```
In [1]:   #importing libraries
          import pandas as pd
          import numpy as np
```

```
In [2]:   url = "C:/Users/Girija/OneDrive/Desktop/fifa21 raw data v2.csv"
```

```
In [3]:   #reading the csv file
          df = pd.read_csv(url)
```

C:\Users\Girija\anaconda3\lib\site-packages\IPython\core\interactiveshell.py:3172: DtypeWarning: Columns (76) hav
e mixed types.Specify dtype option on import or set low_memory=False.
  has_raised = await self.run_ast_nodes(code_ast.body, cell_name,

```
In [4]:   #showingthefirst15rowsofthetable
          df.head(10)
```

Out[4]:

	ID	Name	LongName	photoUrl	playerUrl	Nationality	Age	↓OVA
0	158023	L. Messi	Lionel Messi	https://cdn.sofifa.com/players/158/023/21_60.png	http://sofifa.com/player/158023/lionel-messi/2...	Argentina	33	9:
1	20801	Cristiano Ronaldo	C. Ronaldo dos Santos Aveiro	https://cdn.sofifa.com/players/020/801/21_60.png	http://sofifa.com/player/20801/c-ronaldo-dos-s...	Portugal	35	92
2	200389	J. Oblak	Jan Oblak	https://cdn.sofifa.com/players/200/389/21_60.png	http://sofifa.com/player/200389/jan-oblak/210006/	Slovenia	27	9'
3	192985	K. De Bruyne	Kevin De Bruyne	https://cdn.sofifa.com/players/192/985/21_60.png	http://sofifa.com/player/192985/kevin-de-bruyn...	Belgium	29	9'
4	190871	Neymar Jr	Neymar da Silva Santos Jr.	https://cdn.sofifa.com/players/190/871/21_60.png	http://sofifa.com/player/190871/neymar-da-silv...	Brazil	28	9'
5	188545	R. Lewandowski	Robert Lewandowski	https://cdn.sofifa.com/players/188/545/21_60.png	http://sofifa.com/player/188545/robert-lewando...	Poland	31	9'
6	209331	M. Salah	Mohamed Salah	https://cdn.sofifa.com/players/209/331/21_60.png	http://sofifa.com/player/209331/mohamed-salah/...	Egypt	28	9(
7	212831	Alisson	Alisson Ramses Becker	https://cdn.sofifa.com/players/212/831/21_60.png	http://sofifa.com/player/212831/alisson-ramses...	Brazil	27	9(
8	231747	K. Mbappé	Kylian Mbappé	https://cdn.sofifa.com/players/231/747/21_60.png	http://sofifa.com/player/231747/kylian-mbappe/...	France	21	9(
9	192448	M. ter Stegen	Marc-André ter Stegen	https://cdn.sofifa.com/players/192/448/21_60.png	http://sofifa.com/player/192448/marc-andre-ter...	Germany	28	9(

10 rows × 77 columns

```
In [5]:   #showingthelast10rowsofthetable
          df.tail(10)
```

Out[5]:

	ID	Name	LongName	photoUrl	playerUrl	Nationality	Age	↓OVA
18969	257936	Song Yue	Yue Song	https://cdn.sofifa.com/players/257/936/21_60.png	http://sofifa.com/player/257936/yue-song/210006/	China PR	28	4
18970	257689	Wang Huapeng	Huapeng Wang	https://cdn.sofifa.com/players/257/689/21_60.png	http://sofifa.com/player/257689/huapeng-wang/2...	China PR	20	4
18971	255958	M. Flores	Mateo Flores	https://cdn.sofifa.com/players/255/958/21_60.png	http://sofifa.com/player/255958/mateo-flores/2...	Bolivia	19	4
18972	257710	Zhang Mengxuan	Mengxuan Zhang	https://cdn.sofifa.com/players/257/710/21_60.png	http://sofifa.com/player/257710/mengxuan-zhang...	China PR	21	4
18973	258736	V. Da Silva	Vani Da Silva	https://cdn.sofifa.com/players/258/736/21_60.png	http://sofifa.com/player/258736/vani-da-silva/...	England	17	4
18974	247223	Xia Ao	Ao Xia	https://cdn.sofifa.com/players/247/223/21_60.png	http://sofifa.com/player/247223/ao-xia/210006/	China PR	21	4
18975	258760	B. Hough	Ben Hough	https://cdn.sofifa.com/players/258/760/21_60.png	http://sofifa.com/player/258760/ben-hough/210006/	England	17	4
18976	252757	R. McKinley	Ronan McKinley	https://cdn.sofifa.com/players/252/757/21_60.png	http://sofifa.com/player/252757/ronan-mckinley...	England	18	4
18977	243790	Wang Zhen'ao	Zhen'ao Wang	https://cdn.sofifa.com/players/243/790/21_60.png	http://sofifa.com/player/243790/zhenao-wang/21...	China PR	20	4

| 18978 | 252520 | Zhou Xiao | Xiao Zhou | https://cdn.sofifa.com/players/252/520/21_60.png | | http://sofifa.com/player/252520/xiao-zhou/210006/ | China PR | 21 | 4 |

10 rows × 77 columns

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

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 18979 entries, 0 to 18978
Data columns (total 77 columns):
 #   Column          Non-Null Count  Dtype
---  ------          --------------  -----
 0   ID              18979 non-null  int64
 1   Name            18979 non-null  object
 2   LongName        18979 non-null  object
 3   photoUrl        18979 non-null  object
 4   playerUrl       18979 non-null  object
 5   Nationality     18979 non-null  object
 6   Age             18979 non-null  int64
 7   ↓OVA            18979 non-null  int64
 8   POT             18979 non-null  int64
 9   Club            18979 non-null  object
 10  Contract        18979 non-null  object
 11  Positions       18979 non-null  object
 12  Height          18979 non-null  object
 13  Weight          18979 non-null  object
 14  Preferred Foot  18979 non-null  object
 15  BOV             18979 non-null  int64
 16  Best Position   18979 non-null  object
 17  Joined          18979 non-null  object
 18  Loan Date End   1013 non-null   object
 19  Value           18979 non-null  object
 20  Wage            18979 non-null  object
 21  Release Clause  18979 non-null  object
 22  Attacking       18979 non-null  int64
 23  Crossing        18979 non-null  int64
 24  Finishing       18979 non-null  int64
 25  Heading Accuracy 18979 non-null int64
 26  Short Passing   18979 non-null  int64
 27  Volleys         18979 non-null  int64
 28  Skill           18979 non-null  int64
 29  Dribbling       18979 non-null  int64
 30  Curve           18979 non-null  int64
 31  FK Accuracy     18979 non-null  int64
 32  Long Passing    18979 non-null  int64
 33  Ball Control    18979 non-null  int64
 34  Movement        18979 non-null  int64
 35  Acceleration    18979 non-null  int64
 36  Sprint Speed    18979 non-null  int64
 37  Agility         18979 non-null  int64
 38  Reactions       18979 non-null  int64
 39  Balance         18979 non-null  int64
 40  Power           18979 non-null  int64
 41  Shot Power      18979 non-null  int64
 42  Jumping         18979 non-null  int64
 43  Stamina         18979 non-null  int64
 44  Strength        18979 non-null  int64
 45  Long Shots      18979 non-null  int64
 46  Mentality       18979 non-null  int64
 47  Aggression      18979 non-null  int64
 48  Interceptions   18979 non-null  int64
 49  Positioning     18979 non-null  int64
 50  Vision          18979 non-null  int64
 51  Penalties       18979 non-null  int64
 52  Composure       18979 non-null  int64
 53  Defending       18979 non-null  int64
 54  Marking         18979 non-null  int64
 55  Standing Tackle 18979 non-null  int64
 56  Sliding Tackle  18979 non-null  int64
 57  Goalkeeping     18979 non-null  int64
 58  GK Diving       18979 non-null  int64
 59  GK Handling     18979 non-null  int64
 60  GK Kicking      18979 non-null  int64
 61  GK Positioning  18979 non-null  int64
 62  GK Reflexes     18979 non-null  int64
 63  Total Stats     18979 non-null  int64
 64  Base Stats      18979 non-null  int64
 65  W/F             18979 non-null  object
 66  SM              18979 non-null  object
 67  A/W             18979 non-null  object
 68  D/W             18979 non-null  object
 69  IR              18979 non-null  object
 70  PAC             18979 non-null  int64
 71  SHO             18979 non-null  int64
 72  PAS             18979 non-null  int64
```

```
73  DRI              18979 non-null  int64
74  DEF              18979 non-null  int64
75  PHY              18979 non-null  int64
76  Hits             16384 non-null  object
dtypes: int64(54), object(23)
memory usage: 11.1+ MB
```

In [7]: `df.describe()`

Out[7]:

	ID	Age	↓OVA	POT	BOV	Attacking	Crossing	Finishing	Heading Accuracy	F
count	18979.000000	18979.000000	18979.000000	18979.000000	18979.000000	18979.000000	18979.000000	18979.000000	18979.000000	18979.
mean	226403.384794	25.194109	65.718636	71.136414	66.751726	248.938142	49.688392	45.842405	51.942726	58.
std	27141.054157	4.710520	6.968999	6.114635	6.747193	74.299428	18.131153	19.567081	17.294409	14.
min	41.000000	16.000000	47.000000	47.000000	48.000000	42.000000	6.000000	3.000000	5.000000	7.
25%	210135.000000	21.000000	61.000000	67.000000	62.000000	222.000000	38.000000	30.000000	44.000000	54.
50%	232418.000000	25.000000	66.000000	71.000000	67.000000	263.000000	54.000000	49.000000	55.000000	62.
75%	246922.500000	29.000000	70.000000	75.000000	71.000000	297.000000	63.000000	62.000000	64.000000	68.
max	259216.000000	53.000000	93.000000	95.000000	93.000000	437.000000	94.000000	95.000000	93.000000	94.

8 rows × 54 columns

In [8]: `df.rank()`

Out[8]:

	ID	Age	↓OVA	POT	BOV	Attacking	Crossing	Finishing	Heading Accuracy	Short Passing	...	GK Positioning	GK Reflexes	Total Stats	Ba St
0	311.0	17815.0	18979.0	18975.0	18979.0	18977.5	18942.0	18978.5	16813.0	18974.5	...	14360.5	4434.0	18961.0	1892(
1	13.0	18518.5	18978.0	18969.0	18978.0	18979.0	18927.5	18978.5	18968.0	18769.0	...	14360.5	9494.5	18957.0	1891:
2	3068.0	12331.0	18975.5	18975.0	18975.0	1747.5	993.0	1053.0	1466.5	2591.0	...	18977.0	18977.5	3976.0	1897(
3	2253.0	14731.0	18975.5	18961.5	18975.0	18964.0	18979.0	18861.5	9396.0	18979.0	...	7815.5	12808.5	18978.0	1897:
4	2021.0	13573.0	18975.5	18961.5	18975.0	18966.0	18942.0	18954.5	13254.0	18952.0	...	15787.5	9494.5	18913.0	1881:
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
18974	14357.0	4196.0	7.5	61.5	20.5	2093.0	2200.5	3738.0	4384.0	1052.0	...	14360.5	9494.5	1729.0	5
18975	18710.0	180.0	7.5	4879.5	131.0	3832.5	4716.5	7629.0	3550.5	5616.0	...	1198.0	4434.0	2593.0	57(
18976	15782.0	673.0	7.5	2958.5	20.5	3176.5	3073.0	5633.5	4384.0	4743.0	...	1198.0	9494.5	2856.5	76(
18977	13154.0	2801.5	7.5	148.0	5.0	4143.5	6488.0	10287.5	2514.5	2495.5	...	4433.5	11156.5	2062.5	26(
18978	15702.0	4196.0	7.5	148.0	59.0	2212.0	5222.5	2103.5	3550.5	2825.5	...	7815.5	11156.5	1884.5	15(

18979 rows × 54 columns

In [9]: `df.shape`

Out[9]: `(18979, 77)`

In [10]:
```
missing_value_count = df.isnull().sum()
missing_value_count
```

Out[10]:
```
ID               0
Name             0
LongName         0
photoUrl         0
playerUrl        0
                ...
PAS              0
DRI              0
DEF              0
PHY              0
Hits          2595
Length: 77, dtype: int64
```

```
In [11]: df.dropna()
```

Out[11]:

	ID	Name	LongName	photoUrl	playerUrl	Nationality	Age	↓OVA
205	173731	G. Bale	Gareth Bale	https://cdn.sofifa.com/players/173/731/21_60.png	http://sofifa.com/player/173731/gareth-bale/21...	Wales	30	8:
248	193105	A. Areola	Alphonse Areola	https://cdn.sofifa.com/players/193/105/21_60.png	http://sofifa.com/player/193105/alphonse-areol...	France	27	8:
254	200888	Danilo Pereira	Danilo Luís Hélio Pereira	https://cdn.sofifa.com/players/200/888/21_60.png	http://sofifa.com/player/200888/danilo-luis-he...	Portugal	28	8:
302	216409	M. Politano	Matteo Politano	https://cdn.sofifa.com/players/216/409/21_60.png	http://sofifa.com/player/216409/matteo-politan...	Italy	26	8
306	223959	L. Torreira	Lucas Torreira	https://cdn.sofifa.com/players/223/959/21_60.png	http://sofifa.com/player/223959/lucas-torreira...	Uruguay	24	8
...	...	...	...	...	...	...	...	..
16315	234322	Park Byung Hyun	Byung Hyun Park	https://cdn.sofifa.com/players/234/322/21_60.png	http://sofifa.com/player/234322/byung-hyun-par...	Korea Republic	27	5:
16316	242258	H. Biggins	Harrison Biggins	https://cdn.sofifa.com/players/242/258/21_60.png	http://sofifa.com/player/242258/harrison-biggi...	England	24	5:
16329	243818	M. Sangare	Mohammed Sangare	https://cdn.sofifa.com/players/243/818/21_60.png	http://sofifa.com/player/243818/mohammed-sanga...	Liberia	21	5:
16339	213344	D. Barnes	Dillon Barnes	https://cdn.sofifa.com/players/213/344/21_60.png	http://sofifa.com/player/213344/dillon-barnes/...	England	24	5:
16357	242541	C. Brown	Ciaron Brown	https://cdn.sofifa.com/players/242/541/21_60.png	http://sofifa.com/player/242541/ciaron-brown/2...	Northern Ireland	22	5:

933 rows × 77 columns

```
In [12]: df['Hits']= df['Hits'].dropna(inplace = True)
```

```
In [13]: df['Hits']
```

Out[13]:
```
0        None
1        None
2        None
3        None
4        None
         ...
18974    None
18975    None
18976    None
18977    None
18978    None
Name: Hits, Length: 18979, dtype: object
```

```
In [14]: df['Height']= df['Height'].str.replace('cm','')
```

```
In [15]: df['Height']
```

Out[15]:
```
0        170
1        187
2        188
3        181
4        175
        ...
18974    178
18975    175
18976    179
18977    175
18978    188
Name: Height, Length: 18979, dtype: object
```

```
In [16]: df['Club']=df['Club'].str.replace('\n\n\n\n','')
```

```
In [17]: df['Club']
```

Out[17]:
```
0           FC Barcelona
```

```
1              Juventus
2        Atlético Madrid
3         Manchester City
4    Paris Saint-Germain
                ...
18974          Wuhan Zall
18975      Oldham Athletic
18976          Derry City
18977      Dalian YiFang FC
18978      Dalian YiFang FC
Name: Club, Length: 18979, dtype: object
```

In [18]:
```python
df['Weight'] = df['Weight'].str.replace('kg','')
```

In [19]:
```python
df['Weight']
```

Out[19]:
```
0        72
1        83
2        87
3        70
4        68
         ..
18974    66
18975    65
18976    74
18977    69
18978    75
Name: Weight, Length: 18979, dtype: object
```

In [20]:
```python
df
```

Out[20]:

	ID	Name	LongName	photoUrl	playerUrl	Nationality	Age	↓OVA	
0	158023	L. Messi	Lionel Messi	https://cdn.sofifa.com/players/158/023/21_60.png	http://sofifa.com/player/158023/lionel-messi/2...	Argentina	33	93	
1	20801	Cristiano Ronaldo	C. Ronaldo dos Santos Aveiro	https://cdn.sofifa.com/players/020/801/21_60.png	http://sofifa.com/player/20801/c-ronaldo-dos-s...	Portugal	35	92	
2	200389	J. Oblak	Jan Oblak	https://cdn.sofifa.com/players/200/389/21_60.png	http://sofifa.com/player/200389/jan-oblak/210006/	Slovenia	27	91	
3	192985	K. De Bruyne	Kevin De Bruyne	https://cdn.sofifa.com/players/192/985/21_60.png	http://sofifa.com/player/192985/kevin-de-bruyn...	Belgium	29	91	
4	190871	Neymar Jr	Neymar da Silva Santos Jr.	https://cdn.sofifa.com/players/190/871/21_60.png	http://sofifa.com/player/190871/neymar-da-silv...	Brazil	28	91	
...	...	...	...	...	...	...	...	...	
18974	247223	Xia Ao	Ao Xia	https://cdn.sofifa.com/players/247/223/21_60.png	http://sofifa.com/player/247223/ao-xia/210006/	China PR	21	47	
18975	258760	B. Hough	Ben Hough	https://cdn.sofifa.com/players/258/760/21_60.png	http://sofifa.com/player/258760/ben-hough/210006/	England	17	47	
18976	252757	R. McKinley	Ronan McKinley	https://cdn.sofifa.com/players/252/757/21_60.png	http://sofifa.com/player/252757/ronan-mckinley...	England	18	47	
18977	243790	Wang Zhen'ao	Zhen'ao Wang	https://cdn.sofifa.com/players/243/790/21_60.png	http://sofifa.com/player/243790/zhenao-wang/21...	China PR	20	47	
18978	252520	Zhou Xiao	Xiao Zhou	https://cdn.sofifa.com/players/252/520/21_60.png	http://sofifa.com/player/252520/xiao-zhou/210006/	China PR	21	47	

18979 rows × 77 columns

In [21]:
```python
df.drop('Loan Date End', 1,inplace=True)
```

In [22]:
```python
df.drop('Hits', 1, inplace=True)
```

In [23]:
```python
print(df['Value'])
print(df['Wage'])
print(df['Release Clause'])
```

```
0        €103.5M
1          €63M
```

```
2        €120M
3        €129M
4        €132M
           ...
18974    €100K
18975    €130K
18976    €120K
18977    €100K
18978    €100K
Name: Value, Length: 18979, dtype: object
0        €560K
1        €220K
2        €125K
3        €370K
4        €270K
           ...
18974     €1K
18975    €500
18976    €500
18977     €2K
18978     €1K
Name: Wage, Length: 18979, dtype: object
0        €138.4M
1         €75.9M
2        €159.4M
3          €161M
4        €166.5M
           ...
18974     €70K
18975    €165K
18976    €131K
18977     €88K
18978     €79K
Name: Release Clause, Length: 18979, dtype: object
```

In [24]:
```python
df['Value']=df['Value'].str.replace('M','').str.replace('K','')
df['Wage']=df['Wage'].str.replace('K','')
df['Release Clause']=df['Release Clause'].str.replace('M','')
```

In [25]:
```python
print(df['Value'])
print(df['Wage'])
print(df['Release Clause'])
```

```
0        €103.5
1          €63
2         €120
3         €129
4         €132
           ...
18974     €100
18975     €130
18976     €120
18977     €100
18978     €100
Name: Value, Length: 18979, dtype: object
0         €560
1         €220
2         €125
3         €370
4         €270
           ...
18974      €1
18975    €500
18976    €500
18977      €2
18978      €1
Name: Wage, Length: 18979, dtype: object
0        €138.4
1         €75.9
2        €159.4
3          €161
4        €166.5
           ...
18974     €70K
18975    €165K
18976    €131K
18977     €88K
18978     €79K
Name: Release Clause, Length: 18979, dtype: object
```

In [26]:
```python
print(df['W/F'])
```

```
print(df['SM'])
print(df['IR'])
```

```
0         4 ★
1         4 ★
2         3 ★
3         5 ★
4         5 ★
          ...
18974     2 ★
18975     2 ★
18976     2 ★
18977     3 ★
18978     3 ★
Name: W/F, Length: 18979, dtype: object
0         4★
1         5★
2         1★
3         4★
4         5★
          ..
18974     2★
18975     2★
18976     2★
18977     2★
18978     2★
Name: SM, Length: 18979, dtype: object
0         5 ★
1         5 ★
2         3 ★
3         4 ★
4         5 ★
          ...
18974     1 ★
18975     1 ★
18976     1 ★
18977     1 ★
18978     1 ★
Name: IR, Length: 18979, dtype: object
```

In [27]:
```python
df=df.astype(str).apply(lambda x: x.str.encode('ascii', 'ignore').str.decode('ascii'))
```

In [28]:
```python
print(df['W/F'])
print(df['SM'])
print(df['IR'])
```

```
0         4
1         4
2         3
3         5
4         5
          ..
18974     2
18975     2
18976     2
18977     3
18978     3
Name: W/F, Length: 18979, dtype: object
0         4
1         5
2         1
3         4
4         5
          ..
18974     2
18975     2
18976     2
18977     2
18978     2
Name: SM, Length: 18979, dtype: object
0         5
1         5
2         3
3         4
4         5
          ..
18974     1
18975     1
18976     1
18977     1
18978     1
Name: IR, Length: 18979, dtype: object
```

`df`

	ID	Name	LongName	photoUrl	playerUrl	Nationality	Age	↓OVA
0	158023	L. Messi	Lionel Messi	https://cdn.sofifa.com/players/158/023/21_60.png	http://sofifa.com/player/158023/lionel-messi/2...	Argentina	33	93
1	20801	Cristiano Ronaldo	C. Ronaldo dos Santos Aveiro	https://cdn.sofifa.com/players/020/801/21_60.png	http://sofifa.com/player/20801/c-ronaldo-dos-s...	Portugal	35	92
2	200389	J. Oblak	Jan Oblak	https://cdn.sofifa.com/players/200/389/21_60.png	http://sofifa.com/player/200389/jan-oblak/210006/	Slovenia	27	91
3	192985	K. De Bruyne	Kevin De Bruyne	https://cdn.sofifa.com/players/192/985/21_60.png	http://sofifa.com/player/192985/kevin-de-bruyn...	Belgium	29	91
4	190871	Neymar Jr	Neymar da Silva Santos Jr.	https://cdn.sofifa.com/players/190/871/21_60.png	http://sofifa.com/player/190871/neymar-da-silv...	Brazil	28	91
...	...	...	...	...	...	...	...	...
18974	247223	Xia Ao	Ao Xia	https://cdn.sofifa.com/players/247/223/21_60.png	http://sofifa.com/player/247223/ao-xia/210006/	China PR	21	47
18975	258760	B. Hough	Ben Hough	https://cdn.sofifa.com/players/258/760/21_60.png	http://sofifa.com/player/258760/ben-hough/210006/	England	17	47
18976	252757	R. McKinley	Ronan McKinley	https://cdn.sofifa.com/players/252/757/21_60.png	http://sofifa.com/player/252757/ronan-mckinley...	England	18	47
18977	243790	Wang Zhen'ao	Zhen'ao Wang	https://cdn.sofifa.com/players/243/790/21_60.png	http://sofifa.com/player/243790/zhenao-wang/21...	China PR	20	47
18978	252520	Zhou Xiao	Xiao Zhou	https://cdn.sofifa.com/players/252/520/21_60.png	http://sofifa.com/player/252520/xiao-zhou/210006/	China PR	21	47

18979 rows × 75 columns

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js