PRACTICE-PROJECT2: FIFA_PALYER_2019_ANALYSIS_WITH_EDA

FIFA_PALYER_2019_ANALYSIS_WITH_EDA:

in this project we visualize players Age, Nationality, Value, Preferred Foot, International Reputation, Weak Foot, Position, Wage, Skill Moves, Height, Weight, Work Rate, Special, Potential, Overall, Body Type. we will visualize this with the help of barplot and pie chart

import libraries `

In [1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

Load the Datasets

In [2]:
```
fifa_data = pd.read_csv('Footballer.csv')
```

In [3]:
```
fifa_data.head()
```

Out[3]:

	Unnamed: 0	ID	Name	Age	Photo	Nationality	
0	0	158023	L. Messi	31	https://cdn.sofifa.org/players/4/19/158023.png	Argentina	https://cdn.sofifa.org/fla
1	1	20801	Cristiano Ronaldo	33	https://cdn.sofifa.org/players/4/19/20801.png	Portugal	https://cdn.sofifa.org/fla
2	2	190871	Neymar Jr	26	https://cdn.sofifa.org/players/4/19/190871.png	Brazil	https://cdn.sofifa.org/fla
3	3	193080	De Gea	27	https://cdn.sofifa.org/players/4/19/193080.png	Spain	https://cdn.sofifa.org/fla
4	4	192985	K. De Bruyne	27	https://cdn.sofifa.org/players/4/19/192985.png	Belgium	https://cdn.sofifa.org/fl

5 rows × 89 columns

In [4]:
```
fifa_data.shape
```

Out[4]:
```
(18207, 89)
```

In [5]:
```
fifa_data.describe().T#we will arrange describe function in vertically
```

Out[5]:

	count	mean	std	min	25%	50%	75%	max
Unnamed: 0	18207.0	9103.000000	5256.052511	0.0	4551.5	9103.0	13654.5	18206.0
ID	18207.0	214298.338606	29965.244204	16.0	200315.5	221759.0	236529.5	246620.0
Age	18207.0	25.122206	4.669943	16.0	21.0	25.0	28.0	45.0
Overall	18207.0	66.238699	6.908930	46.0	62.0	66.0	71.0	94.0
Potential	18207.0	71.307299	6.136496	48.0	67.0	71.0	75.0	95.0
Special	18207.0	1597.809908	272.586016	731.0	1457.0	1635.0	1787.0	2346.0
International Reputation	18159.0	1.113222	0.394031	1.0	1.0	1.0	1.0	5.0
Weak Foot	18159.0	2.947299	0.660456	1.0	3.0	3.0	3.0	5.0
Skill Moves	18159.0	2.361308	0.756164	1.0	2.0	2.0	3.0	5.0
Jersey Number	18147.0	19.546096	15.947765	1.0	8.0	17.0	26.0	99.0
Crossing	18159.0	49.734181	18.364524	5.0	38.0	54.0	64.0	93.0
Finishing	18159.0	45.550911	19.525820	2.0	30.0	49.0	62.0	95.0
HeadingAccuracy	18159.0	52.298144	17.379909	4.0	44.0	56.0	64.0	94.0
ShortPassing	18159.0	58.686712	14.699495	7.0	54.0	62.0	68.0	93.0
Volleys	18159.0	42.909026	17.694408	4.0	30.0	44.0	57.0	90.0
Dribbling	18159.0	55.371001	18.910371	4.0	49.0	61.0	68.0	97.0
Curve	18159.0	47.170824	18.395264	6.0	34.0	48.0	62.0	94.0
FKAccuracy	18159.0	42.863153	17.478763	3.0	31.0	41.0	57.0	94.0
LongPassing	18159.0	52.711933	15.327870	9.0	43.0	56.0	64.0	93.0
BallControl	18159.0	58.369459	16.686595	5.0	54.0	63.0	69.0	96.0
Acceleration	18159.0	64.614076	14.927780	12.0	57.0	67.0	75.0	97.0
SprintSpeed	18159.0	64.726967	14.649953	12.0	57.0	67.0	75.0	96.0
Agility	18159.0	63.503607	14.766049	14.0	55.0	66.0	74.0	96.0
Reactions	18159.0	61.836610	9.010464	21.0	56.0	62.0	68.0	96.0
Balance	18159.0	63.966573	14.136166	16.0	56.0	66.0	74.0	96.0
ShotPower	18159.0	55.460047	17.237958	2.0	45.0	59.0	68.0	95.0
Jumping	18159.0	65.089432	11.820044	15.0	58.0	66.0	73.0	95.0
Stamina	18159.0	63.219946	15.894741	12.0	56.0	66.0	74.0	96.0
Strength	18159.0	65.311967	12.557000	17.0	58.0	67.0	74.0	97.0
LongShots	18159.0	47.109973	19.260524	3.0	33.0	51.0	62.0	94.0
Aggression	18159.0	55.868991	17.367967	11.0	44.0	59.0	69.0	95.0
Interceptions	18159.0	46.698276	20.696909	3.0	26.0	52.0	64.0	92.0
Positioning	18159.0	49.958478	19.529036	2.0	38.0	55.0	64.0	95.0
Vision	18159.0	53.400903	14.146881	10.0	44.0	55.0	64.0	94.0
Penalties	18159.0	48.548598	15.704053	5.0	39.0	49.0	60.0	92.0
Composure	18159.0	58.648274	11.436133	3.0	51.0	60.0	67.0	96.0
Marking	18159.0	47.281623	19.904397	3.0	30.0	53.0	64.0	94.0
StandingTackle	18159.0	47.697836	21.664004	2.0	27.0	55.0	66.0	93.0
Tackle	18159.0	45.661435	21.289135	3.0	24.0	52.0	64.0	91.0

	count	mean	std	min	25%	50%	75%	max
GKDiving	18159.0	16.616223	17.695349	1.0	8.0	11.0	14.0	90.0
GKHandling	18159.0	16.391596	16.906900	1.0	8.0	11.0	14.0	92.0
GKKicking	18159.0	16.232061	16.502864	1.0	8.0	11.0	14.0	91.0
GKPositioning	18159.0	16.388898	17.034669	1.0	8.0	11.0	14.0	90.0
GKReflexes	18159.0	16.710887	17.955119	1.0	8.0	11.0	14.0	94.0

show all data-types

In [6]:
```
fifa_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18207 entries, 0 to 18206
Data columns (total 89 columns):
 #   Column                    Non-Null Count  Dtype  
---  ------                    --------------  -----  
 0   Unnamed: 0                18207 non-null  int64  
 1   ID                        18207 non-null  int64  
 2   Name                      18207 non-null  object 
 3   Age                       18207 non-null  int64  
 4   Photo                     18207 non-null  object 
 5   Nationality               18207 non-null  object 
 6   Flag                      18207 non-null  object 
 7   Overall                   18207 non-null  int64  
 8   Potential                 18207 non-null  int64  
 9   Club                      17966 non-null  object 
 10  Club Logo                 18207 non-null  object 
 11  Value                     18207 non-null  object 
 12  Wage                      18207 non-null  object 
 13  Special                   18207 non-null  int64  
 14  Preferred Foot            18159 non-null  object 
 15  International Reputation  18159 non-null  float64
 16  Weak Foot                 18159 non-null  float64
 17  Skill Moves               18159 non-null  float64
 18  Work Rate                 18159 non-null  object 
 19  Body Type                 18159 non-null  object 
 20  Real Face                 18159 non-null  object 
 21  Position                  18147 non-null  object 
 22  Jersey Number             18147 non-null  float64
 23  Joined                    16654 non-null  object 
 24  Loaned From               1264 non-null   object 
 25  Contract Valid Until      17918 non-null  object 
 26  Height                    18159 non-null  object 
 27  Weight                    18159 non-null  object 
 28  LS                        16122 non-null  object 
 29  ST                        16122 non-null  object 
 30  RS                        16122 non-null  object 
 31  LW                        16122 non-null  object 
 32  LF                        16122 non-null  object 
 33  CF                        16122 non-null  object 
 34  RF                        16122 non-null  object 
 35  RW                        16122 non-null  object 
 36  LAM                       16122 non-null  object 
 37  CAM                       16122 non-null  object 
 38  RAM                       16122 non-null  object 
 39  LM                        16122 non-null  object 
 40  LCM                       16122 non-null  object 
 41  CM                        16122 non-null  object 
 42  RCM                       16122 non-null  object 
 43  RM                        16122 non-null  object 
 44  LWB                       16122 non-null  object 
 45  LDM                       16122 non-null  object 
 46  CDM                       16122 non-null  object 
 47  RDM                       16122 non-null  object 
 48  RWB                       16122 non-null  object 
 49  LB                        16122 non-null  object 
 50  LCB                       16122 non-null  object 
 51  CB                        16122 non-null  object 
 52  RCB                       16122 non-null  object 
 53  RB                        16122 non-null  object 
 54  Crossing                  18159 non-null  float64
 55  Finishing                 18159 non-null  float64
 56  HeadingAccuracy           18159 non-null  float64
 57  ShortPassing              18159 non-null  float64
 58  Volleys                   18159 non-null  float64
```

```
 59  Dribbling                 18159 non-null  float64
 60  Curve                     18159 non-null  float64
 61  FKAccuracy                18159 non-null  float64
 62  LongPassing               18159 non-null  float64
 63  BallControl               18159 non-null  float64
 64  Acceleration              18159 non-null  float64
 65  SprintSpeed               18159 non-null  float64
 66  Agility                   18159 non-null  float64
 67  Reactions                 18159 non-null  float64
 68  Balance                   18159 non-null  float64
 69  ShotPower                 18159 non-null  float64
 70  Jumping                   18159 non-null  float64
 71  Stamina                   18159 non-null  float64
 72  Strength                  18159 non-null  float64
 73  LongShots                 18159 non-null  float64
 74  Aggression                18159 non-null  float64
 75  Interceptions             18159 non-null  float64
 76  Positioning               18159 non-null  float64
 77  Vision                    18159 non-null  float64
 78  Penalties                 18159 non-null  float64
 79  Composure                 18159 non-null  float64
 80  Marking                   18159 non-null  float64
 81  StandingTackle            18159 non-null  float64
 82  SlidingTackle             18159 non-null  float64
 83  GKDiving                  18159 non-null  float64
 84  GKHandling                18159 non-null  float64
 85  GKKicking                 18159 non-null  float64
 86  GKPositioning             18159 non-null  float64
 87  GKReflexes                18159 non-null  float64
 88  Release Clause            16643 non-null  object
dtypes: float64(38), int64(6), object(45)
memory usage: 12.4+ MB
```

null values in percentage

In [7]:
```
percent = 100*(fifa_data.isnull().sum()/len(fifa_data))
```

In [8]:
```
percent[:40]
```

Out[8]:
```
Unnamed: 0                  0.000000
ID                          0.000000
Name                        0.000000
Age                         0.000000
Photo                       0.000000
Nationality                 0.000000
Flag                        0.000000
Overall                     0.000000
Potential                   0.000000
Club                        1.323667
Club Logo                   0.000000
Value                       0.000000
Wage                        0.000000
Special                     0.000000
Preferred Foot              0.263635
International Reputation    0.263635
Weak Foot                   0.263635
Skill Moves                 0.263635
Work Rate                   0.263635
Body Type                   0.263635
Real Face                   0.263635
Position                    0.329544
Jersey Number               0.329544
Joined                      8.529686
Loaned From                93.057615
Contract Valid Until        1.587302
Height                      0.263635
Weight                      0.263635
LS                         11.451639
ST                         11.451639
RS                         11.451639
LW                         11.451639
LF                         11.451639
CF                         11.451639
RF                         11.451639
RW                         11.451639
LAM                        11.451639
CAM                        11.451639
RAM                        11.451639
LM                         11.451639
dtype: float64
```

In [9]:
```
percent[41:]
```

Out[9]:
```
CM                 11.451639
RCM                11.451639
RM                 11.451639
LWB                11.451639
LDM                11.451639
CDM                11.451639
RDM                11.451639
RWB                11.451639
LB                 11.451639
LCB                11.451639
CB                 11.451639
RCB                11.451639
RB                 11.451639
Crossing            0.263635
Finishing           0.263635
HeadingAccuracy     0.263635
ShortPassing        0.263635
Volleys             0.263635
Dribbling           0.263635
Curve               0.263635
FKAccuracy          0.263635
LongPassing         0.263635
BallControl         0.263635
Acceleration        0.263635
SprintSpeed         0.263635
Agility             0.263635
Reactions           0.263635
Balance             0.263635
ShotPower           0.263635
Jumping             0.263635
Stamina             0.263635
Strength            0.263635
LongShots           0.263635
Aggression          0.263635
Interceptions       0.263635
Positioning         0.263635
Vision              0.263635
Penalties           0.263635
Composure           0.263635
Marking             0.263635
StandingTackle      0.263635
SlidingTackle       0.263635
GKDiving            0.263635
GKHandling          0.263635
GKKicking           0.263635
GKPositioning       0.263635
GKReflexes          0.263635
Release Clause      8.590103
dtype: float64
```

cleaned dataset

In [10]:
```
fifa_data=fifa_data.dropna(axis=0, subset =['Club'])#club null values is 1.32% so we dro
```

In [11]:
```
fifa_data=fifa_data.dropna(axis=0, subset =['Preferred Foot'])#some columns are related
```

In [12]:
```
null_Val = fifa_data.isnull().sum().sort_values(ascending= False)
```

In [13]:
```
null_Val[:40]#print remaining null values
```

Out[13]:
```
Loaned From       16654
LWB                1992
LCM                1992
RS                 1992
LW                 1992
LF                 1992
CF                 1992
RF                 1992
RW                 1992
LAM                1992
CAM                1992
RAM                1992
LM                 1992
CM                 1992
LS                 1992
RCM                1992
RM                 1992
LDM                1992
CDM                1992
RDM                1992
RWB                1992
LB                 1992
LCB                1992
CB                 1992
RCB                1992
RB                 1992
ST                 1992
Release Clause     1275
Joined             1264
Weight                0
Aggression            0
FKAccuracy            0
LongPassing           0
BallControl           0
Acceleration          0
SprintSpeed           0
Agility               0
Reactions             0
Balance               0
ShotPower             0
dtype: int64
```

fill some important columns null values

In [14]:
```
fifa_data['Loaned From'].dtype
```

Out[14]:
```
dtype('O')
```

In [15]:
```
fifa_data['Loaned From'].fillna('None',inplace=True)
```

In [16]:
```
fifa_data['Loaned From'].value_counts()
```

Out[16]:
```
None                      16654
Atalanta                     20
Sassuolo                     18
Juventus                     17
SL Benfica                   17
                            ...
Sheffield Wednesday           1
Club Necaxa                   1
Royal Excel Mouscron          1
Rayo Vallecano                1
Neuchâtel Xamax               1
Name: Loaned From, Length: 342, dtype: int64
```

In [17]:
```
fifa_data['Joined'].dtype
```

Out[17]:
```
dtype('O')
```

In [18]:
```
fifa_data['Joined'].mode()
```

Out[18]:
```
0    Jul 1, 2018
Name: Joined, dtype: object
```

In [19]:
```
fifa_data['Joined'].fillna('Jul 1, 2018', inplace=True)
```

In [20]:
```
fifa_data['Joined'].value_counts()
```

Out[20]:
```
Jul 1, 2018     2802
Jul 1, 2017     1133
Jan 1, 2018      635
Jul 1, 2016      614
Jul 1, 2015      368
                ...
Jan 4, 2014        1
Apr 20, 2017       1
Nov 10, 2017       1
Dec 11, 2013       1
Jul 4, 2012        1
Name: Joined, Length: 1736, dtype: int64
```

In [21]:
```
fifa_data['Joined'].head(100)
```

Out[21]:
```
0      Jul 1, 2004
1     Jul 10, 2018
2      Aug 3, 2017
3      Jul 1, 2011
4     Aug 30, 2015
          ...
95    Jul 14, 2011
96     Aug 6, 2018
97     Sep 1, 2015
98     Jul 1, 2015
99     Jul 1, 2015
Name: Joined, Length: 100, dtype: object
```

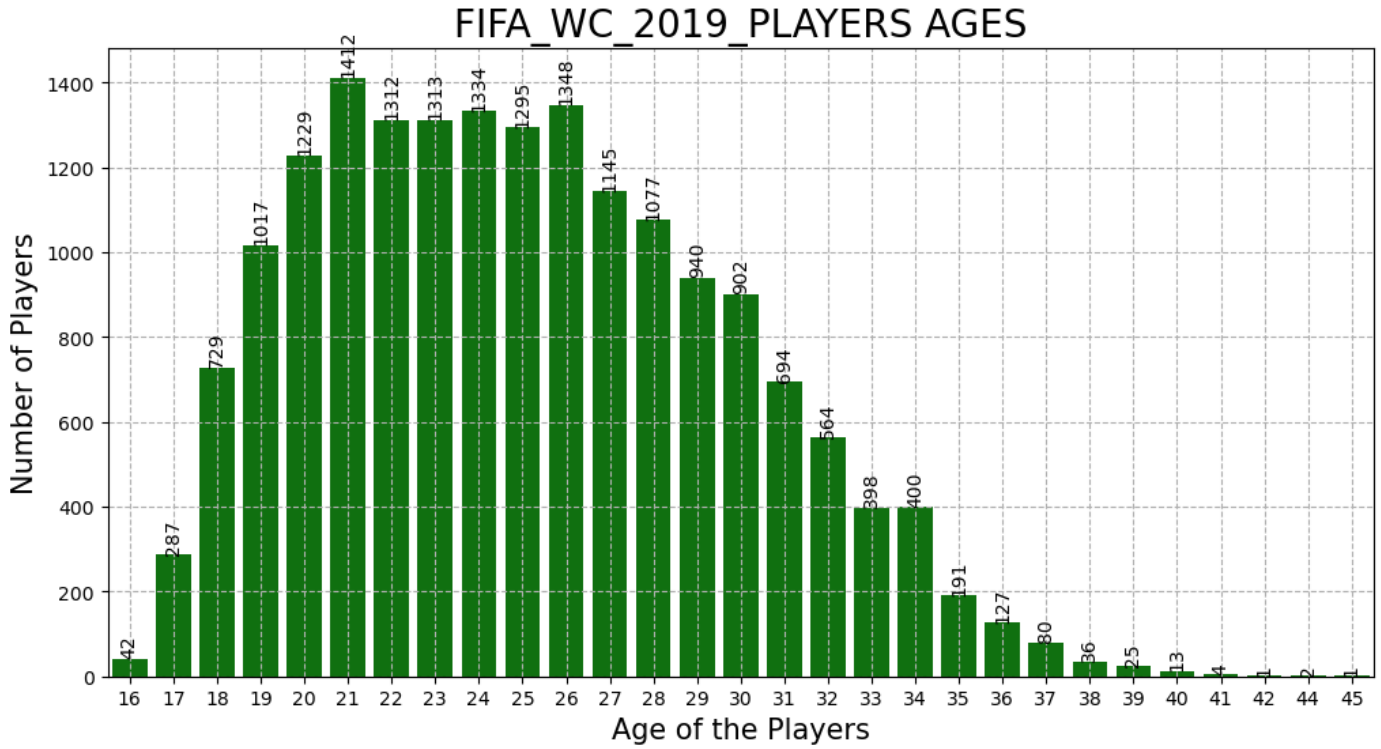
In [22]:
```
fifa_data.columns
```

Out[22]:
```
Index(['Unnamed: 0', 'ID', 'Name', 'Age', 'Photo', 'Nationality', 'Flag',
       'Overall', 'Potential', 'Club', 'Club Logo', 'Value', 'Wage', 'Special',
       'Preferred Foot', 'International Reputation', 'Weak Foot',
       'Skill Moves', 'Work Rate', 'Body Type', 'Real Face', 'Position',
       'Jersey Number', 'Joined', 'Loaned From', 'Contract Valid Until',
       'Height', 'Weight', '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', 'Crossing',
       'Finishing', 'HeadingAccuracy', 'ShortPassing', 'Volleys', 'Dribbling',
       'Curve', 'FKAccuracy', 'LongPassing', 'BallControl', 'Acceleration',
       'SprintSpeed', 'Agility', 'Reactions', 'Balance', 'ShotPower',
       'Jumping', 'Stamina', 'Strength', 'LongShots', 'Aggression',
       'Interceptions', 'Positioning', 'Vision', 'Penalties', 'Composure',
       'Marking', 'StandingTackle', 'SlidingTackle', 'GKDiving', 'GKHandling',
       'GKKicking', 'GKPositioning', 'GKReflexes', 'Release Clause'],
      dtype='object')
```

visualization

1.Age

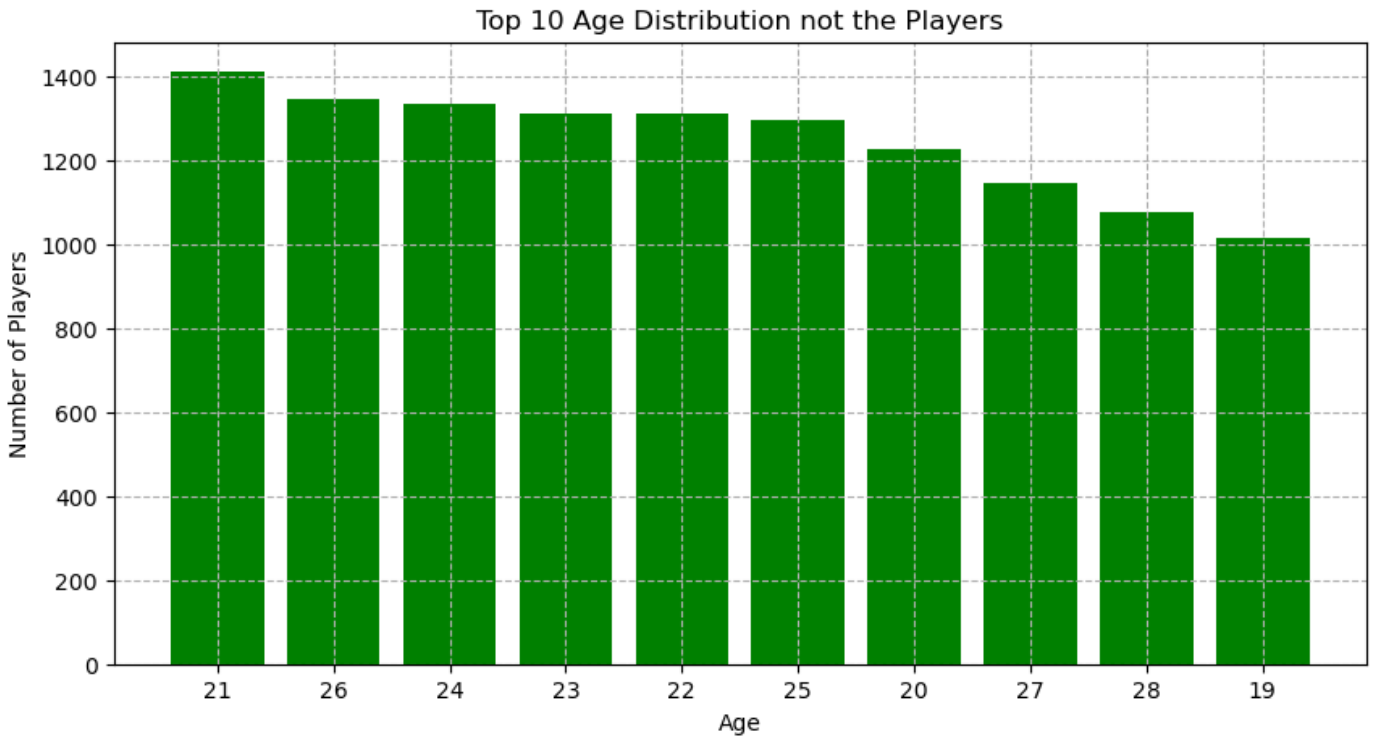
In [23]:
```
plt.figure(figsize=(12,6))
ax = sns.countplot(data = fifa_data, x='Age',color = 'green')
for p in ax.containers:
    ax.bar_label(p, rotation=90)
plt.grid(True, linestyle='--')
plt.xlabel('Age of the Players', fontsize=15)
plt.ylabel('Number of Players', fontsize=15)
plt.title('FIFA_WC_2019_PLAYERS AGES', fontsize=20)
plt.show()
```



In [24]:
```
age = fifa_data['Age'].value_counts()[:10]
age
```

Out[24]:
```
21    1412
26    1348
24    1334
23    1313
22    1312
25    1295
20    1229
27    1145
28    1077
19    1017
Name: Age, dtype: int64
```

In [25]:
```
plt.figure(figsize=(10,5))
plt.bar(age.index.astype(str),age.values, color='green')
plt.xlabel('Age')
plt.ylabel('Number of Players' )
plt.title('Top 10 Age Distribution not the Players')
plt.grid(True, linestyle='--');
```



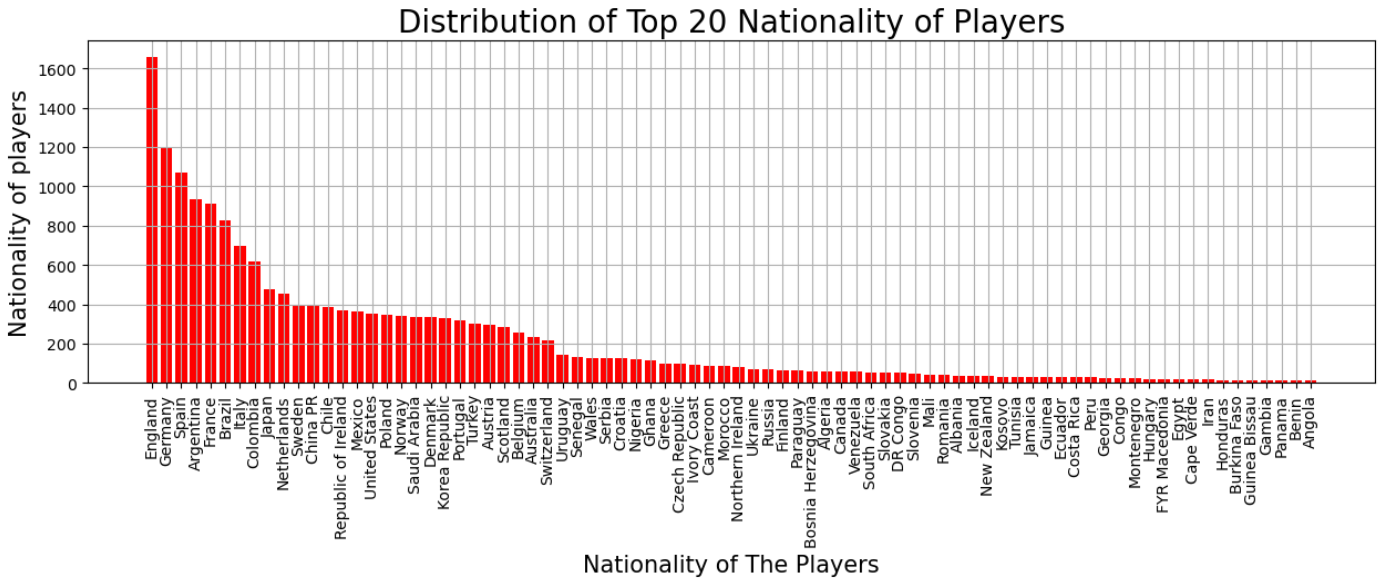
2.Nationality

In [26]:
```
top_Nationality_playes = fifa_data['Nationality'].value_counts().head(80)
top_Nationality_playes
```

Out[26]:
```
England          1657
Germany          1195
Spain            1071
Argentina         935
France            911
                 ...
Guinea Bissau      15
Gambia             15
Panama             15
Benin              15
Angola             15
Name: Nationality, Length: 80, dtype: int64
```

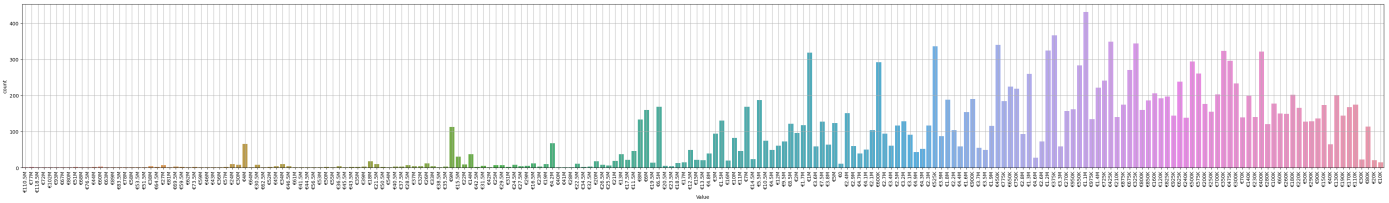
In [27]:
```
plt.figure(figsize=(15,4))
lationality_playes.index.astype(str), top_Nationality_playes.values, color =
```

```
plt.xlabel('Nationality of The Players', fontsize=15)
plt.ylabel('Nationality of players', fontsize=15)
plt.title('Distribution of Top 20 Nationality of Players', fontsize=20)
plt.xticks(rotation=90)
plt.grid(True);
```



3.Value

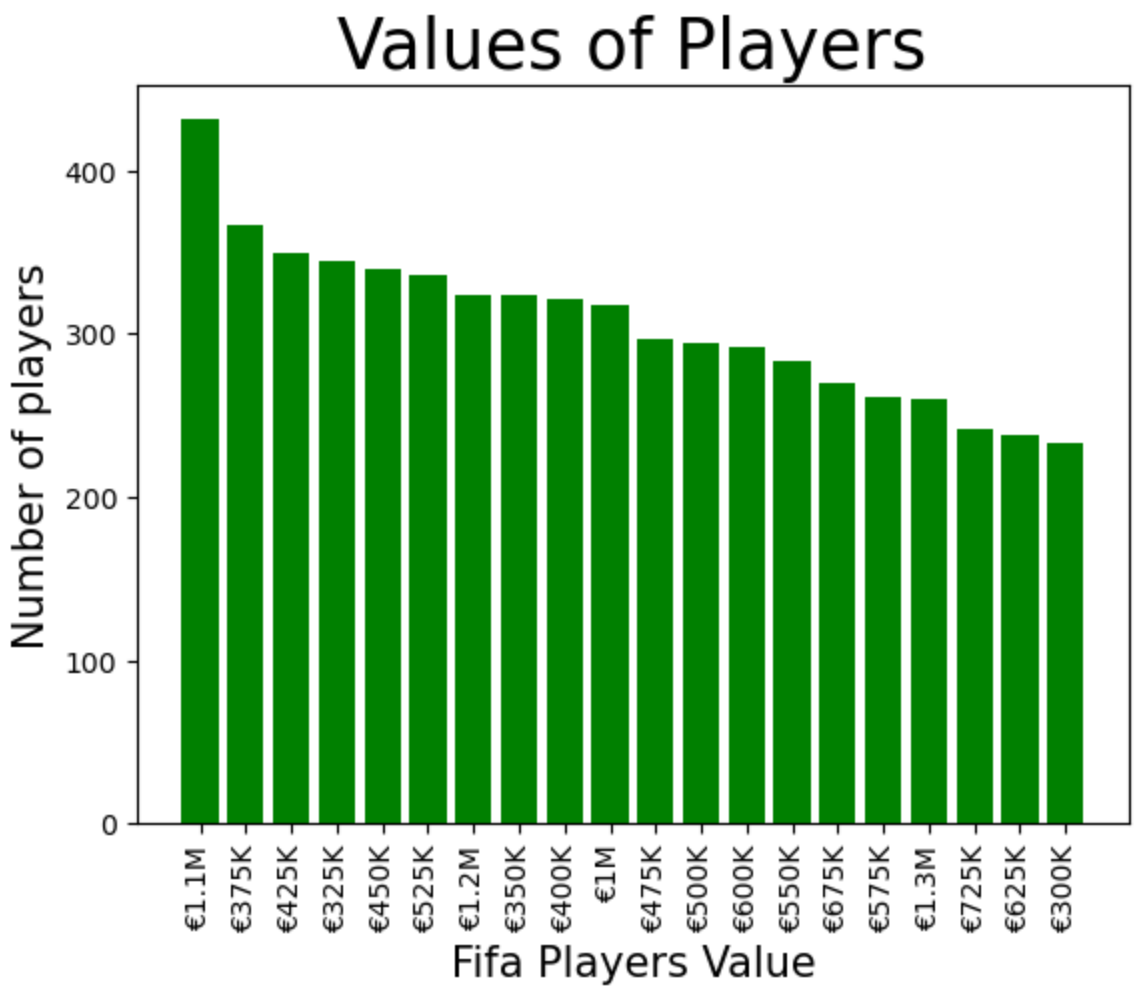
In [28]:
```
plt.figure(figsize=(50,6))
sns.countplot(data=fifa_data, x='Value')
plt.xticks(rotation=90)
plt.grid(True)
```



In [29]:
```
value_of_players = fifa_data['Value'].value_counts()[0:20]
value_of_players
```

Out[29]:
```
€1.1M    431
€375K    366
€425K    349
€325K    344
€450K    340
€525K    336
€1.2M    324
€350K    323
€400K    321
€1M      318
€475K    296
€500K    294
€600K    292
€550K    283
€675K    270
€575K    261
€1.3M    260
€725K    241
€625K    238
€300K    233
          dtype: int64
```

In [30]:
```
plt.bar(value_of_players.index.astype(str), value_of_players.values,color='green')
plt.xlabel('Fifa Players Value', fontsize=15)
plt.ylabel('Number of players', fontsize=15)
plt.title('Values of Players', fontsize=25)
plt.xticks(rotation=90);
```

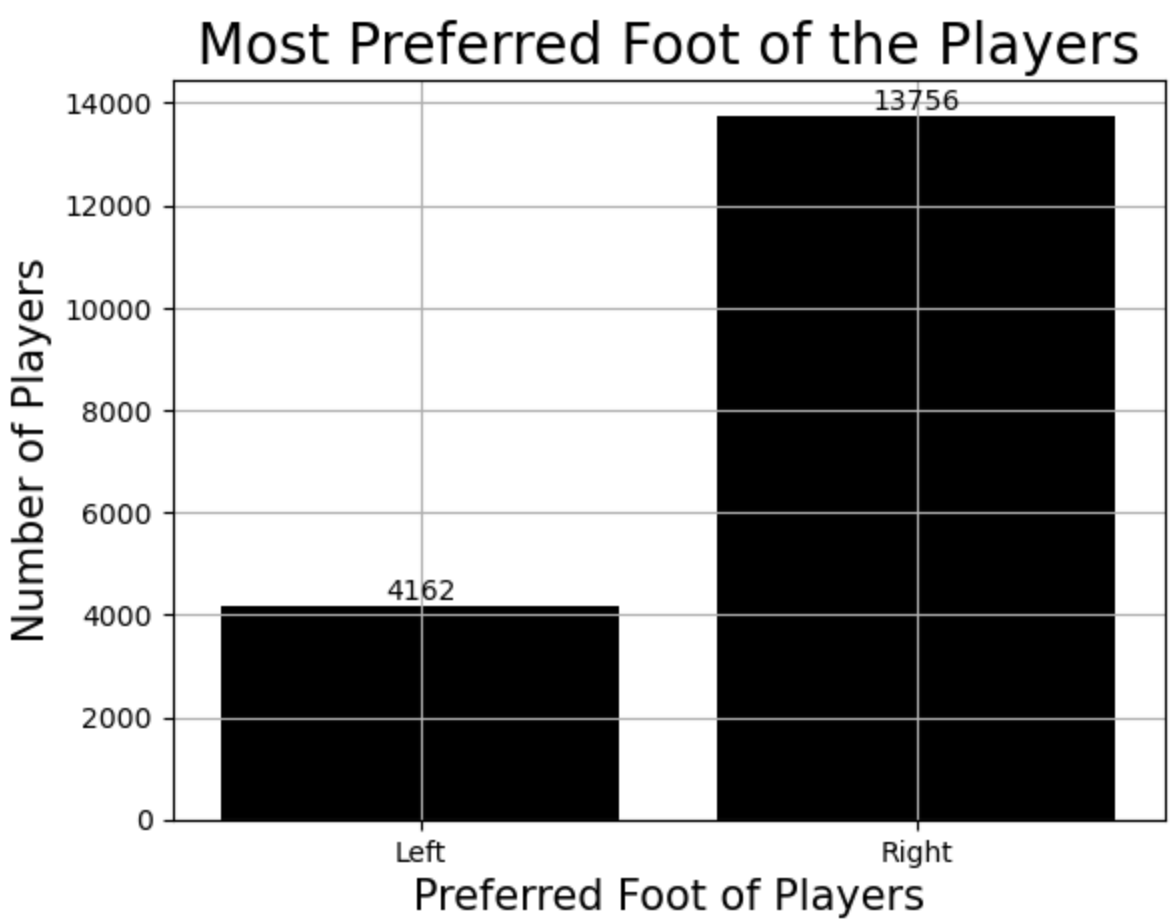


4.Preferred Foot

In [31]:
```
fifa_data['Preferred Foot'].value_counts()
```

Out[31]:
```
Right    13756
Left      4162
Name: Preferred Foot, dtype: int64
```

In [32]:
```
ax = sns.countplot(data=fifa_data,x='Preferred Foot',color='black')
plt.grid(True)
plt.xlabel('Preferred Foot of Players',fontsize=15)
plt.ylabel('Number of Players',fontsize=15)
plt.title('Most Preferred Foot of the Players',fontsize=20)
for p in ax.containers:
    ax.bar_label(p)
```

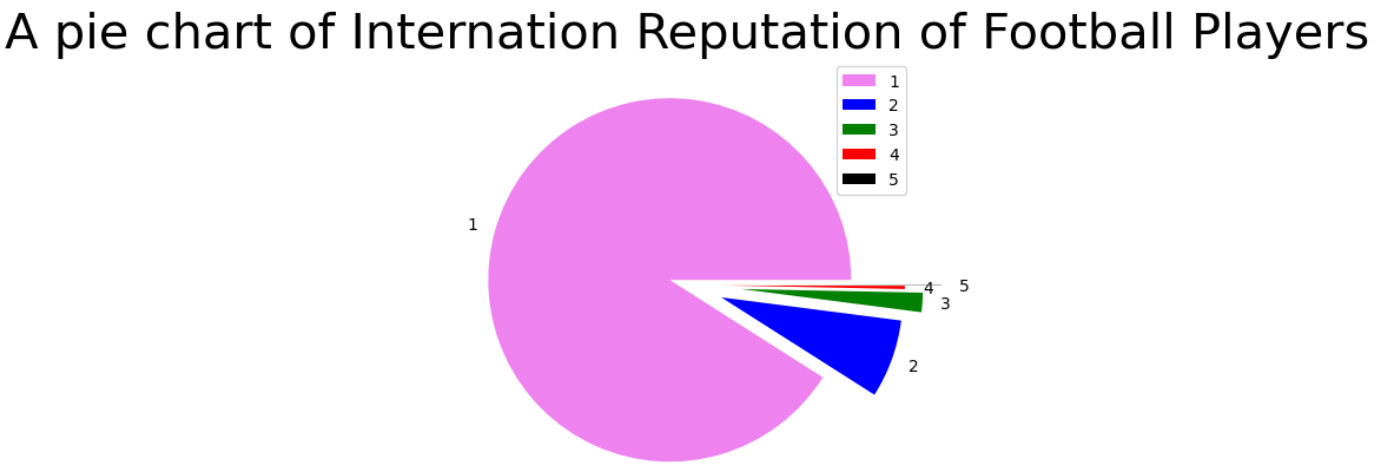


5.International Reputation

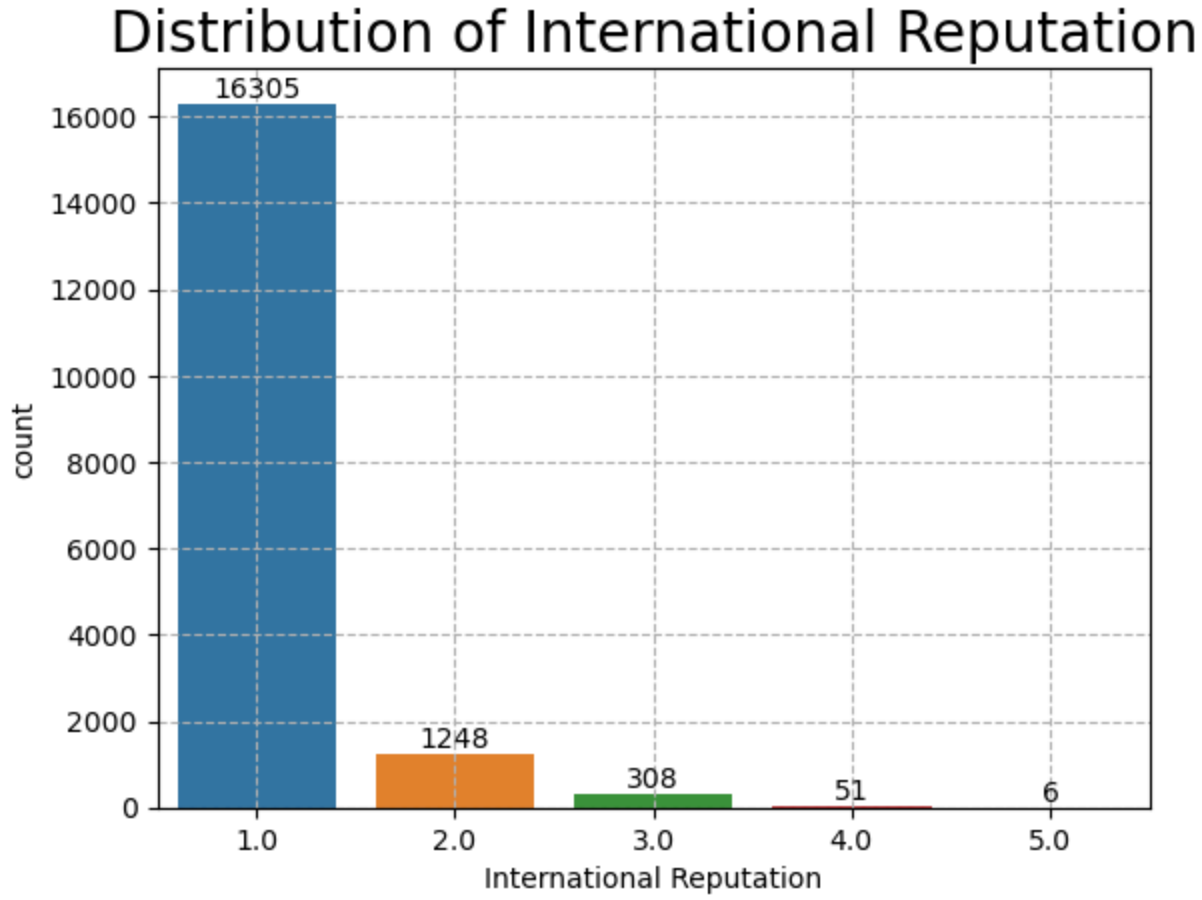
In [33]:
```
fifa_data['International Reputation'].value_counts()
```

Out[33]:
```
1.0    16305
2.0     1248
3.0      308
4.0       51
5.0        6
Name: International Reputation, dtype: int64
```

In [34]:
```
# plot a pie chart of above data
values = ['1','2','3','4','5']
sizes = ['16305','1248','308','51','6']
colors = ['violet','blue','green','red','black']
explode = [.1,.2,.3,.2,.4]
plt.figure(figsize=(5,5))
plt.pie(sizes, labels = values, colors=colors,explode=explode)
plt.title('A pie chart of Internation Reputation of Football Players', fontsize=30)
plt.legend()
plt.show();
```



In [35]:
```
ax = sns.countplot(x='International Reputation', data=fifa_data)
for bars in ax.containers:
    ax.bar_label(bars)
plt.title('Distribution of International Reputation', fontsize=20)
plt.grid(True, linestyle='--')
plt.show()
```



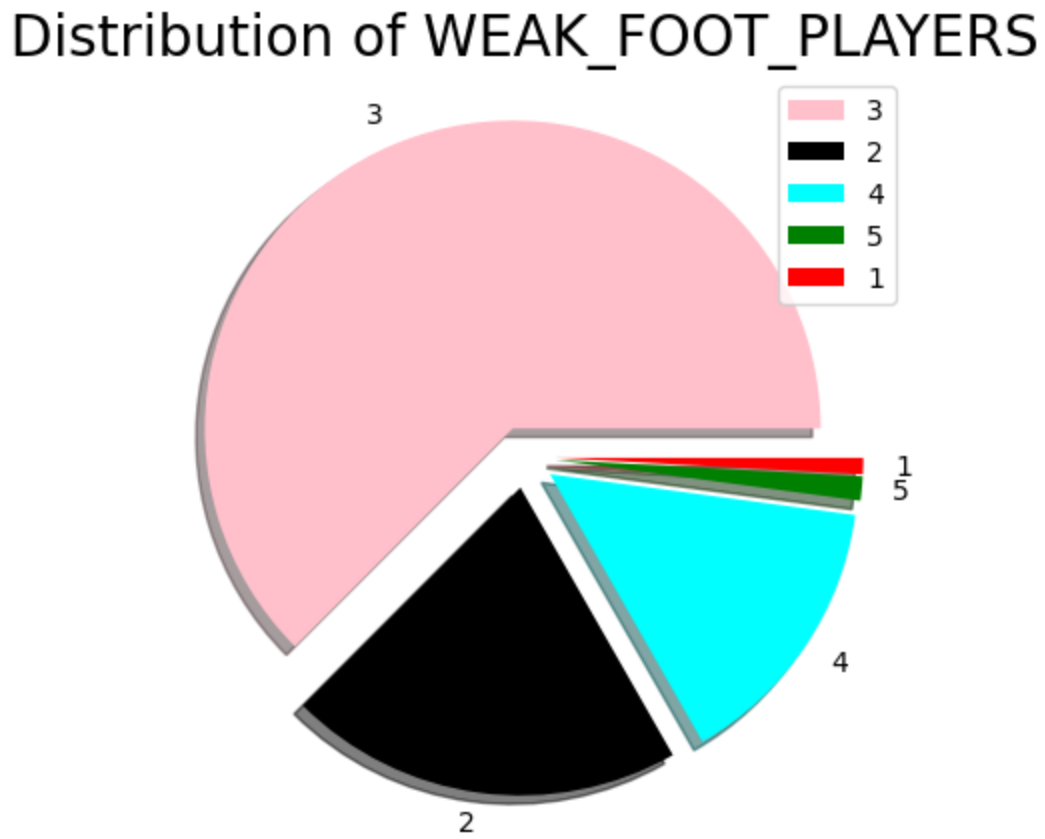
6.Weak Foot

In [36]:
```
fifa_data['Weak Foot'].value_counts()
```

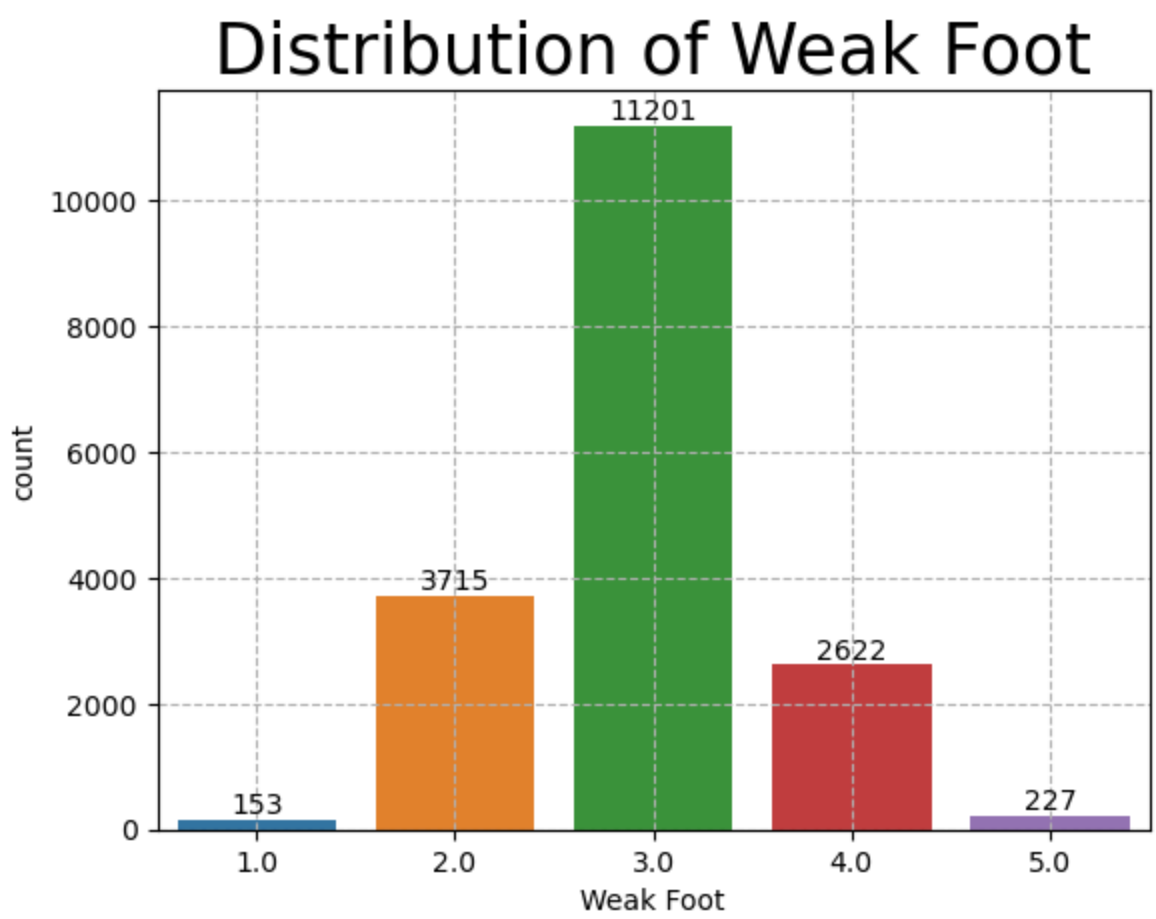
Out[36]:
```
3.0    11201
2.0     3715
4.0     2622
5.0      227
1.0      153
Name: Weak Foot, dtype: int64
```

In [37]:
```
values = ['3','2','4','5','1']
size = ['11201','3715','2622','227','153']
color = ['pink','black','cyan','green','red']
explode = [0.1, 0.1, 0.1, 0.1, 0.1]

plt.figure(figsize=(5,5))
plt.pie(size,labels=values, colors=color, explode=explode,shadow=True)
plt.title('Distribution of WEAK_FOOT_PLAYERS', fontsize=20)
plt.legend()
plt.show();
```



In [38]:
```
ax = sns.countplot(x='Weak Foot', data=fifa_data)
for bars in ax.containers:
    ax.bar_label(bars)
plt.title('Distribution of Weak Foot', fontsize=25)
plt.grid(True, linestyle='--')
```

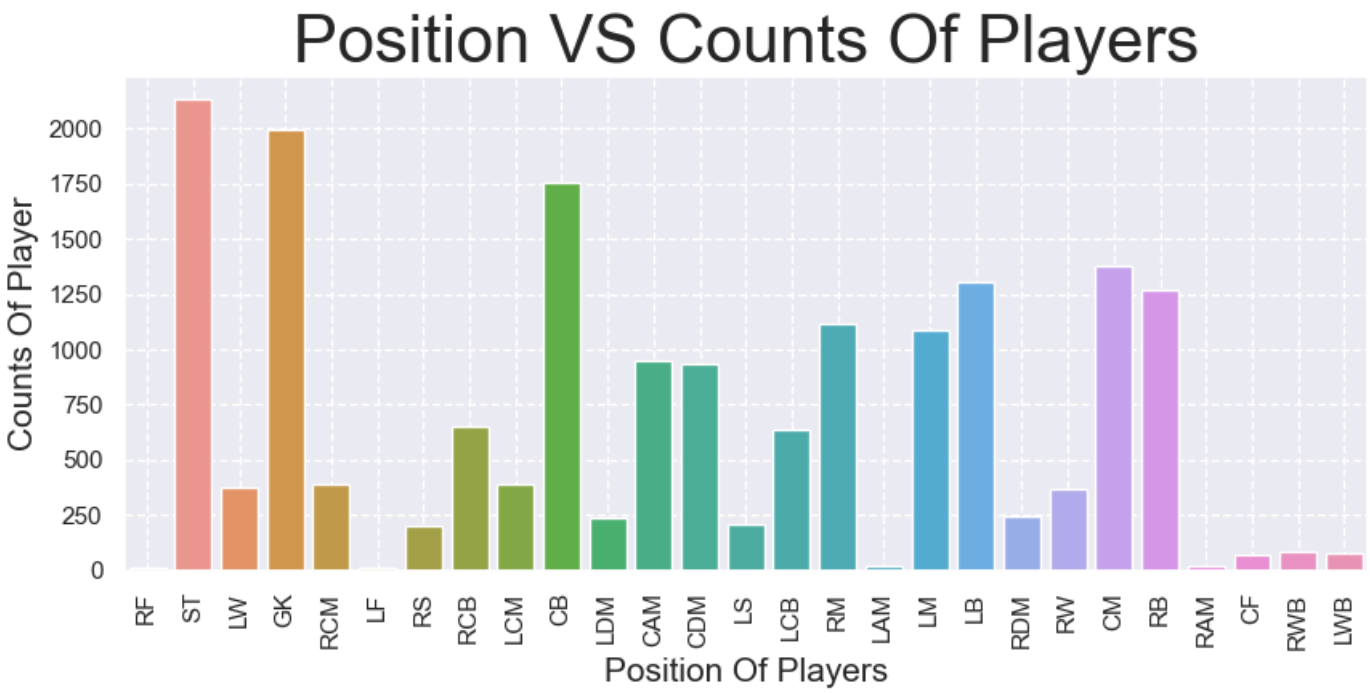


7.Position

In [39]:
```
fifa_data['Position'].value_counts()
```

Out[39]:
```
ST     2130
GK     1992
CB     1754
CM     1377
LB     1305
RB     1268
RM     1114
LM     1086
CAM     948
CDM     936
RCB     652
LCB     637
LCM     389
RCM     387
LW      374
RW      365
RDM     246
LDM     239
LS      206
RS      201
RWB      87
LWB      78
CF       74
LAM      21
RAM      21
RF       16
LF       15
Name: Position, dtype: int64
```

In [40]:
```
plt.figure(figsize=(10,4))
sns.set(style = 'dark', palette = 'colorblind', color_codes = True)
sns.countplot(data = fifa_data, x = 'Position')
plt.xlabel('Position Of Players', fontsize=15)
plt.ylabel('Counts Of Player', fontsize=15)
plt.title('Position VS Counts Of Players', fontsize=30)
plt.grid(True, linestyle='--')
plt.xticks(rotation=90);
```



8.Weight

In [41]:
```
fifa_data['Weight'].head()
```

Out[41]:
```
0    159lbs
1    183lbs
2    150lbs
3    168lbs
4    154lbs
Name: Weight, dtype: object
```

In [42]:
```
# replace 'lbs' from weight column
def extract_value(value):
 out = value.replace('lbs','')
 return float(out)
```

In [43]:
```
fifa_data['Weight'] = fifa_data['Weight'].apply(lambda x : extract_value(x))
```

In [44]:
```
fifa_data['Weight'].head()
```

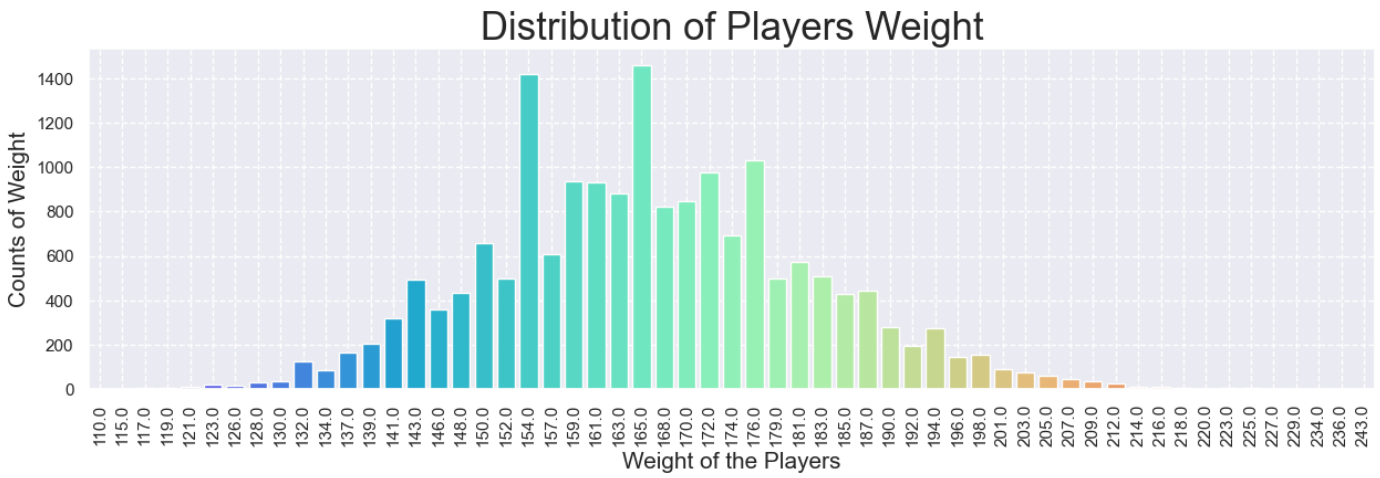
Out[44]:
```
0    159.0
1    183.0
2    150.0
3    168.0
4    154.0
Name: Weight, dtype: float64
```

In [45]:
```
top_weight = fifa_data['Weight'].value_counts()
top_weight
```

Out[45]:
```
165.0    1461
154.0    1418
176.0    1031
172.0     974
159.0     936
161.0     929
163.0     883
170.0     849
168.0     823
174.0     691
150.0     655
157.0     608
181.0     573
183.0     510
179.0     499
152.0     498
143.0     494
187.0     441
148.0     432
185.0     427
146.0     361
141.0     318
190.0     281
194.0     273
139.0     206
192.0     193
137.0     165
198.0     153
196.0     142
132.0     124
201.0      91
134.0      86
203.0      75
205.0      59
207.0      44
209.0      34
130.0      33
128.0      31
212.0      23
123.0      18
126.0      14
214.0      11
121.0      10
216.0       9
117.0       6
218.0       5
119.0       4
223.0       3
225.0       3
227.0       2
236.0       2
110.0       2
243.0       1
220.0       1
229.0       1
115.0       1
234.0       1
Name: Weight, dtype: int64
```

In [46]:
```
plt.figure(figsize=(15,4))
sns.countplot(x='Weight',data=fifa_data, palette='rainbow')
plt.xlabel('Weight of the Players', fontsize=15)
plt.ylabel('Counts of Weight', fontsize=15)
plt.title('Distribution of Players Weight', fontsize=25)
```

```
plt.grid(True, linestyle='--')
plt.xticks(rotation=90);
```



9.Wage

In [47]:
```
fifa_data['Wage'].unique()
```

Out[47]:
```
array(['€565K', '€405K', '€290K', '€260K', '€355K', '€340K', '€420K',
       '€455K', '€380K', '€94K', '€205K', '€125K', '€285K', '€225K',
       '€145K', '€240K', '€315K', '€200K', '€130K', '€300K', '€215K',
       '€100K', '€255K', '€165K', '€265K', '€160K', '€150K', '€245K',
       '€110K', '€77K', '€115K', '€210K', '€195K', '€230K', '€250K',
       '€135K', '€155K', '€180K', '€175K', '€190K', '€185K', '€21K',
       '€82K', '€73K', '€92K', '€88K', '€96K', '€170K', '€66K', '€235K',
       '€28K', '€105K', '€38K', '€81K', '€57K', '€15K', '€63K', '€22K',
       '€84K', '€120K', '€90K', '€72K', '€93K', '€45K', '€74K', '€51K',
       '€42K', '€31K', '€75K', '€25K', '€140K', '€41K', '€78K', '€53K',
       '€95K', '€80K', '€43K', '€60K', '€85K', '€64K', '€67K', '€18K',
       '€70K', '€91K', '€20K', '€49K', '€87K', '€86K', '€26K', '€29K',
       '€55K', '€35K', '€33K', '€56K', '€30K', '€11K', '€59K', '€23K',
       '€46K', '€39K', '€32K', '€36K', '€98K', '€54K', '€68K', '€58K',
       '€27K', '€40K', '€44K', '€19K', '€1K', '€61K', '€50K', '€99K',
       '€17K', '€52K', '€62K', '€12K', '€10K', '€71K', '€14K', '€76K',
       '€48K', '€65K', '€69K', '€24K', '€34K', '€16K', '€37K', '€47K',
       '€89K', '€97K', '€79K', '€13K', '€83K', '€6K', '€3K', '€9K', '€8K',
       '€7K', '€4K', '€2K', '€5K'], dtype=object)
```

In [48]:
```
def replace_value(value):
    out = value.replace('€','')
    if 'M' in out:
        out = float(out.replace('M',''))*1000000
    elif 'K' in value:
        out = float(out.replace('K',''))*1000
    return float(out)
```

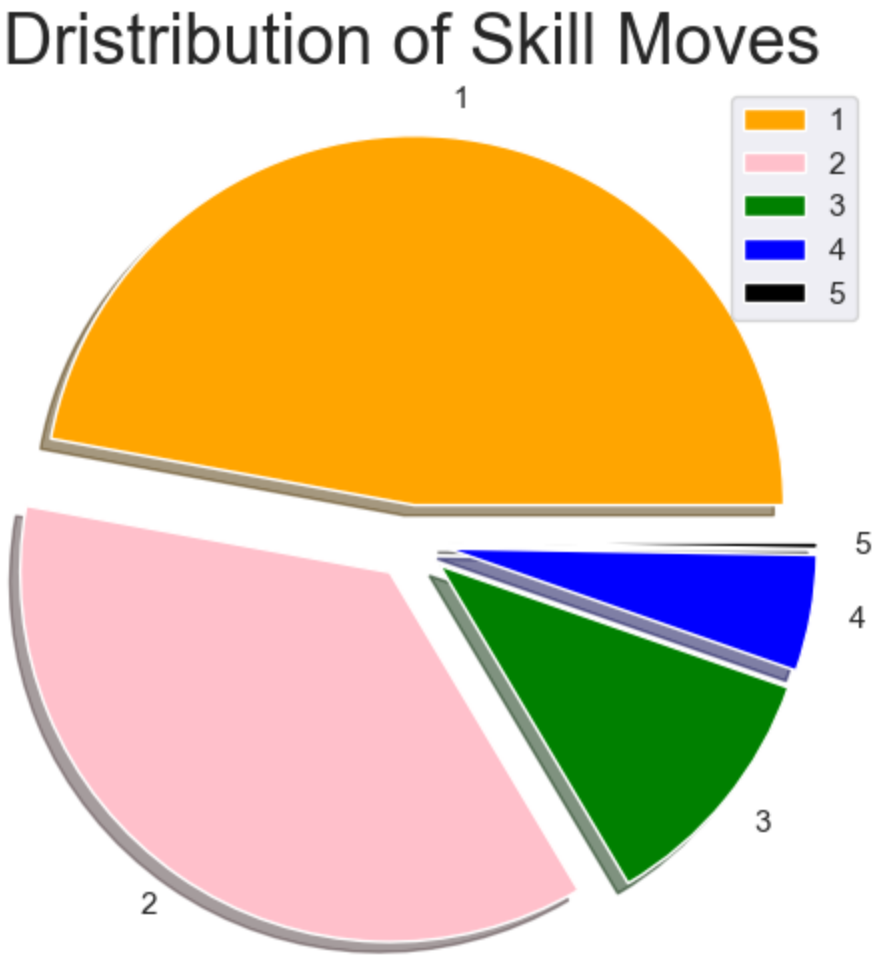
10.Skill Moves

In [49]:
```
fifa_data['Skill Moves'].value_counts()
```

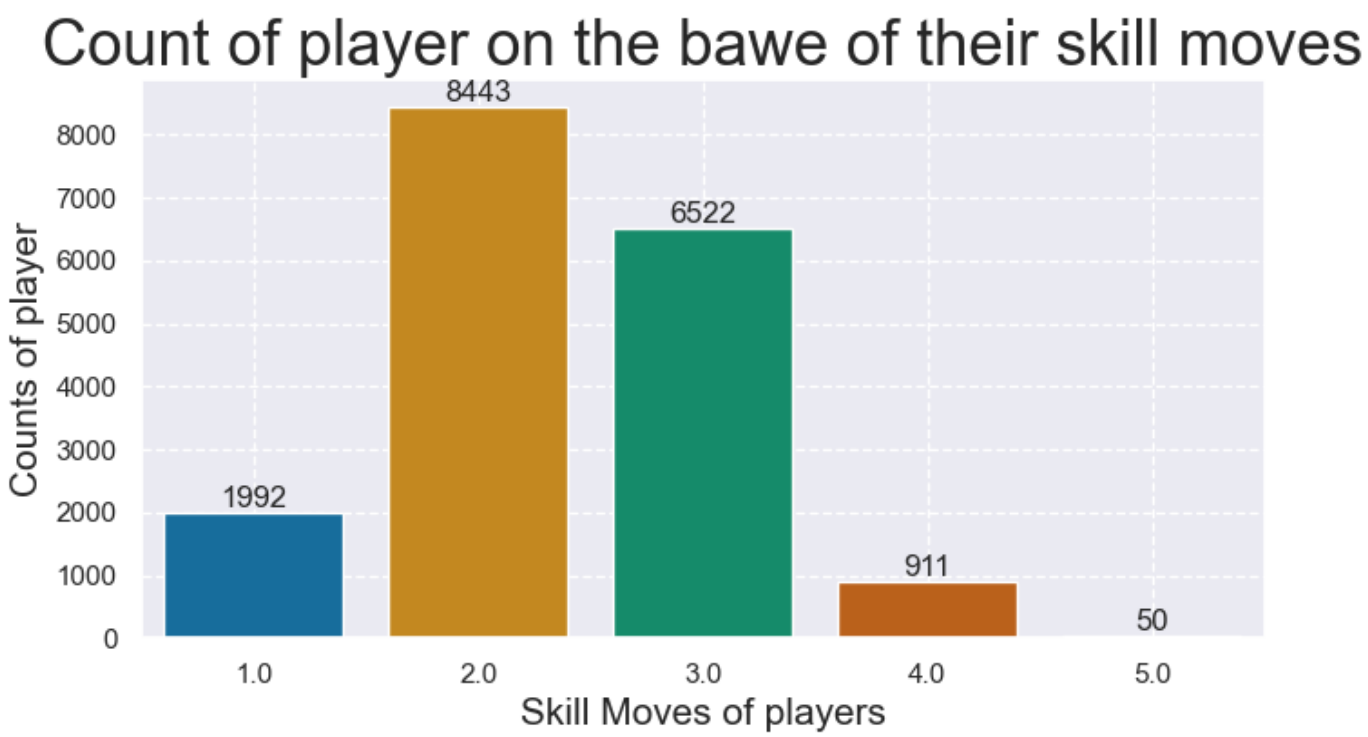
Out[49]:
```
2.0    8443
3.0    6522
1.0    1992
4.0     911
5.0      50
oves, dtype: int64
```

In [50]:
```
values = ['1','2','3','4','5']
size = ['8443','6522','1992','911','50']
color = ['orange','pink','green','blue','black']
explods = [.1,.1,.1,.1,.1]

plt.figure(figsize=(6,6))
plt.pie(size, labels=values, colors=color,explode=explods, shadow=True )
plt.title('Dristribution of Skill Moves', fontsize=25)
plt.legend();
```



In [51]:
```
plt.figure(figsize=(8,4))
ax = sns.countplot(x='Skill Moves',data=fifa_data)
for x in ax.containers:
    ax.bar_label(x)
plt.xlabel('Skill Moves of players', fontsize=15)
plt.ylabel('Counts of player', fontsize=15)
plt.title('Count of player on the bawe of their skill moves', fontsize=25)
plt.grid( linestyle='--')
```



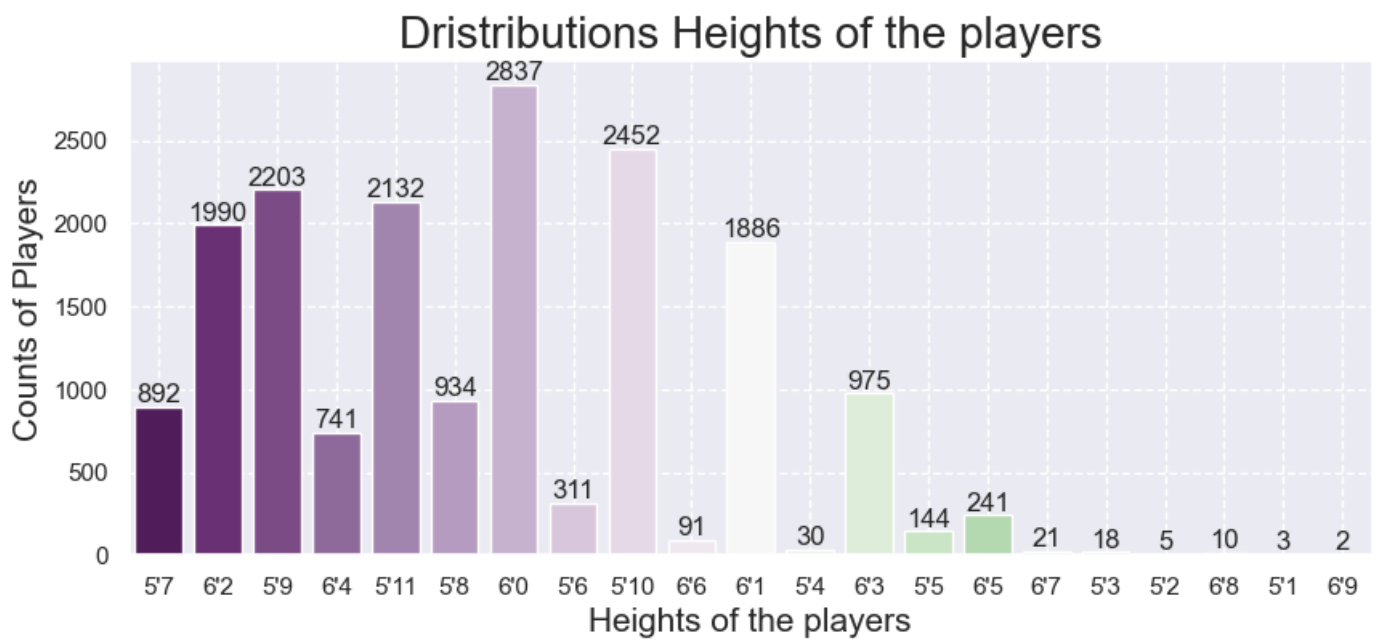
11.Height

In [52]:
```
fifa_data['Height'].value_counts()
```

Out[52]:
```
6'0     2837
5'10    2452
5'9     2203
5'11    2132
6'2     1990
6'1     1886
6'3      975
5'8      934
5'7      892
6'4      741
5'6      311
6'5      241
5'5      144
6'6       91
5'4       30
6'7       21
5'3       18
6'8       10
5'2        5
5'1        3
6'9        2
Name: Height, dtype: int64
```

In [53]:
```
plt.figure(figsize=(10,4))
ax = sns.countplot(x='Height', data=fifa_data, palette='PRGn')
for x in ax.containers:
    ax.bar_label(x)

plt.xlabel('Heights of the players', fontsize=15)
plt.ylabel('Counts of Players', fontsize=15)
plt.title('Dristributions Heights of the players', fontsize=20)
plt.grid(True, linestyle='--')
```



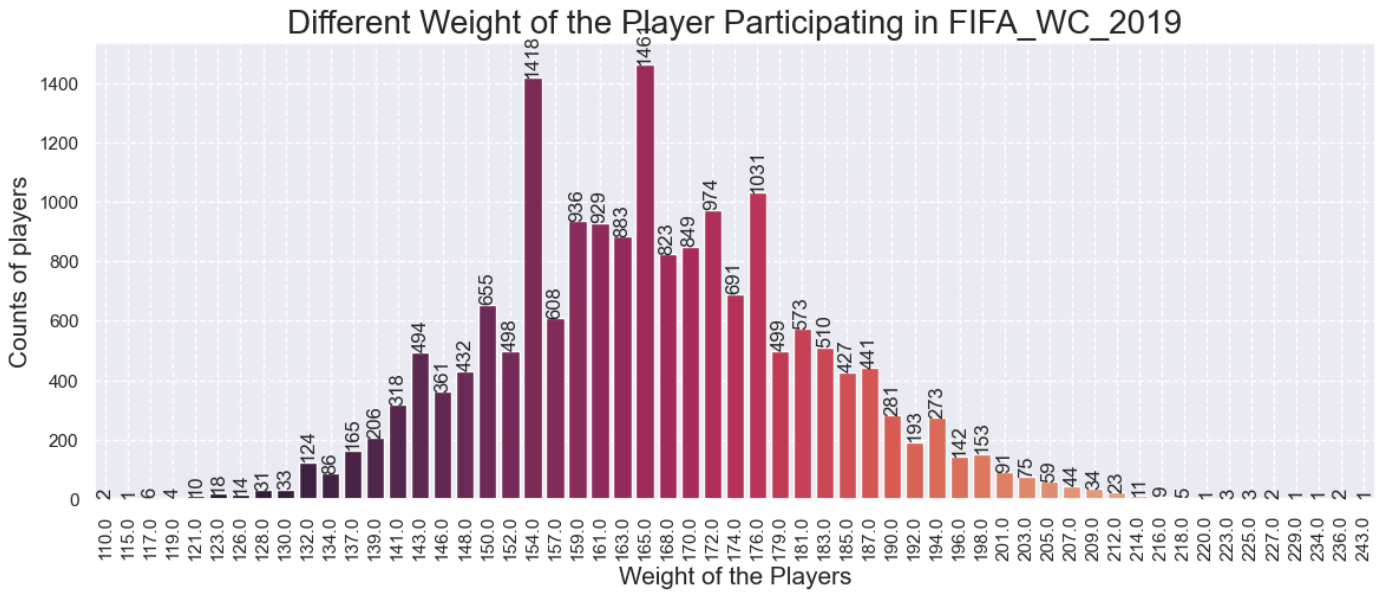
12.Weight

In [54]:
```
fifa_data['Weight'].value_counts()
```

Out[54]:
```
165.0    1461
154.0    1418
176.0    1031
172.0     974
159.0     936
161.0     929
163.0     883
170.0     849
168.0     823
174.0     691
150.0     655
157.0     608
181.0     573
183.0     510
179.0     499
152.0     498
143.0     494
187.0     441
148.0     432
185.0     427
146.0     361
141.0     318
190.0     281
194.0     273
139.0     206
192.0     193
137.0     165
198.0     153
196.0     142
132.0     124
201.0      91
134.0      86
203.0      75
205.0      59
207.0      44
209.0      34
130.0      33
128.0      31
212.0      23
123.0      18
126.0      14
214.0      11
121.0      10
216.0       9
117.0       6
218.0       5
119.0       4
223.0       3
225.0       3
227.0       2
236.0       2
110.0       2
243.0       1
220.0       1
229.0       1
115.0       1
234.0       1
Name: Weight, dtype: int64
```

In [55]:
```
plt.figure(figsize=(14,5))
ax = sns.countplot(x='Weight', data=fifa_data, palette='rocket')
for bar in ax.containers:
    ax.bar_label(bar, rotation=90)
```

```
plt.xlabel('Weight of the Players', fontsize=15)
plt.ylabel('Counts of players', fontsize=15)
plt.title('Different Weight of the Player Participating in FIFA_WC_2019', fontsize=20)
plt.grid(True, linestyle='--')
plt.xticks(rotation=90);
```



13.Work Rate

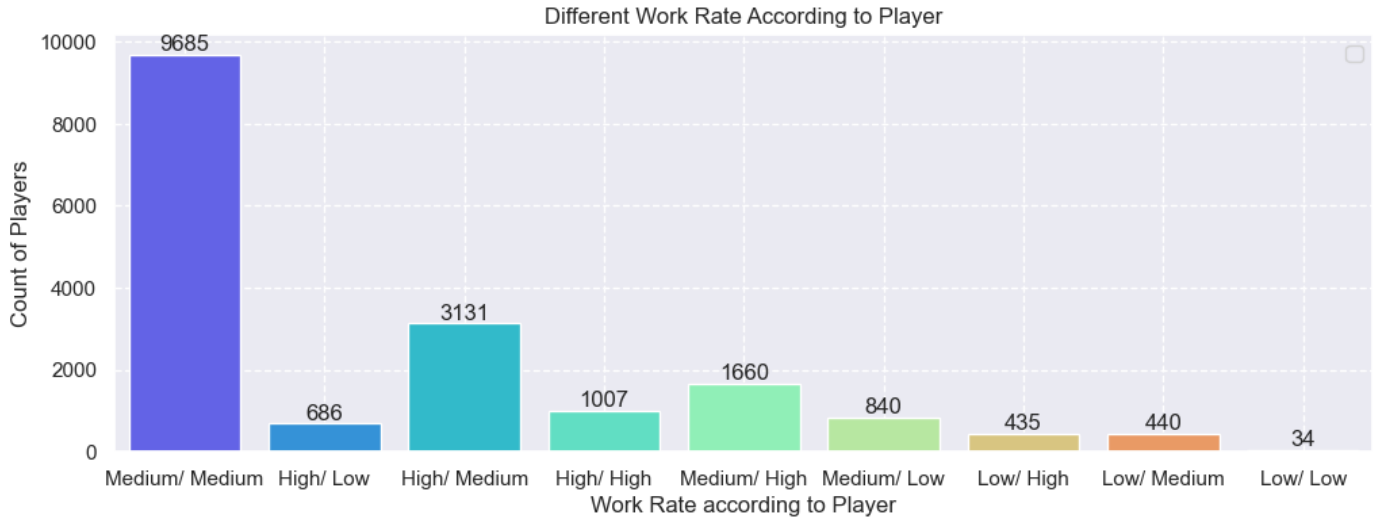
In [56]:
```
fifa_data['Work Rate'].value_counts()
```

Out[56]:
```
Medium/ Medium    9685
High/ Medium      3131
Medium/ High      1660
High/ High        1007
Medium/ Low        840
High/ Low          686
Low/ Medium        440
Low/ High          435
Low/ Low            34
Name: Work Rate, dtype: int64
```

In [57]:
```
plt.figure(figsize=(12,4))
ax = sns.countplot(x='Work Rate', data = fifa_data, palette='rainbow')
for bars in ax.containers:
    ax.bar_label(bars)

plt.xlabel('Work Rate according to Player')
plt.ylabel('Count of Players')
plt.grid(linestyle='--')
plt.title('Different Work Rate According to Player')
plt.legend()
```

Out[57]:
```
No artists with labels found to put in legend.  Note that artists whose label start with an underscore are ignored when legend() is called with no argument.
<matplotlib.legend.Legend at 0x17a1ea92a10>
```

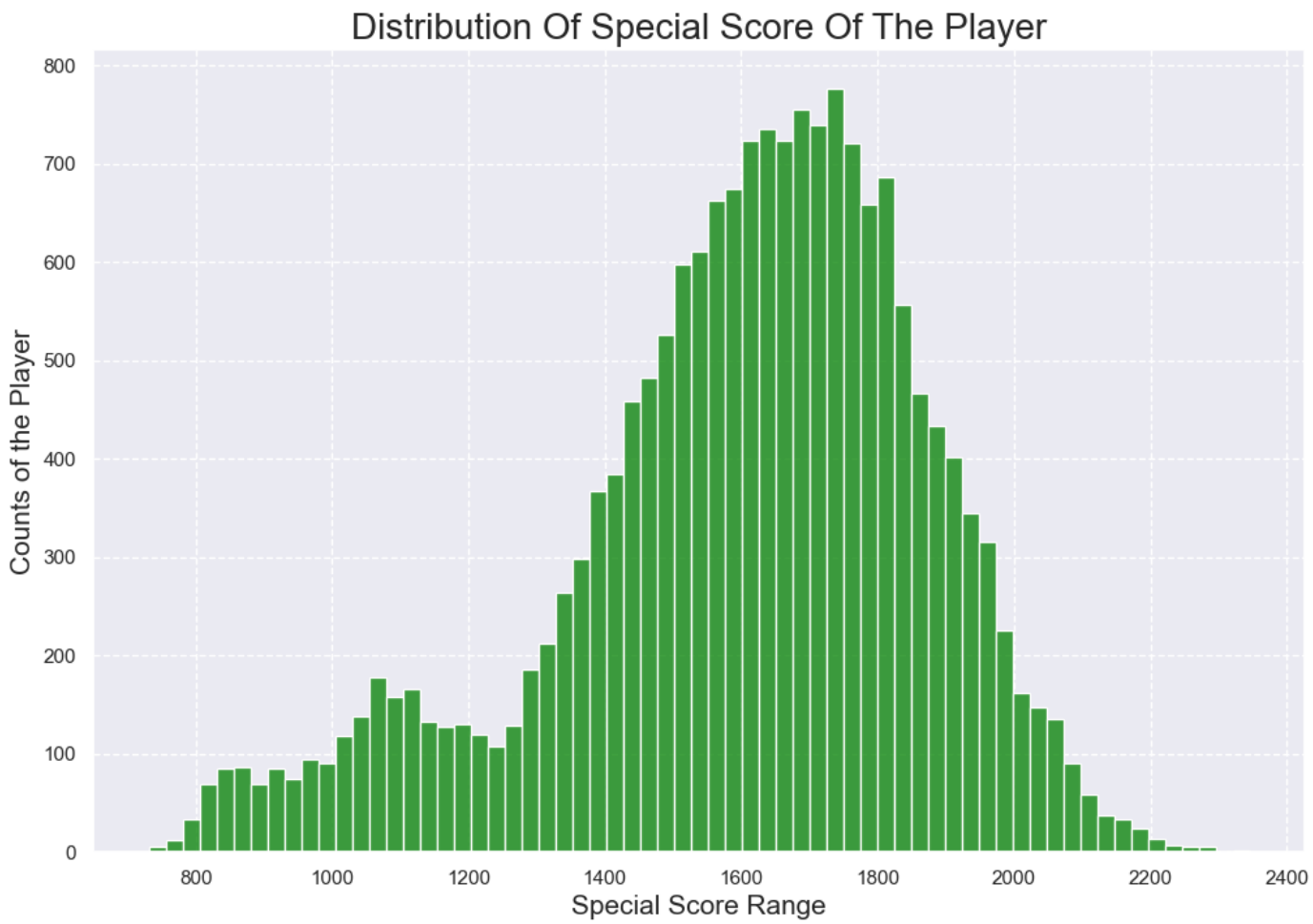


14.Special

In [58]:
```
fifa_data['Special'].value_counts()
```

Out[58]:
```
1745    46
1628    45
1728    44
1655    42
1691    41
        ..
2127     1
2197     1
2104     1
2151     1
731      1
Name: Special, Length: 1421, dtype: int64
```

In [59]:
```
plt.figure(figsize=(12,8))
sns.histplot(x='Special',bins=65,kde=False,data=fifa_data,color='green')
plt.xlabel('Special Score Range', fontsize=15)
plt.ylabel('Counts of the Player', fontsize=15)
plt.grid(linestyle='--')
plt.title('Distribution Of Special Score Of The Player', fontsize=20);
```

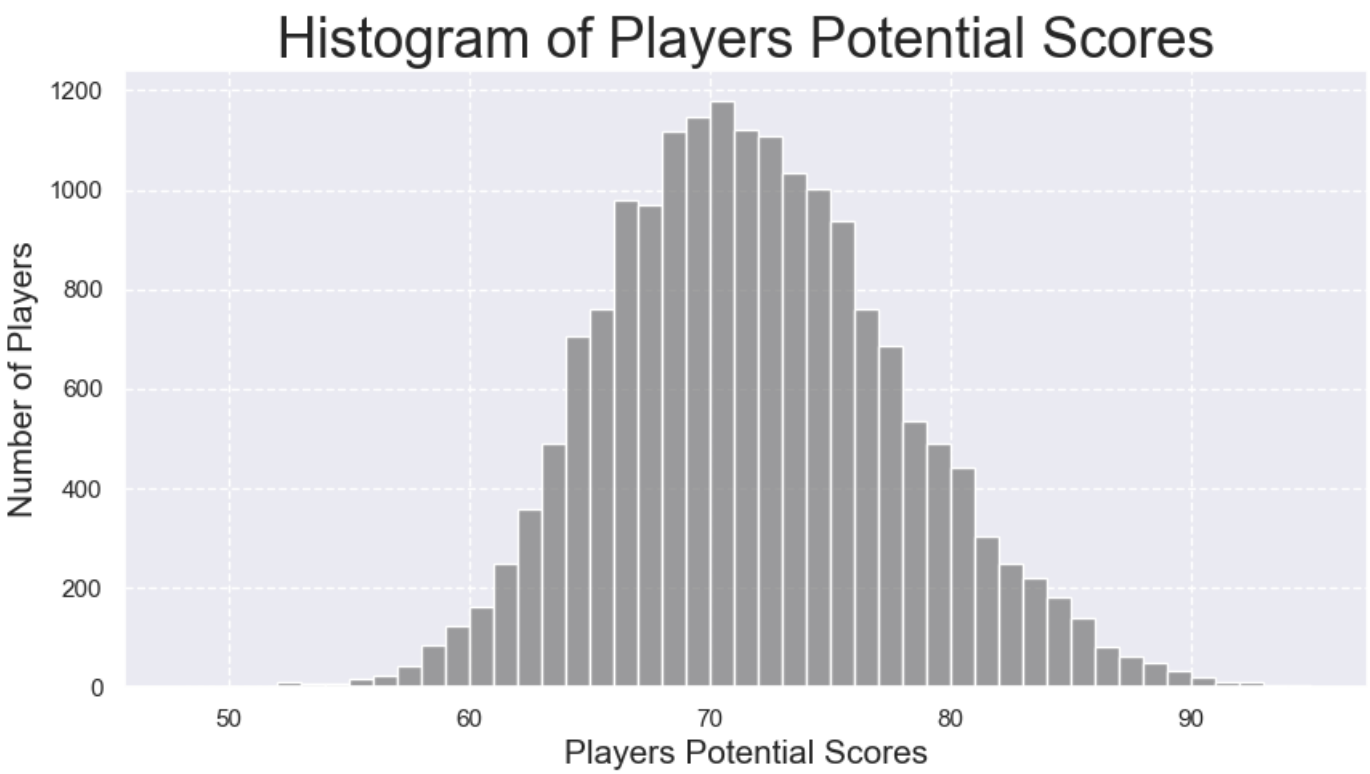


15.Potential

In [60]:
```
fifa_data['Potential'].value_counts()
```

Out[60]:
```
70    1180
69    1147
71    1121
68    1116
72    1108
73    1034
74    1002
66     980
67     970
75     937
76     759
65     759
64     705
77     687
78     536
79     490
63     490
80     442
62     358
81     304
61     249
82     248
83     219
84     180
60     163
85     138
59     124
58      85
86      82
87      61
88      48
57      43
89      33
56      23
90      21
55      18
91      12
52      10
92       9
54       7
53       6
93       4
94       3
51       2
50       2
48       2
95       1
Name: Potential, dtype: int64
```

In [61]:
```
plt.figure(figsize=(10,5))
sns.histplot(x='Potential',bins=47, data=fifa_data, color='gray')
plt.xlabel('Players Potential Scores', fontsize=15)
plt.ylabel('Number of Players', fontsize=15)
plt.title('Histogram of Players Potential Scores', fontsize=25)
plt.grid(linestyle='--')
```



16.Overall

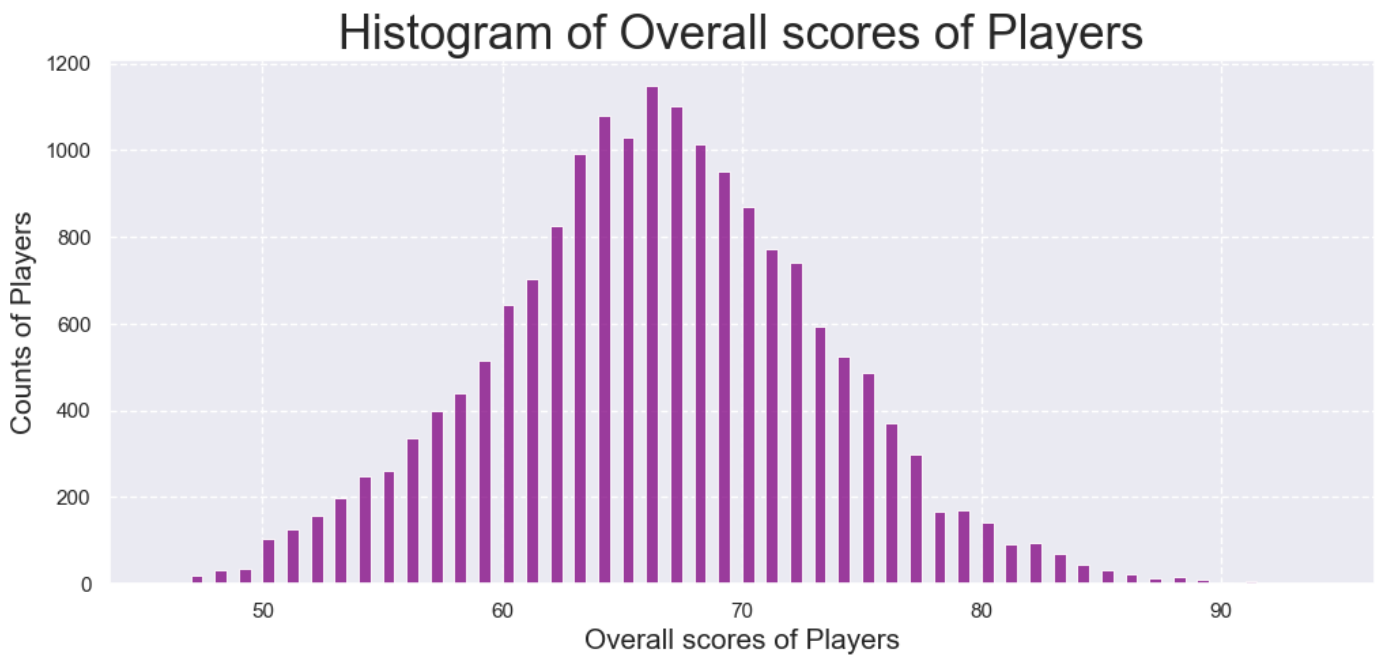
In [62]:
```
fifa_data['Overall'].value_counts()
```

Out[62]:
```
66    1150
67    1101
64    1079
65    1031
68    1015
63     993
69     951
70     870
62     825
71     772
72     742
61     703
60     644
73     594
74     526
59     515
75     488
58     439
57     399
76     371
56     337
77     300
55     262
54     248
53     198
79     169
78     167
52     159
80     141
51     125
50     103
82      94
81      93
83      70
84      45
49      36
85      33
48      32
86      22
47      20
88      17
87      13
89      11
91       6
90       5
94       2
92       1
46       1
Name: Overall, dtype: int64
```

In [63]:
```
len(fifa_data['Overall'].value_counts())
```

Out[63]:
```
48
```

In [64]:
```
plt.figure(figsize=(12,5))
sns.histplot(x='Overall',bins=96,kde=False, data=fifa_data, color='purple')
plt.xlabel('Overall scores of Players', fontsize=15)
plt.ylabel('Counts of Players', fontsize=15)
plt.title('Histogram of Overall scores of Players', fontsize=25)
plt.grid(linestyle='--')
```



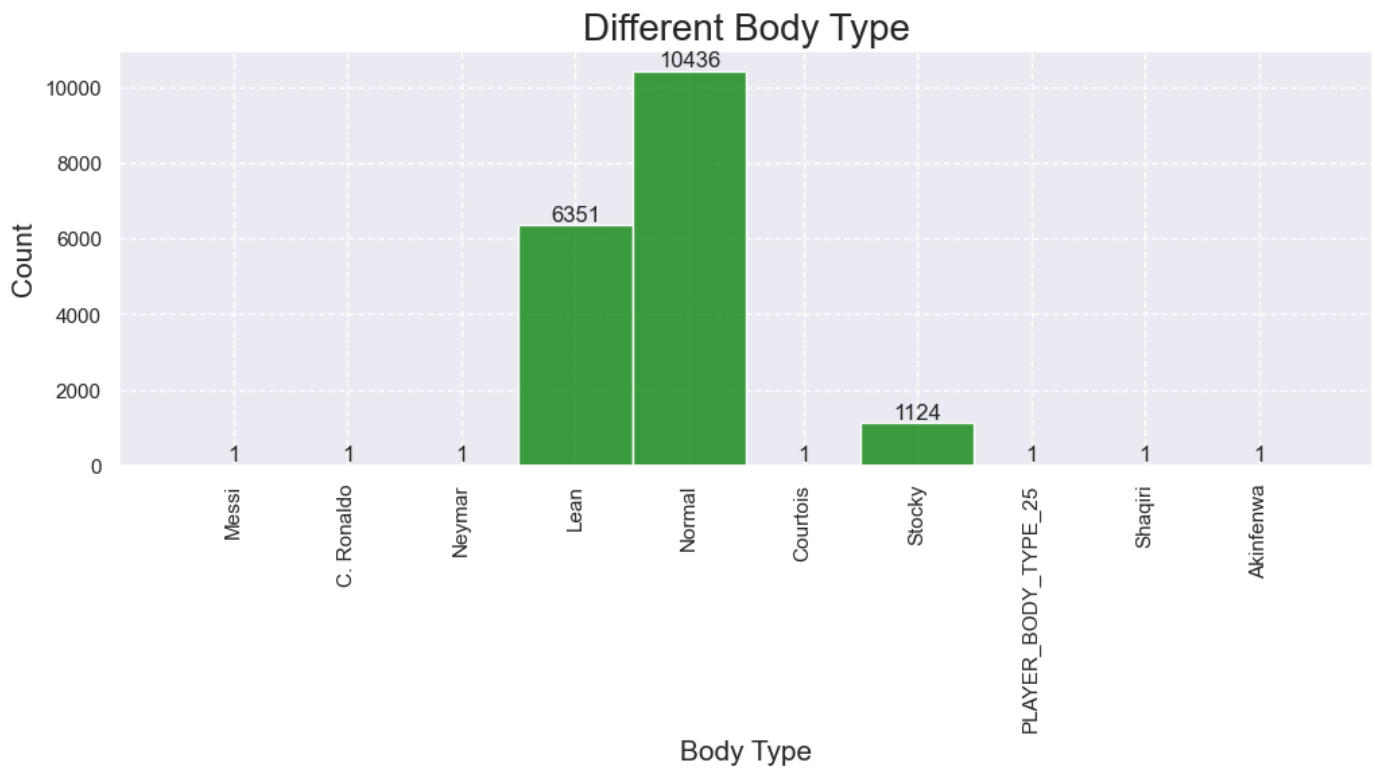
17.Body Type

In [65]:
```
fifa_data['Body Type'].value_counts()
```

Out[65]:
```
Normal                10436
Lean                   6351
Stocky                 1124
Messi                     1
C. Ronaldo                1
Neymar                    1
Courtois                  1
PLAYER_BODY_TYPE_25       1
Shaqiri                   1
Akinfenwa                 1
Name: Body Type, dtype: int64
```

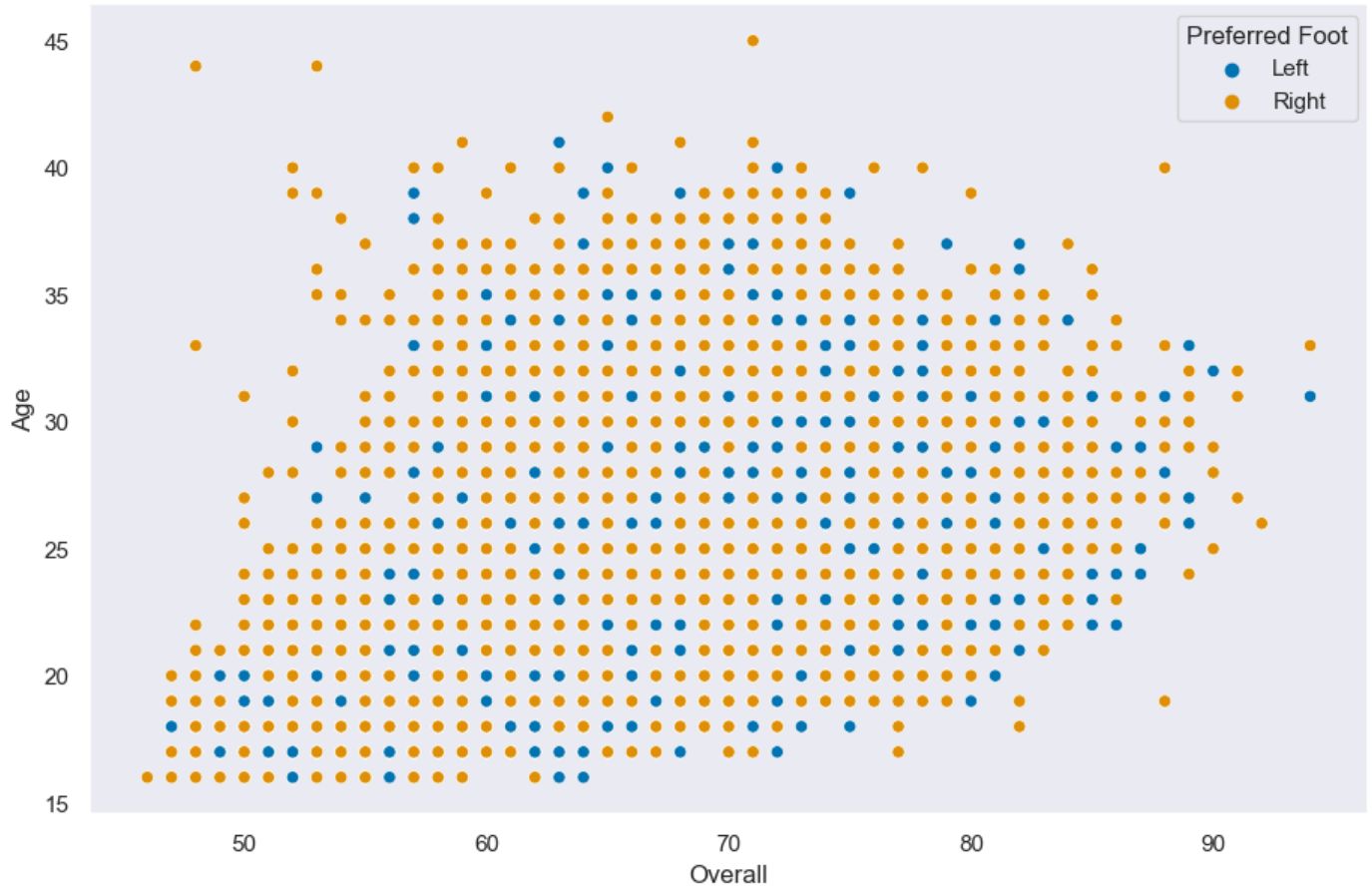
In [66]:
```
plt.figure(figsize=(12,4))
ax = sns.histplot(fifa_data['Body Type'], color='green')
for bar in ax.containers:
    ax.bar_label(bar)

plt.xlabel('Body Type',fontsize=15)
plt.ylabel('Count',fontsize=15)
plt.title('Different Body Type',fontsize=20)
plt.grid(True, linestyle='--')
plt.xticks(rotation=90);
```



In [67]:
```
plt.figure(figsize=(11,7))
sns.scatterplot(y='Age',x='Overall',hue='Preferred Foot',data=fifa_data)
```

Out[67]:
```
<Axes: xlabel='Overall', ylabel='Age'>
```



In [68]:
```
column_selected = ['Name','Age', 'Nationality','Overall', 'Potential', 'Club', 'Value',
                   'Wage', 'Special','Preferred Foot', 'International Reputation', 'Wea
                   'Skill Moves','Work Rate','Body Type','Position','Height', 'Weight',
                   'Finishing', 'HeadingAccuracy', 'ShortPassing','Volleys', 'Dribbling
                   'Curve', 'FKAccuracy', 'LongPassing', 'BallControl', 'Acceleration',
                   'SprintSpeed', 'Agility', 'Reactions', 'Balance', 'ShotPower',
```

```
                    'Jumping', 'Stamina', 'Strength', 'LongShots', 'Aggression',
                    'Interceptions', 'Positioning', 'Vision', 'Penalties', 'Composure',
                    'Marking', 'StandingTackle', 'SlidingTackle', 'GKDiving', 'GKHandlin
                    'GKKicking', 'GKPositioning', 'GKReflexes', 'Release Clause']

selected_data = pd.DataFrame(fifa_data, columns=column_selected)
selected_data.columns
```

Out[68]:
```
Index(['Name', 'Age', 'Nationality', 'Overall', 'Potential', 'Club', 'Value',
       'Wage', 'Special', 'Preferred Foot', 'International Reputation',
       'Weak Foot', 'Skill Moves', 'Work Rate', 'Body Type', 'Position',
       'Height', 'Weight', 'Finishing', 'HeadingAccuracy', 'ShortPassing',
       'Volleys', 'Dribbling', 'Curve', 'FKAccuracy', 'LongPassing',
       'BallControl', 'Acceleration', 'SprintSpeed', 'Agility', 'Reactions',
       'Balance', 'ShotPower', 'Jumping', 'Stamina', 'Strength', 'LongShots',
       'Aggression', 'Interceptions', 'Positioning', 'Vision', 'Penalties',
       'Composure', 'Marking', 'StandingTackle', 'SlidingTackle', 'GKDiving',
       'GKHandling', 'GKKicking', 'GKPositioning', 'GKReflexes',
       'Release Clause'],
      dtype='object')
```

In [69]:
```
selected_data.sample(10)
```

Out[69]:

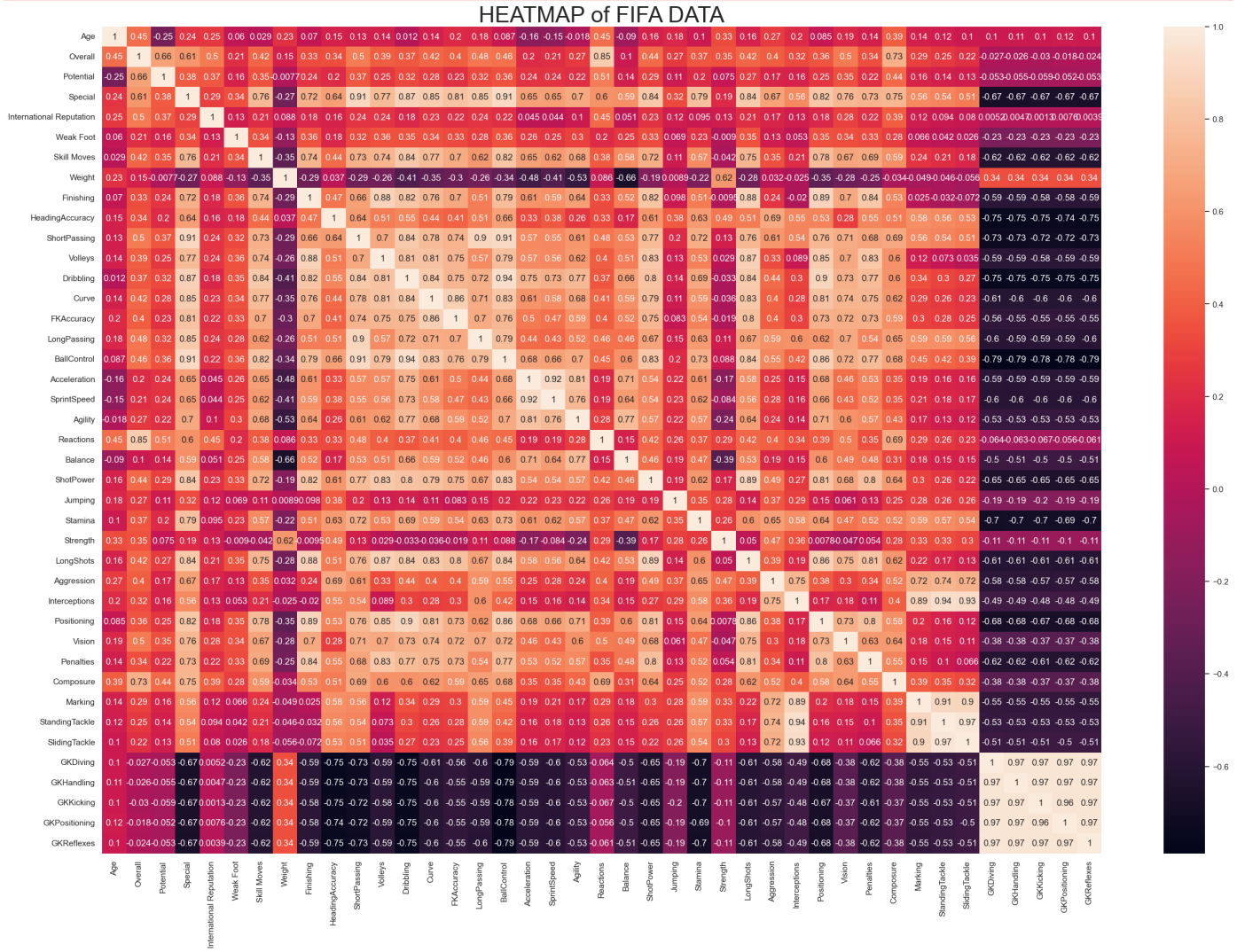
	Name	Age	Nationality	Overall	Potential	Club	Value	Wage	Special	Preferred Foot	...	Composu
9068	D. Balanta	24	Colombia	66	72	Junior FC	€750K	€1K	1474	Right	...	54
7930	R. Di Gennaro	24	Italy	67	70	Inter	€675K	€11K	1068	Right	...	67
16779	C. Rojas	20	Colombia	56	65	Envigado FC	€140K	€1K	1366	Right	...	49
16877	T. Francois	17	Australia	56	74	Fulham	€180K	€3K	1438	Left	...	61
3660	R. Quioto	26	Honduras	72	72	Houston Dynamo	€3.3M	€7K	1818	Right	...	73
1132	L. Acosta	24	Argentina	77	79	DC United	€11M	€8K	1909	Right	...	74
4759	P. Šteinbors	32	Latvia	70	70	Arka Gdynia	€1M	€3K	1163	Right	...	59
16368	M. Ranmark	22	Norway	57	65	Molde FK	€110K	€1K	871	Left	...	38
1839	L. Deaux	29	France	75	75	En Avant de Guingamp	€5M	€21K	1838	Right	...	70
3423	Kim Jin Su	26	Korea Republic	72	73	Jeonbuk Hyundai Motors	€3M	€8K	1829	Left	...	63

10 rows × 52 columns

In [70]:
```
plt.figure(figsize=(30,20))
sns.heatmap(selected_data[['Name','Age', 'Nationality','Overall', 'Potential', 'Club', '
                    'Wage', 'Special','Preferred Foot', 'International Reputation', 'Wea
                    'Skill Moves','Work Rate','Body Type','Position','Height', 'Weight',
                    'Finishing', 'HeadingAccuracy', 'ShortPassing','Volleys', 'Dribbling
                    'Curve', 'FKAccuracy', 'LongPassing', 'BallControl', 'Acceleration',
                    'SprintSpeed', 'Agility', 'Reactions', 'Balance', 'ShotPower',
                    'Jumping', 'Stamina', 'Strength', 'LongShots', 'Aggression',
                    'Interceptions', 'Positioning', 'Vision', 'Penalties', 'Composure',
                    'Marking', 'StandingTackle', 'SlidingTackle', 'GKDiving', 'GKHandlin
```

```
                    'GKKicking', 'GKPositioning', 'GKReflexes', 'Release Clause']].corr(
plt.title('HEATMAP of FIFA DATA', fontsize=30);
```

```
C:\Users\Admin\AppData\Local\Temp\ipykernel_320\933417802.py:11: FutureWarning: The defa
ult value of numeric_only in DataFrame.corr is deprecated. In a future version, it will
default to False. Select only valid columns or specify the value of numeric_only to sile
nce this warning.
  'GKKicking', 'GKPositioning', 'GKReflexes', 'Release Clause']].corr(), annot=True)
```



In [ ]: