

In [1]: 1 `import pandas as pd`

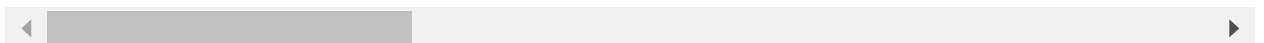
In [2]: 1 `data = pd.read_csv('Fifa.csv', index_col=0)`

In [3]: 1 `data`

Out[3]:

	ID	Name	Age	Photo	Nationality	
0	158023	L. Messi	31	<a href="https://cdn.sofifa.org/players/4/19/158023.png">https://cdn.sofifa.org/players/4/19/158023.png</a>	Argentina	<a href="https://cdn.sofifa.org/players/4/19/158023.png">https://cdn.sofifa.org/players/4/19/158023.png</a>
1	20801	Cristiano Ronaldo	33	<a href="https://cdn.sofifa.org/players/4/19/20801.png">https://cdn.sofifa.org/players/4/19/20801.png</a>	Portugal	<a href="https://cdn.sofifa.org/players/4/19/20801.png">https://cdn.sofifa.org/players/4/19/20801.png</a>
2	190871	Neymar Jr	26	<a href="https://cdn.sofifa.org/players/4/19/190871.png">https://cdn.sofifa.org/players/4/19/190871.png</a>	Brazil	<a href="https://cdn.sofifa.org/players/4/19/190871.png">https://cdn.sofifa.org/players/4/19/190871.png</a>
3	193080	De Gea	27	<a href="https://cdn.sofifa.org/players/4/19/193080.png">https://cdn.sofifa.org/players/4/19/193080.png</a>	Spain	<a href="https://cdn.sofifa.org/players/4/19/193080.png">https://cdn.sofifa.org/players/4/19/193080.png</a>
4	192985	K. De Bruyne	27	<a href="https://cdn.sofifa.org/players/4/19/192985.png">https://cdn.sofifa.org/players/4/19/192985.png</a>	Belgium	<a href="https://cdn.sofifa.org/players/4/19/192985.png">https://cdn.sofifa.org/players/4/19/192985.png</a>
...	...	...	...	...	...	...
18202	238813	J. Lundstram	19	<a href="https://cdn.sofifa.org/players/4/19/238813.png">https://cdn.sofifa.org/players/4/19/238813.png</a>	England	<a href="https://cdn.sofifa.org/players/4/19/238813.png">https://cdn.sofifa.org/players/4/19/238813.png</a>
18203	243165	N. Christoffersson	19	<a href="https://cdn.sofifa.org/players/4/19/243165.png">https://cdn.sofifa.org/players/4/19/243165.png</a>	Sweden	<a href="https://cdn.sofifa.org/players/4/19/243165.png">https://cdn.sofifa.org/players/4/19/243165.png</a>
18204	241638	B. Worman	16	<a href="https://cdn.sofifa.org/players/4/19/241638.png">https://cdn.sofifa.org/players/4/19/241638.png</a>	England	<a href="https://cdn.sofifa.org/players/4/19/241638.png">https://cdn.sofifa.org/players/4/19/241638.png</a>
18205	246268	D. Walker-Rice	17	<a href="https://cdn.sofifa.org/players/4/19/246268.png">https://cdn.sofifa.org/players/4/19/246268.png</a>	England	<a href="https://cdn.sofifa.org/players/4/19/246268.png">https://cdn.sofifa.org/players/4/19/246268.png</a>
18206	246269	G. Nugent	16	<a href="https://cdn.sofifa.org/players/4/19/246269.png">https://cdn.sofifa.org/players/4/19/246269.png</a>	England	<a href="https://cdn.sofifa.org/players/4/19/246269.png">https://cdn.sofifa.org/players/4/19/246269.png</a>

18207 rows × 88 columns



In [4]: 1 data.info()

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 18207 entries, 0 to 18206
Data columns (total 88 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   ID                                     18207 non-null  int64
1   Name                                  18207 non-null  object
2   Age                                   18207 non-null  int64
3   Photo                                 18207 non-null  object
4   Nationality                           18207 non-null  object
5   Flag                                  18207 non-null  object
6   Overall                               18207 non-null  int64
7   Potential                             18207 non-null  int64
8   Club                                  17966 non-null  object
9   Club Logo                             18207 non-null  object
10  Value                                 18207 non-null  object
11  Wage                                  18207 non-null  object
12  Special                               18207 non-null  int64
13  Preferred Foot                         18159 non-null  object
14  ...
```

In [5]: 1 import numpy as np

In [6]: 1 np.unique(data['ID'])

Out[6]: array([ 16, 41, 80, ..., 246616, 246617, 246620], dtype=int64)

In [7]: 1 np.unique(data['Preferred Foot'])

```
-----
TypeError                                Traceback (most recent call last)
<ipython-input-7-da40d52cb153> in <module>
----> 1 np.unique(data['Preferred Foot'])

<__array_function__ internals> in unique(*args, **kwargs)

~\anaconda3\lib\site-packages\numpy\lib\arraysetops.py in unique(ar, return_index, return_inverse, return_counts, axis)
    261     ar = np.asanyarray(ar)
    262     if axis is None:
--> 263         ret = _unique1d(ar, return_index, return_inverse, return_counts)
    264         return _unpack_tuple(ret)
    265

~\anaconda3\lib\site-packages\numpy\lib\arraysetops.py in _unique1d(ar, return_index, return_inverse, return_counts)
    309     aux = ar[perm]
    310     else:
--> 311         ar.sort()
    312         aux = ar
    313     mask = np.empty(aux.shape, dtype=np.bool_)

TypeError: '<' not supported between instances of 'str' and 'float'
```

In [8]: 1 data.isnull().sum()

```
Out[8]: ID                0
        Name              0
        Age              0
        Photo            0
        Nationality      0
        ...
        GKHandling       48
        GKKicking        48
        GKPositioning    48
        GKReflexes       48
        Release Clause   1564
        Length: 88, dtype: int64
```

In [9]: 1 copy = data.copy()

In [10]: 1 copy

Out[10]:

	ID	Name	Age	Photo	Nationality	
0	158023	L. Messi	31	https://cdn.sofifa.org/players/4/19/158023.png	Argentina	https://
1	20801	Cristiano Ronaldo	33	https://cdn.sofifa.org/players/4/19/20801.png	Portugal	https://
2	190871	Neymar Jr	26	https://cdn.sofifa.org/players/4/19/190871.png	Brazil	https://
3	193080	De Gea	27	https://cdn.sofifa.org/players/4/19/193080.png	Spain	https://
4	192985	K. De Bruyne	27	https://cdn.sofifa.org/players/4/19/192985.png	Belgium	https://
...	...	...	...	...	...	...
18202	238813	J. Lundstram	19	https://cdn.sofifa.org/players/4/19/238813.png	England	https://

In [11]: 1 copy['Nationality'].value\_counts()

Out[11]:

England	1662
Germany	1198
Spain	1072
Argentina	937
France	914
...	
South Sudan	1
Ethiopia	1
Malta	1
Mauritius	1
Palestine	1

Name: Nationality, Length: 164, dtype: int64

In [12]: 1 copy['Club'].value\_counts()

Out[12]:

FC Barcelona	33
Manchester United	33
Fortuna Düsseldorf	33
Eintracht Frankfurt	33
Burnley	33
..	
Sport Club do Recife	20
Internacional	20
Sligo Rovers	19
Limerick FC	19
Derry City	18

Name: Club, Length: 651, dtype: int64

In [13]: 1 copy['Club'].fillna(copy['Club'].value\_counts().index[0], inplace = True)

In [14]: 1 copy.loc[568,:]

Out[14]: ID 229909  
 Name A. Lunev  
 Age 26  
 Photo <https://cdn.sofifa.org/players/4/19/229909.png> (<https://cdn.sofifa.org/players/4/19/229909.png>)  
 Nationality Russia  
 ...  
 GKHandling 73  
 GKKicking 65  
 GKPositioning 77  
 GKReflexes 85  
 Release Clause NaN  
 Name: 568, Length: 88, dtype: object

In [15]: 1 copy.isna().sum()

Out[15]: ID 0  
 Name 0  
 Age 0  
 Photo 0  
 Nationality 0  
 ...  
 GKHandling 48  
 GKKicking 48  
 GKPositioning 48  
 GKReflexes 48  
 Release Clause 1564  
 Length: 88, dtype: int64

In [16]: 1 copy['Preferred Foot'].mode()

Out[16]: 0 Right  
 dtype: object

In [17]: 1 copy['Preferred Foot'].fillna(copy['Preferred Foot'].mode().index[0],inplace

In [18]: 1 copy.isna().sum()

Out[18]: ID 0  
 Name 0  
 Age 0  
 Photo 0  
 Nationality 0  
 ...  
 GKHandling 48  
 GKKicking 48  
 GKPositioning 48  
 GKReflexes 48  
 Release Clause 1564  
 Length: 88, dtype: int64

```
In [19]: 1 copy['International Reputation'].mode()
```

```
Out[19]: 0      1.0  
dtype: float64
```

```
In [20]: 1 copy['International Reputation'].value_counts()
```

```
Out[20]: 1.0      16532  
2.0       1261  
3.0        309  
4.0         51  
5.0          6  
Name: International Reputation, dtype: int64
```

```
In [21]: 1 copy['International Reputation'].fillna(copy['International Reputation'].mod
```

```
In [22]: 1 copy.isna().sum()
```

```
Out[22]: ID                0  
Name                0  
Age                0  
Photo              0  
Nationality        0  
...  
GKHandling         48  
GK Kicking         48  
GK Positioning     48  
GK Reflexes        48  
Release Clause    1564  
Length: 88, dtype: int64
```

```
In [23]: 1 copy['Weak Foot'].value_counts()
```

```
Out[23]: 3.0      11349  
2.0       3761  
4.0       2662  
5.0        229  
1.0        158  
Name: Weak Foot, dtype: int64
```

```
In [24]: 1 copy['Weak Foot'].value_counts().index[-1]
```

```
Out[24]: 1.0
```

```
In [25]: 1 copy['Weak Foot'].fillna(copy['Weak Foot'].value_counts().index[-1],inplace
```

```
In [26]: 1 copy.isna().sum()
```

```
Out[26]: ID                0
         Name              0
         Age               0
         Photo             0
         Nationality       0
         ...
         GKHandling        48
         GKKicking         48
         GKPositioning     48
         GKReflexes        48
         Release Clause    1564
         Length: 88, dtype: int64
```

```
In [27]: 1 copy = copy.apply(lambda x:x.fillna(x.mode().index[0]) if x.dtype=='float' else x)
```

```
In [28]: 1 copy.isna().sum()
```

```
Out[28]: ID                0
         Name              0
         Age               0
         Photo             0
         Nationality       0
         ..
         GKHandling        0
         GKKicking         0
         GKPositioning     0
         GKReflexes        0
         Release Clause     0
         Length: 88, dtype: int64
```

In [29]: 1 pd.crosstab(index=copy[''], columns=copy[''], dropna=True)

```
-----
KeyError                                Traceback (most recent call last)
~\anaconda3\lib\site-packages\pandas\core\indexes\base.py in get_loc(self, key,
method, tolerance)
    2645         try:
-> 2646             return self._engine.get_loc(key)
    2647         except KeyError:

pandas\_libs\index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas\_libs\index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashT
able.get_item()

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashT
able.get_item()
```

**KeyError:** ''

During handling of the above exception, another exception occurred:

```
KeyError                                Traceback (most recent call last)
<ipython-input-29-80fcfb422706> in <module>
----> 1 pd.crosstab(index=copy[''], columns=copy[''], dropna=True)

~\anaconda3\lib\site-packages\pandas\core\frame.py in __getitem__(self, key)
    2798         if self.columns.nlevels > 1:
    2799             return self._getitem_multilevel(key)
-> 2800         indexer = self.columns.get_loc(key)
    2801         if is_integer(indexer):
    2802             indexer = [indexer]

~\anaconda3\lib\site-packages\pandas\core\indexes\base.py in get_loc(self, key,
method, tolerance)
    2646         return self._engine.get_loc(key)
    2647         except KeyError:
-> 2648             return self._engine.get_loc(self._maybe_cast_indexer(key))
    2649         indexer = self.get_indexer([key], method=method, tolerance=tolerance)
    2650         if indexer.ndim > 1 or indexer.size > 1:

pandas\_libs\index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas\_libs\index.pyx in pandas._libs.index.IndexEngine.get_loc()

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashT
able.get_item()

pandas\_libs\hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashT
able.get_item()
```

**KeyError:** ''



In [ ]:

1	
---	--