

project-by-jyoti-indore-roll-no-18

September 12, 2024

Jyoti Indore

Artificial Intelligence Branch 2nd year

Roll No. 18

Weather Project Data Visualization

Mount Google Drive

```
[1]: from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

Locate the File in Google Drive

```
[2]: Path = 'drive/My Drive/Dataset 1'
```

Import Dataset

```
[3]: import pandas as pd
import seaborn as sns
sns.set(color_codes=True)
```

```
[4]: weather = pd.read_csv('drive/My Drive/Dataset 1/Test.csv')
```

```
[5]: weather.head()
```

```
[5]:
```

	date_time	is_holiday	air_pollution_index	humidity	wind_speed	\
0	18-05-2017 00:00	NaN	73.0	63.0	1.0	
1	18-05-2017 00:00	NaN	251.0	63.0	1.0	
2	18-05-2017 00:00	NaN	75.0	56.0	1.0	
3	18-05-2017 01:00	NaN	98.0	56.0	1.0	
4	18-05-2017 01:00	NaN	283.0	56.0	1.0	

	wind_direction	visibility_in_miles	dew_point	temperature	rain_p_h	\
0	27.0	4.0	4.0	285.15	0.0	
1	27.0	4.0	4.0	285.15	0.0	
2	0.0	1.0	1.0	285.15	0.0	
3	351.0	2.0	2.0	284.79	0.0	

4	351.0	1.0	1.0	284.79	0.0
---	-------	-----	-----	--------	-----

	snow_p_h	clouds_all	weather_type	weather_description
0	0.0	90.0	Rain	moderate rain
1	0.0	90.0	Mist	mist
2	0.0	90.0	Drizzle	light intensity drizzle
3	0.0	90.0	Rain	heavy intensity rain
4	0.0	90.0	Mist	mist

```
[6]: weather.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 14454 entries, 0 to 14453
Data columns (total 14 columns):
#   Column                Non-Null Count  Dtype
---  -
0   date_time              199 non-null   object
1   is_holiday             0 non-null     float64
2   air_pollution_index   199 non-null   float64
3   humidity               199 non-null   float64
4   wind_speed             199 non-null   float64
5   wind_direction         199 non-null   float64
6   visibility_in_miles    199 non-null   float64
7   dew_point              199 non-null   float64
8   temperature            199 non-null   float64
9   rain_p_h               199 non-null   float64
10  snow_p_h               199 non-null   float64
11  clouds_all             199 non-null   float64
12  weather_type           199 non-null   object
13  weather_description     199 non-null   object
dtypes: float64(11), object(3)
memory usage: 1.5+ MB
```

```
[7]: # check shape of dataset
weather.shape
```

```
[7]: (14454, 14)
```

```
[8]: weather.describe()
```

```
[8]:
```

	is_holiday	air_pollution_index	humidity	wind_speed	\
count	0.0	199.000000	199.000000	199.000000	
mean	NaN	154.130653	75.778894	2.407035	
std	NaN	83.713997	14.921050	1.657597	
min	NaN	10.000000	25.000000	0.000000	
25%	NaN	83.000000	64.000000	1.000000	
50%	NaN	154.000000	75.000000	2.000000	

75%	NaN	231.500000	88.500000	4.000000
max	NaN	298.000000	100.000000	10.000000

	wind_direction	visibility_in_miles	dew_point	temperature	rain_p_h \
count	199.000000	199.000000	199.000000	199.000000	199.0
mean	216.075377	5.236181	5.236181	282.633216	0.0
std	91.932833	2.702318	2.702318	2.311407	0.0
min	0.000000	1.000000	1.000000	279.020000	0.0
25%	164.000000	3.000000	3.000000	281.210000	0.0
50%	205.000000	6.000000	6.000000	282.460000	0.0
75%	302.000000	8.000000	8.000000	283.220000	0.0
max	355.000000	9.000000	9.000000	290.950000	0.0

	snow_p_h	clouds_all
count	199.0	199.000000
mean	0.0	75.507538
std	0.0	29.260217
min	0.0	1.000000
25%	0.0	90.000000
50%	0.0	90.000000
75%	0.0	90.000000
max	0.0	90.000000

