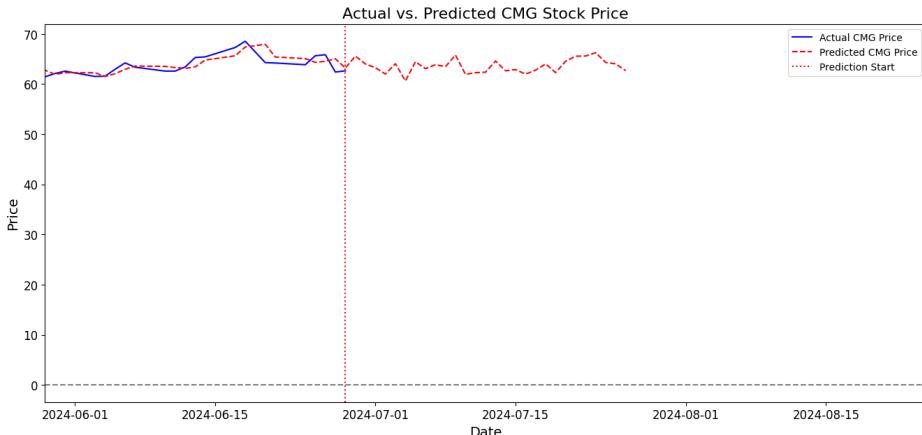
## 5. Predict Future Values

Using a random window approach along with a random shock based on historic data for volatility we predicted from the end of our data to 2024-07-19



## 6. Create the Model

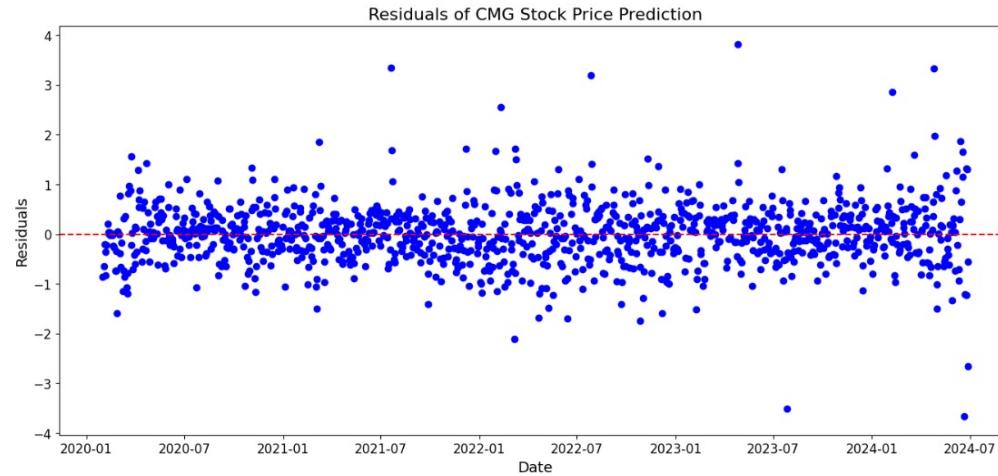
We then predict to 2024-07-26 and obtain a return from the two projected dates



