# Data Science project fifa

September 27, 2021

```python
[1]: import pandas as pd
     from matplotlib import pyplot as plt
     import seaborn as sns
     import numpy as np
     import sklearn as sks
```

```python
[4]: fifa=pd.read_csv('players_20.csv')
```

```python
[5]: fifa.head()
```

```
[5]:    sofifa_id                                        player_url  \
     0     158023  https://sofifa.com/player/158023/lionel-messi/…
     1      20801  https://sofifa.com/player/20801/c-ronaldo-dos-…
     2     190871  https://sofifa.com/player/190871/neymar-da-sil…
     3     200389  https://sofifa.com/player/200389/jan-oblak/20/…
     4     183277  https://sofifa.com/player/183277/eden-hazard/2…

               short_name                          long_name  age         dob  \
     0           L. Messi        Lionel Andrés Messi Cuccittini   32  1987-06-24
     1  Cristiano Ronaldo  Cristiano Ronaldo dos Santos Aveiro   34  1985-02-05
     2          Neymar Jr       Neymar da Silva Santos Junior   27  1992-02-05
     3          J. Oblak                          Jan Oblak   26  1993-01-07
     4         E. Hazard                        Eden Hazard   28  1991-01-07

        height_cm  weight_kg nationality                 club  …   lwb   ldm  \
     0        170         72   Argentina          FC Barcelona  …  68+2  66+2
     1        187         83    Portugal              Juventus  …  65+3  61+3
     2        175         68      Brazil  Paris Saint-Germain  …  66+3  61+3
     3        188         87    Slovenia      Atlético Madrid  …   NaN   NaN
     4        175         74     Belgium          Real Madrid  …  66+3  63+3

         cdm   rdm   rwb    lb   lcb    cb   rcb    rb
     0  66+2  66+2  68+2  63+2  52+2  52+2  52+2  63+2
     1  61+3  61+3  65+3  61+3  53+3  53+3  53+3  61+3
     2  61+3  61+3  66+3  61+3  46+3  46+3  46+3  61+3
     3   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN
     4  63+3  63+3  66+3  61+3  49+3  49+3  49+3  61+3
```

```
[5 rows x 104 columns]
```

```python
[6]:  for terms in fifa.columns:
          print(terms)
```

```
sofifa_id
player_url
short_name
long_name
age
dob
height_cm
weight_kg
nationality
club
overall
potential
value_eur
wage_eur
player_positions
preferred_foot
international_reputation
weak_foot
skill_moves
work_rate
body_type
real_face
release_clause_eur
player_tags
team_position
team_jersey_number
loaned_from
joined
contract_valid_until
nation_position
nation_jersey_number
pace
shooting
passing
dribbling
defending
physic
gk_diving
gk_handling
gk_kicking
gk_reflexes
gk_speed
gk_positioning
```

```
player_traits
attacking_crossing
attacking_finishing
attacking_heading_accuracy
attacking_short_passing
attacking_volleys
skill_dribbling
skill_curve
skill_fk_accuracy
skill_long_passing
skill_ball_control
movement_acceleration
movement_sprint_speed
movement_agility
movement_reactions
movement_balance
power_shot_power
power_jumping
power_stamina
power_strength
power_long_shots
mentality_aggression
mentality_interceptions
mentality_positioning
mentality_vision
mentality_penalties
mentality_composure
defending_marking
defending_standing_tackle
defending_sliding_tackle
goalkeeping_diving
goalkeeping_handling
goalkeeping_kicking
goalkeeping_positioning
goalkeeping_reflexes
ls
st
rs
lw
lf
cf
rf
rw
lam
cam
ram
lm
lcm
```

```
cm
rcm
rm
lwb
ldm
cdm
rdm
rwb
lb
lcb
cb
rcb
rb
```

[7]: `fifa.shape`

[7]: (18278, 104)

[8]: `fifa['nationality'].value_counts`

[8]: <bound method IndexOpsMixin.value_counts of 0          Argentina
    1          Portugal
    2            Brazil
    3          Slovenia
    4           Belgium
                 …
    18273      China PR
    18274      China PR
    18275      China PR
    18276      China PR
    18277      China PR
    Name: nationality, Length: 18278, dtype: object>
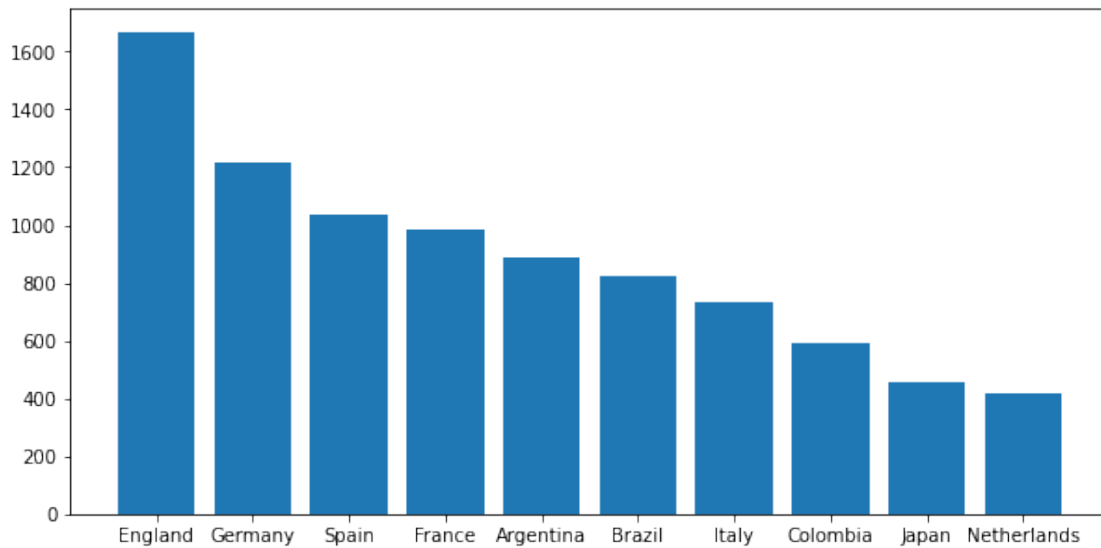
[12]: `fifa['nationality'].value_counts()[0:10]`

[12]: England        1667
    Germany        1216
    Spain          1035
    France          984
    Argentina       886
    Brazil          824
    Italy           732
    Colombia        591
    Japan           453
    Netherlands     416
    Name: nationality, dtype: int64

[14]: `list(fifa['nationality'].value_counts()[0:10].keys())`

```
[14]: ['England',
       'Germany',
       'Spain',
       'France',
       'Argentina',
       'Brazil',
       'Italy',
       'Colombia',
       'Japan',
       'Netherlands']
```

```
[16]: plt.figure(figsize=(10,5))
      plt.bar(list(fifa['nationality'].value_counts()[0:10].
       ↪keys()),fifa['nationality'].value_counts()[0:10])
      plt.show()
```



```
[21]: player_salary=fifa[['short_name','wage_eur']]
```

```
[22]: player_salary.head()
```

```
[22]:         short_name  wage_eur
      0          L. Messi    565000
      1  Cristiano Ronaldo    405000
      2          Neymar Jr    290000
      3          J. Oblak    125000
      4          E. Hazard    470000
```
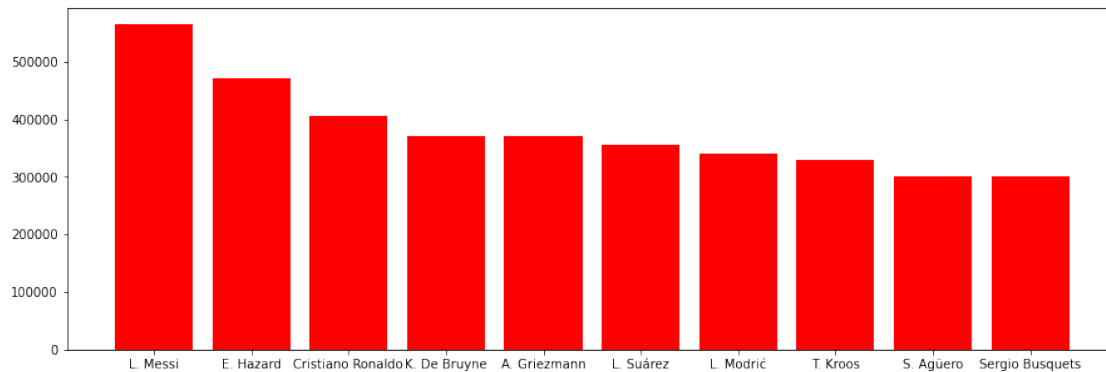
```
[24]: player_salary=player_salary.sort_values(by=['wage_eur'], ascending=False)
```

```
[25]: player_salary.head()
```

```
[25]:              short_name  wage_eur
       0              L. Messi    565000
       4             E. Hazard    470000
       1     Cristiano Ronaldo    405000
       5          K. De Bruyne    370000
       22         A. Griezmann    370000
```

```
[33]: plt.figure(figsize=(15,5))
       plt.bar(list(player_salary['short_name'])[0:
       ↪10],list(player_salary['wage_eur'])[0:10], color='red')
       plt.show()
```



```
[34]: fifa['nationality']=='Germany'
```

```
[34]: 0         False
       1         False
       2         False
       3         False
       4         False
                 …
       18273     False
       18274     False
       18275     False
       18276     False
       18277     False
       Name: nationality, Length: 18278, dtype: bool
```

```
[35]: #working for the country of germany only
      Germany=fifa[fifa['nationality']=='Germany']
```

```
[36]: Germany.head()
```

```
[36]:     sofifa_id                                          player_url  \
      6      192448  https://sofifa.com/player/192448/marc-andre-te…
      31     167495  https://sofifa.com/player/167495/manuel-neuer/…
      36     182521  https://sofifa.com/player/182521/toni-kroos/20…
      37     188350  https://sofifa.com/player/188350/marco-reus/20…
      49     178603  https://sofifa.com/player/178603/mats-hummels/…

            short_name              long_name  age         dob  height_cm  \
      6   M. ter Stegen  Marc-André ter Stegen   27  1992-04-30        187
      31      M. Neuer           Manuel Neuer   33  1986-03-27        193
      36      T. Kroos             Toni Kroos   29  1990-01-04        183
      37       M. Reus             Marco Reus   30  1989-05-31        180
      49    M. Hummels           Mats  Hummels   30  1988-12-16        191

          weight_kg nationality               club  …  lwb   ldm   cdm   rdm  \
      6          85     Germany        FC Barcelona  …  NaN   NaN   NaN   NaN
      31         92     Germany  FC Bayern München  …  NaN   NaN   NaN   NaN
      36         76     Germany         Real Madrid  …  75+3  79+3  79+3  79+3
      37         71     Germany   Borussia Dortmund  …  69+3  65+3  65+3  65+3
      49         94     Germany   Borussia Dortmund  …  77+3  83+3  83+3  83+3

           rwb    lb   lcb    cb   rcb    rb
      6     NaN   NaN   NaN   NaN   NaN   NaN
      31    NaN   NaN   NaN   NaN   NaN   NaN
      36   75+3  72+3  69+3  69+3  69+3  72+3
      37   69+3  64+3  55+3  55+3  55+3  64+3
      49   77+3  78+3  84+3  84+3  84+3  78+3

      [5 rows x 104 columns]
```

```
[38]: Germany.sort_values(by=['height_cm'],ascending=False).head()
```

```
[38]:        sofifa_id                                          player_url  \
      8016      236831  https://sofifa.com/player/236831/aaron-seydel/…
      1217      200212  https://sofifa.com/player/200212/michael-esser…
      1389      199833  https://sofifa.com/player/199833/lars-unnersta…