```
[9]: weather.describe(include = 'all')
```

```
[9]:
```

	date_time	is_holiday	air_pollution_index	humidity \
count	199	0.0	199.000000	199.000000
unique	125	NaN	NaN	NaN
top	18-05-2017 00:00	NaN	NaN	NaN
freq	3	NaN	NaN	NaN
mean	NaN	NaN	154.130653	75.778894
std	NaN	NaN	83.713997	14.921050
min	NaN	NaN	10.000000	25.000000
25%	NaN	NaN	83.000000	64.000000
50%	NaN	NaN	154.000000	75.000000
75%	NaN	NaN	231.500000	88.500000
max	NaN	NaN	298.000000	100.000000

	wind_speed	wind_direction	visibility_in_miles	dew_point \
count	199.000000	199.000000	199.000000	199.000000
unique	NaN	NaN	NaN	NaN
top	NaN	NaN	NaN	NaN
freq	NaN	NaN	NaN	NaN
mean	2.407035	216.075377	5.236181	5.236181
std	1.657597	91.932833	2.702318	2.702318
min	0.000000	0.000000	1.000000	1.000000
25%	1.000000	164.000000	3.000000	3.000000

50%	2.000000	205.000000		6.000000	6.000000
75%	4.000000	302.000000		8.000000	8.000000
max	10.000000	355.000000		9.000000	9.000000

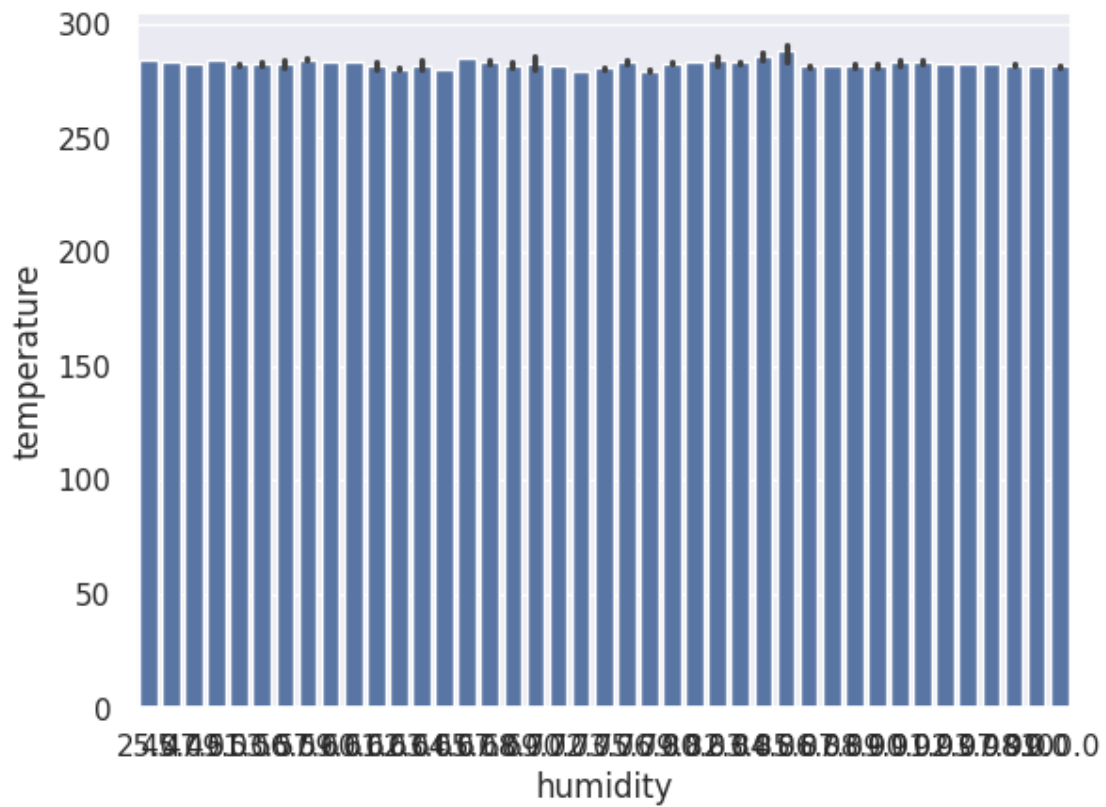
	temperature	rain_p_h	snow_p_h	clouds_all	weather_type \
count	199.000000	199.0	199.0	199.000000	199
unique	NaN	NaN	NaN	NaN	6
top	NaN	NaN	NaN	NaN	Rain
freq	NaN	NaN	NaN	NaN	52
mean	282.633216	0.0	0.0	75.507538	NaN
std	2.311407	0.0	0.0	29.260217	NaN
min	279.020000	0.0	0.0	1.000000	NaN
25%	281.210000	0.0	0.0	90.000000	NaN
50%	282.460000	0.0	0.0	90.000000	NaN
75%	283.220000	0.0	0.0	90.000000	NaN
max	290.950000	0.0	0.0	90.000000	NaN

	weather_description
count	199
unique	12
top	mist
freq	41
mean	NaN
std	NaN
min	NaN
25%	NaN
50%	NaN
75%	NaN
max	NaN

1. Bar Plot

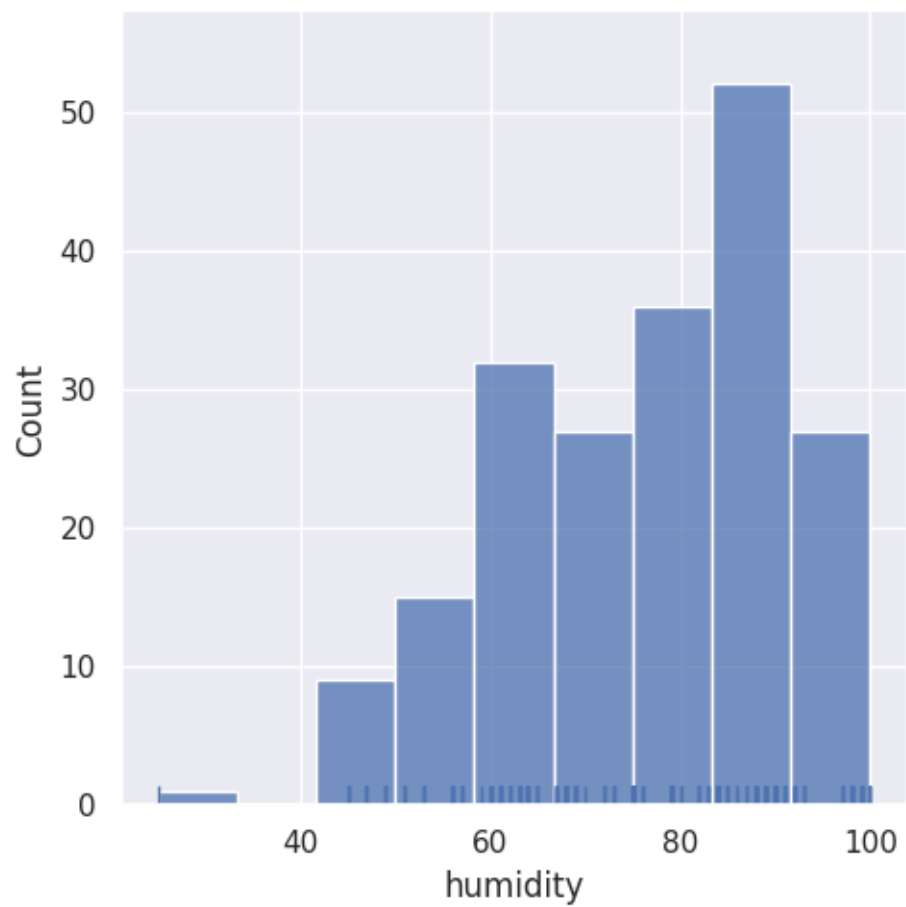
```
[ ]: sns.barplot(x="humidity", y="temperature", data=weather)
```

```
[ ]: <Axes: xlabel='humidity', ylabel='temperature'>
```



2. Dis Plot

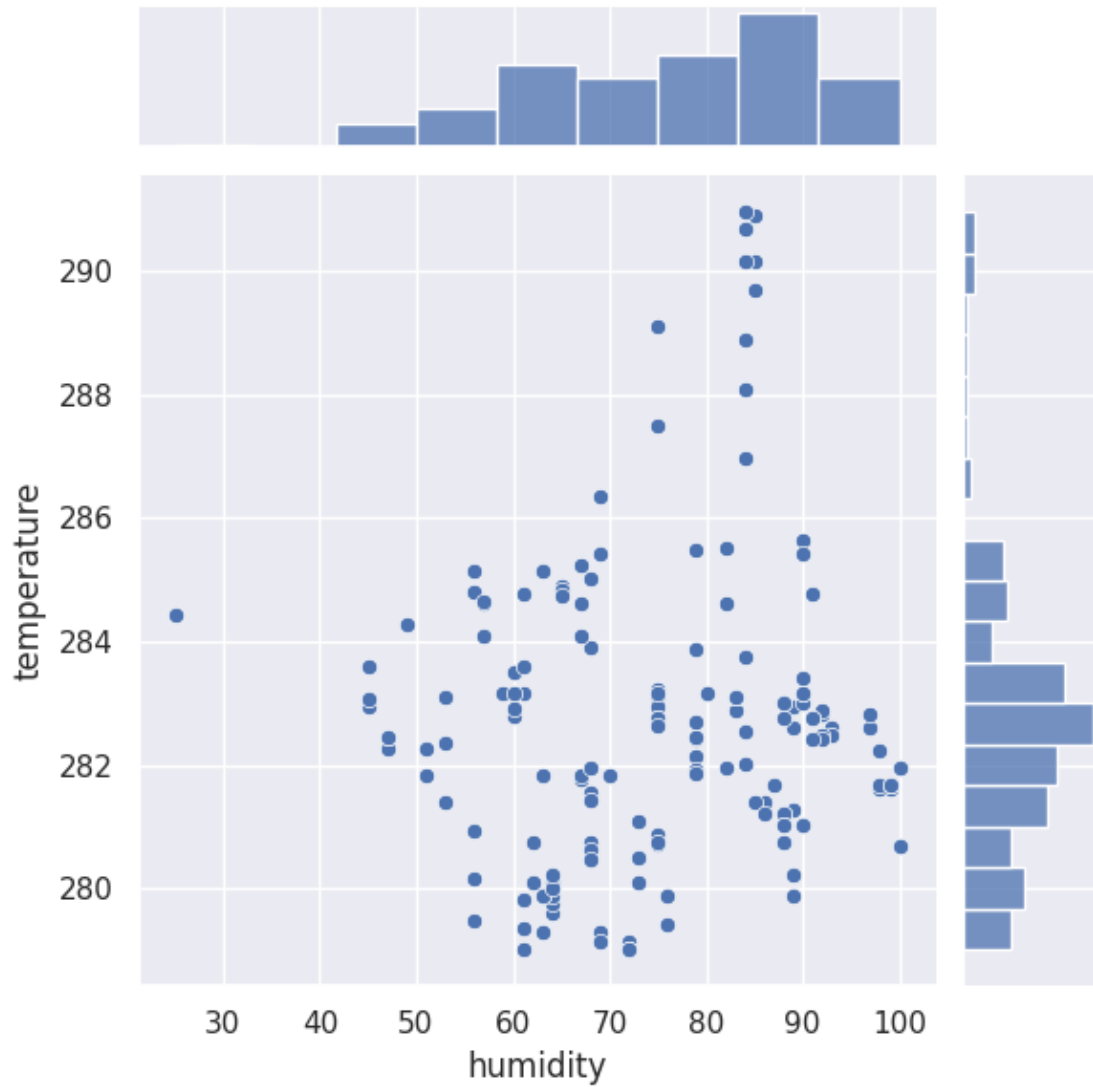
```
[ ]: sns.displot(weather['humidity'], kde=False, rug=True);
```



3. Joint Plot

