

Chapter 4 - Exercise 3: Lọc và sắp xếp dữ liệu

UEFA_European_Championship/Euro 2012

▼ Lấy dữ liệu từ đường dẫn

[Euro 12](#)

Tạo data frame euro12 từ dữ liệu trên. In euro2: type, shape, danh sách các cột

```
import pandas as pd
import numpy as np

# Đọc dữ liệu từ tập tin euro2012.csv vào dataframe euro12
euro12 = pd.read_csv('euro2012.csv', sep=',')

# Hiển thị thông tin type(euro12), shape, columns và 5 dòng đầu của euro12
print(type(euro12))
print(euro12.shape)
print(euro12.columns)
euro12.head(5)
```

```
<class 'pandas.core.frame.DataFrame'>
```

```
(16, 36)
```

```
Index(['Unnamed: 0', 'Team', 'Goals', 'Shots on target', 'Shots off target',  
      'Shooting Accuracy', '% Goals-to-shots', 'Total shots (inc. Blocked)',  
      'Hit Woodwork', 'Penalty goals', 'Penalties not scored', 'Headed goals',  
      'Passes', 'Passes completed', 'Passing Accuracy', 'Touches', 'Crosses',  
      'Dribbles', 'Corners Taken', 'Tackles', 'Clearances', 'Interceptions',  
      'Clearances off line', 'Clean Sheets', 'Blocks', 'Goals conceded',  
      'Saves made', 'Saves-to-shots ratio', 'Fouls Won', 'Fouls Conceded',  
      'Offsides', 'Yellow Cards', 'Red Cards', 'Subs on', 'Subs off',  
      'Players Used'],  
      dtype='object')
```

Unnamed: 0	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	...	Saves made	Saves-to-shots ratio	Fouls Won
0	Croatia	4	13	12	51.9%	16.0%	32	0	0	...	13	81.3%	41

```
# Câu 1: In giá trị cột Goals
```

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

```
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
```

```
# Câu 2: Có bao nhiêu đội tham gia Euro2012?
```

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

```
16
```

```
# Câu 3: In thông tin của Euro2012
```

```
euro12.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 16 entries, 0 to 15
Data columns (total 36 columns):
Unnamed: 0      16 non-null int64
Team            16 non-null object
Goals           16 non-null int64
Shots on target 16 non-null int64
Shots off target 16 non-null int64
Shooting Accuracy 16 non-null object
% Goals-to-shots 16 non-null object
Total shots (inc. Blocked) 16 non-null int64
Hit Woodwork    16 non-null int64
Penalty goals   16 non-null int64
Penalties not scored 16 non-null int64
Headed goals    16 non-null int64
Passes          16 non-null int64
Passes completed 16 non-null int64
Passing Accuracy 16 non-null object
Touches         16 non-null int64
Crosses         16 non-null int64
Dribbles        16 non-null int64
Corners Taken   16 non-null int64
Tackles         16 non-null int64
Clearances      16 non-null int64
Interceptions   16 non-null int64
Clearances off line 15 non-null float64
Clean Sheets    16 non-null int64
Blocks         16 non-null int64
Goals conceded  16 non-null int64
Saves made      16 non-null int64
Saves-to-shots ratio 16 non-null object
Fouls Won       16 non-null int64
Fouls Conceded  16 non-null int64
Offsides        16 non-null int64
```

```
Yellow Cards      16 non-null int64
Red Cards         16 non-null int64
Subs on           16 non-null int64
Subs off          16 non-null int64
Players Used      16 non-null int64
dtypes: float64(1), int64(30), object(5)
memory usage: 4.6+ KB
```

```
# Câu 4: Tạo 1 data frame mới từ euro12 có tên là discipline chỉ chứa 3 cột 'Team', 'Yellow Cards', 'Red Cards'
discipline = euro12[['Team', 'Yellow Cards', 'Red Cards']]
discipline
```

	Team	Yellow Cards	Red Cards
--	------	--------------	-----------

```
# Câu 5: Sắp xếp discipline giảm dần theo 2 cột 'Red Cards', 'Yellow Cards'
discipline.sort_values(['Red Cards', 'Yellow Cards'], ascending = False)
```

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

```
# Câu 6a: Tính trung bình Yellow Cards
discipline[['Yellow Cards']].mean()
```

Yellow Cards 7.4375
dtype: float64

Câu 6b: Lọc các đội đã ghi hơn 6 bàn thắng
euro12[euro12.Goals > 6]

	Unnamed: 0	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	...	Saves made	Saves-to-shots ratio	Fouls Won
5	5	Germany	10	32	32	47.8%	15.6%	80	2	1	...	10	62.6%	63
13	13	Spain	12	42	33	55.9%	16.0%	100	0	1	...	15	93.8%	102

2 rows × 36 columns

Câu 7: In các đội mà tên bắt đầu bằng 'G' (gợi ý: dùng str.startswith())
euro12[euro12.Team.str.startswith('G')]

	Unnamed: 0	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals-to-shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	...	Saves made	Saves-to-shots ratio	Fouls Won
5	5	Germany	10	32	32	47.8%	15.6%	80	2	1	...	10	62.6%	63
6	6	Greece	5	8	18	30.7%	19.2%	32	1	1	...	13	65.1%	67

2 rows × 36 columns

Câu 8: In 7 cột đầu của euro12
euro12.iloc[:, 0:7]

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

Câu 9: In tất cả các cột, trừ 3 cột cuối
euro12.iloc[:, :-3]

	Unnamed: 0	Team	Goals	Shots on target	Shots off target	Shooting Accuracy	% Goals- to- shots	Total shots (inc. Blocked)	Hit Woodwork	Penalty goals	...	Clean Sheets	Blocks	col
0	0	Croatia	4	13	12	51.9%	16.0%	32	0	0	...	0	10	
1	1	Czech Republic	4	13	18	41.9%	12.9%	39	0	0	...	1	10	
2	2	Denmark	4	10	10	50.0%	20.0%	27	1	0	...	1	10	
3	3	England	5	11	18	50.0%	17.2%	40	0	0	...	2	29	
4	4	France	3	22	24	37.9%	6.5%	65	1	0	...	1	7	
5	5	Germany	10	32	32	47.8%	15.6%	80	2	1	...	1	11	
6	6	Greece	5	8	18	30.7%	19.2%	32	1	1	...	1	23	
7	7	Italy	6	34	45	43.0%	7.5%	110	2	0	...	2	18	
8	8	Netherlands	2	12	36	25.0%	4.1%	60	2	0	...	0	9	
9	9	Poland	2	15	23	39.4%	5.2%	48	0	0	...	0	8	
10	10	Portugal	6	22	42	34.3%	9.3%	82	6	0	...	2	11	
11	11	Republic of Ireland	1	7	12	36.8%	5.2%	28	0	0	...	0	23	
12	12	Russia	5	9	31	22.5%	12.5%	59	2	0	...	0	8	

Câu 10: In các cột Team, Goals, Shooting Accuracy, Yellow Cards, Red Cards
euro12[['Team', 'Goals', 'Shooting Accuracy', 'Yellow Cards', 'Red Cards']]

	Team	Goals	Shooting Accuracy	Yellow Cards	Red Cards
0	Croatia	4	51.9%	9	0
1	Czech Republic	4	41.9%	7	0
2	Denmark	4	50.0%	4	0
3	England	5	50.0%	5	0
4	France	3	37.9%	6	0
5	Germany	10	47.8%	4	0
6	Greece	5	30.7%	9	1
7	Italy	6	43.0%	16	0
8	Netherlands	2	25.0%	5	0
9	Poland	2	39.4%	7	1
10	Portugal	6	34.3%	12	0
11	Republic of Ireland	1	36.8%	6	1

Câu 11: In các cột chỉ hiển thị 'Team','Shooting Accuracy' từ 'England', 'Italy', '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%

