

footballdata_analysis

December 6, 2021

```
[1]: #load the required libraries
import pandas as pd
from matplotlib import pyplot as plt
```

```
[6]: #load the file
fifa_data = pd.read_csv("D:\\fifa data.csv")
fifa_data.head()
```

```
[6]:  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  6/24/1987
1  Cristiano Ronaldo  Cristiano Ronaldo dos Santos Aveiro  34  2/5/1985
2      Neymar Jr  Neymar da Silva Santos Junior  27  2/5/1992
3      J. Oblak  Jan Oblak  26  1/7/1993
4      E. Hazard  Eden Hazard  28  1/7/1991

      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]

```
[7]: #printing only columns in this data  
for col in fifa_data:  
    print(col)
```

```
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
```

```
[11]: #printing the shape of this data set
fifa_data.shape
```

```
[11]: (18278, 104)
```

```
[14]: fifa_data['nationality'].value_counts()
```

```
[14]: England      1667
      Germany     1216
      Spain       1035
      France       984
      Argentina    886
      ...
      Ethiopia      1
      Barbados      1
      Vietnam       1
      St Lucia      1
      Chad          1
      Name: nationality, Length: 162, dtype: int64
```

```
[15]: fifa_data['nationality'].value_counts()[0:5]
```

```
[15]: England      1667
      Germany     1216
      Spain       1035
      France       984
      Argentina    886
      Name: nationality, dtype: int64
```

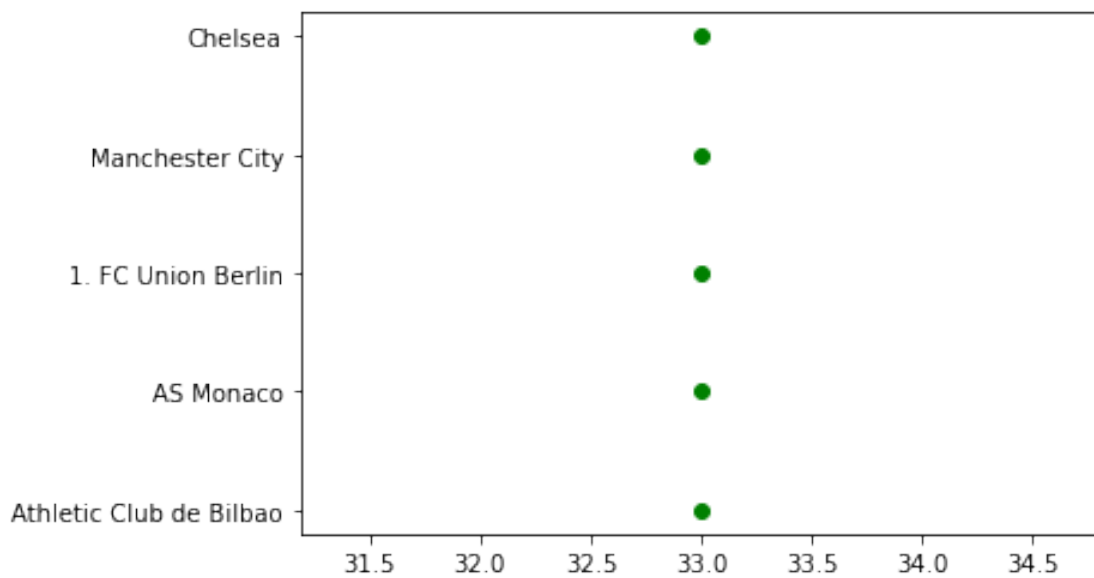
```
[33]: L1=list(fifa_data['club'].value_counts()[0:5])
      L1
```

```
[33]: [33, 33, 33, 33, 33]
```

```
[32]: L2=list(fifa_data['club'].value_counts()[0:5].keys())
      L2
```

```
[32]: ['Athletic Club de Bilbao',
      'AS Monaco',
      '1. FC Union Berlin',
      'Manchester City',
      'Chelsea']
```

```
[34]: #plt.figure(figsize=(5,5))
      plt.scatter(L1,L2,color='g')
      plt.show()
```



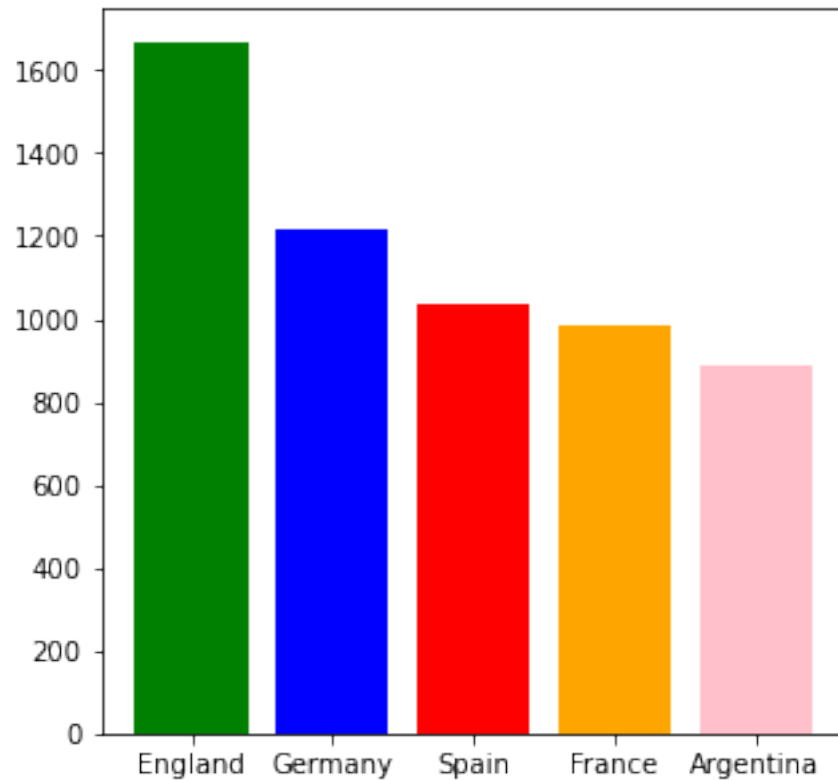
```
[46]: L3=list(fifa_data['nationality'].value_counts()[0:5])
      L3
```

```
[46]: [1667, 1216, 1035, 984, 886]
```

```
[47]: L4 =list( fifa_data['nationality'].value_counts()[0:5].keys())
      L4
```

```
[47]: ['England', 'Germany', 'Spain', 'France', 'Argentina']
```

```
[48]: plt.figure(figsize=(5,5))
      plt.bar(L4,L3,color = ['green','blue','red','orange','pink'])
      plt.show()
```



```
[51]: player_salary = fifa_data[['short_name', 'wage_eur']]
      player_salary.head()
```

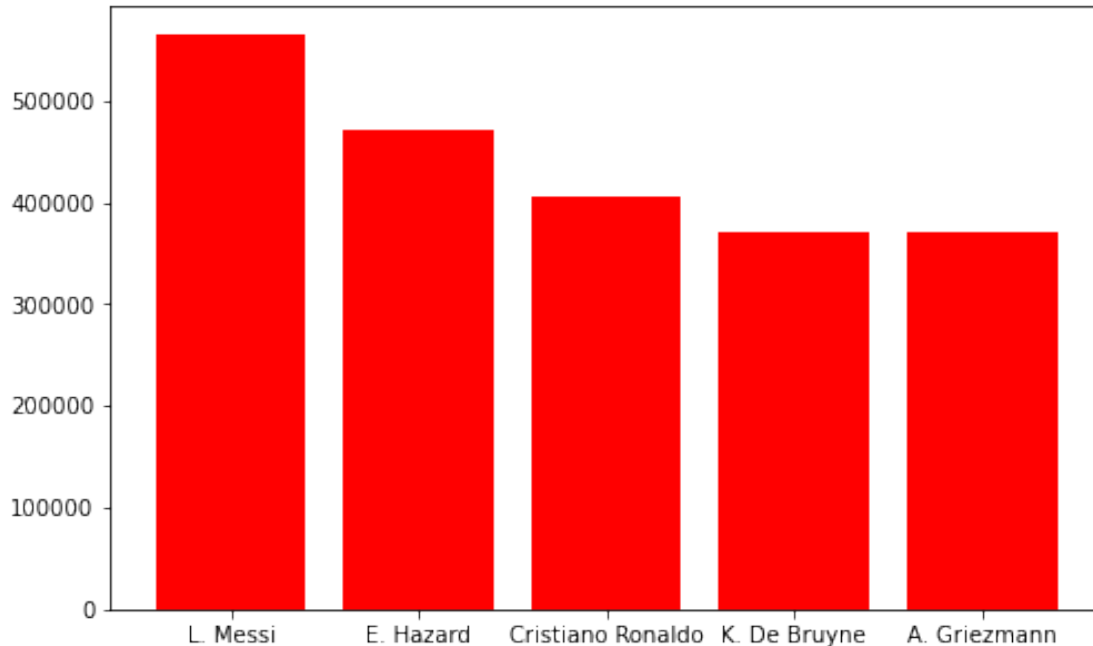
```
[51]:      short_name  wage_eur
0      L. Messi    565000
1  Cristiano Ronaldo  405000
2      Neymar Jr   290000
3      J. Oblak    125000
4      E. Hazard   470000
```

```
[56]: player_salary=player_salary.sort_values(by=['wage_eur'], ascending=False)
```

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

```
[57]:      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
```

```
[59]: plt.figure(figsize=(8,5))
plt.bar((list(player_salary['short_name'])[0:5]),
↪list((player_salary['wage_eur'])[0:5]), color = 'red')
plt.show()
```



```
[61]: #Germany
fifa=fifa_data[fifa_data['nationality']=='Germany']
fifa.head(10)
```

```
[61]:
```

	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/...
55	222492	https://sofifa.com/player/222492/leroy-sane/20...
61	212622	https://sofifa.com/player/212622/joshua-kimmic...
70	189596	https://sofifa.com/player/189596/thomas-muller...
77	212190	https://sofifa.com/player/212190/niklas-sule/2...
109	235790	https://sofifa.com/player/235790/kai-havertz/2...

	short_name	long_name	age	dob	height_cm \
6	M. ter Stegen	Marc-André ter Stegen	27	4/30/1992	187
31	M. Neuer	Manuel Neuer	33	3/27/1986	193
36	T. Kroos	Toni Kroos	29	1/4/1990	183

37	M. Reus	Marco Reus	30	5/31/1989	180
49	M. Hummels	Mats Hummels	30	12/16/1988	191
55	L. Sané	Leroy Sané	23	1/11/1996	183
61	J. Kimmich	Joshua Kimmich	24	2/8/1995	176
70	T. Müller	Thomas Müller	29	9/13/1989	186
77	N. Süle	Niklas Süle	23	9/3/1995	195
109	K. Havertz	Kai Havertz	20	6/11/1999	188

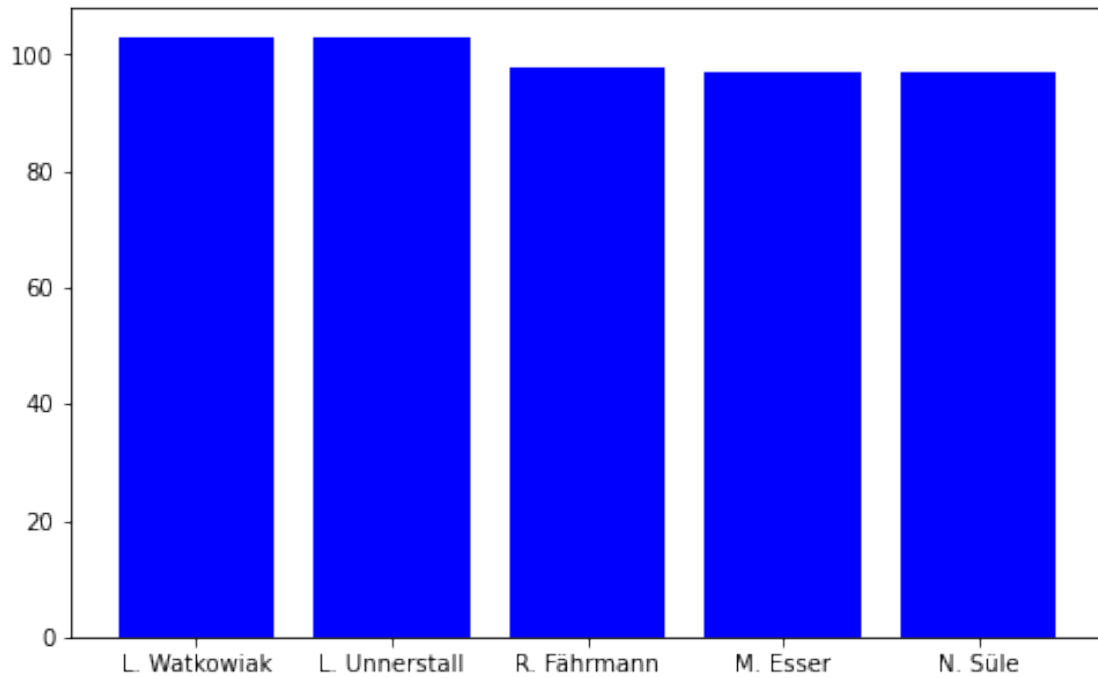
	weight_kg	nationality	club	...	lwb	ldm	cdm	rdm	\
6	85	Germany	FC Barcelona	...	NaN	NaN	NaN	NaN	
31	92	Germany	FC Bayern München	...	NaN	NaN	NaN	NaN	
36	76	Germany	Real Madrid	...	75+3	79+3	79+3	79+3	
37	71	Germany	Borussia Dortmund	...	69+3	65+3	65+3	65+3	
49	94	Germany	Borussia Dortmund	...	77+3	83+3	83+3	83+3	
55	75	Germany	Manchester City	...	65+2	59+2	59+2	59+2	
61	73	Germany	FC Bayern München	...	83+3	82+3	82+3	82+3	
70	75	Germany	FC Bayern München	...	69+3	68+3	68+3	68+3	
77	97	Germany	FC Bayern München	...	71+2	77+2	77+2	77+2	
109	83	Germany	Bayer 04 Leverkusen	...	65+2	63+2	63+2	63+2	

	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
55	65+2	61+2	52+2	52+2	52+2	61+2
61	83+3	81+3	78+3	78+3	78+3	81+3
70	69+3	67+3	62+3	62+3	62+3	67+3
77	71+2	75+2	83+2	83+2	83+2	75+2
109	65+2	62+2	54+2	54+2	54+2	62+2

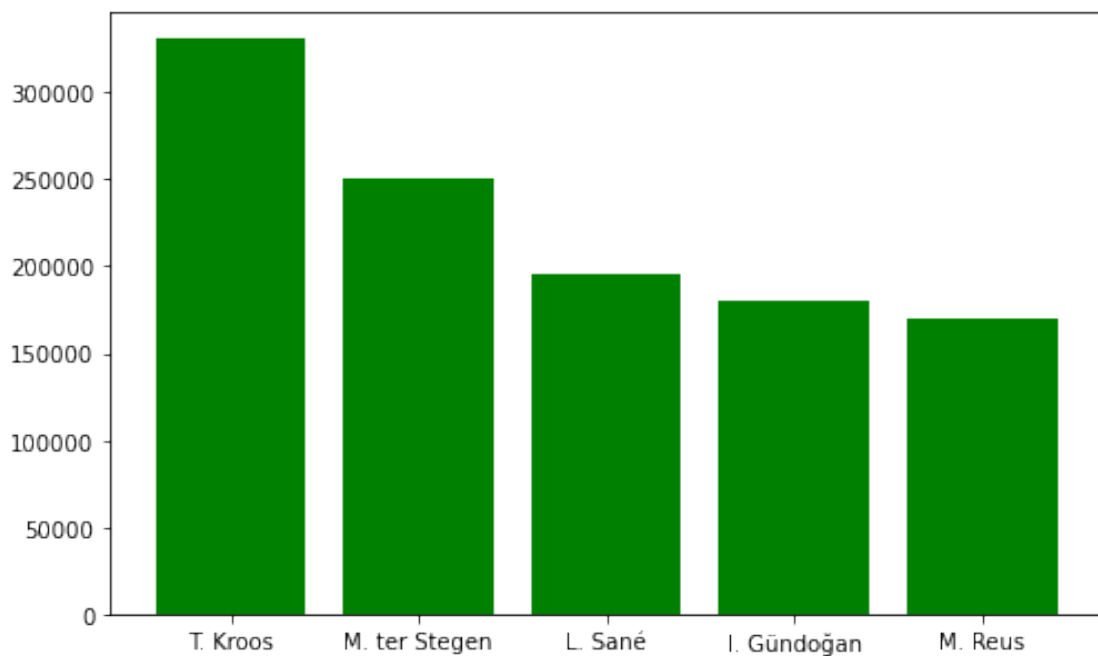
[10 rows x 104 columns]

```
[66]: fifa=fifa.sort_values(by=['weight_kg'],ascending=False)
```

```
[67]: plt.figure(figsize=(8,5))
plt.bar(list(fifa['short_name'])[0:5],list(fifa['weight_kg'])[0:5],color = 'b')
plt.show()
```

```
[69]: fifa=fifa.sort_values(by=['wage_eur'],ascending = False)
plt.figure(figsize=(8,5))
plt.bar(list(fifa['short_name'])[0:5],list(fifa['wage_eur'])[0:5],color = 'g')
plt.show()
```



```
[89]: #prepare a rank ordered list of 10 countries with most players .
      #list 5 countries that are producing the most number of footballers that play_
      ↪at this level
      country_name=list(fifa_data['nationality'].value_counts().keys()[0:10])
      country_players=list(fifa_data['nationality'].value_counts()[0:10])
      country={'name':country_name[0:5], 'players':country_players[0:5]}
      country
```

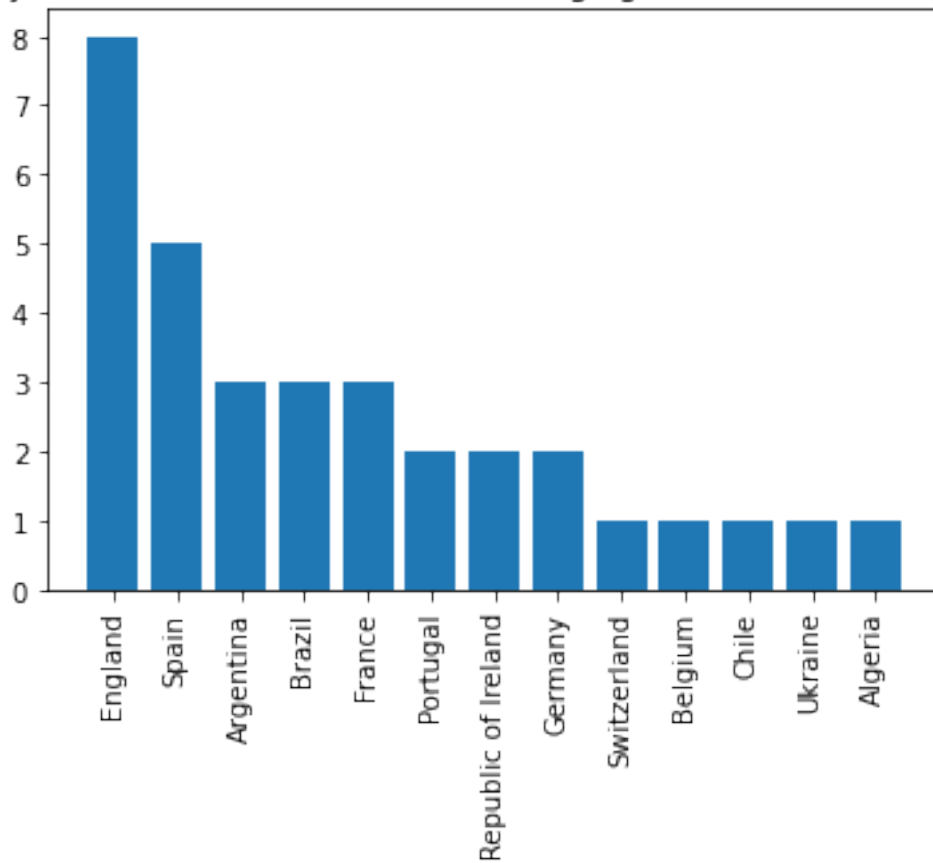
```
[89]: {'name': ['England', 'Germany', 'Spain', 'France', 'Argentina'],
      'players': [1667, 1216, 1035, 984, 886]}
```

```
[90]: fifa_data['club'].value_counts()
```

```
[90]: Athletic Club de Bilbao      33
      AS Monaco                    33
      1. FC Union Berlin           33
      Manchester City              33
      Chelsea                     33
      ..
      Mexico                      1
      Austria                     1
      Seongnam FC                 1
      Poland                     1
      Netherlands                 1
      Name: club, Length: 698, dtype: int64
```

```
[108]: manchester=fifa_data[fifa_data['club']== 'Manchester City']
      manchester['nationality'].value_counts()
      plt.bar(list(manchester['nationality'].value_counts().
      ↪keys()),list(manchester['nationality'].value_counts()))
      plt.xticks(rotation = 90)
      plt.title('players from different countries belonging to manchester city club')
      plt.show()
      plt.figure(figsize=(50,20))
```

players from different countries belonging to manchester city club



[108]: <Figure size 3600x1440 with 0 Axes>

<Figure size 3600x1440 with 0 Axes>

[111]: `fifa_data.sort_values(by=['team_position'],ascending=False)[0:5]`

```
[111]:      sofifa_id      player_url \
18277      233449  https://sofifa.com/player/233449/ximing-pan/20...
11751      209965  https://sofifa.com/player/209965/orhan-ademi/2...
5613       230350  https://sofifa.com/player/230350/flavio-raphae...
11734      199611  https://sofifa.com/player/199611/milan-jurdik/...
5611       230249  https://sofifa.com/player/230249/marlon-romulo...

      short_name      long_name  age  dob \
18277      Pan Ximing                26  1/11/1993
11751         O. Ademi      Orhan Ademi  27  10/28/1991
5613  Flávio Rebeilo  Flávio Raphael Rebeilo Mestres  27  2/29/1992
11734         M. Jurdík      Milan Jurdík  27  11/8/1991
5611  Marlinito Dias  Marlon Rômulo Dias Cavalcanti  35  2/29/1984
```

	height_cm	weight_kg	nationality	club	...	\
18277	182	78	China PR	Hebei China Fortune FC	...	
11751	189	85	Switzerland	Eintracht Braunschweig	...	
5613	175	69	Brazil	Santos	...	
11734	186	73	Czech Republic	WSG Tirol	...	
5611	189	84	Brazil	Cruzeiro	...	

	lwb	ldm	cdm	rdm	rw	lb	lcb	cb	rcb	rb
18277	48+2	49+2	49+2	49+2	48+2	48+2	50+2	50+2	50+2	48+2
11751	44+2	44+2	44+2	44+2	44+2	43+2	44+2	44+2	44+2	43+2
5613	69+2	64+2	64+2	64+2	69+2	68+2	63+2	63+2	63+2	68+2
11734	45+2	45+2	45+2	45+2	45+2	44+2	45+2	45+2	45+2	44+2
5611	62+2	68+2	68+2	68+2	62+2	62+2	68+2	68+2	68+2	62+2

[5 rows x 104 columns]

```
[121]: #which club have maximum players from England
club_data = fifa_data[fifa_data['nationality']=='England']
club_data.head()
club_data.shape
club_data['club'].value_counts()
```

```
[121]: Tranmere Rovers          24
Exeter City                  24
Colchester United           24
Blackburn Rovers            24
Luton Town                  23
..
FC Sochaux-Montbéliard       1
SV Zulte-Waregem             1
DC United                    1
Spezia                       1
Borussia Mönchengladbach     1
Name: club, Length: 153, dtype: int64
```

```
[122]: #which club have maximum players from Spain
club_data1 = fifa_data[fifa_data['nationality']=='Spain']
club_data1.head()
club_data1.shape
club_data1['club'].value_counts()
```

```
[122]: Athletic Club de Bilbao    31
CD Lugo                        26
Real Valladolid CF            26
CA Osasuna                    26
Real Sporting de Gijón        25
```

```

..
Pogoń Szczecin      1
Tiburones Rojos de Veracruz  1
Oxford United      1
Rochdale           1
Sint-Truidense VV   1
Name: club, Length: 157, dtype: int64

```

```

[123]: #which club have maximum players from Germany
club_data1 = fifa_data[fifa_data['nationality']=='Germany']
club_data1.head()
club_data1.shape
club_data1['club'].value_counts()

```

```

[123]: 1. FC Heidenheim 1846      29
SpVgg Unterhaching      27
SV Wehen Wiesbaden      27
TSV 1860 München        27
FC Würzburger Kickers    25
..
Lokomotiv Moscow        1
Antalyaspor             1
WSG Tirol               1
Benevento               1
Preston North End       1
Name: club, Length: 139, dtype: int64

```

```

[169]: #team formation for england(4 defenders , 3 midfielders, 3 forward, 1
↪goalkeeper)
#players with overall values of more than 75 are preferred
#pick out best player based on overall rating
team_england = fifa_data[fifa_data['nationality']=='England']
team_england= team_england[team_england['overall'] > 75]
team_england.sort_values(by =['overall'],ascending=False)
team_england=team_england[['short_name','player_positions','overall']]
team_england_gk=team_england[team_england['player_positions']=='GK']
team_england_gk.head()

```

```

[169]:      short_name player_positions  overall
181   J. Pickford             GK         83
498    B. Foster             GK         80
499    T. Heaton             GK         80
628     N. Pope             GK         79
782    J. Butland             GK         78

```

```

[170]: team_england_fm=team_england[team_england['player_positions']=='CM']
team_england_fm.head()

```

```
[170]:
```

	short_name	player_positions	overall
689	A. Lallana	CM	79
1020	L. Cook	CM	77
1184	D. Drinkwater	CM	77
1192	A. Westwood	CM	77
1461	J. Cork	CM	76

```
[171]: team_england_st=team_england[team_england['player_positions']=='ST']
team_england_st.head()
```

```
[171]:
```

	short_name	player_positions	overall
12	H. Kane	ST	89
321	J. Vardy	ST	82
473	C. Wilson	ST	80
488	W. Rooney	ST	80
904	T. Deeney	ST	78

```
[172]: team_england_D=team_england[team_england['player_positions']=='
↳('CB'or'LB'or'RB')]
team_england_D.head()
```

```
[172]:
```

	short_name	player_positions	overall
172	J. Stones	CB	83
280	H. Maguire	CB	82
457	M. Keane	CB	80
534	C. Smalling	CB	80
626	J. Tarkowski	CB	79

```
[ ]:
```