zerotopandas-course-project

December 29, 2020

1 FIFA-2020 to 2021 Players Data

In this notebook the analysis will be on FIFA players. Here we will find out the top players, top ten clubs, top ten richest clubs and many more. The dataset has been downloaded from Kaggle Dataset. The dataset contains a csv file i.e., players_20.csv. Libraries are used in this notbook are: Numpy, Pandas, Matplotlib and Seaborn. These libraries are covered in the course by Jovian.ml named as Data Analysis with Python: Zero to Pandas. As a first step, let's upload our Jupyter notebook to Jovian.ml.

[10]:
```
project_name = "FIFA-2020 to 2021 Players Data"
```

[11]:
```
!pip install jovian --upgrade -q
```

[12]:
```
import jovian
```

[13]:
```
jovian.commit(project=project_name)
```

```
[jovian] Detected Colab notebook…
[jovian] Uploading colab notebook to Jovian…
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

[13]:
```
'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'
```

1.1 Section-1: Data Preparation and Cleaning

In this step we will perform the following process:

- Install and import the upgraded version of all libraries to * be used in the project
- Analyse the dataset available with us
- Find the missing data, if possible
- Reason of missing data

[ ]:
```
##Installing the upgraded version of all libraries
!pip install numpy pandas matplotlib seaborn --upgrade --quiet
```

[ ]:
```
#Importing the libraries
import numpy as np
import pandas as pd
```

```
import matplotlib
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
```

[ ]:
```
#Global Settings
matplotlib.rcParams['font.size'] = 14
matplotlib.rcParams['figure.figsize'] = (15, 7.5)
matplotlib.rcParams['figure.facecolor'] = '#00000000'
sns.set_style("darkgrid")
```

[ ]:
```
#Reading players_20.csv into FIFA_21_df dataframe
FIFA_21_df = pd.read_csv('players_20.csv')
```

[ ]:
```
#showing FIFA_21_df
FIFA_21_df
```

[ ]:
```
       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…
…            …                                               …
18273     245006  https://sofifa.com/player/245006/shuai-shao/20…
18274     250995  https://sofifa.com/player/250995/mingjie-xiao/…
18275     252332  https://sofifa.com/player/252332/wei-zhang/20/…
18276     251110  https://sofifa.com/player/251110/haijian-wang/…
18277     233449  https://sofifa.com/player/233449/ximing-pan/20…

               short_name                          long_name  age  \
0                L. Messi      Lionel Andrés Messi Cuccittini   32
1      Cristiano Ronaldo  Cristiano Ronaldo dos Santos Aveiro   34
2              Neymar Jr       Neymar da Silva Santos Junior   27
3                J. Oblak                         Jan Oblak   26
4                E. Hazard                      Eden Hazard   28
…                      …                                …  …
18273         Shao Shuai                                     22
18274       Xiao Mingjie                    Mingjie Xiao   22
18275          Zhang Wei                                   19
18276       Wang Haijian                                   18
18277         Pan Ximing                                   26

              dob  height_cm  weight_kg nationality  \
0      1987-06-24        170         72   Argentina
1      1985-02-05        187         83    Portugal
```

```
2      1992-02-05        175         68      Brazil
3      1993-01-07        188         87    Slovenia
4      1991-01-07        175         74     Belgium
…             …        …          …          …
18273  1997-03-10        186         79    China PR
18274  1997-01-01        177         66    China PR
18275  2000-05-16        186         75    China PR
18276  2000-08-02        185         74    China PR
18277  1993-01-11        182         78    China PR

                                  club  …   lwb   ldm   cdm   rdm   rwb    lb  \
0                        FC Barcelona  …  68+2  66+2  66+2  66+2  68+2  63+2
1                            Juventus  …  65+3  61+3  61+3  61+3  65+3  61+3
2                  Paris Saint-Germain  …  66+3  61+3  61+3  61+3  66+3  61+3
3                      Atlético Madrid  …   NaN   NaN   NaN   NaN   NaN   NaN
4                          Real Madrid  …  66+3  63+3  63+3  63+3  66+3  61+3
…                                    …  …   …    …   …    …   …
18273                Beijing Renhe FC  …  43+2  42+2  42+2  42+2  43+2  45+2
18274                Shanghai SIPG FC  …  44+2  43+2  43+2  43+2  44+2  46+2
18275          Hebei China Fortune FC  …  47+2  49+2  49+2  49+2  47+2  47+2
18276  Shanghai Greenland Shenhua FC  …  48+2  48+2  48+2  48+2  48+2  48+2
18277          Hebei China Fortune FC  …  48+2  49+2  49+2  49+2  48+2  48+2

        lcb    cb   rcb    rb
0      52+2  52+2  52+2  63+2
1      53+3  53+3  53+3  61+3
2      46+3  46+3  46+3  61+3
3       NaN   NaN   NaN   NaN
4      49+3  49+3  49+3  61+3
…      …   …    …    …
18273  46+2  46+2  46+2  45+2
18274  47+2  47+2  47+2  46+2
18275  49+2  49+2  49+2  47+2
18276  49+2  49+2  49+2  48+2
18277  50+2  50+2  50+2  48+2

[18278 rows x 104 columns]
```

[ ]:
```
#Getting the information about the FIFA_21 dataframe
FIFA_21_df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18278 entries, 0 to 18277
Columns: 104 entries, sofifa_id to rb
dtypes: float64(16), int64(45), object(43)
memory usage: 14.5+ MB
```

From the above information the following can be observed: 1. There are 18278 rows and 104

columns. 2. We have drop some columns and this can be done by cleaning the dataframe.

[ ]:
```
#Displaying the columns
FIFA_21_df.columns
```

[ ]:
```
Index(['sofifa_id', 'player_url', 'short_name', 'long_name', 'age', 'dob',
       'height_cm', 'weight_kg', 'nationality', 'club',
       …
       'lwb', 'ldm', 'cdm', 'rdm', 'rwb', 'lb', 'lcb', 'cb', 'rcb', 'rb'],
      dtype='object', length=104)
```

We will drop some columns. And we will do this operations for n number of times till when we will get the required results.

[ ]:
```
#Dropping the column of umpires as we will not be using it for any data analysis
FIFA_21_df.drop(columns=['lwb', 'ldm', 'cdm', 'rdm', 'rwb', 'lb', 'lcb', 'cb',␣
↪'rcb', 'rb','ls', 'st', 'rs',
       'lw', 'lf', 'cf', 'rf', 'rw', 'lam', 'cam', 'ram', 'lm', 'lcm', 'cm',
       'rcm', 'rm', 'player_traits', 'attacking_crossing',
       'attacking_finishing', 'attacking_heading_accuracy',
       'attacking_short_passing', 'attacking_volleys','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','work_rate',
       'body_type', 'real_face', 'release_clause_eur',␣
↪'player_tags','gk_handling', 'gk_kicking', 'gk_reflexes', 'gk_speed',
       'gk_positioning', 'skill_dribbling', 'skill_curve', 'skill_fk_accuracy',
       'skill_long_passing', 'skill_ball_control','goalkeeping_reflexes',
       'goalkeeping_handling'], inplace=True)
```

[ ]:
```
#Finding the shape of the matches dataframe
FIFA_21_df.shape
```

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

[ ]:
```
#Finding the columns of matches dataframe
FIFA_21_df.columns
```

[ ]:
```
Index(['sofifa_id', 'player_url', 'short_name', 'long_name', 'age', 'dob',
       'height_cm', 'weight_kg', 'nationality', 'club',
       …
       'lwb', 'ldm', 'cdm', 'rdm', 'rwb', 'lb', 'lcb', 'cb', 'rcb', 'rb'],
      dtype='object', length=104)
```

```
[ ]: #Again we will perform the drop function. Because the dataframe has many columns
     FIFA_21_df.
      ↪drop(columns=['player_url','international_reputation','goalkeeping_diving',␣
      ↪'goalkeeping_kicking'], inplace=True)
```

```
[ ]: #Finding the columns of matches dataframe
     FIFA_21_df.columns
```

```
[ ]: Index(['sofifa_id', 'short_name', 'long_name', 'age', 'dob', 'height_cm',
            'weight_kg', 'nationality', 'club', 'overall', 'potential', 'value_eur',
            'wage_eur', 'player_positions', 'preferred_foot', '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_handling',
            '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'],
           dtype='object')
```

```
[ ]: FIFA_21_df.drop(columns=['physic', 'skill_moves','contract_valid_until',␣
      ↪'loaned_from', 'joined', 'defending_standing_tackle',
           'defending_sliding_tackle'], inplace=True)
```

```
[ ]: FIFA_21_df.columns
```

```
[ ]: Index(['sofifa_id', 'short_name', 'long_name', 'age', 'dob', 'height_cm',
            'weight_kg', 'nationality', 'club', 'overall', 'potential', 'value_eur',
            'wage_eur', 'player_positions', 'preferred_foot', 'weak_foot',
            'team_position', 'team_jersey_number', 'nation_position',
            'nation_jersey_number', 'pace', 'shooting', 'passing', 'dribbling',
            'defending', 'gk_diving', 'defending_marking',
            'goalkeeping_positioning'],
```

```
      dtype='object')
```

[ ]:
```
FIFA_21_df.shape
```

[ ]:
```
(18278, 28)
```

[ ]:
```
FIFA_21_df
```

[ ]:
```
       sofifa_id         short_name                          long_name  age  \
0         158023           L. Messi       Lionel Andrés Messi Cuccittini   32
1          20801  Cristiano Ronaldo  Cristiano Ronaldo dos Santos Aveiro   34
2         190871          Neymar Jr       Neymar da Silva Santos Junior   27
3         200389           J. Oblak                          Jan Oblak   26
4         183277          E. Hazard                       Eden Hazard   28
…            …                 …                                …  …
18273     245006         Shao Shuai                                    22
18274     250995       Xiao Mingjie                      Mingjie Xiao   22
18275     252332          Zhang Wei                                    19
18276     251110       Wang Haijian                                   18
18277     233449         Pan Ximing                                    26

              dob  height_cm  weight_kg nationality  \
0      1987-06-24        170         72   Argentina
1      1985-02-05        187         83    Portugal
2      1992-02-05        175         68      Brazil
3      1993-01-07        188         87    Slovenia
4      1991-01-07        175         74     Belgium
…             …          …          …           …
18273  1997-03-10        186         79    China PR
18274  1997-01-01        177         66    China PR
18275  2000-05-16        186         75    China PR
18276  2000-08-02        185         74    China PR
18277  1993-01-11        182         78    China PR

                            club  overall  …  nation_position  \
0                    FC Barcelona       94  …              NaN
1                        Juventus       93  …               LS
2               Paris Saint-Germain       92  …               LW
3                   Atlético Madrid       91  …               GK
4                      Real Madrid       91  …               LF
…                              …      …  …  …              …
18273             Beijing Renhe FC       48  …              NaN
18274             Shanghai SIPG FC       48  …              NaN
18275       Hebei China Fortune FC       48  …              NaN
18276  Shanghai Greenland Shenhua FC       48  …              NaN
18277       Hebei China Fortune FC       48  …              NaN
```

```
       nation_jersey_number  pace shooting passing  dribbling defending  \
0                       NaN  87.0     92.0    92.0       96.0      39.0
1                       7.0  90.0     93.0    82.0       89.0      35.0
2                      10.0  91.0     85.0    87.0       95.0      32.0
3                       1.0   NaN      NaN     NaN        NaN       NaN
4                      10.0  91.0     83.0    86.0       94.0      35.0
…                         …   …        …       …          …         …
18273                   NaN  57.0     23.0    28.0       33.0      47.0
18274                   NaN  58.0     24.0    33.0       35.0      48.0
18275                   NaN  54.0     35.0    44.0       45.0      48.0
18276                   NaN  59.0     35.0    47.0       47.0      45.0
18277                   NaN  60.0     32.0    51.0       45.0      47.0

       gk_diving defending_marking  goalkeeping_positioning
0            NaN                33                       14
1            NaN                28                       14
2            NaN                27                       15
3           87.0                27                       90
4            NaN                34                        8
…            …                  …                        …
18273        NaN                45                        5
18274        NaN                42                       11
18275        NaN                46                        6
18276        NaN                39                       13
18277        NaN                52                        9

