Forward School

Program Code: J620-002-4:2020

Program Name: FRONT-END SOFTWARE DEVELOPMENT

Title : Exercise 08 Filtering and Sorting Data

Name: Phua Yan Han

IC Number: 050824070059

Date : 27/6/23

Introduction : learning how to arrage data

Conclusion :learned more pandas function

Ex08 - Filtering and Sorting Data

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

Step 1. Import the necessary libraries

In [1]:
```
import numpy as np
import pandas as pd
```

Step 2. Import the dataset from this address (https://raw.githubusercontent.com/guipsamora/pandas_exercises/mas

In [2]:
```
data = pd.read_csv("https://raw.githubusercontent.com/guipsamora/pandas_exercis
data
```

Out[2]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0
3	England	5	11	18	50.0%	17.2%	40	0	0	0
4	France	3	22	24	37.9%	6.5%	65	1	0	0
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0

16 rows × 35 columns

Step 3. Assign it to a variable called euro12.

In [4]:
```
euro12 = data
euro12
```

Out[4]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0
3	England	5	11	18	50.0%	17.2%	40	0	0	0
4	France	3	22	24	37.9%	6.5%	65	1	0	0
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0

16 rows × 35 columns

Step 4. Select only the Goal column.

In [6]:
```
euro12['Goals']
```

Out[6]:
```
0      4
1      4
2      4
3      5
4      3
5     10
6      5
7      6
8      2
9      2
10     6
11     1
12     5
13    12
14     5
15     2
Name: Goals, dtype: int64
```

Step 5. How many team participated in the Euro2012?

In [7]:
```
len(euro12)
```

Out[7]:
```
16
```

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

In [10]:
```
len(euro12.columns)
```

Out[10]:
```
35
```

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

In [11]:
```
discipline = euro12[['Team','Yellow Cards','Red Cards']]
discipline
```

Out[11]:

	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

In [15]:
```
discipline.sort_values(['Red Cards', 'Yellow Cards'],ascending = [True, True])
```

Out[15]:

	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

Step 9. Calculate the mean Yellow Cards given per Team

In [16]:
```
discipline['Yellow Cards'].mean()
```

Out[16]: 7.4375

Step 10. Filter teams that scored more than 6 goals

In [17]:
```
euro12[(euro12['Goals'] > 6)]
```

Out[17]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored	..
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	..
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0	..

2 rows × 35 columns

Step 11. Select the teams that start with G

In [18]:
```
euro12[(euro12['Team'].str.contains("G"))]
```

Out[18]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored	...
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0	...
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1	...

2 rows × 35 columns

Step 12. Select the first 7 columns

In [23]:
```
euro12.iloc[:, 0:6]
```

Out[23]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots
0	Croatia	4	13	12	51.9%	16.0%
1	Czech Republic	4	13	18	41.9%	12.9%
2	Denmark	4	10	10	50.0%	20.0%
3	England	5	11	18	50.0%	17.2%
4	France	3	22	24	37.9%	6.5%
5	Germany	10	32	32	47.8%	15.6%
6	Greece	5	8	18	30.7%	19.2%
7	Italy	6	34	45	43.0%	7.5%
8	Netherlands	2	12	36	25.0%	4.1%
9	Poland	2	15	23	39.4%	5.2%
10	Portugal	6	22	42	34.3%	9.3%
11	Republic of Ireland	1	7	12	36.8%	5.2%
12	Russia	5	9	31	22.5%	12.5%
13	Spain	12	42	33	55.9%	16.0%
14	Sweden	5	17	19	47.2%	13.8%
15	Ukraine	2	7	26	21.2%	6.0%

Step 13. Select all columns except the last 3.

In [24]:
```
euro12.iloc[:, 0:len(euro12.columns)-3]
```

Out[24]:

	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	Penalties not scored
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	0
1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	0
2	Denmark	4	10	10	50.0%	20.0%	27	1	0	0
3	England	5	11	18	50.0%	17.2%	40	0	0	0
4	France	3	22	24	37.9%	6.5%	65	1	0	0
5	Germany	10	32	32	47.8%	15.6%	80	2	1	0
6	Greece	5	8	18	30.7%	19.2%	32	1	1	1
7	Italy	6	34	45	43.0%	7.5%	110	2	0	0
8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	0
9	Poland	2	15	23	39.4%	5.2%	48	0	0	0
10	Portugal	6	22	42	34.3%	9.3%	82	6	0	0
11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	0
12	Russia	5	9	31	22.5%	12.5%	59	2	0	0
13	Spain	12	42	33	55.9%	16.0%	100	0	1	0
14	Sweden	5	17	19	47.2%	13.8%	39	3	0	0
15	Ukraine	2	7	26	21.2%	6.0%	38	0	0	0

16 rows × 32 columns

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

In [27]:
```
euro12[euro12['Team'].isin(["England","Italy","Russia"])][['Team', 'Shooting Ac
```

Out[27]:

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