Cross tab or Contingency table..

Contingency table: A table showing the distribution of one variable in rows and another in columns, used to study the correlation between the two variables.

~~data~~ S = pd.read_excel('survey.xlsx')

Calls

	Nationality	Handedness	Sex	Age
0	American	Right-handed	Male	25
1	"	"	Female	30
2	"	Left-handed	"	35
3	British	Right-handed	Male	28
4	"	Left-handed	Male	32
5	"	"	Female	40

pd.crosstab(s.Nationality, s.Handedness)

Handedness / Nationality	Left-handed	Right-handed
American	1	2
British	2	1

pd.crosstab(s.Nationality, s.Handedness, margin=True)

Handedness / Nationality	Left-handed	Right-handed	All
American	1	2	3
British	2	1	3
All	3	3	6

pd.crosstab(s.Nationality, [s.Sex, s.Handedness], margins = True)

Sex / Handedness / Nationality	Female		Male		All
	Left-handed	Right-handed	Left-handed	Right-handed	
American	1	1	0	1	3
British	1	0	1	1	3
All	2	1	1	2	6

import numpy as np

pd.crosstab(s.Nationality, s.Sex, values=s.Age, aggfunc=np.average)

call datas

Sex / Nationality	Female	Male
American	32.5	25.0
British	40.0	30.0

DateTime Index..

datas

	date	city	temperature	humidity	wind_speed	weather_condition
0	2024-4-1	Manhattan	70	50	10	Sunny
1	4	Los Angeles	75	55	12	Cloudy
2	2024-4-2	N.Y	72	48	9	Partly Cloudy
3	4	L.A	77	60	11	Rainy
4	2024-4-3	N.Y	68	52	8	Sunny
5	4	L.A	73	57	13	Thunderstorms

datas = pd.read_csv('w.csv', parse_dates=['date'])
datas.set_index('date')

(The output will be same only date is change in the indexing)

stock = pd.read_csv('stock_data.csv', parse_dates=['Date'], index_col='Date')

call stock

Date	Open	High	Low	Close	Volume	Name
2006-1-3	39.69	41.22	38.79	40.91	2423...	AABA
2006-1-4	41.22	41.90	40.77	40.97	2055..	
2006-1-5	40.93	41.73	40.85	41.53	1282..	4
2006-1-6	42.88	43.57	42.80	43.21	2942..	4
2006-1-9	43.10	43.66	42.82	43.42	102...	4
. .	. .	. .	. .	. .	. .	. .
2017-12-22	71.42	71.87	71.22	71.58	1097..	4
2017-12-26	70.94	71.39	69.63	69.80	854..	4
2017-12-27	69.77	70.49	69.69	70.06	634..	4
2017-12-28	70.12	70.32	69.51	69.82	755..	4
2017-12-29	69.77					

stock['2017-12-27':'2017-12-30']

Date	Open	High	Low	Close	Volume	Name
2017-12-27	69.77	70.49	69.69	~~69.~~ 70.06	634524	AABA.
2017-12-28	70.12	70.32	69.32	69.82	7556877	AABA
2017-12-29	69.79	70.13	70.13	69.85	6613076	AABA.

stock['2017-12-27':'2017-12-30'].High.max()

Output

70.49

~~to~~_datetime()

dates = ['2017-12-05', 'Jan5 2017', 6/5/2017, '2017.01.05', '2017/01/05', '20170105', '5jan2017', '6 january 2019']

pd.~~to~~_datetime('01/2/15', dayfirst = True)

Output

Timestamp('2015-02-01 00:00:00')

pd.to_datetime('01/2/15', dayfirst = True)

Output

Timestamp('2001-02-15 00:00:00')

Resampling with DatetimeIndex.

stock = pd.read_csv('stock_data.csv', parse_dates=['Date'], index_col='Date')

plot_data = stock.Close.resample("W").mean()

call plot_data

Date	
2006-1-8	41.6550
2006-1-15	41.8120
2006-1-22	35.8400
2006-1-29	34.7580
2006-2-5	34.440
...	
2017-12-3	70.9140
2017-12-10	69.4640
2017-12-17	70.4120
2017-12-24	70.8840
2017-12-31	69.8975