APPENDIX

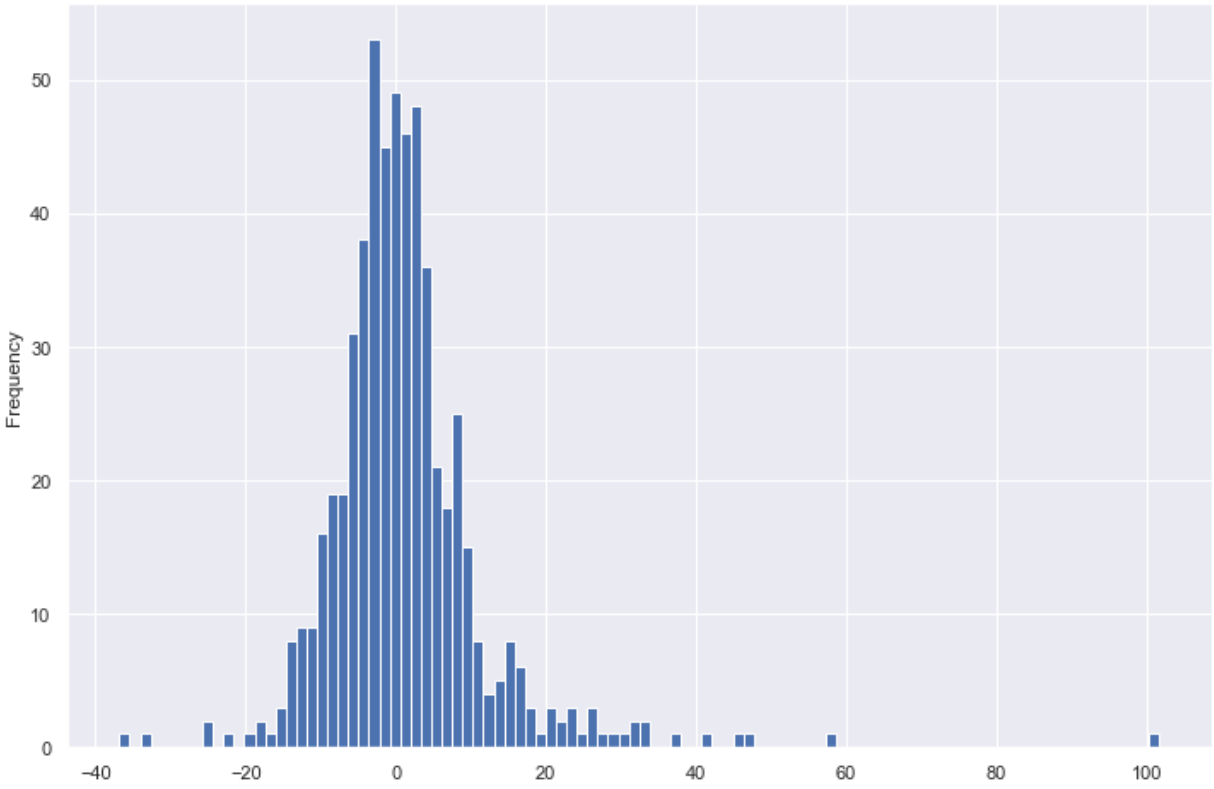
In [1]:
```
import numpy as np
import pandas as pd
from sklearn.preprocessing import StandardScaler, PolynomialFeatures
from sklearn.linear_model import LinearRegression
from sklearn.ensemble import RandomForestRegressor
import pandas_datareader
import matplotlib.pyplot as plt
import seaborn as sns
import datetime
import warnings
import yfinance as yf
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Dense, Dropout, LSTM, RepeatVector
from tensorflow.keras.optimizers import Adam
from tensorflow.keras.losses import MeanSquaredError
warnings.filterwarnings("ignore")
sns.set_theme(style="darkgrid")
%matplotlib inline
```

In [2]:
```
crypto = "SAND-USD"
df = yf.download(crypto, start="2021-01-01", end="2022-08-01")
```

```
[**********100%**********]  1 of 1 completed
```

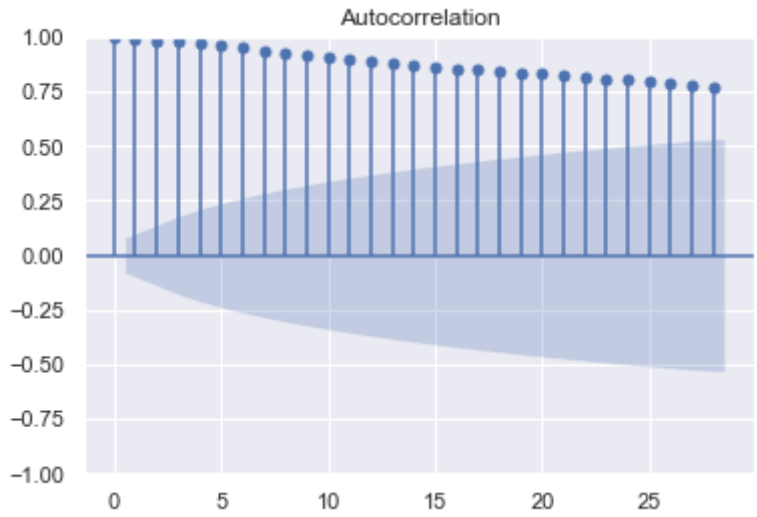
In [3]:
```
plt.figure(figsize=(14,8))
plt.plot(df["Close"], color="black", label="Close Price")
plt.legend()
plt.title(crypto, size=24)
plt.show()
```

In [4]:
```
df["Lag"] = df["Close"].shift(periods=1)
df["Diff"] = df["Close"].sub(df["Lag"])
df["Return"] = ((df["Close"].div(df["Lag"])).sub(1)).mul(100)
df = df[["Close","Return"]]
retn = df["Return"]
retn = retn.dropna()
retn.plot(kind="hist", figsize=(12,8),bins=100)
plt.show()
```



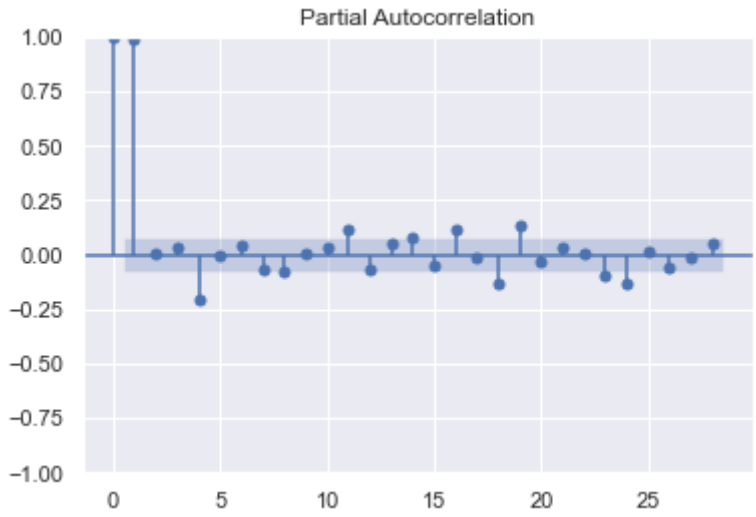
In [5]:
```
plt.figure(figsize=(11, 7))
plot_acf(df["Close"])
plt.show()
```

```
<Figure size 792x504 with 0 Axes>
```



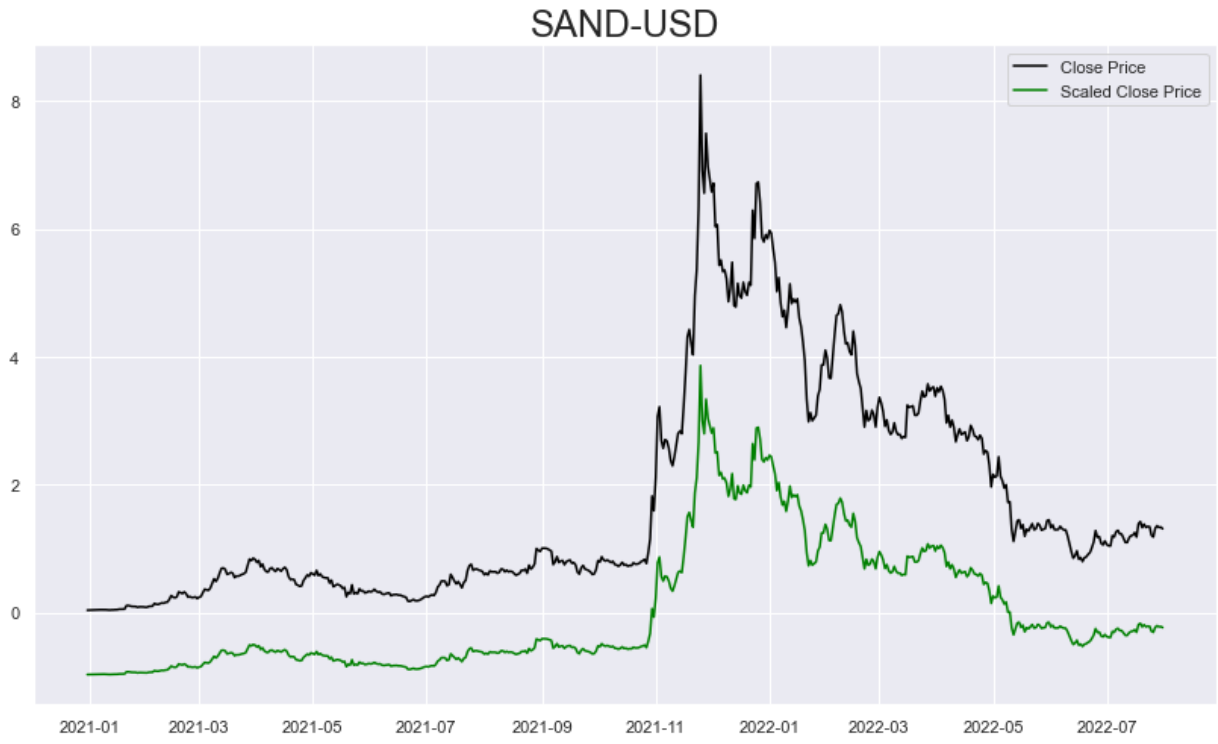
In [6]:
```
plt.figure(figsize=(11, 7))
plot_pacf(df["Close"])
plt.show()
```

```
<Figure size 792x504 with 0 Axes>
```



In [7]:
```
scaler = StandardScaler()
scaled = scaler.fit_transform(df["Close"].values.reshape(-1,1))
df["Scaled Values"] = scaled
df = df[["Close","Scaled Values"]]
```

In [8]:
```
plt.figure(figsize=(14,8))
plt.plot(df["Close"], color="black", label="Close Price")
plt.plot(df["Scaled Values"], color="green", label="Scaled Close Price")
plt.legend()
plt.title(crypto, size=24)
plt.show()
```



In [9]:
```
crypto = "SAND-USD"
lstm_pred = []
lr_pred = []
rf_pred = []
poly = PolynomialFeatures(degree=2)
linearmodel = LinearRegression()
rfmodel = RandomForestRegressor(n_estimators=25, min_samples_split=4, oob_score=
True)
lstm_model = Sequential()
lstm_model.add(LSTM(units=60, input_shape=(2,1)))
lstm_model.add(Dropout(0.1))
lstm_model.add(RepeatVector(1))
lstm_model.add(LSTM(units=30, return_sequences=True))
lstm_model.add(Dense(units=15))
lstm_model.add(Dense(units=7))
lstm_model.add(Dropout(0.1))
lstm_model.add(Dense(units=1))
lstm_model.compile(optimizer=Adam(), loss=MeanSquaredError())
```

In [10]:
```
def start_algorithm(start_date, end_date, iterate=False):
    df = get_data(start_date, end_date)
    x_lr, y_lr = preprocess(df, algo="lr")
    x_rf, y_rf = preprocess(df, algo="rf")
    x_lstm, y_lstm = preprocess(df, algo="lstm")
    if iterate == False:
        pred_lr = fit_lr(x_lr, y_lr)
        pred_rf = fit_rf(x_rf, y_rf)
        pred_lstm = fit_lstm(x_lstm, y_lstm)
    elif iterate == True:
        pred_lr = fit_lr(x_lr, y_lr, refit=True)
        pred_rf = fit_rf(x_rf, y_rf, refit=True)
        pred_lstm = fit_lstm(x_lstm, y_lstm, refit=True)
    token, ans = give_token(y=y_rf[-1], rf=pred_rf, lr=pred_lr, lstm=pred_lstm)
    return token, ans
```

In [11]:
```
def get_data(start_date, end_date):
    return pandas_datareader.DataReader(crypto, 'yahoo', start=start_date, end=e
nd_date)
```

In [12]:
```
def preprocess(df, algo):
    scaler = StandardScaler()
    scaled = scaler.fit_transform(df["Close"].values.reshape(-1,1))
    df["Scaled Values"] = scaled
    df = df[["Close","Scaled Values"]]
    if algo != "lstm":
        df["Lag 1"] = df["Scaled Values"].shift(periods=1)
        df["Lag 2"] = df["Scaled Values"].shift(periods=2)
        df.interpolate(axis=1, inplace=True)
        x1 = df["Lag 1"].to_numpy().reshape(-1,1)
        x2 = df["Lag 2"].to_numpy().reshape(-1,1)
        x = np.concatenate((x1,x2), axis=1)
        if algo == "lr":
            y = df["Scaled Values"].to_numpy().reshape(-1,1)
        elif algo == "rf":
            y = df["Scaled Values"].to_numpy()
    else:
        timestep = 2
        x_temp = []
        y_temp = []
        for i in range(timestep, len(df)):
            x_temp.append(df["Scaled Values"].iloc[i-timestep:i])
            y_temp.append(df["Scaled Values"].iloc[i])
        tempx, y = np.array(x_temp), np.array(y_temp)
        x = tempx.reshape(tempx.shape[0],tempx.shape[1],1)
    return x, y
```

In [13]:
```
def fit_lr(x, y, refit=False, first=False):
    px = poly.fit_transform(x)
    if refit == False:
        linearmodel.fit(px[:-1], y[:-1])
    else:
        linearmodel.fit(px, y)
    ans = linearmodel.predict(np.array([px[-1]]))
    if first == False:
        lr_pred.append(ans[0][0])
    return ans[0][0]
```

In [14]:
```
def fit_rf(x, y, refit=False, first=False):
    if refit == False:
        rfmodel.fit(x[:-1], y[:-1])
    else:
        rfmodel.fit(x, y)
    ans = rfmodel.predict(np.array([x[-1]]))
    if first == False:
        rf_pred.append(ans[0])
    return ans[0]
```

In [15]:
```
def fit_lstm(x, y, refit=False, first=False):
    if refit == False:
        lstm_model.fit(x[:-1], y[:-1], batch_size=12, epochs=18, verbose=0)
    else:
        lstm_model.fit(x, y, batch_size=12, epochs=18, verbose=0)
    ans = lstm_model.predict(np.array([x[-1]]), verbose=0)
    if first == False:
        lstm_pred.append(ans[0][0][0])
    return ans[0][0][0]
```

In [16]:
```
def give_token(y, rf, lr, lstm):
    if abs(y-rf) < abs(y-lr) and abs(y-rf) < abs(y-lstm):
        return "Random Forest", rf
    elif abs(y-lr) < abs(y-rf) and abs(y-lr) < abs(y-lstm):
        return "Polynomial Regression", lr
    elif abs(y-lstm) < abs(y-lr) and abs(y-lstm) < abs(y-rf):
        return "LSTM", lstm
```

In [17]:
```
def make_prediction(token, start_date, end_date):
    df = get_data(start_date, end_date)
    if token == "lr":
        x, y = preprocess(df, algo=token)
        answer = fit_lr(x, y, refit=True, first=True)
        return answer
    elif token == "rf":
        x, y = preprocess(df, algo=token)
        answer = fit_rf(x, y, refit=True, first=True)
        return answer
    elif token == "lstm":
        x, y = preprocess(df, algo="lstm")
        answer = fit_lstm(x, y, refit=True, first=True)
        return answer
```

In [18]:

```
token = ""
start_date = datetime.datetime(2022, 1, 2)
end_date = datetime.datetime(2022, 1, 21)
flag = False
prediction = 0
pred_list = []
days = []
while(end_date != datetime.datetime(2022, 8, 31)):
    if token == "":
        token, _ = start_algorithm(start_date, end_date)
        print(f"Initial token given to {token} for {end_date.date()}")
    else:
        if flag == False:
            start_date += datetime.timedelta(days=1)
            prediction = make_prediction(token, start_date, end_date)
            pred_list.append(prediction)
            days.append(end_date)
            start_date += datetime.timedelta(days=1)
            end_date += datetime.timedelta(days=1)
            token, prediction = start_algorithm(start_date, end_date, iterate=Tr
ue)
            pred_list.append(prediction)
            days.append(end_date)
            flag = True
            print(f"{token} predicts for {end_date.date()}")
        else:
            start_date += datetime.timedelta(days=1)
            end_date += datetime.timedelta(days=1)
            token, prediction = start_algorithm(start_date, end_date, iterate=Tr
ue)
            pred_list.append(prediction)
            days.append(end_date)
            print(f"{token} predicts for {end_date.date()}")
```

Initial token given to Polynomial Regression for 2022-01-21
Polynomial Regression predicts for 2022-01-22
Polynomial Regression predicts for 2022-01-23
Polynomial Regression predicts for 2022-01-24
Random Forest predicts for 2022-01-25
Polynomial Regression predicts for 2022-01-26
Random Forest predicts for 2022-01-27
Polynomial Regression predicts for 2022-01-28
LSTM predicts for 2022-01-29
LSTM predicts for 2022-01-30
Random Forest predicts for 2022-01-31
Random Forest predicts for 2022-02-01
LSTM predicts for 2022-02-02
LSTM predicts for 2022-02-03
Random Forest predicts for 2022-02-04
Polynomial Regression predicts for 2022-02-05
Polynomial Regression predicts for 2022-02-06
Random Forest predicts for 2022-02-07
Random Forest predicts for 2022-02-08
Random Forest predicts for 2022-02-09
Polynomial Regression predicts for 2022-02-10
Polynomial Regression predicts for 2022-02-11
Random Forest predicts for 2022-02-12
Random Forest predicts for 2022-02-13
Random Forest predicts for 2022-02-14
LSTM predicts for 2022-02-15
LSTM predicts for 2022-02-16
Random Forest predicts for 2022-02-17
Polynomial Regression predicts for 2022-02-18
Random Forest predicts for 2022-02-19
Random Forest predicts for 2022-02-20
Polynomial Regression predicts for 2022-02-21
Random Forest predicts for 2022-02-22
Polynomial Regression predicts for 2022-02-23
Polynomial Regression predicts for 2022-02-24
Random Forest predicts for 2022-02-25
Random Forest predicts for 2022-02-26
Random Forest predicts for 2022-02-27
Random Forest predicts for 2022-02-28
LSTM predicts for 2022-03-01
Polynomial Regression predicts for 2022-03-02
Polynomial Regression predicts for 2022-03-03
Random Forest predicts for 2022-03-04
Polynomial Regression predicts for 2022-03-05
Polynomial Regression predicts for 2022-03-06
Polynomial Regression predicts for 2022-03-07
Polynomial Regression predicts for 2022-03-08
Random Forest predicts for 2022-03-09
Random Forest predicts for 2022-03-10
Random Forest predicts for 2022-03-11
Random Forest predicts for 2022-03-12
Polynomial Regression predicts for 2022-03-13
Polynomial Regression predicts for 2022-03-14
Polynomial Regression predicts for 2022-03-15
Random Forest predicts for 2022-03-16
Random Forest predicts for 2022-03-17
Random Forest predicts for 2022-03-18
Random Forest predicts for 2022-03-19
Polynomial Regression predicts for 2022-03-20
Random Forest predicts for 2022-03-21
Random Forest predicts for 2022-03-22
Random Forest predicts for 2022-03-23
Polynomial Regression predicts for 2022-03-24
Random Forest predicts for 2022-03-25
Random Forest predicts for 2022-03-26

Polynomial Regression predicts for 2022-03-27
Random Forest predicts for 2022-03-28
Polynomial Regression predicts for 2022-03-29
Polynomial Regression predicts for 2022-03-30
Random Forest predicts for 2022-03-31
Random Forest predicts for 2022-04-01
Polynomial Regression predicts for 2022-04-02
Random Forest predicts for 2022-04-03
Polynomial Regression predicts for 2022-04-04
LSTM predicts for 2022-04-05
Random Forest predicts for 2022-04-06
Random Forest predicts for 2022-04-07
Polynomial Regression predicts for 2022-04-08
Polynomial Regression predicts for 2022-04-09
Polynomial Regression predicts for 2022-04-10
Polynomial Regression predicts for 2022-04-11
Random Forest predicts for 2022-04-12
Random Forest predicts for 2022-04-13
Random Forest predicts for 2022-04-14
Random Forest predicts for 2022-04-15
Random Forest predicts for 2022-04-16
Random Forest predicts for 2022-04-17
Polynomial Regression predicts for 2022-04-18
LSTM predicts for 2022-04-19
Random Forest predicts for 2022-04-20
Polynomial Regression predicts for 2022-04-21
Random Forest predicts for 2022-04-22
Random Forest predicts for 2022-04-23
Polynomial Regression predicts for 2022-04-24
Random Forest predicts for 2022-04-25
Random Forest predicts for 2022-04-26
Polynomial Regression predicts for 2022-04-27
Polynomial Regression predicts for 2022-04-28
Polynomial Regression predicts for 2022-04-29
Polynomial Regression predicts for 2022-04-30
Random Forest predicts for 2022-05-01
Polynomial Regression predicts for 2022-05-02
Random Forest predicts for 2022-05-03
LSTM predicts for 2022-05-04
Polynomial Regression predicts for 2022-05-05
Polynomial Regression predicts for 2022-05-06
Random Forest predicts for 2022-05-07
Polynomial Regression predicts for 2022-05-08
Random Forest predicts for 2022-05-09
Polynomial Regression predicts for 2022-05-10
Polynomial Regression predicts for 2022-05-11
Polynomial Regression predicts for 2022-05-12
Random Forest predicts for 2022-05-13
Polynomial Regression predicts for 2022-05-14
Polynomial Regression predicts for 2022-05-15
Random Forest predicts for 2022-05-16
Polynomial Regression predicts for 2022-05-17
Random Forest predicts for 2022-05-18
Random Forest predicts for 2022-05-19
Polynomial Regression predicts for 2022-05-20
Polynomial Regression predicts for 2022-05-21
Random Forest predicts for 2022-05-22
Random Forest predicts for 2022-05-23
Random Forest predicts for 2022-05-24
LSTM predicts for 2022-05-25
LSTM predicts for 2022-05-26
Random Forest predicts for 2022-05-27
Random Forest predicts for 2022-05-28
Random Forest predicts for 2022-05-29
Random Forest predicts for 2022-05-30

Random Forest predicts for 2022-05-31
Polynomial Regression predicts for 2022-06-01
Random Forest predicts for 2022-06-02
Random Forest predicts for 2022-06-03
Random Forest predicts for 2022-06-04
Random Forest predicts for 2022-06-05
LSTM predicts for 2022-06-06
Random Forest predicts for 2022-06-07
Polynomial Regression predicts for 2022-06-08
Random Forest predicts for 2022-06-09
Random Forest predicts for 2022-06-10
Polynomial Regression predicts for 2022-06-11
Polynomial Regression predicts for 2022-06-12
Polynomial Regression predicts for 2022-06-13
Polynomial Regression predicts for 2022-06-14
Random Forest predicts for 2022-06-15
Random Forest predicts for 2022-06-16
Random Forest predicts for 2022-06-17
Random Forest predicts for 2022-06-18
Polynomial Regression predicts for 2022-06-19
Random Forest predicts for 2022-06-20
Polynomial Regression predicts for 2022-06-21
Random Forest predicts for 2022-06-22
LSTM predicts for 2022-06-23
LSTM predicts for 2022-06-24
Polynomial Regression predicts for 2022-06-25
Random Forest predicts for 2022-06-26
Random Forest predicts for 2022-06-27
Random Forest predicts for 2022-06-28
Random Forest predicts for 2022-06-29
Random Forest predicts for 2022-06-30
Random Forest predicts for 2022-07-01
Random Forest predicts for 2022-07-02
LSTM predicts for 2022-07-03
Random Forest predicts for 2022-07-04
Random Forest predicts for 2022-07-05
Random Forest predicts for 2022-07-06
Polynomial Regression predicts for 2022-07-07
Random Forest predicts for 2022-07-08
Polynomial Regression predicts for 2022-07-09
LSTM predicts for 2022-07-10
Random Forest predicts for 2022-07-11
Random Forest predicts for 2022-07-12
Random Forest predicts for 2022-07-13
Random Forest predicts for 2022-07-14
Polynomial Regression predicts for 2022-07-15
Random Forest predicts for 2022-07-16
LSTM predicts for 2022-07-17
Random Forest predicts for 2022-07-18
Polynomial Regression predicts for 2022-07-19
Polynomial Regression predicts for 2022-07-20
Polynomial Regression predicts for 2022-07-21
Polynomial Regression predicts for 2022-07-22
Random Forest predicts for 2022-07-23
Random Forest predicts for 2022-07-24
LSTM predicts for 2022-07-25
Random Forest predicts for 2022-07-26
Random Forest predicts for 2022-07-27
Random Forest predicts for 2022-07-28
Random Forest predicts for 2022-07-29
Random Forest predicts for 2022-07-30
Random Forest predicts for 2022-07-31
Random Forest predicts for 2022-08-01
Polynomial Regression predicts for 2022-08-02
Random Forest predicts for 2022-08-03

```
LSTM predicts for 2022-08-04
Random Forest predicts for 2022-08-05
LSTM predicts for 2022-08-06
Polynomial Regression predicts for 2022-08-07
Random Forest predicts for 2022-08-08
LSTM predicts for 2022-08-09
Random Forest predicts for 2022-08-10
Random Forest predicts for 2022-08-11
Random Forest predicts for 2022-08-12
Random Forest predicts for 2022-08-13
Random Forest predicts for 2022-08-14
Random Forest predicts for 2022-08-15
Random Forest predicts for 2022-08-16
Polynomial Regression predicts for 2022-08-17
Polynomial Regression predicts for 2022-08-18
Polynomial Regression predicts for 2022-08-19
Polynomial Regression predicts for 2022-08-20
Random Forest predicts for 2022-08-21
Random Forest predicts for 2022-08-22
Random Forest predicts for 2022-08-23
Random Forest predicts for 2022-08-24
Polynomial Regression predicts for 2022-08-25
Random Forest predicts for 2022-08-26
Polynomial Regression predicts for 2022-08-27
Polynomial Regression predicts for 2022-08-28
Random Forest predicts for 2022-08-29
Polynomial Regression predicts for 2022-08-30
Random Forest predicts for 2022-08-31
```

In [19]:
```
pred_list = np.array(pred_list)
days = np.array(days)
lstm_pred = np.array(lstm_pred)
lr_pred = np.array(lr_pred)
rf_pred = np.array(rf_pred)
```

In [20]:
```
df = get_data(datetime.datetime(2022, 1, 22), datetime.datetime(2022, 8, 31))
x, y = preprocess(df, algo="rf")
```

In [21]:
```
plt.figure(figsize=(14, 8))
plt.plot(days, rf_pred)
plt.plot(days, lr_pred)
plt.plot(days, lstm_pred)
plt.plot(days, y)
plt.xticks(rotation=45)
plt.legend(["Random Forest", "Polynomial Regression", "LSTM", "Actual Price"])
plt.show()
```

In [22]:
```
error_lr = abs(y-lr_pred)
error_rf = abs(y-rf_pred)
error_lstm = abs(y-lstm_pred)
```

In [23]:
```
plt.figure(figsize=(14, 8))
plt.plot(days, error_rf)
plt.plot(days, error_lr)
plt.plot(days, error_lstm)
plt.legend(["Random Forest", "Polynomial Regression", "LSTM"])
plt.show()
```