[18278 rows x 28 columns]
```

[ ]:
```
FIFA_21_df.club.unique()
```

[ ]:
```
array(['FC Barcelona', 'Juventus', 'Paris Saint-Germain',
       'Atlético Madrid', 'Real Madrid', 'Manchester City', 'Liverpool',
       'Napoli', 'Tottenham Hotspur', 'Manchester United', 'Chelsea',
       'FC Bayern München', 'Inter', 'Borussia Dortmund', 'Arsenal',
       'Valencia CF', 'Lazio', 'Milan', 'Sporting CP',
       'Olympique Lyonnais', 'RB Leipzig', 'Ajax', 'LA Galaxy',
       'Atalanta', 'RC Celta', 'Bayer 04 Leverkusen', 'Real Betis',
       'FC Porto', 'SV Werder Bremen', 'West Ham United',
       'Wolverhampton Wanderers', 'AS Saint-Étienne', 'Torino',
       'Dalian YiFang FC', 'Borussia Mönchengladbach', 'Roma',
       'Guangzhou Evergrande Taobao FC', 'SL Benfica',
       'Medipol Başakşehir FK', 'Everton', 'VfL Wolfsburg',
       'Crystal Palace', 'Getafe CF', 'Shanghai SIPG FC',
       'Eintracht Frankfurt', 'Olympique de Marseille', 'Hertha BSC',
       'RSC Anderlecht', 'Villarreal CF', 'Sampdoria', 'Leicester City',
       'AS Monaco', 'Jiangsu Suning FC', 'Los Angeles FC', 'Cagliari',
       'Sevilla FC', 'Fenerbahçe SK', 'Real Sociedad',
```

'TSG 1899 Hoffenheim', 'Atlético Mineiro', 'Grêmio', 'PSV',
'Athletic Club de Bilbao', 'Deportivo Alavés', 'Boca Juniors',
'Lokomotiv Moscow', 'Al Nassr', 'Brescia', 'Shakhtar Donetsk',
'Shanghai Greenland Shenhua FC', 'PFC CSKA Moscow',
'Beijing Sinobo Guoan FC', 'Levante UD', 'Cruzeiro', 'Uruguay',
'Montpellier HSC', 'Atlanta United', 'Watford', '1. FC Köln',
'Bournemouth', 'Beşiktaş JK', 'Real Valladolid CF', 'Racing Club',
'Al Hilal', 'Guangzhou R&F FC', 'Sassuolo',
'FC Girondins de Bordeaux', 'LOSC Lille', 'Galatasaray SK',
'Chicago Fire', 'Fluminense', 'Ecuador', 'RCD Espanyol',
'Dinamo Zagreb', 'FC Nantes', 'River Plate', 'OGC Nice',
'Newcastle United', 'Brighton & Hove Albion', 'Club Brugge KV',
'FC Schalke 04', 'SD Eibar', 'DC United', 'Orlando City SC',
'Hebei China Fortune FC', 'Tigres U.A.N.L.', 'Aston Villa',
'Montreal Impact', 'Olympiacos CFP', 'Norwich City', 'Feyenoord',
'Toronto FC', 'KRC Genk', 'Fiorentina', 'Spartak Moscow',
'Dynamo Kyiv', 'SK Slavia Praha', 'Southampton', 'Burnley',
'SC Braga', 'Russia', 'RC Strasbourg Alsace', 'Wuhan Zall',
'Vissel Kobe', 'Portland Timbers', 'Genoa', 'Beijing Renhe FC',
'Toulouse Football Club', 'Girona FC', 'Real Zaragoza',
'CD Leganés', 'Shenzhen FC', 'Internacional', 'CSA - AL', 'Santos',
'1. FSV Mainz 05', 'Stoke City', 'Udinese', 'Colombia',
'Angers SCO', 'FC Augsburg', 'Netherlands', 'Fulham',
'FC København', 'KAA Gent', 'SC Freiburg', 'Stade Rennais FC',
'Club América', 'Trabzonspor', 'BSC Young Boys', 'Helsingborgs IF',
'Kaizer Chiefs', 'Parma', 'Mexico', 'Royal Antwerp FC',
'Tianjin TEDA FC', 'Hannover 96', 'Tianjin Quanjian FC', 'Al Ahli',
'Bologna', 'VfB Stuttgart', 'Seattle Sounders FC', 'Godoy Cruz',
'Sparta Praha', 'Independiente Medellín', 'Sivasspor',
'Independiente', 'Nîmes Olympique', 'Club Tijuana', 'SPAL',
'Monterrey', 'CD Tondela', 'Fortuna Düsseldorf',
'Vitória Guimarães', 'AZ Alkmaar', 'Atlético de San Luis', 'PAOK',
'Nagoya Grampus', 'Club Atlético Banfield',
'Shandong Luneng TaiShan FC', 'New York Red Bulls',
'FC Red Bull Salzburg', 'Sweden', 'Amiens SC', 'Hellas Verona',
'Cruz Azul', 'Gençlerbirliği SK', '1. FC Union Berlin',
'Standard de Liège', 'Chongqing Dangdai Lifan FC SWM Team',
'New England Revolution', 'Club Atlético Colón', 'Celtic',
'Club Atlas', 'Botafogo', 'En Avant de Guingamp',
'West Bromwich Albion', 'Pachuca', 'AEK Athens', 'Portimonense SC',
'Real Salt Lake', 'FC Utrecht', 'Sheffield United',
"Newell's Old Boys", 'Club Atlético Talleres',
'Philadelphia Union', 'Rosenborg BK', 'FC Basel 1893', 'Brentford',
'Club León', 'Unión de Santa Fe', 'Deportivo de La Coruña',
'Rangers FC', 'Turkey', 'Atiker Konyaspor', 'Granada CF',
'Perth Glory', 'Club Atlético Lanús', 'Hamburger SV', 'Al Ittihad',
'Santos Laguna', 'Western United FC', 'Columbus Crew SC',

'Deportivo Toluca', 'Cardiff City', 'CA Osasuna', 'Swansea City',
'Melbourne Victory', 'Leeds United', 'Göztepe SK', 'Hungary',
'New York City FC', 'Bulgaria', 'Kayserispor',
'Minnesota United FC', 'Guadalajara', 'FC Groningen', 'Paraguay',
'Junior FC', 'Al Taawoun', 'Huddersfield Town', 'Ettifaq FC',
'Stade de Reims', 'Rayo Vallecano', 'San Lorenzo de Almagro',
'Bahia', 'Atlético Paranaense', 'Goiás', 'Avaí FC', 'Fortaleza',
'Kawasaki Frontale', 'Vélez Sarsfield', 'Hull City',
'Houston Dynamo', 'Birmingham City', 'Al Wehda',
'Nottingham Forest', 'CD Tenerife', 'Aalborg BK',
'Preston North End', 'Bristol City', 'Lecce',
'Gimnasia y Esgrima La Plata', 'CD Aves', 'Viktoria Plzeň',
'1. FC Nürnberg', 'Slovenia', 'FC Metz', 'FC Midtjylland',
'Molde FK', 'Colo-Colo', 'RCD Mallorca', 'FC Sion', 'Wisła Kraków',
'Denizlispor', 'Middlesbrough', 'Universidad Católica',
'Alanyaspor', 'Moreirense FC', "Côte d'Ivoire",
'Stade Brestois 29', 'Chievo Verona', 'Boavista FC',
'MKE Ankaragücü', 'Sydney FC', 'Al Ain FC', 'Yeni Malatyaspor',
'Urawa Red Diamonds', 'Querétaro', 'Sporting Kansas City',
'Málaga CF', '1. FC Heidenheim 1846', 'Atlético Tucumán',
'Gazişehir Gaziantep F.K.', 'Clube Sport Marítimo', 'Chapecoense',
'Atlético Nacional', 'Vitesse', 'Australia', 'Henan Jianye FC',
'Panathinaikos FC', 'Blackburn Rovers', 'Santa Clara', 'Cameroon',
'Puebla FC', 'U.N.A.M.', 'SD Huesca', 'Estudiantes de La Plata',
'Austria', 'KAS Eupen', 'LASK Linz', 'Sporting de Charleroi',
'Daegu FC', 'Real Sporting de Gijón', 'Derby County', 'Rio Ave FC',
'South Africa', 'Famalicão', 'Neuchâtel Xamax', 'Benevento',
'Cádiz CF', 'AIK', 'Sheffield Wednesday', 'Empoli',
'Colorado Rapids', 'Os Belenenses', 'Unión Magdalena',
'Real Oviedo', 'Peru', 'Antalyaspor', 'Dijon FCO', 'SV Sandhausen',
'Rosario Central', 'Reading', 'DSC Arminia Bielefeld', 'Malmö FF',
'Jeonbuk Hyundai Motors', 'Frosinone', 'FC Tokyo', 'Canada',
'Çaykur Rizespor', 'FCSB (Steaua)', 'Defensa y Justicia',
'Monarcas Morelia', 'Club Atlético Huracán', 'Ceará Sporting Club',
'Argentinos Juniors', 'Al Shabab', 'Legia Warszawa',
'Shimizu S-Pulse', 'Millonarios FC', 'Lechia Gdańsk', 'Brøndby IF',
'Albacete BP', 'FC Lorient', 'Universitatea Craiova',
'Deportivo Cali', 'SK Rapid Wien', 'Kashima Antlers', 'Poland',
'Elche CF', 'Club Atlético Aldosivi', 'Deportes Tolima',
'Cúcuta Deportivo', 'Club Necaxa', 'Piast Gliwice', 'Pescara',
'Kasimpaşa SK', 'Egypt', 'Holstein Kiel', 'Livorno',
'Coquimbo Unido', 'UD Las Palmas', 'Górnik Zabrze', 'FC Twente',
'Paris FC', 'Racing Club de Lens', 'VfL Bochum 1848', 'Sunderland',
'Aberdeen', 'Heart of Midlothian', 'Romania', 'Crotone', 'Iceland',
'CFR Cluj', 'FK Austria Wien', 'UD Almería', 'SK Brann',
'BK Häcken', 'Al Fateh', 'CD Everton de Viña del Mar',
'FC St. Pauli', 'Gamba Osaka', 'Yokohama F. Marinos',

'Djurgårdens IF', 'FC Dallas', 'Sagan Tosu', 'Al Fayha', 'Chile',
'Once Caldas', 'Atlético Bucaramanga', 'FC Cincinnati',
'San Jose Earthquakes', 'América de Cali', 'Perugia', 'La Equidad',
'SV Darmstadt 98', 'FC Zürich', 'SC Paderborn 07',
'SV Zulte-Waregem', 'IFK Norrköping', 'FC Viitorul', 'Pordenone',
'AD Alcorcón', 'Sint-Truidense VV', 'KV Kortrijk',
'Royal Excel Mouscron', 'FC Lugano', 'Hokkaido Consadole Sapporo',
'Spezia', 'FC Paços de Ferreira', 'Servette FC',
'US Salernitana 1919', 'Lech Poznań', 'ESTAC Troyes',
'Fortuna Sittard', 'KV Oostende', 'VVV-Venlo', 'KSV Cercle Brugge',
'FC Seoul', 'Cerezo Osaka', 'Stade Malherbe Caen',
'Universidad de Chile', 'Kalmar FF', 'KV Mechelen',
'Deportes Iquique', 'US Cremonese', 'Bolivia', 'Valenciennes FC',
'Ulsan Hyundai FC', 'Vegalta Sendai', 'ADO Den Haag',
'CD Palestino', 'Cittadella', 'Vitória de Setúbal',
'FC Nordsjælland', 'Charlton Athletic', 'Al Raed',
'Jagiellonia Białystok', 'Odense Boldklub', 'FC Thun',
'SG Dynamo Dresden', 'Hammarby IF', 'Central Córdoba',
'Queens Park Rangers', ' SSV Jahn Regensburg',
'Tiburones Rojos de Veracruz', 'Patronato', 'Virtus Entella',
'SK Sturm Graz', 'Kilmarnock', 'Extremadura UD', 'Willem II',
'Gil Vicente FC', 'Randers FC', 'Ascoli', 'Vancouver Whitecaps FC',
'Millwall', 'Pohang Steelers', 'Suwon Samsung Bluewings',
'Heracles Almelo', 'Al Hazem', 'FC Juárez', 'SC Heerenveen',
'Dinamo Bucureşti', 'Gyeongnam FC', 'KFC Uerdingen 05',
'PEC Zwolle', 'Al Faisaly', 'Arsenal de Sarandí',
'FC Ingolstadt 04', 'Wolfsberger AC', 'Wigan Athletic',
'Júbilo Iwata', 'CD Lugo', 'SpVgg Greuther Fürth', 'FC Emmen',
'Cosenza', 'AJ Auxerre', 'Cracovia', 'FC Erzgebirge Aue',
'Grenoble Foot 38', 'Korona Kielce', 'Alianza Petrolera',
'CD Universidad de Concepción', 'Melbourne City FC',
'CD Antofagasta', 'Independiente Santa Fe', 'Sanfrecce Hiroshima',
'Damac FC', 'Chamois Niortais Football Club', 'Atlético Huila',
'CD Numancia', 'Luton Town', 'Unión La Calera', 'Le Havre AC',
'AS Nancy Lorraine', 'Audax Italiano', 'Hibernian', 'Barnsley',
'Pogoń Szczecin', 'Sarpsborg 08 FF', 'Kristiansund BK',
'FC Luzern', 'Unión Española', 'IFK Göteborg', 'Envigado FC',
'Jeju United FC', 'Patriotas Boyacá FC', 'Peterborough United',
'Deportivo Pasto', 'Wales', 'Clermont Foot 63', 'Northern Ireland',
'Astra Giurgiu', 'Racing Santander', 'Pisa', 'Newcastle Jets',
'Castellammare di Stabia', 'Rionegro Águilas', 'Vålerenga Fotball',
"CD O'Higgins", 'AC Ajaccio', 'CF Fuenlabrada', 'VfL Osnabrück',
'Arka Gdynia', 'FC Sochaux-Montbéliard', 'SpVgg Unterhaching',
'Central Coast Mariners', 'SV Wehen Wiesbaden', 'Fleetwood Town',
'Venezia FC', 'Karlsruher SC', 'Wisła Płock', 'Gangwon FC',
'IF Elfsborg', 'Sangju Sangmu FC', 'Zagłębie Lubin', 'Abha Club',
'Jaguares de Córdoba', 'FC Botoşani', 'Motherwell',

```
'FC St. Gallen', 'Bayern München II', 'CD Mirandés',
'St. Johnstone FC', 'Lillestrøm SK', 'Odds BK', 'SD Ponferradina',
'Doncaster Rovers', 'Örebro SK', 'Shamrock Rovers', 'FSV Zwickau',
'1. FC Kaiserslautern', 'Portsmouth', 'Waasland-Beveren',
'Ipswich Town', 'Dundalk', 'CD Huachipato', 'Le Mans FC',
'Strømsgodset IF', 'Sepsi OSK', 'Sparta Rotterdam',
'FC Admira Wacker Mödling', 'SønderjyskE', 'Hallescher FC',
'SKN St. Pölten', 'MSV Duisburg', 'La Berrichonne de Châteauroux',
'FC Hansa Rostock', 'Incheon United FC', 'Shonan Bellmare',
'Orlando Pirates', 'FK Bodø/Glimt', 'Lincoln City',
'Brisbane Roar', 'Eintracht Braunschweig', 'Curicó Unido',
'Falkenbergs FF', 'Venezuela', '1. FC Magdeburg', 'CD Cobresal',
'Aarhus GF', 'Salford City', 'HJK Helsinki', 'Oxford United',
'Hobro IK', 'Esbjerg fB', 'Bolton Wanderers', 'SV Mattersburg',
'Rotherham United', 'SCR Altach', 'Gaz Metan Mediaş', 'Tromsø IL',
'WSG Tirol', 'IK Sirius', 'Shrewsbury', 'Oldham Athletic',
'Southend United', 'Blackpool', 'Coventry City', 'Adelaide United',
'Ranheim Fotball', 'SG Sonnenhof Großaspach', 'Matsumoto Yamaga',
'Burton Albion', 'FC Hermannstadt', 'TSV 1860 München', 'LKS Lodz',
'FC Würzburger Kickers', 'TSV Hartberg', 'Politehnica Iaşi',
'Wellington Phoenix', 'Milton Keynes Dons', 'Stabæk Fotball',
'Wycombe Wanderers', 'SV Waldhof Mannheim', 'AC Horsens',
'Scunthorpe United', 'RKC Waalwijk', 'Oita Trinita', 'St. Mirren',
'Bristol Rovers', 'Rodez Aveyron Football', 'SV Meppen',
'Viking FK', 'Östersunds FK', 'FK Haugesund', 'Rochdale',
'Colchester United', 'Trapani', 'Stevenage', 'Bradford City',
'Livingston FC', 'FC Chambly Oise', 'Mansfield Town',
'SC Preußen Münster', 'FC Voluntari', 'Accrington Stanley',
'Al Adalah', 'Lyngby BK', 'FC Carl Zeiss Jena', 'Viktoria Köln',
'Tranmere Rovers', 'Silkeborg IF', 'Gillingham', 'Plymouth Argyle',
'Chemnitzer FC', 'Mjøndalen IF', 'Walsall', 'Northampton Town',
'Hamilton Academical FC', 'Grimsby Town', 'Exeter City',
'Swindon Town', 'Raków Częstochowa', 'Chindia Târgovişte',
'US Orléans Loiret Football', 'Forest Green Rovers',
'AFC Wimbledon', 'Carlisle United', 'Morecambe', 'Port Vale',
'Cheltenham Town', 'Academica Clinceni', 'Crawley Town',
'Ross County FC', 'AFC Eskilstuna', 'Macclesfield Town',
'Cork City', 'Newport County', 'Crewe Alexandra', 'Leyton Orient',
'Cambridge United', "St. Patrick's Athletic", 'Bohemian FC',
'India', 'Finland', 'Waterford FC', 'Derry City', 'Bury',
'New Zealand', 'GIF Sundsvall', 'Sligo Rovers', 'Finn Harps',
'Seongnam FC', 'UCD AFC', 'Śląsk Wrocław'], dtype=object)
```

[ ]:
```
#drop rows and store in the new variable named as FIFA_df
FIFA_df = FIFA_21_df[:-18183]
```

[ ]:
```
FIFA_df
```

[ ]:
```
    sofifa_id         short_name                          long_name  age  \
0      158023           L. Messi       Lionel Andrés Messi Cuccittini   32
1       20801  Cristiano Ronaldo  Cristiano Ronaldo dos Santos Aveiro   34
2      190871          Neymar Jr       Neymar da Silva Santos Junior   27
3      200389           J. Oblak                          Jan Oblak   26
4      183277           E. Hazard                       Eden Hazard   28
..        …                 …                                 …  …
90     192774         K. Manolas                     Kostas Manolas   28
91     194404              Neto              Norberto Murara Neto   29
92     197445           D. Alaba           David Olatukunbo Alaba   27
93     201399           M. Icardi       Mauro Emanuel Icardi Rivero   26
94     204963           Carvajal              Daniel Carvajal Ramos   27

           dob  height_cm  weight_kg nationality                club  \
0   1987-06-24        170         72   Argentina         FC Barcelona
1   1985-02-05        187         83    Portugal             Juventus
2   1992-02-05        175         68      Brazil  Paris Saint-Germain
3   1993-01-07        188         87    Slovenia      Atlético Madrid
4   1991-01-07        175         74     Belgium          Real Madrid
..         …        …        …         …                   …
90  1991-06-14        189         83      Greece               Napoli
91  1989-07-19        190         84      Brazil         FC Barcelona
92  1992-06-24        180         78     Austria    FC Bayern München
93  1993-02-19        181         75   Argentina                Inter
94  1992-01-11        173         73       Spain          Real Madrid

    overall  …   lwb   ldm   cdm   rdm   rwb    lb   lcb    cb   rcb    rb
0        94  …  68+2  66+2  66+2  66+2  68+2  63+2  52+2  52+2  52+2  63+2
1        93  …  65+3  61+3  61+3  61+3  65+3  61+3  53+3  53+3  53+3  61+3
2        92  …  66+3  61+3  61+3  61+3  66+3  61+3  46+3  46+3  46+3  61+3
3        91  …   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN
4        91  …  66+3  63+3  63+3  63+3  66+3  61+3  49+3  49+3  49+3  61+3
..      …  …   …   …   …   …   …   …   …   …   …   …
90       85  …  71+3  74+3  74+3  74+3  71+3  75+3  83+3  83+3  83+3  75+3
91       85  …   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN   NaN
92       85  …  81+3  80+3  80+3  80+3  81+3  81+3  78+3  78+3  78+3  81+3
93       85  …  53+3  52+3  52+3  52+3  53+3  51+3  50+3  50+3  50+3  51+3
94       85  …  83+3  82+3  82+3  82+3  83+3  83+3  80+3  80+3  80+3  83+3

[95 rows x 100 columns]
```

[ ]:
```
FIFA_df.club.unique()
```

[ ]:
```
array(['FC Barcelona', 'Juventus', 'Paris Saint-Germain',
       'Atlético Madrid', 'Real Madrid', 'Manchester City', 'Liverpool',
       'Napoli', 'Tottenham Hotspur', 'Manchester United', 'Chelsea',
       'FC Bayern München', 'Inter', 'Borussia Dortmund', 'Arsenal',
```

```
       'Valencia CF', 'Lazio', 'Milan', 'Sporting CP',
       'Olympique Lyonnais', 'RB Leipzig'], dtype=object)
```

[ ]:
```
FIFA_df.columns
```

[ ]:
```
Index(['sofifa_id', 'short_name', 'long_name', 'age', 'dob', 'height_cm',
       'weight_kg', 'nationality', 'club', 'overall', 'potential', 'value_eur',
       'wage_eur', 'player_positions', 'preferred_foot', '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_handling',
       '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'],
      dtype='object')
```

We create new dataframe and store the FIFA_df dataframe in the newly created dataframe i.e., New_FIFA_df. Here we select 15 columns to analyze the data.

[ ]:
```
New_FIFA_df = FIFA_df[['sofifa_id','short_name', 'age', 'dob', 'nationality',␣
↪'club','overall','potential','value_eur','player_positions','nation_jersey_number','team_je
```

[ ]:
```
New_FIFA_df
```

[ ]:

	sofifa_id	short_name	age	dob	nationality	\
0	158023	L. Messi	32	1987-06-24	Argentina	
1	20801	Cristiano Ronaldo	34	1985-02-05	Portugal	
2	190871	Neymar Jr	27	1992-02-05	Brazil	
3	200389	J. Oblak	26	1993-01-07	Slovenia	
4	183277	E. Hazard	28	1991-01-07	Belgium	
..	…	…	…	…	…	
90	192774	K. Manolas	28	1991-06-14	Greece	

```
91     194404              Neto   29  1989-07-19      Brazil
92     197445          D. Alaba   27  1992-06-24     Austria
93     201399         M. Icardi   26  1993-02-19   Argentina
94     204963          Carvajal   27  1992-01-11       Spain

                  club  overall  potential  value_eur player_positions  \
0          FC Barcelona       94         94   95500000       RW, CF, ST
1              Juventus       93         93   58500000           ST, LW
2   Paris Saint-Germain       92         92  105500000          LW, CAM
3       Atlético Madrid       91         93   77500000               GK
4          Real Madrid       91         91   90000000           LW, CF
..                 …      …        …         …               …
90              Napoli       85         86   37500000               CB
91         FC Barcelona       85         86   31000000               GK
92    FC Bayern München       85         86   38000000           LB, CB
93                Inter       85         86   46000000               ST
94         Real Madrid       85         86   38000000               RB

    nation_jersey_number  team_jersey_number  pace  shooting  wage_eur
0                    NaN                10.0  87.0      92.0    565000
1                    7.0                 7.0  90.0      93.0    405000
2                   10.0                10.0  91.0      85.0    290000
3                    1.0                13.0   NaN       NaN    125000
4                   10.0                 7.0  91.0      83.0    470000
..                   …                  …    …       …       …
90                   4.0                44.0  82.0      25.0    105000
91                   NaN                13.0   NaN       NaN    175000
92                   8.0                27.0  83.0      73.0    135000
93                   NaN                36.0  75.0      84.0    115000
94                   2.0                 2.0  81.0      47.0    205000

[95 rows x 15 columns]
```

[ ]:
```
#Crafting a new Parameter for Growth = Potential - Overall
New_FIFA_df['Growth'] = New_FIFA_df['potential'] - New_FIFA_df['overall']
New_FIFA_df.head()
```

[ ]:
```
   sofifa_id         short_name  age         dob nationality  \
0     158023           L. Messi   32  1987-06-24   Argentina
1      20801  Cristiano Ronaldo   34  1985-02-05    Portugal
2     190871          Neymar Jr   27  1992-02-05      Brazil
3     200389           J. Oblak   26  1993-01-07    Slovenia
4     183277          E. Hazard   28  1991-01-07     Belgium

                club  overall  potential  value_eur player_positions  \
0        FC Barcelona       94         94   95500000       RW, CF, ST
1            Juventus       93         93   58500000           ST, LW
```

```
2  Paris Saint-Germain       92         92  105500000          LW, CAM
3      Atlético Madrid       91         93   77500000               GK
4          Real Madrid       91         91   90000000           LW, CF

