

```

1 C:\Users\adity\PycharmProjects\
  pythonProject\.venv\Scripts\python.exe
  "C:\Users\adity\PycharmProjects\
  pythonProject\MachineLearningProjects\
  DAILY WEATHER DATA ANALYSIS USING
  DECISION TREE CLASSIFICATION.py"
2 Columns are:  Index(['number', '
  air_pressure_9am', 'air_temp_9am', '
  avg_wind_direction_9am',
3      'avg_wind_speed_9am', '
  max_wind_direction_9am', '
  max_wind_speed_9am',
4      'rain_accumulation_9am', '
  rain_duration_9am', '
  relative_humidity_9am',
5      'relative_humidity_3pm'],
6      dtype='object')
7 Data:
8      number  air_pressure_9am  ...
  relative_humidity_9am
  relative_humidity_3pm
9 0          0          918.060000
  ...          42.420000
          36.160000
10 1          1          917.347688
  ...          24.328697
          19.426597
11 2          2          923.040000
  ...          8.900000
          14.460000
12 3          3          920.502751

```

```

12      ...      12.189102
      12.742547
13  4      4      921.160000
      ...      92.410000
      76.740000
14  ...      ...      ...      ...
      ...
      ...
15  1090      1090      918.900000
      ...      26.020000
      38.180000
16  1091      1091      918.710000
      ...      90.350000
      73.340000
17  1092      1092      916.600000
      ...      45.590000
      52.310000
18  1093      1093      912.600000
      ...      64.840000
      58.280000
19  1094      1094      921.530000
      ...      14.560000
      15.100000
20
21 [1095 rows x 11 columns]
22 Null Data:
23      number  air_pressure_9am  ...
      relative_humidity_9am
      relative_humidity_3pm
24  16      16      917.890000
      ...      48.990000

```

24			51.190000
25	111	111	915.290000
	...		21.500000
			29.690000
26	177	177	915.900000
	...		29.260000
			46.500000
27	262	262	923.596607
	...		17.990876
			16.461685
28	277	277	920.480000
	...		52.580000
			54.030000
29	334	334	916.230000
	...		31.880000
			32.900000
30	358	358	917.440000
	...		13.880000
			25.930000
31	361	361	920.444946
	...		12.278715
			7.618649
32	381	381	918.480000
	...		20.640000
			14.350000
33	409	409	NaN
	...		18.487385
			20.356594
34	517	517	920.570000
	...		79.880000
			84.530000

35	519	519	916.250000
	...		72.550000
			74.390000
36	546	546	NaN
	...		87.870000
			70.770000
37	620	620	921.200000
	...		59.790000
			77.750000
38	625	625	912.400000
	...		86.840000
			64.740000
39	656	656	920.830000
	...		23.770000
			51.630000
40	670	670	910.920000
	...		80.560000
			88.220000
41	672	672	922.448945
	...		16.753670
			17.804720
42	705	705	911.900000
	...		77.630000
			59.130000
43	731	731	922.970167
	...		34.807753
			18.418179
44	737	737	917.895130
	...		13.771311
			16.792455
45	788	788	917.923442

45	...		6.939692
			18.793825
46	840	840	918.043767
	...		11.911222
			18.154358
47	848	848	915.250000
	...		91.000000
			90.780000
48	861	861	919.065408
	...		12.497839
			13.438518
49	869	869	NaN
	...		85.270000
			90.260000
50	998	998	914.140000
	...		24.200000
			41.380000
51	1031	1031	922.669195
	...		18.920805
			19.641841
52	1035	1035	919.670000
	...		56.860000
			50.650000
53	1063	1063	917.300185
	...		14.972668
			20.966267
54	1066	1066	919.564869
	...		11.657314
			17.331823
55			
56	[31 rows x 11 columns]		

```

57 1095
58 1064
59 Total rows dropped: 31
60 0      1
61 1      0
62 2      0
63 3      0
64 4      1
65      ..
66 1090    1
67 1091    1
68 1092    1
69 1093    1
70 1094    0
71 Name: high_humidity_label, Length: 1064
   , dtype: int32
72 Y Data:
73     high_humidity_label
74 0                      1
75 1                      0
76 2                      0
77 3                      0
78 4                      1
79 Columns in X: Index(['air_pressure_9am', 'air_temp_9am', '
   avg_wind_direction_9am',
80     'avg_wind_speed_9am', '
   max_wind_direction_9am', '
   max_wind_speed_9am',
81     'rain_accumulation_9am', '
   rain_duration_9am'],

```

```

82         dtype='object')
83 Columns in Y:  Index(['
    high_humidity_label'], dtype='object')
84 X_train is as under:
85         air_pressure_9am  air_temp_9am
    ...  rain_accumulation_9am
    rain_duration_9am
86 841          918.370000      72.932000
    ...
    0.0
87 75          920.100000      53.492000
    ...
    0.0
88 95          927.610000      54.896000
    ...
    0.0
89 895          919.235153      65.951112
    ...
    0.0
90 699          919.888128      68.687822
    ...
    0.0
91
92 [5 rows x 8 columns]
93 X_test is as under:
94         air_pressure_9am  air_temp_9am
    ...  rain_accumulation_9am
    rain_duration_9am
95 456          918.800000      80.384000
    ...
    0.0

```

```

96 845          921.613372      68.658973
    ...                      0.0
                                0.0
97 693          917.900000      76.802000
    ...                      0.0
                                0.0
98 259          914.830000      74.570000
    ...                      0.0
                                0.0
99 723          917.010000      51.836000
    ...                      0.0
                                0.0

100
101 [5 rows x 8 columns]
102 y_train is as under:
103     high_humidity_label
104 841                      0
105 75                       1
106 95                      0
107 895                     0
108 699                      0
109 y_test is as under:
110     high_humidity_label
111 456                      0
112 845                      0
113 693                      1
114 259                      1
115 723                      1
116 Let us describe y_train
117 Sample Predictions:
118 [0 0 1 1 1 1 0 0 0 1]

```



```
119 Sample Y Test(Actual Data):
120 456      0
121 845      0
122 693      1
123 259      1
124 723      1
125 224      1
126 300      1
127 442      0
128 585      1
129 1057     1
130 Name: high_humidity_label, dtype:
    int32
131 Accuracy:
132 0.8153409090909091
133
134 Process finished with exit code 0
135
```