Consider the rainfall dataset. This data contains region(district) wise rainfall across India. Perform the following operations for the dataset

In [63]:
```
import numpy as np
import pandas as pd
```

In [64]:
```
import pandas as pd
import numpy as np
df = pd.read_csv('rainfall.csv')
```

In [65]:
```
df.head()
```

Out[65]:

	STATE_UT_NAME	DISTRICT	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb	Mar-May	
0	ANDAMAN And NICOBAR ISLANDS	NICOBAR	107.3	57.9	65.2	117.0	358.5	295.5	285.0	271.9	354.8	326.0	315.2	250.9	2805.2	165.2	540.7	12
1	ANDAMAN And NICOBAR ISLANDS	SOUTH ANDAMAN	43.7	26.0	18.6	90.5	374.4	457.2	421.3	423.1	455.6	301.2	275.8	128.3	3015.7	69.7	483.5	17
2	ANDAMAN And NICOBAR ISLANDS	N & M ANDAMAN	32.7	15.9	8.6	53.4	343.6	503.3	465.4	460.9	454.8	276.1	198.6	100.0	2913.3	48.6	405.6	18
3	ARUNACHAL PRADESH	LOHIT	42.2	80.8	176.4	358.5	306.4	447.0	660.1	427.8	313.6	167.1	34.1	29.8	3043.8	123.0	841.3	18
4	ARUNACHAL PRADESH	EAST SIANG	33.3	79.5	105.9	216.5	323.0	738.3	990.9	711.2	568.0	206.9	29.5	31.7	4034.7	112.8	645.4	30

Check for missing values, if any and drop the corresponding rows.

In [66]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 641 entries, 0 to 640
Data columns (total 19 columns):
 #   Column         Non-Null Count  Dtype  
---  ------         --------------  -----  
 0   STATE_UT_NAME  641 non-null    object 
 1   DISTRICT       641 non-null    object 
 2   JAN            641 non-null    float64
 3   FEB            641 non-null    float64
 4   MAR            641 non-null    float64
 5   APR            641 non-null    float64
 6   MAY            641 non-null    float64
 7   JUN            641 non-null    float64
 8   JUL            641 non-null    float64
 9   AUG            641 non-null    float64
 10  SEP            641 non-null    float64
 11  OCT            641 non-null    float64
 12  NOV            641 non-null    float64
 13  DEC            641 non-null    float64
 14  ANNUAL         641 non-null    float64
 15  Jan-Feb        641 non-null    float64
 16  Mar-May        641 non-null    float64
 17  Jun-Sep        641 non-null    float64
 18  Oct-Dec        641 non-null    float64
dtypes: float64(17), object(2)
memory usage: 95.3+ KB
```

Find the district that gets the highest annual rainfall.

In [67]:
```
sorted_df = df.sort_values(by = 'ANNUAL', ascending=False)
highest = sorted_df.iloc[0,1]
print("District that gets the highest annual rainfall:",highest)
```

