

DOP: / /2023 DOS: / /2023

#### **Experiment No: 08**

Aim: - Design and Implement IoT Data Processing using Pandas.

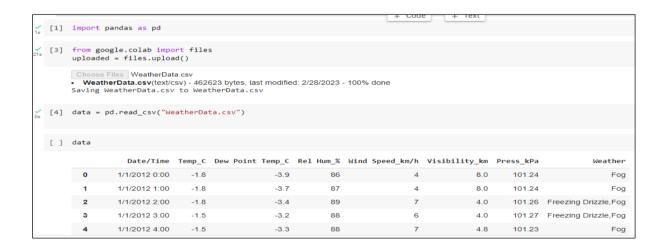
#### Theory:

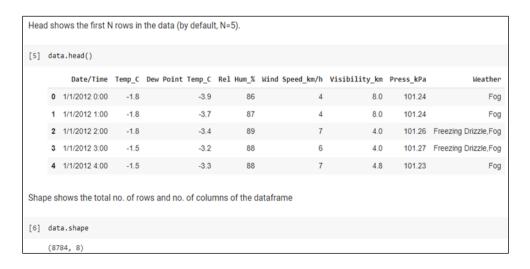
#### What is Pandas?

Pandas is an open-source Python library used for data manipulation, analysis, and visualization. It provides data structures for efficiently storing and manipulating large datasets, and a wide range of functions for data processing and analysis.

The main data structures in pandas are the Series (one-dimensional labeled array) and Data Frame (two-dimensional labeled data structure with columns of potentially different types). Pandas also provides tools for working with time series data, handling missing data, and merging, grouping, and reshaping datasets.

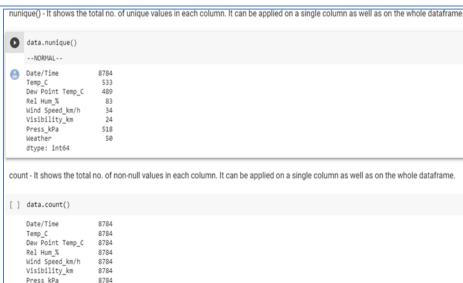
Pandas is widely used in data science and analytics projects, as it offers a convenient and powerful way to work with tabular data. It can be used in conjunction with other Python libraries such as NumPy, Matplotlib, and Scikit-learn for more complex data analysis and modeling tasks.







Index attribute provides t	he index of the dataframe
[ ] data.index	
RangeIndex(start=0	, stop=8784, step=1)
Columns shows the name	e of each column
data.columns	
	, 'Temp_C', 'Dew Point Temp_C', 'Rel Hum_%', _km/h', 'Visibility_km', 'Press_kPa', 'Weather'], t')
dtypes shows the data-ty	pe of each column
[ ] data.dtypes	
Date/Time Temp_C Dew Point Temp_C Rel Hum_% Wind Speed_km/h Visibility_km	object float64 float64 int64 int64 float64





```
value_counts - In a column, it shows all the unique values with their count. It can be applied on a single column only.
[ ] data.value_counts

        cbound method DataFrame.value_counts of
        Date/Time

        0
        1/1/2012 0:00
        -1.8
        -3.9
        86

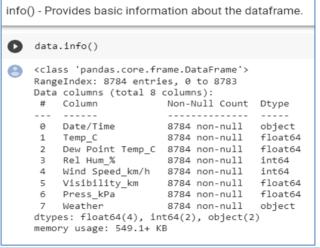
        1
        1/1/2012 1:00
        -1.8
        -3.7
        87

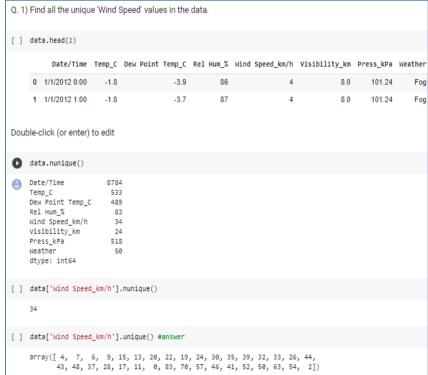
        2
        1/1/2012 2:00
        -1.8
        -3.4
        89

        3
        1/1/2012 3:00
        -1.5
        -3.2
        88

        4
        1/1/2012 4:00
        -1.5
        -3.3
        88

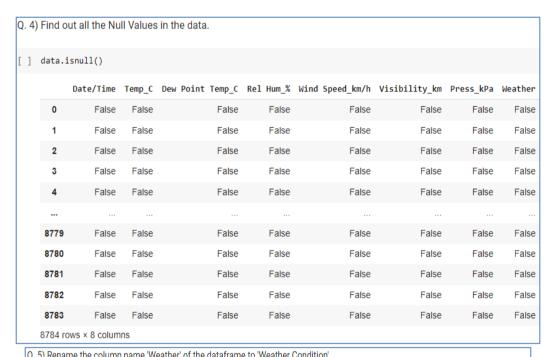
                                                                                                    Date/Time Temp_C Dew Point Temp_C Rel Hum_% Wind Speed_km/h
                                                                                                       88
88
                                                                                                                                           7
                                                                    -2.7 81
-2.4 83
-1.5 93
-1.8 89
-2.1 86
        8779 12/31/2012 19:00 0.1
       8789 12/31/2012 20:00 0.2
8781 12/31/2012 21:00 -0.5
8782 12/31/2012 21:00 -0.2
8783 12/31/2012 21:00 0.0
                  Visibility km Press kPa
                                                                                   Weather
                                   8.0 101.24
8.0 101.24
                                                                                      Fog
                                                                                             Fog
                                                  101.24
                                    4.0 101.26 Freezing Drizzle,Fog
                                    4 0
                                                 101.27 Freezing Drizzle,Fog
        4
                                    4.8 101.23
                                            100.13
                                    9.7
                                                                                             Snow
                                     9.7
                                                 100.03
        8780
                                                                                             Snow
```





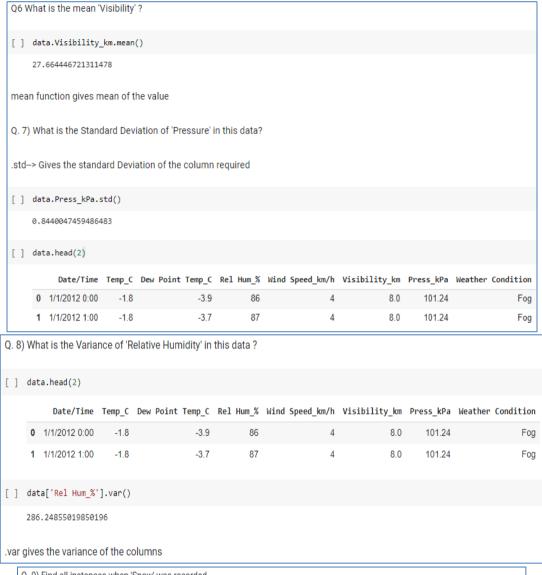


). 2) F	. 2) Find the number of times when the 'Weather is exactly Clear													
] d	] data.groupby('Weather').get_group('Clear') #answer													
		Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather					
	67	1/3/2012 19:00	-16.9	-24.8	50	24	25.0	101.74	Clear					
	114	1/5/2012 18:00	-7.1	-14.4	56	11	25.0	100.71	Clear					
	115	1/5/2012 19:00	-9.2	-15.4	61	7	25.0	100.80	Clear					
	116	1/5/2012 20:00	-9.8	-15.7	62	9	25.0	100.83	Clear					
	117	1/5/2012 21:00	-9.0	-14.8	63	13	25.0	100.83	Clear					
	8646	12/26/2012 6:00	-13.4	-14.8	89	4	25.0	102.47	Clear					
	8698	12/28/2012 10:00	-6.1	-8.6	82	19	24.1	101.27	Clear					
	8713	12/29/2012 1:00	-11.9	-13.6	87	11	25.0	101.31	Clear					
	8714	12/29/2012 2:00	-11.8	-13.1	90	13	25.0	101.33	Clear					
	8756	12/30/2012 20:00	-13.8	-16.5	80	24	25.0	101.52	Clear					
1	1326 ro	ws × 8 columns												



data.r	data.rename(columns={'Weather':'Weather Condition'})												
	Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather Conditio					
0	1/1/2012 0:00	-1.8	-3.9	86	4	8.0	101.24	Fo					
1	1/1/2012 1:00	-1.8	-3.7	87	4	8.0	101.24	Fo					
2	1/1/2012 2:00	-1.8	-3.4	89	7	4.0	101.26	Freezing Drizzle,Fo					
3	1/1/2012 3:00	-1.5	-3.2	88	6	4.0	101.27	Freezing Drizzle,Fo					
4	1/1/2012 4:00	-1.5	-3.3	88	7	4.8	101.23	Fo					
							***						
8779	12/31/2012 19:00	0.1	-2.7	81	30	9.7	100.13	Sno					
8780	12/31/2012 20:00	0.2	-2.4	83	24	9.7	100.03	Sno					
8781	12/31/2012 21:00	-0.5	-1.5	93	28	4.8	99.95	Sno					
8782	12/31/2012 22:00	-0.2	-1.8	89	28	9.7	99.91	Sno					
8783	12/31/2012 23:00	0.0	-2.1	86	30	11.3	99.89	Sno					





). 9) Fin	nd all	instances when	'Snow' w	as recorded.					
] dat	ta[da	ata['Weather Cor	ndition']	]=='Snow']					
		Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather Condition
5	55	1/3/2012 7:00	-14.0	-19.5	63	19	25.0	100.95	Snov
8	34	1/4/2012 12:00	-13.7	-21.7	51	11	24.1	101.25	Snov
8	36	1/4/2012 14:00	-11.3	-19.0	53	7	19.3	100.97	Snov
8	37	1/4/2012 15:00	-10.2	-16.3	61	11	9.7	100.89	Snov
8	38	1/4/2012 16:00	-9.4	-15.5	61	13	19.3	100.79	Snov
87	779	12/31/2012 19:00	0.1	-2.7	81	30	9.7	100.13	Sno
87	780	12/31/2012 20:00	0.2	-2.4	83	24	9.7	100.03	Sno
87	781	12/31/2012 21:00	-0.5	-1.5	93	28	4.8	99.95	Sno
87	782	12/31/2012 22:00	-0.2	-1.8	89	28	9.7	99.91	Sno
87	783	12/31/2012 23:00	0.0	-2.1	86	30	11.3	99.89	Sno
390	) row	s × 8 columns							



Q. 10) F	ind a	all instances whe	n 'Wind S	Speed is above 24' a	nd 'Visibility	/ is 25'.						
[] da	ta[(	data['Wind Speed	1 km/h'l:	>24)&(data['Visibi	litv km'l==	25)]						
. ,	Date/Time Temp_C Dew Point Temp_C Rel Hum_% Wind Speed_km/h Visibility_km Press_kPa Weather Condition											
	23	1/1/2012 23:00	5.3	2.0	79	niina spec	30	25.0	99.31	neacher	Cloudy	
	24	1/2/2012 0:00	5.2	1.5	77		35	25.0	99.26	R	ain Showers	
2	25	1/2/2012 1:00	4.6	0.0	72		39	25.0	99.26		Cloudy	
:	26	1/2/2012 2:00	3.9	-0.9	71		32	25.0	99.26	M	ostly Cloudy	
2	27	1/2/2012 3:00	3.7	-1.5	69		33	25.0	99.30	M	ostly Cloudy	
87	705	12/28/2012 17:00	-8.6	-12.0	76		26	25.0	101.34		Mainly Clea	
87	753	12/30/2012 17:00	-12.1	-15.8	74		28	25.0	101.26	1	Mainly Clea	
87	755	12/30/2012 19:00	-13.4	-16.5	77		26	25.0	101.47	ı	Mainly Clea	
87	759	12/30/2012 23:00	-12.1	-15.1	78		28	25.0	101.52	M	ostly Cloudy	
	760 hat is	12/31/2012 0:00 s the Mean value (	-11.1 of each o	-14.4 column against each	77 I 'Weather C	ondition?	26	25.0	101.51		Cloudy	
11) Wh	hat is		of each o	column against each		ondition?	26	25.0	101.51		Cloud	
11) Wh	hat is	s the Mean value o	of each o	column against each	ı 'Weather C			25.0 Wind Speed_km/		ility_km		
11) Wh	hat is	s the Mean value o	of each o	column against each ').mean() Temp_	"Weather C	t Temp_C ∣		Wind Speed_km/	/h Visibi		Press_kP	
11) Wh	hat is	s the Mean value o	of each o	column against each ').mean()  Temp_ ondition	"Weather C C Dew Poin	t Temp_C	Rel Hum_%	Wind Speed_km/	/h Visibi			
. 11) Wh	hat is	s the Mean value of pupby('Weather Co W Clear	of each o	column against each ').mean()  Temp_ ondition  6.82571	T Weather C  C Dew Poin  6	t Temp_C 0.089367 2.375810	Rel Hum_% 64.497738	Wind Speed_km/ 10.55731 16.12731	/h Visibi	0.153243	Press_kP	
. 11) Wh	hat is	s the Mean value of puppy('Weather Co W Clear Cloudy	of each C	column against each ').mean()  Temp_ ondition  6.82571 7.97054	n Weather C  C Dew Poin  6  4	t Temp_C   0.089367   2.375810   5.504878	Rel Hum_% 64.497738 69.592593	Wind Speed_km/ 10.55731 16.12731	/h Visibi	0.153243 6.625752	Press_kP 101.58744 100.91144	
11) Wh	hat is	s the Mean value of puppy ('Weather Co We Clear Cloudy Drizzle	of each condition	column against each ').mean()  Temp_ ondition  6.82571  7.97054  7.35365	'Weather C  C Dew Poin  6  4  9	t Temp_C   0.089367   2.375810   5.504878   7.033750	Rel Hum_% 64.497738 69.592593 88.243902	Wind Speed_km/ 10.55731 16.12731 16.09756 11.86250	/h Visibi	0.153243 6.625752 7.931707	Press_kF 101.58744 100.91144 100.43536	
11) Wh	hat is	s the Mean value of puppy('Weather Co We Clear Cloudy Drizzle,Fo	ondition eather C	column against each ').mean()  Temp_ ondition  6.82571 7.97054 7.35365 8.06750	'Weather C  C Dew Poin  6  4  9  0	t Temp_C 0.089367 2.375810 5.504878 7.033750 -0.700000	Rel Hum_% 64.497738 69.592593 88.243902 93.275000	Wind Speed_km/ 10.55731 16.12731 16.09756 11.86250 20.00000	7h Visibi	0.153243 6.625752 7.931707 5.257500	Press_kF 101.58744 100.91144 100.43536 100.78662	
11) Wh	hat is	s the Mean value of Dupby ('Weather Co	of each condition eather C	olumn against each ').mean()  Temp_ ondition  6.82571  7.97054  7.35365  8.06750  0.40000	C Dew Poin  G Dew Poin	t Temp_C   0.089367   2.375810   5.504878   7.033750   0.150000   0.150000	Rel Hum_% 64.497738 69.592593 88.243902 93.275000 92.000000	Wind Speed_km/ 10.55731 16.12731 16.09756 11.86250 20.00000	7h Visibi	0.153243 6.625752 7.931707 5.257500 4.000000	Press_kf 101.58744 100.91144 100.43536 100.78662 100.79000 100.89000	
11) Wh	hat is	s the Mean value of the Mean v	of each condition eather C	Temp_ondition  6.82571 7.97054 7.35365 8.06750 0.40000 1.05000	C Dew Poin  G  G  G  G  G  G  G  G  G  G  G  G  G	t Temp_C   0.089367   2.375810   5.504878   7.033750   0.150000   0.1200000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.120000   0.1200000   0.1200000   0.1200000   0.12000000   0.1200000000   0.1200000000   0.1200000000000000   0.120000000000000000	Rel Hum_% 64.497738 69.592593 88.243902 93.275000 92.000000	Wind Speed_km/ 10.55731 16.12731 16.09756 11.86250 20.00000 14.00000	7h Visibi	0.153243 6.625752 7.931707 5.257500 4.000000 0.500000	Press_kf 101.58744 100.91144 100.43536 100.78662 100.79000 100.89000 99.28133	
11) Wh	hat is	s the Mean value of Dupby ('Weather Co	of each o	Temp_ondition  6.82571 7.97054 7.35365 8.06750 0.40000 1.05000 0.69333	C Dew Poin  C Dew	t Temp_C 0.089367 2.375810 5.504878 7.033750 0.700000 0.150000 0.120000 3.159333	Rel Hum_% 64.497738 69.592593 88.243902 93.275000 92.000000 93.500000 95.866667	Wind Speed_km/ 10.55731 16.12731 16.09756 11.86250 20.00000 14.00000 15.53333	7h Visibi	0.153243 6.625752 7.931707 5.257500 4.00000 0.500000 5.513333	Press_kF 101.58744 100.91144 100.43536 100.78662 100.79000	

Q. 12) What is the Minimum & Maximum value of each column against each "Weather Condition ?											
0	data.groupby('Weather Condition	').min()									
D-	Паде	1/22/2012 12:00	-11.0	-10.0	00	U	4.0	เบบ.จอ			
L,	Mainly Clear	1/10/2012 11:00	-22.8	-28.0	20	0	12.9	98.67			
	Moderate Rain,Fog	12/10/2012 8:00	1.7	0.8	94	17	6.4	99.98			
	Moderate Snow	1/12/2012 15:00	-6.3	-7.6	83	26	0.6	99.88			
	Moderate Snow, Blowing Snow	12/27/2012 10:00	-5.5	-6.6	92	39	0.6	100.50			
	Mostly Cloudy	1/1/2012 16:00	-23.2	-28.5	18	0	11.3	98.36			
	Rain	1/1/2012 18:00	0.3	-5.7	40	0	4.0	97.52			
	Rain Showers	1/1/2012 22:00	1.6	-7.2	37	0	6.4	98.51			
	Rain Showers,Fog	10/20/2012 3:00	12.8	12.1	96	13	6.4	99.83			
	Rain Showers, Snow Showers	11/4/2012 8:00	2.1	-1.8	75	17	19.3	101.09			
	Rain,Fog	1/23/2012 18:00	0.0	-1.2	83	0	2.0	98.61			
	Rain,Haze	3/13/2012 7:00	4.0	1.0	81	7	4.0	100.50			
	Rain,Ice Pellets	12/18/2012 5:00	0.6	-0.6	92	24	9.7	100.12			
	Rain, Snow	1/10/2012 5:00	0.6	-1.7	81	13	2.4	98.18			
	Rain, Snow Grains	12/21/2012 0:00	1.9	-2.1	75	26	25.0	100.60			

-8.000000 82.000000

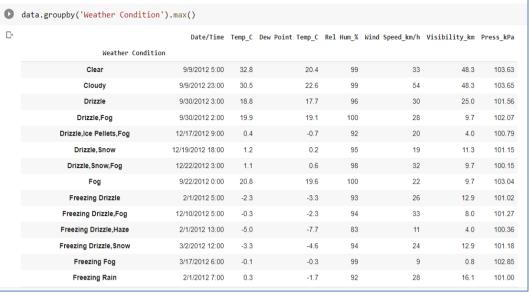
10.333333

2.666667 100.316667

-5.433333

Freezing Drizzle,Haze







Q. 14) Find all instances when 'Weather is Clear' or 'Visibility is above 40'. data[(data['Weather Condition']=='Clear')&(data['Visibility\_km']>40)] Гэ Date/Time Temp\_C Dew Point Temp\_C Rel Hum\_% Wind Speed\_km/h Visibility\_km Press\_kPa Weather Condition 351 1/15/2012 15:00 -22.8 102.71 352 1/15/2012 16:00 -22.8 48.3 102.79 Clear 425 1/18/2012 17:00 -11.3 -18.8 54 48.3 101.54 Clear -18.4 19 440 1/19/2012 8:00 -13.7 68 48.3 101.84 Clear 17 1/19/2012 9:00 -12.7 -17.2 69 48.3 101.73 441 Clear 12/15/2012 8:00 -10.7 67 13 48.3 102.69 8384 -15.6 Clear 12/15/2012 9:00 -15.9 8389 12/15/2012 13:00 -14.7 60 19 48.3 102.64 Clear 8631 12/25/2012 15:00 -13.7 59 17 48.3 101.98 Clear -13.9 8632 12/25/2012 16:00 -7.5 60 48.3 102.03 Clear 313 rows × 8 columns



Q. 1	5) Find	all instances whe	n : A. 'We	ather is Clear' and '	Relative Hun	nidity is greater th	an 50' or B. 'Visil	oility is abov	re 40'
0	data[	(data['Weathe	er Condi	ition']=='Clear	')&(data[	'Rel Hum_%']==	:50) (data['V	isibility	_km']>40)]
₽		Date/Time	Temp_C	Dew Point Temp_C	Rel Hum_%	Wind Speed_km/h	Visibility_km	Press_kPa	Weather Condition
	67	1/3/2012 19:00	-16.9	-24.8	50	24	25.0	101.74	Clear
	106	1/5/2012 10:00	-6.0	-10.0	73	17	48.3	100.45	Mainly Clear
	107	1/5/2012 11:00	-5.6	-10.2	70	22	48.3	100.41	Mainly Clear
	108	1/5/2012 12:00	-4.7	-9.6	69	20	48.3	100.38	Mainly Clear
	109	1/5/2012 13:00	-4.4	-9.7	66	26	48.3	100.40	Mainly Clear
	8748	12/30/2012 12:00	-12.2	-15.7	75	26	48.3	100.91	Mostly Cloudy
	8749	12/30/2012 13:00	-12.4	-16.2	73	37	48.3	100.92	Mostly Cloudy
	8750	12/30/2012 14:00	-11.8	-16.1	70	37	48.3	100.96	Mainly Clear
	8751	12/30/2012 15:00	-11.3	-15.6	70	32	48.3	101.05	Mainly Clear
	8752	12/30/2012 16:00	-11.4	-15.5	72	26	48.3	101.15	Mainly Clear

#### **Conclusion**:

Pandas is a powerful and widely-used Python library for data manipulation, analysis, and visualization. Its data structures and functions allow for efficient handling of large datasets and provide a wide range of tools for data processing, manipulation, and analysis. Pandas is a must-have library for anyone working with tabular data in Python and is commonly used in data science and analytics projects.