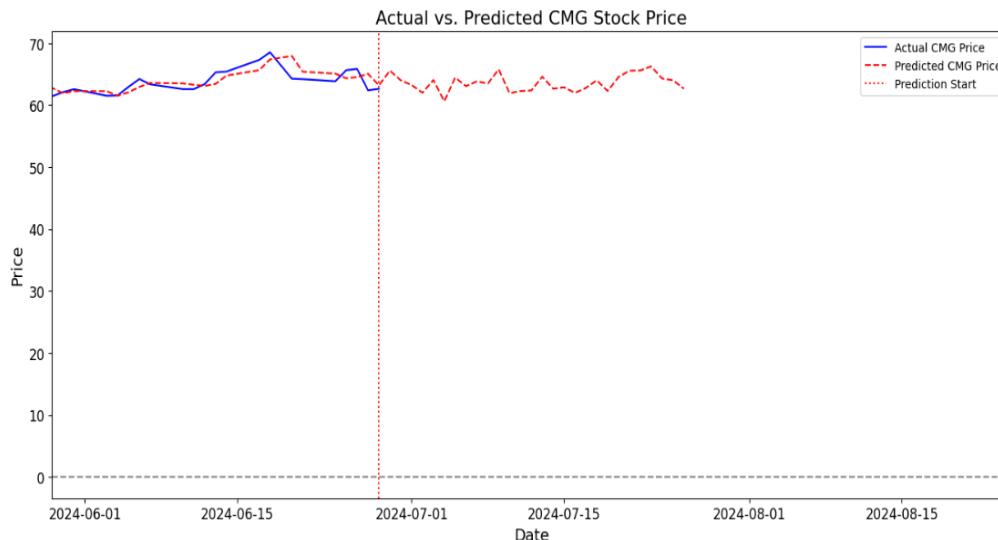
# Historical Model Performance

Company	2024-07-22 Opening Price	2024-07-26 Projected Close	Return (%)
 CHIPOTLE MEXICAN GRILL	\$53.48	\$62.68	+17.20%
 Kraft Heinz	\$33.15	\$31.78	+4.31%
 Uber	\$67.53	\$69.73	+3.26%
 ESTĒE LAUDER	\$98.89	\$103.67	+4.83%
 Expedia®	\$134.83	\$123.22	-8.61%
 salesforce	\$248.75	\$247.71	-0.42%

Residual plot shows the model predictions are unbiased and consistent over time



The model closely followed actual prices and continued the volatility into the predictions



# Forward Predictive Model Building

**Objective:** Forecast stock prices 21 trading days ahead (between June 26 and July 26) with minimal error margins to inform portfolio allocation

**Create Forward Lag Variable:** Captures the gap between the end of historical data and the sale date

- Calls the Adjusted Closing Price of the stock
- Allows for direct comparison and prediction of a forward price based on a given day's data

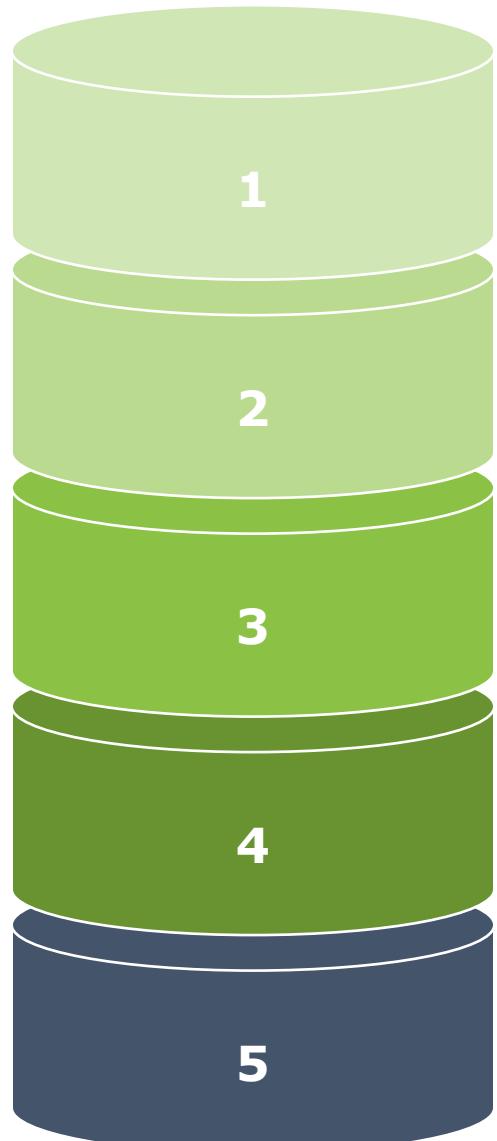
**Variable Selection Process:** Focused on the most influential predictors to enhance accuracy

- Defined function to select best performing variables through a random forest regressor
- Ranked variables based on importance in explaining change to target variable

**Best-Performing Model:** Iterated through various models to identify the top performer.

- Tested all models covered in class with the selected predictive variables
- Selected the model with the lowest MSE score to be used in price predictions

**Price Prediction:** Used the last 5 days of available data in the dataset to provide a detailed forecast for the week of July 22, 2024, to July 26, 2024, to offer insights into expected market trends for the week



# Forward Predictive Model Performance

We forecasted stock prices 21 trading days ahead using a forward lag variable and the most influential predictors. Our model, based on the latest five days of data, provided detailed forecasts for the week of July 22, 2024, to July 26, 2024. These forecasts offer valuable insights into market trends, aiding in portfolio allocation.



Key Variables	'OBV', 'EMA_12', 'Bollinger_Low', 'Bollinger_High'	'EMA_12', 'MACD', 'ADX', 'MACD*EMA_12'	'Bollinger_High', 'OBV', 'EMA_12'	'EMA_12', 'Bollinger_Low', 'OBV'	'VPT'	'EMA_12', 'MACD*EMA_12' , 'MACD'
Model Used	Linear Regression	Linear Regression	Linear Regression	Linear Regression	Linear Regression	Linear Regression
Predicted Price	\$65.90	\$33.02	\$68.92	\$105.77	\$86.10	\$224.09
Expected Return	+23.09%	-0.30%	+2.39%	+6.64%	-36.64%	-9.51%
Model MSE	15.05	9.71	55.03	668.32	603.81	517.44

# Five Day Rolling Window Model Building

Training on 5-Day Rolling Window

## Feature Selection

Stoch\_Osc

Stoch\_Osc\*ADX

ADX

## Model Creation: KNN

Euclidean

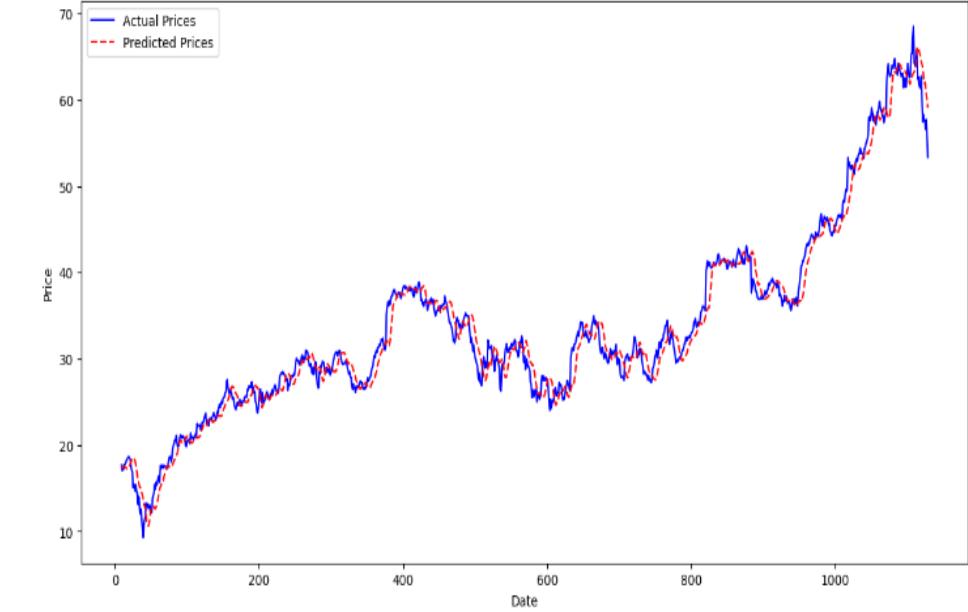
Uniform

N = 30

### Actual vs Predicted CMG Stock Price



### Actual vs Predicted CMG Stock Price (5-Day FWD)



Uses the last 5 days to predict the 6<sup>th</sup> day

MSE: 0.7494

R<sup>2</sup> Score: 0.9939

Uses the last 5 days to predict 5-day FWD

MSE: 1.327

R<sup>2</sup> Score: 0.9759

# Five Day Rolling Window Implementation

## Expanded Feature Selection

RandomForestRegressor

Stoch\_Osc, Stoch\_Osc\*ADX, RSI, MACD, ATR, Fourier, EMA\_12\*Volume, CMF, FI, and Trend\_Strength

## Best Hyperparameters

GridSearchCV

- Learning Rate: 0.05
- Max Depth: 5
- Min Leaf: 4
- Min Split: 10
- N = 100
- Subsample: 0.8

### CMG: Next 20 Trading Days Compared to Previous 252



### CMG Predicted Results

Rolling Window Model	Results
Latest Closing Price (July 19)	\$53.54
Predicted Price (July 22)	\$53.66
Predicted Price (July 29)	\$49.74
Expected Return	(7.09%)
R^2 / MSE	0.4747 / 0.09

# Evaluation + Deployment

CRISP-DM Model Part 5+6



# Evaluation + Deployment

Allocate, Implement and Monitor the best predictive models and communicate findings to all key stakeholders.

## Goal Alignment

- Allocate \$1 million across the six selected stocks
- Maximize short-term profit based on predictive model results

## Select Stocks

- Very high-risk tolerance
- Looking for maximum profits over short investment horizon

## Fund Allocation

- Which stocks meet our goal of positive short-term returns?
- How will the fund be allocated among those stock(s) to maximize profit?

## Projected Value

- Test the different portfolio options
- Pick optimal method to allocate the fund that aligns with our risk tolerance

# Fund Allocation – Combined Models

Company	2024-07-19 Closing Price	2024-07-26 Projected Close	Projected Return (%)
	\$53.48	\$59.45	+11.04%
<b>KraftHeinz</b>	\$33.12	\$32.40	-2.18%
	\$67.53	\$67.11	-0.29%
 ESTĒE LAUDER	\$99.18	\$104.72	+5.59%
	\$135.88	\$117.32	-13.66%
	\$247.63	\$235.89	-4.74%

## Fund Allocation

Company % Allocation Formula:  $\frac{\text{Company Return Projection}}{\text{Total Return Projection}} \times 100\%$



ESTĒE LAUDER

EL Return Projection: +5.59%

Total Return Projection: +16.63%

EL % Allocation: +33.60%

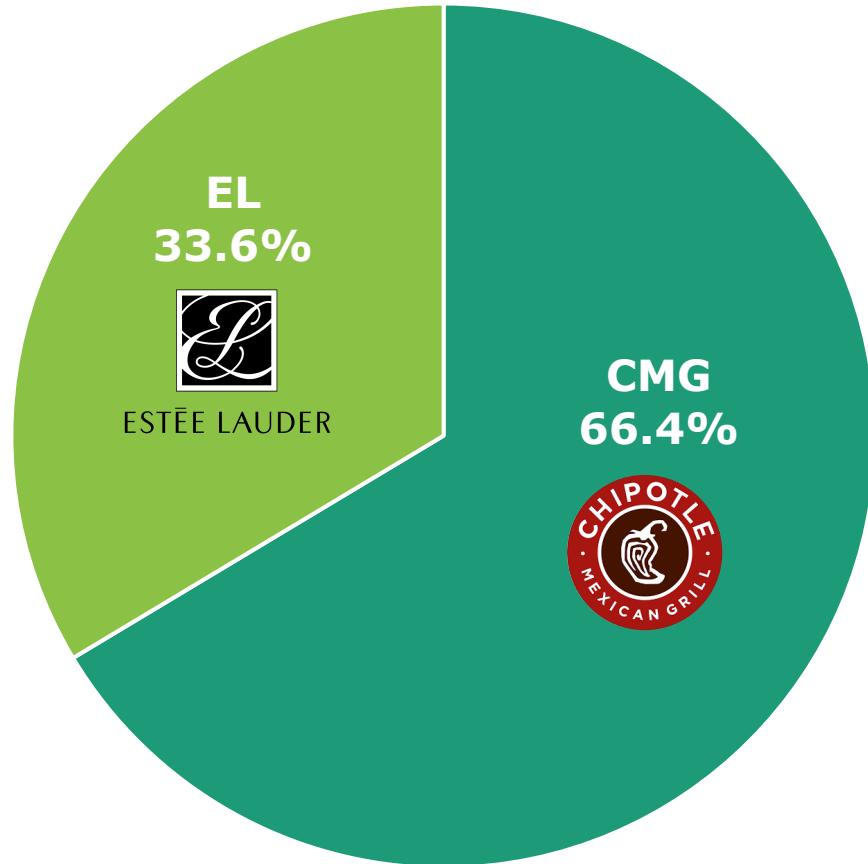


CMG Return Projection: +11.04%

Total Return Projection: +16.63%

CMG % Allocation: +66.40%

# Fund Allocation – Chipotle Mexican Grill and Estée Lauder



EL % Allocation: **33.60%**

Total Fund Amount: **\$1 Million**

EL Dollar Allocation: **\$336,000**

CMG % Allocation: **66.40%**

Total Fund Amount: **\$1 Million**

CMG Dollar Allocation: **\$664,000**

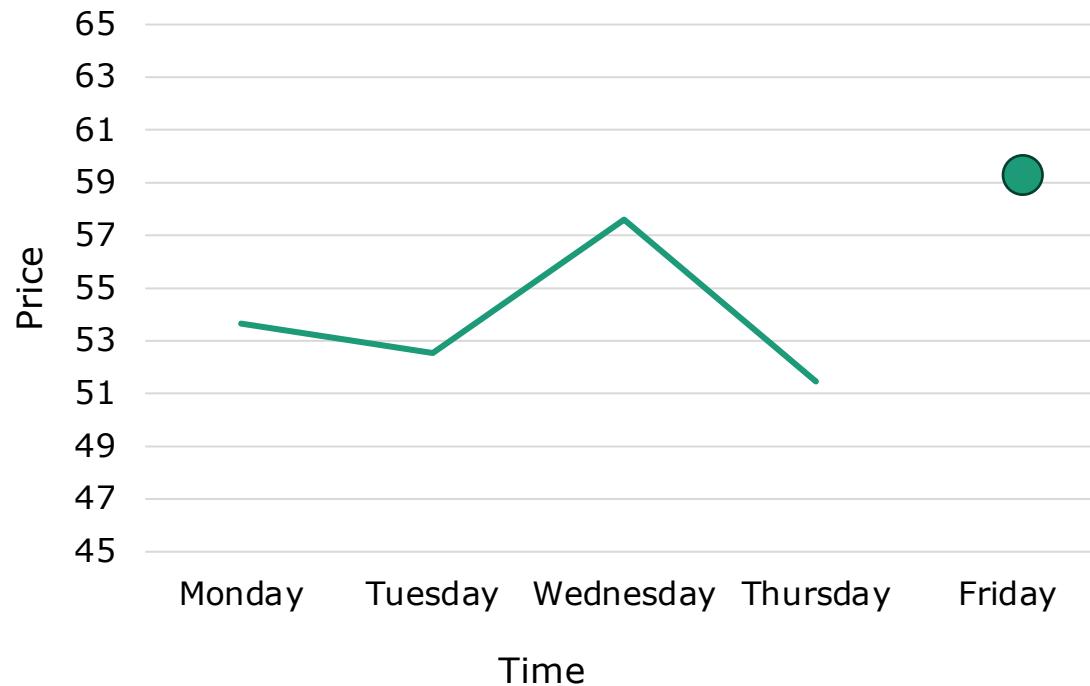
# Allocation Projection

Company	Portfolio %	Allocation \$	Projected Appreciation	Projected Investment Value
<b>Estée Lauder</b>  ESTĒE LAUDER	<b>33.6%</b>	<b>\$336,000</b>	<b>+5.59%</b>	<b>\$354,782</b>
<b>Chipotle Mexican Grill</b> 	<b>66.4%</b>	<b>\$664,000</b>	<b>+11.04%</b>	<b>\$737,306</b>
<b>Total</b>	<b>100%</b>	<b>\$1,000,000</b>	<b>+9.21%</b>	<b>\$1,092,088</b>

# Monitoring

Target Price Marker: 

## Chipotle Mexican Grill

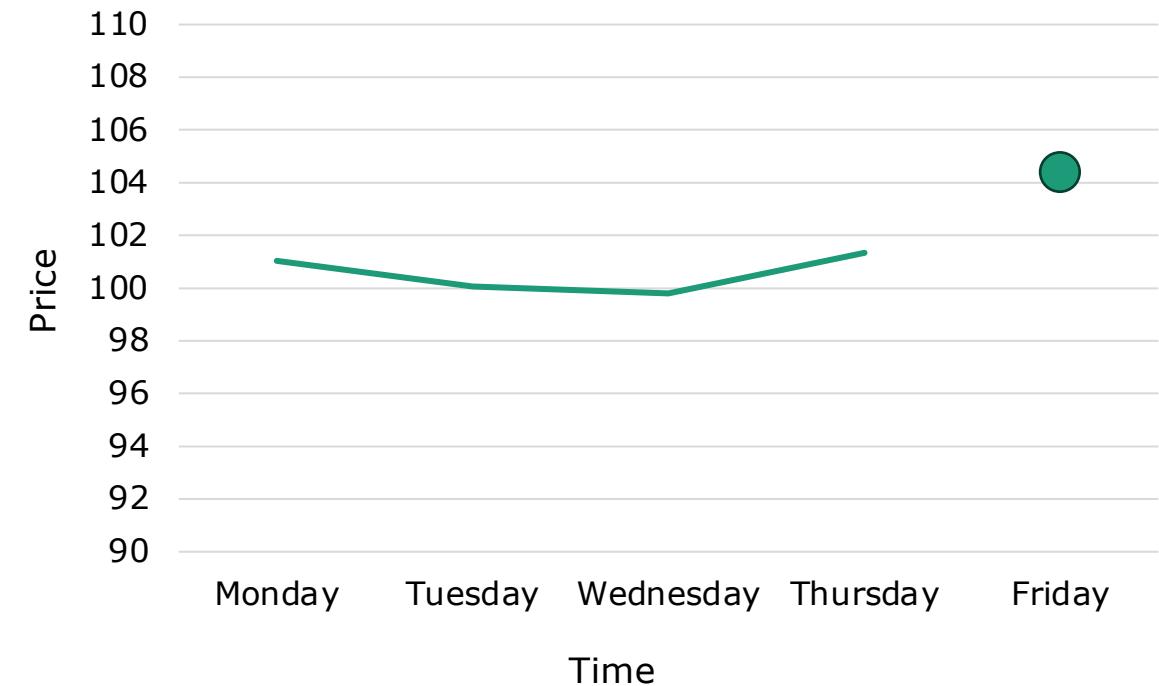


CMG Target Price: **\$59.45**

CMG Return Projection: **+11.04%**

CMG Current Return: **-4.58%**

## Estée Lauder

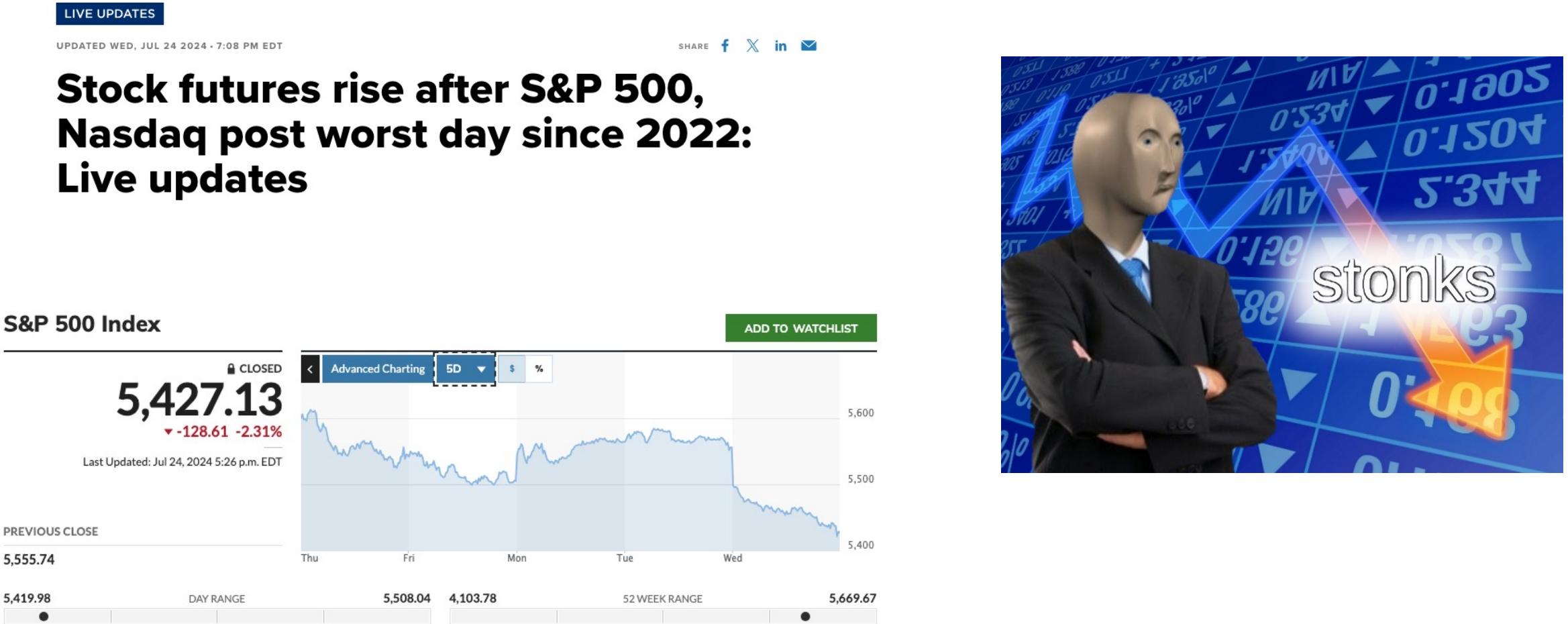


EL Target Price: **\$104.72**

EL Return Projection: **+5.59%**

EL Current Return: **+3.03%**

# What Happened?



# Allocation Reality

Company	Portfolio %	Allocation \$	Actual Appreciation	Actual Investment Value
<b>Estée Lauder</b>  ESTĒE LAUDER	<b>33.6%</b>	<b>\$336,000</b>	<b>+3.03%</b>	<b>\$346,180</b>
<b>Chipotle Mexican Grill</b> 	<b>66.4%</b>	<b>\$664,000</b>	<b>-4.58%</b>	<b>\$633,589</b>
<b>Total</b>	<b>100%</b>	<b>\$1,000,000</b>	<b>-2.02%</b>	<b>\$979,769</b>

**Total Loss = \$20,231**

# Thank you

- Nate's Group

