jupyter 1 Last Checkpoint: 7 hours ago (autosaved)

Logout

File | Edit | View | Insert | Cell | Kernel | Help

Not Trusted | Python 3 (ipykernel) ○

```
mencari data di kaggle
```

```
In [1]: !kaggle datasets list -s "goal leagues"
```

```
ref                                              title                                            size   lastUpdated
downloadCount  voteCount  usabilityRating
--------------------------------------------     -------------------------------------------      -----  --------------
-----  -------------  ---------  ---------------
slehkyi/extended-football-stats-for-european-leagues-xg  Football Data: Expected Goals and Other Metrics  1MB    2020-08-02 17:
28:39          5751       150  1.0
vivovinco/nba-player-stats                       2021-2022 NBA Player Stats                       48KB   2022-06-18 07:
52:52          6535       122  1.0
sanjeetsinghnaik/most-expensive-footballers-2021 Most Expensive Footballers 2021                  13KB   2021-12-03 20:
28:03          1940       46  1.0
shreyanshkhandelwal/goal-dataset-top-5-european-leagues  Goal Dataset - Top 5 European Leagues    174KB  2021-05-23 21:
20:09          4005       21  0.5294118
secareanualin/football-events                    Football Events                                  21MB   2017-01-25 01:
19:19          25804      629  0.7647059
vivovinco/20222023-nba-player-stats-regular      2022-2023 NBA Player Stats                       33KB   2023-04-23 19:
56:49          2163       55  1.0
vivovinco/20212022-football-team-stats           2021-2022 Football Team Stats                    6KB    2022-05-29 07:
48:12          1802       66  1.0
vivovinco/20212022-football-player-stats         2021-2022 Football Player Stats                  627KB  2022-05-29 12:
13:07          3620       101  1.0
cuecacuela/england-football-matches              England Football Matches                         389KB  2023-04-25 16:
08:27          565        27  1.0
yamaerenay/most-popular-soccer-leagues           Most Popular Soccer Leagues                      30KB   2020-08-01 16:
59:30          205        6  1.0
vivovinco/20222023-football-player-stats         2022-2023 Football Player Stats                  470KB  2023-02-12 17:
01:33          1961       60  1.0
chaibapat/fantasy-premier-league                 Fantasy Premier League - 2016/2017               476MB  2017-05-16 18:
56:26          1545       32  0.85294116
sanjeetsinghnaik/premier-league-matches-20142020 Premier League Matches 2014-2020                 227KB  2021-12-21 19:
08:15          423        22  0.9117647
mcarujo/european-football-season-202223          European Football Season 2022/23                 909KB  2023-05-01 13:
13:58          702        36  0.8235294
eladsi1/football-games-odds                      Football Matches Odds                            5MB    2018-10-15 08:
15:14          904        28  0.8235294
sanjeetsinghnaik/la-liga-match-data              La Liga Match Data                               226KB  2022-01-02 18:
21:11          460        21  0.9117647
andrew1456/202122-nba-season-per-game-stats      2021-22 NBA Season Per Game Stats                29KB   2022-09-01 15:
34:59          121        11  0.9411765
vivovinco/20222023-football-team-stats           2022-2023 Football Team Stats                    6KB    2023-04-23 19:
50:32          537        24  1.0
jashsheth5/indepth-soccer-statistics-xg-xa-and-more  In-depth soccer statistics: xG, xA and more  1MB    2020-09-26 05:
36:58          389        9  0.9411765
cuecacuela/brazilian-football-matches            Brazilian Football Matches                       275KB  2023-04-25 12:
29:45          77         12  0.8235294
```

#mendownload data yang di inginkan

```
In [2]: !kaggle datasets download shreyanshkhandelwal/goal-dataset-top-5-european-leagues
```

```
goal-dataset-top-5-european-leagues.zip: Skipping, found more recently modified local copy (use --force to force download)
```

```
In [ ]: menampilkan data
```

```
In [6]: import pandas as pd

        g = pd.read_csv(r'C:\Users\AinXD\.kaggle\Serie_A-goalScorer(20-21).csv')
```

```
In [7]: g
```

Out[7]:

	Unnamed: 0	id	player_name	games	time	goals	xG	assists	xA	shots	key_passes	yellow_cards	red_cards	position	team_title	np
0	0	2371.0	Cristiano Ronaldo	NaN	2807.0	29.0	29.838081	3.0	3.854639	167.0	36.0	3.0	0.0	F S	Juventus	23.
1	1	594.0	Romelu Lukaku	NaN	2885.0	24.0	23.425768	11.0	8.387602	97.0	53.0	4.0	0.0	F S	Inter	18.
2	2	1229.0	Luis Muriel	36.0	1434.0	22.0	16.692218	8.0	4.887805	86.0	46.0	0.0	0.0	F M S	Atalanta	20.
3	3	7084.0	Dusan Vlahovic	37.0	2945.0	21.0	18.476590	3.0	2.296164	86.0	18.0	1.0	0.0	F S	Fiorentina	15.
4	4	1209.0	Ciro Immobile	35.0	2887.0	20.0	19.771116	6.0	7.202586	119.0	54.0	3.0	1.0	F S	Lazio	16.
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
589	589	9503.0	Kacper Urbanski	1.0	1.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	0.0	S	Bologna	0.
590	590	9518.0	Edoardo Bove	1.0	8.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	0.0	S	Roma	0.
591	591	9539.0	Antonio Raimondo	1.0	1.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	0.0	S	Bologna	0.
592	592	9550.0	Yayah Kallon	1.0	42.0	0.0	0.122062	0.0	0.021998	2.0	1.0	0.0	0.0	S	Genoa	0.
593	593	9559.0	Marco Bertini	1.0	3.0	0.0	0.000000	0.0	0.000000	0.0	0.0	0.0	0.0	S	Lazio	0.

594 rows × 19 columns

```
In [ ]: search team lazio dengan posisis CF
```

```
In [8]: g.loc[(g['team_title'] == 'Lazio') & (g['position'] == 'F S')].head()
```

Out[8]:

	Unnamed: 0	id	player_name	games	time	goals	xG	assists	xA	shots	key_passes	yellow_cards	red_cards	position	team_title	np
4	4	1209.0	Ciro Immobile	35.0	2887.0	20.0	19.771116	6.0	7.202586	119.0	54.0	3.0	1.0	F S	Lazio	16.
35	35	1809.0	Joaquín Correa	28.0	1948.0	8.0	7.407989	3.0	2.480826	54.0	29.0	1.0	1.0	F S	Lazio	7.
36	36	2116.0	Felipe Caicedo	25.0	915.0	8.0	5.328443	0.0	0.725062	31.0	12.0	5.0	0.0	F S	Lazio	7.
277	277	9002.0	Vedat Muriqi	27.0	809.0	1.0	3.270341	0.0	0.618362	31.0	9.0	1.0	0.0	F S	Lazio	1.

```
In [ ]: mencari data pemain dengan goal lebih dari 20
```

```
In [9]: g.loc[(g['goals'] > 20)].head()
```

Out[9]:

	Unnamed: 0	id	player_name	games	time	goals	xG	assists	xA	shots	key_passes	yellow_cards	red_cards	position	team_title	npg

0	0	2371.0	Cristiano Ronaldo	NaN	2807.0	29.0	29.838081	3.0	3.854639	167.0	36.0	3.0	0.0	F S	Juventus	23.0
1	1	594.0	Romelu Lukaku	NaN	2885.0	24.0	23.425768	11.0	8.387602	97.0	53.0	4.0	0.0	F S	Inter	18.0
2	2	1229.0	Luis Muriel	36.0	1434.0	22.0	16.692218	8.0	4.887805	86.0	46.0	0.0	0.0	F M S	Atalanta	20.0
3	3	7084.0	Dusan Vlahovic	37.0	2945.0	21.0	18.476590	3.0	2.296164	86.0	18.0	1.0	0.0	F S	Fiorentina	15.0

In [ ]: deteksi missing data

In [10]:
```python
print(g.isna().sum())
```

```
Unnamed: 0       0
id               24
player_name      24
games            26
time             24
goals            24
xG               24
assists          24
xA               24
shots            24
key_passes       24
yellow_cards     24
red_cards        24
position         24
team_title       24
npg              24
npxG             24
xGChain          24
xGBuildup;       24
dtype: int64
```

In [ ]:
```python
print( g.isnull())
```

In [ ]:

In [ ]:
```python
g.fillna(10)
```

In [ ]:
```python
g
```

In [ ]:

In [ ]: