04-11-2024

STOCK PRICE ANALYSIS

A DEEP LEARNING APPROACH TO FORECASTING STOCK PRICES

ANJANA P | SULEKHA P S

PROJECT OVERVIEW

Objective

To analyze and predict stock prices for Amazon, Google, Microsoft, and Apple using machine learning.

Problem Statement

To analyze the stock prices of companies based on historical data.

Approach Summary

Using an LSTM model to capture complex temporal patterns in stock prices, leveraging deep learning to enhance prediction accuracy.

SCOPE OF WORK

01	02	03
Data Exploration	Data Preprocessing	Feature Selection

04	05	06
Model Building and Training	Evaluation and Visualization	Reporting Findings

DATA COLLECTION

DATA SOURCE

Yahoo Finance library for reliable and timely stock data.

FEATURES

Opening price
Closing price
High
Low
Adjusted closing price
Trading volume

SELECTED STOCKS

Amazon
Google
Microsoft
Apple

TIME FRAME

Historical data from the previous year was utilized to train and test the model, capturing important price patterns.

DATA PREPROCESSING

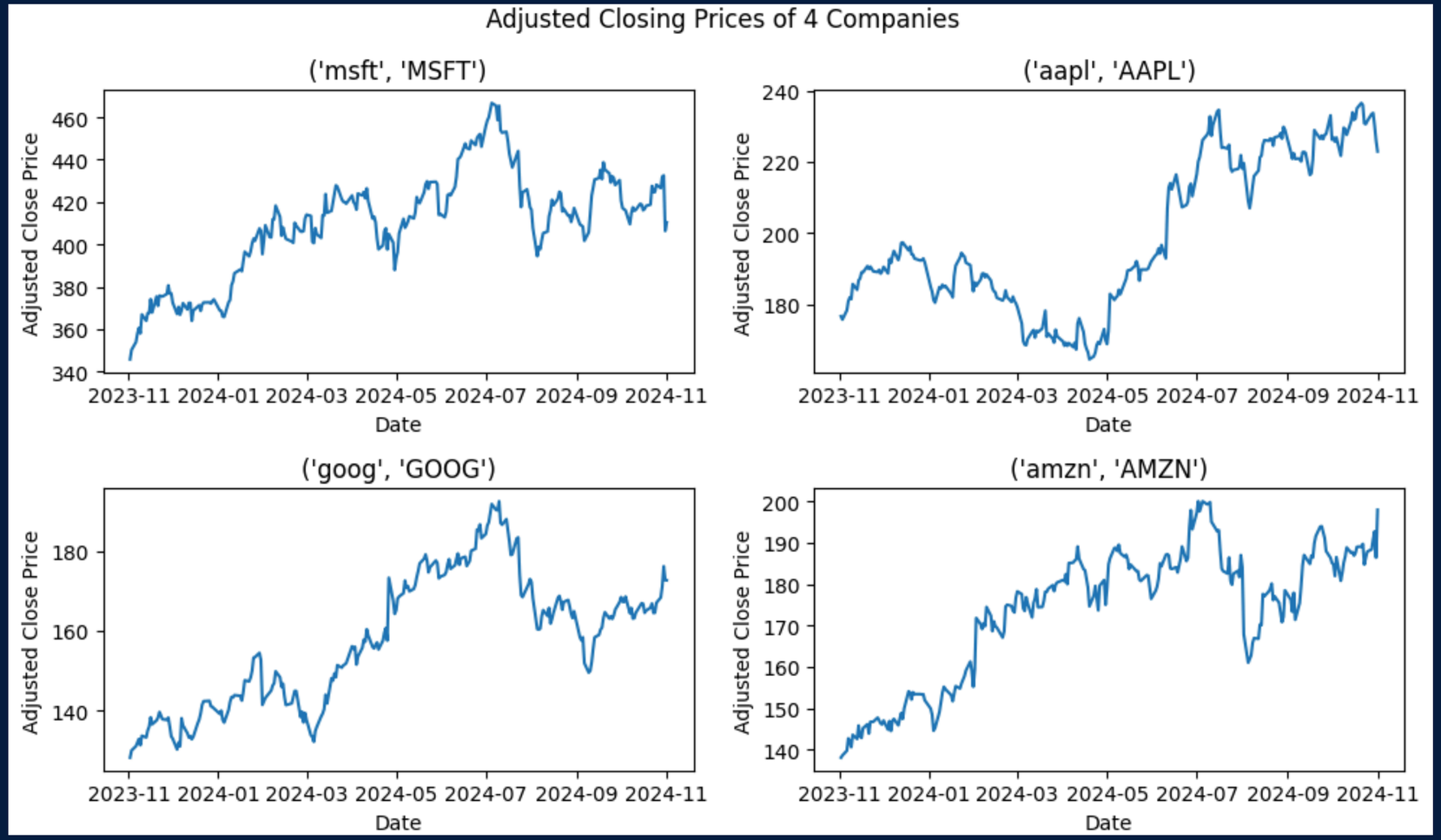
Ensure the data is accurate, consistent, and scaled properly to improve model performance.

01 Handling Missing Values: Forward Fill Method

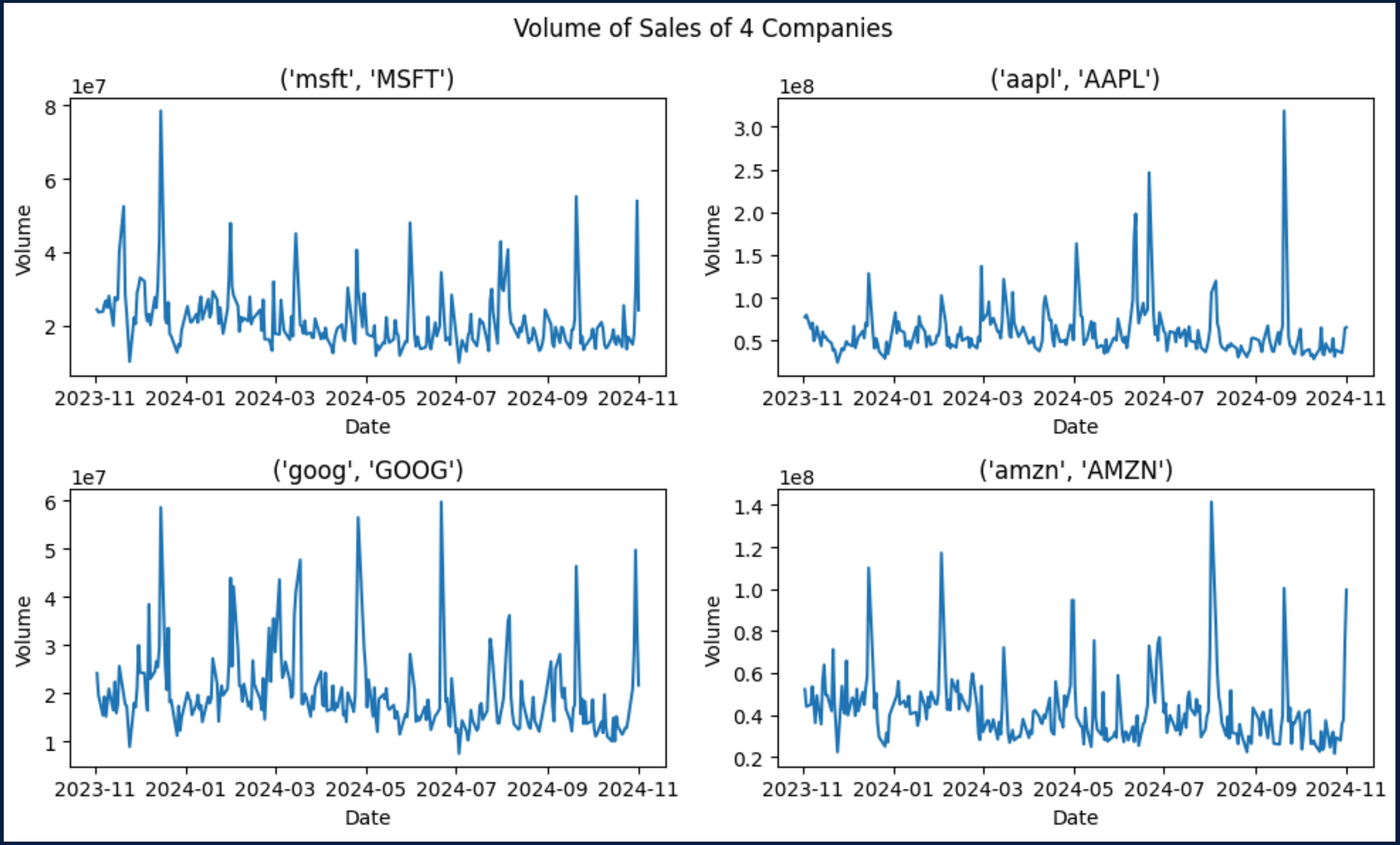
02 Duplicates Removal

03 Normalization & Scaling

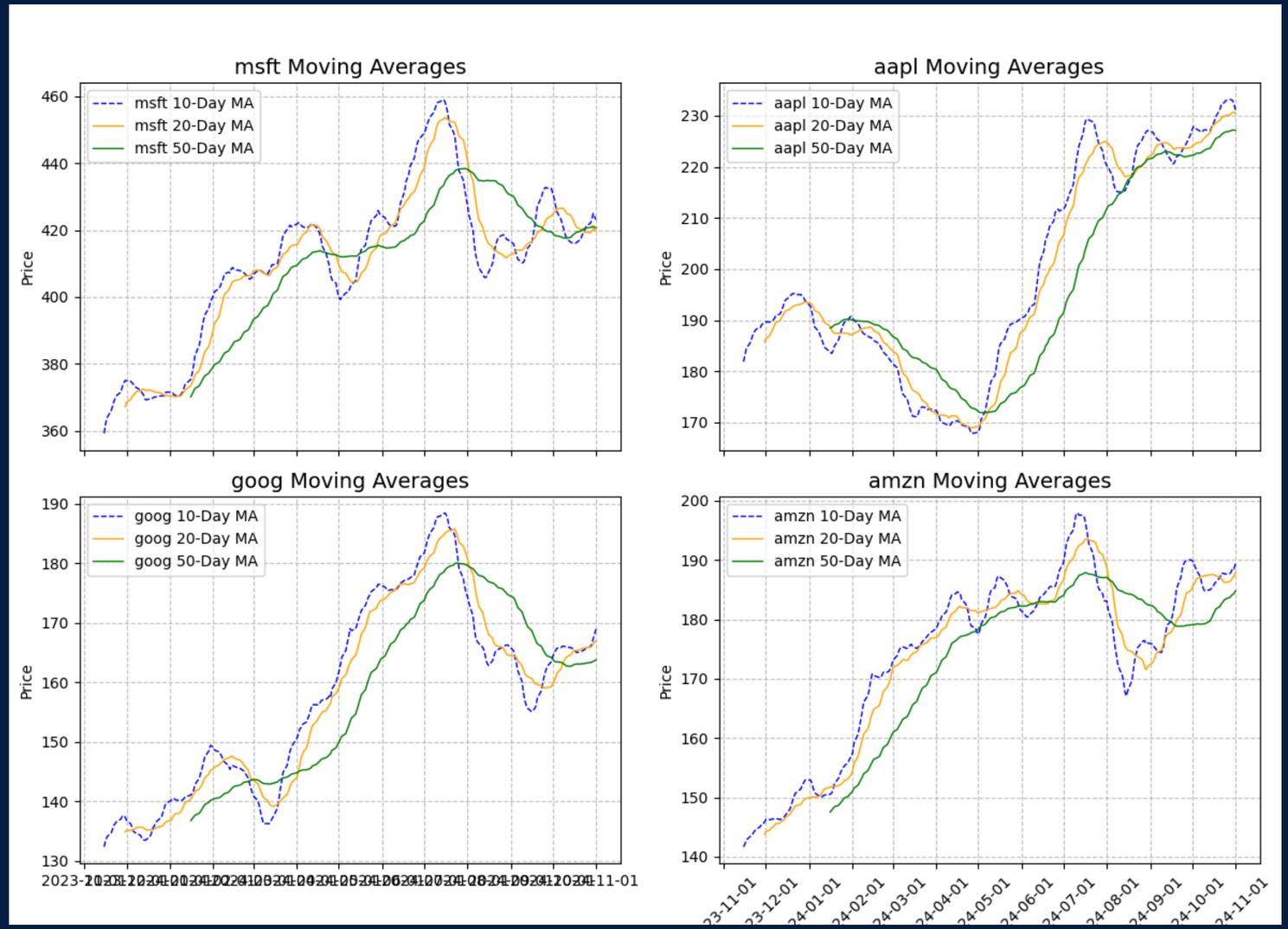
Adjusted Closing Prices



Volume of Sales

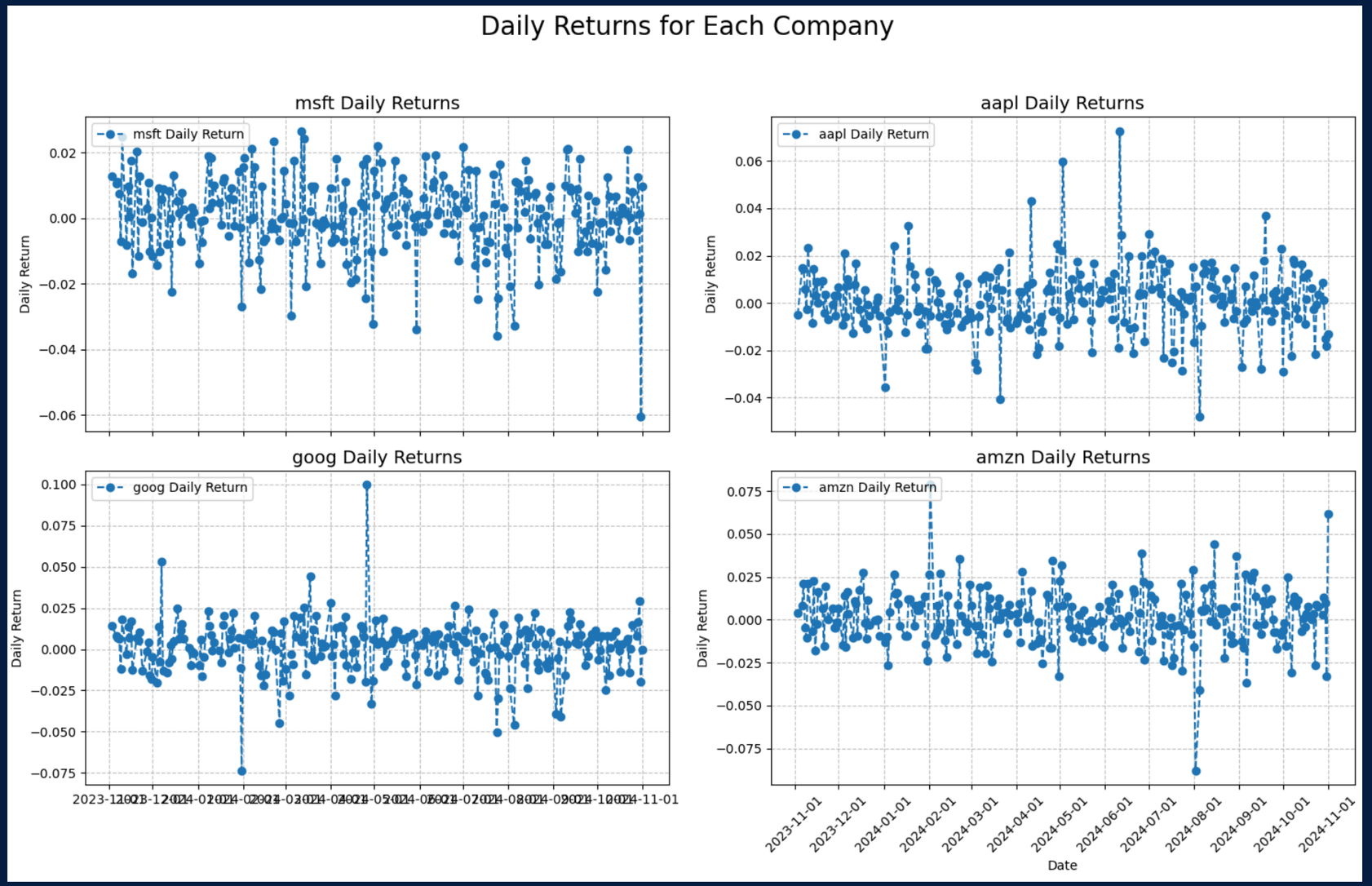


Moving Averages



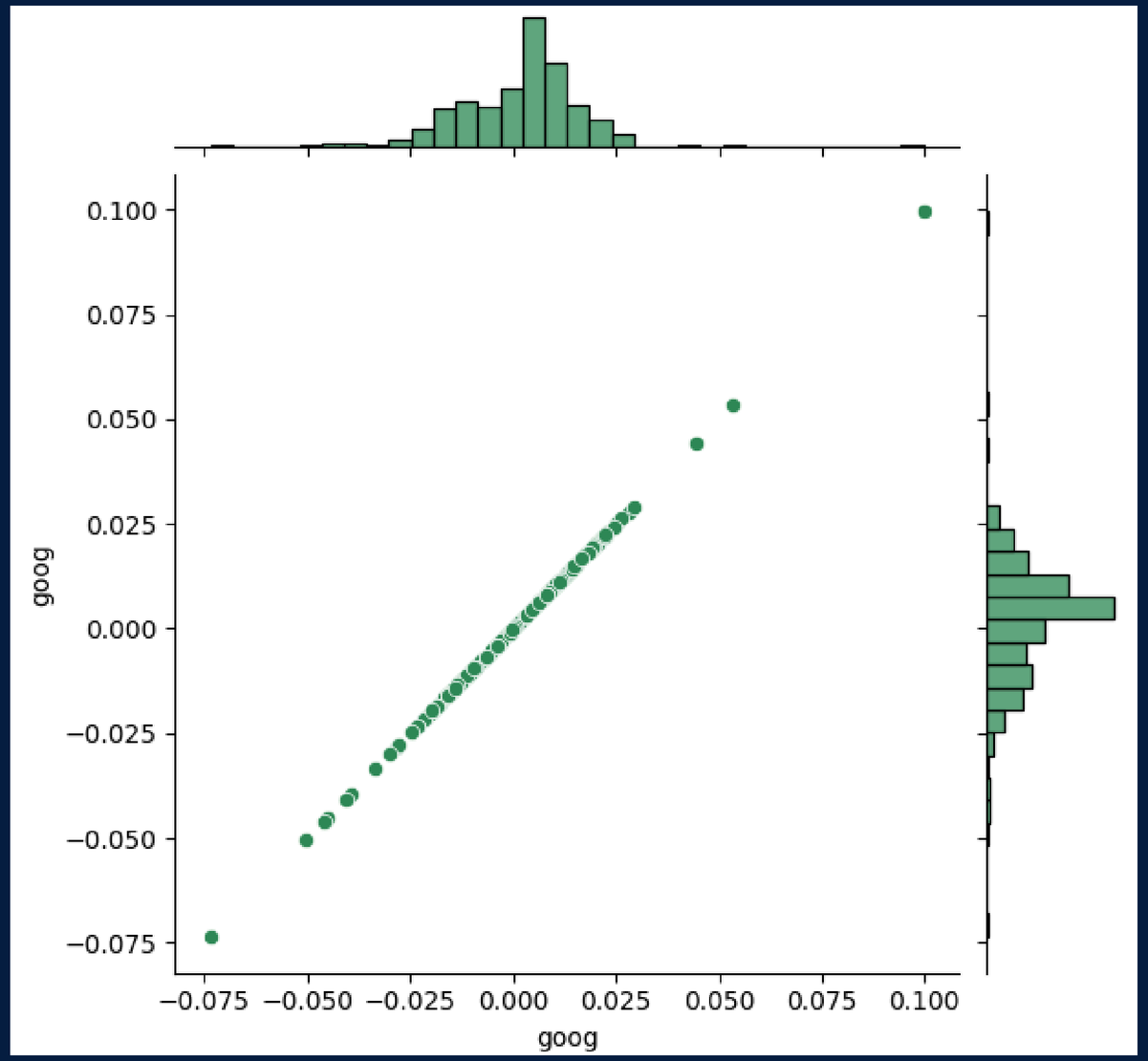
$$SMA = (P1 + P2 + \ldots + Pn) / n$$

Daily Returns

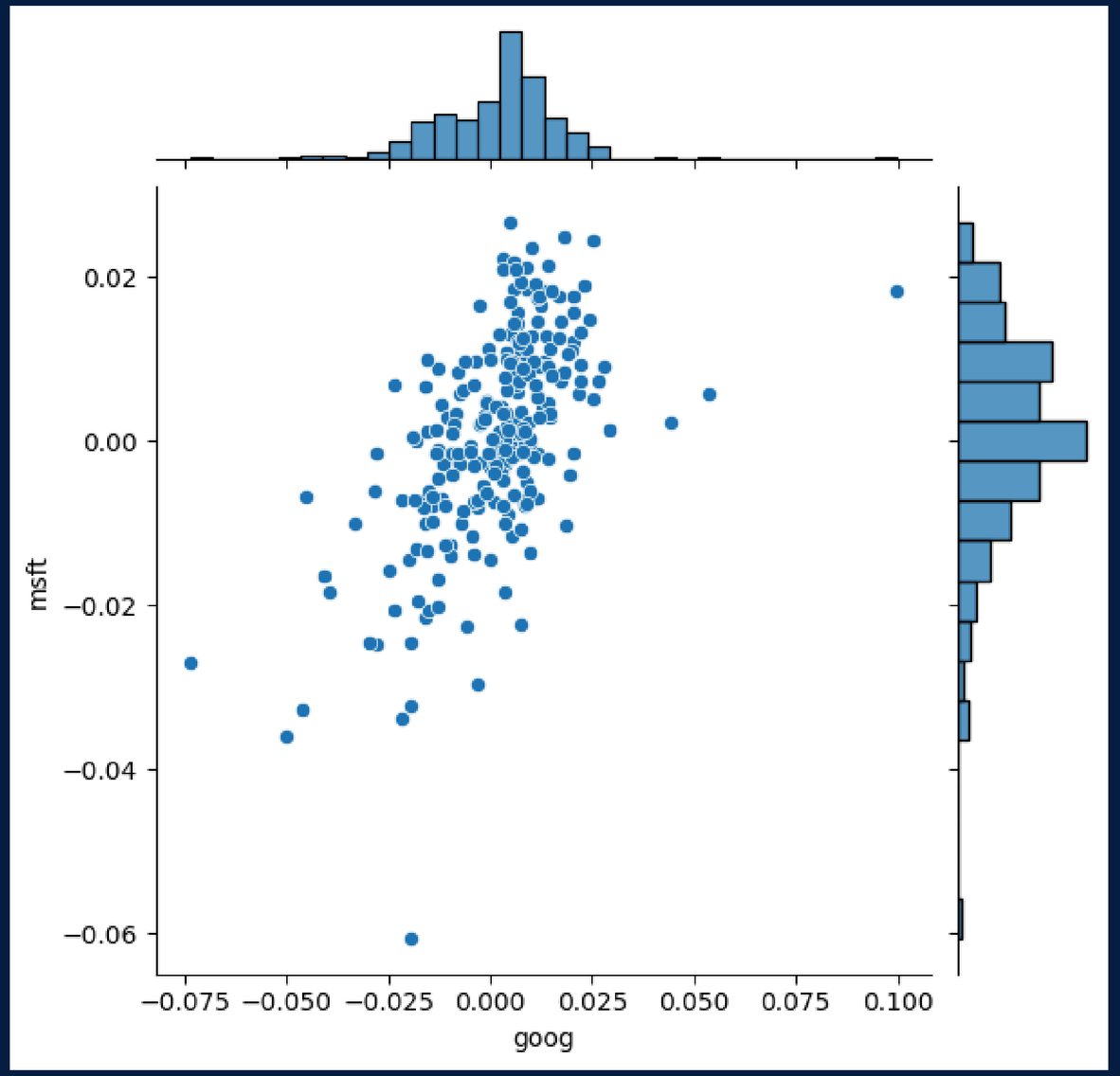


Daily Return(t)
=
Price(t)–Price(t–1)/
Price(t–1)

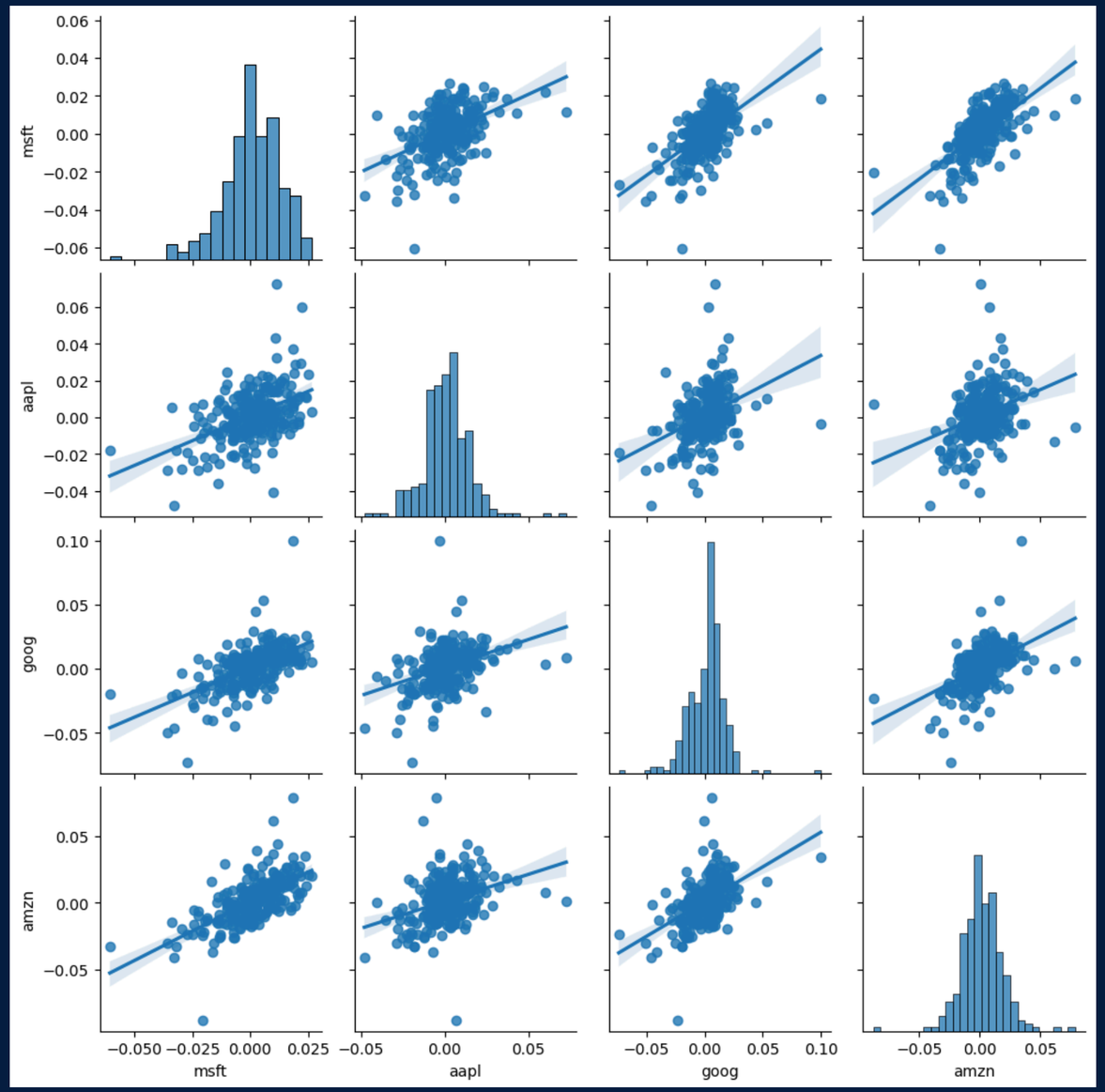
Joint Plot - Google Vs Google



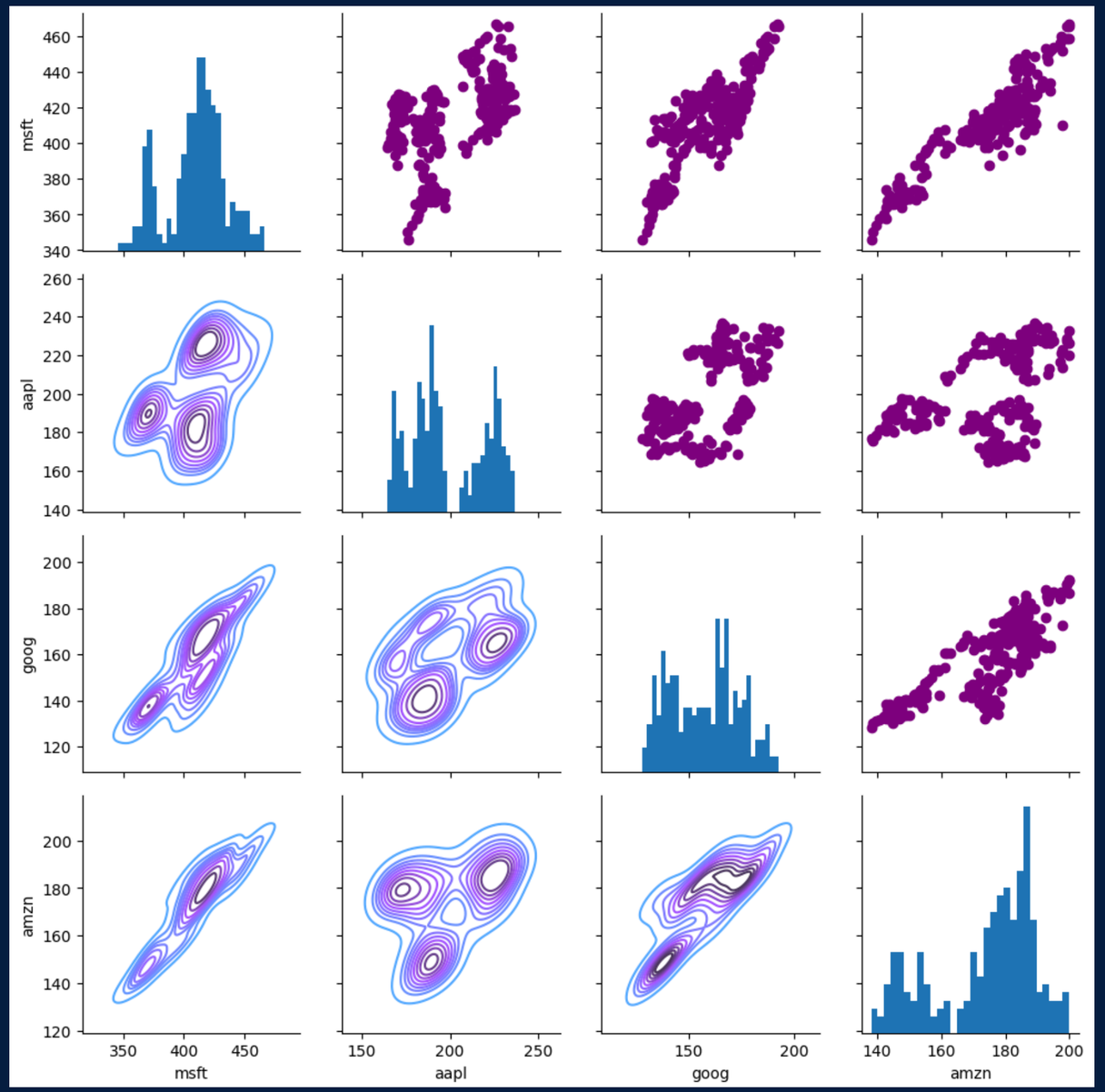
Joint Plot - Google Vs Microsoft



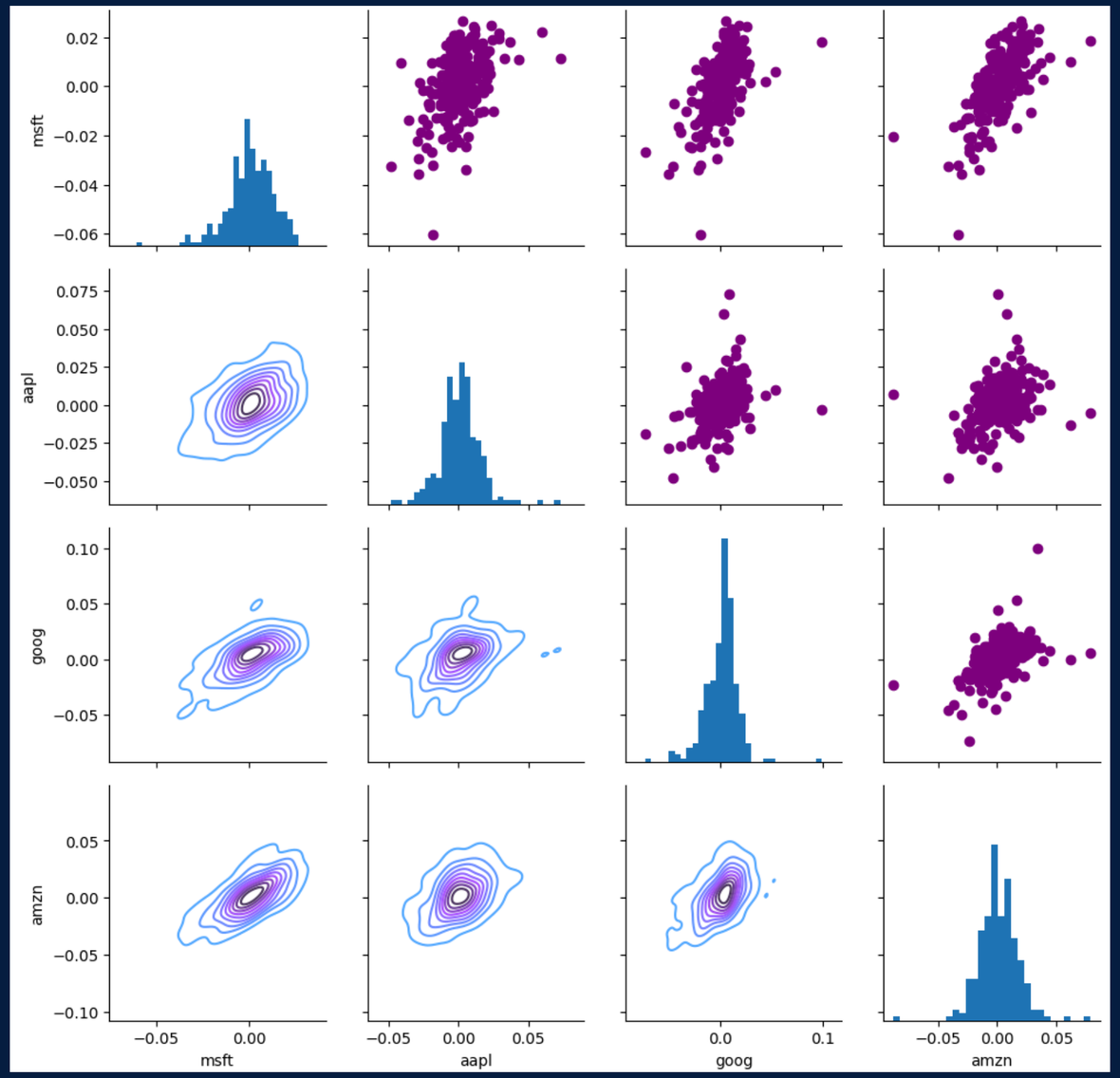
Pair Plot


msft
aapl
goog
amzn
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
0.10
0.05
−0.05
−0.050
−0.025
0.000
0.025
0.05
0.10

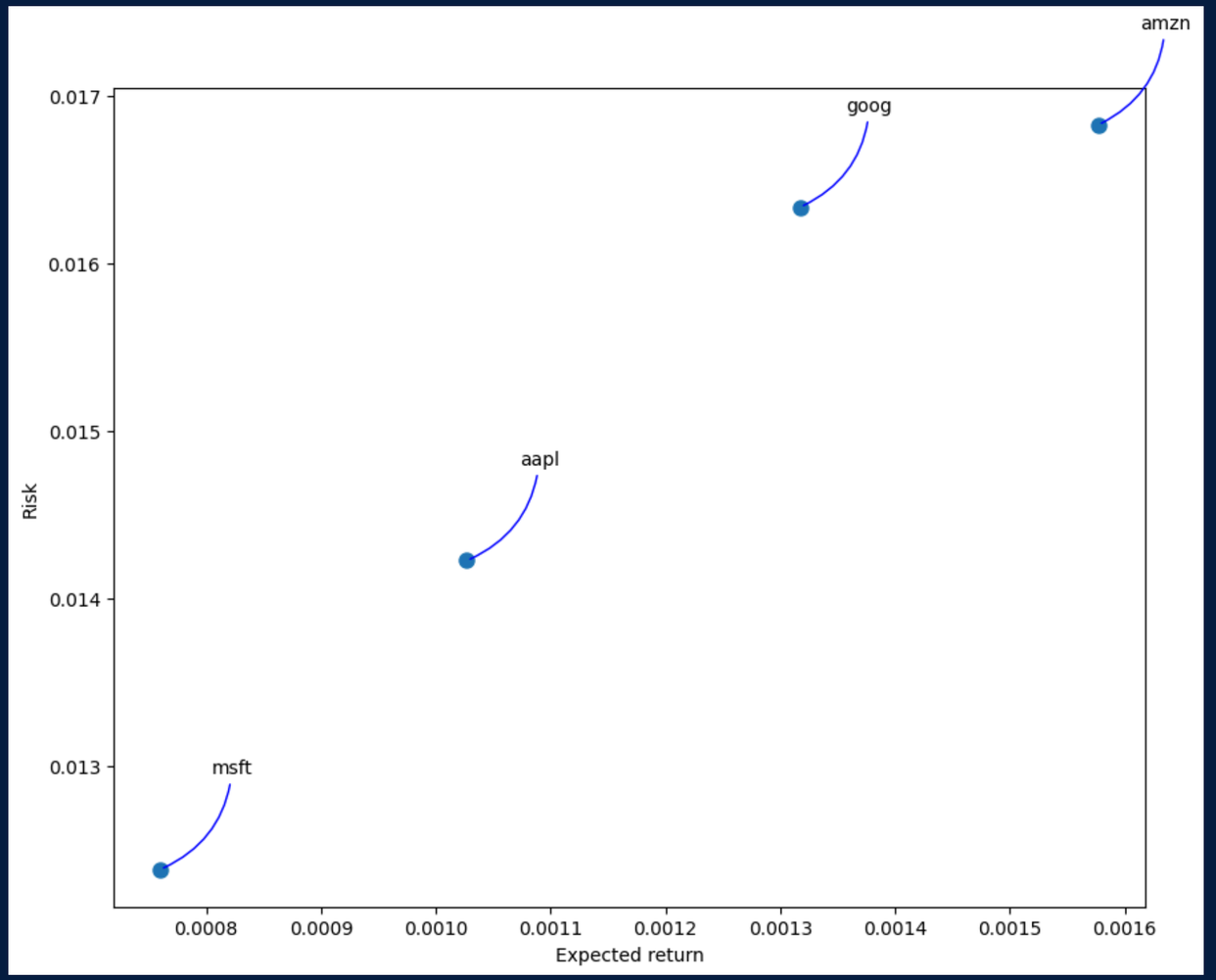
Pair Grid- Closing Prices



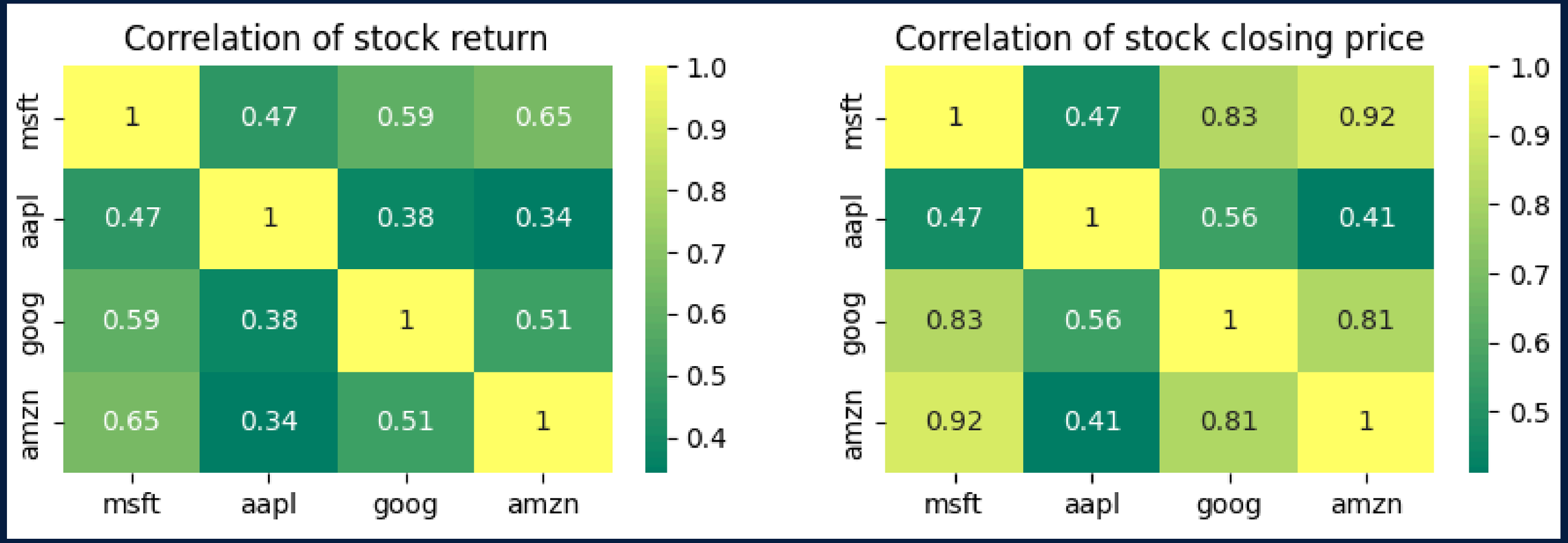
Pair Grid- Daily Returns



Risk Analysis

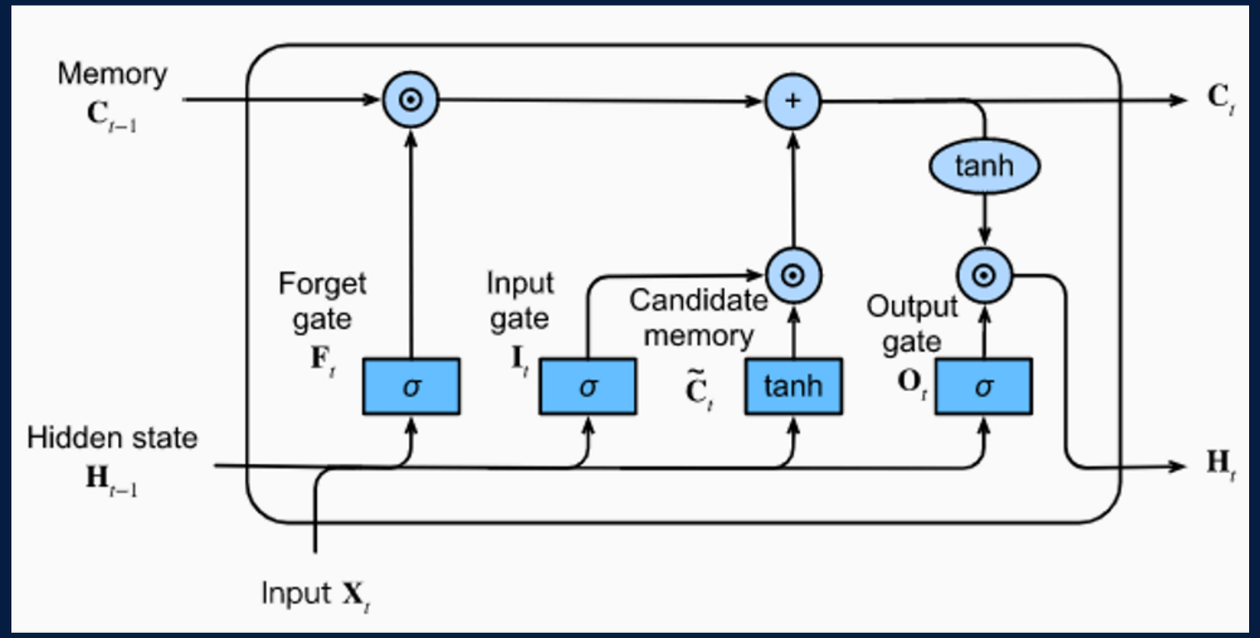


Correlation Heatmaps

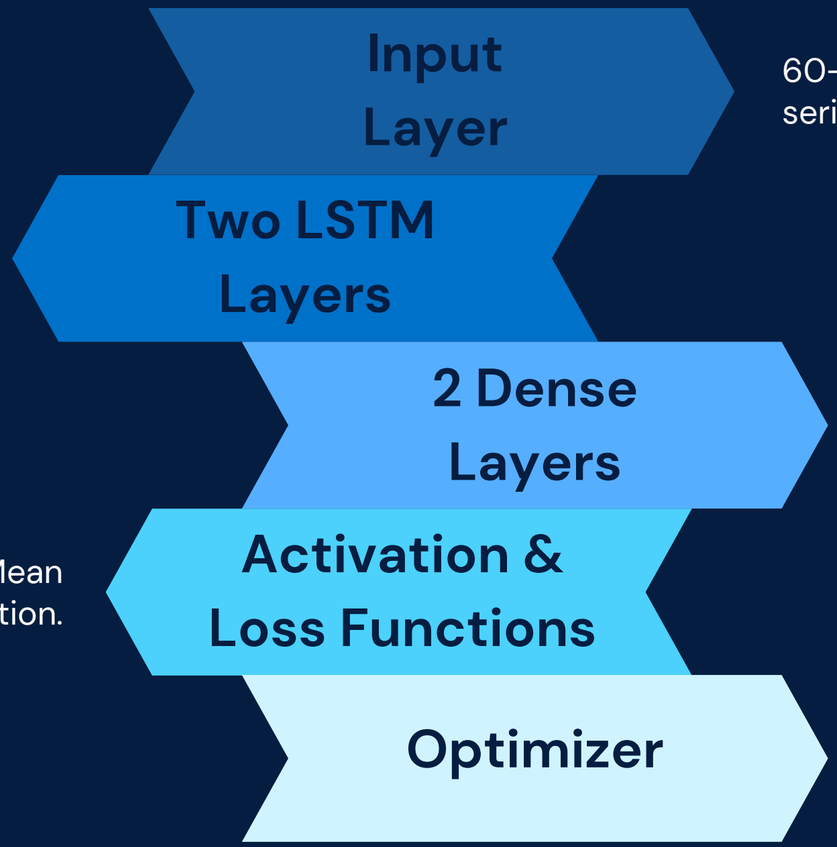


MODEL SELECTION

- Algorithm Chosen: Long Short-Term Memory (LSTM)
- Why LSTM?



MODEL ARCHITECTURE



60-day sliding window used for time-series analysis.

128 units and 64 units respectively, with return sequences for deeper learning.

Final output layer for price prediction.

LSTM layers use ReLU activation; Mean Squared Error (MSE) as the loss function.

Adam optimizer for efficient convergence.

MODEL TRAINING & EVALUATION

95%

Data used for training

5%

Used to evaluate the model's prediction accuracy.

Data Splitting

30

Epochs

32

Batch Size

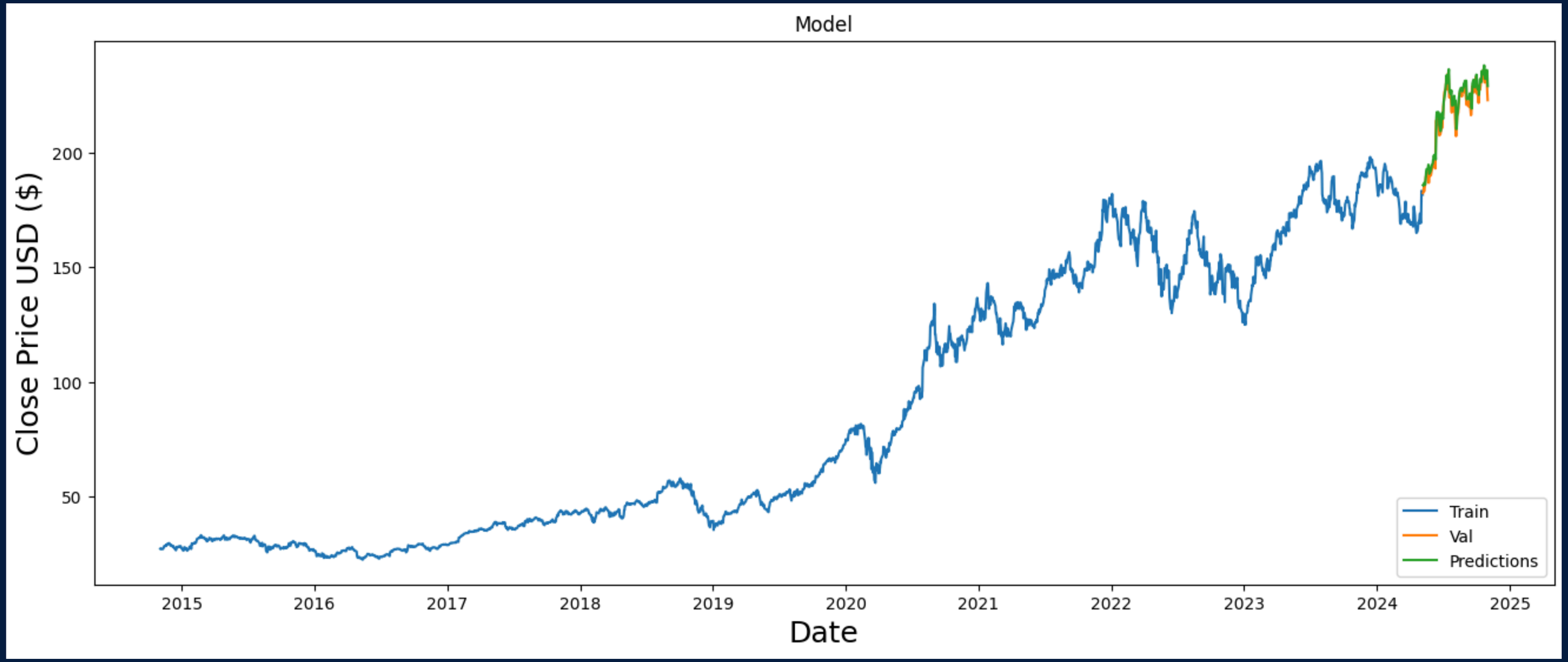
Training Process

Root Mean Squared Error (RMSE)

Evaluation Metric

APPLE - Closing Price

Prediction Graph



THANK YOU