```
[ ]: sns.jointplot(x="humidity", y="temperature", data=weather)
```

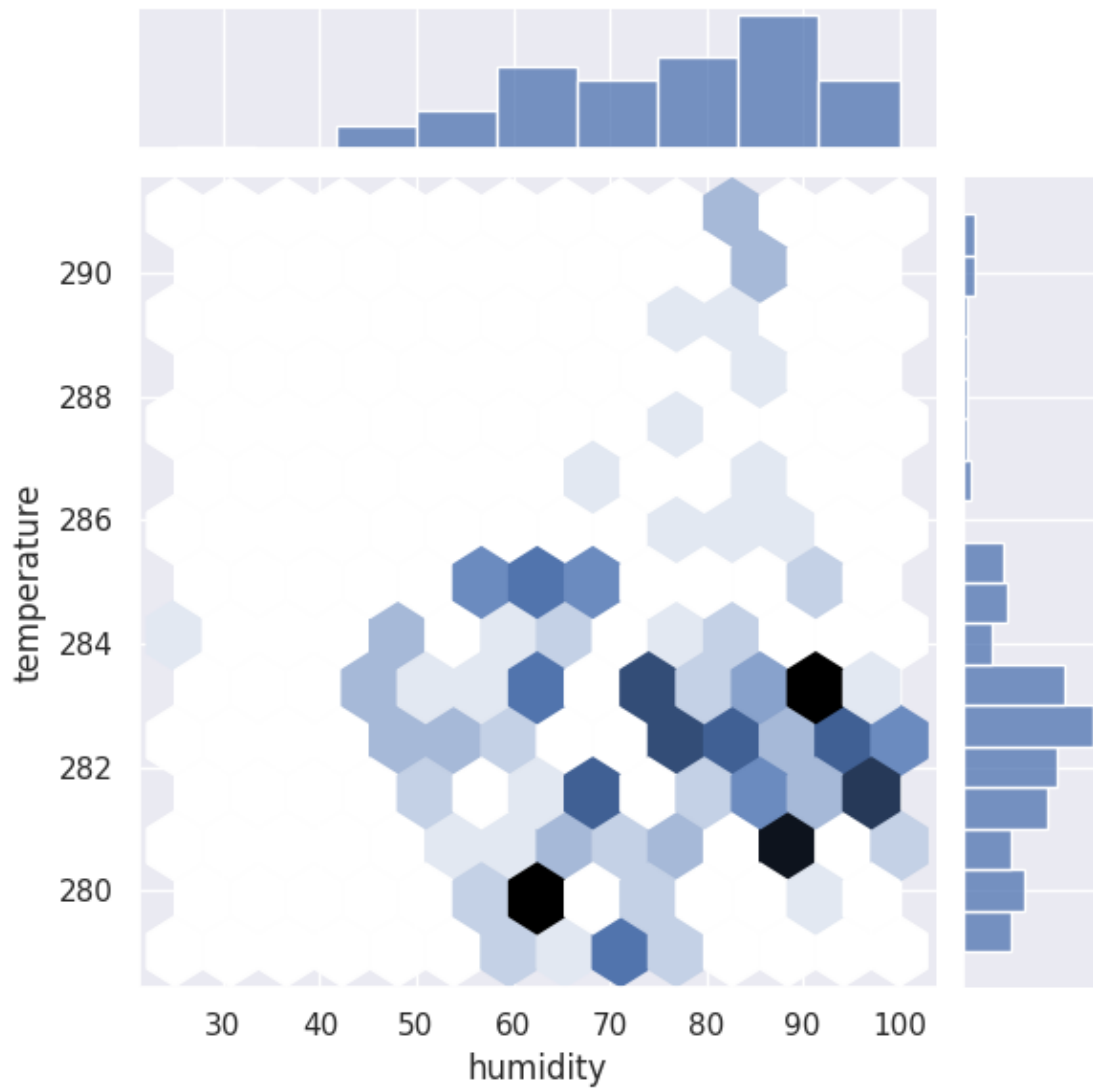
```
[ ]: <seaborn.axisgrid.JointGrid at 0x7b0c199bbbb0>
