ANALYSIS OF FPT CORPORATION BUSINESS PERFORMANCE AND STOCK FORECAST

Lecturer: Dr. Tran Van Hai Trieu

Nguyen Minh Ha - 21520216
Thieu Vinh Tien - 21521533
Tran Anh Thy - 21521517
Truong Phuong Thy - 21520477
Nguyen Le Won Bin - 21521870
Vu Dinh Tuan Kiet - 21521039

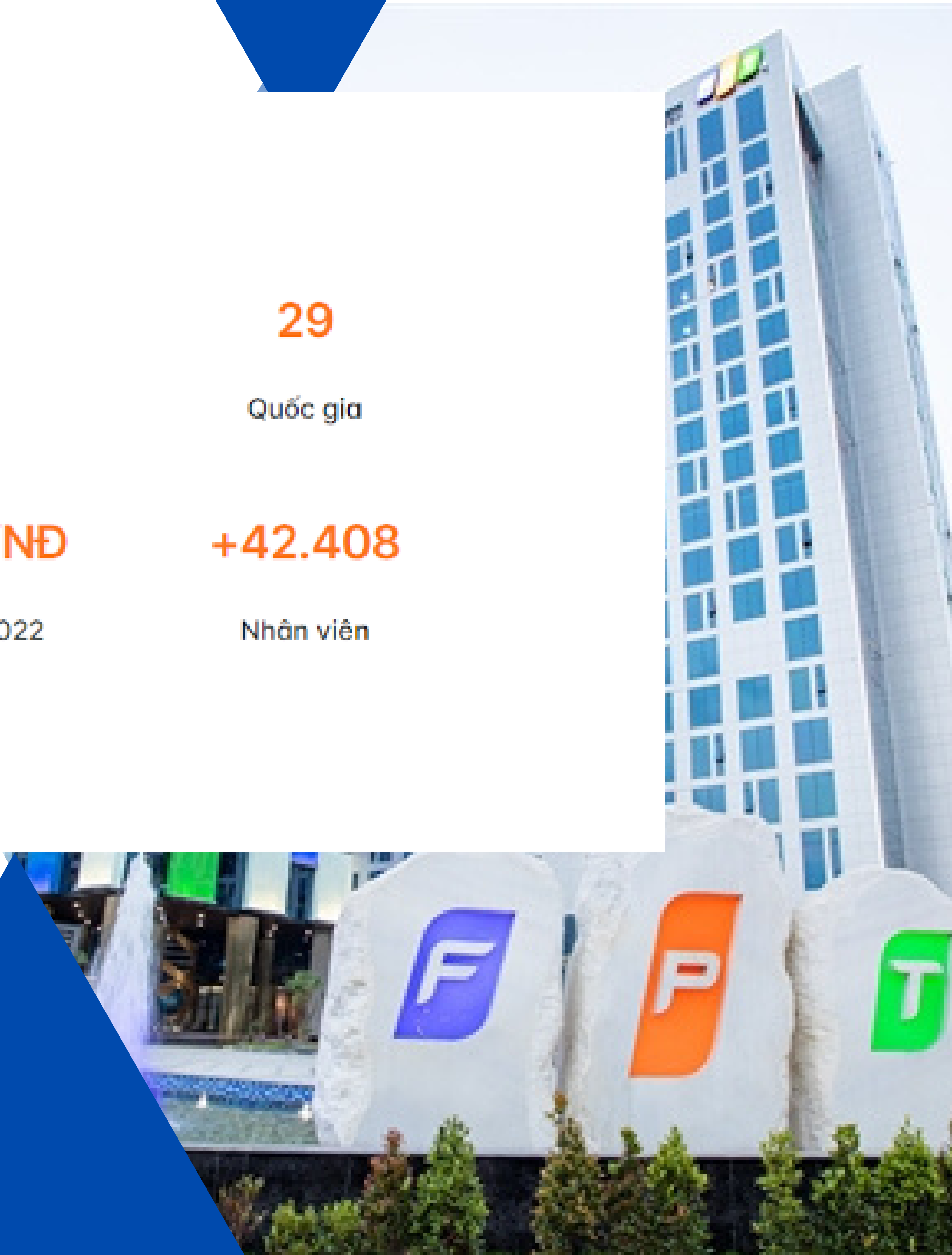
ABOUT COMPANY

Tập đoàn công nghệ hàng đầu Việt Nam

Tiên phong trong xu hướng công nghệ mới

Khẳng định vị thế Việt Nam trên toàn cầu

35 năm	29
Kinh nghiệm	Quốc gia
44.010 tỷ VNĐ	**+42.408**
Doanh thu năm 2022	Nhân viên



ABOUT COMPANY

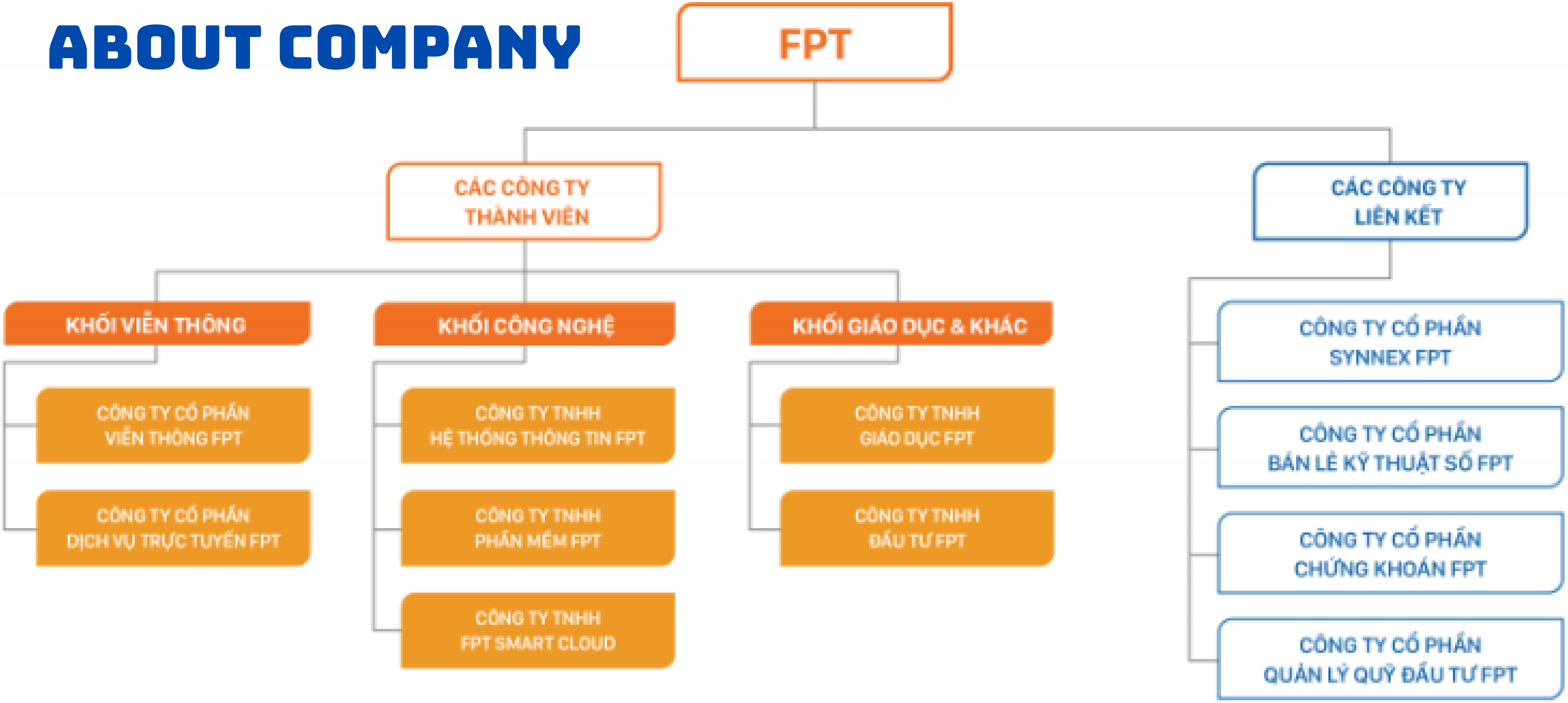
Thành lập	Phủ khắp	Hiện diện	Gần	Doanh thu 2019
32	63	46	29.000	27.717
NĂM	TỈNH THÀNH	CHI NHÁNH, VĂN PHÒNG NGOÀI VIỆT NAM	NHÂN VIÊN	TỶ VNĐ

- Strongly promote education and modernize the country's key economic sectors.
- Efficiently provide services/solutions to customers globally.

ABOUT COMPANY

- Provide solutions and services to customers in Vietnam and worldwide, helping them transform into digital enterprises.
- Engage in the construction and provision of smart technology services for the Government, Healthcare, Transportation, Education, Energy, Telecommunications, and Manufacturing sectors.
- Meet customer needs with just a touch of a hand.

ABOUT COMPANY



The subsidiary company model of FPT Corporation

THEORETICAL BASIS

- The economic growth rate can affect the business activities of both the manufacturing and service sectors (Lee, 2014).
- Financial leverage has a negative relationship with financial business activities (Nguyen, T.T.C., Le, A.T.H., and Nguyen, C.V., 2023).
- The financial performance of companies is significantly influenced by financial structure and economic crises (Siminica et al., 2011).

RESEARCH METHODOLOGY

Hypotheses

- H1: GDP index has a positive impact on the operational efficiency of FPT.
- H2: Depreciation rate has a negative impact on the operational efficiency of FPT.
- H3: Current ratio has a positive impact on the operational efficiency of FPT.
- H4: Debt-to-asset ratio has a positive impact on the operational efficiency of FPT.

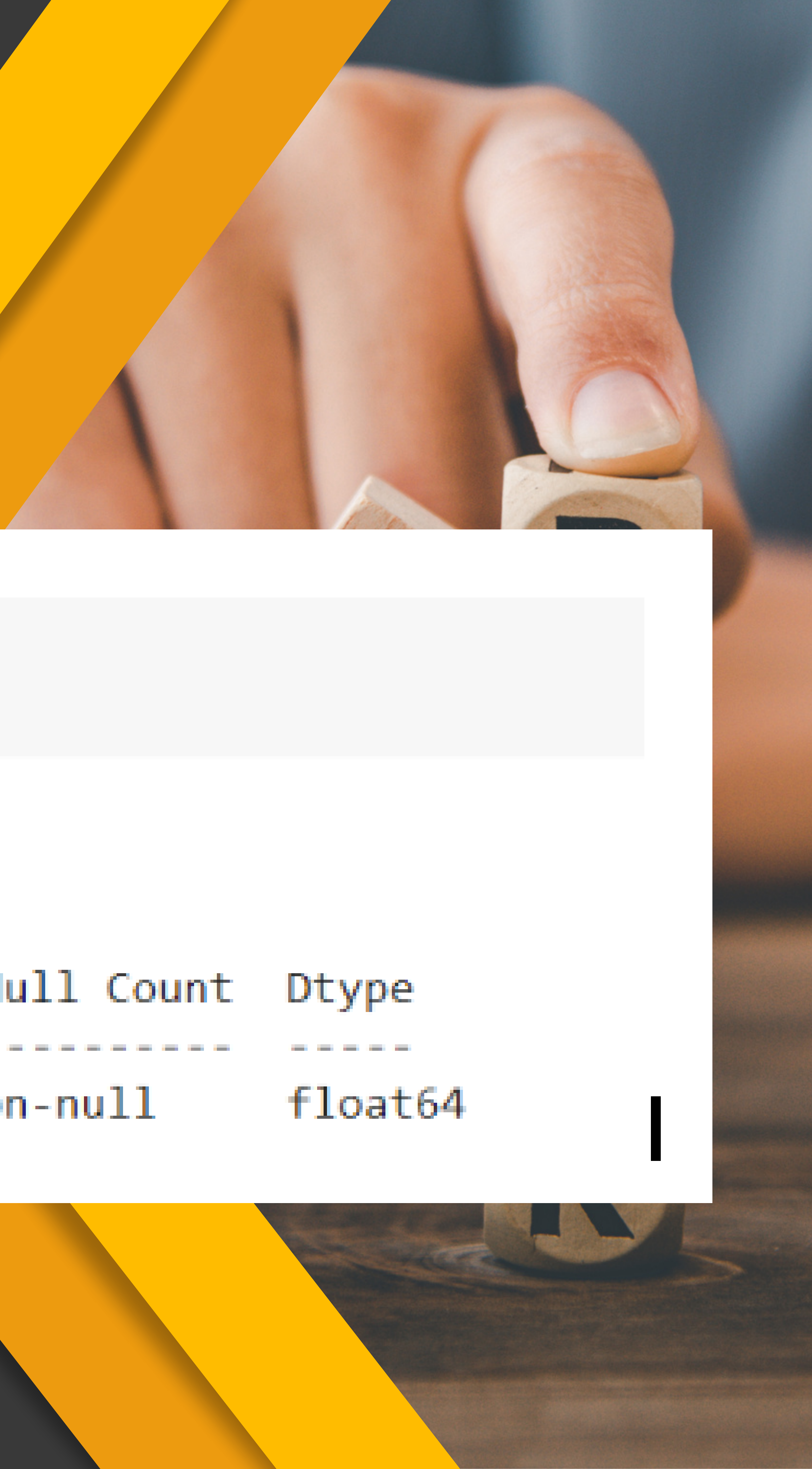
Research data

- The financial reports of FPT Corporation from 2004 to 2022.
- The operational performance of FPT during the period of 2004-2022.
- The data on GDP and the depreciation rate were obtained from the WorldBank website.

MODEL, RESEARCH METHODOLOGY

- **ARIMA model**
- **Linear Regression model**
- **LSTM (Long Short-Term Memory) model**
- **Descriptive statistics**

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022



```
df.set_index('Giai đoạn', inplace = True)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
DatetimeIndex: 19 entries, 2004-01-01 to 2022-01-01
Data columns (total 10 columns):
 #   Column                                       Non-Null Count  Dtype
---  ------                                       --------------  -----
 0   Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)  19 non-null     float64
```

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

```
# Tính giá trị trung bình ROAA
mean_roaa = np.mean(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Giá trị trung bình ROAA:", mean_roaa)

# Tính trung vị ROAA
median_roaa = np.median(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Trung vị ROAA:", median_roaa)

# Tính phạm vi ROAA
range_roaa = np.max(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)']) - np.min(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Phạm vi ROAA:", range_roaa)

# Tính độ lệch chuẩn ROAA
std_deviation_roaa = np.std(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Độ lệch chuẩn ROAA:", std_deviation_roaa)
# Tính min, max của ROAA
min_roaa = np.min(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Giá trị nhỏ nhất của ROAA: ", min_roaa)
max_roaa = np.max(df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)'])
print("Giá trị lớn nhất của ROAA: ",max_roaa)
data = df['Tỷ suất sinh lợi trên tổng tài sản bình quân (ROAA)']
# Tính skewness
print("Giá trị skewness: ", data.skew())
# Tính Standard Error of Mean
print("Giá trị standard Error of Mean: ", data.sem())
# Tính Kurtosis
```

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

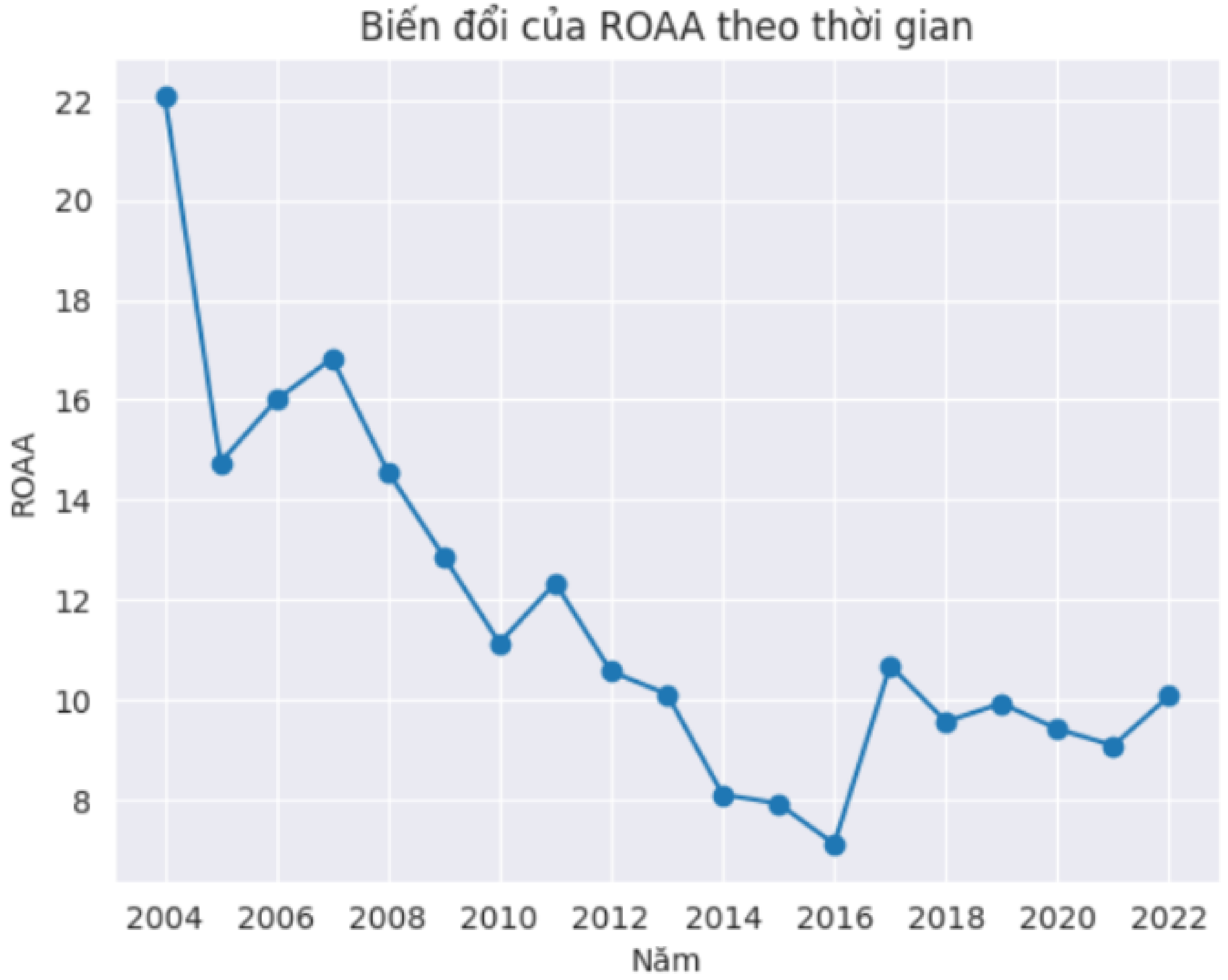
```
Giá trị trung bình ROAA: 11.747894736842106
Trung vị ROAA: 10.57
Phạm vi ROAA: 14.959999999999997
Độ lệch chuẩn ROAA: 3.60784755621959
Giá trị nhỏ nhất của ROAA:  7.12
Giá trị lớn nhất của ROAA:  22.08
Giá trị skewness:  1.340641610136987З
Giá trị standard Error of Mean:  0.8503778241633951
Giá trị Kurtosis:  1.982381362603521Z
```

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

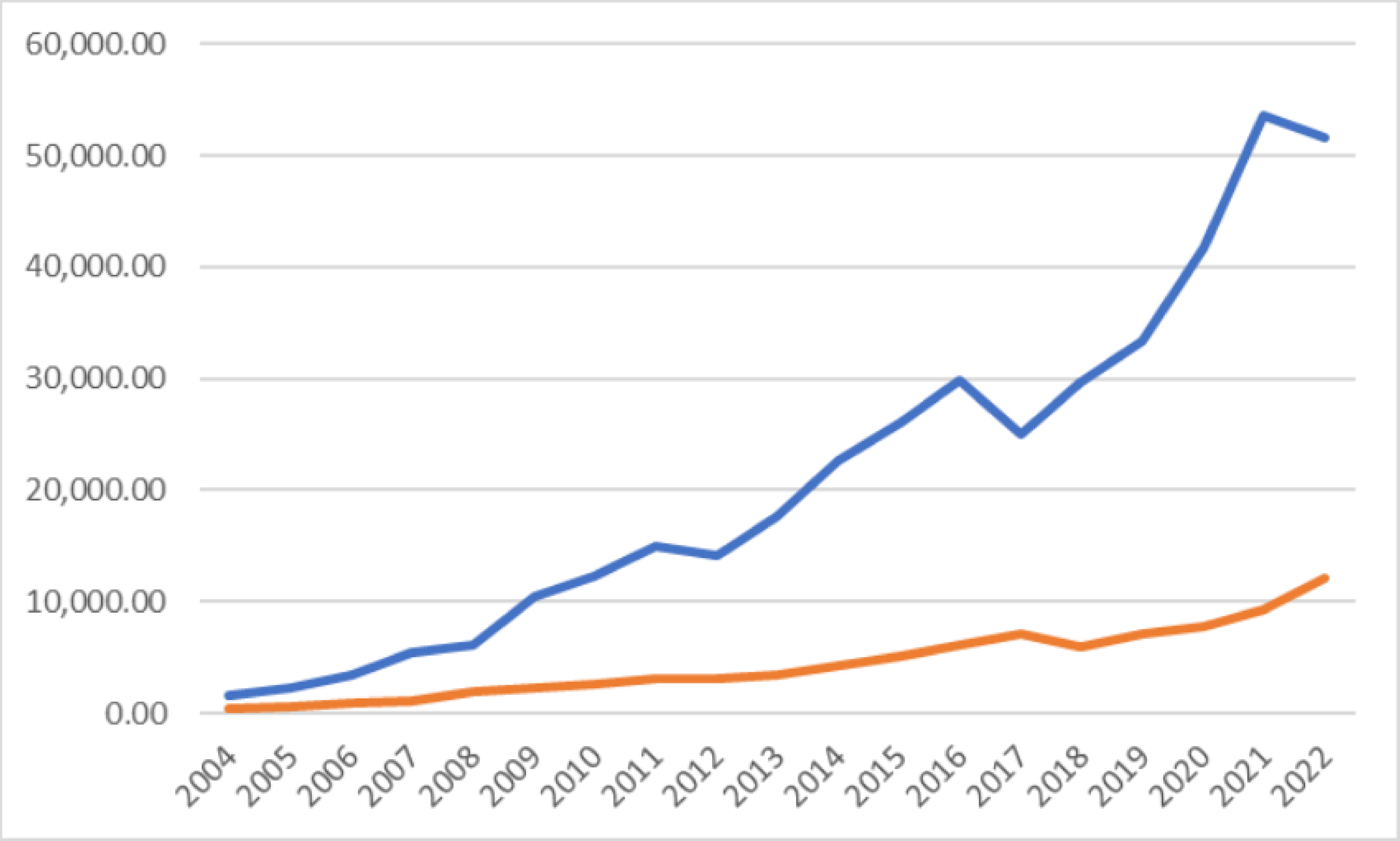
Overall assessment

- The average value of ROAA = 11.75%: Relatively high profitability -> Utilizing their assets to generate profits is an effective strategy.
- Variance of ROAA = 13,74 -> The relatively moderate fluctuation in FPT's financial performance (market volatility, changes in business strategies, as well as economic and political factor.)
- Skewness = 1,34 > 0 -> The success in certain projects or specific business activities in the past indicates FPT's ability to create long-term value, effectively manage business operations, and demonstrate good risk management capabilities.
- Kurtosis = 1,98 < 3 -> There are not many outliers in the distribution of FPT's ROAA, indicating a relatively moderate and stable concentration of profitability. This contributes to the overall stability of the business.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022



ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022



ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- **Sharp decrease from the beginning**

Kết quả kinh doanh - FPT	2004	2005	2006	2007	2008	2009
Giai đoạn	01/01-31/12	01/01-31/12	01/01-31/12	01/01-31/12	01/01-31/12	01/01-31/12
Hợp nhất	Đơn lẻ	Đơn lẻ	Hợp nhất	Hợp nhất	Hợp nhất	Hợp nhất
Kiểm toán	Kiểm toán	Kiểm toán	Kiểm toán	Kiểm toán	Kiểm toán	Kiểm toán
Công ty kiểm toán		AASC	AASC	KPMG	KPMG	Deloitte
Ý kiến kiểm toán			Chấp nhận từng phần - Ngoại trừ	Chấp nhận từng phần - Ngoại trừ	Chấp nhận toàn phần	Chấp nhận toàn phần
Doanh thu thuần về bán hàng và cung cấp dịch vụ	**8,735.00**	**14,101.00**	**21,400.00**	**13,499.00**	**16,382.00**	**18,404.00**
Giá vốn hàng bán	8,197.00	13,180.00	20,049.00	11,537.00	13,403.00	14,719.00
Lợi nhuận gộp về bán hàng và cung cấp dịch vụ	**538.00**	**921.00**	**1,351.00**	**1,961.00**	**2,978.00**	**3,685.00**
Doanh thu hoạt động tài chính	4.00	7.00	11.00	49.00	197.00	188.00
Chi phí tài chính	28.00	40.00	74.00	72.00	495.00	445.00
Chi phí bán hàng	158.00	284.00	358.00	385.00	527.00	527.00
Chi phí quản lý doanh nghiệp	164.00	269.00	436.00	600.00	963.00	1,306.00
Lợi nhuận thuần từ hoạt động kinh doanh	**192.00**	**335.00**	**495.00**	**953.00**	**1,191.00**	**1,594.00**
Lợi nhuận khác	1.00	9.00	114.00	72.00	89.00	33.00
Phần lợi nhuận/lỗ từ công ty liên kết liên doanh	0.00	0.00	0.00	4.00	(40.00)	70.00
Tổng lợi nhuận kế toán trước thuế	**193.00**	**344.00**	**609.00**	**1,029.00**	**1,240.00**	**1,698.00**
Lợi nhuận sau thuế thu nhập doanh nghiệp	**175.00**	**301.00**	**536.00**	**880.00**	**1,051.00**	**1,406.00**
Lợi nhuận sau thuế của cổ đông Công ty mẹ	**175.00**	**280.00**	**450.00**	**737.00**	**836.00**	**1,063.00**
Lãi cơ bản trên cổ phiếu (VNĐ)	92,130.00	5,125.00	8,008.00	0.00	5,959.00	7,498.00

FPT và Microsoft đã chính thức ký biên bản ghi nhớ về phát triển các ứng dụng công nghệ thông tin tại khu vực châu Á – Thái Bình Dương.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- Trending downward

In 2006, FPT received a $36.5 million investment, and this increase in capital may have expanded the company's total assets. While beneficial for long-term growth, it may have temporarily reduced the Return on Average Assets (ROAA) in the short term.

FPT has made significant investments in the financial-banking and real estate sectors through capital contributions to establish FPTS with a charter capital of 440 billion VND, of which FPT contributed 110 billion VND, TPBank with a charter capital of 1,100 billion VND, of which FPT contributed 150 billion VND, and FPT Land with a charter capital of 30 billion VND.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- Trending downward

2007 - 2010, the profitability ratio showed a downward trend, partially reflecting the inefficiency of FPT's relatively diversified investments outside its core business sectors. At this time, a new strategy focused on key business areas was implemented, but FPT believed that more time was needed for the restructuring process to achieve the expected efficiency.

