

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
```

```
data = pd.read_csv('climate_data.csv')
```

```
data
```

	Date	Average_temperature_°F	Average_humidity_%	\
0	NaN	NaN	NaN	
1	NaN	NaN	NaN	
2	NaN	NaN	NaN	
3	NaN	NaN	NaN	
4	NaN	NaN	NaN	
...	...	...	...	
3897	24-07-2020	64.1	62.0	
3898	25-07-2020	62.8	60.0	
3899	26-07-2020	60.6	68.0	
3900	27-07-2020	61.7	64.0	
3901	28-07-2020	60.5	61.0	

	Average_dewpoint_°F	Average_barometer_in	Average_windspeed_mph
\			
0	NaN	NaN	NaN
1	NaN	NaN	NaN
2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	NaN
...	...	...	...
3897	49.8	29.6	3.6
3898	48.1	29.7	2.5
3899	48.9	29.8	1.7
3900	47.4	29.9	2.2
3901	45.3	29.7	4.0

	Average_gustspeed_mph	Average_direction_°deg
Rainfall_for_month_in		
0	NaN	NaN

NaN		
1	NaN	NaN
NaN		
2	NaN	NaN
NaN		
3	NaN	NaN
NaN		
4	NaN	NaN
NaN		
...	...	...
...		
3897	5.8	240.0
0.24		
3898	4.0	242.0
0.33		
3899	2.9	357.0
0.33		
3900	4.0	66.0
0.33		
3901	6.2	248.0
0.35		

	Rainfall_for_year_in	...	Maximum_humidity_%	Minimum_humidity_
% \				
0	NaN	...	NaN	
NaN				
1	NaN	...	NaN	
NaN				
2	NaN	...	NaN	
NaN				
3	NaN	...	NaN	
NaN				
4	NaN	...	NaN	
NaN				
...	...	...	...	..
.				
3897	4.38	...	86.0	
35.0				
3898	4.47	...	90.0	
36.0				
3899	4.47	...	90.0	
40.0				
3900	4.47	...	96.0	
35.0				
3901	4.49	...	94.0	
35.0				

	Maximum_pressure	Minimum_pressure	Maximum_windspeed_mph	\
0	NaN	NaN	NaN	
1	NaN	NaN	NaN	

2	NaN	NaN	NaN
3	NaN	NaN	NaN
4	NaN	NaN	NaN
...	...	...	...
3897	29.686	29.577	15.0
3898	29.781	29.645	8.1
3899	29.930	29.745	11.5
3900	29.941	29.781	13.8
3901	29.792	29.675	17.3

	Maximum_gust_speed_mph	Maximum_heat_index_°F	Date1	Month
\				
0	NaN	NaN	NaN	NaN
1	NaN	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN
3	NaN	NaN	NaN	NaN
4	NaN	NaN	NaN	NaN
...	...	...	...	...
3897	25.3	77.4	24-07-2020	7.0
3898	17.3	77.5	25-07-2020	7.0
3899	15.0	77.5	26-07-2020	7.0
3900	18.4	78.2	27-07-2020	7.0
3901	26.5	77.6	28-07-2020	7.0

	diff_pressure
0	NaN
1	NaN
2	NaN
3	NaN
4	NaN
...	...
3897	0.109
3898	0.136
3899	0.185
3900	0.160
3901	0.117

[3902 rows x 22 columns]

```
data=data.dropna()
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
Int64Index: 3353 entries, 11 to 3901
```

```
Data columns (total 22 columns):
```

#	Column	Non-Null Count	Dtype
0	Date	3353 non-null	object
1	Average_temperature_°F	3353 non-null	float64
2	Average_humidity_%	3353 non-null	float64
3	Average_dewpoint_°F	3353 non-null	float64
4	Average_barometer_in	3353 non-null	float64
5	Average_windspeed_mph	3353 non-null	float64
6	Average_gustspeed_mph	3353 non-null	float64
7	Average_direction_°deg	3353 non-null	float64
8	Rainfall_for_month_in	3353 non-null	float64
9	Rainfall_for_year_in	3353 non-null	float64
10	Maximum_temperature_°F	3353 non-null	float64
11	Minimum_temperature_°F	3353 non-null	float64
12	Maximum_humidity_%	3353 non-null	float64
13	Minimum_humidity_%	3353 non-null	float64
14	Maximum_pressure	3353 non-null	float64
15	Minimum_pressure	3353 non-null	float64
16	Maximum_windspeed_mph	3353 non-null	float64
17	Maximum_gust_speed_mph	3353 non-null	float64
18	Maximum_heat_index_°F	3353 non-null	float64
19	Date1	3353 non-null	object
20	Month	3353 non-null	float64
21	diff_pressure	3353 non-null	float64

```
dtypes: float64(20), object(2)
```

```
memory usage: 602.5+ KB
```

```
inp=data.drop(['Date1','Date','Rainfall_for_year_in'],axis=1)
```

```
inp
```

\	Average_temperature_°F	Average_humidity_%	Average_dewpoint_°F
11	23.4	60.0	11.0
12	32.4	47.0	14.6
13	33.1	45.0	14.2
14	35.7	42.0	14.5
15	41.6	28.0	9.7
...	...	...	...

3897	64.1	62.0	49.8
3898	62.8	60.0	48.1
3899	60.6	68.0	48.9
3900	61.7	64.0	47.4
3901	60.5	61.0	45.3

	Average_barometer_in	Average_windspeed_mph
Average_gustspeed_mph \		
11	30.3	7.4
11.0		
12	30.2	16.3
22.6		
13	30.2	9.9
14.0		
14	30.3	7.9
11.8		
15	30.3	11.9
16.7		
...	...	...
..		
3897	29.6	3.6
5.8		
3898	29.7	2.5
4.0		
3899	29.8	1.7
2.9		
3900	29.9	2.2
4.0		
3901	29.7	4.0
6.2		

	Average_direction_deg	Rainfall_for_month_in
Maximum_temperature_°F \		
11	351.0	0.03
31.0		
12	254.0	0.03
38.0		
13	259.0	0.03
39.0		
14	270.0	0.03
43.0		
15	269.0	0.03
53.0		

...	...	...
3897	240.0	0.24
74.9		
3898	242.0	0.33
69.2		
3899	357.0	0.33
71.9		
3900	66.0	0.33
77.3		
3901	248.0	0.35
75.6		

	Minimum_temperature_°F	Maximum_humidity_%	Minimum_humidity_
% \			
11	16.0	8.0	41.0
12	28.0	5.0	34.0
13	28.0	5.0	34.0
14	27.0	5.0	28.0
15	32.0	4.0	14.0
...	...	...	...
3897	55.3	86.0	35.0
3898	55.1	90.0	36.0
3899	50.5	90.0	40.0
3900	43.6	96.0	35.0
3901	46.0	94.0	35.0

	Maximum_pressure	Minimum_pressure	Maximum_windspeed_mph \
11	30.524	30.012	25.3
12	30.340	30.046	32.2
13	30.271	30.115	29.9
14	30.405	30.201	19.6
15	30.432	30.127	26.5
...	...	...	...
3897	29.686	29.577	15.0
3898	29.781	29.645	8.1
3899	29.930	29.745	11.5
3900	29.941	29.781	13.8

3901	29.792	29.675	17.3
------	--------	--------	------

	Maximum_gust_speed_mph	Maximum_heat_index_°F	Month
diff_pressure			
11	33.0	32.0	1.0
0.512			
12	47.0	38.0	1.0
0.294			
13	32.0	39.0	1.0
0.156			
14	26.0	43.0	1.0
0.204			
15	37.0	53.0	1.0
0.305			
...	...	...	...
...			
3897	25.3	77.4	7.0
0.109			
3898	17.3	77.5	7.0
0.136			
3899	15.0	77.5	7.0
0.185			
3900	18.4	78.2	7.0
0.160			
3901	26.5	77.6	7.0
0.117			

[3353 rows x 19 columns]

