

Ex2 - Filtering and Sorting Data

This time we are going to pull data directly from the internet.

Step 1. Import the necessary libraries

```
import pandas as pd
```

Step 2. Import the dataset from this [address](#).

Step 3. Assign it to a variable called euro12.

```
euro12 = pd.read_csv('https://raw.githubusercontent.com/thieu1995/csv-files/main/data/pandas/Euro_2012_stats_TEAM.csv')
euro12
```

Table with 35 columns: Team, Goals, Shots on target, Shots off target, Shooting Accuracy, Goals-to-shots, % Total shots (inc. Blocked), Hit Woodwork, Penalty goals, Penalties not scored, Saves made, Saves-to-shots ratio, Fouls Won, Fouls Conceded, Offsides, Yellow Cards, Red Cards, Subs on, Subs off. The table lists statistics for 16 teams: Croatia, Czech Republic, Denmark, England, France, Germany, Greece, Italy, Netherlands, Poland, Portugal, Republic of Ireland, Russia, Spain, Sweden, and Ukraine.

Step 4. Select only the Goal column.

```
euro12['Goals']
```

Table with 2 columns: Index (0-15), Goals. It shows the number of goals scored by each team: Croatia (4), Czech Republic (4), Denmark (4), England (5), France (3), Germany (10), Greece (5), Italy (6), Netherlands (2), Poland (2), Portugal (6), Republic of Ireland (1), Russia (5), Spain (12), Sweden (5), and Ukraine (2).

Step 5. How many team participated in the Euro2012?

```
euro12.shape[0]
```

16

Step 6. What is the number of columns in the dataset?

```
euro12.shape[1]
```

35

Step 7. View only the columns Team, Yellow Cards and Red Cards and assign them to a dataframe called discipline

```
print(euro12[['Team', 'Yellow Cards', 'Red Cards']])
discipline = euro12[['Team', 'Yellow Cards', 'Red Cards']]
discipline
```

```
0      Croatia      9      0
1  Czech Republic      7      0
2      Denmark      4      0
3      England      5      0
4      France      6      0
5      Germany      4      0
6      Greece      9      1
7      Italy     16      0
8  Netherlands      5      0
9      Poland      7      1
10     Portugal     12      0
11 Republic of Ireland      6      1
12      Russia      6      0
13      Spain     11      0
14      Sweden      7      0
15     Ukraine      5      0
```

	Team	Yellow Cards	Red Cards
0	Croatia	9	0
1	Czech Republic	7	0
2	Denmark	4	0
3	England	5	0
4	France	6	0
5	Germany	4	0
6	Greece	9	1
7	Italy	16	0
8	Netherlands	5	0
9	Poland	7	1
10	Portugal	12	0
11	Republic of Ireland	6	1
12	Russia	6	0
13	Spain	11	0
14	Sweden	7	0
15	Ukraine	5	0

Step 8. Sort the teams by Red Cards, then to Yellow Cards

```
discipline.sort_values(by=['Red Cards', 'Yellow Cards'])
```

```
0
```

	Team	Yellow Cards	Red Cards
2	Denmark	4	0
5	Germany	4	0
3	England	5	0
8	Netherlands	5	0
15	Ukraine	5	0
4	France	6	0
12	Russia	6	0
1	Czech Republic	7	0
14	Sweden	7	0
0	Croatia	9	0
13	Spain	11	0
10	Portugal	12	0
7	Italy	16	0
11	Republic of Ireland	6	1
9	Poland	7	1
6	Greece	9	1

Step 9. Calculate the mean Yellow Cards given per Team

```
print(discipline['Yellow Cards'].mean())

7.4375
```

Step 10. Filter teams that scored more than 6 goals

```
euro12[euro12['Goals'] > 6]
```

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored	...	Saves made	Saves-to-shots ratio	Fouls Won	Fouls Conceded	Offsides	Yellow Cards	Red Cards	Subs on	Subs off
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	10	62.6%	63	49	12	4	0	15	15
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	...	15	93.8%	102	83	19	11	0	17	17

Step 11. Select the teams that start with G

```
euro12[euro12['Team'].str.startswith('G')]
```

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored	...	Saves made	Saves-to-shots ratio	Fouls Won	Fouls Conceded	Offsides	Yellow Cards	Red Cards	Subs on	Subs off	P
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	10	62.6%	63	49	12	4	0	15	15	
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1	...	13	65.1%	67	48	12	9	1	12	12	

Step 12. Select the first 7 columns

```
euro12.iloc[:, :7]
```

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)
0	Croatia	4	13	12	51.9%	16.0%	32
1	Czech Republic	4	13	18	41.9%	12.9%	39
2	Denmark	4	10	10	50.0%	20.0%	27
3	England	5	11	18	50.0%	17.2%	40
4	France	3	22	24	37.9%	6.5%	65
5	Germany	10	32	32	47.8%	15.6%	80
6	Greece	5	8	18	30.7%	19.2%	32
7	Italy	6	34	45	43.0%	7.5%	110
8	Netherlands	2	12	36	25.0%	4.1%	60
9	Poland	2	15	23	39.4%	5.2%	48
10	Portugal	6	22	42	34.3%	9.3%	82
11	Republic of Ireland	1	7	12	36.8%	5.2%	28
12	Russia	5	9	31	22.5%	12.5%	59
13	Spain	12	42	33	55.9%	16.0%	100
14	Sweden	5	17	19	47.2%	13.8%	39
15	Ukraine	2	7	26	24.0%	6.0%	28

Step 13. Select all columns except the last 3.

```
euro12.iloc[:, :-3]
```



	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored	...	Clean Sheets	Blocks	Goals conceded	Saves made	Saves-to-shots ratio	Fouls Won	Fouls Conceded	Offsides
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0	...	0	10	3	13	81.3%	41	62	2
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0	...	1	10	6	9	60.1%	53	73	8
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0	...	1	10	5	10	66.7%	25	38	8
3	England	5	11	18	50.0%	17.2%	40	0	0	0	...	2	29	3	22	88.1%	43	45	6
4	France	3	22	24	37.9%	6.5%	65	1	0	0	...	1	7	5	6	54.6%	36	51	5
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	1	11	6	10	62.6%	63	49	12
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1	...	1	23	7	13	65.1%	67	48	12
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0	...	2	18	7	20	74.1%	101	89	16
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0	...	0	9	5	12	70.6%	35	30	3
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0	...	0	8	3	6	66.7%	48	56	3
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0	...	2	11	4	10	71.5%	73	90	10
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0	...	0	23	9	17	65.4%	43	51	11
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0	...	0	8	3	10	77.0%	34	43	4
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	...	5	8	1	15	93.8%	102	83	19
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0	...	1	12	5	8	61.6%	35	51	7
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0	...	0	4	4	13	76.5%	48	31	4

16 rows × 32 columns



▼ Step 14. Present only the Shooting Accuracy from England, Italy and Russia

```
euro12.loc[euro12['Team'].isin(['England','Italy','Russia']),['Team','Shooting Accuracy']]
```



	Team	Shooting Accuracy
3	England	50.0%
7	Italy	43.0%
12	Russia	22.5%