The slow development of the macroeconomic environment led to a decline in the distribution of electronic goods by FPT due to reduced sales. The financial crisis in 2007 also resulted in FPT's stock profits becoming zero. Vietnam's competitiveness ranking according to the World Economic Forum (WEF) also decreased from 64th to 75th during the period of 2006-2009.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- Trending downward

In the telecommunications sector, FPT invested approximately 870 billion VND, specifically allocating 220 billion VND for the Asia Pacific Gateway (APG) international submarine fiber optic cable system, which was expected to be operational in 2014 and land in Da Nang. Additionally, 350 billion VND was invested in backbone telecommunications network infrastructure, and 100 billion VND was allocated for providing telecommunications services to eight provinces and cities along the North-South axis.

FPT expanded its international presence by investing approximately 200 billion VND in providing telecommunications services in four major cities in Cambodia. Additionally, the retail network would be more systematically developed with an investment of 250 billion VND for the FPT Digital Retail chain, which aimed to have 150 stores nationwide.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- **Stability and recovery.**

During the period from 2014 to the present, FPT has experienced a fruitful phase. The corporation has established a stable development direction by focusing on meeting global demands and trends. This is evident in the face of the complex COVID-19 pandemic. In foreign markets, FPT achieved three contracts worth over 100 million USD in 2020, which was an unprecedented achievement prior to the pandemic.

They have developed a solution to achieve three objectives: utilizing technology for task execution and monitoring, project-based activity control with leaders at the center, and implementing internal communication activities to maintain employee enthusiasm.

In 2021, there were 43 internal digital transformation projects related to automation, business process digitization, customer care management, human resources management, and business target and plan management implemented across the entire corporation. These initiatives resulted in cost savings of 98 billion VND and contributed 141 billion VND in revenue. For instance, the Customer Data Platform analyzed data from over 48 million customers using FPT's service platforms.

ANALYZING THE OPERATIONAL PERFORMANCE OF FPT 2004 - 2022

- The profit margin over the past 18 years has shown a consistent trend with continuous fluctuations.
- Always seeking innovation and ongoing development.
- A reliable conglomerate, a major player in the technology sector in Vietnam.
- Based on this analysis, we can formulate the following hypothesis: H5: Stock prices affect ROAA (Return on Average Assets).

Conclusion

ANALYSIS OF THE IMPACT ON RETURN ON AVERAGE ASSETS

Explain the variables in the model

	Variable name	Notation
Dependent variable	ROAA	ROAA
Independent variable	GDP	X1
	Exchange rate	X2
	Current ratio	X3
	Debt ratio	X4
	Stock price	X5

ANALYSIS OF THE IMPACT ON RETURN ON AVERAGE ASSETS

Descriptive Statistics

	ROAA	GDP Growth	exchange rate	Current ratio	Debt ratio
ROAA	1.000000	0.285354	-0.837405	0.425994	0.506241
GDP Growth	0.285354	1.000000	-0.261565	0.177339	0.082710
exchange rate	-0.837405	-0.261565	1.000000	-0.648484	-0.502023
Current ratio	0.425994	0.177339	-0.648484	1.000000	-0.092473
Debt ratio	0.506241	0.082710	-0.502023	-0.092473	1.000000

Utilizing the Linear Regression model to explore the impact

The linear regression model between the dependent variable Return on Average Assets (ROAA) and two independent variables, GDP Growth, and Exchange, is established.

```
                               OLS Regression Results
==========================================================================================
Dep. Variable:     Return on Average Assets (ROAA)   R-squared:                       0.706
Model:                                         OLS   Adj. R-squared:                  0.669
Method:                              Least Squares   F-statistic:                     19.21
Date:                             Thu, 16 Nov 2023   Prob (F-statistic):           5.59e-05
Time:                                     23:03:15   Log-Likelihood:                -39.710
No. Observations:                               19   AIC:                             85.42
Df Residuals:                                   16   BIC:                             88.25
Df Model:                                        2
Covariance Type:                         nonrobust
=================================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
---------------------------------------------------------------------------------
const            31.4550      4.704      6.687      0.000      21.483      41.427
GDP Growth        0.1794      0.354      0.507      0.619      -0.571       0.930
Exchange rate    -0.0010      0.000     -5.830      0.000      -0.001      -0.001
==============================================================================
Omnibus:                        6.643   Durbin-Watson:                   1.278
Prob(Omnibus):                  0.036   Jarque-Bera (JB):                3.983
Skew:                           0.914   Prob(JB):                        0.137
Kurtosis:                       4.301   Cond. No.                     1.95e+05
==============================================================================
```

Utilizing the Linear Regression model to explore the impact

The linear regression model between the dependent variable Return on Average Assets (ROAA) and two independent variables, Current Ratio (short term), and Financial Leverage Ratio (Debt Ratio)

```
                                  OLS Regression Results
==============================================================================================
Dep. Variable:     Return on Average Assets (ROAA)   R-squared:                       0.193
Model:                                         OLS   Adj. R-squared:                  0.168
Method:                              Least Squares   F-statistic:                     7.788
Date:                             Thu, 16 Nov 2023   Prob (F-statistic):           0.000929
Time:                                     23:07:50   Log-Likelihood:                -121.81
No. Observations:                               68   AIC:                             249.6
Df Residuals:                                   65   BIC:                             256.3
Df Model:                                        2
Covariance Type:                         nonrobust
==================================================================================================
                                 coef    std err          t      P>|t|      [0.025      0.975]
--------------------------------------------------------------------------------------------------
const                          0.1150      2.575      0.045      0.965      -5.027       5.257
Current ratio (short term)     2.7781      0.794      3.498      0.001       1.192       4.364
Finace leverage (Debt ratio)  -0.0233      0.037     -0.633      0.529      -0.097       0.050
==============================================================================
Omnibus:                       86.297   Durbin-Watson:                   1.595
Prob(Omnibus):                  0.000   Jarque-Bera (JB):             1363.481
Skew:                           3.634   Prob(JB):                    8.39e-297
Kurtosis:                      23.698   Cond. No.                         786.
==============================================================================
```

Utilizing the Linear Regression model to explore the impact

The linear regression model between the dependent variable Return on Average Assets (ROAA) and the independent variable Stock Price

```
                            OLS Regression Results
==============================================================================
Dep. Variable:     Return on Average Assets (ROAA)   R-squared:                       0.016
Model:                                         OLS   Adj. R-squared:                  0.000
Method:                              Least Squares   F-statistic:                     1.028
Date:                             Fri, 17 Nov 2023   Prob (F-statistic):              0.314
Time:                                     15:01:51   Log-Likelihood:                -93.695
No. Observations:                               67   AIC:                             191.4
Df Residuals:                                   65   BIC:                             195.8
Df Model:                                        1
Covariance Type:                         nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const          2.8538      0.172     16.546      0.000       2.509       3.198
Price      -4.944e-06   4.88e-06     -1.014      0.314   -1.47e-05    4.79e-06
==============================================================================
Omnibus:                       23.938   Durbin-Watson:                   1.444
Prob(Omnibus):                  0.000   Jarque-Bera (JB):               38.332
Skew:                           1.325   Prob(JB):                     4.74e-09
Kurtosis:                       5.589   Cond. No.                     5.02e+04
==============================================================================
```

Utilize ANOVA

Select an appropriate model

Conclusion

STOCK PREDICTION

Utilize the LSTM model

Utilize the LSTM model

STOCK PREDICTION

LTSM Model

```
# Read the CSV file
df = pd.read_csv('D:/Desktop/FPT Historical Data.csv')
```

```
# some format data
df.set_index('Date',drop=True,inplace=True)
```

+ Code

```
df.head(20)
```

```
# Plot using hvplot
plot = df.hvplot.line(title='Time Series Plot', xlabel='Date', ylabel='Price')

# Show the plot
plot
```

Step 1: Data Preparation

STOCK PREDICTION

LSTM Model

```
# Extract the 'Close' prices for modeling
prices = df['Price'].values.reshape(-1, 1)
```

```
# Normalize the data
scaler = MinMaxScaler(feature_range=(0, 1))
prices_scaled = scaler.fit_transform(prices)
```

```
prices_scaled
```

```
array([[0.14617319],
       [0.15503653],
       [0.16432273],
       ...,
       [0.92188598],
       [0.92084446],
       [0.93334271]])
```

Step 1: Data Preparation

STOCK PREDICTION

LSTM Model

Step 2: Sequence Creation

```
# Function to create sequences for LSTM
def create_sequences(data, seq_length):
    X, y = [], []
    for i in range(len(data) - seq_length):
        X.append(data[i:i + seq_length, 0])
        y.append(data[i + seq_length, 0])
    return np.array(X), np.array(y)
```

```
# Set the sequence length (number of time steps to look back)
sequence_length = 10
```

```
# Create sequences for training
X, y = create_sequences(prices_scaled, sequence_length)
```

Mô hình LSTM

```
# function to plit the data into train, dev, and test sets
def train_dev_test_split(data, seq_length, train_ratio=0.75, dev_ratio=0.15):
    total_size = len(data)
    train_size = int(total_size * train_ratio)
    dev_size = int(total_size * dev_ratio)

    X, y = create_sequences(data, seq_length)

    X_train, y_train = X[:train_size], y[:train_size]
    X_dev, y_dev = X[train_size:train_size + dev_size], y[train_size:train_size + dev_size]
    X_test, y_test = X[train_size + dev_size:], y[train_size + dev_size:]

    return (X_train, y_train), (X_dev, y_dev), (X_test, y_test)
```

```
# X_train, y_train = X[:train_size], y[:train_size]
# X_dev, y_dev = X[train_size:train_size + dev_size], y[train_size:train_size + dev_size]
# X_test, y_test = X[train_size + dev_size:], y[train_size + dev_size:]

(X_train, y_train), (X_dev, y_dev), (X_test, y_test) = train_dev_test_split(prices_scaled, SEQUENCE_LENGTH)
```

Step 3: Preprocessing

STOCK PREDICTION

```python
# Reshape input data to be 3D for LSTM input (samples, time steps, features)
X_train = np.reshape(X_train, (X_train.shape[0], X_train.shape[1], 1))
X_dev = np.reshape(X_dev, (X_dev.shape[0], X_dev.shape[1], 1))
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))
```

```
array([[[0.14617319],
        [0.15503653],
        [0.16432273],
        ...,
        [0.15377004],
        [0.16263442],
        [0.17149776]],

       [[0.15503653],
        [0.16432273],
        [0.17403074],
        ...,
        [0.16263442],
        [0.17149776],
        [0.17149776]],

       [[0.16432273],
        [0.17403074],
        [0.18416056],
        ...,
        [0.17149776],
        [0.17149776],
        [0.16305624]],

       ...,

...
        [0.30187213],
        ...,
        [0.30373229],
        [0.30435199],
        [0.31923323]]])
```

Step 4: Reshape data for LSTM

STOCK PREDICTION

LSTM Model

```
model = Sequential([
    LSTM(50, return_sequences=True, input_shape=(SEQUENCE_LENGTH, 1)),
    LSTM(50, return_sequences=True),
    LSTM(50),
    Dense(1)
])

model.compile(optimizer='adam', loss='mean_squared_error')
```

```
# Train the model
history = model.fit(X_train, y_train, epochs=100, batch_size=32, validation_data=(X_dev, y_dev))
```

Step 5: LSTM model definition and Training

STOCK PREDICTION

LSTM Model

```
Epoch 77/100
99/99 [==============================] - 1s 11ms/step - loss: 1.1200e-05 - val_loss: 9.7285e-05
Epoch 78/100
99/99 [==============================] - 1s 12ms/step - loss: 1.1241e-05 - val_loss: 1.2983e-04
Epoch 79/100
99/99 [==============================] - 1s 10ms/step - loss: 1.1641e-05 - val_loss: 2.0910e-04
Epoch 80/100
99/99 [==============================] - 1s 12ms/step - loss: 9.8998e-06 - val_loss: 1.0814e-04
Epoch 81/100
99/99 [==============================] - 1s 11ms/step - loss: 1.0621e-05 - val_loss: 1.1567e-04
Epoch 82/100
99/99 [==============================] - 1s 11ms/step - loss: 9.6715e-06 - val_loss: 1.0763e-04
Epoch 83/100
99/99 [==============================] - 1s 12ms/step - loss: 1.0937e-05 - val_loss: 9.6378e-05
Epoch 84/100
99/99 [==============================] - 1s 11ms/step - loss: 9.4299e-06 - val_loss: 9.5888e-05
Epoch 85/100
99/99 [==============================] - 1s 12ms/step - loss: 9.2853e-06 - val_loss: 1.0034e-04
Epoch 86/100
99/99 [==============================] - 1s 11ms/step - loss: 1.0348e-05 - val_loss: 1.0532e-04
...
Epoch 99/100
99/99 [==============================] - 1s 13ms/step - loss: 9.2419e-06 - val_loss: 1.6366e-04
Epoch 100/100
99/99 [==============================] - 1s 11ms/step - loss: 1.0562e-05 - val_loss: 1.0573e-04
```

Output is truncated. View as a scrollable element or open in a text editor. Adjust cell output settings...

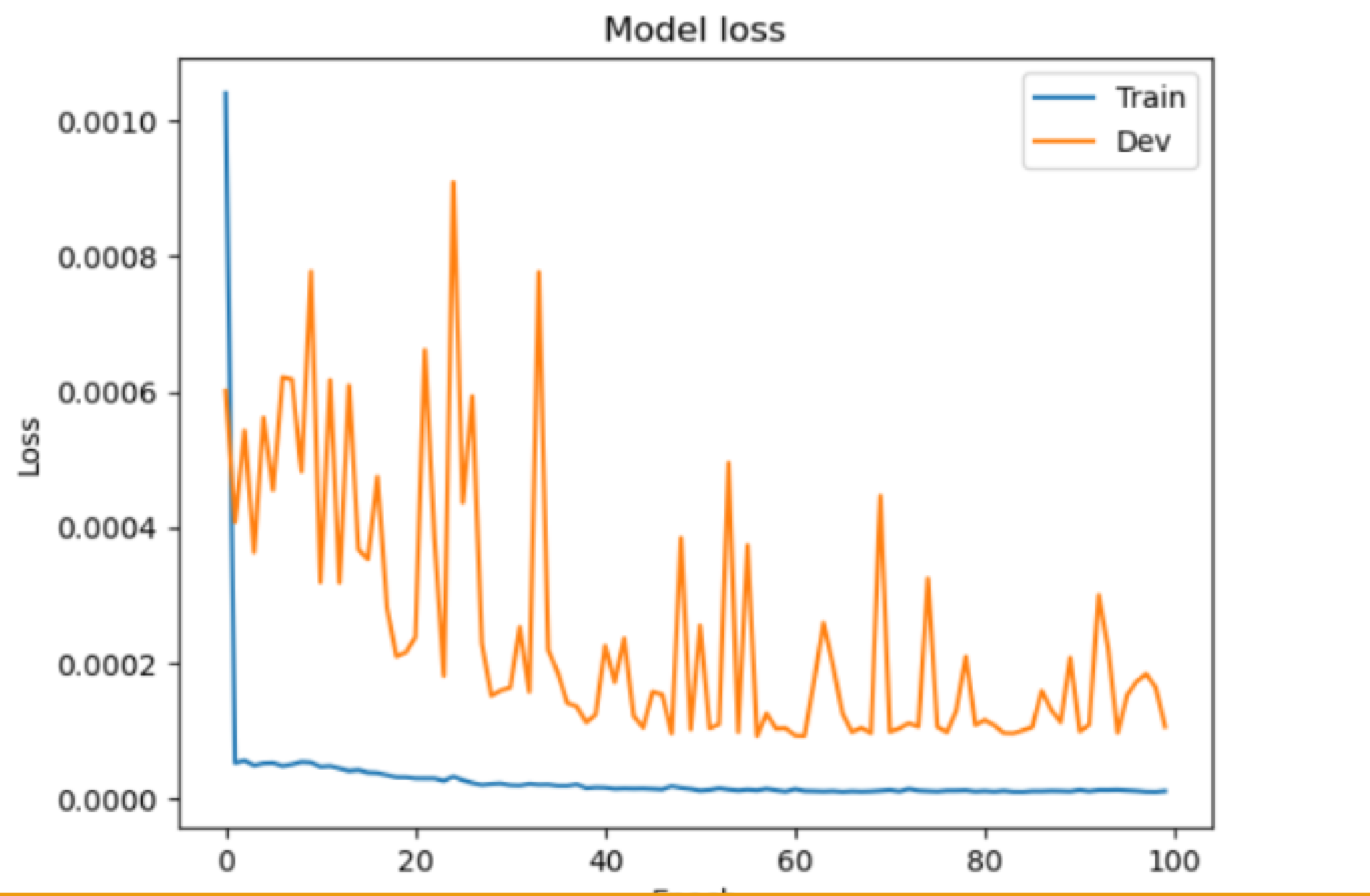
Step 5: LSTM model definition and Training

STOCK PREDICTION

LSTM Model

Step 6: Model Evaluation and Visualization

```python
# Plot training & validation loss values
plt.plot(history.history['loss'])
plt.plot(history.history['val_loss'])
plt.title('Model loss')
plt.ylabel('Loss')
plt.xlabel('Epoch')
plt.legend(['Train', 'Dev'], loc='upper right')
plt.show()
```



STOCK PREDICTION

Mô hình LSTM

```
# Make predictions on the test set
predictions = model.predict(X_test)
```

```
13/13 [==============================] - 1s 4ms/step
```

```
# Inverse transform the predictions and actual values to the original scale
predictions_inv = scaler.inverse_transform(predictions)
y_test_inv = scaler.inverse_transform(y_test.reshape(-1, 1))
```

```
# Compare predictions with actual values
comparison_df = pd.DataFrame({'Actual': y_test_inv.flatten(), 'Predicted': predictions_inv.flatten()})
print(comparison_df)
```

	Actual	Predicted
0	78833.0	78229.328125
1	80810.0	78054.054688
2	86412.0	80674.695312
3	87318.0	86834.726562
4	88142.0	85498.054688
..	...	...

	Actual	Predicted
0	78833.0	78229.328125
1	80810.0	78054.054688
2	86412.0	80674.695312
3	87318.0	86834.726562
4	88142.0	85498.054688
..	...	...

