EXERCISE NO. 03
LINEAR REGRESSION MODEL FOR FORECASTING TIME SERIES DATA

AIM:
To develop a linear regression model for forecasting time series data.

ALGORITHM:
1. Import necessary libraries.
2. Load the dataset.
3. Preprocess the dataset.
4. Prepare the data for the linear regression by creating lagged features, creating X and y variables, and split the dataset into training and test data.
5. Initialise the linear regression model and fit the model into training data.
6. Evaluate the model performance.
7. Visualise the actual vs predicted performance.

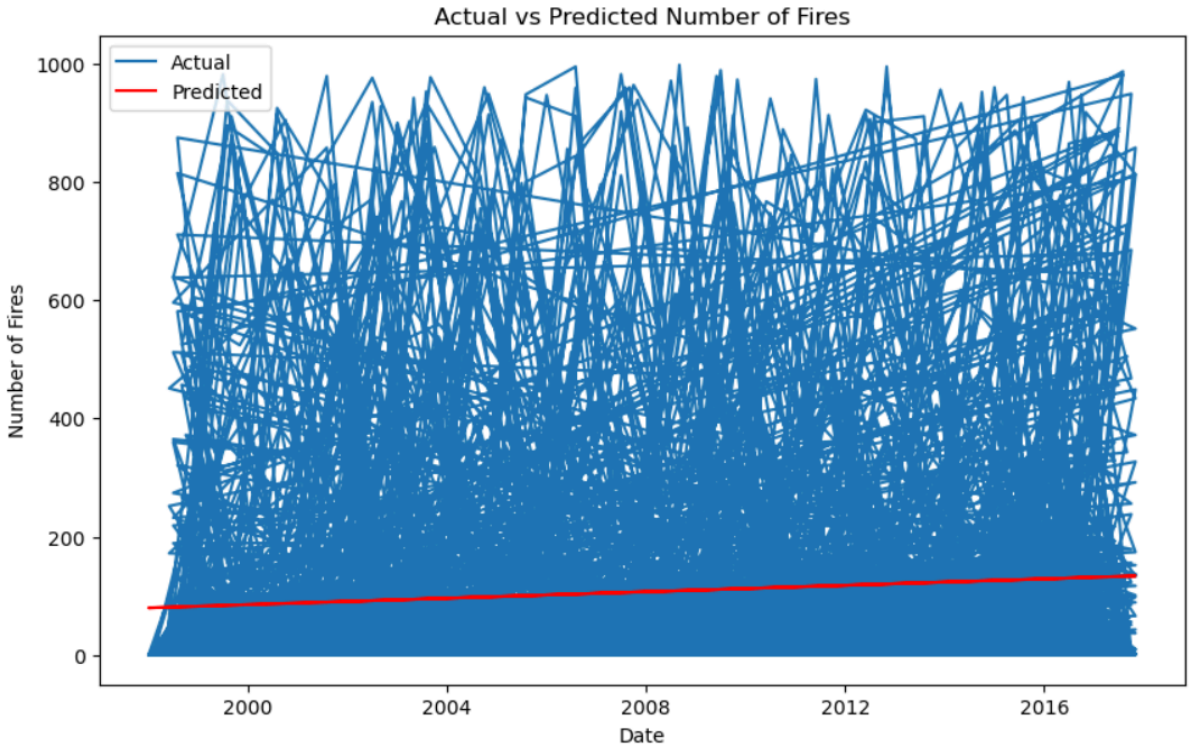
PROGRAM:

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error
df = pd.read_csv('.../amazon.csv', encoding='latin1')
month_map = {
  'Janeiro': 'January', 'Fevereiro': 'February', 'Março': 'March',
  'Abril': 'April', 'Maio': 'May', 'Junho': 'June',
  'Julho': 'July', 'Agosto': 'August', 'Setembro': 'September',
  'Outubro': 'October', 'Novembro': 'November', 'Dezembro': 'December'
}

df['month'] = df['month'].map(month_map)
df['date'] = pd.to_datetime(df['month'] + ' ' + df['year'].astype(str), format='%B %Y')
df.set_index('date', inplace=True)
 df['Month'] = df.index.map(pd.Timestamp.toordinal)

X = df['Month'].values.reshape(-1, 1)
y = df['number']

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, shuffle=False)
model = LinearRegression()
model.fit(X_train, y_train)
y_pred = model.predict(X_test)
```

```
mse = mean_squared_error(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
plt.figure(figsize=(10, 6))
plt.plot(df.index, df['number'], label='Actual')
plt.plot(df.index[-len(y_test):], y_pred, label='Predicted', color='red')
plt.title('Actual vs Predicted Number of Fires')
plt.xlabel('Date')
plt.ylabel('Number of Fires')
plt.legend()
plt.show()
```

OUTPUT:

Mean Squared Error: 37347.68350433537



RESULT:

Thus the program to develop a linear regression model for forecasting time series data has been successfully implemented and verified.