```
x=inp.drop(['Rainfall_for_month_in'],axis=1)
y=pd.DataFrame(inp.Rainfall_for_month_in)

from sklearn.model_selection import train_test_split
x_train,x_test,y_train,y_test= train_test_split(x,y,test_size=0.30)
x_train
```

\	Average_temperature_°F	Average_humidity_%	Average_dewpoint_°F
566	62.8	39.0	31.0
3622	33.6	66.0	22.5
1084	57.9	44.0	35.4
1728	39.0	44.0	17.7
2462	58.7	54.0	40.7
...	...	...	...

849	63.0	48.0	40.0
2916	48.3	61.0	33.2
494	48.0	71.0	37.7
2130	52.0	89.0	48.8
2194	57.5	30.0	21.9

	Average_barometer_in	Average_windspeed_mph
Average_gustspeed_mph \		
566	29.9	2.9
4.7		
3622	29.9	8.3
12.6		
1084	29.7	4.5
7.3		
1728	30.3	5.0
8.1		
2462	30.0	2.4
4.4		
...	...	...
..		
849	29.8	3.6
5.7		
2916	29.9	4.8
7.4		
494	30.0	4.5
5.4		
2130	29.8	1.3
2.4		
2194	29.9	2.8
4.9		

	Average_direction_deg	Maximum_temperature_°F
Minimum_temperature_°F \		
566	274.0	83.4
39.6		
3622	269.0	44.3
22.1		
1084	69.0	66.3
44.9		
1728	297.0	50.9
29.4		
2462	112.0	71.6
42.1		



...	...	...
849	82.0	80.1
40.5		
2916	256.0	62.9
35.2		
494	5.0	61.3
38.3		
2130	104.0	59.3
48.3		
2194	309.0	80.3
34.9		

	Maximum_humidity_%	Minimum_humidity_%	Maximum_pressure	\
566	77.0	12.0	30.108	
3622	90.0	39.0	30.207	
1084	68.0	19.0	29.981	
1728	72.0	21.0	30.457	
2462	86.0	31.0	30.112	
...	...	...	...	
849	83.0	13.0	29.967	
2916	94.0	22.0	30.029	
494	94.0	36.0	30.131	
2130	95.0	70.0	29.871	
2194	63.0	9.0	29.984	

	Minimum_pressure	Maximum_windspeed_mph	Maximum_gust_speed_mph
\			
566	29.758	17.3	25.3
3622	29.723	28.8	46.0
1084	29.331	18.4	21.9
1728	30.187	21.9	32.2
2462	29.898	11.5	18.4
...	...	...	...
849	29.750	19.6	24.2
2916	29.626	19.6	26.5
494	29.861	15.0	22.0
2130	29.737	11.5	17.3
2194	29.790	13.8	20.7

	Maximum_heat_index_°F	Month	diff_pressure
566	80.7	8.0	0.350
3622	44.3	10.0	0.484
1084	79.1	6.0	0.650
1728	50.9	5.0	0.270
2462	77.3	6.0	0.214
...	...	...	...
849	79.1	8.0	0.217
2916	62.9	9.0	0.403
494	61.3	6.0	0.270
2130	59.3	7.0	0.134
2194	78.3	9.0	0.194

[2347 rows x 18 columns]

x\_test

\	Average_temperature_°F	Average_humidity_%	Average_dewpoint_°F
13	33.1	45.0	14.2
1130	63.5	49.0	39.7
249	62.1	46.0	32.9
3406	30.3	48.0	9.4
3286	36.1	52.0	20.1
...	...	...	...
3134	52.3	57.0	35.7
1772	65.0	30.0	28.8
1802	56.9	83.0	51.4
2440	52.5	43.0	26.5
2436	45.4	62.0	31.5

	Average_barometer_in	Average_windspeed_mph
Average_gustspeed_mph \		
13	30.2	9.9
14.0		
1130	30.0	2.2

3.9		
249	30.0	1.4
123.8		
3406	30.2	3.3
5.2		
3286	29.7	8.3
12.6		
...	...	...
...		
3134	29.7	3.9
6.6		
1772	29.6	4.7
7.4		
1802	29.9	1.3
2.5		
2440	29.9	5.7
8.9		
2436	29.5	6.1
9.0		

	Average_direction_°deg	Maximum_temperature_°F
Minimum_temperature_°F \		
13	259.0	39.0
28.0		
1130	93.0	82.0
41.2		
249	9.0	77.7
39.8		
3406	283.0	43.6
20.2		
3286	282.0	41.3
29.1		
...	...	...
...		
3134	281.0	63.4
39.6		
1772	250.0	83.9
38.9		
1802	190.0	71.6
50.0		
2440	109.0	66.8
36.6		
2436	70.0	56.7
31.6		

	Maximum_humidity_%	Minimum_humidity_%	Maximum_pressure \
13	5.0	34.0	30.271
1130	87.0	17.0	30.114
249	90.0	15.0	30.256
3406	85.0	14.0	30.285

3286	73.0	43.0	29.994
...	...	...	...
3134	91.0	34.0	29.801
1772	68.0	12.0	29.784
1802	93.0	48.0	29.991
2440	83.0	14.0	30.046
2436	94.0	35.0	29.570

	Minimum_pressure	Maximum_windspeed_mph	Maximum_gust_speed_mph
\			
13	30.115	29.9	32.0
1130	29.791	12.7	17.3
249	29.695	8.1	255.0
3406	29.990	16.1	24.2
3286	29.478	26.5	34.5
...	...	...	...
3134	29.541	16.1	24.2
1772	29.487	20.7	31.1
1802	29.820	10.4	13.8
2440	29.803	18.4	25.3
2436	29.414	19.6	31.1

	Maximum_heat_index_°F	Month	diff_pressure
13	39.0	1.0	0.156
1130	79.9	8.0	0.323
249	78.1	9.0	0.561
3406	43.6	3.0	0.295
3286	41.3	11.0	0.516
...	...	...	...
3134	63.4	5.0	0.260
1772	81.0	6.0	0.297
1802	77.0	7.0	0.171
2440	66.8	5.0	0.243
2436	56.7	5.0	0.156

[1006 rows x 18 columns]

y\_train

