EDS Assignment no. 2

Team Members:- Mohit Muley (276), Manas Kale (275), Chinmay Parite(278), Darshan Thatsingar (264)

Div:- B

Batch:- B4

Assignment on rainfall dataset with 20 problem statements.

```
import pandas as pd

# Load the CSV file into a DataFrame
data = pd.read_csv('/content/rainfall.csv')

# 1.Calculate the average monthly rainfall for each month.
data['MONTHLY_AVG'] = data.iloc[:, 3:15].mean(axis=0)
print("Average monthly rainfall for each month:")
print(data[['DISTRICT', 'JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN',
'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC']])

# 2.Identify the month with the highest rainfall and provide the
corresponding measurement.
max_rainfall_month = data.iloc[:, 3:15].max().idxmax()
max_rainfall_measurement = data[max_rainfall_month].max()
print("Month with the highest rainfall:", max_rainfall_month)
print("Highest rainfall measurement:", max_rainfall_measurement)

# 3.Determine the month with the lowest rainfall and provide the
corresponding measurement.
min_rainfall_month = data.iloc[:, 3:15].min().idxmin()
min_rainfall_measurement = data[min_rainfall_month].min()
print("Month with the lowest rainfall:", min_rainfall_month)
print("Lowest rainfall measurement:", min_rainfall_measurement)

# 4.Calculate the total rainfall for the year.
total_rainfall_year = data.iloc[:, 3:15].sum().sum()
print("Total rainfall for the year:", total_rainfall_year)

# 5.Identify the months with above-average rainfall.
above_avg_rainfall_months = data.iloc[:, 3:15].columns[data.iloc[:,
3:15].mean() > data['ANNUAL'].mean()]
```

```
print("Months with above-average rainfall:", above_avg_rainfall_months)

# 6.Identify the months with below-average rainfall.
below_avg_rainfall_months = data.iloc[:, 3:15].columns[data.iloc[:,
3:15].mean() < data['ANNUAL'].mean()]
print("Months with below-average rainfall:", below_avg_rainfall_months)

# 7.Calculate the average quarterly rainfall.
data['Q1'] = data[['JAN', 'FEB', 'MAR']].mean(axis=1)
data['Q2'] = data[['APR', 'MAY', 'JUN']].mean(axis=1)
data['Q3'] = data[['JUL', 'AUG', 'SEP']].mean(axis=1)
data['Q4'] = data[['OCT', 'NOV', 'DEC']].mean(axis=1)
avg_quarterly_rainfall = data[['DISTRICT', 'Q1', 'Q2', 'Q3', 'Q4']]
print("Average quarterly rainfall:")
print(avg_quarterly_rainfall)

# 8.Determine the wettest quarter of the year and provide the total
rainfall for that quarter.
wettest_quarter = avg_quarterly_rainfall.iloc[:,
1:].sum(axis=1).idxmax()
wettest_quarter_rainfall = avg_quarterly_rainfall.iloc[:,
1:].sum(axis=1).max()
print("Wettest quarter:", wettest_quarter)
print("Total rainfall for the wettest quarter:",
wettest_quarter_rainfall)

# 9.Determine the driest quarter of the year and provide the total
rainfall for that quarter.
driest_quarter = avg_quarterly_rainfall.iloc[:,
1:].sum(axis=1).idxmin()
driest_quarter_rainfall = avg_quarterly_rainfall.iloc[:,
1:].sum(axis=1).min()
print("Driest quarter:", driest_quarter)
print("Total rainfall for the driest quarter:",
driest_quarter_rainfall)

# 10.Calculate the rainfall deviation for each month from the annual
average.
data['RAINFALL_DEVIATION'] = data.iloc[:, 3:15].sub(data['ANNUAL'],
axis=0)

# 11.Identify any months with significant rainfall deviations (above or
below average) and provide the deviation value.
significant_deviation_months = data[['DISTRICT'] +
list(data.columns[3:15])].melt(id_vars=['DISTRICT'], var_name='MONTH',
value_name='RAINFALL')
significant_deviation_months['DEVIATION'] =
significant_deviation_months['RAINFALL'] - data['ANNUAL']
```

```
significant_deviation_months =
significant_deviation_months[(significant_deviation_months['DEVIATION']
> 100) | (significant_deviation_months['DEVIATION'] < -100)]
print("Months with significant rainfall deviations:")
print(significant_deviation_months)


# 12.Calculate the cumulative rainfall for each quarter.
data['CUMULATIVE_Q1'] = data[['JAN', 'FEB', 'MAR']].sum(axis=1)
data['CUMULATIVE_Q2'] = data[['JAN', 'FEB', 'MAR', 'APR', 'MAY',
'JUN']].sum(axis=1)
data['CUMULATIVE_Q3'] = data[['JAN', 'FEB', 'MAR', 'APR', 'MAY', 'JUN',
'JUL', 'AUG', 'SEP']].sum(axis=1)
data['CUMULATIVE_Q4'] = data.iloc[:, 3:15].sum(axis=1)

# 13.Identify any months with a significant increase or decrease in
rainfall compared to the previous month.
significant_change_months = data[['DISTRICT', 'JAN', 'FEB', 'MAR',
'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV',
'DEC']].diff(axis=1).abs().gt(100)
print("Months with significant increase or decrease in rainfall:")
print(significant_change_months)

# 14.Calculate the rainfall variance for each month.
data['VARIANCE'] = data.iloc[:, 3:15].var(axis=0)

# 15.Determine the month with the highest variance in rainfall and
provide the variance value.
highest_variance_month = data['VARIANCE'].idxmax()
highest_variance = data['VARIANCE'].max()
print("Month with the highest variance in rainfall:",
highest_variance_month)
print("Variance value:", highest_variance)

# 16.Calculate the monthly rainfall trend (increase or decrease) over
the year.
monthly_trend = data.iloc[:, 3:15].diff(axis=1).apply(lambda x:
'Increase' if x > 0 else 'Decrease')
print("Monthly rainfall trend:")
print(monthly_trend)

# 17.Identify any months with extreme rainfall events (e.g., heavy
storms) and provide the corresponding measurement.
extreme_rainfall_months = data.iloc[:, 3:15].columns[data.iloc[:,
3:15].max() > 500]
print("Months with extreme rainfall events:")
print(extreme_rainfall_months)
```

```
# 18.Determine the season with the highest rainfall and provide the
total rainfall for that season.
data['SEASONAL_TOTAL'] = data[['JAN', 'FEB', 'MAR']].sum(axis=1) +
data[['APR', 'MAY', 'JUN']].sum(axis=1) + data[['JUL', 'AUG',
'SEP']].sum(axis=1) + data[['OCT', 'NOV', 'DEC']].sum(axis=1)
wettest_season = data[['DISTRICT',
'SEASONAL_TOTAL']].sort_values(by='SEASONAL_TOTAL',
ascending=False).iloc[0]
print("Season with the highest rainfall:", wettest_season['DISTRICT'])
print("Total rainfall for the wettest season:",
wettest_season['SEASONAL_TOTAL'])

# 19.Identify any consecutive months with increasing or decreasing
rainfall and provide the corresponding measurements.
consecutive_increasing_months = data.iloc[:, 3:15].apply(lambda x:
all(x[i] <= x[i+1] for i in range(len(x)-1)), axis=1)
consecutive_decreasing_months = data.iloc[:, 3:15].apply(lambda x:
all(x[i] >= x[i+1] for i in range(len(x)-1)), axis=1)

print("Consecutive months with increasing rainfall:")
print(data[consecutive_increasing_months][['DISTRICT', 'JAN', 'FEB',
'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC']])

print("Consecutive months with decreasing rainfall:")
print(data[consecutive_decreasing_months][['DISTRICT', 'JAN', 'FEB',
'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV', 'DEC']])

# 20.Calculate the monthly rainfall anomaly by comparing each month's
rainfall measurement to the long-term average for that month.
monthly_anomaly = data.iloc[:, 3:15].subtract(data['MONTHLY_AVG'],
axis=1)
print("Monthly rainfall anomaly:")
print(monthly_anomaly)
```

Output

Average monthly rainfall for each month:

	DISTRICT	JAN	FEB	MAR	APR	MAY	JUN	JUL	AUG	SEP \
0	SANGLI	107.3	57.9	65.2	117.0	358.5	295.5	285.0	271.9	354.8
1	SATARA	43.7	26.0	18.6	90.5	374.4	457.2	421.3	423.1	455.6
2	PUNE	32.7	15.9	8.6	53.4	343.6	503.3	465.4	460.9	454.8
3	MUMBAI	42.2	80.8	176.4	358.5	306.4	447.0	660.1	427.8	313.6

4	AMRAVATI	33.3	79.5	105.9	216.5	323.0	738.3	990.9	711.2	568.0
5	NASHIK	28.0	48.3	85.3	101.5	140.5	228.4	217.4	182.8	159.8
6	AURANGABAD	42.2	72.7	141.0	316.9	328.7	614.7	851.9	500.6	418.3
7	HINGOLI	42.2	80.8	176.4	358.5	306.4	447.0	660.1	427.8	313.6
8	NANDURBAR	83.7	153.9	303.5	383.6	268.0	374.2	272.0	160.5	266.7
9	DHULE	70.3	170.9	367.9	554.4	334.2	526.2	460.8	291.5	353.6
10	SOLAPUR	33.5	67.8	106.1	226.9	453.0	640.5	609.5	503.4	492.3
11	AHMEDNAGAR	97.5	109.3	92.4	204.3	266.2	284.1	248.9	270.5	192.7
12	RATNAGIRI	74.3	176.7	362.6	397.5	408.7	801.9	653.0	417.9	686.0
13	NANDED	26.0	66.7	76.8	229.2	239.5	416.6	592.4	312.4	291.1
14	NAGPUR	83.7	153.9	303.5	383.6	268.0	374.2	272.0	160.5	266.7
15	CHANDRAPUR	35.2	43.5	58.9	134.3	341.1	665.3	749.9	579.1	490.9
16	LATUR	49.0	74.4	96.5	156.9	208.0	345.7	368.5	256.2	275.9
17	YAVATMAL	35.2	43.5	58.9	134.3	341.1	665.3	749.9	579.1	490.9
18	AKOLA	82.7	70.0	128.2	245.7	271.4	292.7	404.0	276.3	283.5

	OCT	NOV	DEC
0	326.0	315.2	250.9
1	301.2	275.8	128.3
2	276.1	198.6	100.0
3	167.1	34.1	29.8
4	206.9	29.5	31.7
5	75.9	20.9	11.6
6	218.7	42.9	22.9
7	167.1	34.1	29.8
8	167.2	64.0	56.0
9	275.0	64.9	74.2
10	214.7	19.2	11.3
11	78.5	49.5	27.2

12 264.9 86.9 71.7

13 126.8 33.7 29.5

14 167.2 64.0 56.0

15 233.9 40.3 27.0

16 138.2 34.4 27.2

17 233.9 40.3 27.0

18 92.3 32.3 42.4

Month with the highest rainfall: ANNUAL

Highest rainfall measurement: 4402.1

Month with the lowest rainfall: MAR

Lowest rainfall measurement: 8.6

Total rainfall for the year: 110101.7

Months with above-average rainfall: Index([], dtype='object')

Months with below-average rainfall: Index(['FEB', 'MAR', 'APR', 'MAY', 'JUN', 'JUL', 'AUG', 'SEP', 'OCT', 'NOV',

'DEC'],

dtype='object')

Average quarterly rainfall:

	DISTRICT	Q1	Q2	Q3	Q4
0	SANGLI	76.800000	257.000000	303.900000	297.366667
1	SATARA	29.433333	307.366667	433.333333	235.100000
2	PUNE	19.066667	300.100000	460.366667	191.566667
3	MUMBAI	99.800000	370.633333	467.166667	77.000000
4	AMRAVATI	72.900000	425.933333	756.700000	89.366667
5	NASHIK	53.866667	156.800000	186.666667	36.133333
6	AURANGABAD	85.300000	420.100000	590.266667	94.833333
7	HINGOLI	99.800000	370.633333	467.166667	77.000000
8	NANDURBAR	180.366667	341.933333	233.066667	95.733333
9	DHULE	203.033333	471.600000	368.633333	138.033333
10	SOLAPUR	69.133333	440.133333	535.066667	81.733333

11 AHMEDNAGAR 99.733333 251.533333 237.366667 51.733333

12 RATNAGIRI 204.533333 536.033333 585.633333 141.166667

13 NANDED 56.500000 295.100000 398.633333 63.333333

14 NAGPUR 180.366667 341.933333 233.066667 95.733333

15 CHANDRAPUR 45.866667 380.233333 606.633333 100.400000

16 LATUR 73.300000 236.866667 300.200000 66.600000

17 YAVATMAL 45.866667 380.233333 606.633333 100.400000

18 AKOLA 93.633333 269.933333 321.266667 55.666667

Wettest quarter: 12

Total rainfall for the wettest quarter: 1467.3666666666666

Driest quarter: 5

Total rainfall for the driest quarter: 433.46666666666664