```
District that gets the highest annual rainfall: TAMENGLONG
```

Display the top 5 states that get the highest annual rainfall.

In [68]:
```
sorted_df.head(5)
```

Out[68]:

	STATE_UT_NAME	DISTRICT	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL	Jan-Feb
55	MANIPUR	TAMENGLONG	48.5	229.6	224.5	431.5	539.9	1158.7	1820.9	1522.1	726.3	376.1	144.0	7.2	7229.3	278.1
47	MEGHALAYA	JAINTIA HILLS	33.8	44.1	115.1	282.3	598.8	1316.1	1591.3	933.8	826.3	517.7	110.9	9.7	6379.9	77.9
46	MEGHALAYA	EAST KHASI HI	15.4	24.1	129.7	312.5	733.7	1476.2	1518.4	1019.4	607.8	277.9	40.3	10.7	6166.1	39.5
12	ARUNACHAL PRADESH	UPPER SIANG	74.3	176.7	362.6	397.5	408.7	801.9	653.0	417.9	686.0	264.9	86.9	71.7	4402.1	251.0
598	KARNATAKA	UDUPI	1.4	0.4	4.1	29.4	193.8	1081.0	1371.6	902.2	404.9	223.8	74.6	18.8	4306.0	1.8

Drop the columns 'Jan-Feb', 'Mar-May', 'Jun-Sep', 'Oct-Dec'.

In [69]:
```
new_df = df.drop(['Jan-Feb','Mar-May','Jun-Sep','Oct-Dec'], axis=1)
new_df
```

Out[69]:

	STATE_UT_NAME	DISTRICT	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP	OCT	NOV	DEC	ANNUAL
0	ANDAMAN And NICOBAR ISLANDS	NICOBAR	107.3	57.9	65.2	117.0	358.5	295.5	285.0	271.9	354.8	326.0	315.2	250.9	2805.2
1	ANDAMAN And NICOBAR ISLANDS	SOUTH ANDAMAN	43.7	26.0	18.6	90.5	374.4	457.2	421.3	423.1	455.6	301.2	275.8	128.3	3015.7
2	ANDAMAN And NICOBAR ISLANDS	N & M ANDAMAN	32.7	15.9	8.6	53.4	343.6	503.3	465.4	460.9	454.8	276.1	198.6	100.0	2913.3
3	ARUNACHAL PRADESH	LOHIT	42.2	80.8	176.4	358.5	306.4	447.0	660.1	427.8	313.6	167.1	34.1	29.8	3043.8
4	ARUNACHAL PRADESH	EAST SIANG	33.3	79.5	105.9	216.5	323.0	738.3	990.9	711.2	568.0	206.9	29.5	31.7	4034.7
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
636	KERALA	IDUKKI	13.4	22.1	43.6	150.4	232.6	651.6	788.9	527.3	308.4	343.2	172.9	48.1	3302.5
637	KERALA	KASARGOD	2.3	1.0	8.4	46.9	217.6	999.6	1108.5	636.3	263.1	234.9	84.6	18.4	3621.6
638	KERALA	PATHANAMTHITTA	19.8	45.2	73.9	184.9	294.7	556.9	539.9	352.7	266.2	359.4	213.5	51.3	2958.4
639	KERALA	WAYANAD	4.8	8.3	17.5	83.3	174.6	698.1	1110.4	592.9	230.7	213.1	93.6	25.8	3253.1
640	LAKSHADWEEP	LAKSHADWEEP	20.8	14.7	11.8	48.9	171.7	330.2	287.7	217.5	163.1	157.1	117.7	58.8	1600.0

641 rows × 15 columns

Display the state-wise mean rainfall for all the months using a pivot table

In [70]:
```
new_df = new_df.drop(['ANNUAL'],axis=1)
table = pd.pivot_table(new_df,index=['STATE_UT_NAME'])
table
```

Out[70]:

STATE_UT_NAME	APR	AUG	DEC	FEB	JAN	JUL	JUN	MAR	MAY	NOV	
ANDAMAN And NICOBAR ISLANDS	86.966667	385.300000	159.733333	33.266667	61.233333	390.566667	418.666667	30.800000	358.833333	263.200000	3
ANDHRA PRADESH	19.873913	179.426087	15.565217	7.352174	6.321739	185.365217	114.369565	10.095652	48.765217	58.965217	1
ARUNACHAL PRADESH	275.162500	378.600000	35.956250	93.293750	53.687500	547.581250	491.381250	165.018750	300.262500	43.187500	1
ASSAM	181.266667	377.370370	11.440741	31.714815	15.733333	494.844444	465.185185	77.762963	333.870370	24.922222	1
BIHAR	16.865789	289.481579	5.786842	9.278947	13.134211	340.836842	168.781579	9.873684	51.673684	6.715789	
CHANDIGARH	14.800000	287.500000	23.400000	38.900000	44.300000	282.400000	120.000000	33.200000	30.100000	9.900000	
CHATISGARH	13.116667	375.338889	5.811111	10.472222	10.377778	375.405556	180.583333	12.977778	17.483333	8.494444	
DADAR NAGAR HAVELI	0.000000	655.900000	0.000000	0.300000	0.400000	884.500000	385.100000	0.000000	7.400000	10.500000	
DAMAN AND DUI	0.100000	394.600000	0.450000	0.500000	0.550000	583.100000	276.500000	0.200000	4.150000	12.400000	
DELHI	8.900000	245.500000	8.600000	16.300000	16.400000	220.700000	59.800000	15.300000	19.300000	5.600000	
GOA	7.800000	683.800000	10.200000	0.050000	0.550000	1108.100000	908.100000	0.550000	87.750000	35.000000	1
GUJARAT	0.507692	257.630769	1.592308	0.392308	0.784615	333.838462	139.246154	1.142308	4.803846	10.826923	
HARYANA	7.619048	190.909524	7.914286	16.457143	19.485714	180.361905	51.009524	13.738095	14.642857	5.266667	
HIMACHAL	47.683333	322.325000	38.225000	80.450000	81.925000	343.825000	108.683333	87.633333	54.358333	16.908333	
JAMMU AND KASHMIR	82.268182	167.918182	46.395455	91.645455	77.977273	172.090909	53.604545	119.986364	65.136364	27.159091	
JHARKHAND	18.662500	310.316667	6.704167	16.320833	15.837500	333.854167	198.775000	16.516667	45.875000	10.212500	
KARNATAKA	36.773333	209.256667	11.170000	2.696667	2.026667	280.700000	204.880000	7.163333	88.166667	44.350000	1
KERALA	109.021429	417.950000	38.242857	16.200000	9.542857	724.328571	658.707143	31.071429	244.728571	151.535714	2
LAKSHADWEEP	48.900000	217.500000	58.800000	14.700000	20.800000	287.700000	330.200000	11.800000	171.700000	117.700000	1
MADHYA PRADESH	3.270000	331.048000	8.790000	9.158000	12.892000	311.088000	114.686000	7.486000	7.006000	10.042000	
MAHARASHTRA	6.974286	314.585714	7.417143	3.474286	4.791429	388.894286	240.980000	5.997143	19.925714	18.588571	
MANIPUR	150.766667	451.800000	11.788889	55.122222	22.600000	498.055556	487.088889	82.411111	213.377778	56.000000	1
MEGHALAYA	211.228571	584.371429	11.042857	21.685714	14.900000	857.742857	757.228571	74.757143	430.042857	39.571429	2
MIZORAM	152.600000	440.588889	15.288889	29.944444	11.566667	452.311111	429.833333	96.255556	321.322222	64.633333	2
NAGALAND	134.227273	350.872727	10.354545	27.672727	18.481818	395.036364	340.318182	63.018182	213.381818	38.554545	1
ORISSA	36.653333	363.346667	5.136667	22.370000	10.810000	332.316667	212.516667	27.453333	70.723333	30.400000	1
PONDICHERRY	12.275000	116.425000	227.350000	25.425000	26.750000	78.025000	47.675000	16.725000	40.825000	395.150000	2
PUNJAB	12.160000	172.415000	13.905000	24.480000	25.965000	190.610000	46.325000	25.900000	16.165000	6.085000	
RAJASTHAN	3.303030	194.554545	3.021212	4.721212	5.348485	195.278788	54.096970	3.815152	10.627273	6.254545	
SIKKIM	206.900000	434.600000	20.900000	77.300000	47.550000	499.200000	483.800000	130.600000	323.550000	30.950000	2
TAMIL NADU	42.596875	91.571875	96.487500	14.021875	18.906250	72.606250	50.321875	18.068750	67.531250	184.625000	1
TRIPURA	220.750000	356.475000	11.125000	33.650000	11.225000	414.975000	465.425000	93.625000	391.575000	43.300000	1
UTTAR PRADESH	5.318310	291.232394	6.870423	13.157746	17.183099	280.067606	90.770423	10.107042	15.561972	4.576056	
UTTARANCHAL	29.815385	426.784615	20.830769	49.592308	49.892308	432.792308	165.715385	51.669231	58.392308	9.238462	
WEST BENGAL	56.647368	361.573684	7.363158	19.084211	15.031579	412.989474	308.531579	27.973684	139.489474	19.389474	1

Display the count of districts in each state.

In [71]:
```
df.groupby(['STATE_UT_NAME']).count()['DISTRICT']
```

Out[71]:
```
STATE_UT_NAME
ANDAMAN And NICOBAR ISLANDS     3
ANDHRA PRADESH                 23
ARUNACHAL PRADESH              16
ASSAM                          27
BIHAR                          38
CHANDIGARH                      1
CHATISGARH                     18
DADAR NAGAR HAVELI              1
DAMAN AND DUI                   2
DELHI                           9
GOA                             2
GUJARAT                        26
HARYANA                        21
HIMACHAL                       12
JAMMU AND KASHMIR              22
JHARKHAND                      24
KARNATAKA                      30
KERALA                         14
LAKSHADWEEP                     1
MADHYA PRADESH                 50
MAHARASHTRA                    35
MANIPUR                         9
MEGHALAYA                       7
MIZORAM                         9
NAGALAND                       11
ORISSA                         30
PONDICHERRY                     4
PUNJAB                         20
RAJASTHAN                      33
SIKKIM                          4
TAMIL NADU                     32
TRIPURA                         4
UTTAR PRADESH                  71
UTTARANCHAL                    13
WEST BENGAL                    19
Name: DISTRICT, dtype: int64
```

For each state, display the district that gets the highest rainfall in May. Also display the recorded rainfall.

In [72]:
```
#df.groupby(['STATE_UT_NAME'])['DISTRICT'].max()
pivot = pd.pivot_table(data=df,index='STATE_UT_NAME',values=['DISTRICT','MAY'],aggfunc=['max','sum'])
print(pivot)
```

	max		sum
	DISTRICT	MAY	MAY
STATE_UT_NAME			
ANDAMAN And NICOBAR ISLANDS	SOUTH ANDAMAN	374.4	1076.5
ANDHRA PRADESH	WEST GODAVARI	96.6	1121.6
ARUNACHAL PRADESH	WEST SIANG	453.0	4804.2
ASSAM	UDALGURI(DARA	604.0	9014.5
BIHAR	WEST CHAMPARAN	155.7	1963.6
CHANDIGARH	CHANDIGARH	30.1	30.1
CHATISGARH	SURGUJA	38.6	314.7
DADAR NAGAR HAVELI	DNH	7.4	7.4
DAMAN AND DUI	DIU	7.4	8.3
DELHI	WEST DELHI	19.3	173.7
GOA	SOUTH GOA	94.3	175.5
GUJARAT	VALSAD	12.5	124.9
HARYANA	YAMUNANAGAR	27.9	307.5
HIMACHAL	UNA	91.7	652.3
JAMMU AND KASHMIR	UDHAMPUR	111.4	1433.0
JHARKHAND	WEST SINGHBHUM	86.1	1101.0
KARNATAKA	YADGIR	193.8	2645.0
KERALA	WAYANAD	300.4	3426.2
LAKSHADWEEP	LAKSHADWEEP	171.7	171.7
MADHYA PRADESH	VIDISHA	19.9	350.3
MAHARASHTRA	YAVATMAL	60.2	697.4
MANIPUR	UKHRUL	539.9	1920.4
MEGHALAYA	WEST GARO HIL	733.7	3010.3
MIZORAM	SERCHHIP	351.4	2891.9
NAGALAND	ZUNHEBOTO	325.6	2347.2
ORISSA	SUNDARGARH	136.8	2121.7
PONDICHERRY	YANAM	43.6	163.3
PUNJAB	TARN TARAN	25.6	323.3
RAJASTHAN	UDAIPUR	20.7	350.7
SIKKIM	WEST SIKKIM	355.4	1294.2
TAMIL NADU	VIRUDHUNAGAR	141.5	2161.0
TRIPURA	WEST TRIPURA	440.1	1566.3
UTTAR PRADESH	VARANASI	38.6	1104.9
UTTARANCHAL	UTTARKASHI	102.1	759.1
WEST BENGAL	WEST MIDNAPOR	345.4	2650.3