```

        Rainfall_for_month_in
566          1.06
3622         0.17
1084         0.27
1728         1.06
2462         0.07
...
849          0.01
2916         0.92
494          1.95
2130         2.30
2194         0.08

```

[2347 rows x 1 columns]

y\_test

```

        Rainfall_for_month_in
13          0.03
1130         0.08
249         0.08
3406         0.69
3286         0.02
...
3134         2.03
1772         0.12
1802         3.08
2440         0.89
2436         0.85

```

[1006 rows x 1 columns]

```
from sklearn.linear_model import LinearRegression
```

```
inp.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
Int64Index: 3353 entries, 11 to 3901
```

```
Data columns (total 19 columns):
```

#	Column	Non-Null Count	Dtype
0	Average_temperature_°F	3353 non-null	float64
1	Average_humidity_%	3353 non-null	float64
2	Average_dewpoint_°F	3353 non-null	float64
3	Average_barometer_in	3353 non-null	float64
4	Average_windspeed_mph	3353 non-null	float64
5	Average_gustspeed_mph	3353 non-null	float64
6	Average_direction_°deg	3353 non-null	float64
7	Rainfall_for_month_in	3353 non-null	float64
8	Maximum_temperature_°F	3353 non-null	float64
9	Minimum_temperature_°F	3353 non-null	float64

```

10 Maximum_humidity_%      3353 non-null    float64
11 Minimum_humidity_%      3353 non-null    float64
12 Maximum_pressure        3353 non-null    float64
13 Minimum_pressure        3353 non-null    float64
14 Maximum_windspeed_mph   3353 non-null    float64
15 Maximum_gust_speed_mph  3353 non-null    float64
16 Maximum_heat_index_°F   3353 non-null    float64
17 Month                   3353 non-null    float64
18 diff_pressure           3353 non-null    float64
dtypes: float64(19)
memory usage: 523.9 KB

climate_data=LinearRegression().fit(x,y)

climate_data.intercept_

array([-6.14415618])

climate_data.coef_

array([[ 2.75068028e-02,  8.10660565e-03,  3.49531224e-03,
         1.79447184e-01,  7.19085422e-03,  4.43567819e-04,
         3.67087815e-04, -1.23087757e-02, -4.25932038e-03,
         2.50915790e-03, -1.56525795e-04, -1.00214005e-02,
         1.03081938e-02, -1.09684545e-03, -4.70216189e-04,
         3.70396413e-03, -9.74058546e-03, -2.03295943e-02]])

predicted=pd.DataFrame(climate_data.predict(x),columns=['prediction'])
predicted

   prediction
0      0.271058
1      0.312402
2      0.271224
3      0.309561
4      0.248662
...
3348    0.972073
3349    1.015922
3350    1.058602
3351    0.944395
3352    0.921656

[3353 rows x 1 columns]

predicted_data=pd.concat([y,x,predicted],axis=1)
predicted_data=predicted_data.dropna()

predicted_data

   Rainfall_for_month_in  Average_temperature_°F  Average_humidity_
% \

```

11	0.03	23.4	
60.0			
12	0.03	32.4	
47.0			
13	0.03	33.1	
45.0			
14	0.03	35.7	
42.0			
15	0.03	41.6	
28.0			
...	...	...	..
.			
3348	0.06	31.6	
42.0			
3349	0.06	40.8	
37.0			
3350	0.06	36.2	
44.0			
3351	0.10	20.4	
42.0			
3352	0.10	23.5	
39.0			

\	Average_dewpoint_°F	Average_barometer_in	Average_windspeed_mph
11	11.0	30.3	7.4
12	14.6	30.2	16.3
13	14.2	30.2	9.9
14	14.5	30.3	7.9
15	9.7	30.3	11.9
...	...	...	...
3348	10.9	30.1	13.6
3349	15.8	29.8	10.8
3350	12.6	29.4	8.6
3351	0.5	30.0	7.7
3352	2.0	30.0	14.4

Average_gustspeed_mph	Average_direction_°deg
-----------------------	------------------------

Maximum_temperature_°F \		
11	11.0	351.0
31.0		
12	22.6	254.0
38.0		
13	14.0	259.0
39.0		
14	11.8	270.0
43.0		
15	16.7	269.0
53.0		
...	...	...
...		
3348	18.4	283.0
38.9		
3349	15.5	283.0
47.6		
3350	12.7	283.0
47.7		
3351	11.6	283.0
28.8		
3352	20.4	283.0
29.0		

% \	Minimum_temperature_°F	Maximum_humidity_%	Minimum_humidity_
11	16.0	8.0	41.0
12	28.0	5.0	34.0
13	28.0	5.0	34.0
14	27.0	5.0	28.0
15	32.0	4.0	14.0
...	...	...	...
3348	24.1	54.0	31.0
3349	35.3	53.0	26.0
3350	19.2	86.0	16.0
3351	16.1	72.0	28.0
3352	16.5	47.0	27.0



	Maximum_pressure	Minimum_pressure	Maximum_windspeed_mph	\
11	30.524	30.012	25.3	
12	30.340	30.046	32.2	
13	30.271	30.115	29.9	
14	30.405	30.201	19.6	
15	30.432	30.127	26.5	
...	...	...	...	
3348	30.245	29.969	26.5	
3349	30.029	29.551	25.3	
3350	29.615	29.225	26.5	
3351	30.207	29.615	27.6	
3352	30.184	29.852	31.1	

	Maximum_gust_speed_mph	Maximum_heat_index_°F	Month
diff_pressure \			
11	33.0	32.0	1.0
0.512			
12	47.0	38.0	1.0
0.294			
13	32.0	39.0	1.0
0.156			
14	26.0	43.0	1.0
0.204			
15	37.0	53.0	1.0
0.305			
...	...	...	...
...			
3348	40.3	38.9	1.0
0.276			
3349	40.3	47.6	1.0
0.478			
3350	31.1	47.7	1.0
0.390			
3351	38.0	28.8	1.0
0.592			
3352	42.6	29.0	1.0
0.332			

	prediction
11	0.321235
12	0.505641
13	0.125537
14	-0.046477
15	-0.021502
...	...
3348	0.972073
3349	1.015922
3350	1.058602
3351	0.944395
3352	0.921656

[2872 rows x 20 columns]