```



4. Joint Plot with Hexagonal Shape

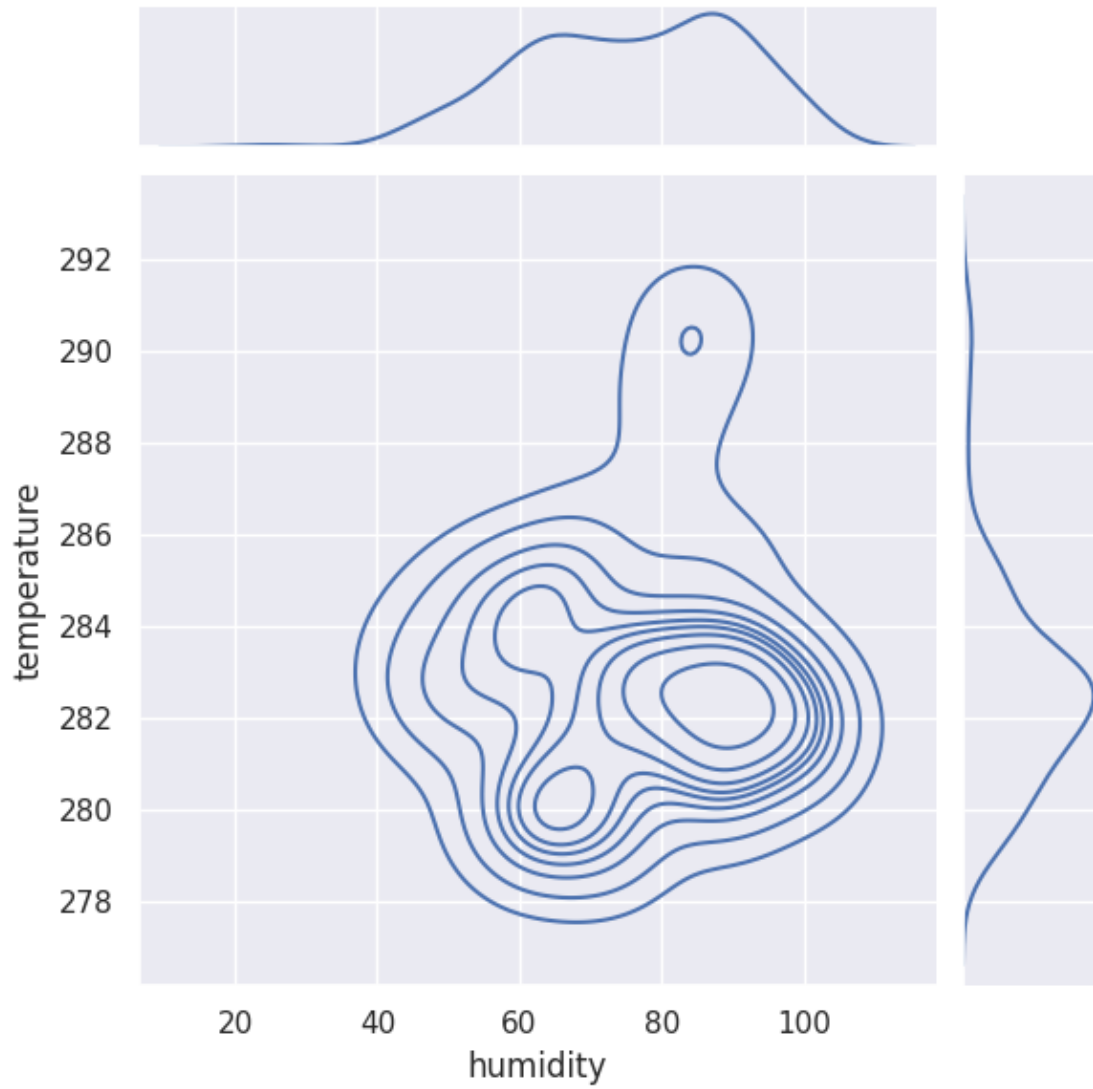
```
[ ]: sns.jointplot(x="humidity", y="temperature", data=weather, kind="hex")
```