   nation_jersey_number  team_jersey_number  pace  shooting  wage_eur  Growth
0                   NaN                10.0  87.0      92.0    565000       0
1                   7.0                 7.0  90.0      93.0    405000       0
2                  10.0                10.0  91.0      85.0    290000       0
3                   1.0                13.0   NaN       NaN    125000       2
4                  10.0                 7.0  91.0      83.0    470000       0
```

[ ]:
```
#Finding the rows where the nation_jersey_number is NaN
New_FIFA_df[New_FIFA_df.nation_jersey_number.isna()]
```

[ ]:
```
    sofifa_id           short_name  age         dob         nationality  \
0      158023             L. Messi   32  1987-06-24           Argentina
8      177003            L. Modrić   33  1985-09-09             Croatia
11     201024          K. Koulibaly   28  1991-06-20             Senegal
13     212831               Alisson   26  1992-10-02              Brazil
19     176580            L. Suárez   32  1987-01-24             Uruguay
25     210257              Ederson   25  1993-08-17              Brazil
29     152729                Piqué   32  1987-02-02               Spain
30     162835         S. Handanovič   34  1984-07-14            Slovenia
33     168542           David Silva   33  1986-01-08               Spain
34     179813             E. Cavani   32  1987-02-14             Uruguay
35     182493              D. Godín   33  1986-02-16             Uruguay
38     188567         P. Aubameyang   30  1989-06-18               Gabon
39     208722               S. Mané   27  1992-04-10             Senegal
40     212218            A. Laporte   25  1994-05-27              France
42     200145              Casemiro   27  1992-02-23              Brazil
43     200104                H. Son   26  1992-07-08       Korea Republic
44     135507           Fernandinho   34  1985-05-04              Brazil
45     164240          Thiago Silva   34  1984-09-22              Brazil
46     165153            K. Benzema   31  1987-12-19              France
49     178603            M. Hummels   30  1988-12-16             Germany
53     193041              K. Navas   32  1986-12-15          Costa Rica
56     207865            Marquinhos   25  1994-05-14              Brazil
57     232363          M. Škriniar   24  1995-02-11            Slovakia
63     201942      Roberto Firmino   27  1991-10-02              Brazil
64     168651            I. Rakitić   31  1988-03-10             Croatia
65     180206             M. Pjanić   29  1990-04-02  Bosnia Herzegovina
68     189242              Coutinho   27  1992-06-12              Brazil
69     189513                Parejo   30  1989-04-16               Spain
70     189596             T. Müller   29  1989-09-13             Germany
72     193301          A. Lacazette   28  1991-05-28              France
78     223848  S. Milinković-Savić   24  1995-02-27              Serbia
81     209499               Fabinho   25  1993-10-23              Brazil
```

83	216460	J. Giménez	24	1995-01-20	Uruguay	
84	220440	C. Lenglet	24	1995-06-17	France	
86	199482	A. Lopes	28	1990-10-01	Portugal	
88	189117	R. Bürki	28	1990-11-14	Switzerland	
91	194404	Neto	29	1989-07-19	Brazil	
93	201399	M. Icardi	26	1993-02-19	Argentina	

	club	overall	potential	value_eur	player_positions	\
0	FC Barcelona	94	94	95500000	RW, CF, ST	
8	Real Madrid	90	90	45000000	CM	
11	Napoli	89	91	67500000	CB	
13	Liverpool	89	91	58000000	GK	
19	FC Barcelona	89	89	53000000	ST	
25	Manchester City	88	91	54500000	GK	
29	FC Barcelona	88	88	38000000	CB	
30	Inter	88	88	26000000	GK	
33	Manchester City	88	88	36000000	CAM, CM	
34	Paris Saint-Germain	88	88	47000000	ST	
35	Inter	88	88	28000000	CB	
38	Arsenal	88	88	57000000	ST, LM	
39	Liverpool	88	88	62000000	LW, LM	
40	Manchester City	87	90	56500000	CB, LB	
42	Real Madrid	87	89	53500000	CDM	
43	Tottenham Hotspur	87	88	60000000	CF, LM	
44	Manchester City	87	87	19500000	CDM	
45	Paris Saint-Germain	87	87	18500000	CB	
46	Real Madrid	87	87	45000000	CF, ST	
49	Borussia Dortmund	87	87	41000000	CB	
53	Real Madrid	87	87	30500000	GK	
56	Paris Saint-Germain	86	90	51500000	CB, CDM	
57	Inter	86	90	52000000	CB	
63	Liverpool	86	87	52000000	CF, ST, CAM	
64	FC Barcelona	86	86	38000000	CM, CDM	
65	Juventus	86	86	42500000	CM, CDM	
68	FC Bayern München	86	86	46500000	LW, CM	
69	Valencia CF	86	86	41000000	CM	
70	FC Bayern München	86	86	43500000	CAM, RM, RW	
72	Arsenal	86	86	46000000	ST	
78	Lazio	85	90	50500000	CM, CAM	
81	Liverpool	85	89	45000000	CDM	
83	Atlético Madrid	85	89	45000000	CB	
84	FC Barcelona	85	89	45000000	CB	
86	Olympique Lyonnais	85	87	33500000	GK	
88	Borussia Dortmund	85	86	32000000	GK	
91	FC Barcelona	85	86	31000000	GK	
93	Inter	85	86	46000000	ST	

	nation_jersey_number	team_jersey_number	pace	shooting	wage_eur	Growth
0	NaN	10.0	87.0	92.0	565000	0
8	NaN	10.0	74.0	76.0	340000	0
11	NaN	26.0	71.0	28.0	150000	2
13	NaN	1.0	NaN	NaN	155000	2
19	NaN	9.0	73.0	89.0	355000	0
25	NaN	31.0	NaN	NaN	185000	3
29	NaN	3.0	56.0	61.0	285000	0
30	NaN	1.0	NaN	NaN	110000	0
33	NaN	21.0	65.0	74.0	265000	0
34	NaN	9.0	75.0	86.0	195000	0
35	NaN	2.0	60.0	48.0	135000	0
38	NaN	14.0	94.0	85.0	205000	0
39	NaN	10.0	94.0	83.0	220000	0
40	NaN	14.0	64.0	50.0	195000	3
42	NaN	14.0	62.0	72.0	240000	2
43	NaN	7.0	88.0	86.0	185000	1
44	NaN	25.0	66.0	74.0	200000	0
45	NaN	2.0	62.0	54.0	135000	0
46	NaN	9.0	76.0	82.0	285000	0
49	NaN	15.0	51.0	58.0	130000	0
53	NaN	1.0	NaN	NaN	195000	0
56	NaN	5.0	74.0	41.0	120000	4
57	NaN	37.0	69.0	41.0	89000	4
63	NaN	9.0	77.0	82.0	170000	1
64	NaN	4.0	62.0	81.0	245000	0
65	NaN	5.0	68.0	68.0	180000	0
68	NaN	10.0	79.0	80.0	175000	0
69	NaN	10.0	41.0	80.0	69000	0
70	NaN	25.0	72.0	83.0	170000	0
72	NaN	9.0	82.0	85.0	165000	0
78	NaN	21.0	67.0	79.0	73000	5
81	NaN	3.0	68.0	69.0	125000	4
83	NaN	2.0	69.0	45.0	69000	4
84	NaN	15.0	75.0	45.0	175000	4
86	NaN	1.0	NaN	NaN	88000	2
88	NaN	1.0	NaN	NaN	92000	1
91	NaN	13.0	NaN	NaN	175000	1
93	NaN	36.0	75.0	84.0	115000	1

We put values to nan values in the dataframe.

[ ]:
```
New_FIFA_df.
↪loc[[0,8,11,13,19,25,93,91,88,86,84,83,81,78,72,70,69,68,65,64,63,57,56,53,42,43,45,46,49,3
↪= '10'
```

[ ]:
```
New_FIFA_df
```

[ ]:
```
    sofifa_id         short_name  age         dob nationality  \
0      158023           L. Messi   32  1987-06-24   Argentina
1       20801  Cristiano Ronaldo   34  1985-02-05    Portugal
2      190871          Neymar Jr   27  1992-02-05      Brazil
3      200389           J. Oblak   26  1993-01-07    Slovenia
4      183277          E. Hazard   28  1991-01-07     Belgium
..        …                …  …         …         …
90     192774         K. Manolas   28  1991-06-14      Greece
91     194404               Neto   29  1989-07-19      Brazil
92     197445           D. Alaba   27  1992-06-24     Austria
93     201399          M. Icardi   26  1993-02-19   Argentina
94     204963           Carvajal   27  1992-01-11       Spain

                   club  overall  potential  value_eur player_positions  \
0          FC Barcelona       94         94   95500000       RW, CF, ST
1              Juventus       93         93   58500000           ST, LW
2   Paris Saint-Germain       92         92  105500000          LW, CAM
3       Atlético Madrid       91         93   77500000               GK
4           Real Madrid       91         91   90000000           LW, CF
..                  …      …        …         …              …
90              Napoli       85         86   37500000               CB
91         FC Barcelona       85         86   31000000               GK
92    FC Bayern München       85         86   38000000           LB, CB
93                Inter       85         86   46000000               ST
94          Real Madrid       85         86   38000000               RB

   nation_jersey_number  team_jersey_number  pace  shooting  wage_eur  Growth
0                    10                10.0  87.0      92.0    565000       0
1                     7                 7.0  90.0      93.0    405000       0
2                    10                10.0  91.0      85.0    290000       0
3                     1                13.0   NaN       NaN    125000       2
4                    10                 7.0  91.0      83.0    470000       0
..                  …                …   …       …       …       …
90                    4                44.0  82.0      25.0    105000       1
91                   10                13.0   NaN       NaN    175000       1
92                    8                27.0  83.0      73.0    135000       1
93                   10                36.0  75.0      84.0    115000       1
94                    2                 2.0  81.0      47.0    205000       1

[95 rows x 16 columns]
```

[ ]:
```
New_FIFA_df[New_FIFA_df.pace.isna()]
```

[ ]:
```
    sofifa_id     short_name  age         dob  nationality  \
3      200389       J. Oblak   26  1993-01-07     Slovenia
6      192448  M. ter Stegen   27  1992-04-30      Germany
13     212831        Alisson   26  1992-10-02       Brazil
```

14	193080	De Gea	28	1990-11-07	Spain
25	210257	Ederson	25	1993-08-17	Brazil
28	192119	T. Courtois	27	1992-05-11	Belgium
30	162835	S. Handanovič	34	1984-07-14	Slovenia
31	167495	M. Neuer	33	1986-03-27	Germany
32	167948	H. Lloris	32	1986-12-26	France
53	193041	K. Navas	32	1986-12-15	Costa Rica
60	186153	W. Szczęsny	29	1990-04-18	Poland
74	230621	G. Donnarumma	20	1999-02-25	Italy
86	199482	A. Lopes	28	1990-10-01	Portugal
87	185122	P. Gulácsi	29	1990-05-06	Hungary
88	189117	R. Bürki	28	1990-11-14	Switzerland
91	194404	Neto	29	1989-07-19	Brazil

	club	overall	potential	value_eur	player_positions	\
3	Atlético Madrid	91	93	77500000	GK	
6	FC Barcelona	90	93	67500000	GK	
13	Liverpool	89	91	58000000	GK	
14	Manchester United	89	90	56000000	GK	
25	Manchester City	88	91	54500000	GK	
28	Real Madrid	88	89	48000000	GK	
30	Inter	88	88	26000000	GK	
31	FC Bayern München	88	88	32000000	GK	
32	Tottenham Hotspur	88	88	36000000	GK	
53	Real Madrid	87	87	30500000	GK	
60	Juventus	86	88	37500000	GK	
74	Milan	85	92	41500000	GK	
86	Olympique Lyonnais	85	87	33500000	GK	
87	RB Leipzig	85	86	31000000	GK	
88	Borussia Dortmund	85	86	32000000	GK	
91	FC Barcelona	85	86	31000000	GK	

	nation_jersey_number	team_jersey_number	pace	shooting	wage_eur	Growth
3	1	13.0	NaN	NaN	125000	2
6	22	1.0	NaN	NaN	250000	3
13	10	1.0	NaN	NaN	155000	2
14	1	1.0	NaN	NaN	205000	1
25	10	31.0	NaN	NaN	185000	3
28	1	13.0	NaN	NaN	235000	1
30	NaN	1.0	NaN	NaN	110000	0
31	1	1.0	NaN	NaN	155000	0
32	1	1.0	NaN	NaN	150000	0
53	10	1.0	NaN	NaN	195000	0
60	1	1.0	NaN	NaN	135000	2
74	22	99.0	NaN	NaN	34000	7
86	10	1.0	NaN	NaN	88000	2
87	1	1.0	NaN	NaN	69000	1

```
88                  10                1.0   NaN       NaN     92000       1
91                  10               13.0   NaN       NaN    175000       1
```

[ ]:
```
#placing pace value to 68 where there is no value
New_FIFA_df.loc[[3,6,13,14,91,88,86,87],'pace'] = '68'
```

[ ]:
```
New_FIFA_df
```

[ ]:
```
    sofifa_id         short_name  age         dob nationality  \
0      158023           L. Messi   32  1987-06-24   Argentina
1       20801  Cristiano Ronaldo   34  1985-02-05    Portugal
2      190871          Neymar Jr   27  1992-02-05      Brazil
3      200389           J. Oblak   26  1993-01-07    Slovenia
4      183277          E. Hazard   28  1991-01-07     Belgium
..        …                 …  …         …          …
90     192774         K. Manolas   28  1991-06-14      Greece
91     194404               Neto   29  1989-07-19      Brazil
92     197445          D. Alaba   27  1992-06-24     Austria
93     201399          M. Icardi   26  1993-02-19   Argentina
94     204963           Carvajal   27  1992-01-11       Spain

                  club  overall  potential  value_eur player_positions  \
0          FC Barcelona       94         94   95500000       RW, CF, ST
1              Juventus       93         93   58500000           ST, LW
2   Paris Saint-Germain       92         92  105500000          LW, CAM
3       Atlético Madrid       91         93   77500000               GK
4          Real Madrid       91         91   90000000           LW, CF
..                 …      …         …         …                …
90             Napoli       85         86   37500000               CB
91         FC Barcelona       85         86   31000000               GK
92    FC Bayern München       85         86   38000000           LB, CB
93                Inter       85         86   46000000               ST
94         Real Madrid       85         86   38000000               RB

   nation_jersey_number  team_jersey_number pace  shooting  wage_eur  Growth
0                    10                10.0   87      92.0    565000       0
1                     7                 7.0   90      93.0    405000       0
2                    10                10.0   91      85.0    290000       0
3                     1                13.0   68       NaN    125000       2
4                    10                 7.0   91      83.0    470000       0
..                  …                 …  …       …       …        …
90                    4                44.0   82      25.0    105000       1
91                   10                13.0   68       NaN    175000       1
92                    8                27.0   83      73.0    135000       1
93                   10                36.0   75      84.0    115000       1
94                    2                 2.0   81      47.0    205000       1
```

```
[95 rows x 16 columns]
```

[ ]:
```
New_FIFA_df.loc[[74,60,53,32],'pace'] = '75'
```

[ ]:
```
New_FIFA_df
```

[ ]:
```
    sofifa_id         short_name  age         dob nationality  \
0      158023           L. Messi   32  1987-06-24   Argentina
1       20801  Cristiano Ronaldo   34  1985-02-05    Portugal
2      190871          Neymar Jr   27  1992-02-05      Brazil
3      200389           J. Oblak   26  1993-01-07    Slovenia
4      183277          E. Hazard   28  1991-01-07     Belgium
..        ...                ...  ...         ...         ...
90     192774         K. Manolas   28  1991-06-14      Greece
91     194404               Neto   29  1989-07-19      Brazil
92     197445           D. Alaba   27  1992-06-24     Austria
93     201399          M. Icardi   26  1993-02-19   Argentina
94     204963           Carvajal   27  1992-01-11       Spain

                   club  overall  potential  value_eur player_positions  \
0           FC Barcelona       94         94   95500000       RW, CF, ST
1               Juventus       93         93   58500000           ST, LW
2    Paris Saint-Germain       92         92  105500000          LW, CAM
3         Atlético Madrid       91         93   77500000               GK
4            Real Madrid       91         91   90000000           LW, CF
..                   ...      ...        ...        ...              ...
90               Napoli       85         86   37500000               CB
91          FC Barcelona       85         86   31000000               GK
92     FC Bayern München       85         86   38000000           LB, CB
93                 Inter       85         86   46000000               ST
94          Real Madrid       85         86   38000000               RB

    nation_jersey_number  team_jersey_number pace  shooting  wage_eur  Growth
0                     10                10.0   87      92.0    565000       0
1                      7                 7.0   90      93.0    405000       0
2                     10                10.0   91      85.0    290000       0
3                      1                13.0   68       NaN    125000       2
4                     10                 7.0   91      83.0    470000       0
..                   ...                ...  ...       ...       ...     ...
90                     4                44.0   82      25.0    105000       1
91                    10                13.0   68       NaN    175000       1
92                     8                27.0   83      73.0    135000       1
93                    10                36.0   75      84.0    115000       1
94                     2                 2.0   81      47.0    205000       1

[95 rows x 16 columns]
```

[ ]:
```
New_FIFA_df.loc[[25,28,30,31],'pace'] = '70'
```

[ ]:
```
New_FIFA_df[New_FIFA_df.nation_jersey_number.isna()]
```

[ ]:

	sofifa_id	short_name	age	dob	nationality	club	\
29	152729	Piqué	32	1987-02-02	Spain	FC Barcelona	
30	162835	S. Handanovič	34	1984-07-14	Slovenia	Inter	
40	212218	A. Laporte	25	1994-05-27	France	Manchester City	
44	135507	Fernandinho	34	1985-05-04	Brazil	Manchester City	

	overall	potential	value_eur	player_positions	nation_jersey_number	\
29	88	88	38000000	CB	NaN	
30	88	88	26000000	GK	NaN	
40	87	90	56500000	CB, LB	NaN	
44	87	87	19500000	CDM	NaN	

	team_jersey_number	pace	shooting	wage_eur	Growth
29	3.0	56	61.0	285000	0
30	1.0	70	NaN	110000	0
40	14.0	64	50.0	195000	3
44	25.0	66	74.0	200000	0

Here we adding the values to different players. We will add and run following cells. Because every players in the team have unique jersey number.

[ ]:
```
New_FIFA_df.loc[[29],'nation_jersey_number'] = '11'
```

[ ]:
```
New_FIFA_df.loc[[30],'nation_jersey_number'] = '16'
```

[ ]:
```
New_FIFA_df.loc[[40],'nation_jersey_number'] = '18'
```

[ ]:
```
New_FIFA_df.loc[[44],'nation_jersey_number'] = '9'
```

[ ]:
```
New_FIFA_df
```

[ ]:

	sofifa_id	short_name	age	dob	nationality	\
0	158023	L. Messi	32	1987-06-24	Argentina	
1	20801	Cristiano Ronaldo	34	1985-02-05	Portugal	
2	190871	Neymar Jr	27	1992-02-05	Brazil	
3	200389	J. Oblak	26	1993-01-07	Slovenia	
4	183277	E. Hazard	28	1991-01-07	Belgium	
..	…	…	…	…	…	
90	192774	K. Manolas	28	1991-06-14	Greece	
91	194404	Neto	29	1989-07-19	Brazil	
92	197445	D. Alaba	27	1992-06-24	Austria	
93	201399	M. Icardi	26	1993-02-19	Argentina	
94	204963	Carvajal	27	1992-01-11	Spain	

```
                   club  overall  potential  value_eur player_positions  \
0          FC Barcelona       94         94   95500000       RW, CF, ST
1              Juventus       93         93   58500000           ST, LW
2   Paris Saint-Germain       92         92  105500000          LW, CAM
3       Atlético Madrid       91         93   77500000               GK
4           Real Madrid       91         91   90000000           LW, CF
..                  …      …        …        …               …
90              Napoli       85         86   37500000               CB
91         FC Barcelona       85         86   31000000               GK
92    FC Bayern München       85         86   38000000           LB, CB
93                Inter       85         86   46000000               ST
94          Real Madrid       85         86   38000000               RB

   nation_jersey_number  team_jersey_number pace  shooting  wage_eur  Growth
0                    10                10.0   87      92.0    565000       0
1                     7                 7.0   90      93.0    405000       0
2                    10                10.0   91      85.0    290000       0
3                     1                13.0   68       NaN    125000       2
4                    10                 7.0   91      83.0    470000       0
..                  …                 …  …       …       …      …
90                    4                44.0   82      25.0    105000       1
91                   10                13.0   68       NaN    175000       1
92                    8                27.0   83      73.0    135000       1
93                   10                36.0   75      84.0    115000       1
94                    2                 2.0   81      47.0    205000       1

[95 rows x 16 columns]
```

[ ]:
```
import jovian
```

[ ]:
```
jovian.commit()
```

```
<IPython.core.display.Javascript object>
```

```
[jovian] Attempting to save notebook..
[jovian] Updating notebook "xdiad47/fifa-2020-to-2021-players-data" on
https://jovian.ml/
[jovian] Uploading notebook..
[jovian] Capturing environment..
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

[ ]:
```
'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'
```

1.2 Section-2: Exploratory Analysis and Visualization

This section will consist of general analysis of the New_FIFA_df dataframe

[ ]:
```
New_FIFA_df
```

[ ]:
```
    sofifa_id         short_name  age         dob nationality  \
0      158023           L. Messi   32  1987-06-24   Argentina
1       20801  Cristiano Ronaldo   34  1985-02-05    Portugal
2      190871          Neymar Jr   27  1992-02-05      Brazil
3      200389           J. Oblak   26  1993-01-07    Slovenia
4      183277          E. Hazard   28  1991-01-07     Belgium
..        …                 …  …         …         …
90     192774         K. Manolas   28  1991-06-14      Greece
91     194404               Neto   29  1989-07-19      Brazil
92     197445           D. Alaba   27  1992-06-24     Austria
93     201399          M. Icardi   26  1993-02-19   Argentina
94     204963           Carvajal   27  1992-01-11       Spain

                   club  overall  potential  value_eur player_positions  \
0          FC Barcelona       94         94   95500000       RW, CF, ST
1              Juventus       93         93   58500000           ST, LW
2   Paris Saint-Germain       92         92  105500000          LW, CAM
3       Atlético Madrid       91         93   77500000               GK
4          Real Madrid       91         91   90000000           LW, CF
..                 …      …        …         …               …
90              Napoli       85         86   37500000               CB
91         FC Barcelona       85         86   31000000               GK
92    FC Bayern München       85         86   38000000           LB, CB
93                Inter       85         86   46000000               ST
94         Real Madrid       85         86   38000000               RB

   nation_jersey_number  team_jersey_number pace  shooting  wage_eur  Growth
0                    10                10.0   87      92.0    565000       0
1                     7                 7.0   90      93.0    405000       0
2                    10                10.0   91      85.0    290000       0
3                     1                13.0   68       NaN    125000       2
4                    10                 7.0   91      83.0    470000       0
..                  …                  …  …       …       …
90                    4                44.0   82      25.0    105000       1
91                   10                13.0   68       NaN    175000       1
92                    8                27.0   83      73.0    135000       1
93                   10                36.0   75      84.0    115000       1
94                    2                 2.0   81      47.0    205000       1

[95 rows x 16 columns]
```

[ ]:
```
#Number of players in the data frame
New_FIFA_df.sofifa_id.count()
```

[ ]:
```
95
```

[ ]:
```
New_FIFA_df.overall.unique()
```

[ ]:
```
array([94, 93, 92, 91, 90, 89, 88, 87, 86, 85])
```

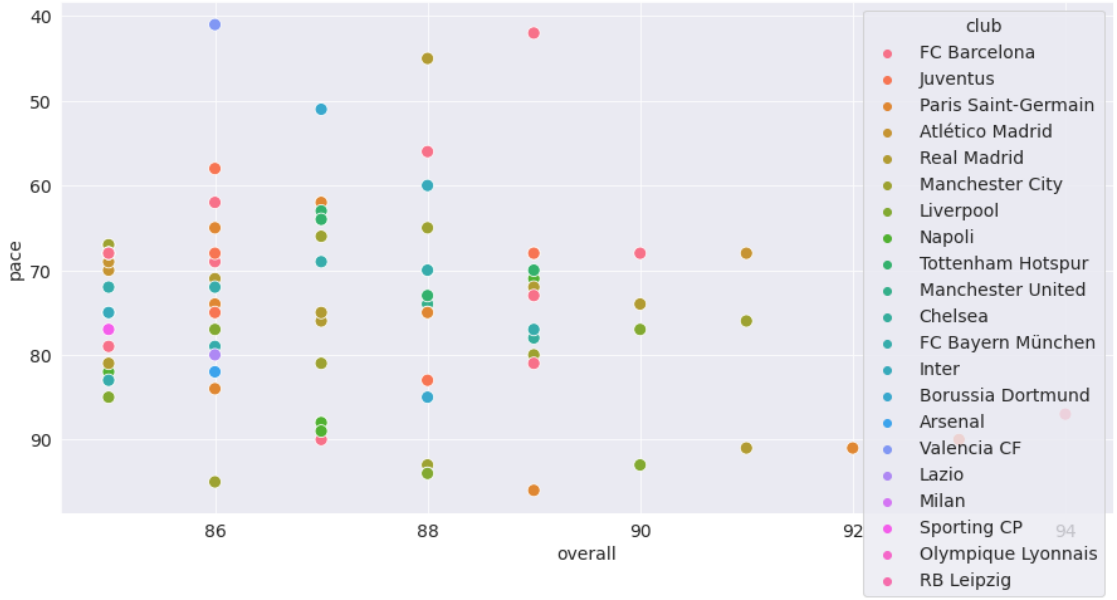
[ ]:
```
#Finding the number Football nation in the wolrd that are affilated to FIFA␣
 ↪federation
#Datas are fetch from original dataset i.e., FIFA_21_df
print(FIFA_21_df.nationality.unique())
print('{} countries are affilated to FIFA federation in 2020-2021'.
 ↪format(FIFA_21_df.nationality.unique().shape[0]))
```

```
['Argentina' 'Portugal' 'Brazil' 'Slovenia' 'Belgium' 'Germany'
 'Netherlands' 'Croatia' 'Egypt' 'France' 'Senegal' 'England' 'Spain'
 'Italy' 'Uruguay' 'Poland' 'Denmark' 'Gabon' 'Korea Republic'
 'Costa Rica' 'Slovakia' 'Bosnia Herzegovina' 'Serbia' 'Scotland'
 'Hungary' 'Switzerland' 'Greece' 'Austria' 'Morocco' 'Sweden' 'Wales'
 'Colombia' 'Czech Republic' 'Chile' 'Algeria' 'Ivory Coast' 'Togo'
 'Norway' 'Mexico' 'Iceland' 'Finland' 'Jamaica' 'Albania' 'Guinea'
 'Cameroon' 'Ghana' 'Montenegro' 'Ukraine' 'Russia' 'DR Congo'
 'Central African Rep.' 'Venezuela' 'Nigeria' 'Armenia' 'Israel' 'Ecuador'
 'Paraguay' 'Australia' 'Turkey' 'Romania' 'Japan' 'Mali' 'United States'
 'Kosovo' 'Dominican Republic' 'Tanzania' 'China PR' 'Northern Ireland'
 'Republic of Ireland' 'Tunisia' 'Cape Verde' 'FYR Macedonia'
 'Burkina Faso' 'Kenya' 'Angola' 'South Africa' 'Peru' 'Syria' 'Gambia'
 'New Zealand' 'Equatorial Guinea' 'Zimbabwe' 'Georgia' 'Canada' 'Estonia'
 'Benin' 'Bulgaria' 'Mozambique' 'Honduras' 'Guinea Bissau' 'Iran'
 'Philippines' 'Cyprus' 'Madagascar' 'Uzbekistan' 'Moldova' 'Cuba'
 'Sierra Leone' 'Curacao' 'Zambia' 'Congo' 'Bolivia' 'Comoros' 'Iraq'
 'Chad' 'Lithuania' 'Saudi Arabia' 'Panama' 'Libya' 'Bahrain'
 'St Kitts Nevis' 'New Caledonia' 'Luxembourg' 'Trinidad & Tobago'
 'Thailand' 'United Arab Emirates' 'Eritrea' 'Korea DPR' 'El Salvador'
 'Azerbaijan' 'Latvia' 'Montserrat' 'Puerto Rico' 'Bermuda'
 'São Tomé & Príncipe' 'Antigua & Barbuda' 'Burundi' 'Kazakhstan'
 'Liberia' 'Guyana' 'Haiti' 'Jordan' 'Faroe Islands' 'Mauritania'
 'Namibia' 'Rwanda' 'Uganda' 'Hong Kong' 'Chinese Taipei' 'Belize'
 'Palestine' 'Mauritius' 'Guam' 'Suriname' 'Lebanon' 'Guatemala' 'Sudan'
 'Liechtenstein' 'Grenada' 'St Lucia' 'Afghanistan' 'Ethiopia' 'Barbados'
 'India' 'Malta' 'Niger' 'Vietnam' 'Malawi' 'Gibraltar' 'Macau'
 'South Sudan' 'Indonesia']
162 countries are affilated to FIFA federation in 2020-2021
```

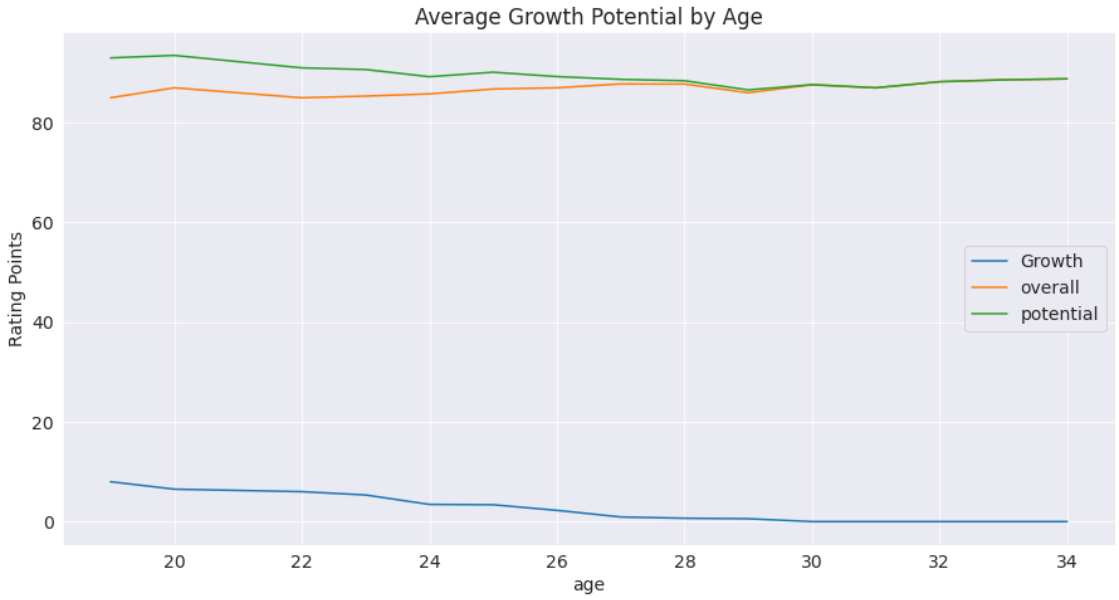
1.2.1 Displaying the pace and overall stats in scatterplot distribution of Clubs.

[ ]:
```
sns.scatterplot(New_FIFA_df.overall,
                New_FIFA_df.pace,
                hue=New_FIFA_df.club,
                s=100,
```

```
                data=New_FIFA_df);
```

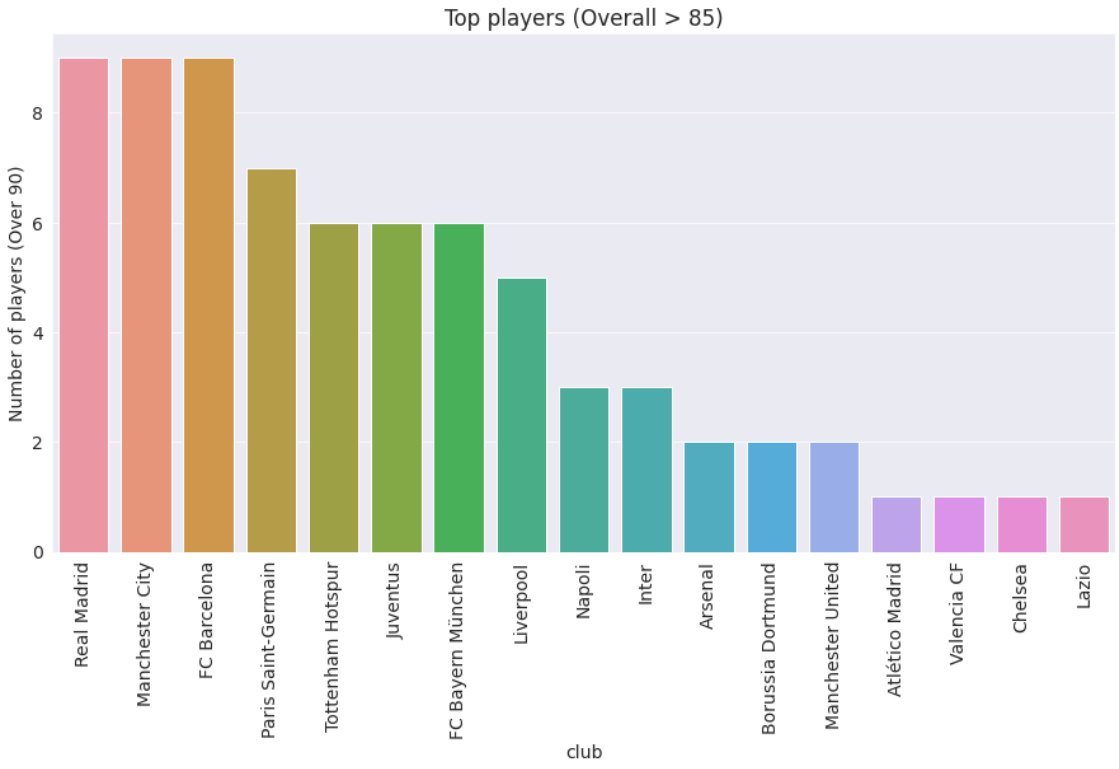


[ ]:
```
fifa_growth = New_FIFA_df.groupby(['age'])['Growth'].mean()
fifa_overall = New_FIFA_df.groupby(['age'])['overall'].mean()
fifa_potential = New_FIFA_df.groupby(['age'])['potential'].mean()

summary = pd.concat([fifa_growth, fifa_overall, fifa_potential], axis=1)

axis = summary.plot()
axis.set_ylabel('Rating Points')
axis.set_title('Average Growth Potential by Age');
```



1.2.2 Let's find clubs with most players rated over 85

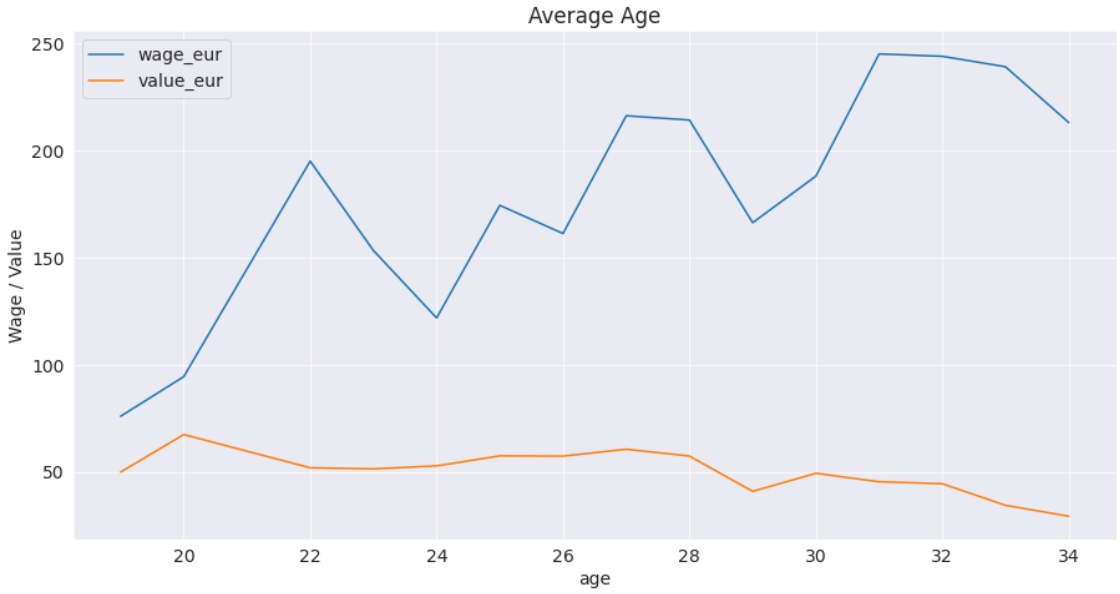
[ ]:
```
#clubs with most players rated over 85
#Finding the best players by overall rating

cutoff = 85
players = New_FIFA_df[New_FIFA_df['overall']>cutoff]
grouped_players = New_FIFA_df[New_FIFA_df['overall']>cutoff].groupby('club')
number_of_players = grouped_players.count()['short_name'].sort_values(ascending␣
↪= False)

ax = sns.countplot(x = 'club', data = players, order = number_of_players.index)

ax.set_xticklabels(labels = number_of_players.index, rotation='vertical')
ax.set_ylabel('Number of players (Over 90)')
ax.set_xlabel('club')
ax.set_title('Top players (Overall > %.i)' %cutoff);
```



[ ]:

1.2.3 Let's explore value against age

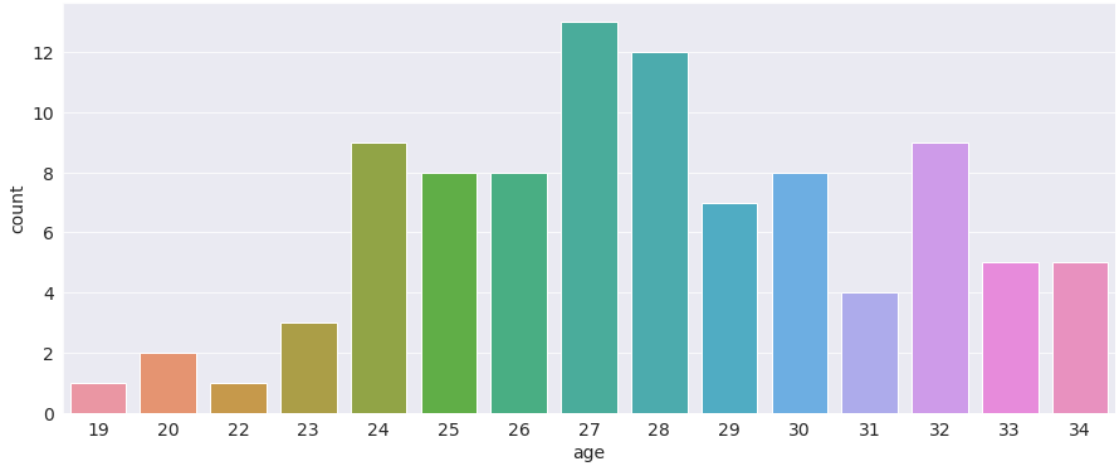
[ ]:
```
fifa_wage_a = New_FIFA_df.groupby(['age'])['wage_eur'].mean()
fifa_value_a = New_FIFA_df.groupby(['age'])['value_eur'].mean()
fifa_wage_a = fifa_wage_a.apply(lambda x: x/1000)
fifa_value_a = fifa_value_a.apply(lambda x: x/1000000)
summary = pd.concat([fifa_wage_a, fifa_value_a], axis=1)

axis = summary.plot()
axis.set_ylabel('Wage / Value')
axis.set_title('Average Age');
```



Here the above distribution we can see that between 20 to 28 age of players has highest market value.

1.2.4 let's find out the majority of player's age.

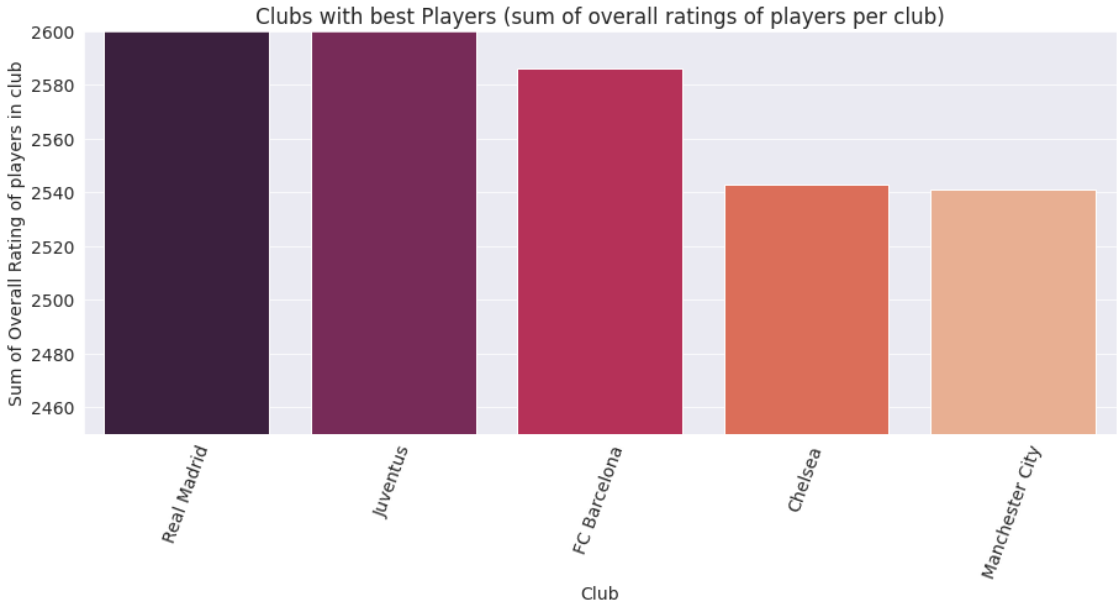
[ ]:
```
plt.figure(figsize=(15,6))
sns.countplot(x="age",data=New_FIFA_df);
```



From the above visualization, that the majority of players are between the age of 24 and 28, with the largest peak of 27 years.

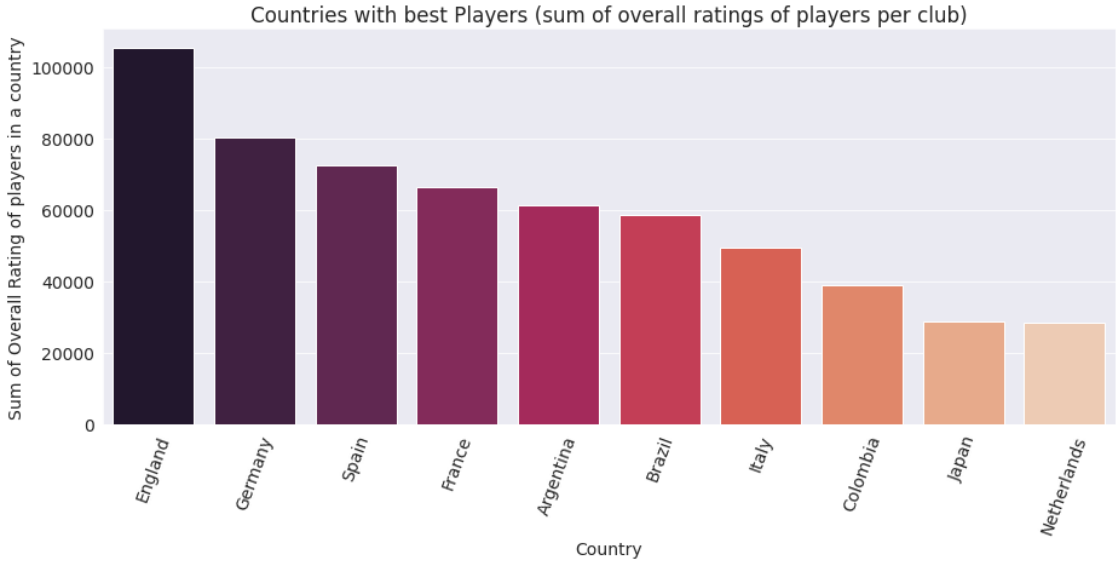
1.2.5 Let's find out the clubs with best players.

[ ]:
```
best_dict = {}
for club in FIFA_21_df['club'].unique():
    overall_rating = FIFA_21_df['overall'][FIFA_21_df['club'] == club].sum()
    best_dict[club] = overall_rating
best_club = pd.DataFrame.from_dict(best_dict,orient='index', columns =␣
 ↪['overall'])
best_club['club'] = best_club.index
best_club = best_club.sort_values(by = 'overall' , ascending =  False)

plt.figure(1 , figsize = (15 , 6))
sns.barplot(x = 'club' , y  = 'overall' , data = best_club.
 ↪head(5),palette='rocket')
plt.xticks(rotation = 70)
plt.xlabel("Club")
plt.ylabel('Sum of Overall Rating of players in club')
plt.title('Clubs with best Players (sum of overall ratings of players per␣
 ↪club)')
plt.ylim(2450 , 2600)
plt.show()
```



1.2.6 Countries with best Players

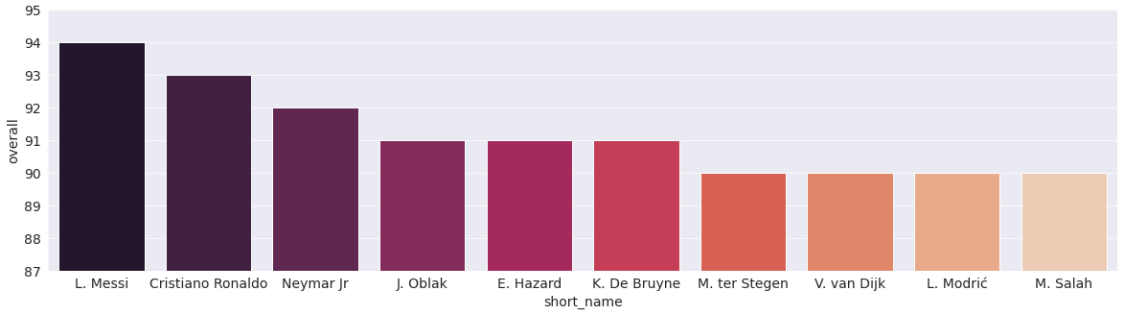
```
[ ]: best_dict = {}
for country in FIFA_21_df['nationality'].unique():
    overall_rating = FIFA_21_df['overall'][FIFA_21_df['nationality'] ==␣
 ↪country].sum()
    best_dict[country] = overall_rating
best_country = pd.DataFrame.from_dict(best_dict,orient='index', columns =␣
 ↪['overall'])
best_country['club'] = best_country.index
best_country = best_country.sort_values(by = 'overall' , ascending =  False)

plt.figure(1 , figsize = (15 , 6))
sns.barplot(x = 'club' , y  = 'overall' , data = best_country.
 ↪head(10),palette='rocket')
plt.xticks(rotation = 70)
plt.xlabel("Country")
plt.ylabel('Sum of Overall Rating of players in a country')
plt.title('Countries with best Players (sum of overall ratings of players per␣
 ↪club)')
plt.show()
```



By the above visualization we realise that England just has lot of players.

1.2.7 Top ten best players in the world

[ ]:
```
#Top ten best players in the world
fifa_best_players = pd.DataFrame.copy(FIFA_21_df.sort_values(by = 'overall' ,␣
 ↪ascending = False ).head(10))

plt.figure(1 , figsize = (20 , 5))
sns.barplot(x ='short_name' , y = 'overall' , data =␣
 ↪fifa_best_players,palette='rocket')

plt.ylim(87 , 95)
plt.show()
```



Here we can see L.Messi is the best player in the world for current year i.e., 2020

[ ]:
```

```

[14]:
```
import jovian
```

[16]:
```
jovian.commit(project=project_name)
```

```
[jovian] Detected Colab notebook…
[jovian] Uploading colab notebook to Jovian…
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

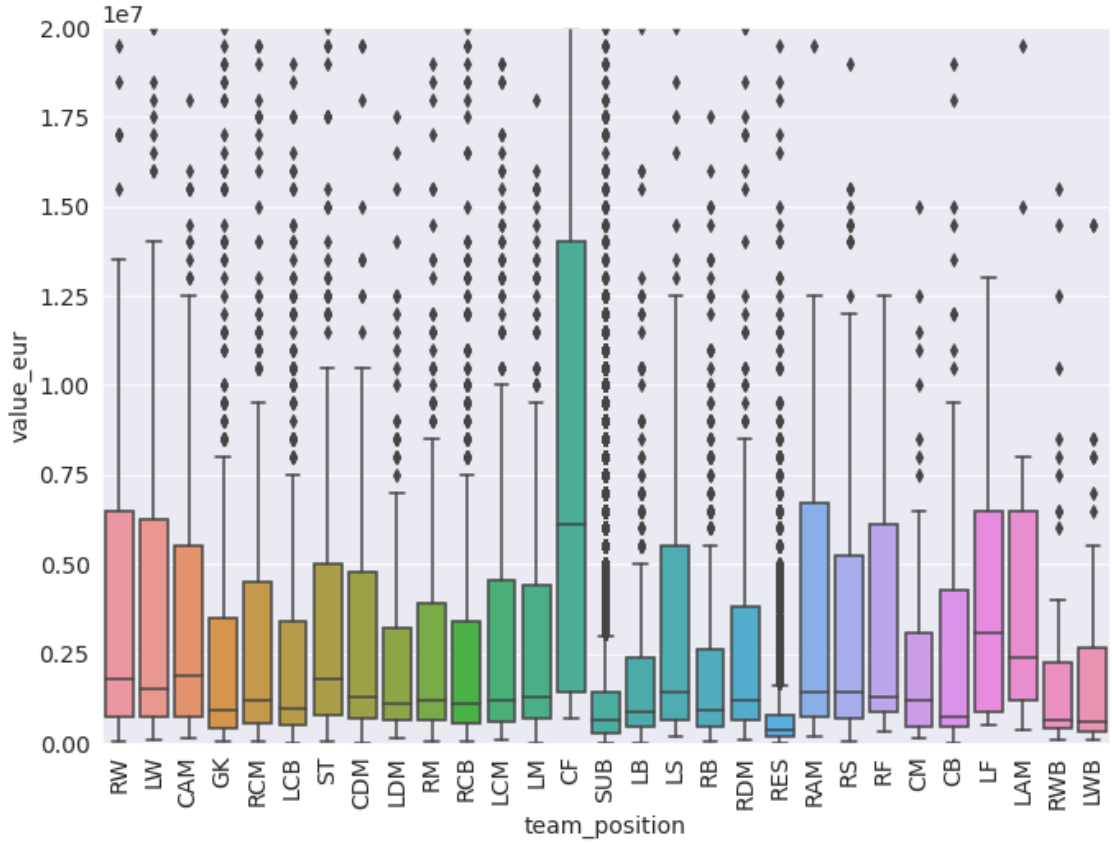
[16]:
```
'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'
```

1.3 Section-3: Interesting Questions

In this section the following questions will be asked: 1. Does position have a say in a player's value? 2. Display top 5 highest earner in Football world, along with their club name, nationality. 3. Count the overall rating in age distribution in famous clubs by country. 4. By the above distribution show the overall rating distribution in famous clubs by country. 5. Find the eldest players and youngest players in the club. 6. How do market values of players compare across top clubs?

1.3.1 Q1. Does position have a say in a player's value?

```
[ ]: avg_value_by_position = FIFA_21_df.groupby('team_position')['value_eur'].mean()
plt.figure(figsize=(11,8))
p = sns.boxplot(x = 'team_position', y = 'value_eur', data = FIFA_21_df)
p = plt.xticks(rotation=90)
p = plt.ylim(0, 20000000)
```



Turns out, Players playing at CF i.e., Center forward, LW i.e., left winger and RW i.e., Right winger have highest market value, followed by midfielders positions.

1.3.2 Q2. Display top 5 highest earner in Football world, along with their club name, nationality.

```
[ ]: #function for Highest earner

def normalizing_wage(x):
    if '€' in str(x) and 'M' in str(x):
        c = str(x).replace('€' , '')
        c = str(c).replace('M' , '')
        c = float(c) * 1000000
```

```
    else:
        c = str(x).replace('€' , '')
        c = str(c).replace('K' , '')
        c = float(c) * 1000

    return c

New_FIFA_df['Normalized_Wage'] = New_FIFA_df['wage_eur'].apply(lambda x :␣
↪normalizing_wage(x))
```

[ ]:
```
New_FIFA_df.sort_values(by = 'Normalized_Wage' , ascending =␣
↪False)[['short_name','club','nationality','overall',
'age','Normalized_Wage','wage_eur']].head(5)
```

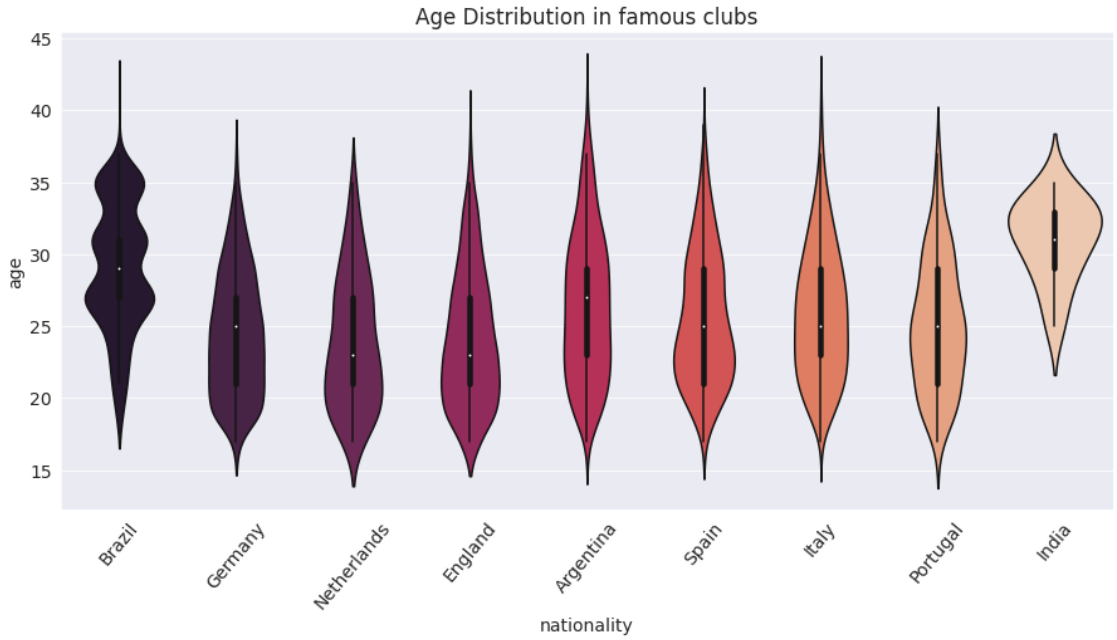
[ ]:
```
           short_name             club nationality  overall  age  \
0            L. Messi     FC Barcelona   Argentina       94   32
4          E. Hazard      Real Madrid     Belgium       91   28
1   Cristiano Ronaldo         Juventus    Portugal       93   34
22       A. Griezmann     FC Barcelona      France       89   28
5        K. De Bruyne  Manchester City     Belgium       91   28

