

Weekly Pandas 🐼 Challenge #2

```
In [1]: import pandas as pd
file = "Euro_2012_stats_TEAM.csv"
```

```
In [2]: df = pd.read_csv(file)
```

Take A Peek At Dataset

```
In [3]: df.head()
```

```
Out[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	...	Saves made	Saves-to-shots ratio	Fouls Won	Fouls Conceded	Offsides	Yellow Cards	Red Cards
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0	...	13	81.3%	41	62	2	9	0
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0	...	9	60.1%	53	73	8	7	0
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0	...	10	66.7%	25	38	8	4	0
3	England	5	11	18	50.0%	17.2%	40	0	0	0	...	22	88.1%	43	45	6	5	0
4	France	3	22	24	37.9%	6.5%	65	1	0	0	...	6	54.6%	36	51	5	6	0

5 rows × 35 columns

Q1. How many teams participated in the Euro2012?

```
In [4]: print(f"Total {len(df['Team'])} teams participated in the Euro2012")
Total 16 teams participated in the Euro2012
```

Q2. What is the number of columns in the dataset?

```
In [5]: print(f"Total {len(df.columns)} number of columns in the dataset")
Total 35 number of columns in the dataset
```

Q3. View only the columns Team, Yellow Cards and Red Cards and assign them to a dataframe called discipline.

```
In [6]: discipline = df[['Team', 'Yellow Cards', 'Red Cards']]
discipline.head()
```

```
Out[6]:
```

	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

Q4. Sort the teams by Red Cards, then to Yellow Cards.

```
In [7]: df.sort_values(by=['Red Cards', 'Yellow Cards'], axis=0, ascending=True, inplace=False)
```

Out[7]:

	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	I Ca
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0	...	10	66.7%	25	38	8	4	
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	10	62.6%	63	49	12	4	
3	England	5	11	18	50.0%	17.2%	40	0	0	0	...	22	88.1%	43	45	6	5	
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0	...	12	70.6%	35	30	3	5	
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0	...	13	76.5%	48	31	4	5	
4	France	3	22	24	37.9%	6.5%	65	1	0	0	...	6	54.6%	36	51	5	6	
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0	...	10	77.0%	34	43	4	6	
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0	...	9	60.1%	53	73	8	7	
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0	...	8	61.6%	35	51	7	7	
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0	...	13	81.3%	41	62	2	9	
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	...	15	93.8%	102	83	19	11	
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0	...	10	71.5%	73	90	10	12	
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0	...	20	74.1%	101	89	16	16	
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0	...	17	65.4%	43	51	11	6	
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0	...	6	66.7%	48	56	3	7	
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1	...	13	65.1%	67	48	12	9	

16 rows × 35 columns

Q5. Calculate the mean Yellow Cards given per Team.

```
In [8]: df['Avg Per Team'] = df['Yellow Cards'].apply(lambda x : x/sum(df['Yellow Cards']))
df[['Team', 'Yellow Cards', 'Avg Per Team']]
```

Out[8]:

	Team	Yellow Cards	Avg Per Team
0	Croatia	9	0.075630
1	Czech Republic	7	0.058824
2	Denmark	4	0.033613
3	England	5	0.042017
4	France	6	0.050420
5	Germany	4	0.033613
6	Greece	9	0.075630
7	Italy	16	0.134454
8	Netherlands	5	0.042017
9	Poland	7	0.058824
10	Portugal	12	0.100840
11	Republic of Ireland	6	0.050420
12	Russia	6	0.050420
13	Spain	11	0.092437
14	Sweden	7	0.058824
15	Ukraine	5	0.042017

Q6. Filter teams that scored more than 6 goals.

```
In [9]: df[df['Goals']>6]
```

Out[9]:

	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-to-shots ratio	Fouls Won	Fouls Conceded	Offsides	Yellow Cards	Red Cards	Subs on
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	62.6%	63	49	12	4	0	15
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	...	93.8%	102	83	19	11	0	17

2 rows × 36 columns

Q7. Select the teams that start with the letter G.

```
In [10]: df['Team'][df['Team'].apply(lambda x : x.startswith('G'))]
```

```
Out[10]: 5    Germany
6    Greece
Name: Team, dtype: object
```

Q8. Select the first 7 columns.

```
In [11]: df[df.columns[:7]]
```

Out[11]:

	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	21.2%	6.0%	38

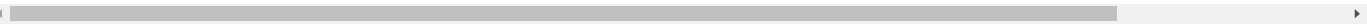
Q9. Select all columns except the last 3.

```
In [12]: df[df.columns[:-4]]
```

Out[12]:

	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	Con
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0	...	0	10	3	13	81.3%	41	
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0	...	1	10	6	9	60.1%	53	
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0	...	1	10	5	10	66.7%	25	
3	England	5	11	18	50.0%	17.2%	40	0	0	0	...	2	29	3	22	88.1%	43	
4	France	3	22	24	37.9%	6.5%	65	1	0	0	...	1	7	5	6	54.6%	36	
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...	1	11	6	10	62.6%	63	
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1	...	1	23	7	13	65.1%	67	
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0	...	2	18	7	20	74.1%	101	
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0	...	0	9	5	12	70.6%	35	
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0	...	0	8	3	6	66.7%	48	
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0	...	2	11	4	10	71.5%	73	
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0	...	0	23	9	17	65.4%	43	
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0	...	0	8	3	10	77.0%	34	
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	...	5	8	1	15	93.8%	102	
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0	...	1	12	5	8	61.6%	35	
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0	...	0	4	4	13	76.5%	48	

16 rows × 32 columns



Q10. Present only the Shooting Accuracy from England, Italy and Russia.

```
In [13]: df[['Team', 'Shooting Accuracy']]\n[ df['Team'].apply(lambda x : x in ['Italy', 'England', 'Russia']) ]
```

Out[13]:

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

Challenge Completed Successfully 🌟 Ready For More 🍷