# Stock Market Prediction Project:

## Data Exploration:

Displayed the first and last five rows of the dataset to get an overview

```
data_set.head()
```

```
data_set.tail()
```

## Dataset Overview:

```
data_set.shape
```

```
data_set.info()
```

```
data_set.describe()
```

```
data_set.columns
```

```
data_set.isnull().sum()
```

Analyzed data set  about shape,information, columns,summary statistics, and found there is some null data points in some featured columns.

## Visualizing Key Statistics and Trends

Visualized the relationship between 'open' and 'close' prices, and 'low' and 'high' prices.(scatter plots)

```
sns.relplot(x="open", y="close", data=data_set)
```

```
sns.relplot(x="low", y="high", data=data_set)
```

Plotted histograms for 'open,' 'close,' 'high,' and 'low' prices.(histograms)

```
sns.histplot(data_set["open"], bins=50)
```

```
sns.histplot(data_set["close"], bins=50)
```

```
sns.histplot(data_set["high"], bins=50)
```

```
sns.histplot(data_set["low"], bins=50)
```

## Data Preparation and Model Building

### Features and Target Variable Selection:

Selected features 'open,' 'high,' 'low,' and 'volume' as independent variables.

'Close' price is chosen as the dependent variable.

```
X = data_set[["open", "high", "low", "volume"]]

Y = data_set['close']
```

### Model Selection and Training:

Used Linear Regression for stock price prediction:

```
regressor = LinearRegression()

regressor.fit(X_train, Y_train)
```

Predicted stock prices on the test set using the trained model.

```
test_data_prediction = regressor.predict(X_test)
```

## Model Evaluation Metrics:

```
error_score = metrics.r2_score(Y_test, test_data_prediction)

print("R squared error : ", error_score)
```

Calculated the R-squared error as the evaluation metric

Got R-squared as `R squared error :  0.9998677583231753`

The Linear Regression model was employed for predicting stock prices based on the selected features.

R-squared error was used to evaluate the model's performance.

The actual and predicted values were visualized in a plot for comparison.