STOCK PRICE PREDICTION

Certainly, here are the steps to perform the tasks you mentioned with the dataset named "MSFT.csv" and columns "date, open, high, low, close, adj close, volume":

1. Importing the required libraries and reading the dataset:

python

import pandas as pd

data = pd.read\_csv('MSFT.csv')

2. Handling missing data using the `SimpleImputer` from `sklearn.preprocessing`:

python

from sklearn.impute import SimpleImputer

imputer = SimpleImputer(strategy='mean')

data[['open', 'high', 'low', 'close', 'adj close', 'volume']] = imputer.fit\_transform(data[['open', 'high', 'low', 'close', 'adj close', 'volume']])

3. Encoding Categorical Data (if any, based on your dataset):

- If there are categorical columns, you can encode them using one-hot encoding. Assuming there are no categorical columns in your dataset, you can skip this step.

4. Splitting the dataset into training and testing sets:

python

from sklearn.model\_selection import train\_test\_split

X = data[['open', 'high', 'low', 'close', 'adj close', 'volume']]

Y = data['target\_column\_name'] # Replace 'target\_column\_name' with the actual target column name

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size=0.2, random\_state=42)

5. Feature Scaling using `StandardScaler`:

python

from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()

X\_train = scaler.fit\_transform(X\_train)

X\_test = scaler.transform(X\_test)

Now you've completed the steps for handling your dataset 'MSFT.csv' with the specified columns and tasks. Make sure to replace 'target\_column\_name' with the actual name of your target variable if you have one.