In [1]:

```
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
sns.set(style="ticks")
```

In [2]:

```
data = pd.read_csv('C:\\Users\\Albriht\\Desktop\\FIFA 2018 Statistics.csv' , sep=",")
```

In [3]:

```
# Размер набора
data.shape
```

Out[3]:

```
(128, 27)
```

In [4]:

```
# Типы колонок
data.dtypes
```

Out[4]:

```
Date                          object
Team                          object
Opponent                      object
Goal Scored                    int64
Ball Possession %              int64
Attempts                       int64
On-Target                      int64
Off-Target                     int64
Blocked                        int64
Corners                        int64
Offsides                       int64
Free Kicks                     int64
Saves                          int64
Pass Accuracy %                int64
Passes                         int64
Distance Covered (Kms)         int64
Fouls Committed                int64
Yellow Card                    int64
Yellow & Red                   int64
Red                            int64
Man of the Match              object
1st Goal                     float64
Round                         object
PSO                           object
Goals in PSO                   int64
Own goals                    float64
Own goal Time                float64
dtype: object
```

In [5]:

```
total_count = data.shape[0]
print('Всего строк: {}'.format(total_count))
```

```
Всего строк: 128
```

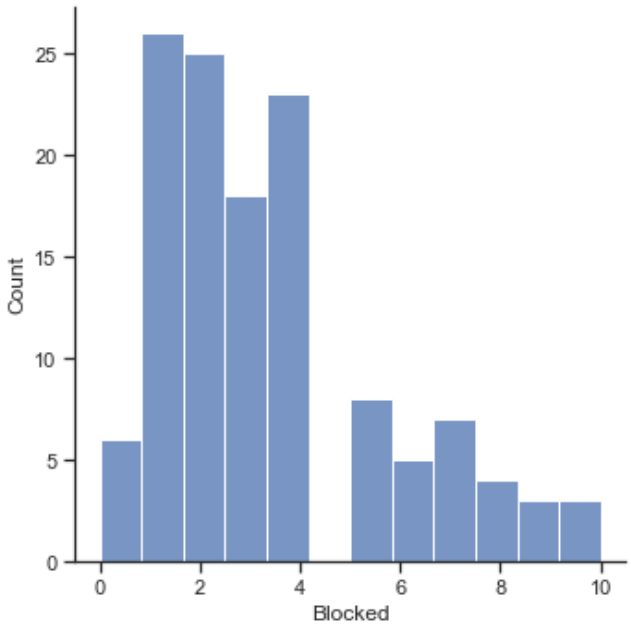
In [9]:

```
sns.displot (x=data['Blocked'])
```

Out[9]:

Out[9]:

```
<seaborn.axisgrid.FacetGrid at 0x255874934f0>
```



In [10]:

```
# Найдем пропуски в столбцах
data.isnull().sum()
```

Out[10]:

```
Date                          0
Team                          0
Opponent                      0
Goal Scored                   0
Ball Possession %             0
Attempts                      0
On-Target                     0
Off-Target                    0
Blocked                       0
Corners                       0
Offsides                      0
Free Kicks                    0
Saves                         0
Pass Accuracy %               0
Passes                        0
Distance Covered (Kms)        0
Fouls Committed               0
Yellow Card                   0
Yellow & Red                  0
Red                           0
Man of the Match              0
1st Goal                     34
Round                         0
PSO                           0
Goals in PSO                  0
Own goals                   116
Own goal Time               116
dtype: int64
```

In [11]:

```
num_cols = []
for col in data.columns:
    # Количество пустых значений
    temp_null_count = data[data[col].isnull()].shape[0]
    dt = str(data[col].dtype)
    if temp_null_count>0 and (dt=='float64' or dt=='object' or dt=='int64'):
        num_cols.append(col)
        temp_perc = round((temp_null_count / total_count) * 100.0, 2)
        print('Колонка {}. Тип данных {}. Количество пустых значений {}, {}%.'.format(co
```

```
l, dt, temp_null_count, temp_perc))
```

```
Колонка 1st Goal. Тип данных float64. Количество пустых значений 34, 26.56%.
Колонка Own goals. Тип данных float64. Количество пустых значений 116, 90.62%.
Колонка Own goal Time. Тип данных float64. Количество пустых значений 116, 90.62%.
```

In [12]:

```
for col in data.columns:
    temp_null_count = data[data[col].isnull()].shape[0]
    dt = str(data[col].dtype)
    if temp_null_count > 0 and (dt=='float64'):
        data[col]=data[col].fillna('0')
```

In [13]:

```
num_cols = []
for col in data.columns:
    # Количество пустых значений
    temp_null_count = data[data[col].isnull()].shape[0]
    dt = str(data[col].dtype)
    if temp_null_count>0 and (dt=='float64' or dt=='object' or dt=='int64'):
        num_cols.append(col)
        temp_perc = round((temp_null_count / total_count) * 100.0, 2)
        print('Колонка {}. Тип данных {}. Количество пустых значений {}, {}%.'.format(co
l, dt, temp_null_count, temp_perc))
```

In [14]:

```
data.isnull().sum()
```

Out[14]:

```
Date                       0
Team                       0
Opponent                   0
Goal Scored                0
Ball Possession %          0
Attempts                   0
On-Target                  0
Off-Target                 0
Blocked                    0
Corners                    0
Offsides                   0
Free Kicks                 0
Saves                      0
Pass Accuracy %            0
Passes                     0
Distance Covered (Kms)     0
Fouls Committed            0
Yellow Card                0
Yellow & Red               0
Red                        0
Man of the Match           0
1st Goal                   0
Round                      0
PSO                        0
Goals in PSO               0
Own goals                  0
Own goal Time              0
dtype: int64
```

In [ ]: