import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

from sklearn import preprocessing, svm

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

from sklearn.linear\_model import LinearRegression

df = pd.read\_excel("tomato\_data.xlsx")

prices = []

for i in df["Today price"].tolist():

price = i.split()

prices.append(price[0])

supply = df["Today arrival"].tolist()

supply\_lm = df["Last month arrival"].tolist()

supply\_lw = df["Last week arrival"].tolist()

df = pd.read\_excel("tomato\_data.xlsx", "FEB 2020")

for i in df["Today price"].tolist():

price = i.split()

prices.append(price[0])

supply = supply + df["Today arrival"].tolist()

supply\_lm = supply\_lm + df["Last month arrival"].tolist()

supply\_lw = supply\_lw + df["Last week arrival"].tolist()

df = pd.read\_excel("tomato\_data.xlsx", "MAR 2020")

for i in df["Today price"].tolist():

price = i.split()

prices.append(price[0])

supply = supply + df["Today arrival"].tolist()

supply\_lw = supply\_lw + df["Last week arrival"].tolist()

supply\_lm = supply\_lm + df["Last month arrival"].tolist()

dates = list(range(31+29+31))

new\_supply\_lm = []

for i in range(len(supply\_lm)):

new\_supply\_lm.append(float(supply\_lm[i])\*\*3)

plt.scatter(new\_supply\_lm, prices)

regr = LinearRegression()

X = np.array(new\_supply\_lm).reshape(-1, 1)

Y = np.array(prices).reshape(-1, 1)

regr.fit(X, Y)

print(regr.score(X, Y))

y\_pred = regr.predict(X)

plt.plot(new\_supply\_lm, y\_pred, color ='k')

plt.scatter(new\_supply\_lm, prices, color ='b')