      11859     167437  https://sofifa.com/player/167437/dominik-stroh…
      13576     239746  https://sofifa.com/player/239746/lukas-watkowi…

                 short_name           long_name  age         dob  height_cm  \
      8016        A. Seydel         Aaron Seydel   23  1996-02-07        199
      1217         M. Esser        Michael Esser   31  1987-11-22        198
```

```
1389      L. Unnerstall        Lars Unnerstall   28  1990-07-20           198
11859  D. Stroh-Engel  Dominik Stroh-Engel   33  1985-11-27           197
13576      L. Watkowiak       Lukas Watkowiak   23  1996-03-06           197

       weight_kg nationality                 club  …   lwb   ldm   cdm   rdm  \
8016          90     Germany       1. FSV Mainz 05  …  46+2  43+2  43+2  43+2
1217          97     Germany          Hannover 96  …   NaN   NaN   NaN   NaN
1389         103     Germany                  PSV  …   NaN   NaN   NaN   NaN
11859         94     Germany   SpVgg Unterhaching  …  39+2  45+2  45+2  45+2
13576        103     Germany   SV Wehen Wiesbaden  …   NaN   NaN   NaN   NaN

        rwb    lb   lcb    cb   rcb    rb
8016   46+2  44+2  43+2  43+2  43+2  44+2
1217    NaN   NaN   NaN   NaN   NaN   NaN
1389    NaN   NaN   NaN   NaN   NaN   NaN
11859  39+2  39+2  47+2  47+2  47+2  39+2
13576   NaN   NaN   NaN   NaN   NaN   NaN

[5 rows x 104 columns]
```

```
[39]: Germany.sort_values(by=['weight_kg'],ascending=False).head()
```

```
[39]:        sofifa_id                                        player_url  \
      13576     239746  https://sofifa.com/player/239746/lukas-watkowi…
      1389      199833  https://sofifa.com/player/199833/lars-unnersta…
      518       179783  https://sofifa.com/player/179783/ralf-fahrmann…
      1217      200212  https://sofifa.com/player/200212/michael-esser…
      77        212190  https://sofifa.com/player/212190/niklas-sule/2…

              short_name        long_name  age         dob  height_cm  weight_kg  \
      13576  L. Watkowiak  Lukas Watkowiak   23  1996-03-06        197        103
      1389   L. Unnerstall  Lars Unnerstall   28  1990-07-20        198        103
      518      R. Fährmann    Ralf Fährmann   30  1988-09-27        197         98
      1217        M. Esser    Michael Esser   31  1987-11-22        198         97
      77          N. Süle       Niklas Süle   23  1995-09-03        195         97

            nationality                 club  …   lwb   ldm   cdm   rdm   rwb  \
      13576     Germany   SV Wehen Wiesbaden  …   NaN   NaN   NaN   NaN   NaN
      1389      Germany                  PSV  …   NaN   NaN   NaN   NaN   NaN
      518       Germany         Norwich City  …   NaN   NaN   NaN   NaN   NaN
      1217      Germany          Hannover 96  …   NaN   NaN   NaN   NaN   NaN
      77        Germany    FC Bayern München  …  71+2  77+2  77+2  77+2  71+2

             lb   lcb    cb   rcb    rb
      13576  NaN   NaN   NaN   NaN   NaN
      1389   NaN   NaN   NaN   NaN   NaN
      518    NaN   NaN   NaN   NaN   NaN
```

```
1217     NaN    NaN    NaN    NaN    NaN
77       75+2   83+2   83+2   83+2   75+2

[5 rows x 104 columns]
```

[40]: `Germany.sort_values(by=['wage_eur'],ascending=False).head()`

[40]:
```
      sofifa_id                                    player_url  \
36       182521  https://sofifa.com/player/182521/toni-kroos/20…
6        192448  https://sofifa.com/player/192448/marc-andre-te…
55       222492  https://sofifa.com/player/222492/leroy-sane/20…
146      186942  https://sofifa.com/player/186942/ilkay-gundoga…
70       189596  https://sofifa.com/player/189596/thomas-muller…

         short_name              long_name  age         dob  height_cm  \
36          T. Kroos             Toni Kroos   29  1990-01-04        183
6     M. ter Stegen  Marc-André ter Stegen   27  1992-04-30        187
55          L. Sané             Leroy Sané   23  1996-01-11        183
146    I. Gündoğan         İlkay Gündoğan   28  1990-10-24        180
70        T. Müller          Thomas Müller   29  1989-09-13        186

      weight_kg nationality               club  …   lwb   ldm   cdm   rdm  \
36           76     Germany        Real Madrid  …  75+3  79+3  79+3  79+3
6            85     Germany       FC Barcelona  …   NaN   NaN   NaN   NaN
55           75     Germany    Manchester City  …  65+2  59+2  59+2  59+2
146          80     Germany    Manchester City  …  76+3  79+3  79+3  79+3
70           75     Germany  FC Bayern München  …  69+3  68+3  68+3  68+3

       rwb    lb   lcb    cb   rcb    rb
36    75+3  72+3  69+3  69+3  69+3  72+3
6      NaN   NaN   NaN   NaN   NaN   NaN
55    65+2  61+2  52+2  52+2  52+2  61+2
146   76+3  74+3  72+3  72+3  72+3  74+3
70    69+3  67+3  62+3  62+3  62+3  67+3

[5 rows x 104 columns]
```

[42]:
```
#we cant see the wages in the dataset so we have to do some more work to make
↪it visible
Germany[['long_name','wage_eur']].sort_values(by=['wage_eur'],ascending=False).
↪head()
```

[42]:
```
                 long_name  wage_eur
36              Toni Kroos    330000
6    Marc-André ter Stegen    250000
55              Leroy Sané    195000
146         İlkay Gündoğan    180000
```

```
70         Thomas Müller     170000
```

```
[52]: #extracting the player the best shooting capabilities
      player_shooting=Germany[['short_name','shooting']]
```

```
[58]: player_shooting.head()
```

```
[58]:        short_name  shooting
      6   M. ter Stegen       NaN
      31       M. Neuer       NaN
      36       T. Kroos      80.0
      37        M. Reus      88.0
      49    M. Hummels      58.0
```

```
[50]: player_shooting.sort_values(by=['shooting'], ascending=False).head()
```

```
[50]:        short_name  shooting
      37        M. Reus      88.0
      276   K. Volland      84.0
      70     T. Müller      83.0
      217     M. Kruse      83.0
      120    S. Gnabry      82.0
```

```
[67]: #defending
      player_defending=fifa[['short_name','nationality','club','defending']]
```

```
[68]: player_defending.head()
```

```
[68]:          short_name nationality                 club  defending
      0           L. Messi   Argentina         FC Barcelona       39.0
      1  Cristiano Ronaldo    Portugal             Juventus       35.0
      2           Neymar Jr      Brazil  Paris Saint-Germain       32.0
      3           J. Oblak    Slovenia      Atlético Madrid        NaN
      4           E. Hazard     Belgium          Real Madrid       35.0
```

```
[71]: player_defending.sort_values(by=['defending'], ascending=False).head()
```

```
[71]:        short_name   nationality                 club  defending
      16  G. Chiellini         Italy             Juventus       90.0
      7    V. van Dijk   Netherlands            Liverpool       90.0
      11  K. Koulibaly       Senegal               Napoli       89.0
      35      D. Godín       Uruguay                Inter       89.0
      49    M. Hummels       Germany    Borussia Dortmund       89.0
```

```
[75]: #extrecting out the player belonging to the real madrid