```
[ ]: <seaborn.axisgrid.JointGrid at 0x7b0c1907f940>
```



```
[ ]: sns.jointplot(x="humidity", y="temperature", data=weather, kind="kde")
```

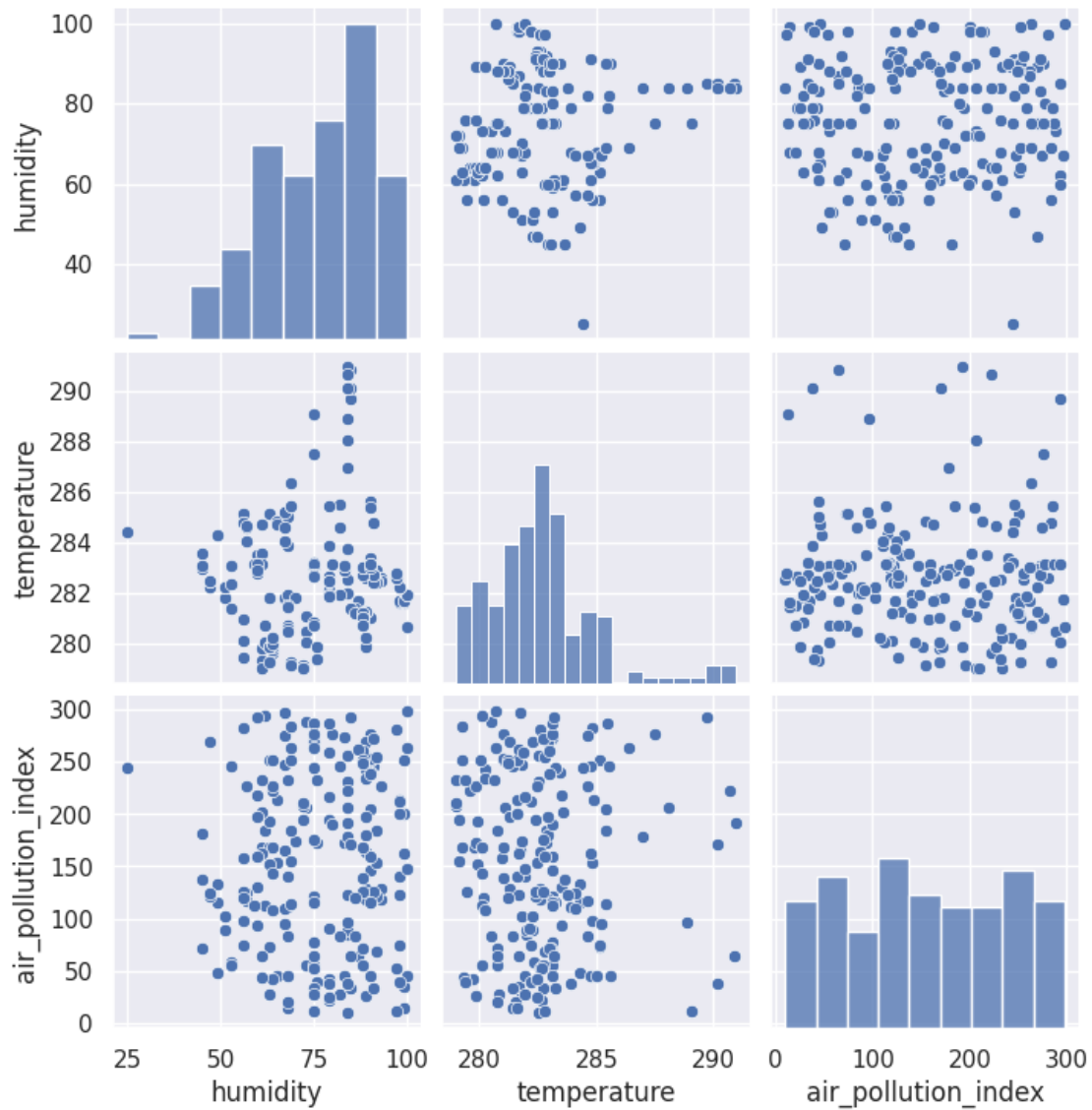
```
[ ]: <seaborn.axisgrid.JointGrid at 0x7b0c18dcead0>
```

5. Pair Plot

```
[ ]: sns.pairplot(weather[['humidity', 'temperature', 'air_pollution_index']])
```

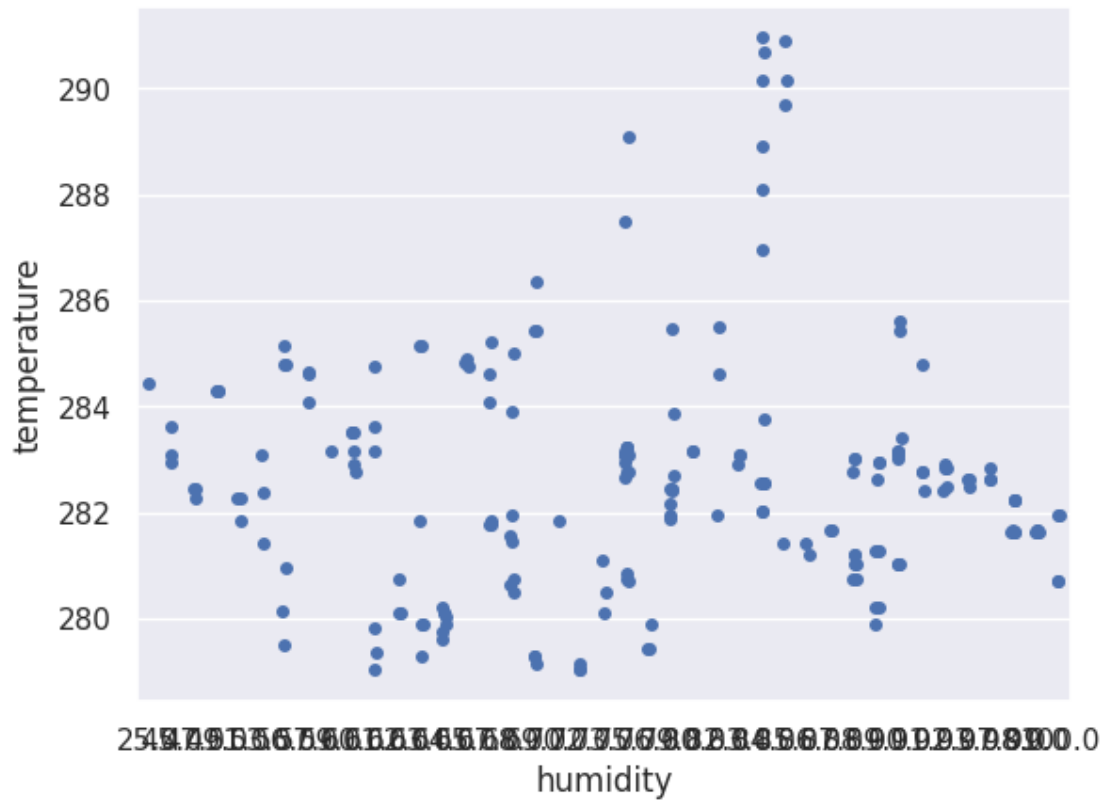
```
[ ]: <seaborn.axisgrid.PairGrid at 0x7b0c18d570a0>
```



6. Strip Plot

```
[ ]: sns.stripplot(x="humidity", y="temperature", data=weather)
```

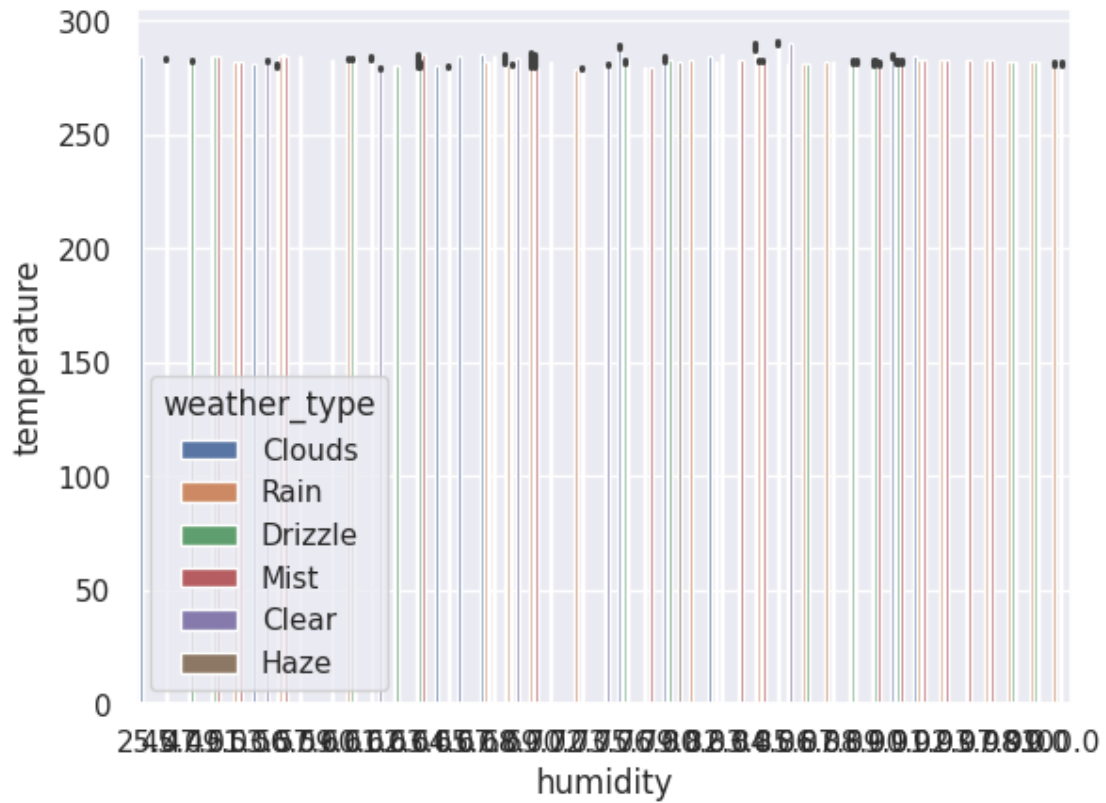
```
[ ]: <Axes: xlabel='humidity', ylabel='temperature'>
```



7.. Bar Plot

```
[ ]: sns.barplot(x="humidity", y="temperature", data=weather, hue=weather['weather_type'])
```

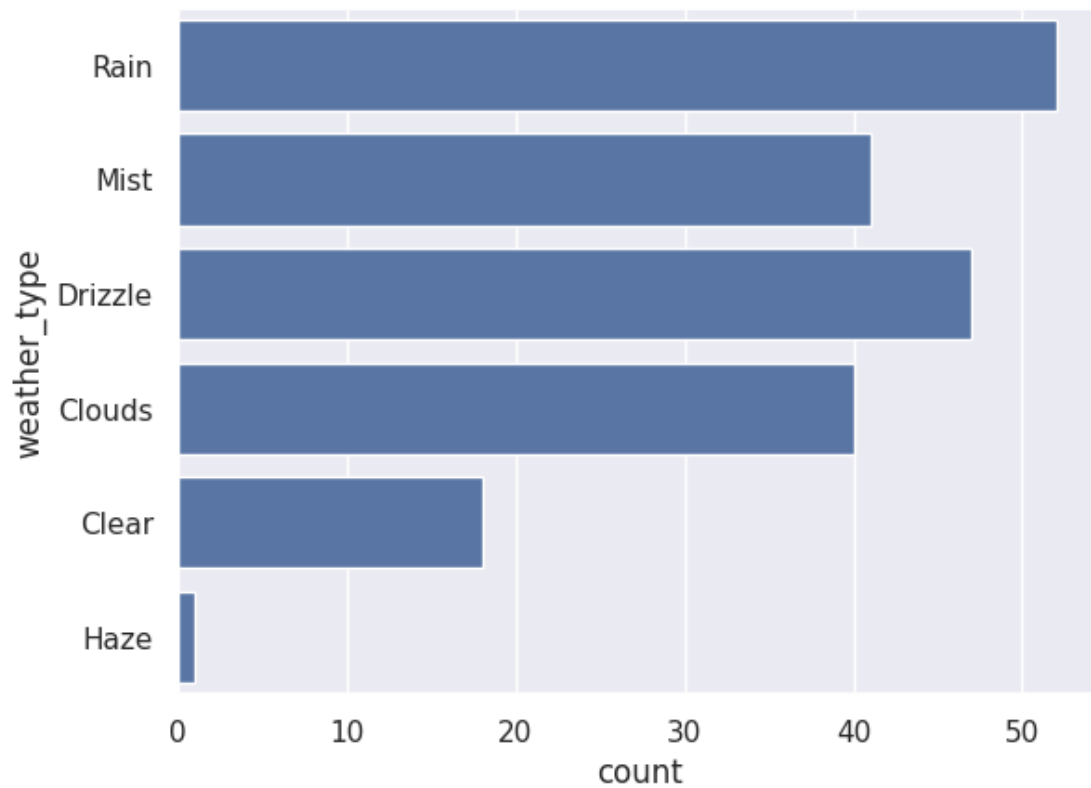
```
[ ]: <Axes: xlabel='humidity', ylabel='temperature'>
```



8. Count Plot

```
[ ]: sns.countplot(weather['weather_type'])
```

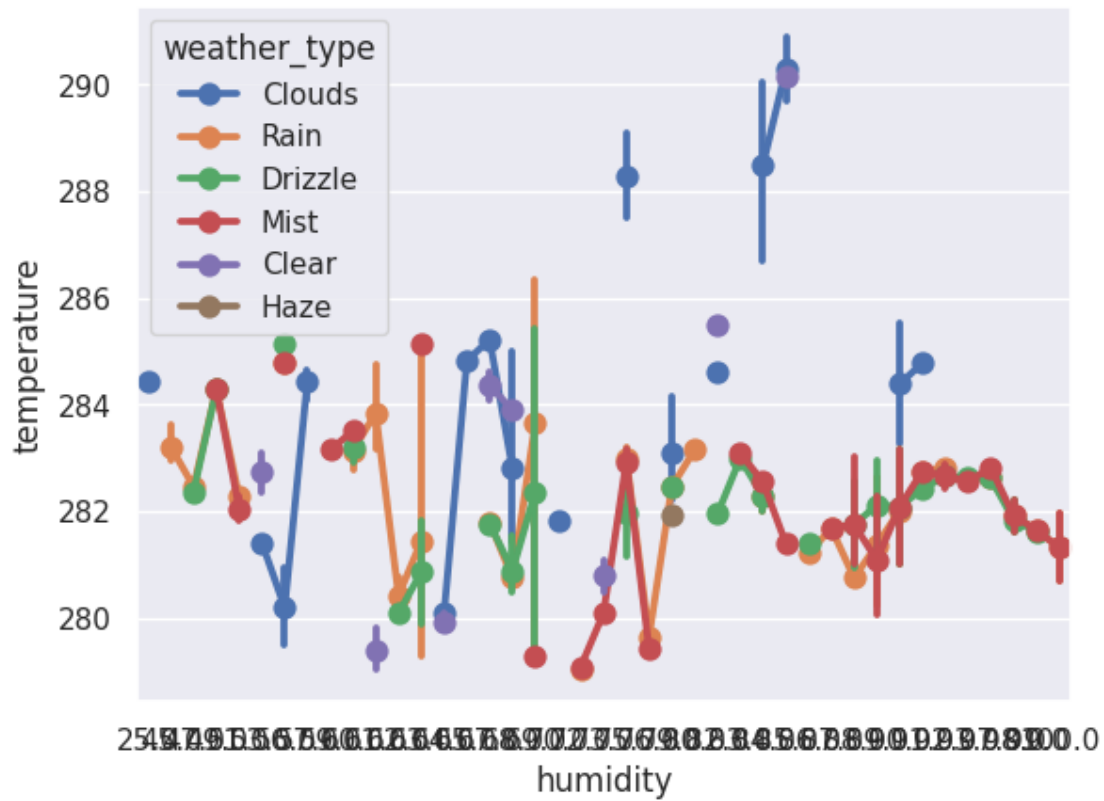
```
[ ]: <Axes: xlabel='count', ylabel='weather_type'>
```



9. Point Plot

```
[ ]: sns.pointplot(x="humidity", y="temperature", data=weather, hue=weather['weather_type'])
```

```
[ ]: <Axes: xlabel='humidity', ylabel='temperature'>
```



10. Lm Plot

```
[ ]: sns.lmplot(x="humidity", y="temperature", hue="weather_type", data=weather)
```

```
[ ]: <seaborn.axisgrid.FacetGrid at 0x7b0c0a2b8970>
```

