

In [1]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

epl = pd.read_excel('EPL Data.xlsx', sheet_name=1)
```

In [2]:

```
epl.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 65 entries, 0 to 64
Data columns (total 7 columns):
 #   Column          Non-Null Count  Dtype
---  -
 0   Team            65 non-null    object
 1   City            65 non-null    object
 2   Stadium         65 non-null    object
 3   Capacity        65 non-null    int64
 4   In Big Six      65 non-null    object
 5   In EPL          65 non-null    object
 6   Region         20 non-null    object
dtypes: int64(1), object(6)
memory usage: 3.7+ KB
```

In [3]:

```
epl.head()
```

Out[3]:

	Team	City	Stadium	Capacity	In Big Six	In EPL	Region
0	Arsenal	London	Emirates Stadium	60361	Big Six	EPL	London
1	Aston Villa	Birmingham	Villa Park	42785	Other	EPL	Central
2	Bournemouth	Bournemouth	Vitality Stadium	11464	Other	EPL	South
3	Brighton	Brighton	The Amex	30500	Other	EPL	South
4	Burnley	Burnley	Turf Moor	21401	Other	EPL	North

In [4]:

```
ep1[ep1['Region'].isnull()]
```

Out[4]:

	Team	City	Stadium	Capacity	In Big Six	In EPL	Region
20	Barnsley	Barnsley	Oakwell Stadium	23287	Other	Other	NaN
21	Birmingham City	Birmingham	St Andrew's Stadium	30009	Other	Other	NaN
22	Blackburn Rovers	Blackburn	Ewood Park	31154	Other	Other	NaN
23	Blackpool	Blackpool	Bloomfield Road	16007	Other	Other	NaN
24	Bolton Wanderers	Bolton	University of Bolton Stadium	28100	Other	Other	NaN
25	Bradford City	Bradford	Coral Windows Stadium	25136	Other	Other	NaN
26	Brentford	London	Griffin Park	12763	Other	Other	NaN
27	Bristol City	Bristol	Ashton Gate Stadium	27000	Other	Other	NaN
28	Bristol Rovers	Bristol	Memorial Stadium	11000	Other	Other	NaN
29	Bury	Bury	Gigg Lane	11840	Other	Other	NaN
30	Charlton Athletic	London	The Valley	27111	Other	Other	NaN
31	Coventry City	Coventry	Ricoh Arena	32604	Other	Other	NaN
32	Dagenham & Redbridge	Dagenham	Victoria Road	6078	Other	Other	NaN
33	Derby County	Derby	The iPro Stadium	33010	Other	Other	NaN
34	Doncaster Rovers	Doncaster	Keepmoat Stadium	15231	Other	Other	NaN
35	Fleetwood Town	Fleetwood	Highbury Stadium	5094	Other	Other	NaN
36	Fulham	London	Craven Cottage	25700	Other	Other	NaN
37	Gillingham FC	Gillingham	Priestfield Stadium	10500	Other	Other	NaN
38	Huddersfield Town	Huddersfield	John Smith's Stadium	24554	Other	Other	NaN
39	Hull City	Hull	KCOM Stadium	25586	Other	Other	NaN
40	Ipswich Town	Ipswich	Portman Road	30311	Other	Other	NaN
41	Leeds United	Leeds	Elland Road	39460	Other	Other	NaN
42	Leyton Orient	London	The Matchroom Stadium	9271	Other	Other	NaN
43	Luton Town	Luton	Kenilworth Road	10226	Other	Other	NaN
44	Middlesbrough	Middlesbrough	Riverside Stadium	34742	Other	Other	NaN
45	Millwall	London	The Den	19734	Other	Other	NaN
46	Nottingham Forest	Nottingham	City Ground	30567	Other	Other	NaN
47	Notts County	Nottingham	Meadow Lane	20300	Other	Other	NaN
48	Oxford United	Oxford	The Kassam Stadium	12500	Other	Other	NaN
49	Peterborough	Peterborough	ABAX Stadium	14084	Other	Other	NaN
50	Portsmouth	Portsmouth	Fratton Park	21178	Other	Other	NaN
51	Preston North End	Preston	Deepdale	23404	Other	Other	NaN

	Team	City	Stadium	Capacity	In Big Six	In EPL	Region
52	Queens Park Rangers	London	Loftus Road	18439	Other	Other	NaN
53	Reading	Reading	Madejski Stadium	24161	Other	Other	NaN
54	Rochdale AFC	Rochdale	Crown Oil Arena	12249	Other	Other	NaN
55	Rotherham United	Rotherham	New York Stadium	12000	Other	Other	NaN
56	Scunthorpe United	Scunthorpe	Glanford Park	9088	Other	Other	NaN
57	Sheffield Wednesday	Sheffield	Hillsborough	39732	Other	Other	NaN
58	Shrewsbury Town	Shrewsbury	Montgomery Waters Stadium	9875	Other	Other	NaN
59	Southend United	Southend-on-Sea	Roots Hall Stadium	13392	Other	Other	NaN
60	Stevenage FC	Stevenage	The Lamex Stadium	6490	Other	Other	NaN
61	Stoke City	Stoke-on-Trent	bet365 Stadium	27743	Other	Other	NaN
62	Sunderland	Sunderland	Stadium of Light	49000	Other	Other	NaN
63	West Bromwich Albion	West Bromwich	The Hawthorns	26287	Other	Other	NaN
64	Wigan Athletic	Wigan	DW Stadium	25133	Other	Other	NaN

In [5]:

```
ep1['Region'].value_counts()
```

Out[5]:

Region

Central 6

London 5

North 5

South 4

Name: count, dtype: int64

In [6]:

```
ep1[ep1['Region']=="London"]
```

Out[6]:

	Team	City	Stadium	Capacity	In Big Six	In EPL	Region
0	Arsenal	London	Emirates Stadium	60361	Big Six	EPL	London
5	Chelsea	London	Stamford Bridge	41837	Big Six	EPL	London
6	Crystal Palace	London	Selhurst Park	26225	Other	EPL	London
16	Tottenham	London	Tottenham Hotspur Stadium	62303	Big Six	EPL	London
18	West Ham	London	London Stadium	60000	Other	EPL	London

In [3]:

```
ep1.describe()
```

Out[3]:

Capacity	
count	65.000000
mean	27280.107692
std	15343.715712
min	5094.000000
25%	14084.000000
50%	25700.000000
75%	32609.000000
max	75811.000000

In [4]:

```
ep1.groupby(['Region']).aggregate({'Capacity':['sum','min','max']})
```

Out[4]:

	Capacity		
	sum	min	max
Region			
Central	228038	27010	54074
London	250726	26225	62303
North	237160	21401	75811
South	96153	11464	32689

In [6]:

```
ep1['Region'] = ep1['Region'].fillna('Default')
```

In [7]:

```
ep1.groupby(['Region']).aggregate({'Capacity':['sum','min','max']})
```

Out[7]:

Region	Capacity		
	sum	min	max
Central	228038	27010	54074
Default	961130	5094	49000
London	250726	26225	62303
North	237160	21401	75811
South	96153	11464	32689

In [8]:

```
ep1=ep1.sort_values(by=['Capacity'], ascending=False)  
ax = sns.swarmplot(x='Region',y='Capacity', data = ep1)
```

