

```

import pandas as pd
import numpy as np
data = pd.read_csv('/content/austin_weather (1).csv')
data = data.drop(['Events', 'Date', 'SeaLevelPressureHighInches',
                  'SeaLevelPressureLowInches'], axis = 1)
data = data.replace('T', 0.0)
data = data.replace('-', 0.0)

data.to_csv('austin_final.csv')

import pandas as pd
import numpy as np
import sklearn as sk
from sklearn.linear_model import LinearRegression
import matplotlib.pyplot as plt

data = pd.read_csv("austin_final.csv")
X = data.drop(['PrecipitationSumInches'], axis = 1)
Y = data['PrecipitationSumInches']

Y = Y.values.reshape(-1, 1)

day_index = 798
days = [i for i in range(Y.size)]
clf = LinearRegression()
clf.fit(X, Y)
inp = np.array([[74], [60], [45], [67], [49], [43], [33], [45],
                [57], [29.68], [10], [7], [2], [0], [20], [4], [31]])
inp = inp.reshape(1, -1)
print('The precipitation in inches for the input is:', clf.predict(inp))
print("the precipitation trend graph: ")
plt.scatter(days, Y, color = 'g')
plt.scatter(days[day_index], Y[day_index], color = 'r')
plt.title("Precipitation level")
plt.xlabel("Days")
plt.ylabel("Precipitation in inches")

plt.show()
x_vis = X.filter(['TempAvgF', 'DewPointAvgF', 'HumidityAvgPercent',
                  'SeaLevelPressureAvgInches', 'VisibilityAvgMiles',
                  'WindAvgMPH'], axis = 1)
print("Precipitation vs selected attributes graph: ")

for i in range(x_vis.columns.size):
    plt.subplot(3, 2, i + 1)
    plt.scatter(days, x_vis[x_vis.columns.values[i]][:100]),
                color = 'g')

plt.scatter(days[day_index],

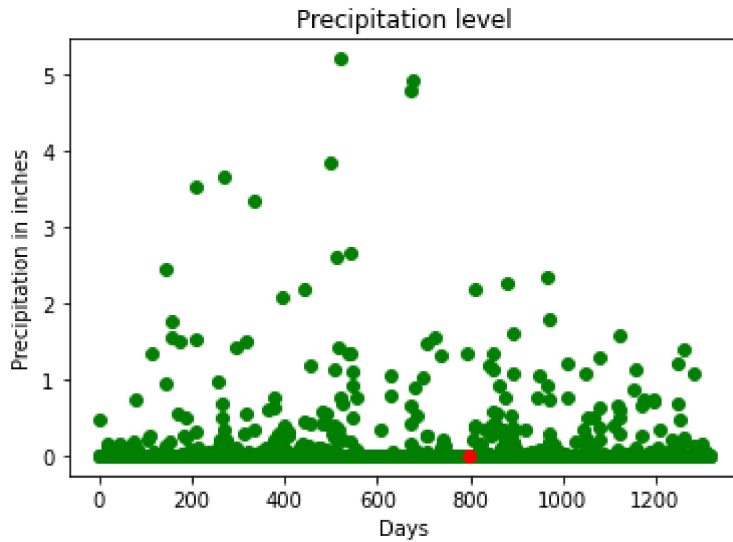
```

```
x_vis[x_vis.columns.values[i]][day_index],
color = 'r')
```

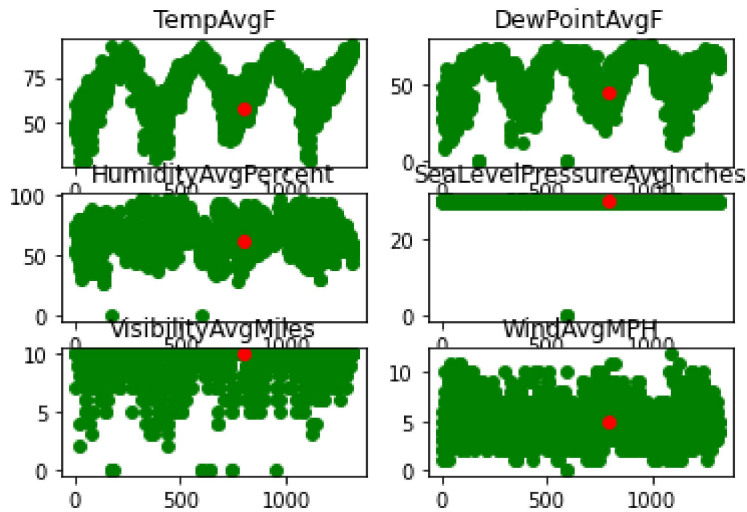
```
plt.title(x_vis.columns.values[i])
```

```
plt.show()
```

The precipitation in inches for the input is: `[[1.33868402]]`
the precipitation trend graph:



Precipitation vs selected attributes graph:



✓ 1s completed at 6:48 PM

● ✕