      madrid=fifa[fifa['club']=='Real Madrid']
```

```
[76]: madrid.head()
```

```
[76]:      sofifa_id                                          player_url  \
     4       183277   https://sofifa.com/player/183277/eden-hazard/2…
     8       177003   https://sofifa.com/player/177003/luka-modric/2…
     18      155862   https://sofifa.com/player/155862/sergio-ramos-…
     28      192119   https://sofifa.com/player/192119/thibaut-court…
     36      182521   https://sofifa.com/player/182521/toni-kroos/20…

         short_name            long_name  age         dob  height_cm  weight_kg  \
     4      E. Hazard         Eden Hazard   28  1991-01-07        175         74
     8      L. Modrić         Luka Modrić   33  1985-09-09        172         66
     18   Sergio Ramos  Sergio Ramos García   33  1986-03-30        184         82
     28   T. Courtois     Thibaut Courtois   27  1992-05-11        199         96
     36      T. Kroos          Toni Kroos   29  1990-01-04        183         76

        nationality         club  …   lwb   ldm   cdm   rdm   rwb    lb   lcb  \
     4      Belgium  Real Madrid  …  66+3  63+3  63+3  63+3  66+3  61+3  49+3
     8      Croatia  Real Madrid  …  81+3  81+3  81+3  81+3  81+3  79+3  72+3
     18       Spain  Real Madrid  …  80+3  84+3  84+3  84+3  80+3  82+3  87+3
     28     Belgium  Real Madrid  …   NaN   NaN   NaN   NaN   NaN   NaN   NaN
     36     Germany  Real Madrid  …  75+3  79+3  79+3  79+3  75+3  72+3  69+3

           cb   rcb    rb
     4   49+3  49+3  61+3
     8   72+3  72+3  79+3
     18  87+3  87+3  82+3
     28   NaN   NaN   NaN
     36  69+3  69+3  72+3

     [5 rows x 104 columns]
```

```
[77]:  madrid.sort_values(by=['wage_eur'],ascending=False).head()
```

```
[77]:      sofifa_id                                          player_url  \
     4       183277   https://sofifa.com/player/183277/eden-hazard/2…
     8       177003   https://sofifa.com/player/177003/luka-modric/2…
     36      182521   https://sofifa.com/player/182521/toni-kroos/20…
     18      155862   https://sofifa.com/player/155862/sergio-ramos-…
     46      165153   https://sofifa.com/player/165153/karim-benzema…

         short_name            long_name  age         dob  height_cm  weight_kg  \
     4      E. Hazard         Eden Hazard   28  1991-01-07        175         74
     8      L. Modrić         Luka Modrić   33  1985-09-09        172         66
     36      T. Kroos          Toni Kroos   29  1990-01-04        183         76
     18   Sergio Ramos  Sergio Ramos García   33  1986-03-30        184         82
     46    K. Benzema        Karim Benzema   31  1987-12-19        185         81

        nationality         club  …   lwb   ldm   cdm   rdm   rwb    lb   lcb  \
```

```
4       Belgium  Real Madrid  …  66+3  63+3  63+3  63+3  66+3  61+3  49+3
8       Croatia  Real Madrid  …  81+3  81+3  81+3  81+3  81+3  79+3  72+3
36      Germany  Real Madrid  …  75+3  79+3  79+3  79+3  75+3  72+3  69+3
18        Spain  Real Madrid  …  80+3  84+3  84+3  84+3  80+3  82+3  87+3
46       France  Real Madrid  …  62+3  62+3  62+3  62+3  62+3  58+3  54+3


        cb    rcb     rb
4    49+3   49+3   61+3
8    72+3   72+3   79+3
36   69+3   69+3   72+3
18   87+3   87+3   82+3
46   54+3   54+3   58+3

[5 rows x 104 columns]
```

`[80]:` `madrid.sort_values(by=['age'],ascending=False).head()`

```
[80]:      sofifa_id                                      player_url  \
     18      155862  https://sofifa.com/player/155862/sergio-ramos-…
     8       177003  https://sofifa.com/player/177003/luka-modric/2…
     53      193041  https://sofifa.com/player/193041/keylor-navas/…
     46      165153  https://sofifa.com/player/165153/karim-benzema…
     101     176676  https://sofifa.com/player/176676/marcelo-vieir…

           short_name                      long_name  age         dob  height_cm  \
     18   Sergio Ramos        Sergio Ramos García      33  1986-03-30        184
     8      L. Modrić                Luka Modrić        33  1985-09-09        172
     53      K. Navas               Keylor Navas       32  1986-12-15        185
     46    K. Benzema              Karim Benzema       31  1987-12-19        185
     101      Marcelo  Marcelo Vieira da Silva Júnior  31  1988-05-12        174

           weight_kg nationality         club  …   lwb   ldm   cdm   rdm   rwb  \
     18            82       Spain  Real Madrid  …  80+3  84+3  84+3  84+3  80+3
     8             66     Croatia  Real Madrid  …  81+3  81+3  81+3  81+3  81+3
     53            80  Costa Rica  Real Madrid  …   NaN   NaN   NaN   NaN   NaN
     46            81      France  Real Madrid  …  62+3  62+3  62+3  62+3  62+3
     101           75      Brazil  Real Madrid  …  82+3  81+3  81+3  81+3  82+3

           lb   lcb    cb   rcb    rb
     18   82+3  87+3  87+3  87+3  82+3
     8    79+3  72+3  72+3  72+3  79+3
     53    NaN   NaN   NaN   NaN   NaN
     46   58+3  54+3  54+3  54+3  58+3
     101  81+3  79+3  79+3  79+3  81+3

[5 rows x 104 columns]
```

```
[79]: madrid.sort_values(by=['shooting'],ascending=False).head()
```

```
[79]:      sofifa_id                                        player_url  \
      100     173731  https://sofifa.com/player/173731/gareth-bale/2…
      106     198710  https://sofifa.com/player/198710/james-rodrigu…
      4       183277  https://sofifa.com/player/183277/eden-hazard/2…
      46      165153  https://sofifa.com/player/165153/karim-benzema…
      161     232432  https://sofifa.com/player/232432/luka-jovic/20…

            short_name                   long_name  age         dob  height_cm  \
      100      G. Bale           Gareth Frank Bale   29  1989-07-16        185
      106  J. Rodríguez  James David Rodríguez Rubio   27  1991-07-12        180
      4      E. Hazard                 Eden Hazard   28  1991-01-07        175
      46    K. Benzema               Karim Benzema   31  1987-12-19        185
      161     L. Jović                  Luka Jović   21  1997-12-23        182

           weight_kg nationality          club  …   lwb   ldm   cdm   rdm   rwb  \
      100         82       Wales  Real Madrid  …  72+3  69+3  69+3  69+3  72+3
      106         75    Colombia  Real Madrid  …  67+3  67+3  67+3  67+3  67+3
      4           74     Belgium  Real Madrid  …  66+3  63+3  63+3  63+3  66+3
      46          81      France  Real Madrid  …  62+3  62+3  62+3  62+3  62+3
      161         85      Serbia  Real Madrid  …  55+2  53+2  53+2  53+2  55+2

            lb   lcb    cb   rcb    rb
      100  70+3  66+3  66+3  66+3  70+3
      106  62+3  57+3  57+3  57+3  62+3
      4    61+3  49+3  49+3  49+3  61+3
      46   58+3  54+3  54+3  54+3  58+3
      161  51+2  47+2  47+2  47+2  51+2

      [5 rows x 104 columns]
```

```
[89]: madrid[madrid['age']<=27]['age'].value_counts().sort_values(ascending=True)
```

```
[89]: 24    1
      25    1
      22    1
      26    2
      23    2
      19    2
      21    3
      18    3
      20    3
      27    5
      Name: age, dtype: int64
```

```
[ ]:
```