```
import numpy as np
import pandas as pd
from sklearn import linear_model
import matplotlib.pyplot as plt
# np.random.seed(2)

df = pd.read_csv("SATGPA.csv")

# ============== Plot Data ===========

df.plot.scatter('SAT','GPA', color = 'red',
marker='+')
plt.show()
# ====================

train_x = df[['SAT']]
train_y = df[['GPA']]
# print(train_x.shape)
# print(train_x)

# new_df = df.drop('price', axis='columns')
# print(new_df)
model = linear_model.LinearRegression()
model.fit(train_x, train_y)

check = [[2101], [2221], [2331], [2341],
[2001]]
predictedValues = model.predict(check)
print(predictedValues)
4

coef = model.coef_
inter = model.intercept_

#plt.plot(check,predictedValues)
plt.plot(train_x, coef+inter*train_x)
plt.show()
```