    Normalized_Wage  wage_eur
0        565000000.0    565000
4        470000000.0    470000
1        405000000.0    405000
22       370000000.0    370000
5        370000000.0    370000
```

1.3.3 Q3. Count the overall rating in age distribution in famous clubs by country.

[ ]:
```
#let's count the overall rating by country
countries = ['England' , 'Brazil' , 'Portugal' ,'Argentina',
             'Italy' , 'Spain' , 'Germany' ,'Netherlands','India']
```

We took these above countries to see the age distribution in famous clubs.

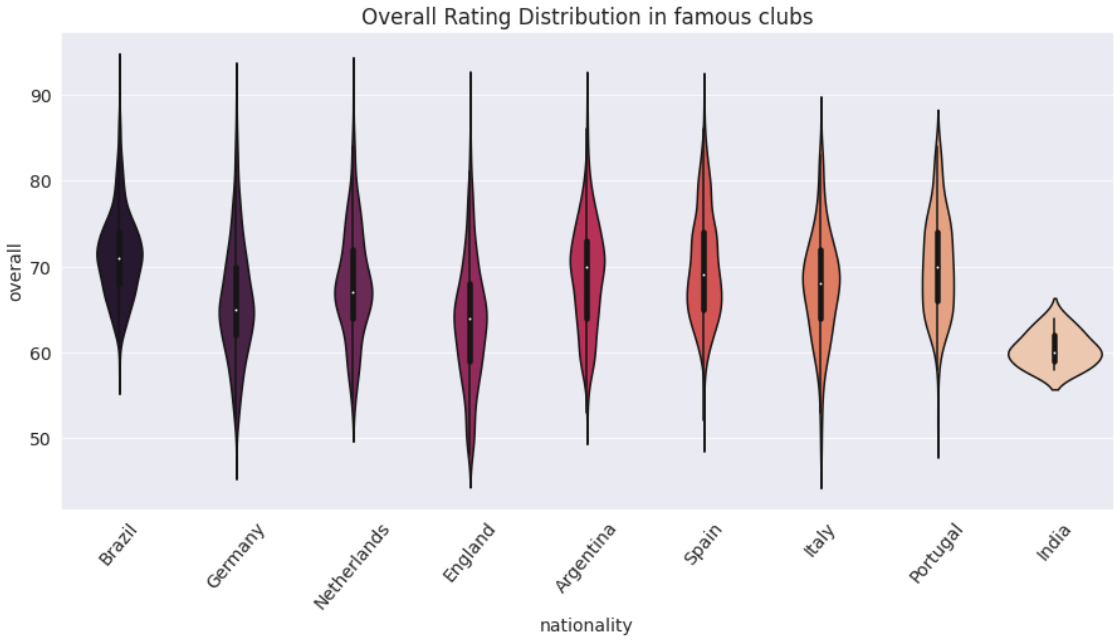
[ ]:
```
fifa_country_age = FIFA_21_df.loc[FIFA_21_df['nationality'].isin(countries) &␣
↪FIFA_21_df['age']]
plt.figure(1 , figsize = (15 ,7))
sns.violinplot(x = 'nationality' , y = 'age' , data = fifa_country_age,␣
↪palette='rocket')
plt.title('Age Distribution in famous clubs')
plt.xticks(rotation = 50)
plt.show()
```



As we can see from the above distibution, Argentina and Italy aren't too far behind.

1.3.4 Q4. By the above distribution show the overall rating distribution in famous clubs by country.

[ ]:
```
fifa_country_rating = FIFA_21_df.loc[FIFA_21_df['nationality'].isin(countries)␣
 ↪& FIFA_21_df['overall']]
plt.figure(1 , figsize = (15 ,7))
sns.violinplot(x = 'nationality' , y = 'overall' , data = fifa_country_age,␣
 ↪palette='rocket')
plt.title('Overall Rating Distribution in famous clubs')
plt.xticks(rotation = 50)
plt.show()
```



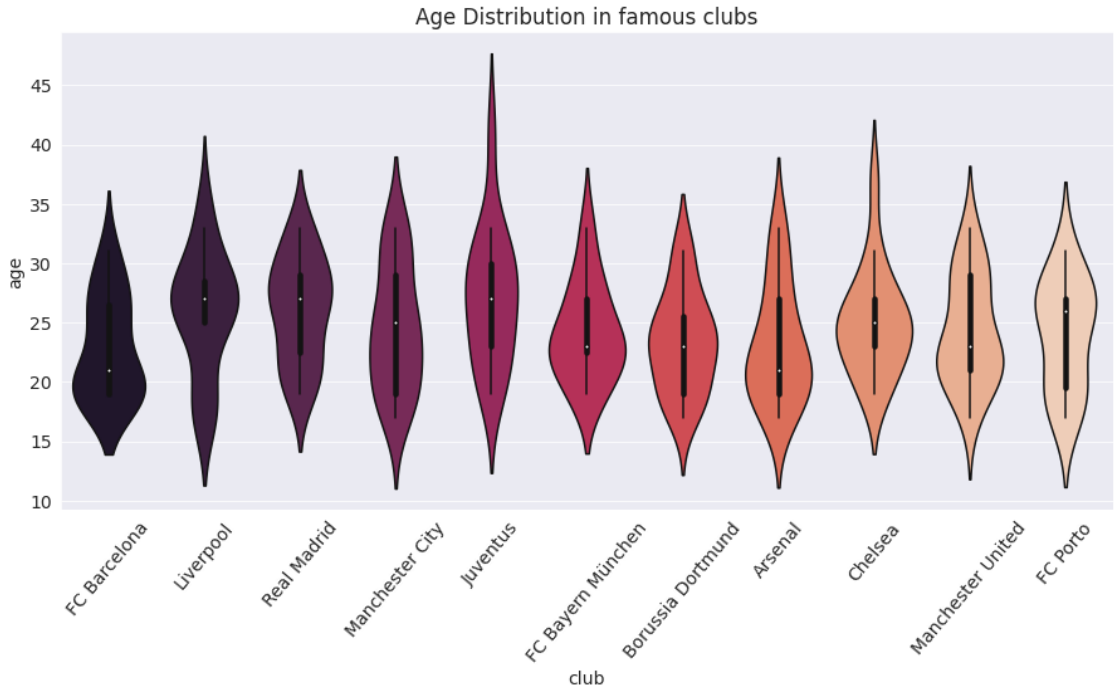
Italy, Portugal, spain really seems to be pretty dominant here. Must be the experienced players.

1.3.5 Q5. Find the eldest players and youngest players in the club.

We will find out the youngest talent as well as experienced players. We will do by age distribution for the following top european clubs.

[ ]:
```
#Age Distribution
#Let's start with the age distribution in these clubs.
clubs = ['Chelsea' , 'Arsenal', 'Juventus', 'Paris Sain-Germain' ,'FC Bayern␣
↪München',
      'Real Madrid' , 'FC Barcelona' , 'Borussia Dortmund' , 'Manchester␣
↪United' ,
      'FC Porto', 'Liverpool', 'Manchester City']
```

[ ]:
```
fifa_club_age = FIFA_21_df.loc[FIFA_21_df['club'].isin(clubs) &␣
↪FIFA_21_df['age']]
plt.figure(1 , figsize = (15 ,7))
sns.violinplot(x = 'club' , y = 'age' , data = fifa_club_age,palette='rocket')
plt.title('Age Distribution in famous clubs')
plt.xticks(rotation = 50)
plt.show()
```

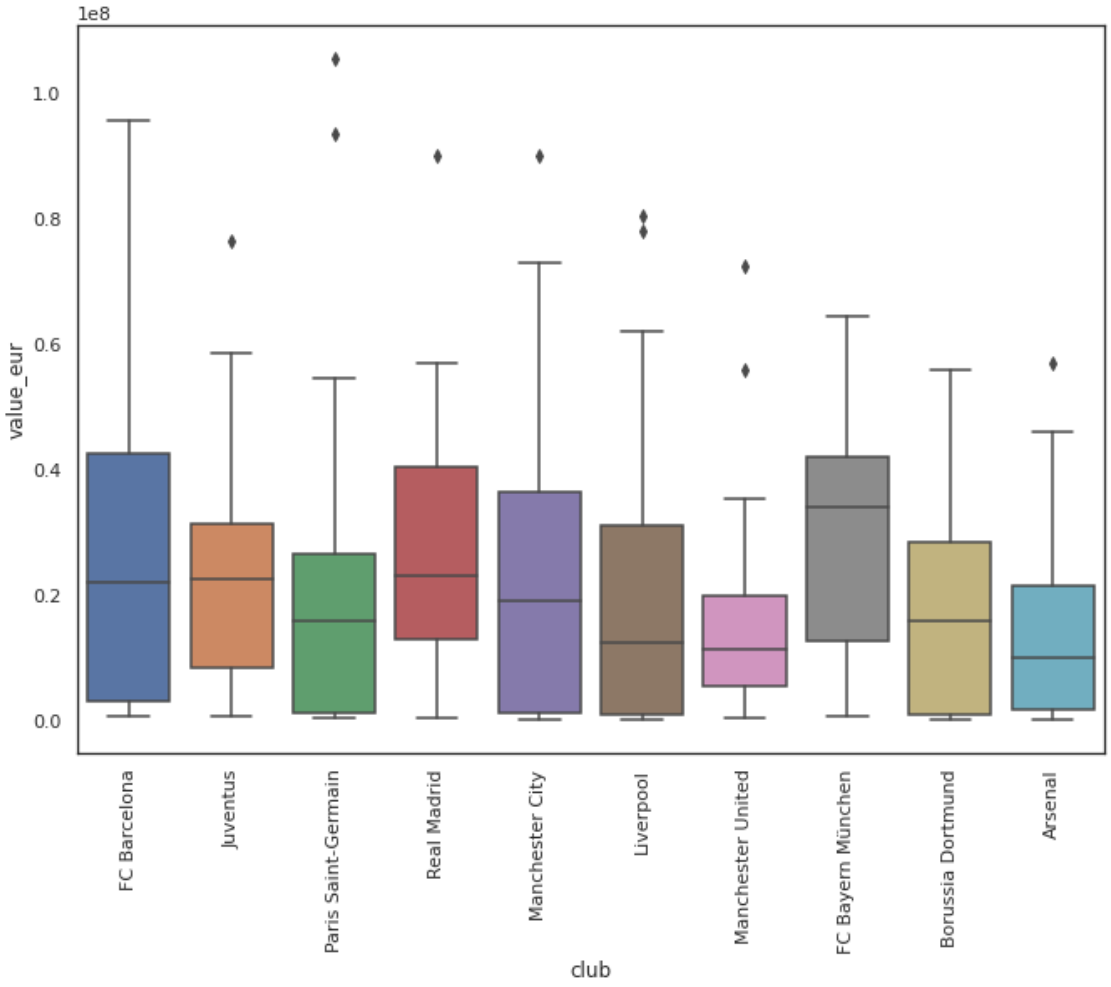


As we can see FC Barcelona, FC Porto and Borussia Dortmund have the most young talent. Juventus, liverpool and chelsea have most experienced players.

1.3.6 Q6. How do market values of players compare across top clubs?

We will take some of the top 10 famous clubs in the world.

[ ]:
```
# getting top ten clubs
top_ten_clubs = ['FC Barcelona', 'Real Madrid', 'Manchester City', 'Arsenal',␣
 ↪'Liverpool', 'Manchester United', 'Borussia Dortmund', 'FC Bayern München',␣
 ↪'Juventus', 'Paris Saint-Germain']
top_ten_clubs_data = FIFA_21_df.loc[FIFA_21_df['club'].isin(top_ten_clubs), :]
sns.set(style="white")
plt.figure(figsize=(11,8))
p = sns.boxplot(x = 'club', y = 'value_eur', data = top_ten_clubs_data)
p = plt.xticks(rotation=90)
```



Currently FC Barcelona have the most expensive players followed by manchester city, FC Bayern Munchen and Liverpool in the year 2020

[20]:
```
import jovian
```

[21]:
```
jovian.commit(project=project_name)
```

```
[jovian] Detected Colab notebook…
[jovian] Uploading colab notebook to Jovian…
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

[21]:
```
'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'
```

1.4 Section-4: Summary

The above analysis on player_20.csv can be summarised as below:

1. Fc barcelona followed by Juventus, paris-saint germain has good pace and overall rating players.
2. Players between 20 to 28 of age have growth as well potential.
3. We also get clubs with most players rated over 85 and they are Real madrid, Manchester City and FC Barcelona.
4. Between 20 to 28 age of players has highest market value.
5. From the visualization, the majority of players are between the age of 24 to 28, with the largest peak of 27 years.
6. Real madrid, Juventus, FC Barcelona clubs have the best players in the football world.
7. From the visualization England followed by germany and spain have the highest number of football players in the world.
8. L.Messi is the best player in the world for current year i.e., 2020.
9. L. Messi is the highest earner followed by E.Hazard and Cristiano Ronaldo.
10. Italy, Portugal, spain really seems to be pretty dominant in football game and have the experienced players.
11. Currently FC Barcelona have the most expensive players followed by manchester city, FC Bayern Munchen and Liverpool in the year 2020

[22]:
```
import jovian
```

[23]:
```
jovian.commit(project=project_name)
```

```
[jovian] Detected Colab notebook…
[jovian] Uploading colab notebook to Jovian…
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

[23]:
```
'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'
```

1.5 Future Work

The dataset for FIFA(2020-2021) consisted of two .csv files i.e., players_20.csv and data.csv. Only players_20.csv is considered in this project for the analysis. For data related to overall ratings, potential ratings, prefered position etc can be retrieved from data.csv ## References Check out the various resource used in this project:

ZerotoPandas Tutorial by JovianML- Data Analysis with Python: Zero to Pandas

Numpy Documentation: https://numpy.org/devdocs/user/quickstart.html

User guide for Pandas: https://pandas.pydata.org/docs/user_guide/index.html

Seaborn gallery: https://seaborn.pydata.org/examples/index.html

Matplotlib gallery: https://matplotlib.org/3.1.1/gallery/index.html

[24]:
```
import jovian
```

[25]:
```
jovian.commit(project=project_name)
```

```
[jovian] Detected Colab notebook…
[jovian] Uploading colab notebook to Jovian…
[jovian] Committed successfully!
https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data
```

[25]: 'https://jovian.ml/xdiad47/fifa-2020-to-2021-players-data'

[ ]: