

```
In [1]: import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import pandas as pd
%matplotlib inline
import warnings
warnings.filterwarnings('ignore')
```

```
In [2]: data=pd.read_csv(r"C:\Users\S SHYAMILI\OneDrive\Desktop\data science\28th - Seaborn
```

```
In [3]: 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	Argentinian
1	1	20801	Cristiano Ronaldo	33	https://cdn.sofifa.org/players/4/19/20801.png	Portuguese
2	2	190871	Neymar Jr	26	https://cdn.sofifa.org/players/4/19/190871.png	Brazilian
3	3	193080	De Gea	27	https://cdn.sofifa.org/players/4/19/193080.png	Spanish
4	4	192985	K. De Bruyne	27	https://cdn.sofifa.org/players/4/19/192985.png	Belgian

5 rows × 89 columns



```
In [4]: data.tail()
```

Out[4]:

	Unnamed: 0	ID	Name	Age	Photo
18202	18202	238813	J. Lundstram	19	https://cdn.sofifa.org/players/4/19/238813.png
18203	18203	243165	N. Christoffersson	19	https://cdn.sofifa.org/players/4/19/243165.png
18204	18204	241638	B. Worman	16	https://cdn.sofifa.org/players/4/19/241638.png
18205	18205	246268	D. Walker-Rice	17	https://cdn.sofifa.org/players/4/19/246268.png
18206	18206	246269	G. Nugent	16	https://cdn.sofifa.org/players/4/19/246269.png

5 rows × 89 columns



In [5]: `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
```

```
In [11]: data.shape
```

```
Out[11]: (18207, 89)
```

```
In [13]: data.isnull()
```

Out[13]:

	Unnamed: 0	ID	Name	Age	Photo	Nationality	Flag	Overall	Potential	Club
0	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	False
4	False	False	False	False	False	False	False	False	False	False
...
18202	False	False	False	False	False	False	False	False	False	False
18203	False	False	False	False	False	False	False	False	False	False
18204	False	False	False	False	False	False	False	False	False	False
18205	False	False	False	False	False	False	False	False	False	False
18206	False	False	False	False	False	False	False	False	False	False

18207 rows × 89 columns

In [15]: `data.isnull().sum()`

```
Out[15]: Unnamed: 0          0
ID              0
Name            0
Age             0
Photo           0
...
GKHandling      48
GKKicking       48
GKPositioning   48
GKReflexes      48
Release Clause  1564
Length: 89, dtype: int64
```

In [17]: `data.columns`

```
Out[17]: 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')
```

```
In [19]: data['Body Type'].value_counts()
```

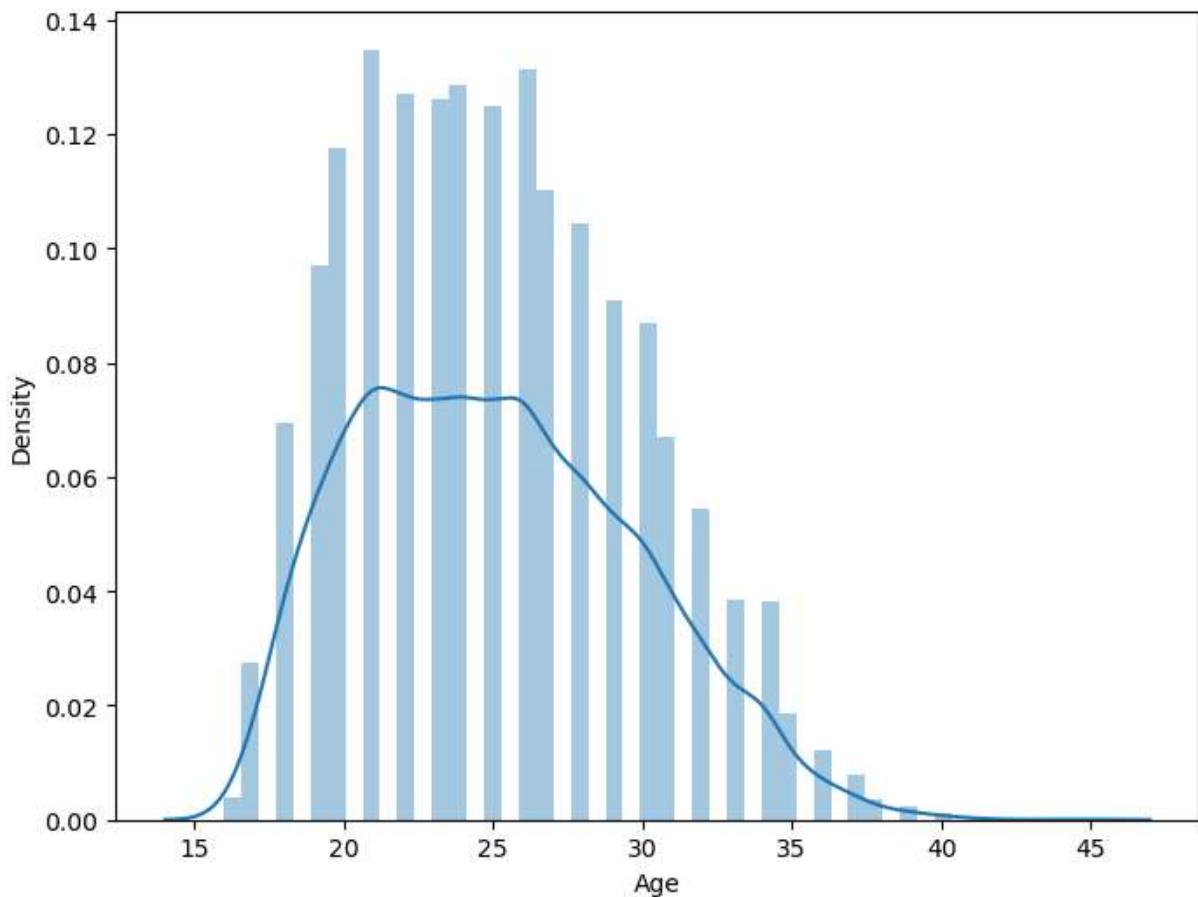
```
Out[19]: Body Type
Normal           10595
Lean              6417
Stocky             1140
Messi                1
C. Ronaldo          1
Neymar                1
Courtois               1
PLAYER_BODY_TYPE_25        1
Shaqiri                1
Akinfenwa               1
Name: count, dtype: int64
```

```
In [21]: data.columns.value_counts()
```

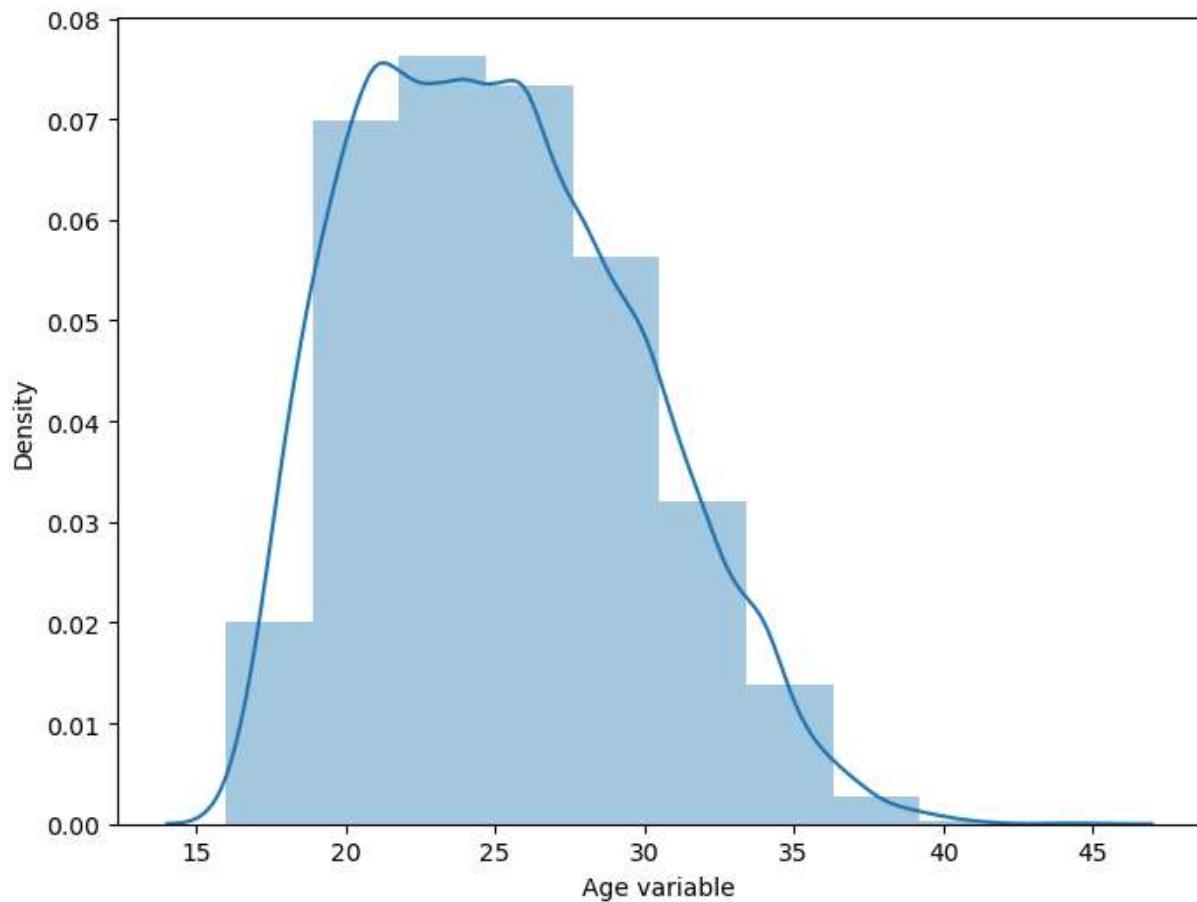
```
Out[21]: Unnamed: 0           1
LDM                 1
SprintSpeed          1
Acceleration          1
BallControl           1
..
LS                  1
Weight                1
Height                1
Contract Valid Until    1
Release Clause          1
Name: count, Length: 89, dtype: int64
```

```
In [23]: plt.subplots(figsize=(8,6))
sns.distplot(data['Age'])
```

```
Out[23]: <Axes: xlabel='Age', ylabel='Density'>
```

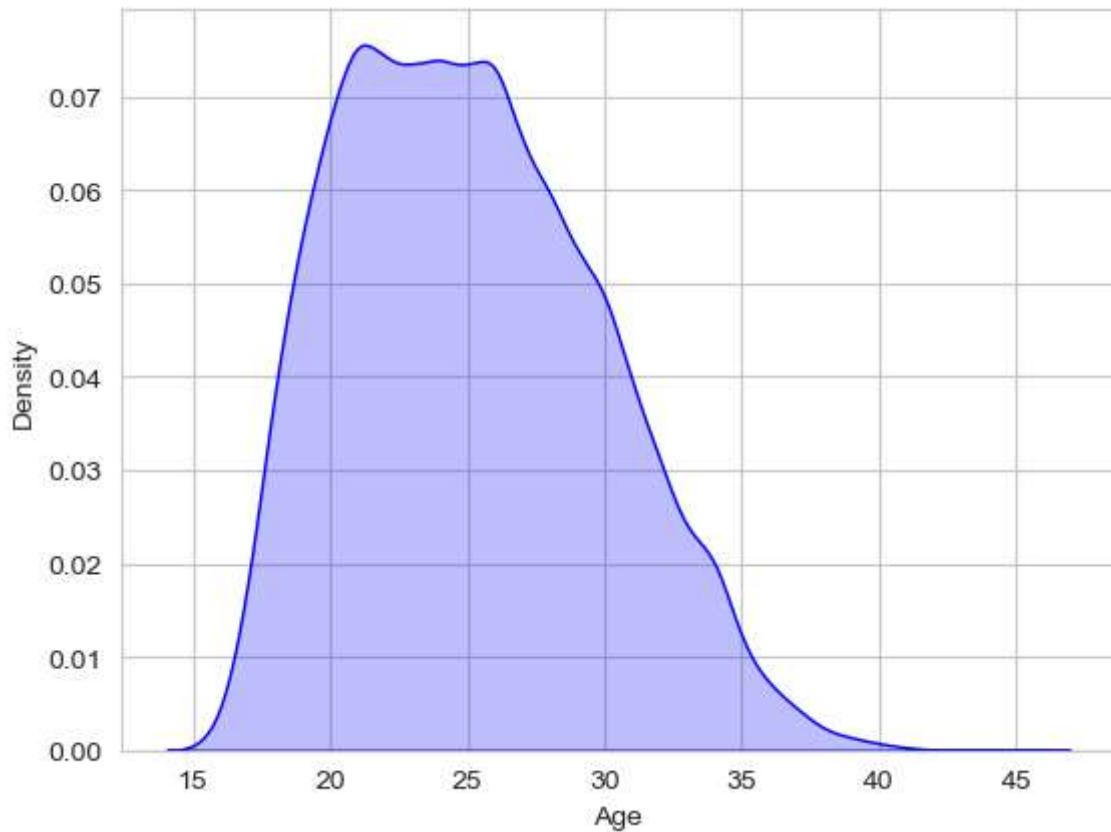


```
In [24]: f, ax = plt.subplots(figsize=(8,6))
x = data['Age']
x = pd.Series(x, name="Age variable")
ax = sns.distplot(x, bins=10)
plt.show()
```



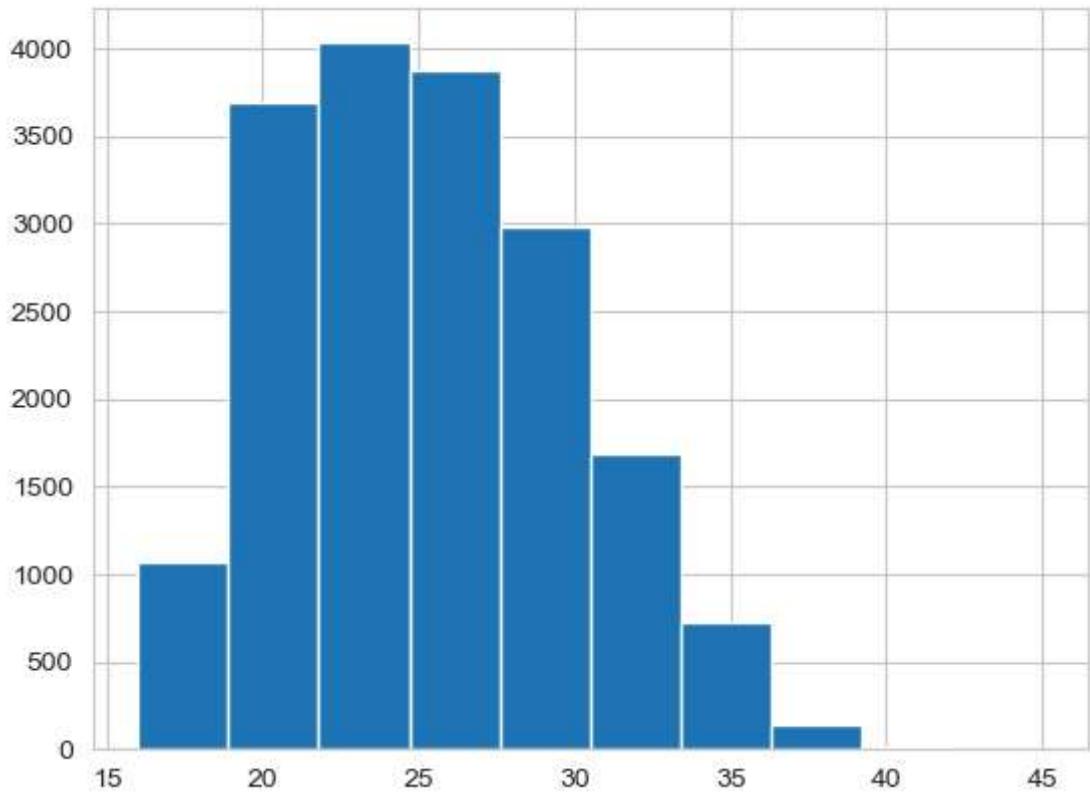
```
In [25]: sns.set_style('whitegrid')
sns.kdeplot(data=data,x='Age',shade=True,color='b')
```

```
Out[25]: <Axes: xlabel='Age', ylabel='Density'>
```

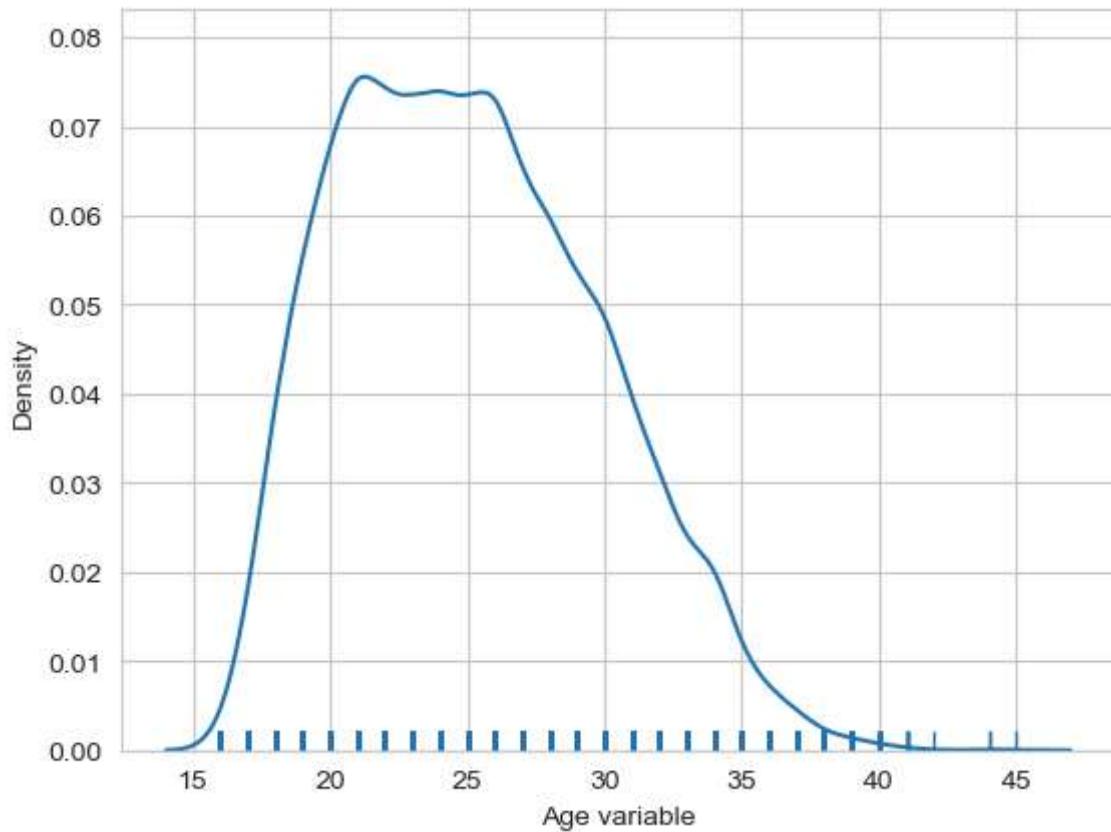


```
In [27]: plt.hist(data.Age)
```

```
Out[27]: (array([1.063e+03, 3.687e+03, 4.030e+03, 3.868e+03, 2.977e+03, 1.689e+03,
   7.270e+02, 1.440e+02, 1.900e+01, 3.000e+00]),  
 array([16. , 18.9, 21.8, 24.7, 27.6, 30.5, 33.4, 36.3, 39.2, 42.1, 45. ]),  
 <BarContainer object of 10 artists>)
```



```
In [29]: ax = sns.distplot(x, hist=False, rug=True, bins=10)
```



```
In [30]: data['Preferred Foot'].value_counts()
```

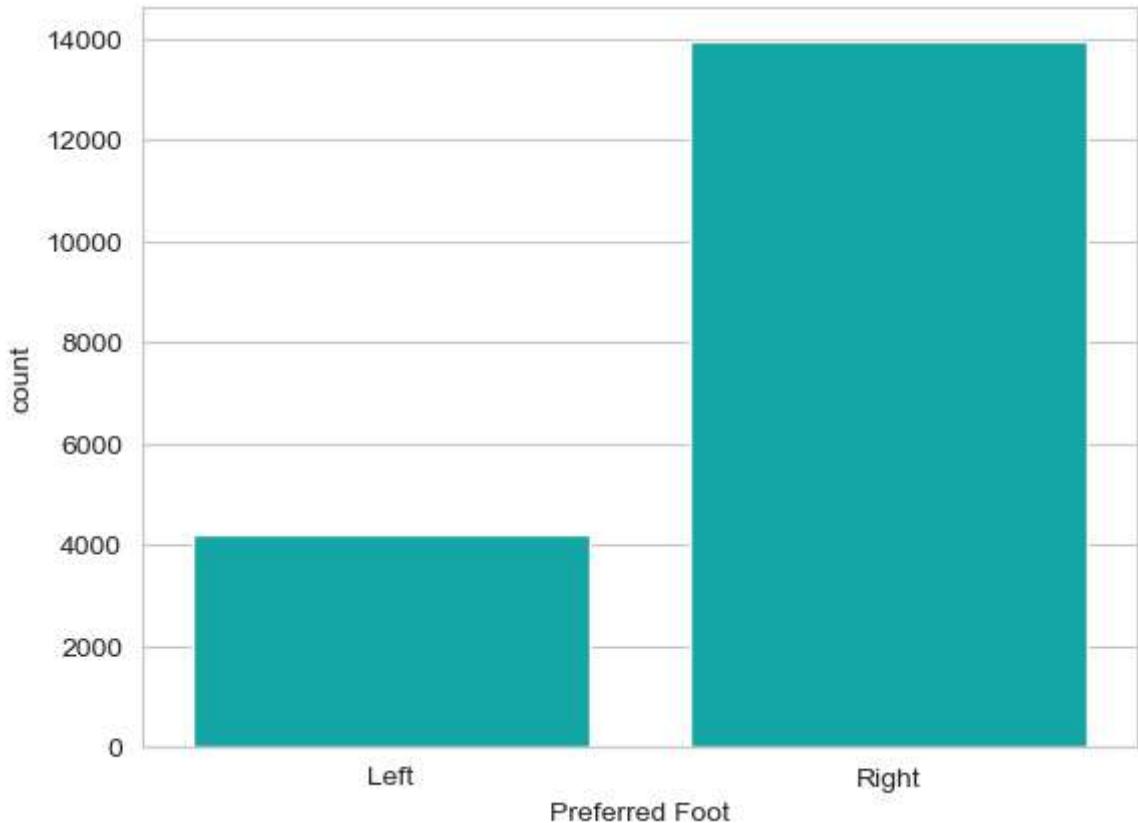
```
Out[30]: Preferred Foot
Right    13948
Left     4211
Name: count, dtype: int64
```

```
In [32]: data['Preferred Foot'].nunique()
```

```
Out[32]: 2
```

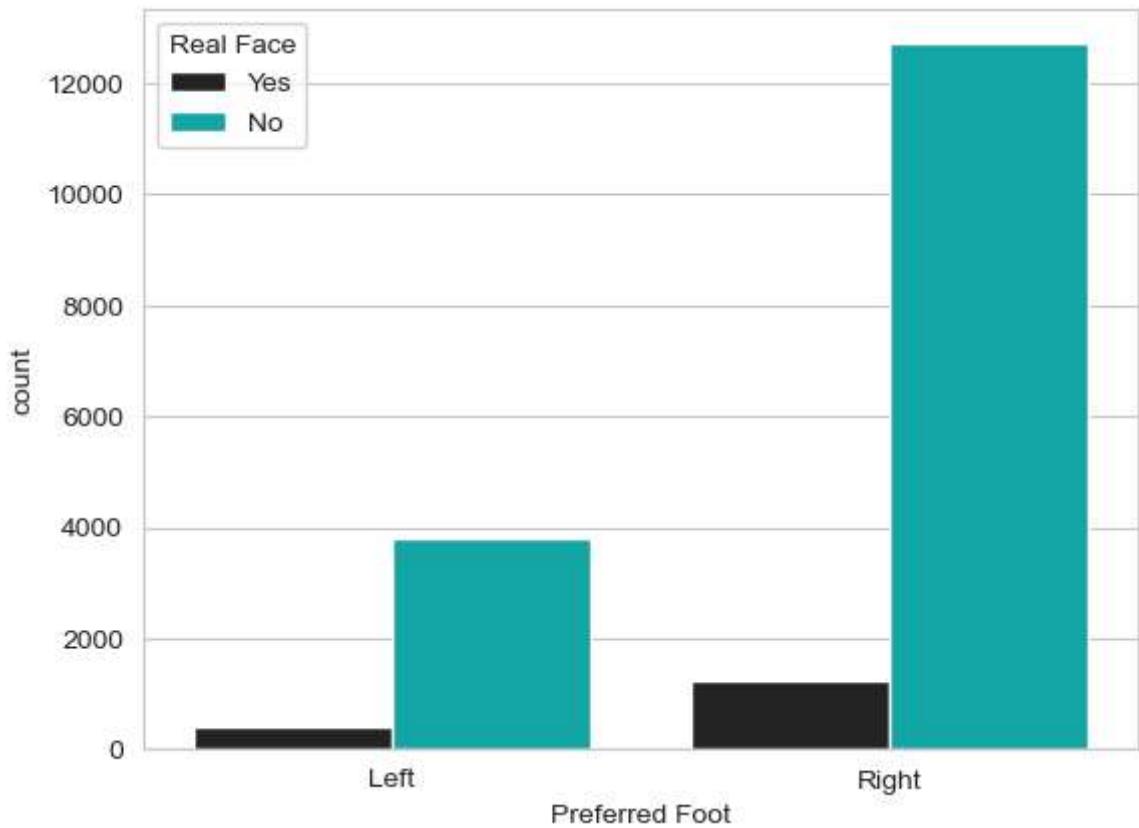
```
In [34]: sns.countplot(x='Preferred Foot', data=data, color='c')
```

```
Out[34]: <Axes: xlabel='Preferred Foot', ylabel='count'>
```



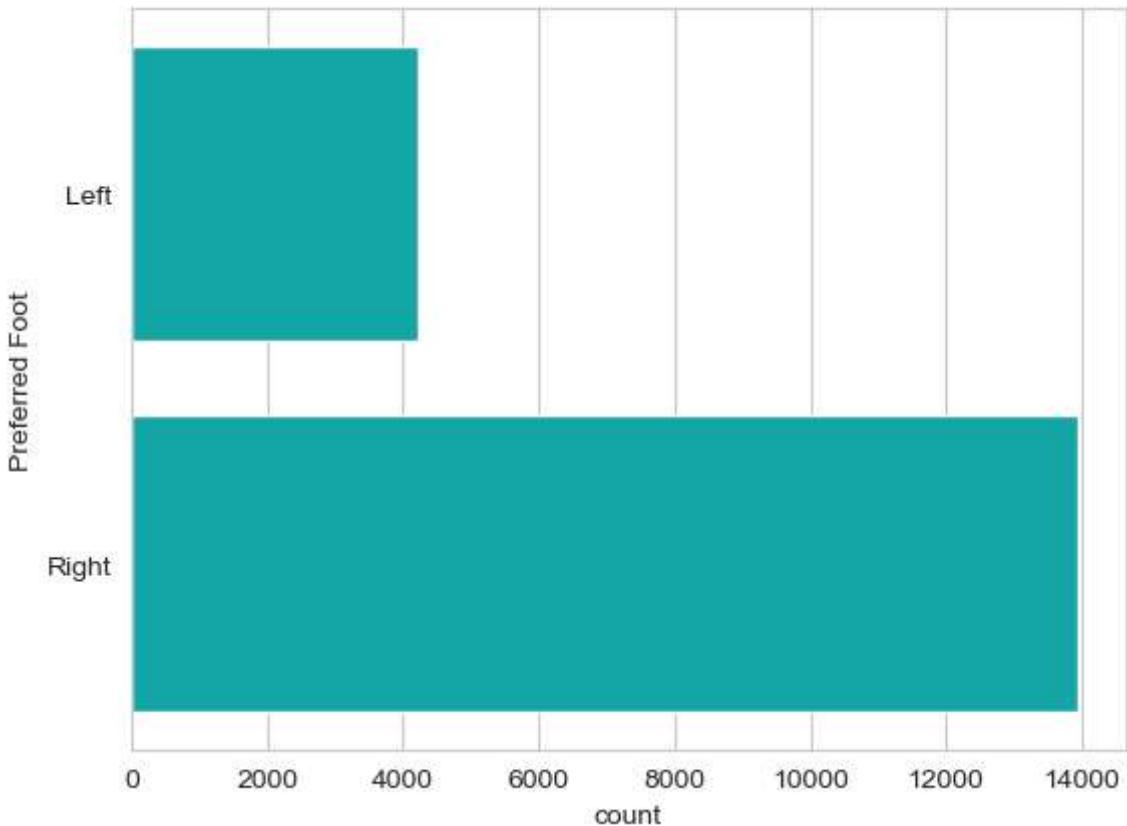
```
In [35]: sns.countplot(x='Preferred Foot', data=data, color='c', hue='Real Face')
```

```
Out[35]: <Axes: xlabel='Preferred Foot', ylabel='count'>
```



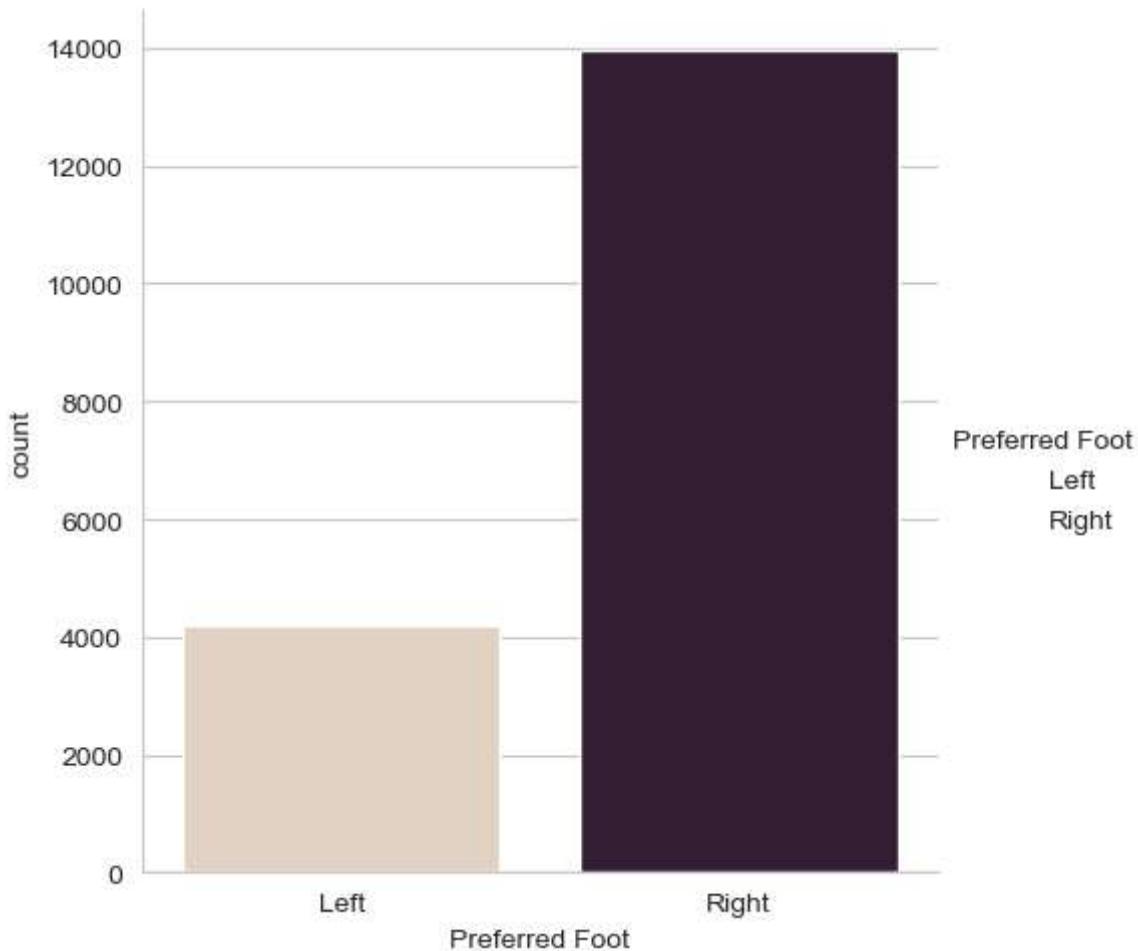
```
In [49]: sns.countplot(y="Preferred Foot", data=data, color="c")
```

```
Out[49]: <Axes: xlabel='count', ylabel='Preferred Foot'>
```



```
In [55]: sns.catplot(x="Preferred Foot", kind="count", palette="ch:.25", data=data)
```

```
Out[55]: <seaborn.axisgrid.FacetGrid at 0x25398537140>
```



```
In [57]: data['International Reputation'].nunique()
```

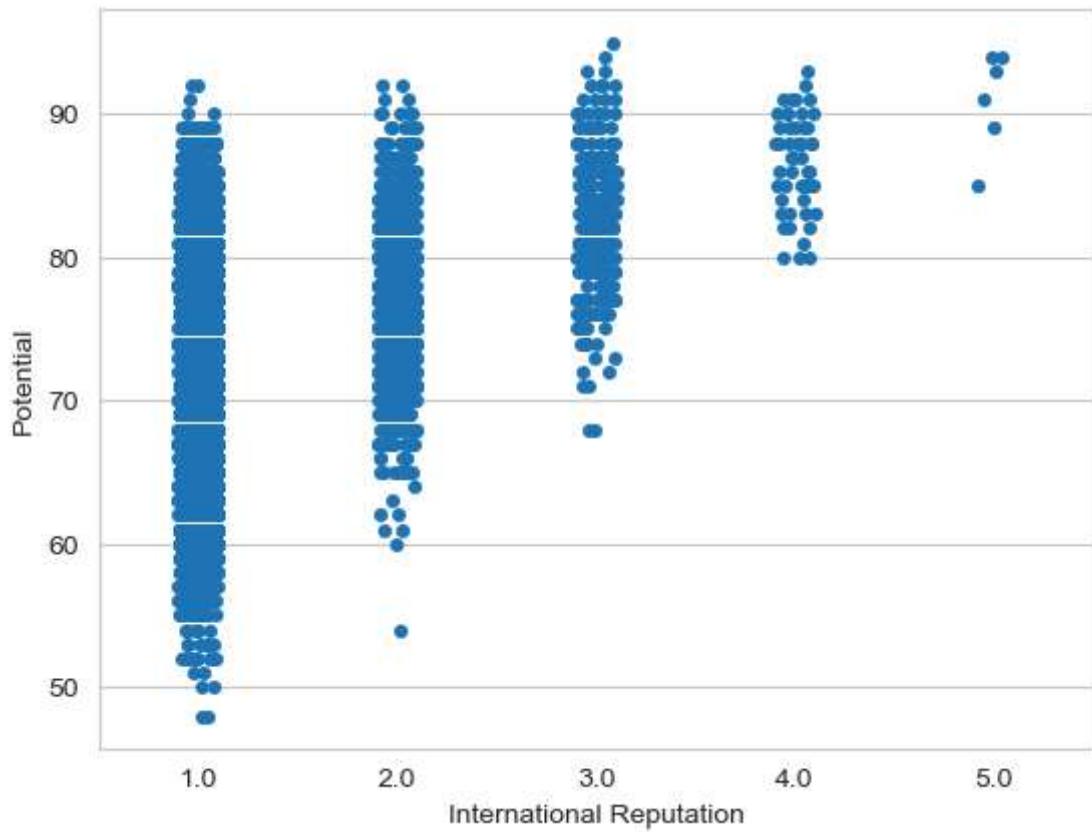
```
Out[57]: 5
```

```
In [59]: data['International Reputation'].value_counts()
```

```
Out[59]: International Reputation
1.0    16532
2.0    1261
3.0    309
4.0     51
5.0      6
Name: count, dtype: int64
```

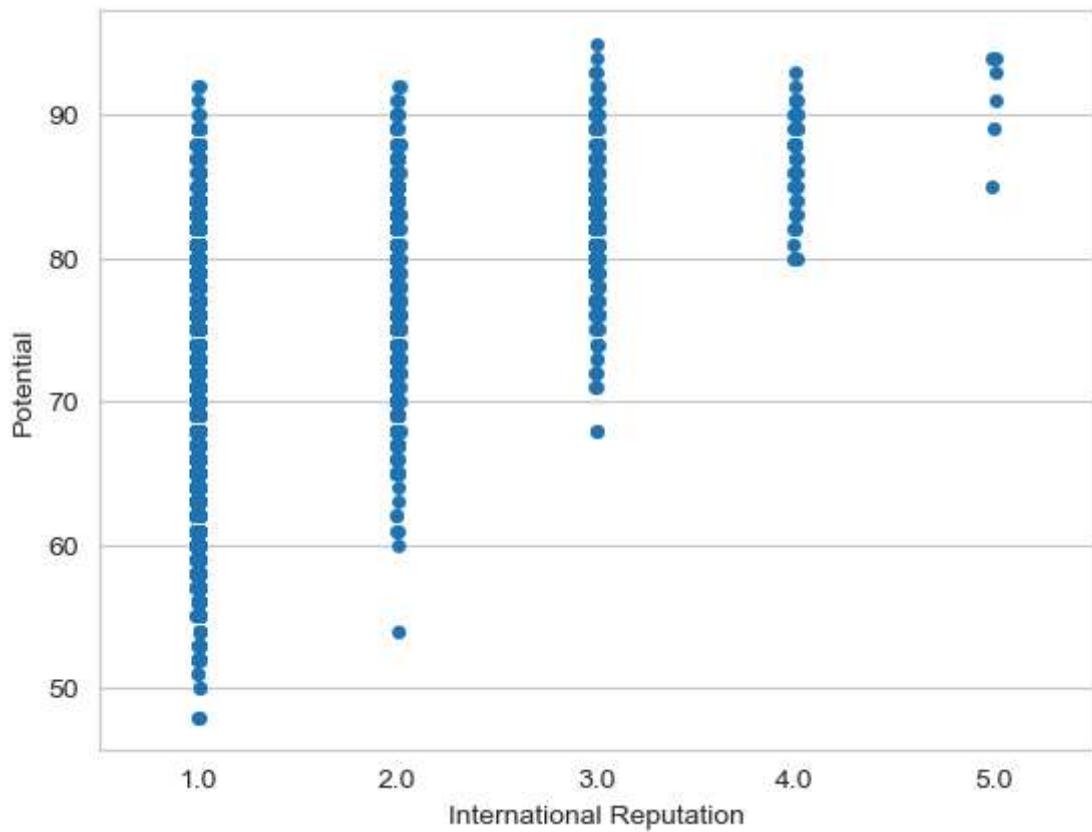
```
In [61]: sns.stripplot(x="International Reputation", y="Potential", data=data)
```

```
Out[61]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



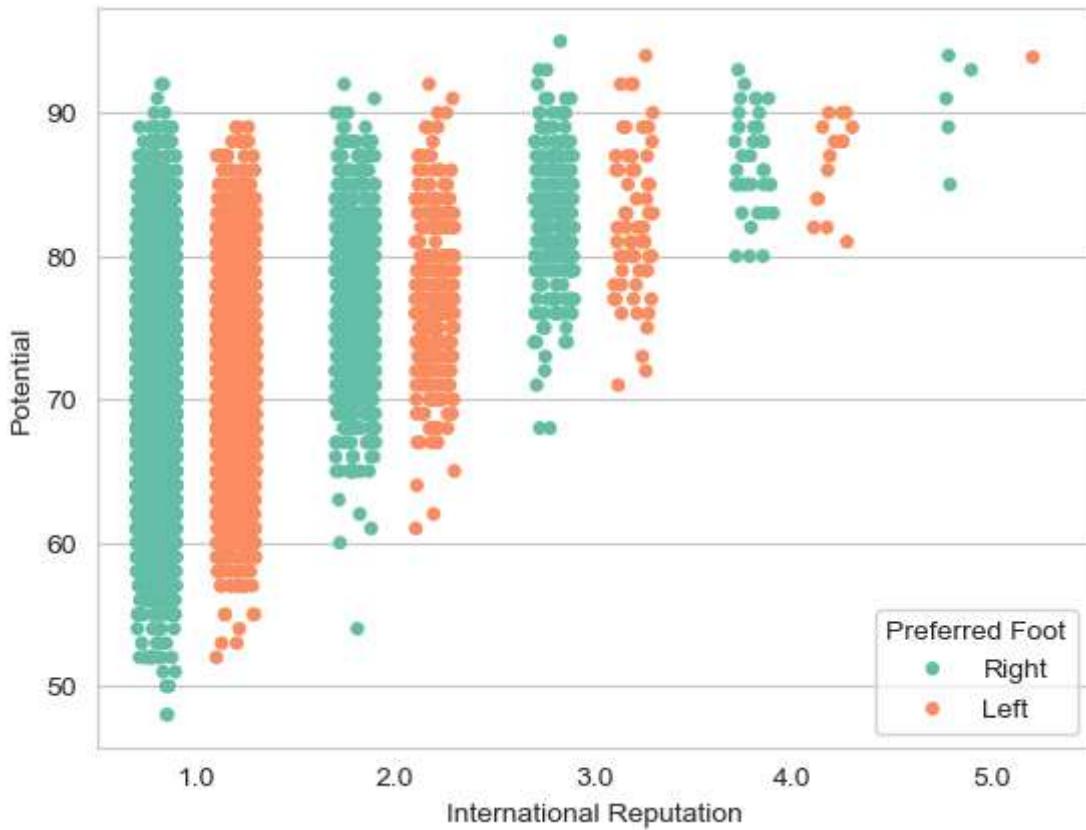
```
In [65]: sns.stripplot(x="International Reputation", y="Potential", data=data, jitter=0.01)
```

```
Out[65]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



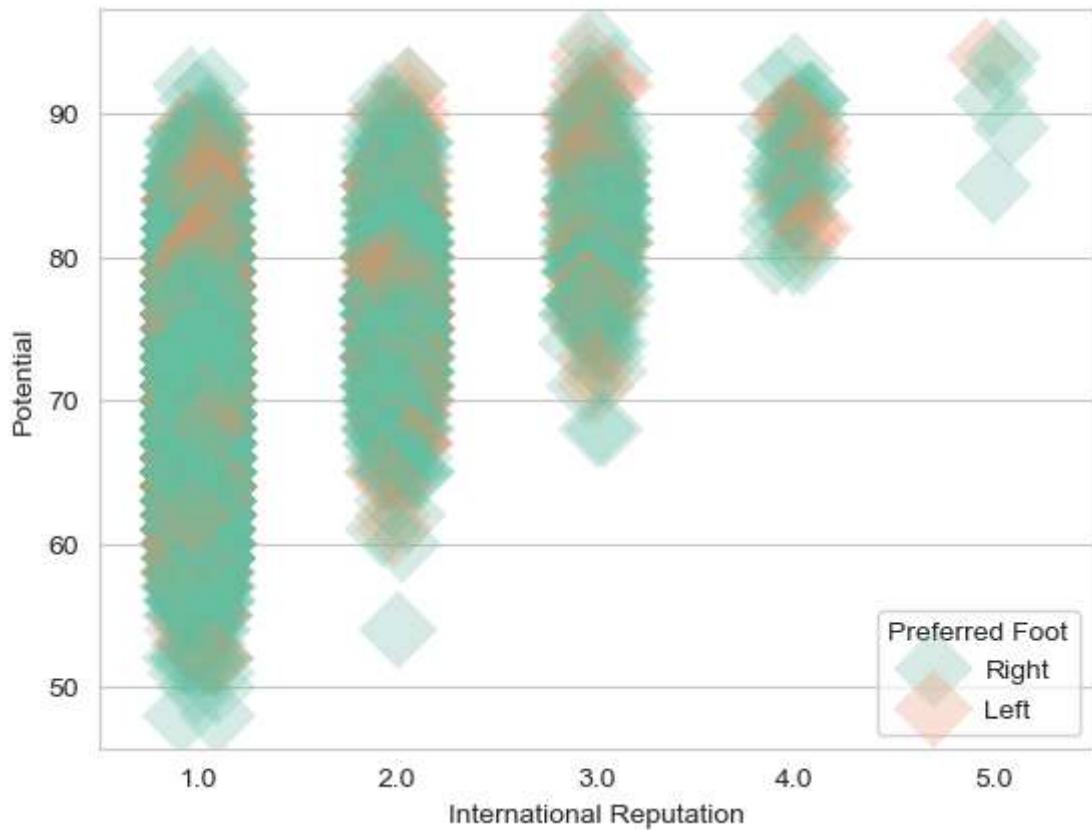
```
In [67]: sns.stripplot(x="International Reputation", y="Potential", hue="Preferred Foot",
                      data=data, jitter=0.2, palette="Set2", dodge=True)
```

```
Out[67]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



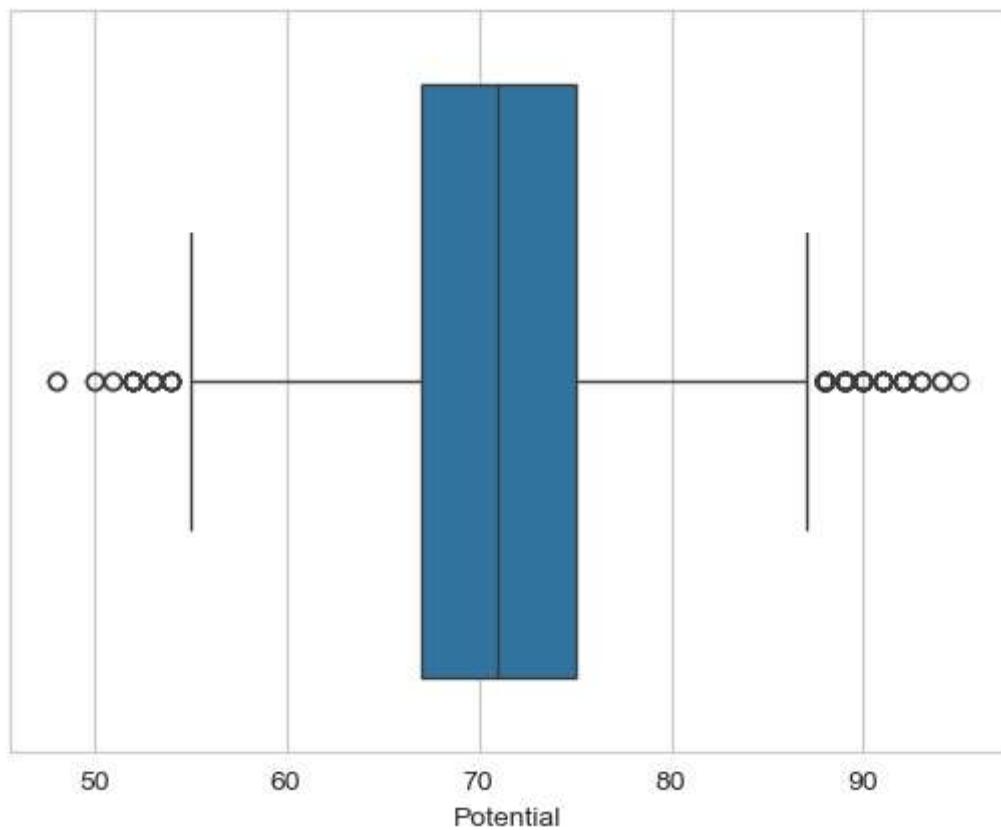
```
In [71]: sns.stripplot(x="International Reputation", y="Potential", hue="Preferred Foot",
                      data=data, palette="Set2", size=20, marker="D",
                      edgecolor="gray", alpha=.25)
```

```
Out[71]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



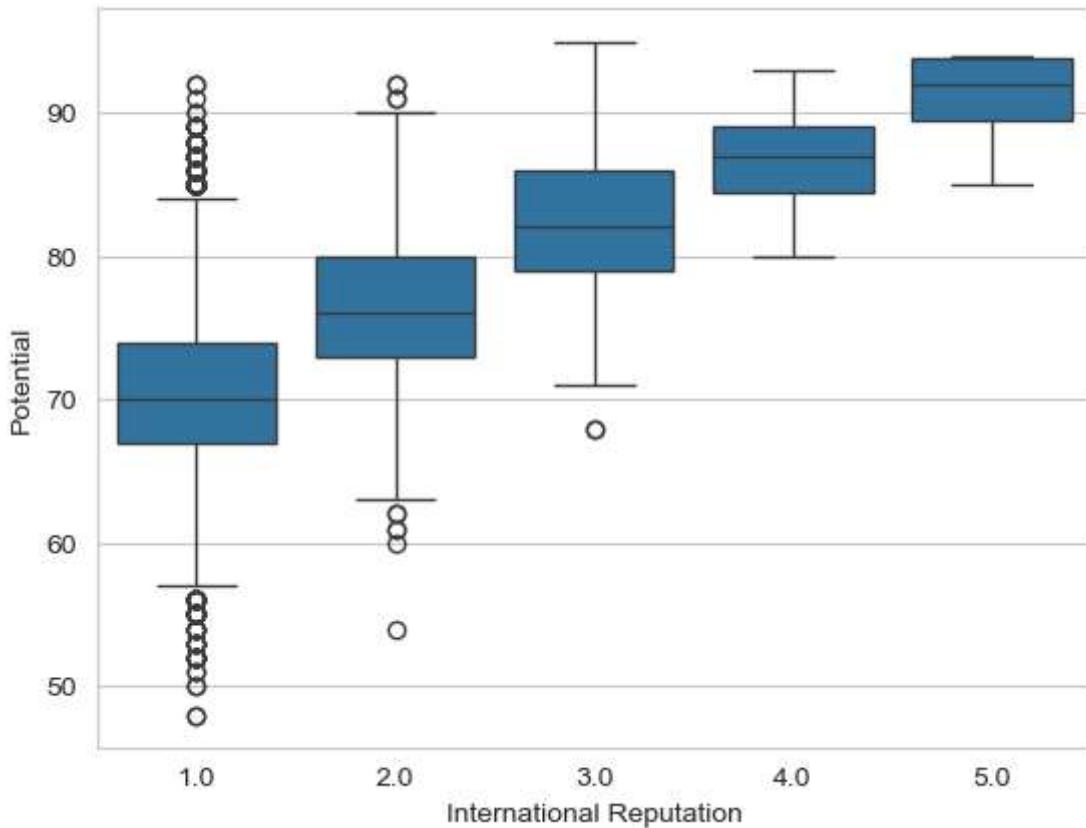
```
In [75]: sns.boxplot(x=data["Potential"])
```

```
Out[75]: <Axes: xlabel='Potential'>
```



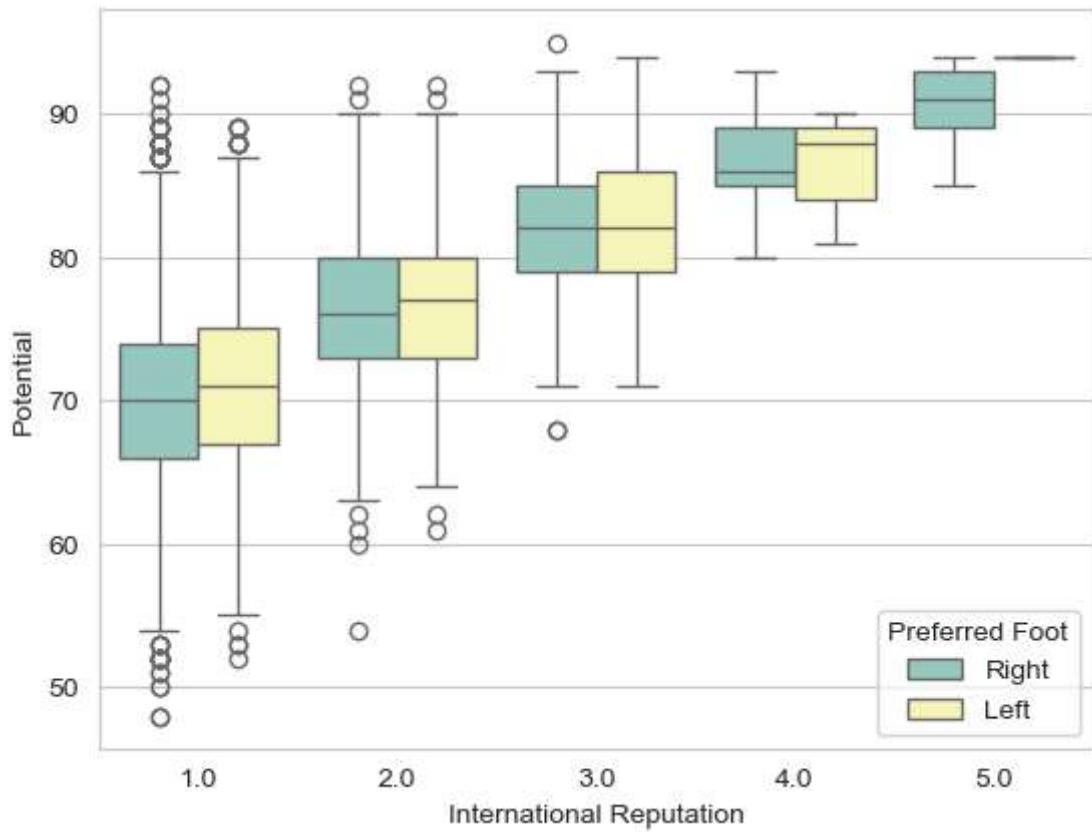
```
In [79]: sns.boxplot(x="International Reputation", y="Potential", data=data)
```

```
Out[79]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



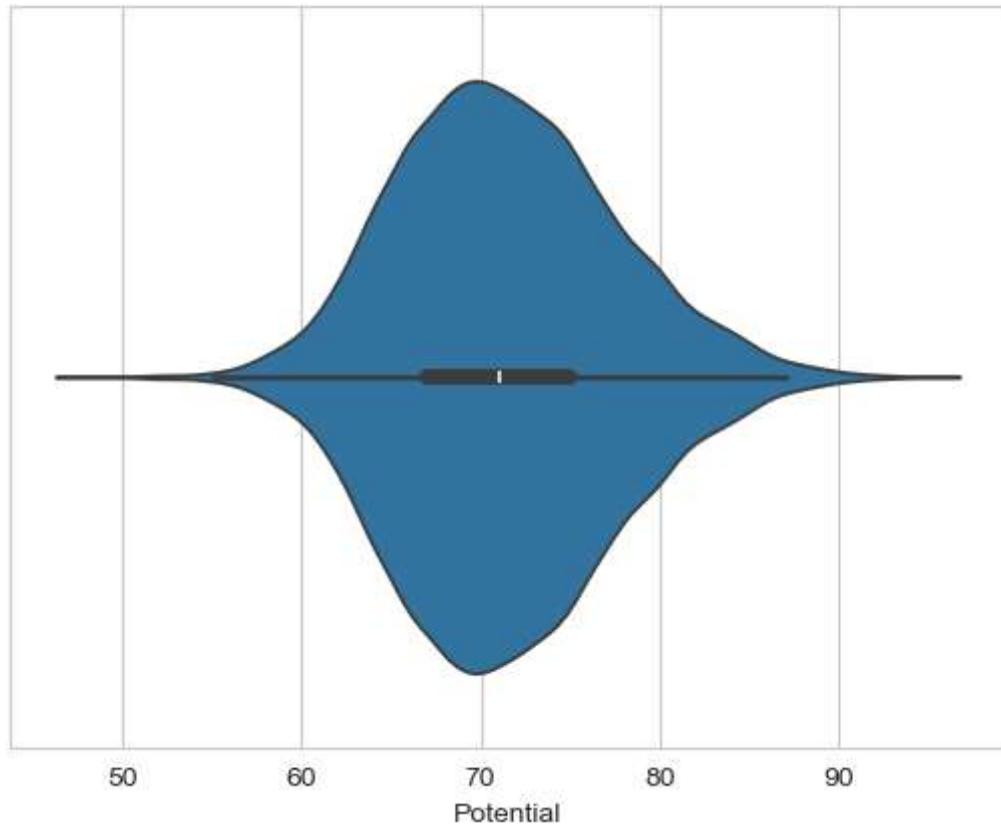
```
In [83]: sns.boxplot(x="International Reputation", y="Potential", hue="Preferred Foot", data
```

```
Out[83]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



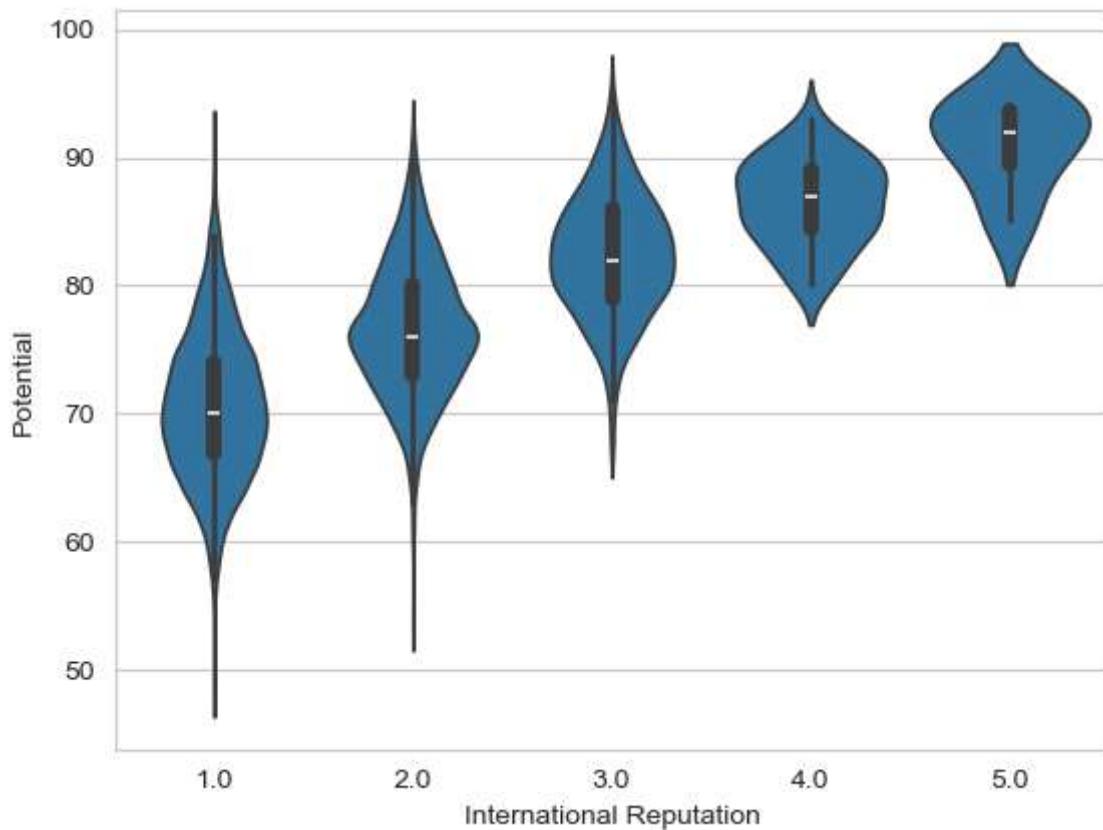
```
In [87]: sns.violinplot(x=data["Potential"])
```

```
Out[87]: <Axes: xlabel='Potential'>
```



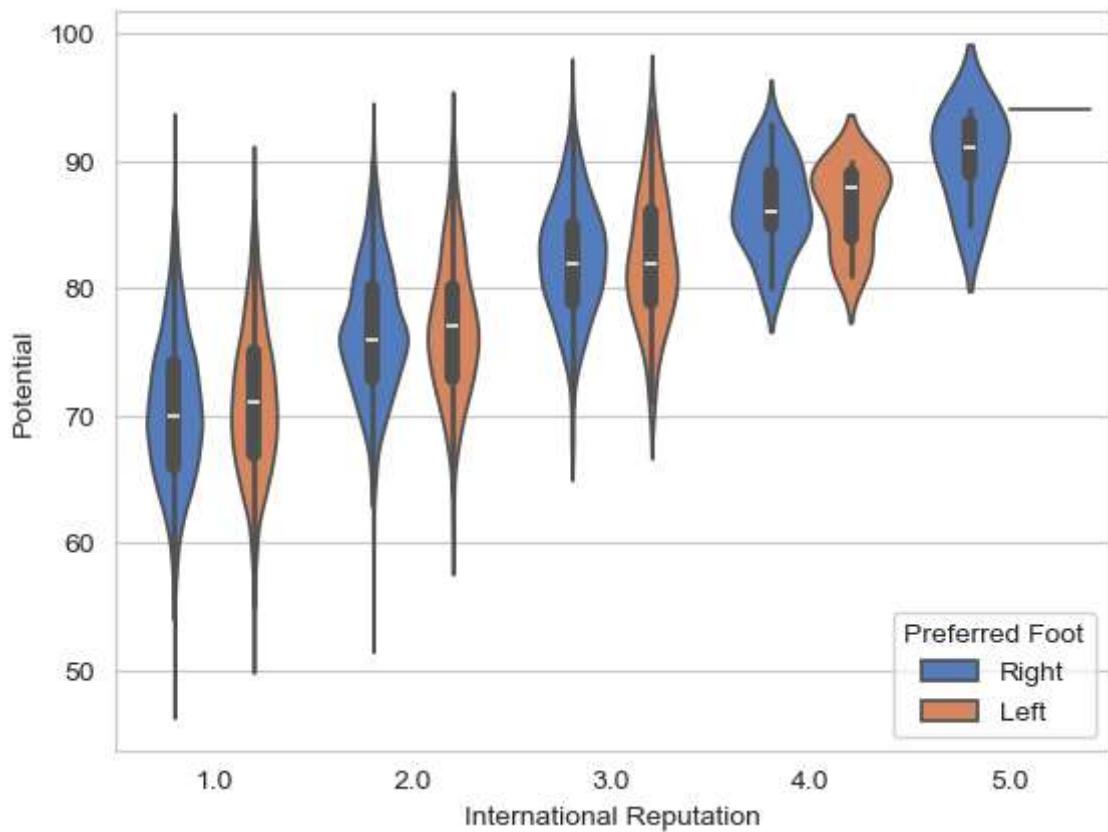
```
In [93]: sns.violinplot(x="International Reputation", y="Potential", data=data)
```

```
Out[93]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



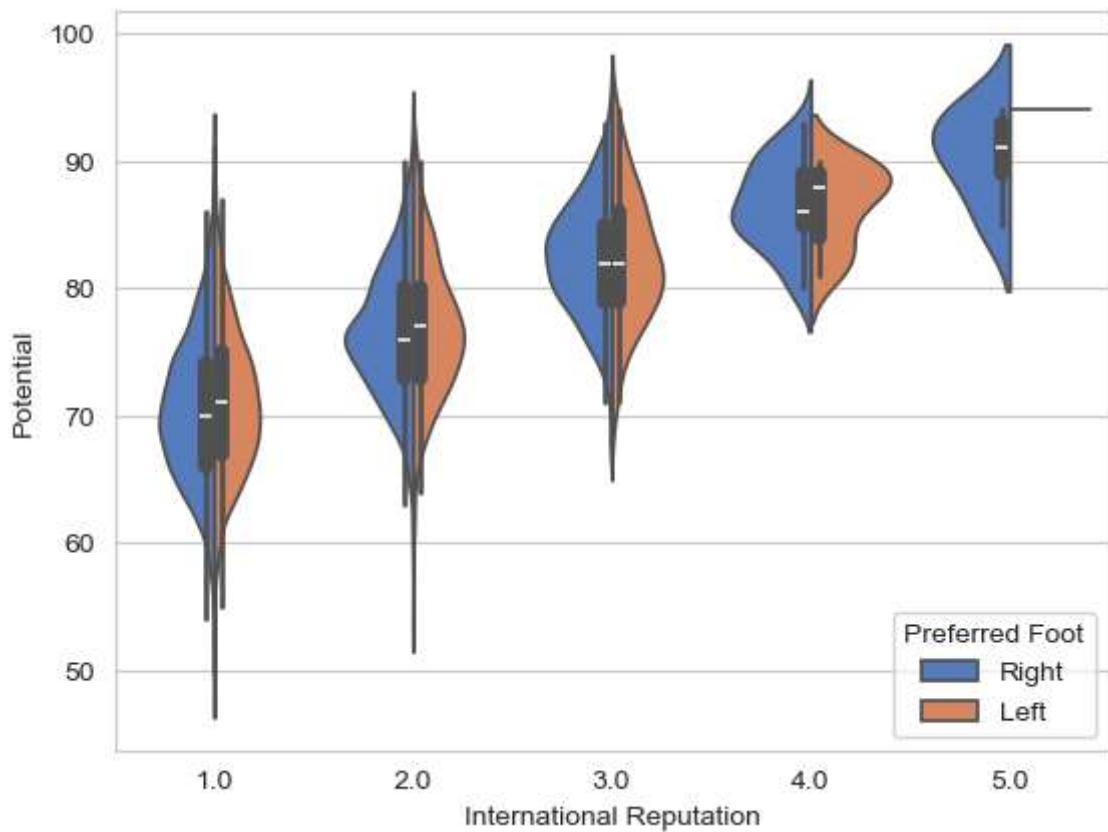
```
In [97]: sns.violinplot(x="International Reputation", y="Potential", hue="Preferred Foot", d
```

```
Out[97]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



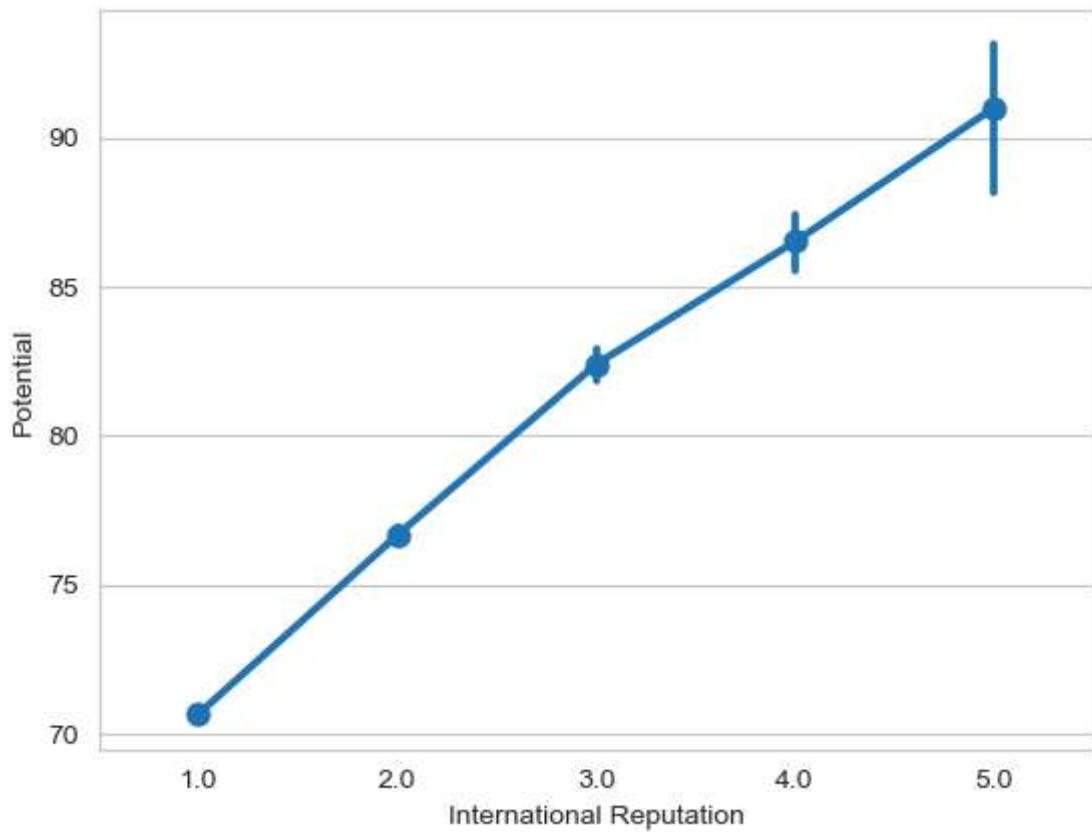
```
In [99]: sns.violinplot(x="International Reputation", y="Potential", hue="Preferred Foot",
                      data=data, palette="muted", split=True)
```

```
Out[99]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



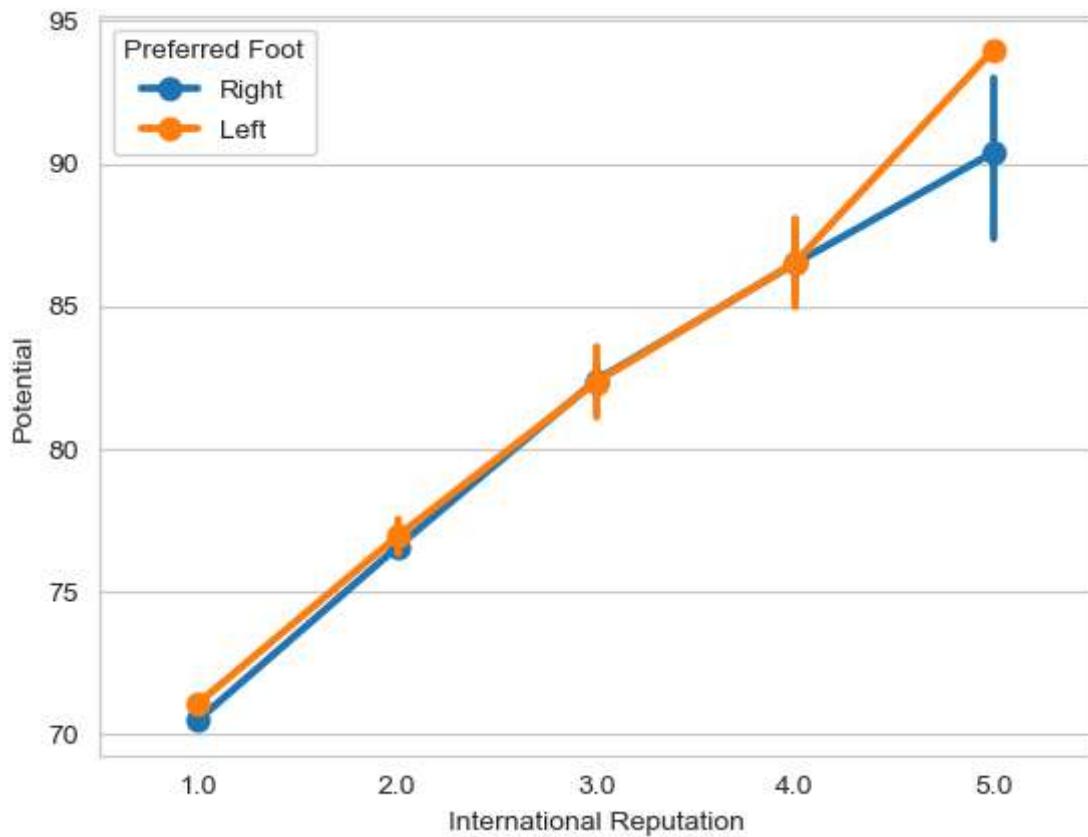
```
In [103]: sns.pointplot(x="International Reputation", y="Potential", data=data)
```

```
Out[103]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



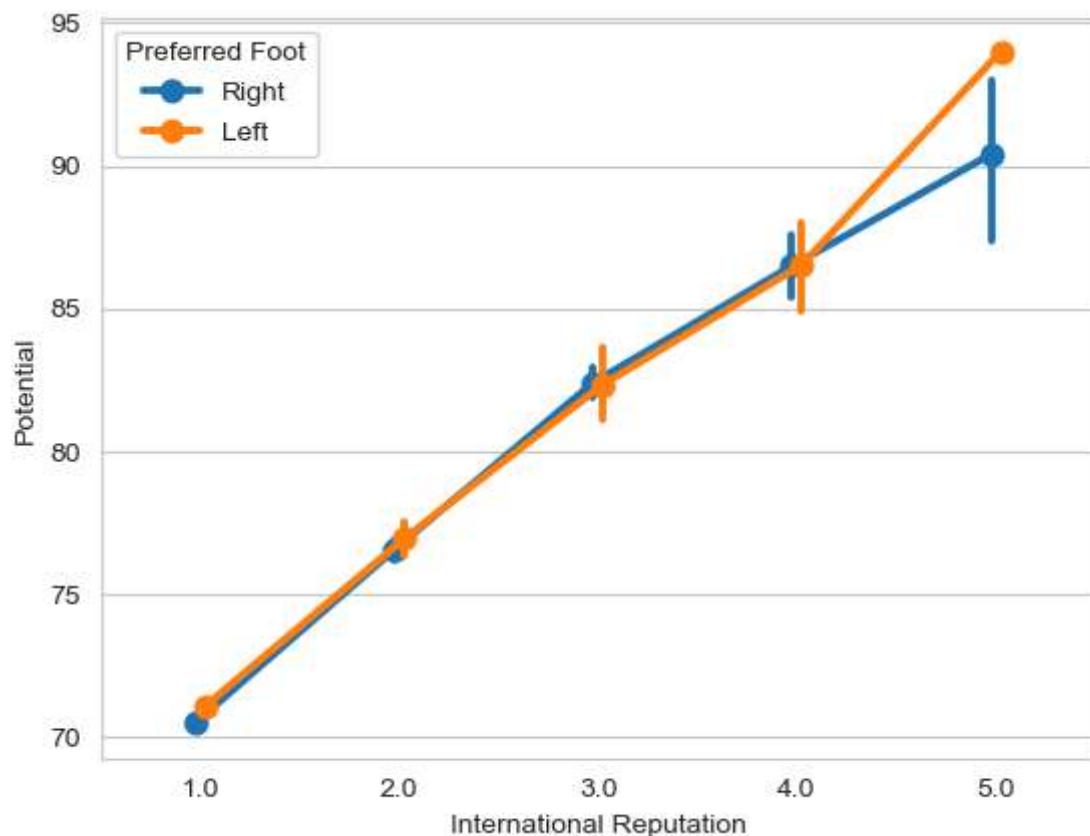
```
In [107... sns.pointplot(x="International Reputation", y="Potential", hue="Preferred Foot", da
```

```
Out[107... <Axes: xlabel='International Reputation', ylabel='Potential'>
```



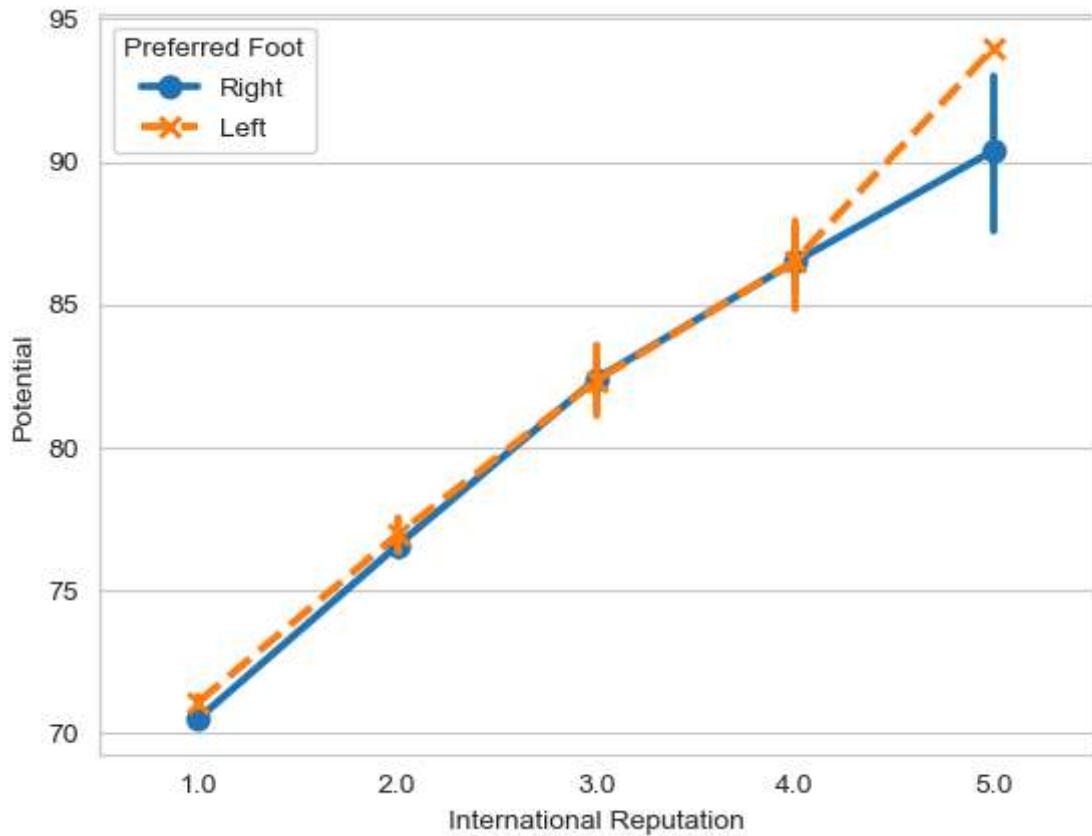
```
In [111... sns.pointplot(x="International Reputation", y="Potential", hue="Preferred Foot", da
```

```
Out[111... <Axes: xlabel='International Reputation', ylabel='Potential'>
```



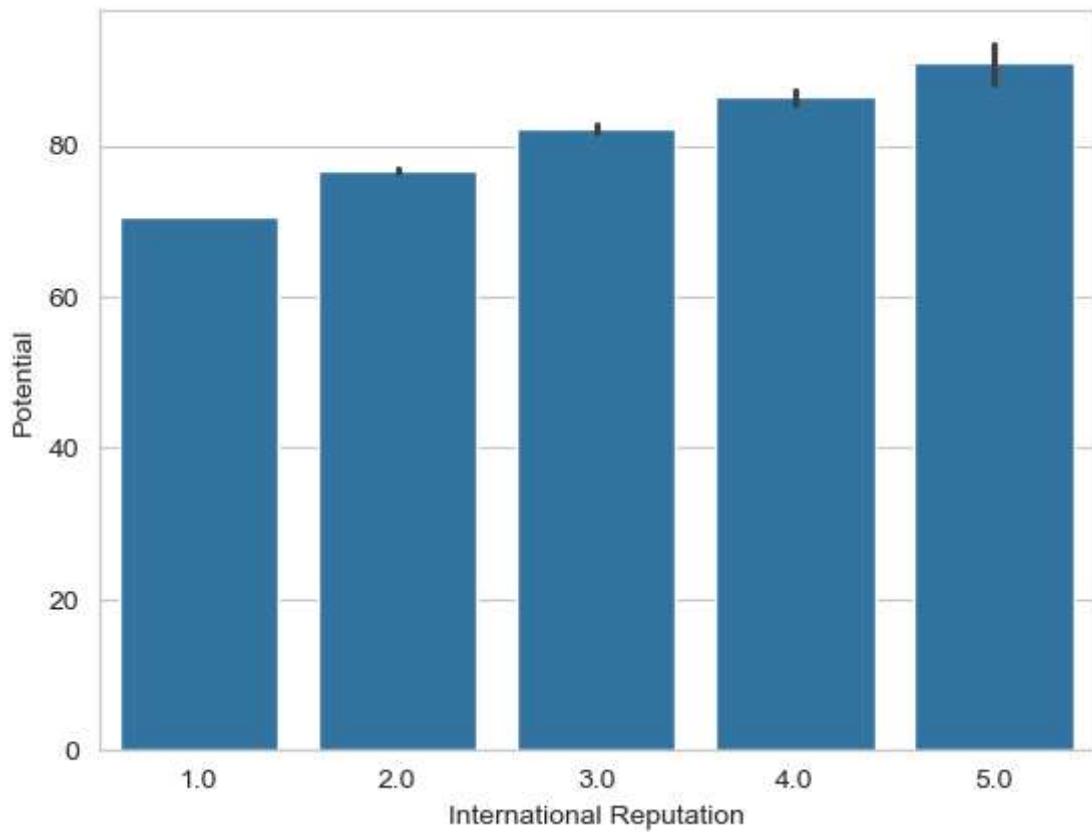
```
In [115...]: sns.pointplot(x="International Reputation", y="Potential", hue="Preferred Foot",
                     data=data, markers=["o", "x"], linestyles=[ "-", "--"])
```

```
Out[115...]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



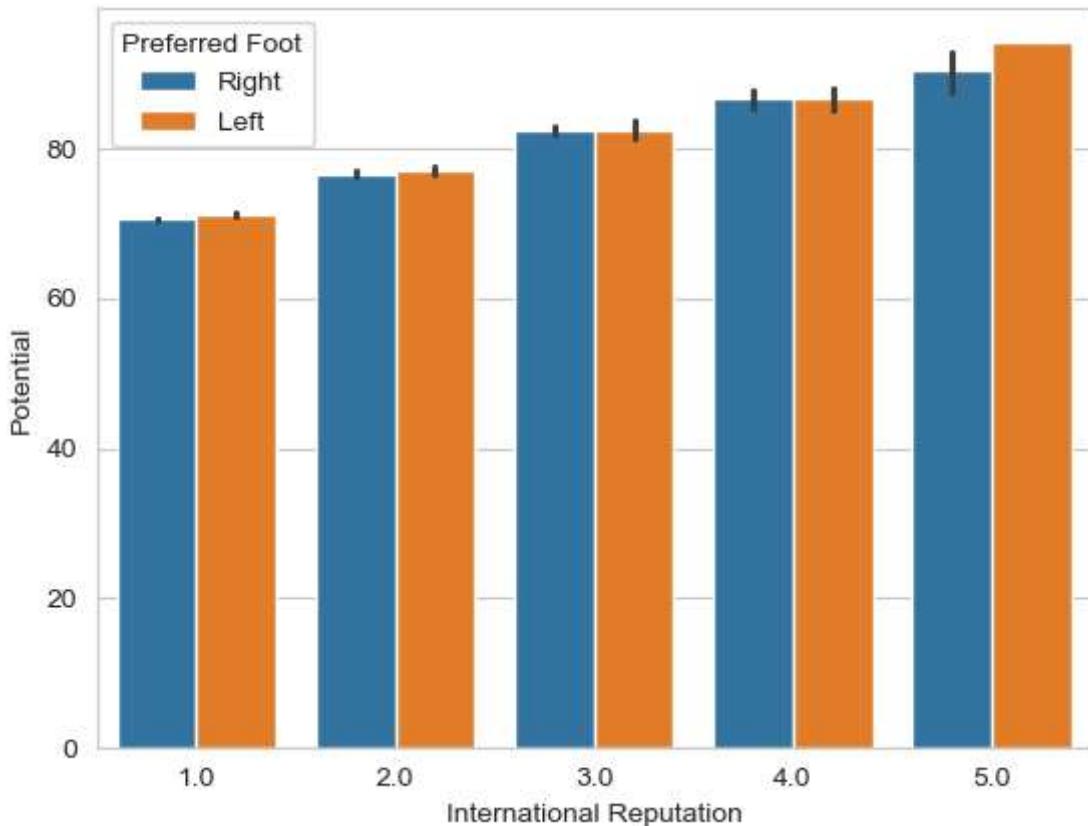
```
In [117]: sns.barplot(x="International Reputation", y="Potential", data=data)
```

```
Out[117]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```



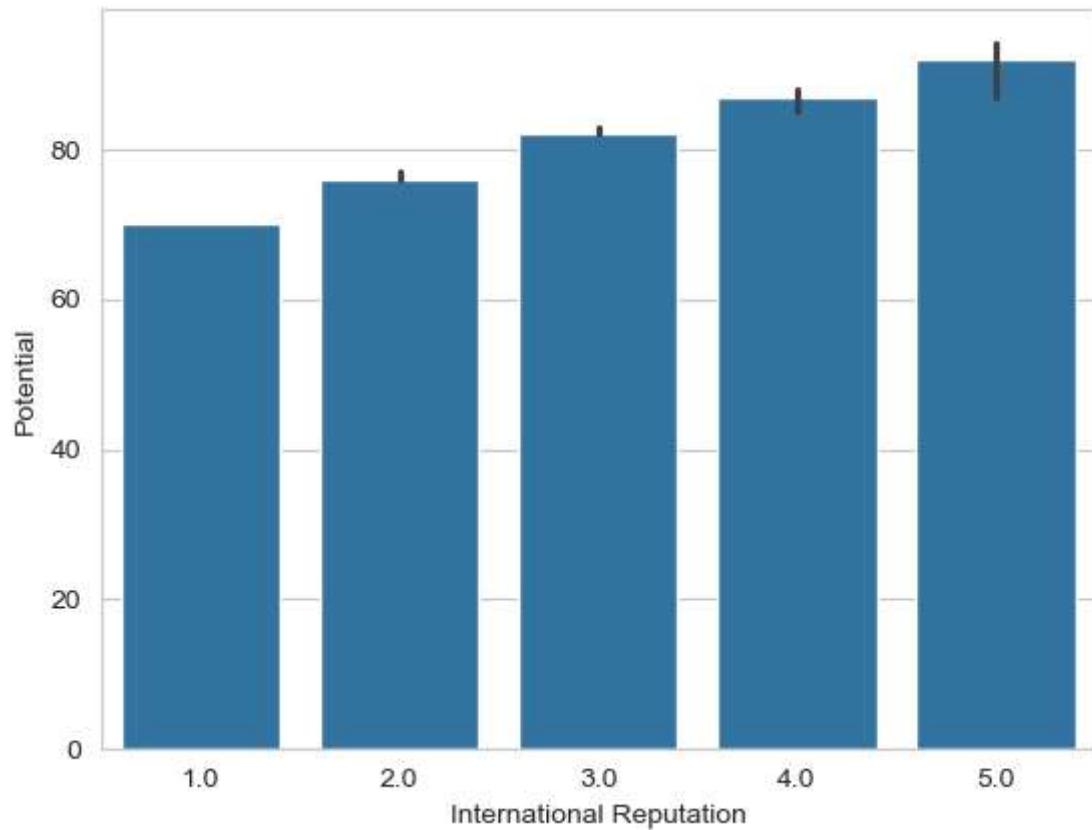
```
In [123... sns.barplot(x="International Reputation", y="Potential", hue="Preferred Foot", data
```

```
Out[123... <Axes: xlabel='International Reputation', ylabel='Potential'>
```



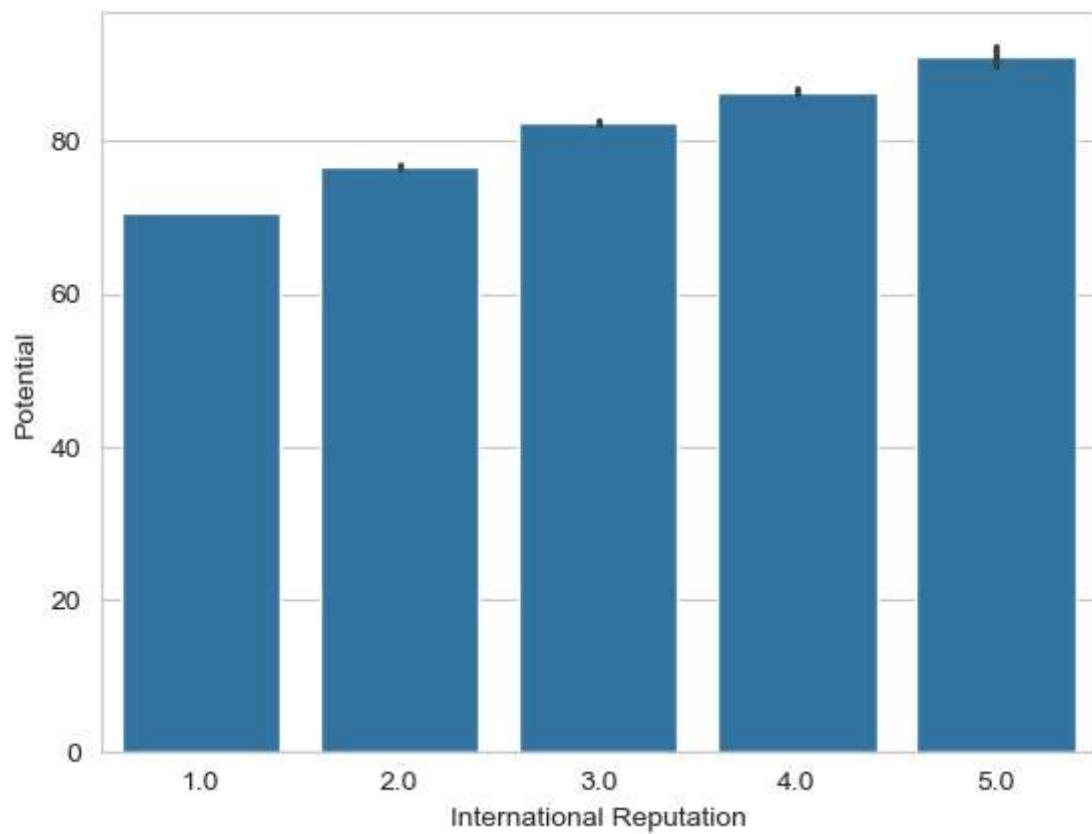
```
In [131... from numpy import median  
sns.barplot(x="International Reputation", y="Potential", data=data, estimator=median
```

```
Out[131... <Axes: xlabel='International Reputation', ylabel='Potential'>
```



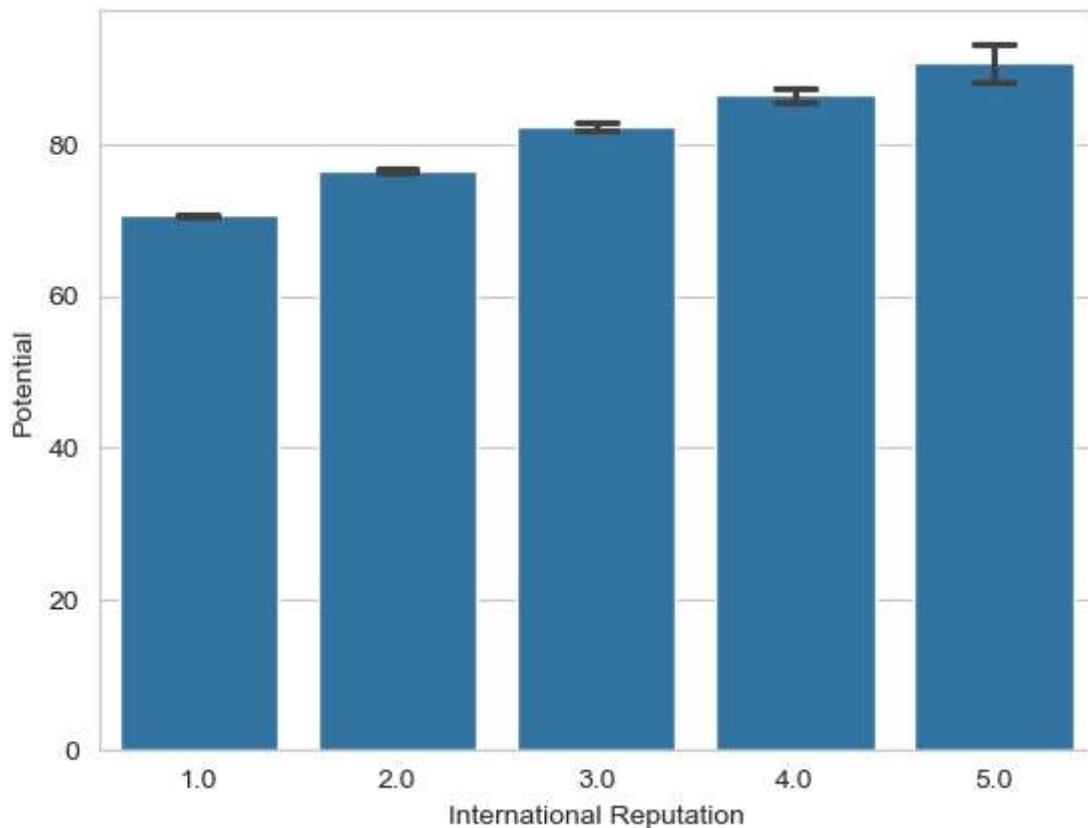
```
In [135...]: sns.barplot(x="International Reputation", y="Potential", data=data, ci=68)
```

```
Out[135...]: <Axes: xlabel='International Reputation', ylabel='Potential'>
```

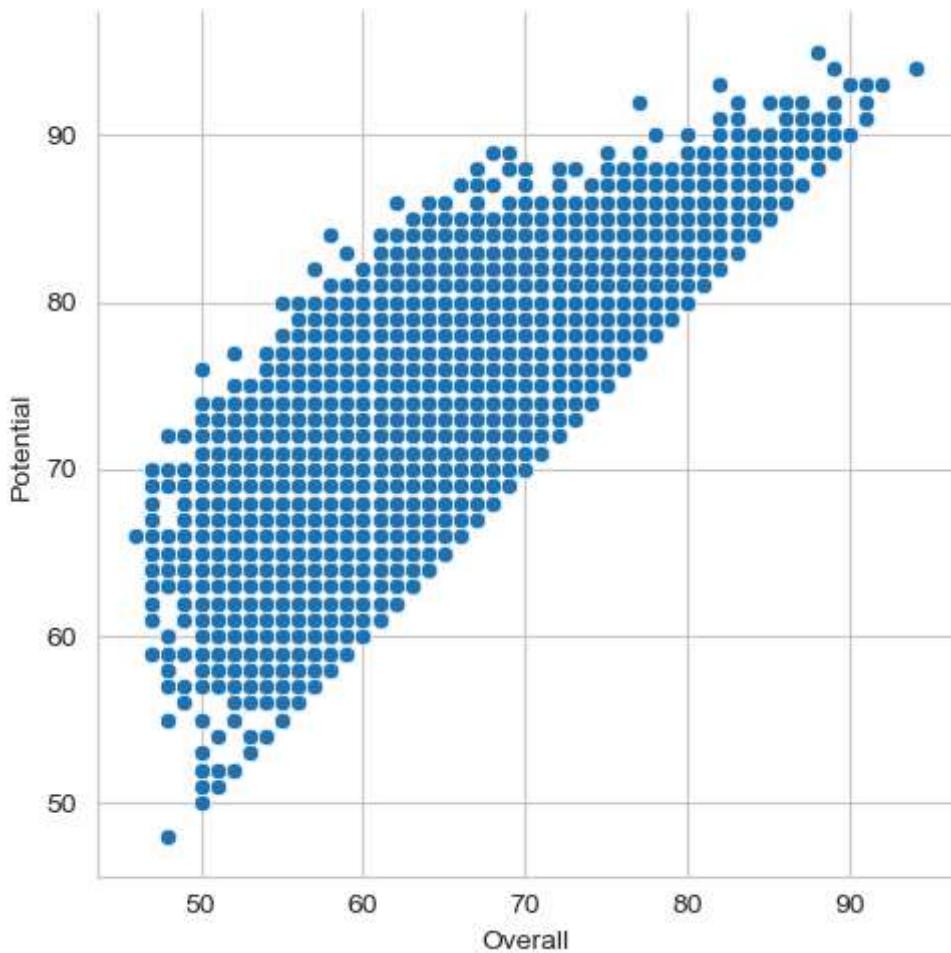


```
In [139... sns.barplot(x="International Reputation", y="Potential", data=data, capsize=0.2)
```

```
Out[139... <Axes: xlabel='International Reputation', ylabel='Potential'>
```

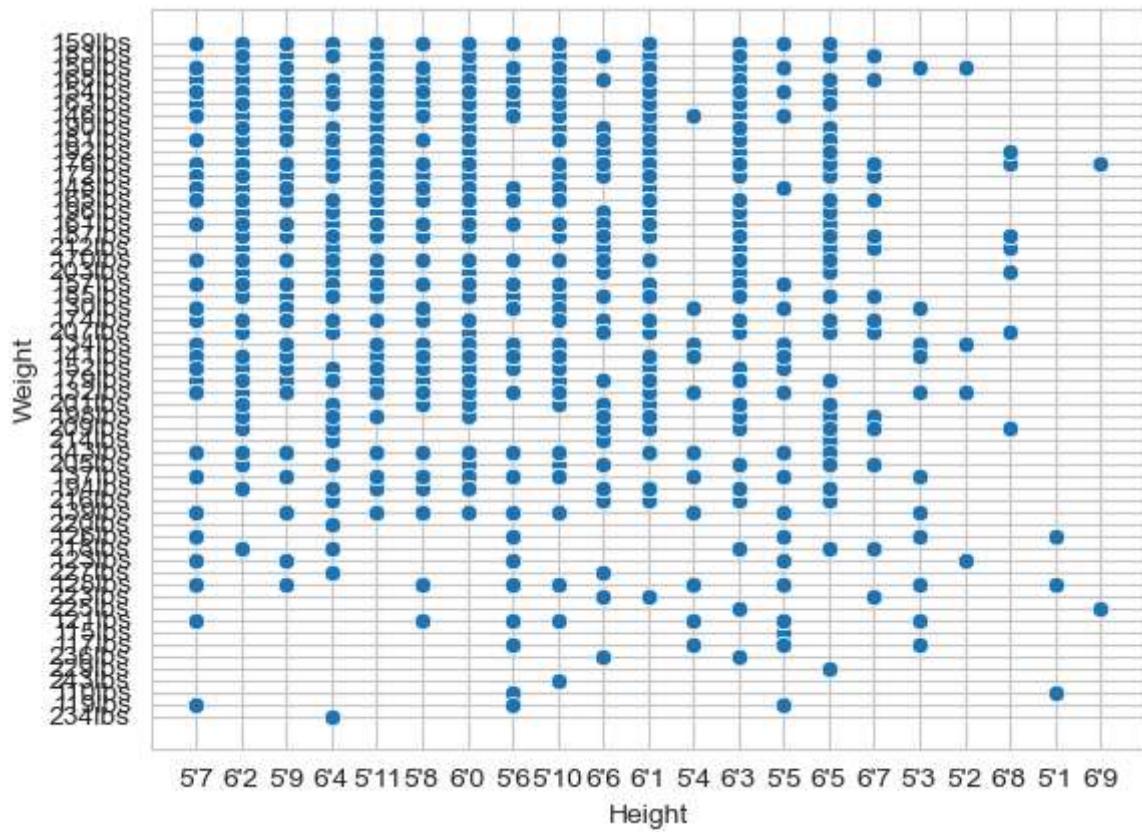


```
In [143... g = sns.relplot(x="Overall", y="Potential", data=data)
```

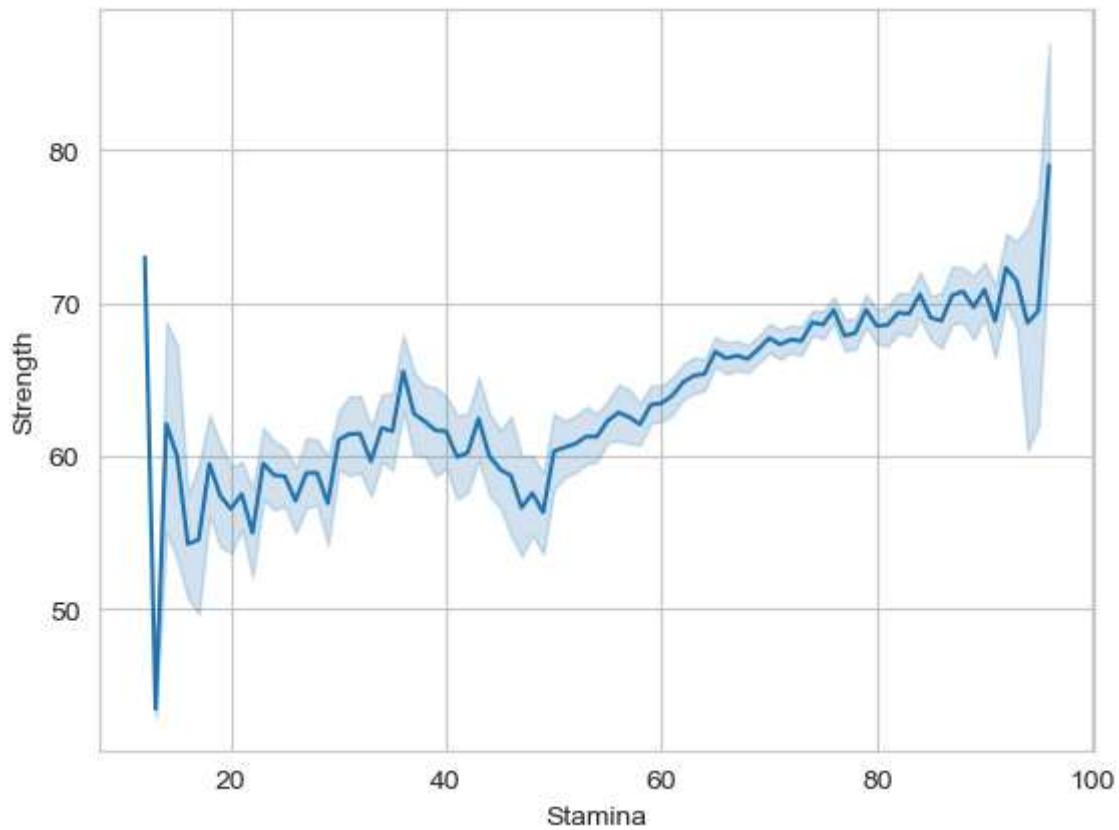


```
In [147]: sns.scatterplot(x="Height", y="Weight", data=data)
```

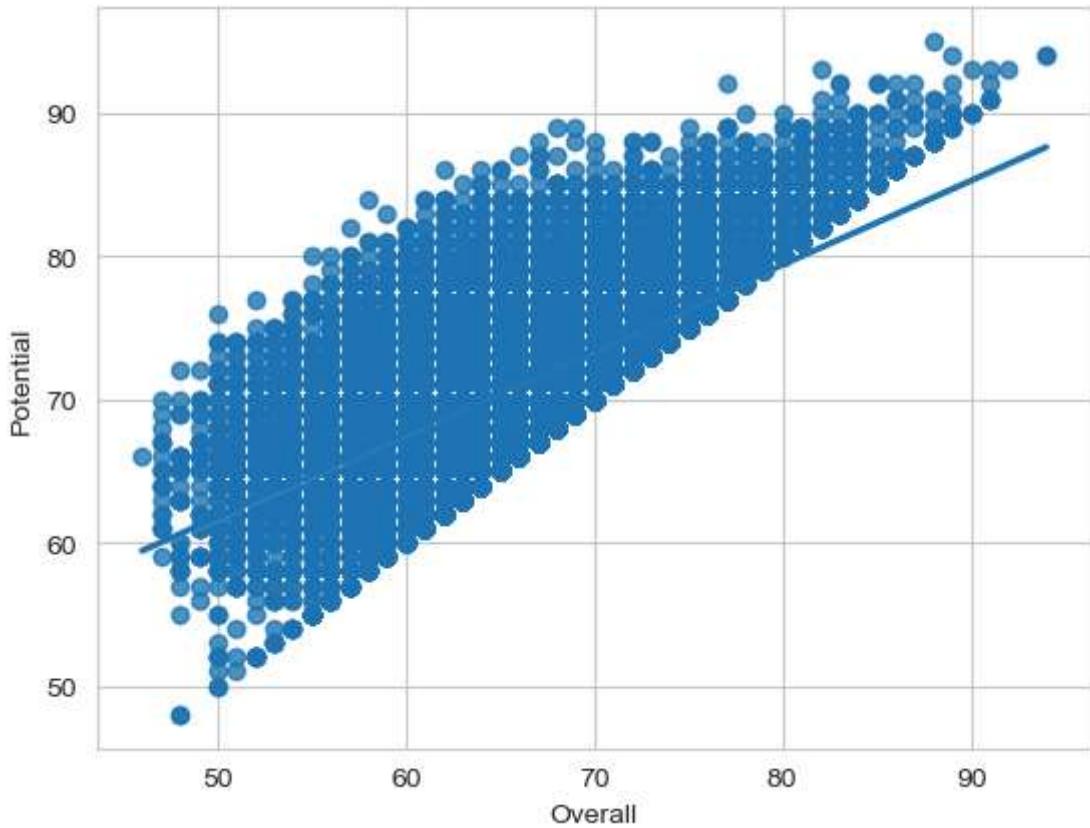
```
Out[147]: <Axes: xlabel='Height', ylabel='Weight'>
```



```
In [151]: ax = sns.lineplot(x="Stamina", y="Strength", data=data)
```

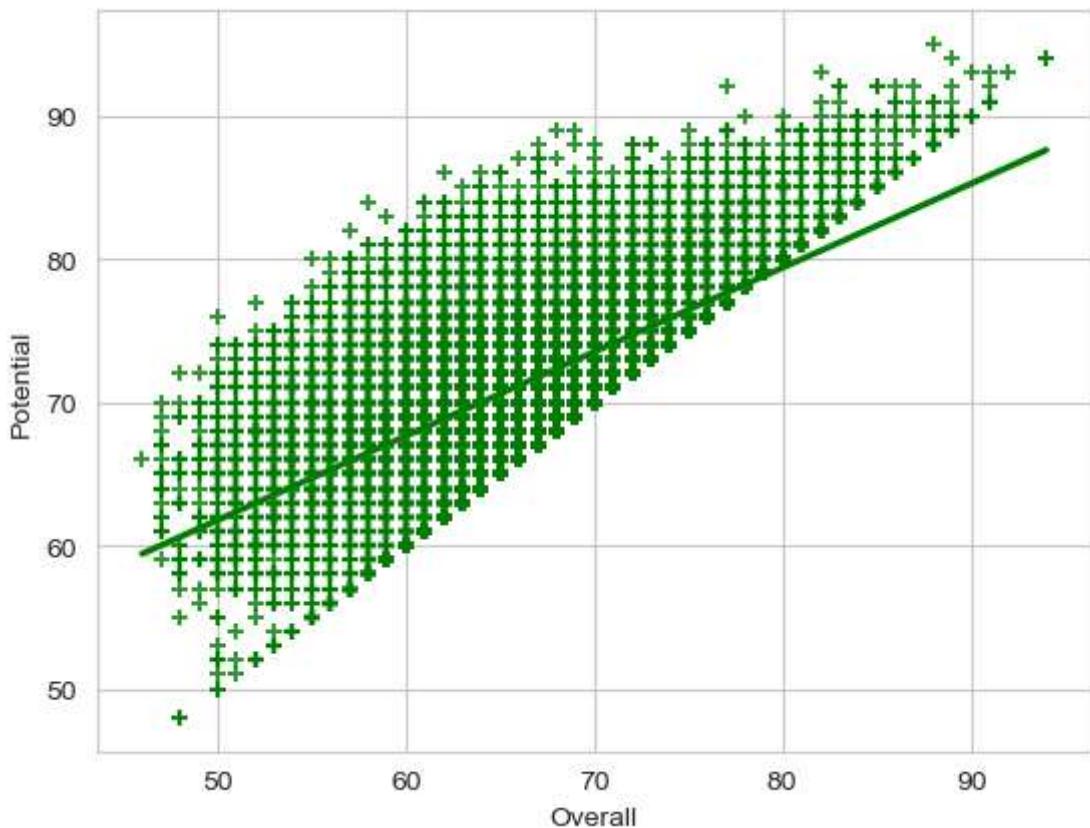


```
In [155]: ax = sns.regplot(x="Overall", y="Potential", data=data)
```



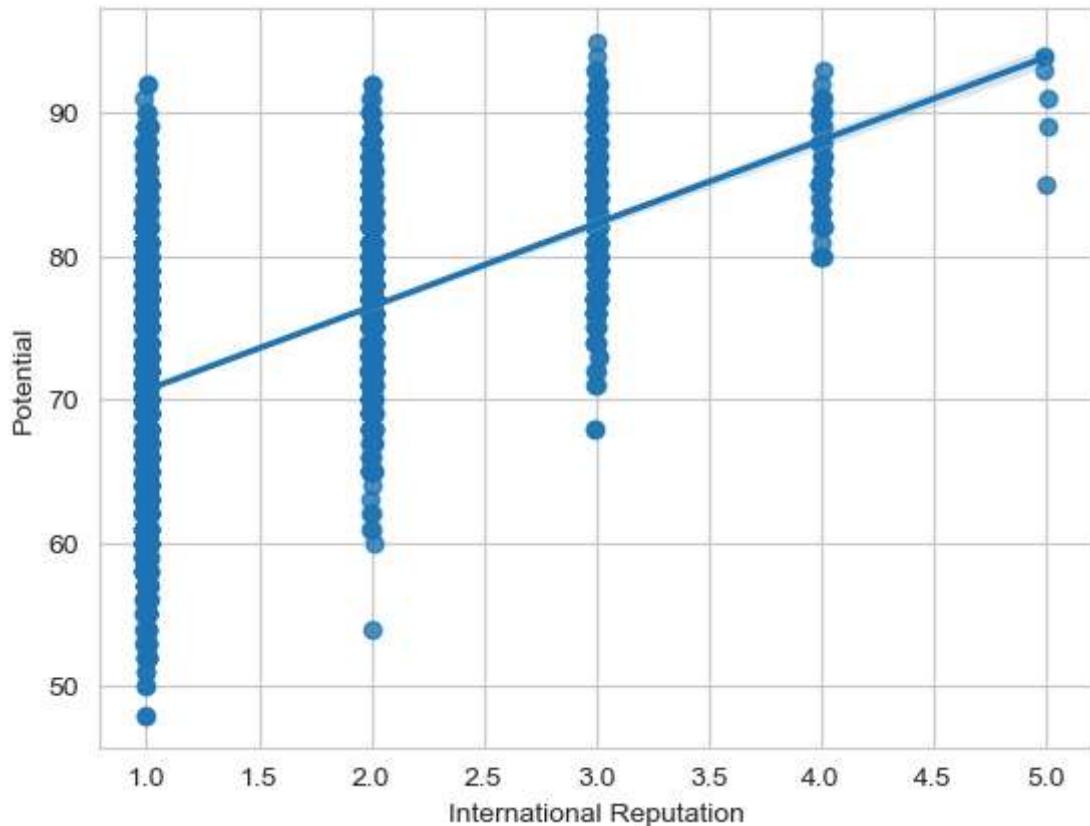
```
In [159]: sns.regplot(x="Overall", y="Potential", data=data, color= "g", marker="+")
```

```
Out[159]: <Axes: xlabel='Overall', ylabel='Potential'>
```



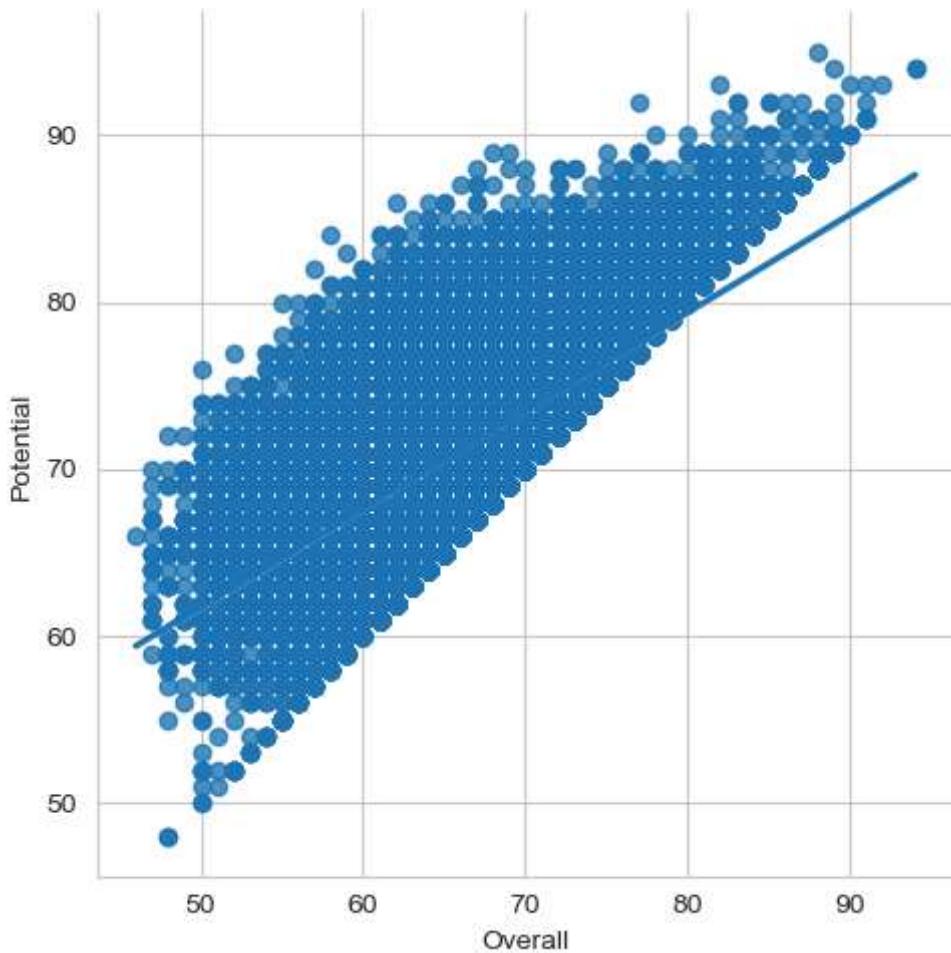
```
In [163... sns.regplot(x="International Reputation", y="Potential", data=data, x_jitter=.01)
```

```
Out[163... <Axes: xlabel='International Reputation', ylabel='Potential'>
```



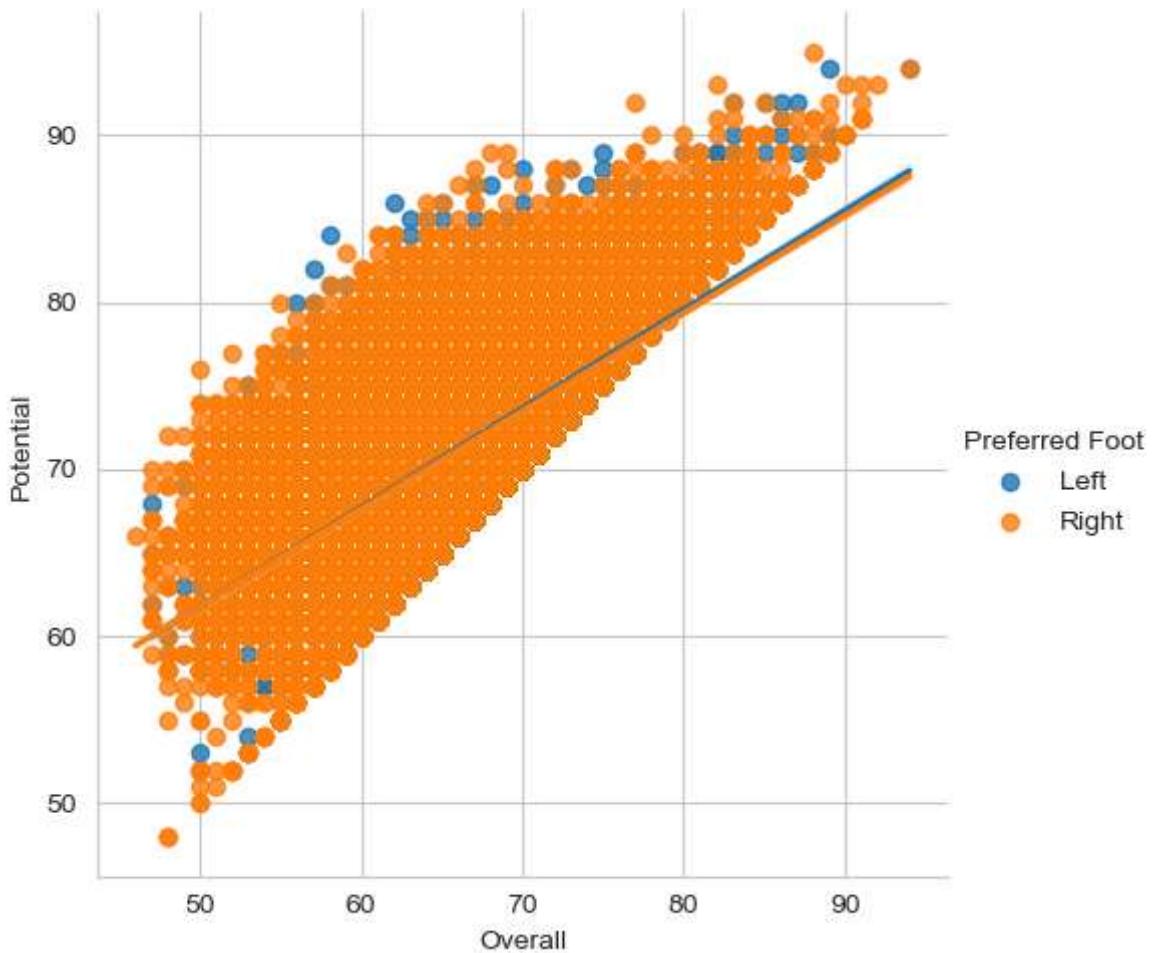
```
In [167... sns.lmplot(x="Overall", y="Potential", data=data)
```

```
Out[167... <seaborn.axisgrid.FacetGrid at 0x253a06d91c0>
```



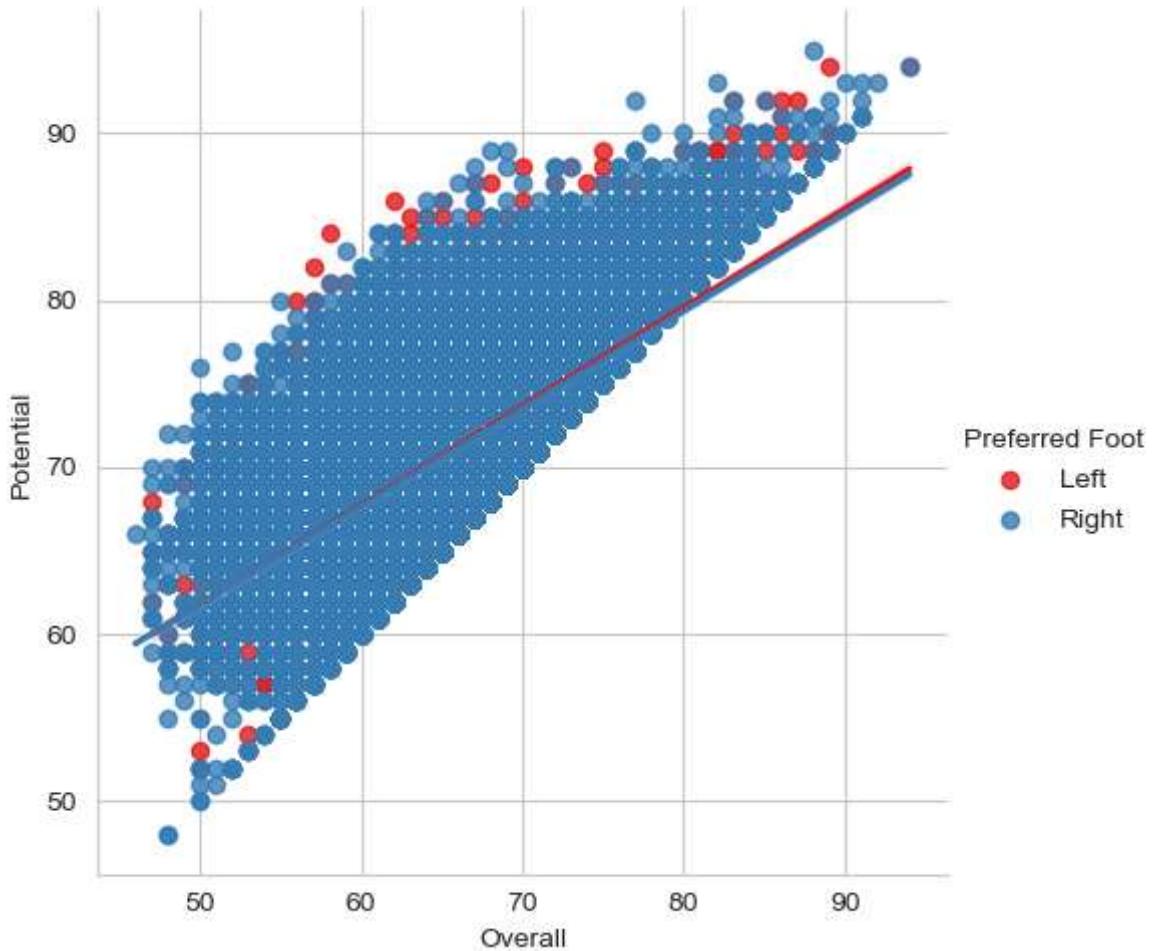
```
In [171]: sns.lmplot(x="Overall", y="Potential", hue="Preferred Foot", data=data)
```

```
Out[171]: <seaborn.axisgrid.FacetGrid at 0x253a0b477a0>
```



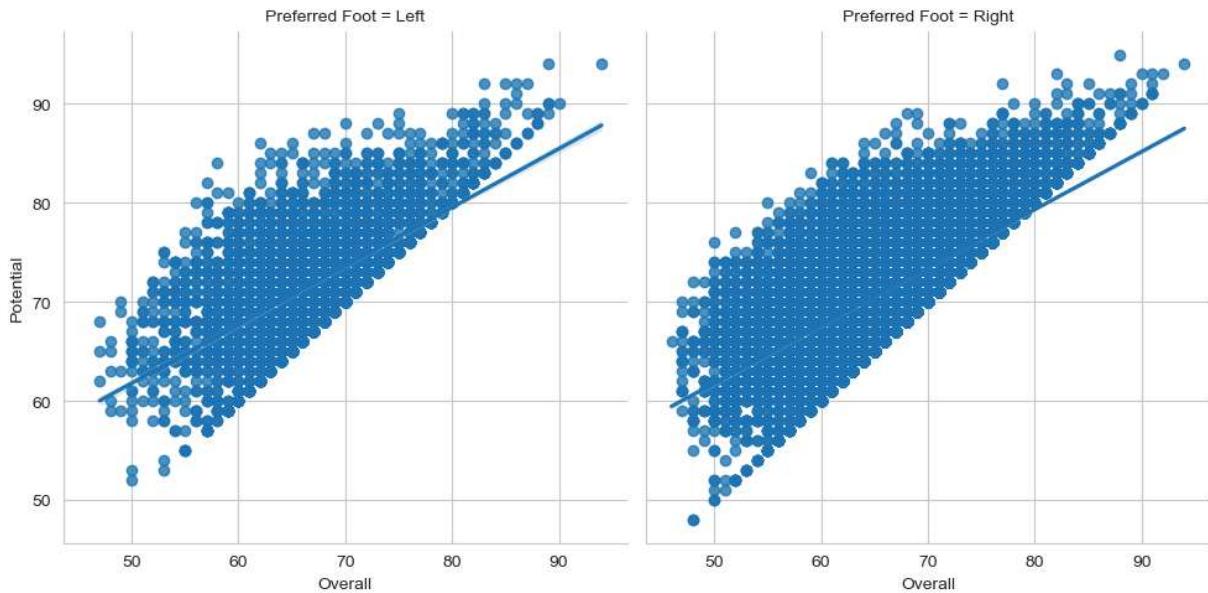
```
In [175]: sns.lmplot(x="Overall", y="Potential", hue="Preferred Foot", data=data, palette="S
```

```
Out[175]: <seaborn.axisgrid.FacetGrid at 0x253a09614f0>
```

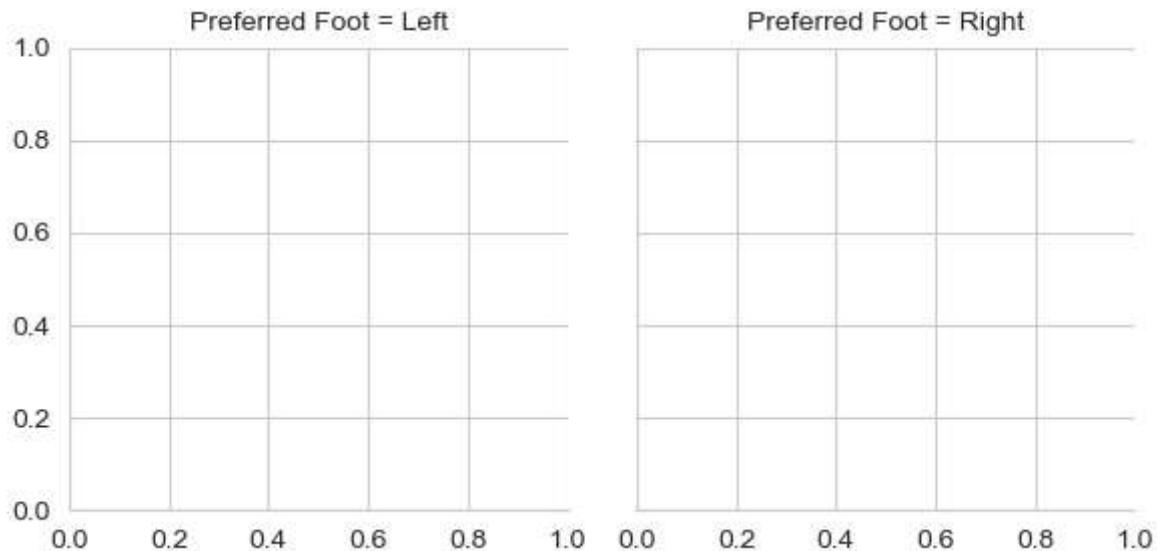


```
In [179]: sns.lmplot(x="Overall", y="Potential", col="Preferred Foot", data=data)
```

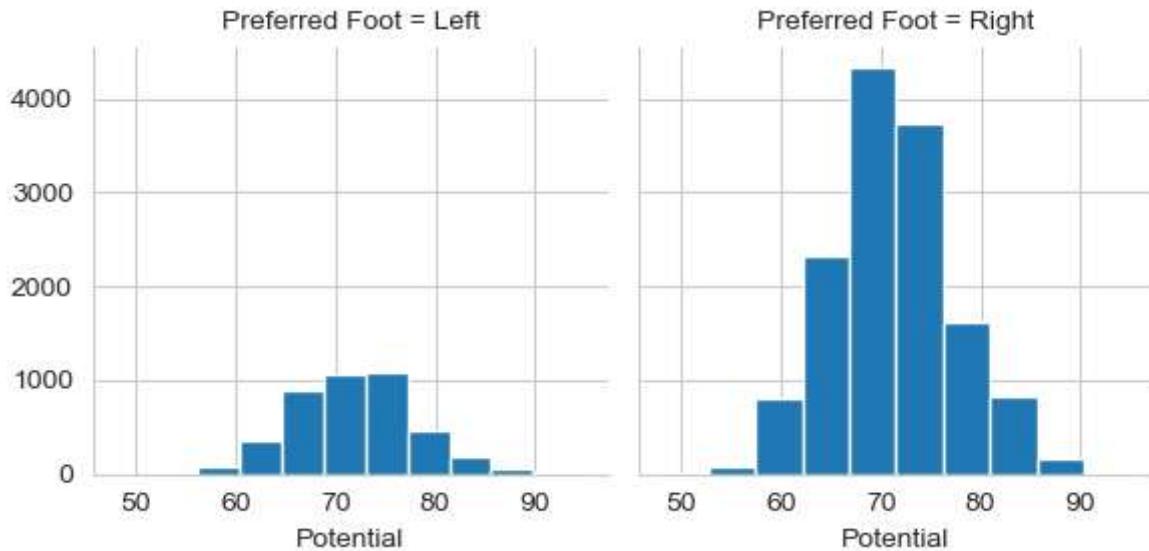
```
Out[179]: <seaborn.axisgrid.FacetGrid at 0x2539ef4d5e0>
```



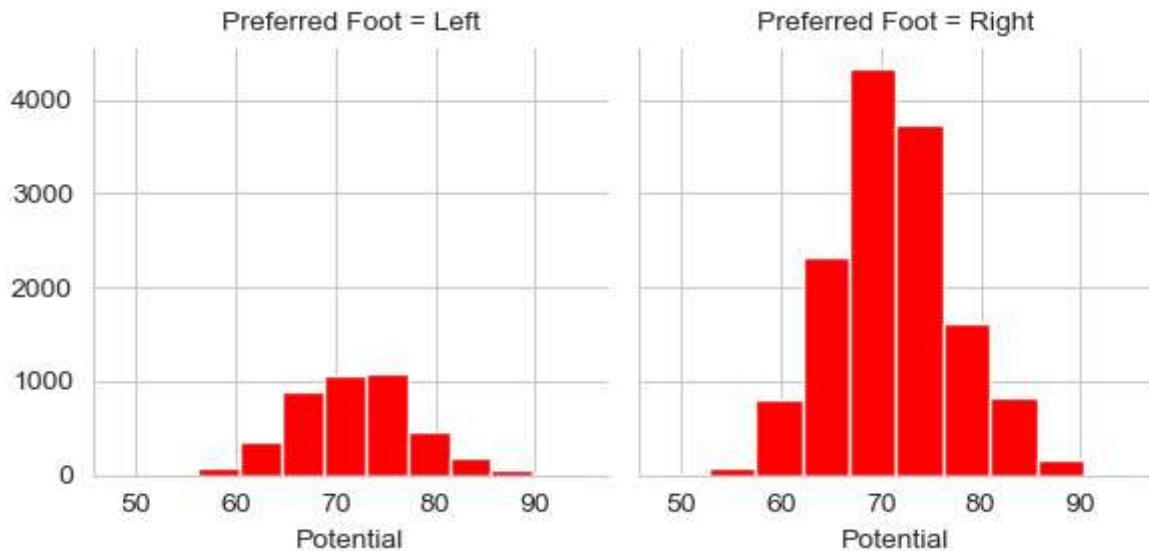
```
In [183]: g = sns.FacetGrid(data, col="Preferred Foot")
```



```
In [187...  
g = sns.FacetGrid(data, col="Preferred Foot")  
g = g.map(plt.hist, "Potential")
```

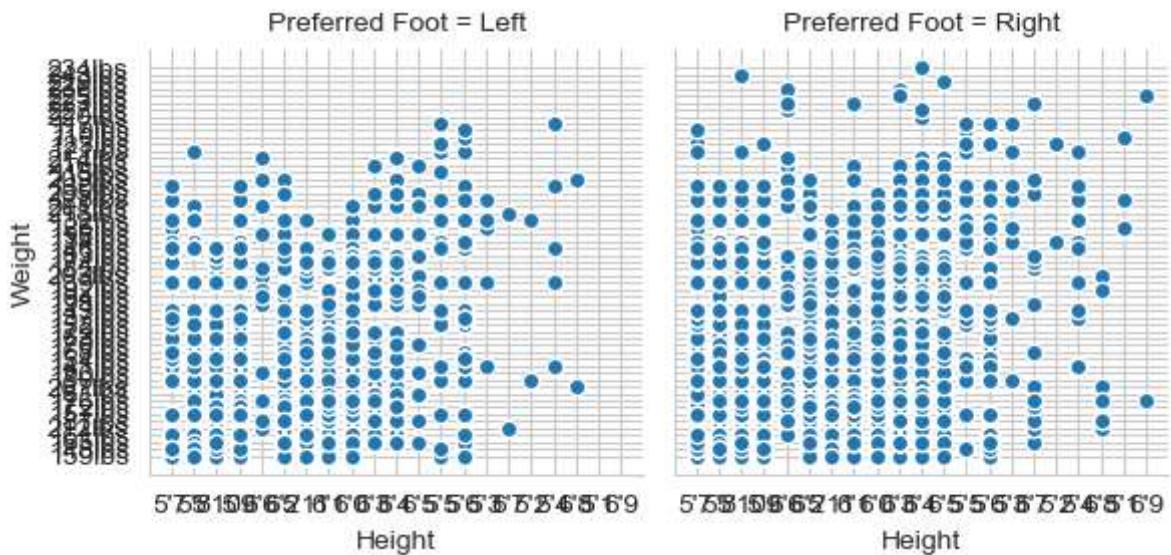


```
In [191...  
g = sns.FacetGrid(data, col="Preferred Foot")  
g = g.map(plt.hist, "Potential", bins=10, color="r")
```



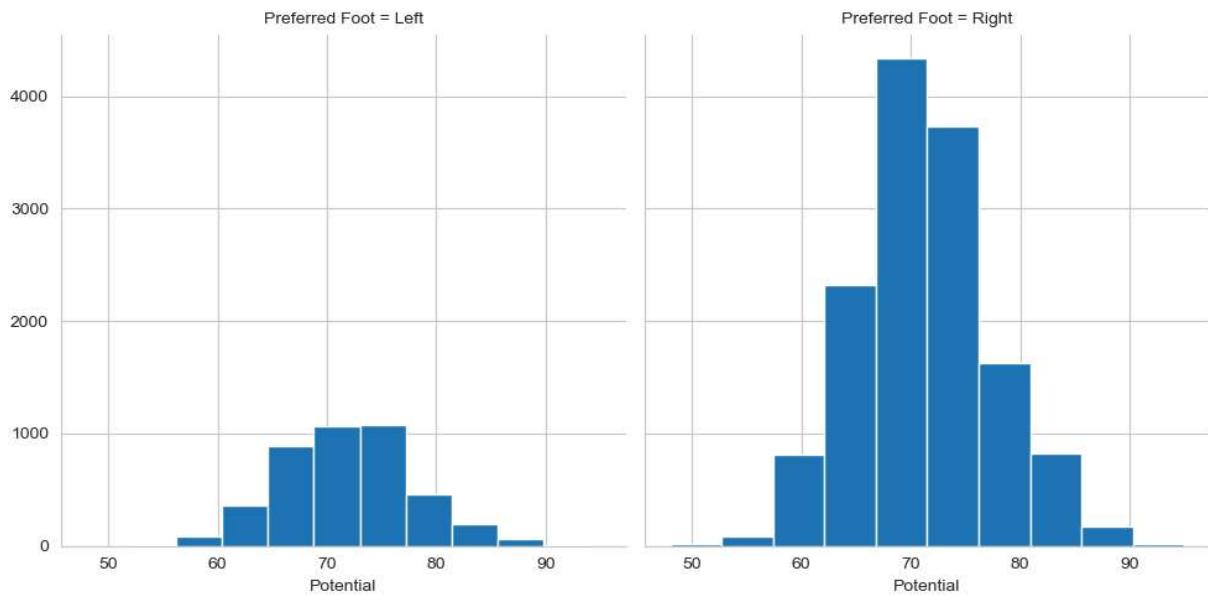
In [193...]

```
g = sns.FacetGrid(data, col="Preferred Foot")
g = (g.map(plt.scatter, "Height", "Weight", edgecolor="w").add_legend())
```



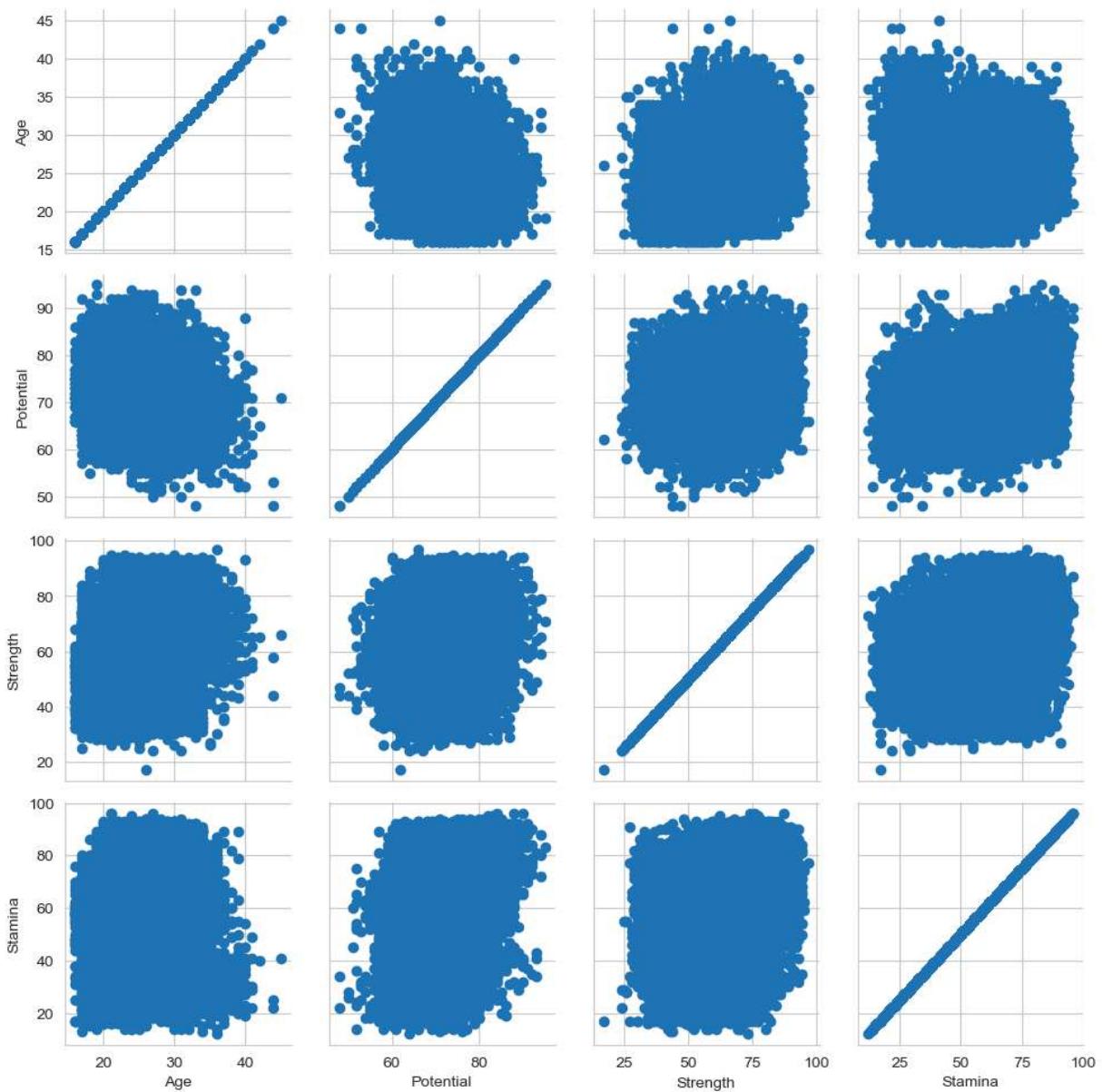
In [197...]

```
g = sns.FacetGrid(data, col="Preferred Foot", height=5, aspect=1)
g = g.map(plt.hist, "Potential")
```



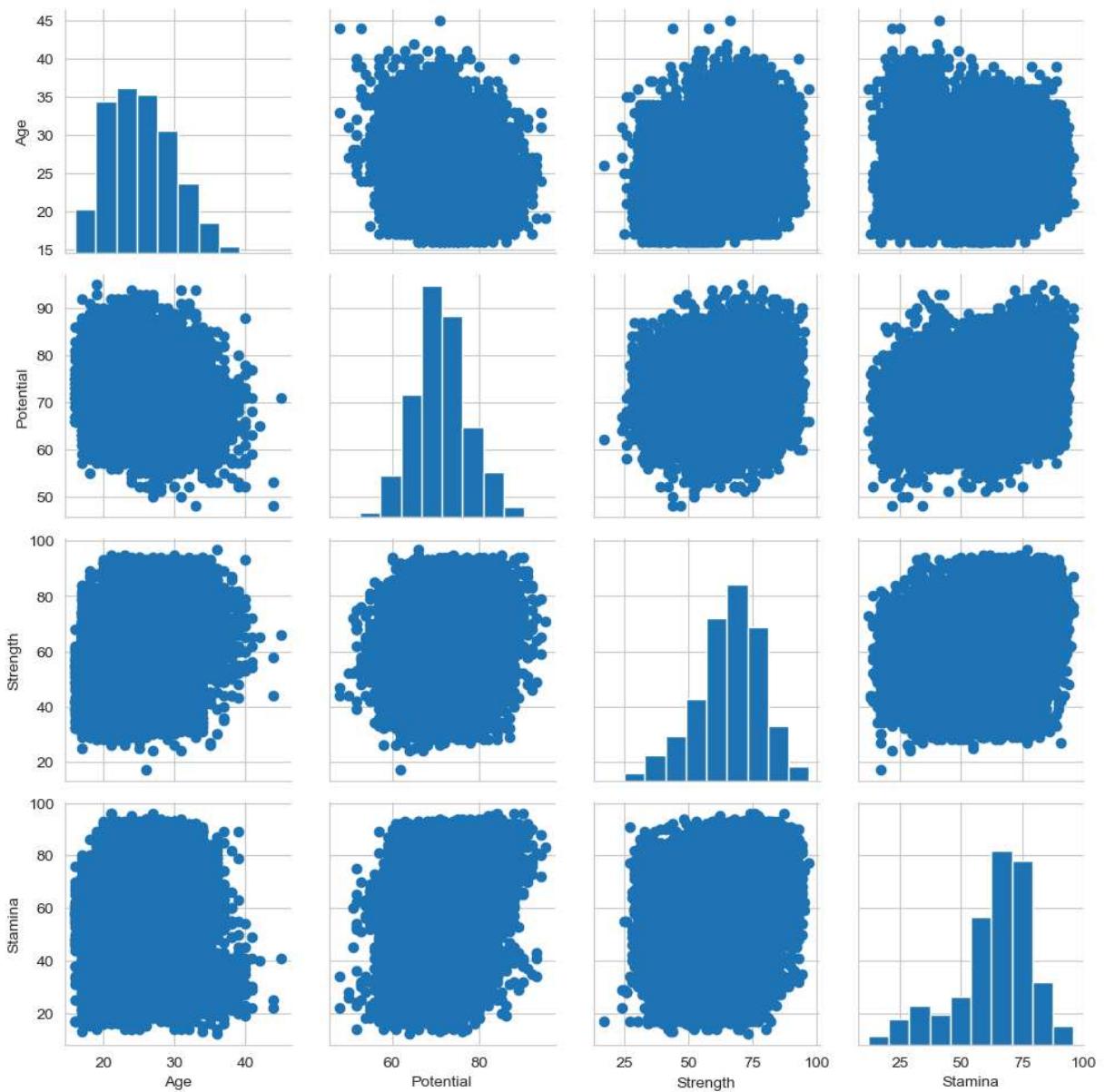
```
In [201...]: data_new = data[['Age', 'Potential', 'Strength', 'Stamina', 'Preferred Foot']]
```

```
In [203...]: g = sns.PairGrid(data_new)
g = g.map(plt.scatter)
```



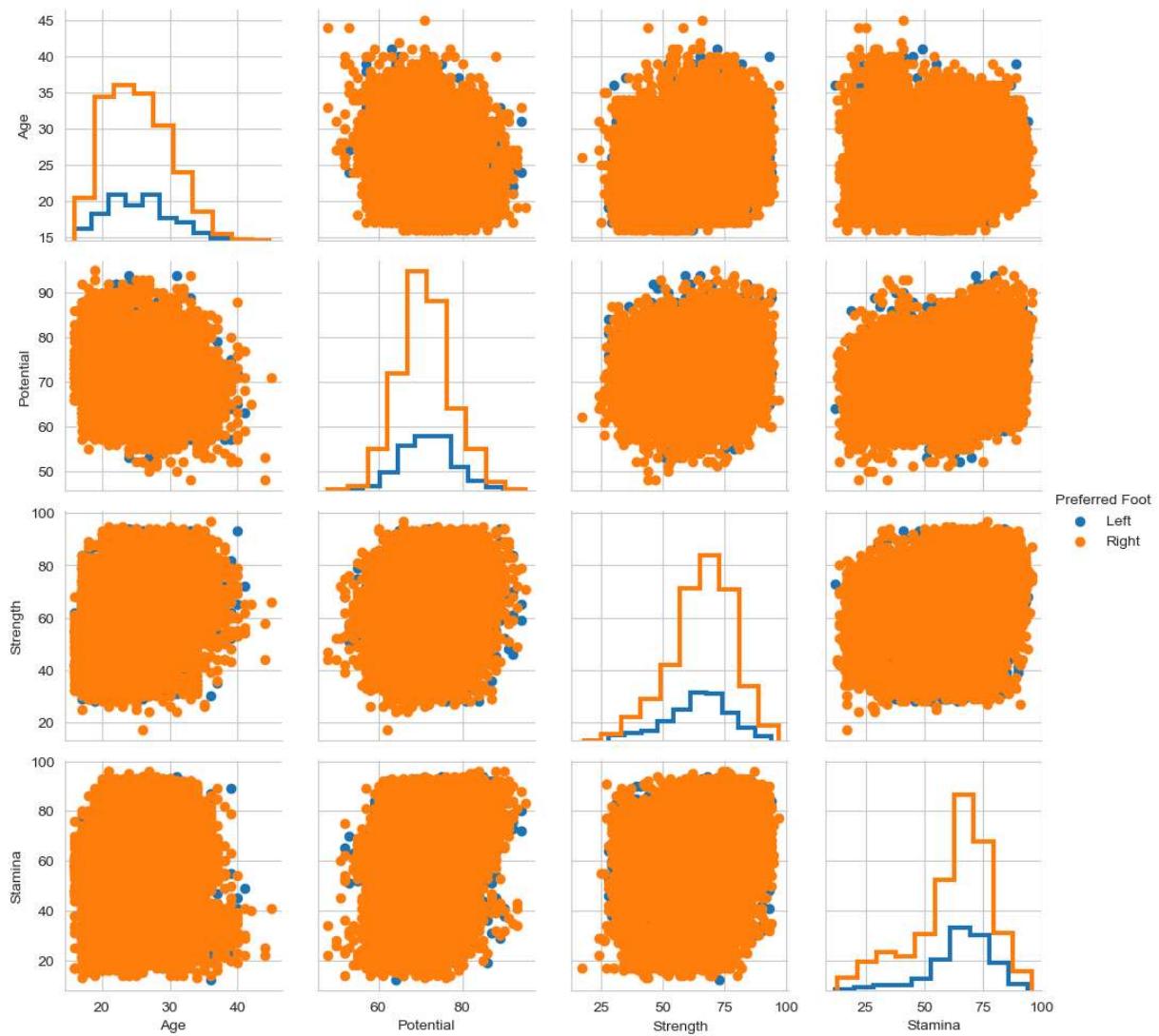
In [205]:

```
g = sns.PairGrid(data_new)
g = g.map_diag(plt.hist)
g = g.map_offdiag(plt.scatter)
```

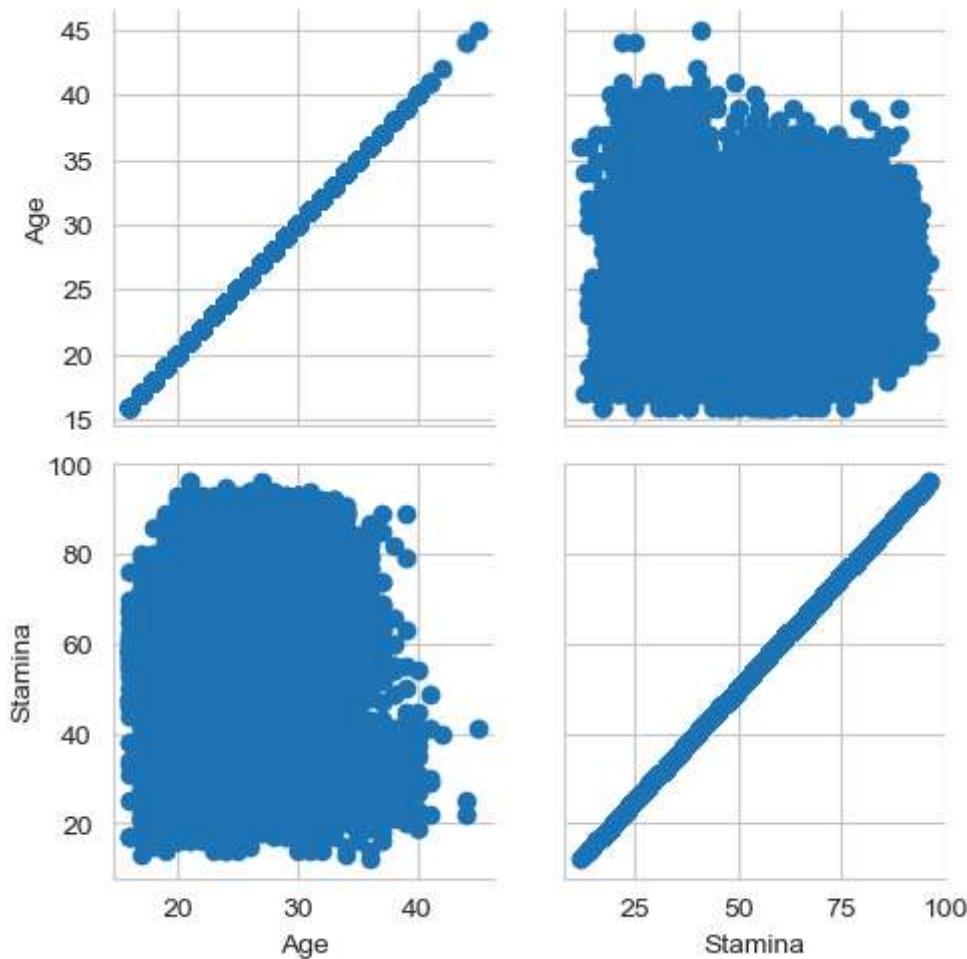


In [207]:

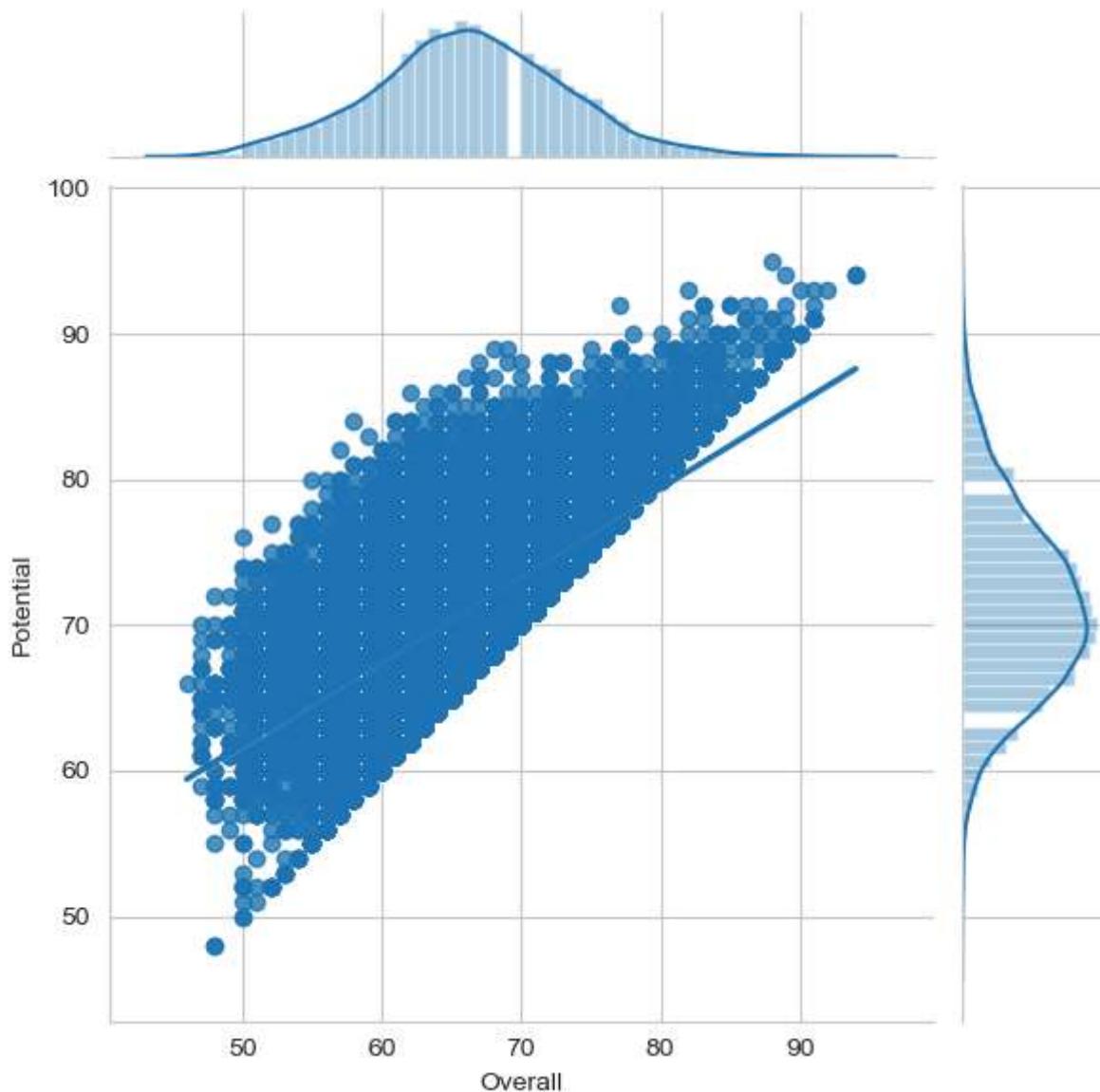
```
g = sns.PairGrid(data_new, hue="Preferred Foot")
g = g.map_diag(plt.hist, histtype="step", linewidth=3)
g = g.map_offdiag(plt.scatter)
g = g.add_legend()
```



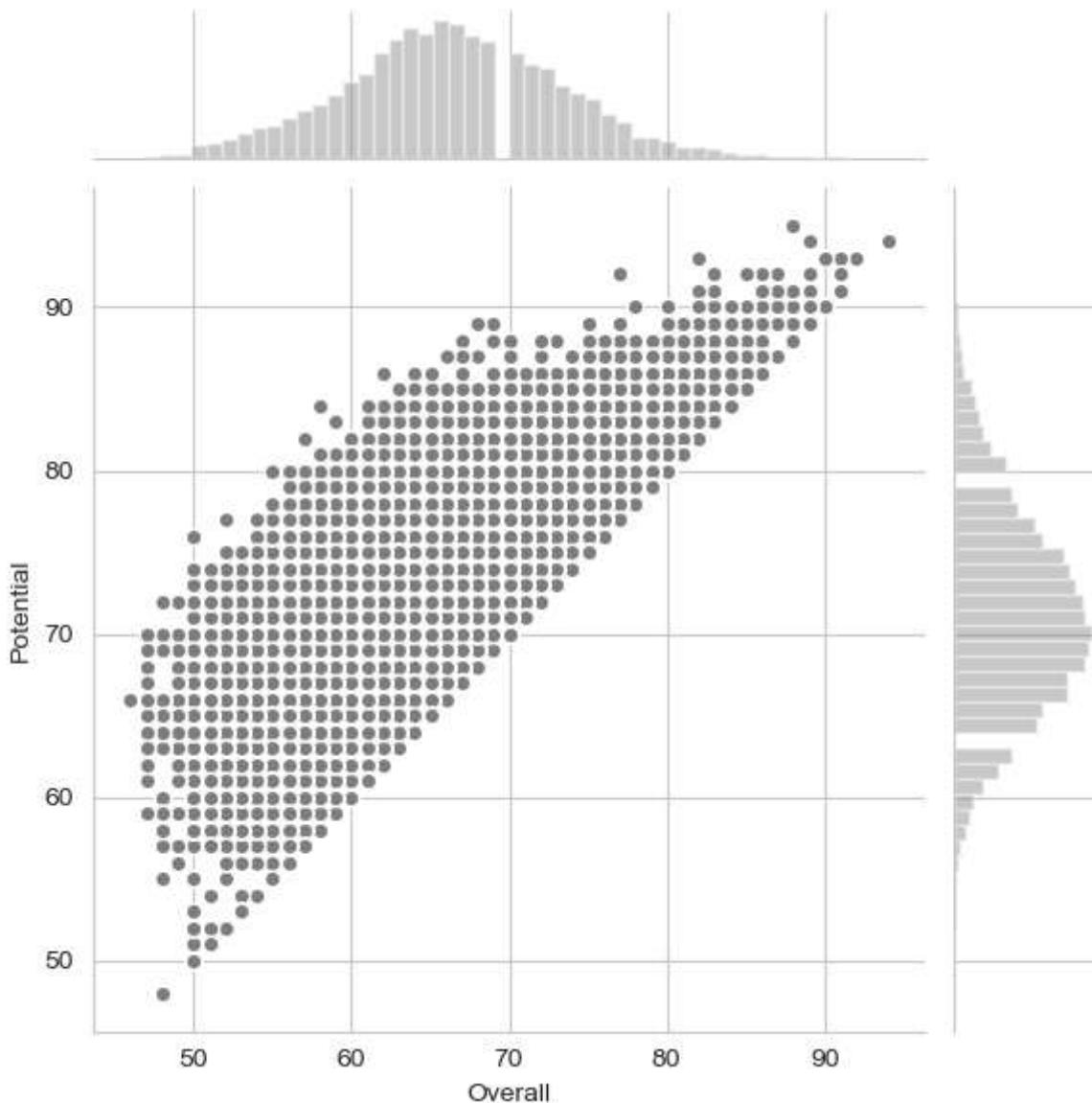
```
In [209]:  
g = sns.PairGrid(data_new, vars=['Age', 'Stamina'])  
g = g.map(plt.scatter)
```



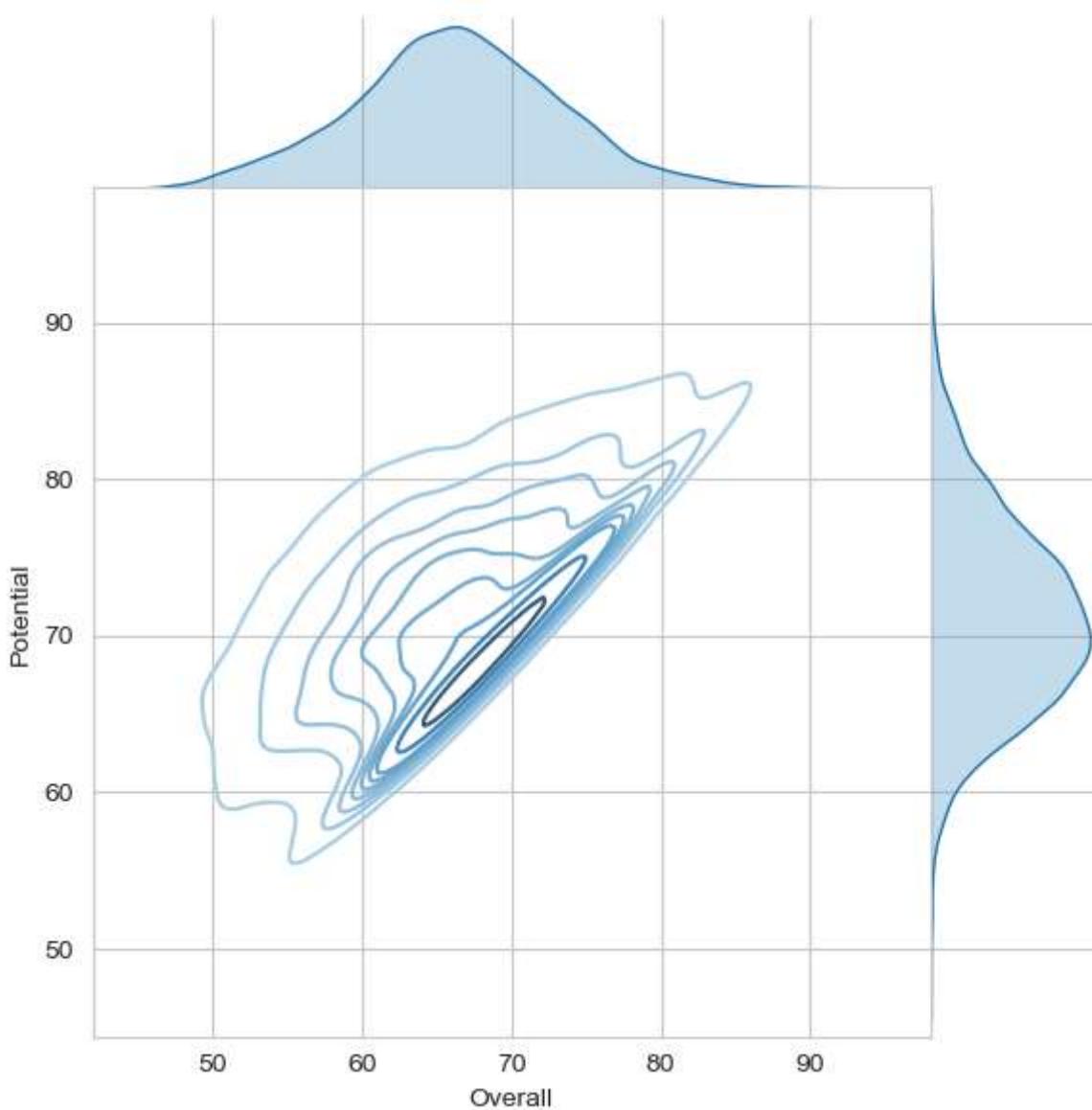
```
In [211]:  
g = sns.JointGrid(x="Overall", y="Potential", data=data)  
g = g.plot(sns.regplot, sns.distplot)
```



```
In [215]:  
g = sns.JointGrid(x="Overall", y="Potential", data=data)  
g = g.plot_joint(plt.scatter, color=".5", edgecolor="white")  
g = g.plot_marginals(sns.distplot, kde=False, color=".5")
```



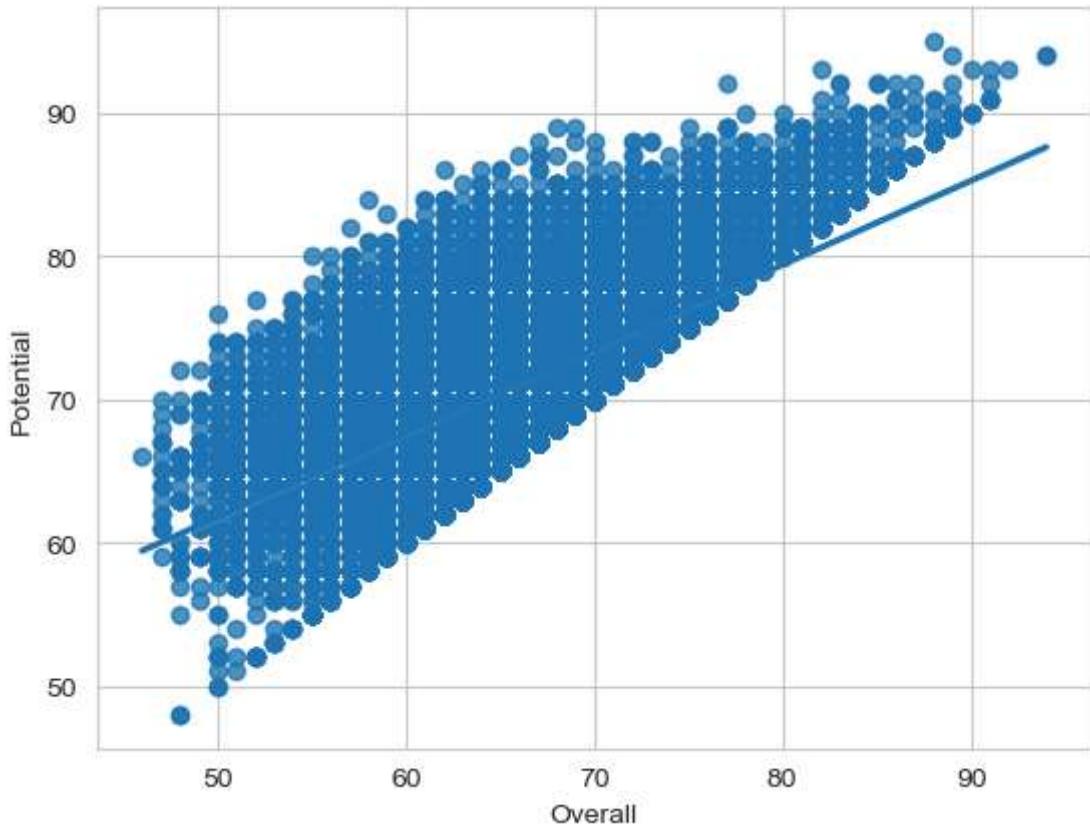
```
In [217]:  
g = sns.JointGrid(x="Overall", y="Potential", data=data, space=0)  
g = g.plot_joint(sns.kdeplot, cmap="Blues_d")  
g = g.plot_marginals(sns.kdeplot, shade=True)
```



```
In [ ]: g = sns.JointGrid(x="Overall", y="Potential", data=data, height=5, ratio=2)
g = g.plot_joint(sns.kdeplot, cmap="Reds_d")
g = g.plot_marginals(sns.kdeplot, color="r", shade=True)
```

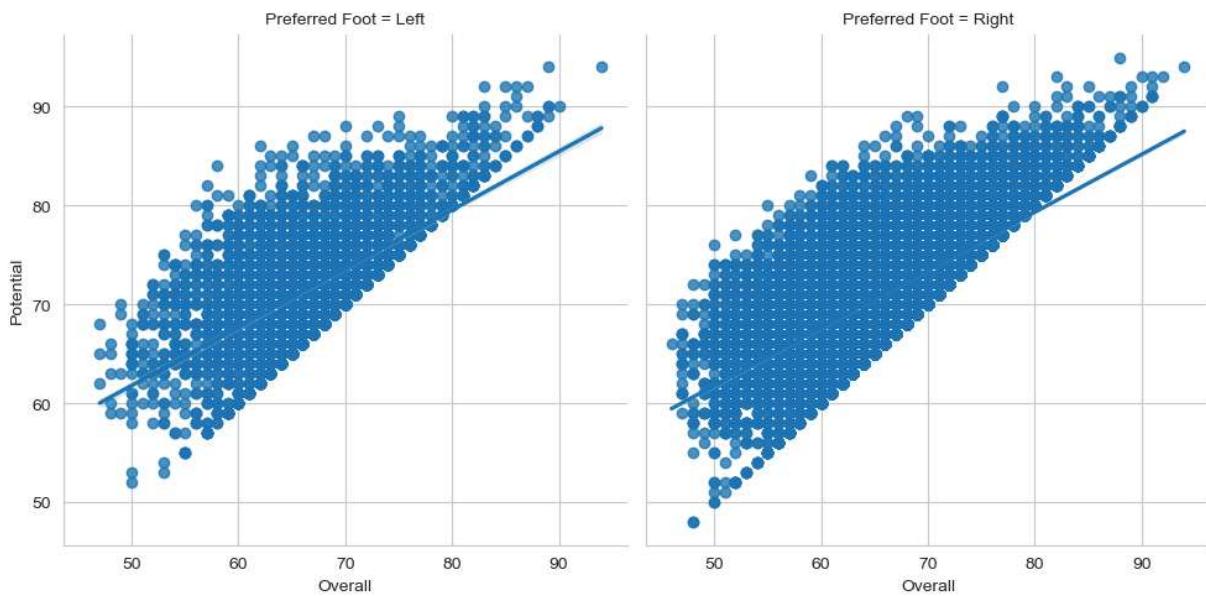
```
In [222]: sns.regplot(x="Overall", y="Potential", data=data)
```

```
Out[222]: <Axes: xlabel='Overall', ylabel='Potential'>
```

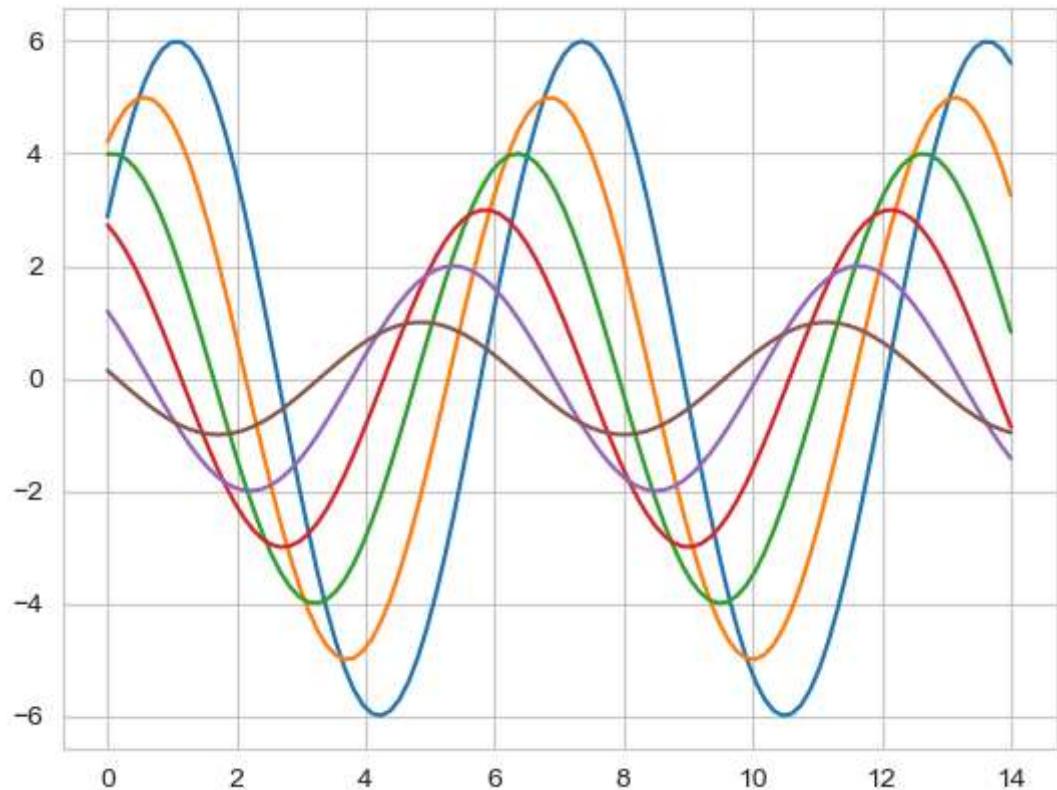


```
In [224]: sns.lmplot(x="Overall", y="Potential", col="Preferred Foot", data=data, col_wrap=2,
```

```
Out[224]: <seaborn.axisgrid.FacetGrid at 0x253a5b5f170>
```



```
In [228]: def sinplot(flip=1):
    x = np.linspace(0, 14, 100)
    for i in range(1, 7):
        plt.plot(x, np.sin(x + i * .5) * (7 - i) * flip)
sns.set_style("whitegrid")
sinplot()
```



In []: