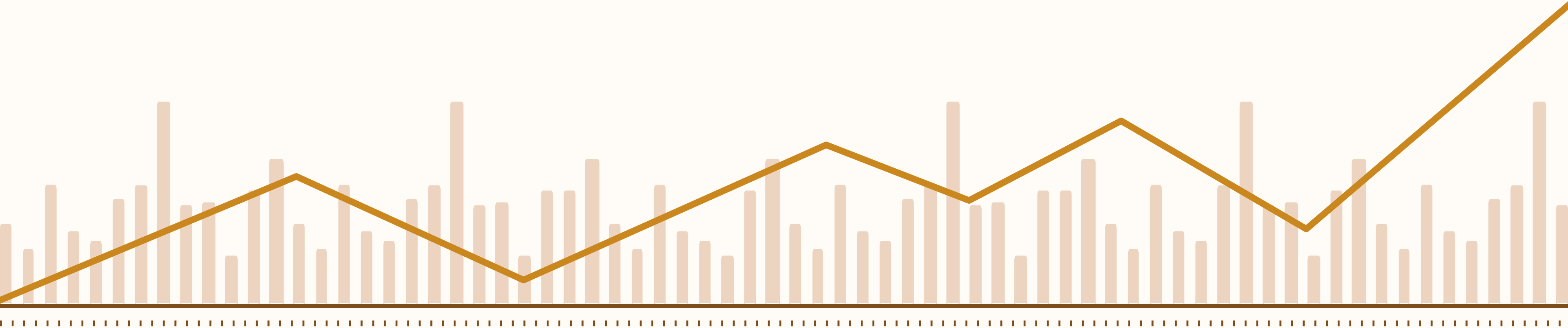


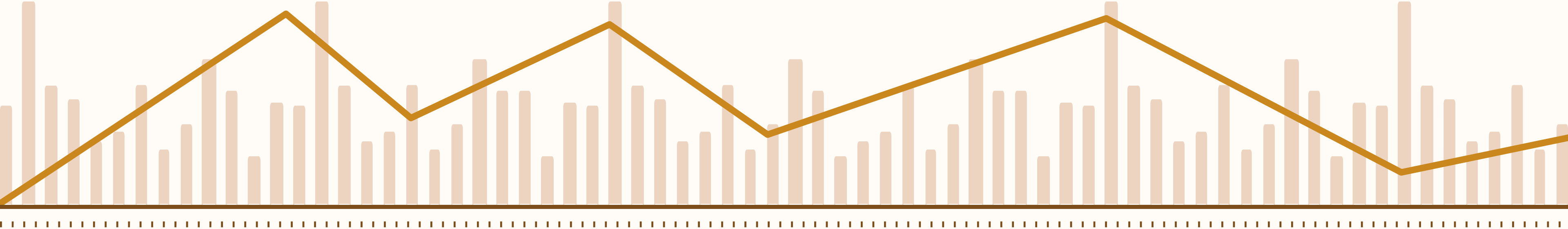
Stock Price Anomaly Detection & Forecasting

Cynthia Du

1

Statistical Deviations

Statistically an anomaly.... but is it actually?



2

Analyzing Stock Volatility

Hidden Markov Model

Used to identify market regimes + trends, good in detecting shifts in market behavior

State 0

Periods of low volatility and stable returns

Represents calm market / minimal price fluctuations

State 1

Periods of high returns or increased volatility

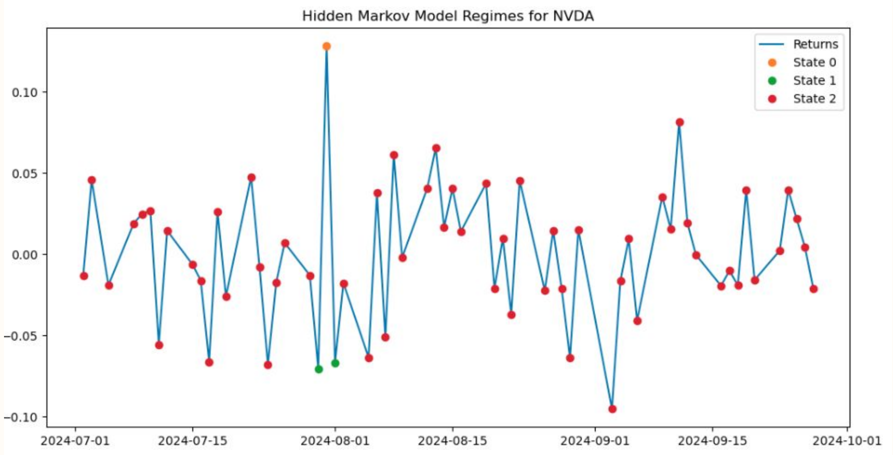
Indicates bullish trends or recovery phases with larger upward price movements.

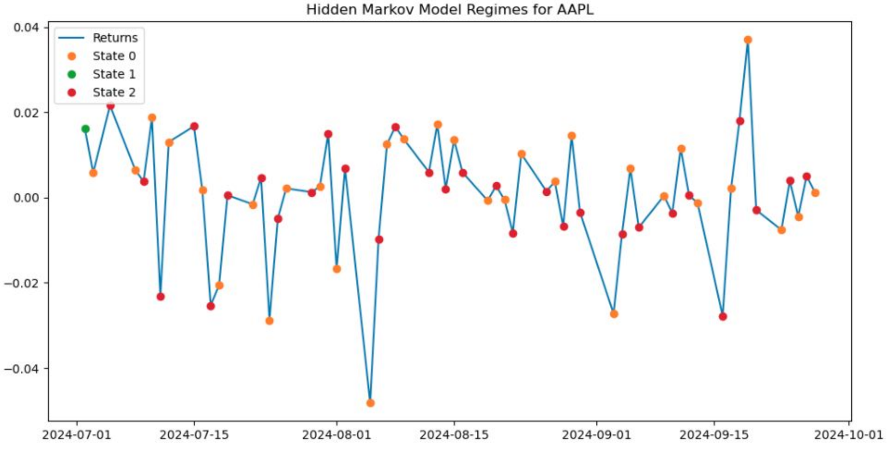
State 2

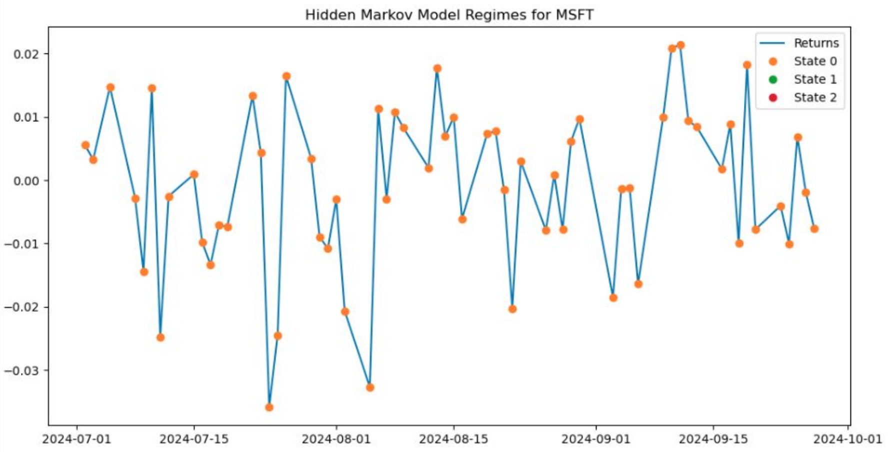
Periods of high volatility or bearish activity / negative returns

Associated with sharp declines, corrections, or market uncertainty.

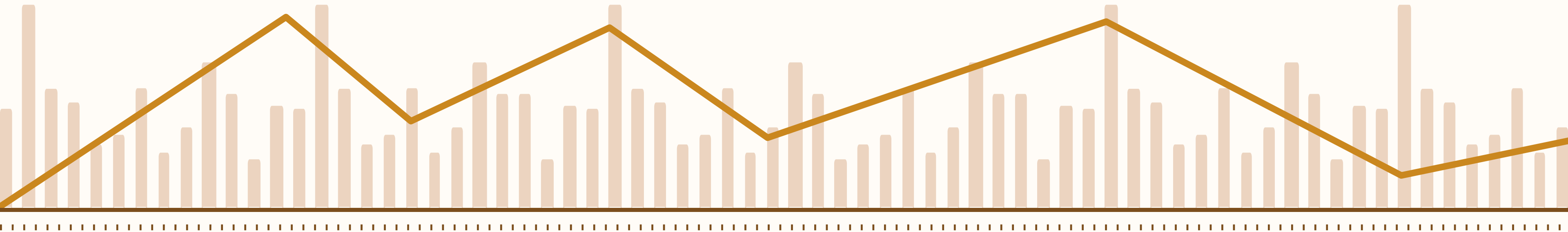
Volatility: degree of variation in a stock's price over certain period. High volatility indicates that the stock's price changes significantly over a shorter period







NVDA: Showcases distinct volatility patterns with most periods in high-return states (State 1) and bearish states (State 2). Risky but potential for high returns :)
AAPL: Reflects relatively stable transitions with brief shifts to bearish states.
MSFT: Demonstrates consistency in low volatility states, with limited bearish activity. Most stable :0



3

STL Decomposition + ARIMA Forecasting

STL Decomposition

(Seasonal and Trend decomposition using Loess)

Decompose the time series into its components (seasonal, trend, and residual) and use the residuals to detect anomalies

Trend: Reflects the underlying movement over time, which is typically smooth.
Seasonal: Reflects short-term fluctuations that repeat over a fixed period (e.g., yearly, monthly, etc.).
Residual: Represents randomness or noise that cannot be explained by trend or seasonality.

ARIMA Forecasting

(Auto-Regressive Integrated Moving Average)

Used for time series forecasting

Captures three key components:
Auto-Regression (AR): Incorporates dependence between an observation and a specified number of lagged observations.
Integration (I): Makes the time series stationary by differencing it (removing trends).
Moving Average (MA): Models the error of the time series as a linear combination of error terms from past observations.

STL Decomposition

(Seasonal and Trend decomposition using

ARIMA Forecasting

(Auto-Regressive Integrated Moving

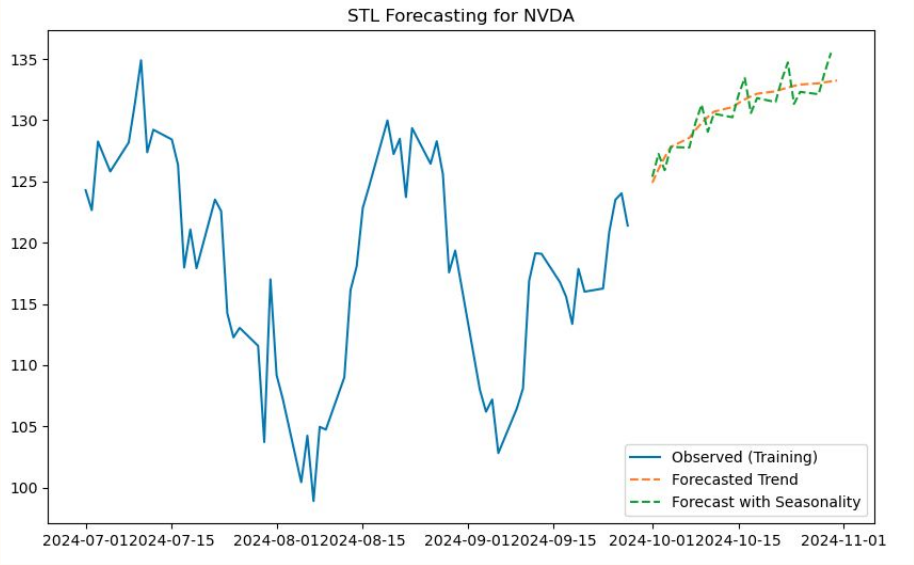
```python
# Forecasting with ARIMA (on the trend component)
trend = result.trend.dropna()
arima_model = ARIMA(trend, order=(1, 1, 1))
arima_fit = arima_model.fit()

# Forecasting future values
forecast_steps = pd.date_range(start=forecast_start_date, end=forecast_end_date, freq='B')
forecast = arima_fit.get_forecast(len(forecast_steps)).predicted_mean
forecast = pd.Series(forecast.values, index=forecast_steps)
```

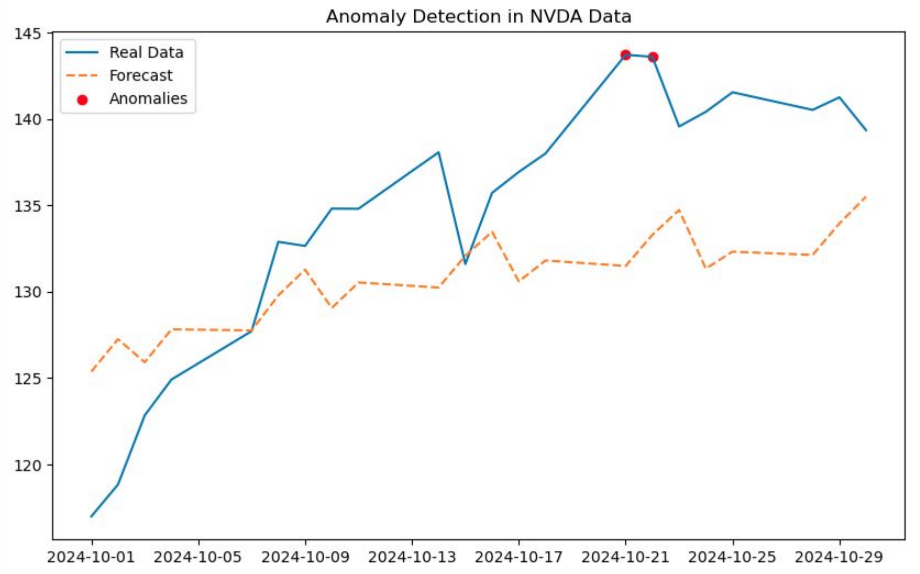
```python
# Combine forecasted trend with the seasonal pattern (from the last cycle)
seasonal_cycle = result.seasonal[-len(forecast):]
forecast_with_seasonality = forecast + seasonal_cycle.values
forecast_with_seasonality = forecast_with_seasonality.loc[aapl_test.index]
```

seasonality.

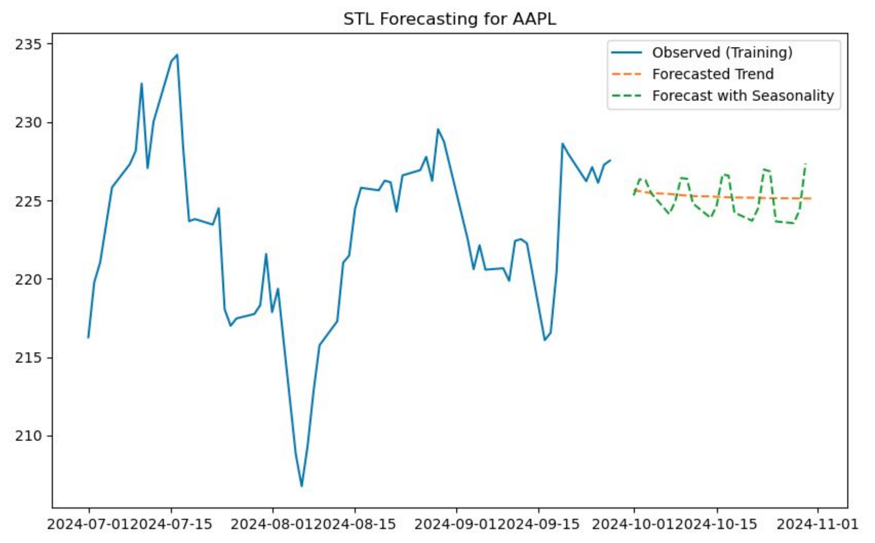
error terms from past observations.

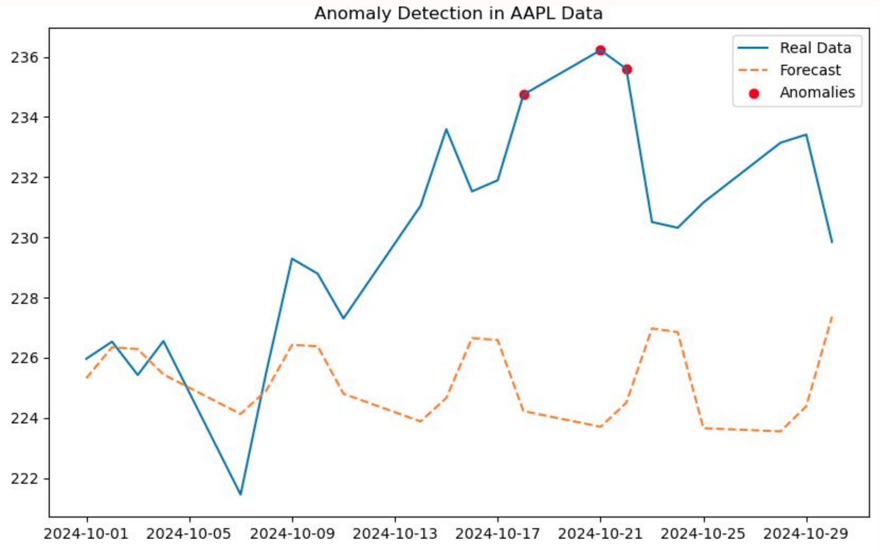
NVDA


STL Forecasting for NVDA
135
130
125
120
115
110
105
100
2024-07-01
2024-07-15
2024-08-01
2024-08-15
2024-09-01
2024-09-15
2024-10-01
2024-10-15
2024-11-01
Observed (Training)
Forecasted Trend
Forecast with Seasonality

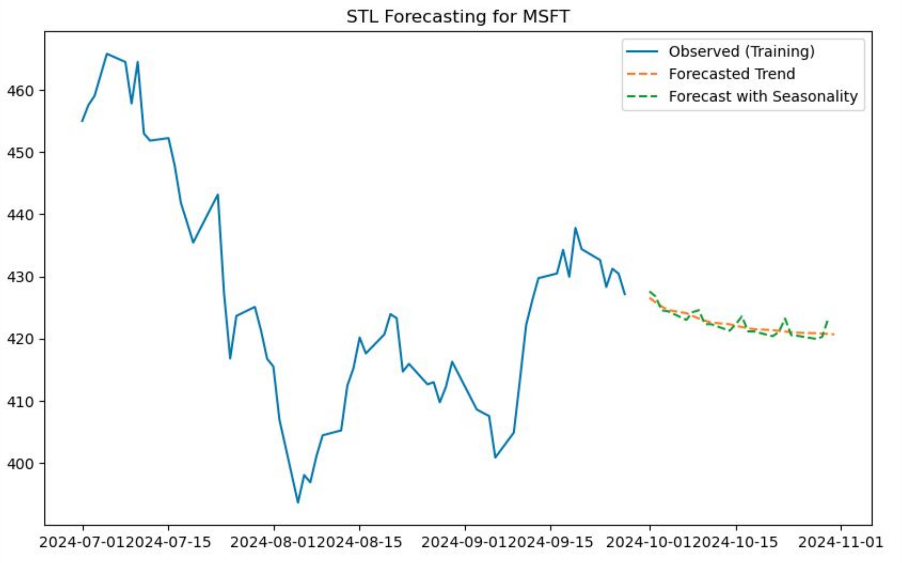

Anomaly Detection in NVDA Data
145
140
135
130
125
120
Real Data
Forecast
Anomalies
2024-10-01
2024-10-05
2024-10-09
2024-10-13
2024-10-17
2024-10-21
2024-10-25
2024-10-29

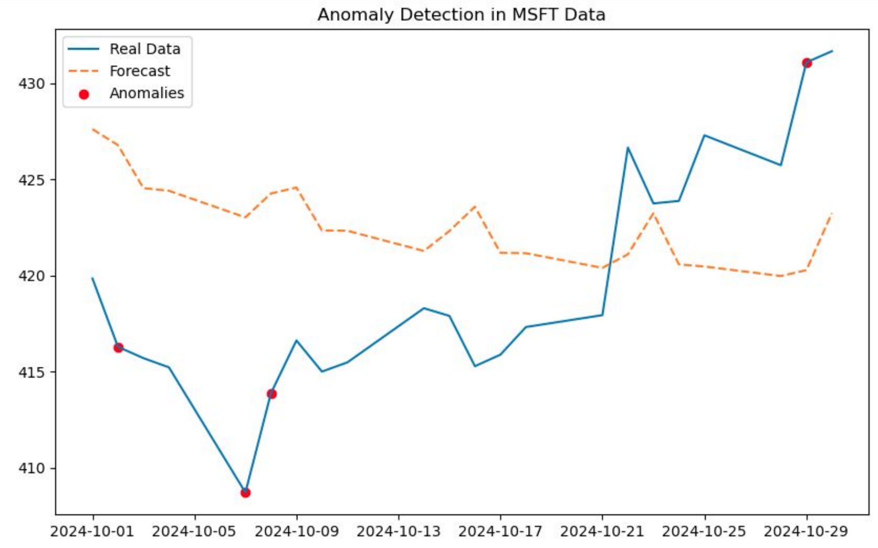
AAPL


STL Forecasting for AAPL
Observed (Training)
Forecasted Trend
Forecast with Seasonality
235
230
225
220
215
210
2024-07-01
2024-07-15
2024-08-01
2024-08-15
2024-09-01
2024-09-15
2024-10-01
2024-10-15
2024-11-01


Anomaly Detection in AAPL Data
Real Data
Forecast
Anomalies
236
234
232
230
228
226
224
222
2024-10-01
2024-10-05
2024-10-09
2024-10-13
2024-10-17
2024-10-21
2024-10-25
2024-10-29

MSFT


STL Forecasting for MSFT
Observed (Training)
Forecasted Trend
Forecast with Seasonality
460
450
440
430
420
410
400
2024-07-01
2024-07-15
2024-08-01
2024-08-15
2024-09-01
2024-09-15
2024-10-01
2024-10-15
2024-11-01


Anomaly Detection in MSFT Data
Real Data
Forecast
Anomalies
430
425
420
415
410
2024-10-01
2024-10-05
2024-10-09
2024-10-13
2024-10-17
2024-10-21
2024-10-25
2024-10-29

Thank you!

Shoutout to the best mentor Justin and Bree for their help !!!