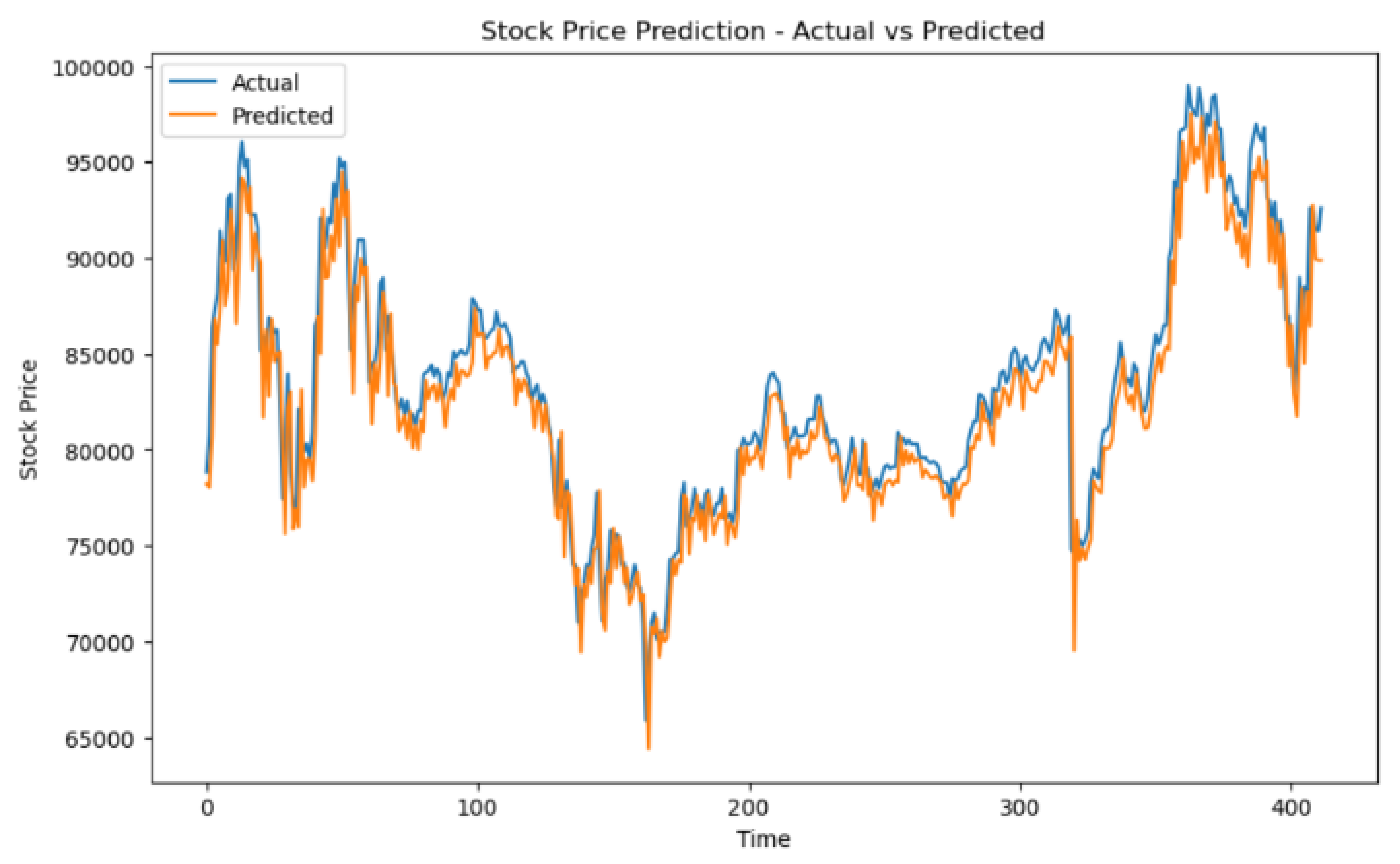
	Actual	Predicted
407	92600.0	86426.726562
408	92500.0	92745.398438
409	91500.0	89931.062500
410	91400.0	89872.906250
411	92600.0	89879.257812

[412 rows x 2 columns]

Step 6: Model Evaluation and Visualization

STOCK PREDICTION

LSTM Model



Step 6: Model Evaluation and Visualization

LSTM Model

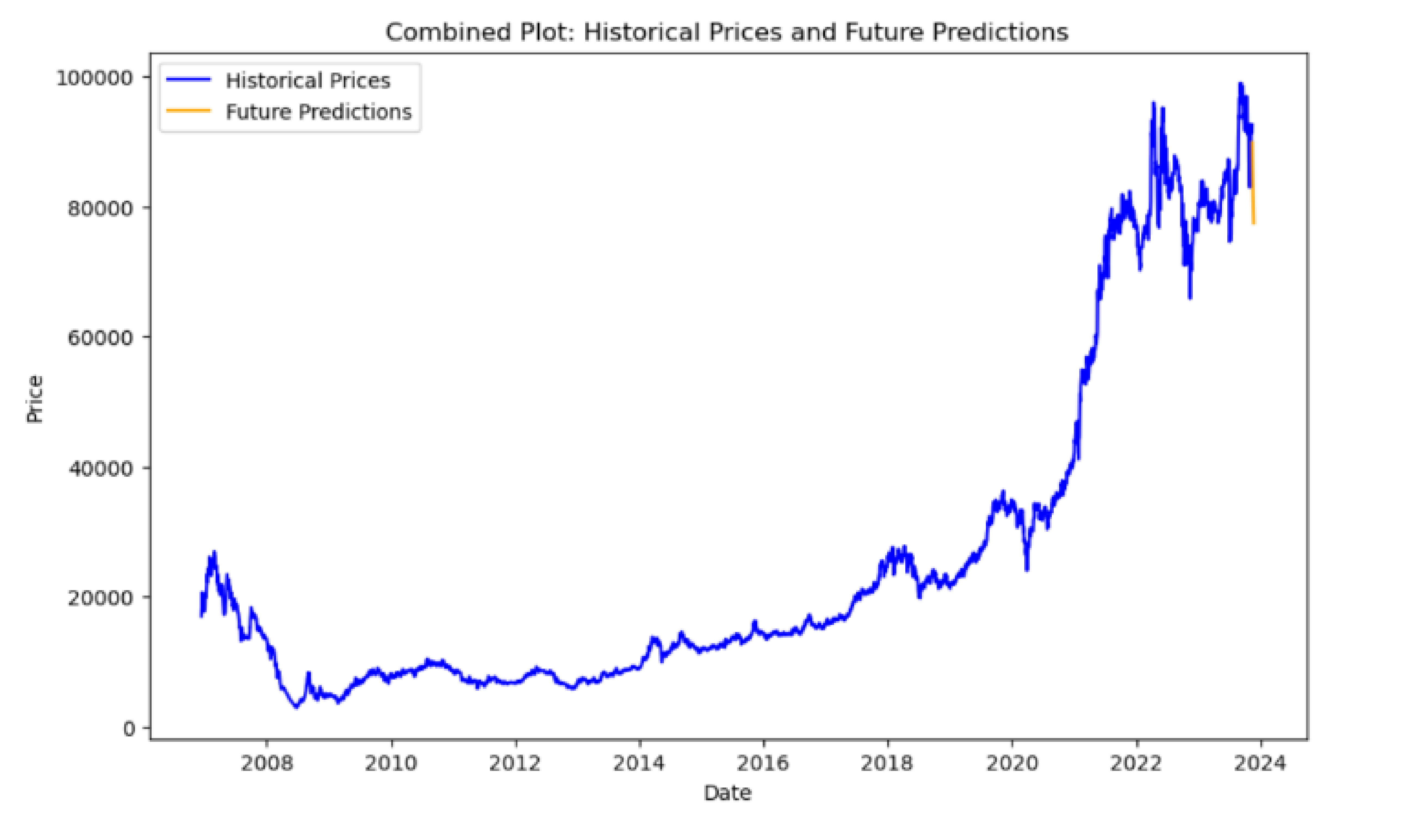
```
# Calculate evaluation metrics
mse = mean_squared_error(y_test_inv, predictions_inv)
mae = mean_absolute_error(y_test_inv, predictions_inv)
r2 = r2_score(y_test_inv, predictions_inv)

# Print the evaluation metrics
print(f'Mean Squared Error (MSE): {mse:.2f}')
print(f'Mean Absolute Error (MAE): {mae:.2f}')
print(f'R-squared (R2): {r2:.2f}')
```

```
Mean Squared Error (MSE): 4149204.31
Mean Absolute Error (MAE): 1529.53
R-squared (R2): 0.90
```

Step 6: Model Evaluation and Visualization

STOCK PREDICTION

Mô hình LSTM



Step 7 : Predict future price

STOCK PREDICTION

Arima Model

```
In [2]: # Read the CSV file
        df = pd.read_csv('D:/Desktop/FPT Historical Data Monthly.csv')
```

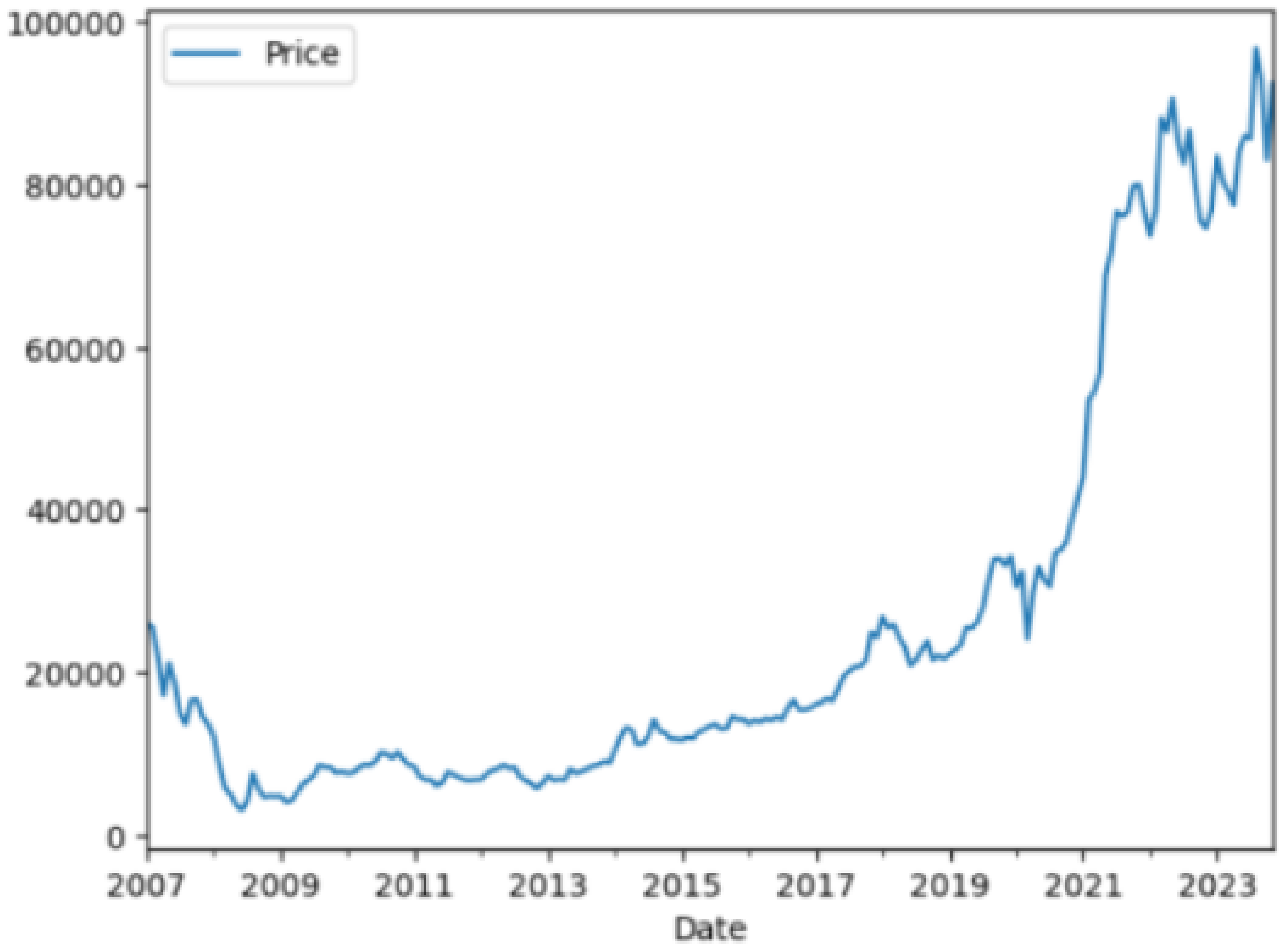
	Price
Date	
2007-01-01	26139.5
2007-02-01	25612.7
2007-03-01	21945.5
2007-04-01	17240.0
2007-05-01	21235.6
...	...
2023-07-01	85600.0
2023-08-01	96700.0
2023-09-01	92800.0
2023-10-01	83000.0
2023-11-01	92400.0

203 rows × 1 columns

Step 1: collect data

STOCK PREDICTION

Arima Model



Step 2: Plot and consider

STOCK PREDICTION

Arima Model

Step 3: Estimate whether this is non-stationary data or stationary data by using ADF

```
# adfuller for data train

adf_0 = adfuller(df)
print('ADF stationary is:', adf_0[0])
print('p-value of ADF test is:', adf_0[1])
print('Critical values:')
for key, value in adf_0[4].items():
    print('\t%s: .%.3f'%(key,value))
```

```
ADF stationary is: 0.1614096592298381
p-value of ADF test is: 0.970012201805755
Critical values:
	1%: .-3.466
	5%: .-2.877
	10%: .-2.575
```

STOCK PREDICTION

Arima Model

```python
# Function to perform differencing
def difference(data, d=1):
    """
    Perform differencing on a time series data for 'd' orders.
    data: Pandas Series or DataFrame
    d: Order of differencing (default = 1 for first difference)
    """
    if isinstance(data, pd.Series):
        # If data is a Pandas Series
        diff_series = data.diff(periods=d).dropna()
        return pd.DataFrame({f'Difference_{d}': diff_series})
    elif isinstance(data, pd.DataFrame):
        # If data is a Pandas DataFrame, operate on the specified column
        diff_series = data.iloc[:, 0].diff(periods=d).dropna()
        return pd.DataFrame({f'Difference_{d}': diff_series})
    else:
        raise ValueError("Input should be a Pandas Series or DataFrame")
```

Step 3: Estimate whether this is non-stationary data or stationary data by using ADF

STOCK PREDICTION

Arima Model

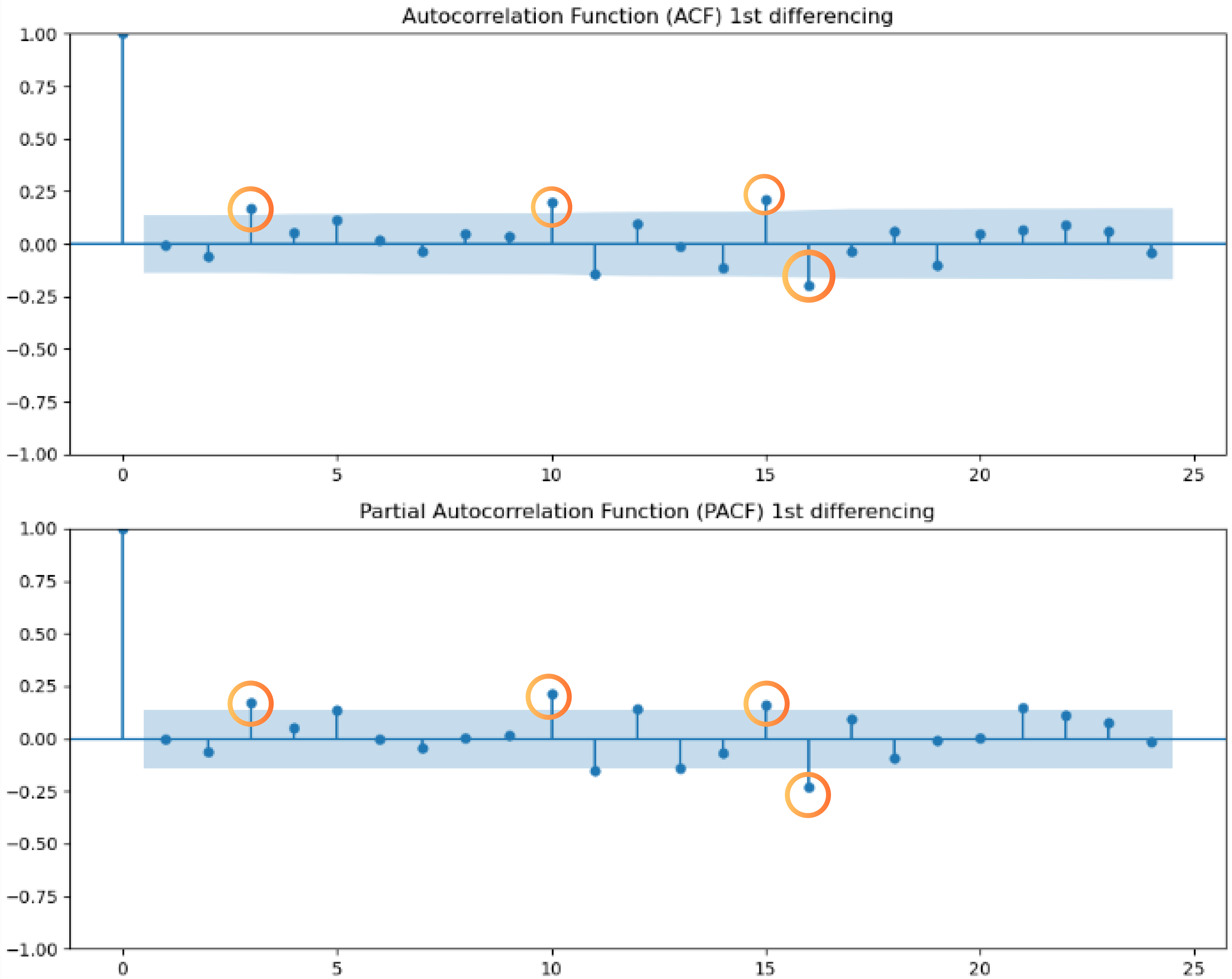
```
adf_1 = adfuller(first_difference)
print('ADF 1diff stationary is:', adf_1[0])
print('p-value of ADF 1diff test is:', adf_1[1])
print('Critical values 1diff:')
for key, value in adf_1[4].items():
    print('\t%s: .%.3f'%(key,value))
```

```
ADF 1diff stationary is: -3.788044302133533
p-value of ADF 1diff test is: 0.003032033582000458
Critical values 1diff:
	1%: .-3.466
	5%: .-2.877
	10%: .-2.575
```

Step 3: Estimate whether this is non-stationary data or stationary data by using ADF

STOCK PREDICTION

Arima Model



Step 4: Identify the order of the AR term (p) and MA term (q).

STOCK PREDICTION

Arima Model

```
# Define the p, d, q values
p_values = [3, 10, 15, 16]
d_value = 1
q_values = [3, 10, 15, 16]
```

```
# Create an empty list to store the results
results = []
```

Step 4: Identify the order of the AR term (p) and MA term (q).

```
# Loop over the p and q values
for p in p_values:
    for q in q_values:
        # Try to fit the ARIMA model
        try:
            model = ARIMA(train, order=(p, d_value, q))
            model_fit = model.fit()

            # Forecast on the test set
            forecast = model_fit.forecast(steps=len(test))

            # Calculate MAE and RMSE
            mae = mean_absolute_error(test, forecast)
            rmse = np.sqrt(mean_squared_error(test, forecast))

            # Append the model summary and the criteria to the results list
            results.append((p, q, model_fit.summary(), model_fit.llf,
                                    model_fit.aic, model_fit.bic, model_fit.hqic, mae, rmse))
        # Handle the exception if the model fails to converge
        except Exception as e:
            print(f'Failed to fit ARIMA({p}, {d_value}, {q}) model due to {e}')

# Sort the results by the log likelihood in descending order
results.sort(key=lambda x: x[3], reverse=True)
```

STOCK PREDICTION

Arima Model

```
---------------------------------------
ARIMA(10, 1, 3) model
Log likelihood: -1616.4888186699739
AIC: 3260.9776373399477
BIC: 3305.756595777669
HQIC: 3279.1319847780437
MAE: 9691.878929143617
RMSE: 11256.540774993191
---------------------------------------
ARIMA(3, 1, 3) model
Log likelihood: -1619.524612227854
AIC: 3253.049224455708
BIC: 3275.438703674569
HQIC: 3262.126398174756
MAE: 9029.761385807647
RMSE: 10656.67611423437
---------------------------------------
```

Step 4: Identify the order of the AR term (p) and MA term (q).

STOCK PREDICTION

Arima Model

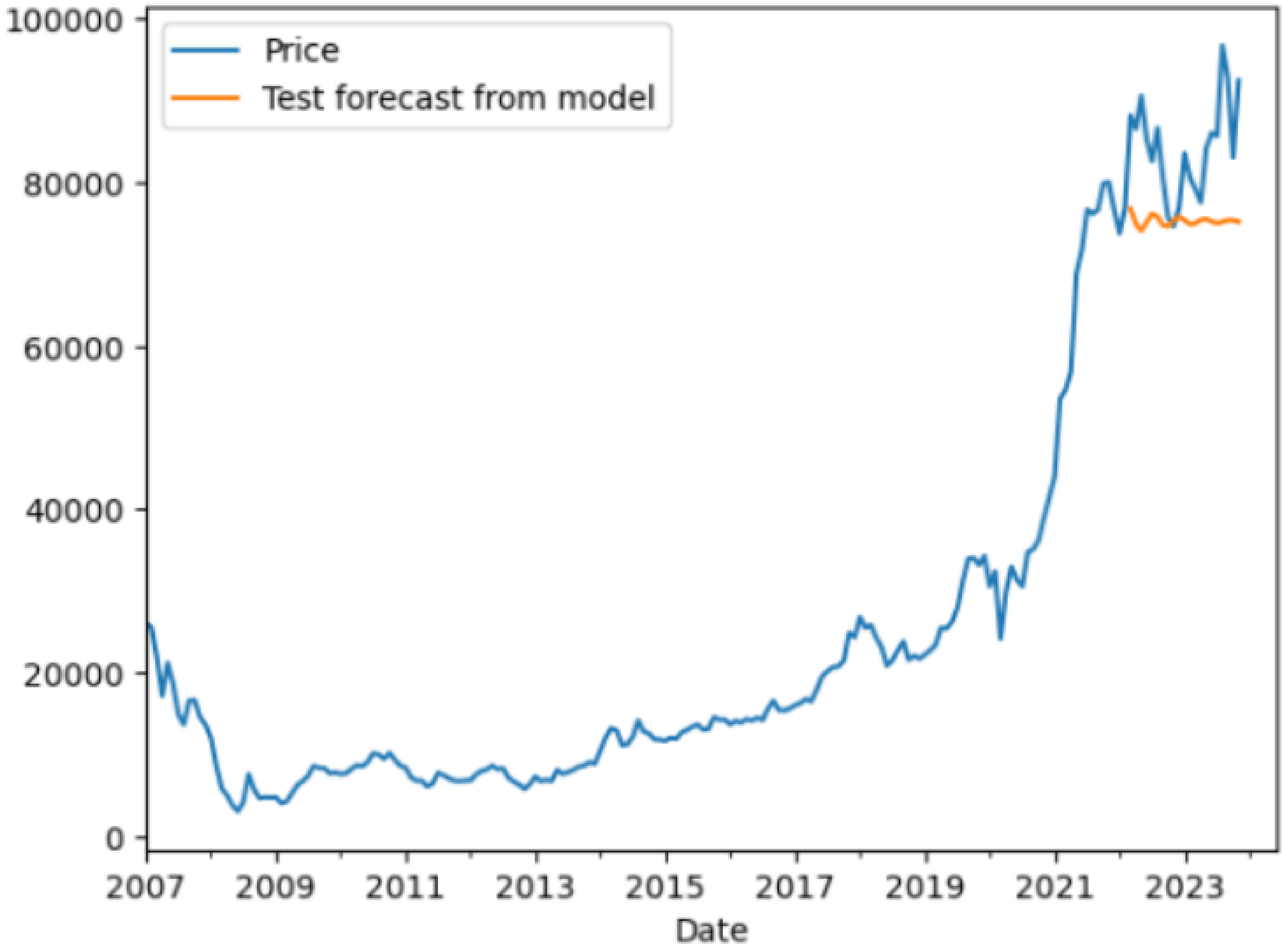
```
                               SARIMAX Results
==============================================================================
Dep. Variable:                  Price   No. Observations:                  182
Model:                 ARIMA(3, 1, 3)   Log Likelihood               -1619.525
Date:                Fri, 17 Nov 2023   AIC                           3253.049
Time:                        14:03:11   BIC                           3275.439
Sample:                    01-01-2007   HQIC                          3262.126
                         - 02-01-2022
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
ar.L1          1.0269      0.164      6.259      0.000       0.705       1.348
ar.L2         -1.0448      0.122     -8.593      0.000      -1.283      -0.806
ar.L3          0.4739      0.146      3.246      0.001       0.188       0.760
ma.L1         -1.0310      0.185     -5.586      0.000      -1.393      -0.669
ma.L2          1.0985      0.150      7.343      0.000       0.805       1.392
ma.L3         -0.2901      0.186     -1.561      0.119      -0.654       0.074
sigma2      3.179e+06   1.83e+05     17.325      0.000    2.82e+06    3.54e+06
===================================================================================
Ljung-Box (L1) (Q):                   1.23   Jarque-Bera (JB):               493.06
Prob(Q):                              0.27   Prob(JB):                         0.00
Heteroskedasticity (H):               3.20   Skew:                             0.54
Prob(H) (two-sided):                  0.00   Kurtosis:                        11.01
===================================================================================
```

Step 4: Identify the order of the AR term (p) and MA term (q).

The limitations of the model

- The coefficient for ma.L3 is not statistically significant, indicating the presence of residual component in MA

- The errors exhibit higher kurtosis than the normal distribution, suggesting an influence from outliers or abnormal fluctuations



Draw predictions

```python
# predict for6 month in future
forecast__313 = model_fit_313.forecast(len(test)+3)
print("\nForecast:")
print(forecast__313)
```

```
Forecast:
2022-03-01    76743.630573
2022-04-01    74873.368964
2022-05-01    74041.129553
2022-06-01    75087.070675
2022-07-01    76144.246725
2022-08-01    75742.677361
2022-09-01    74721.525206
2022-10-01    74593.481478
2022-11-01    75338.529426
2022-12-01    75753.431133
2023-01-01    75340.415497
2023-02-01    74835.924879
2023-03-01    74946.001609
2023-04-01    75390.365413
```

Draw predictions

	ARIMA	LSTM
RMSE	10656.68	2036.9
MAE	9020.76	1529.53

Compare two models

CONCLUSION AND RECOMMENDATIONS

General Conclusion

- The Linear Regression model is suitable for analyzing this dataset.
- From the results, it can be concluded that the independent variables X1, X2, X3, X4, and H5 exhibit mutual influence on the variable ROAA. The obtained results support the hypotheses H1, H2, H3, H4, and H5

Recommendations

- Enterprises should learn from FPT on how to effectively manage risks and allocate financial resources.
- Utilizing financial tools such as risk management and default rate management can minimize negative impacts.
- Diversifying business operations and consumer markets can reduce dependence.

Thank You

For Your Attention