In [1]:

```
import pandas as pd
```

In [21]:

```
ipl_data = pd.read_csv('IPL_Dataset.csv')
ipl_data
```

Out[21]:

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	1	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
1	2	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
2	3	David Warner	16	16	2	548	85*	39.14	407	134.64	0	4	52	14
3	4	Shreyas Iyer	17	17	2	519	88*	34.60	421	123.27	0	3	40	16
4	5	Ishan Kishan	14	13	4	516	99	57.33	354	145.76	0	4	36	30
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
128	129	Khaleel Ahmed	7	1	0	0	0*	0.00	2	0.00	0	0	0	0
129	130	Arshdeep Singh	8	1	0	0	0*	0.00	3	0.00	0	0	0	0
130	131	Daniel Sams	3	1	0	0	0*	0.00	2	0.00	0	0	0	0
131	132	Shreevats Goswami	2	2	0	0	0*	0.00	4	0.00	0	0	0	0
132	133	Trent Boult	15	1	0	0	0*	0.00	1	0.00	0	0	0	0

133 rows × 14 columns

In [7]:

```
#Q1. What is the maximum number of matches played by an individual player in a season?
max_matches_in_a_season = ipl_data[ipl_data['Mat']==ipl_data['Mat'].max()]
max_matches_in_a_season
```

Out[7]:

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
1	2	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
3	4	Shreyas Iyer	17	17	2	519	88*	34.60	421	123.27	0	3	40	16
19	20	Marcus Stoinis	17	17	3	352	65	25.14	237	148.52	0	3	31	16
67	68	Kagiso Rabada	17	8	4	56	15*	14.00	49	114.28	0	0	4	2

In [8]:

```
#Q2. What is the highest average of a player in a season?
avg_highest_in_a_seasion = ipl_data[ipl_data['Avg']==ipl_data['Avg'].max()]
avg_highest_in_a_seasion
```

Out[8]:

	POS	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
57	58	Deepak Hooda	7	5	4	101	62*	101.0	71	142.25	0	1	5	5

In [16]:

```
#Q3. What is the average number of matches played by the players in a season?
avg_match_played_by_players = ipl_data['Mat'].mean().round(2)
avg_match_played_by_players
```

Out[16]:

```
9.63
```

In [22]:

```
ipl_data = ipl_data.drop(columns = 'POS')
```

In [25]:

```
#Q4. Please check the correlation between all the features.
corr_matrix = ipl_data.corr(numeric_only=True)
corr_matrix
```

Out[25]:

	Mat	Inns	NO	Runs	Avg	BF	SR	100	50	4s	6s
Mat	1.000000	0.728358	0.555261	0.565237	0.306128	0.555102	0.371328	0.128405	0.455437	0.522628	0.523367
Inns	0.728358	1.000000	0.464356	0.895914	0.585148	0.889753	0.581468	0.224328	0.755027	0.846749	0.810922
NO	0.555261	0.464356	1.000000	0.308747	0.429888	0.271021	0.415060	0.019726	0.174184	0.204779	0.391472
Runs	0.565237	0.895914	0.308747	1.000000	0.687475	0.989948	0.517678	0.382577	0.912264	0.959576	0.861166
Avg	0.306128	0.585148	0.429888	0.687475	1.000000	0.664173	0.591306	0.233726	0.638326	0.623643	0.646224
BF	0.555102	0.889753	0.271021	0.989948	0.664173	1.000000	0.474226	0.358550	0.900547	0.955994	0.810916
SR	0.371328	0.581468	0.415060	0.517678	0.591306	0.474226	1.000000	0.135331	0.406029	0.465132	0.566317
100	0.128405	0.224328	0.019726	0.382577	0.233726	0.358550	0.135331	1.000000	0.292812	0.456622	0.186767
50	0.455437	0.755027	0.174184	0.912264	0.638326	0.900547	0.406029	0.292812	1.000000	0.892025	0.757501
4s	0.522628	0.846749	0.204779	0.959576	0.623643	0.955994	0.465132	0.456622	0.892025	1.000000	0.731094
6s	0.523367	0.810922	0.391472	0.861166	0.646224	0.810916	0.566317	0.186767	0.757501	0.731094	1.000000

In [31]:

```
#Q5. Check the list of players who has an average greater than 50 as well strike rate above 120.
pl_avg_n_sr = ipl_data[(ipl_data['Avg']>50) & (ipl_data['SR']>120)]
pl_avg_n_sr[['Player', 'Avg', 'SR']]
```

Out[31]:

	Player	Avg	SR
0	KL Rahul	55.83	129.34
4	Ishan Kishan	57.33	145.76
31	Kieron Pollard	53.60	191.42
36	Wriddhiman Saha	71.33	139.86
37	Ruturaj Gaikwad	51.00	120.71
57	Deepak Hooda	101.00	142.25
60	Tom Curran	83.00	133.87

In [33]:

```
#Q6. Please check the list of players who has an average greater than 40 and balls faced above 100.
pl_abv40_n_bf_abv100 = ipl_data[(ipl_data['Avg']>40) & (ipl_data['BF']>100)]
pl_abv40_n_bf_abv100[['Player', 'Avg', 'BF']]
```

Out[33]:

	Player	Avg	BF
0	KL Rahul	55.83	518
1	Shikhar Dhawan	44.14	427
4	Ishan Kishan	57.33	354
8	Virat Kohli	42.36	384
9	AB de Villiers	45.40	286
10	Faf du Plessis	40.81	319
14	Eoin Morgan	41.80	302
24	Kane Williamson	45.28	237
27	Chris Gayle	41.14	210
28	Ben Stokes	40.71	200
31	Kieron Pollard	53.60	140
32	Rahul Tewatia	42.50	183
33	Ravindra Jadeja	46.40	135
36	Wriddhiman Saha	71.33	153
37	Ruturaj Gaikwad	51.00	169

In [34]:

```
#Q7. Please check the list of players who scored more than 300 runs with an average greater than 45 &
pl_run300_avg45_sr130 = ipl_data[(ipl_data['Runs']>300) & (ipl_data['Avg']>45) &(ipl_data['SR']>130)]
pl_run300_avg45_sr130
```

Out[34]:

	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
4	Ishan Kishan	14	13	4	516	99	57.33	354	145.76	0	4	36	30
9	AB de Villiers	15	14	4	454	73*	45.40	286	158.74	0	5	33	23
24	Kane Williamson	12	11	4	317	67	45.28	237	133.75	0	3	26	10

In [35]:

```
#Q8. Players who scroed alteast one century in this season.
pl_atleast_one_100 = ipl_data[ipl_data['100']>=1]
pl_atleast_one_100
```

Out[35]:

	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
1	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
13	Mayank Agarwal	11	11	0	424	106	38.54	271	156.45	1	2	44	15
28	Ben Stokes	8	8	1	285	107*	40.71	200	142.50	1	1	36	7

In [37]:

```
#Q9. Players who scored atleast 4 half centuries in this season.
pl_atleast_4_half_century = ipl_data[ipl_data['50']>=4]
pl_atleast_4_half_century
```

Out[37]:

	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
1	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
2	David Warner	16	16	2	548	85*	39.14	407	134.64	0	4	52	14
4	Ishan Kishan	14	13	4	516	99	57.33	354	145.76	0	4	36	30
5	Quinton de Kock	16	16	2	503	78*	35.92	358	140.50	0	4	46	22
6	Suryakumar Yadav	16	15	3	480	79*	40.00	331	145.01	0	4	61	11
7	Devdutt Padikkal	15	15	0	473	74	31.53	379	124.80	0	5	51	8
9	AB de Villiers	15	14	4	454	73*	45.40	286	158.74	0	5	33	23
10	Faf du Plessis	13	13	2	449	87*	40.81	319	140.75	0	4	42	14

In [44]:

```
#Q10. Check the list of players who hit more than 45 boundARIES and more than 10 sixes in this season
ipl_data[(ipl_data['4s'] >45)&(ipl_data['6s'] >10)]
```

Out[44]:

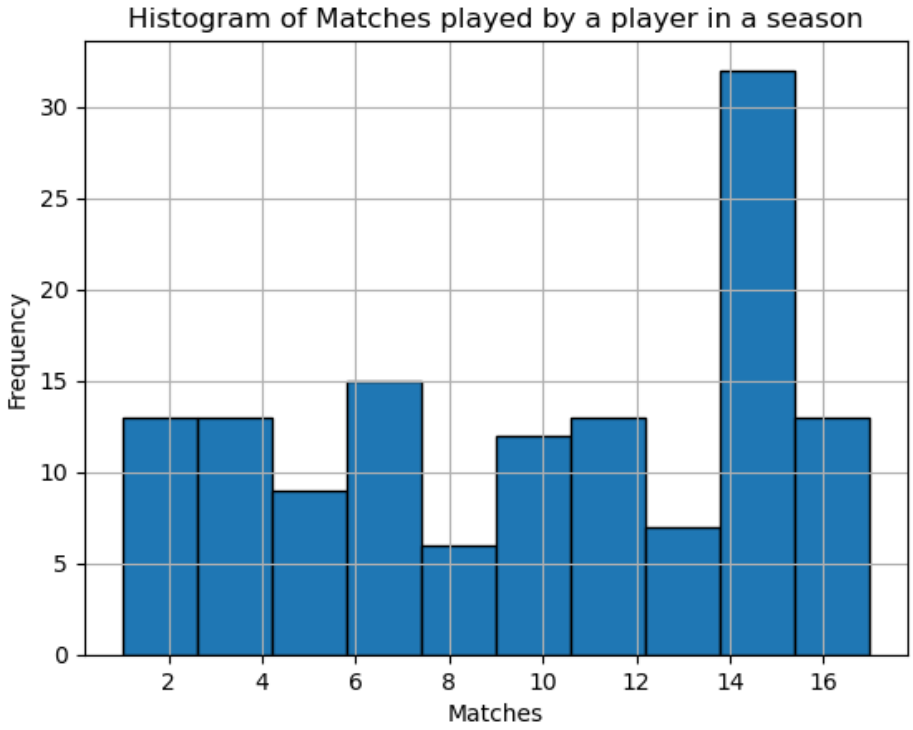
	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
1	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
2	David Warner	16	16	2	548	85*	39.14	407	134.64	0	4	52	14
5	Quinton de Kock	16	16	2	503	78*	35.92	358	140.50	0	4	46	22
6	Suryakumar Yadav	16	15	3	480	79*	40.00	331	145.01	0	4	61	11

In [49]:

```
#Q11. Plot a histogram of number of matches played in a season by players.
import matplotlib.pyplot as plt
```

In [51]:

```
ipl_data.hist('Mat', bins=10, edgecolor='black')
plt.xlabel('Matches')
plt.ylabel('Frequency')
plt.title('Histogram of Matches played by a player in a season')
plt.show()
```

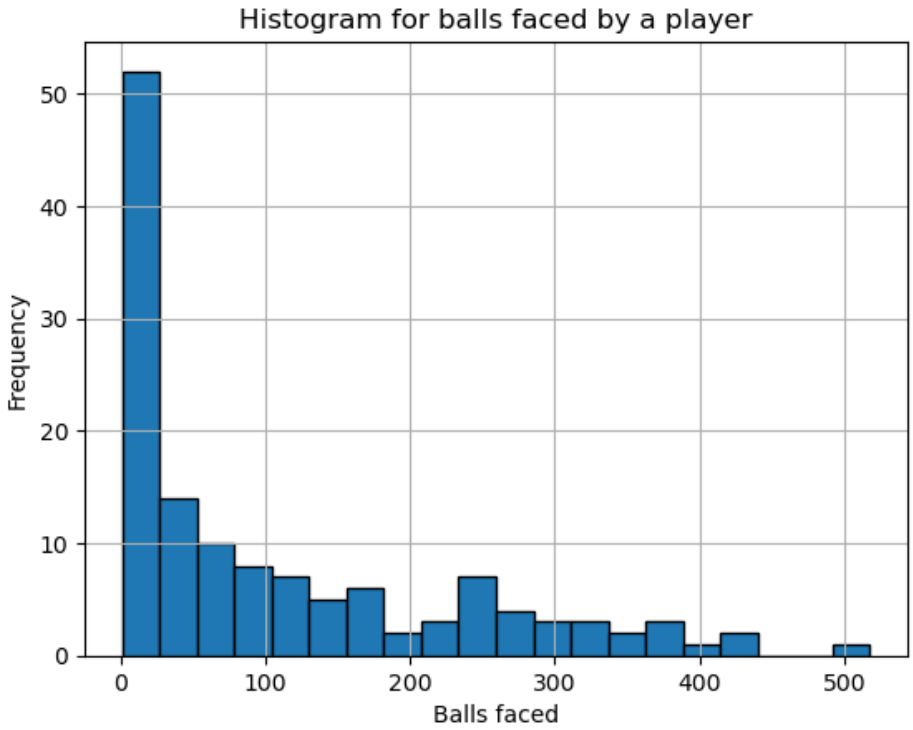


In [54]:

```
#Plot the histogram of balls faced by players.
ipl_data.hist(column='BF', bins=20, edgecolor='black')
plt.xlabel('Balls faced')
plt.ylabel('Frequency')
plt.title('Histogram for balls faced by a player')
```

Out[54]:

```
Text(0.5, 1.0, 'Histogram for balls faced by a player')
```



In [59]:

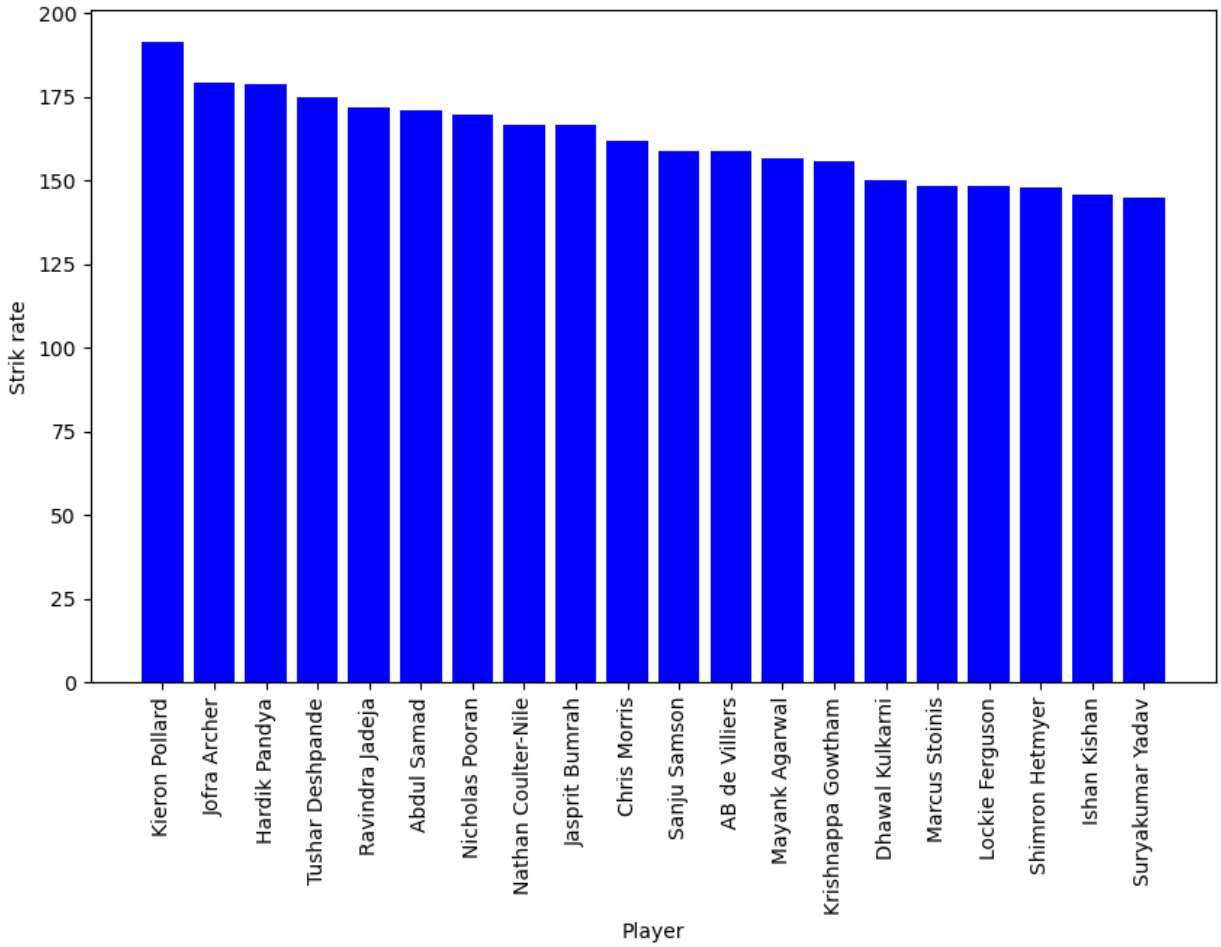
```
#Q13. Top 20 players with most runs in a season.
top_20_batter = ipl_data.sort_values(by='Runs', ascending=False)
top_20_batter.head(20)
```

Out[59]:

	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
0	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
1	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
2	David Warner	16	16	2	548	85*	39.14	407	134.64	0	4	52	14
3	Shreyas Iyer	17	17	2	519	88*	34.60	421	123.27	0	3	40	16
4	Ishan Kishan	14	13	4	516	99	57.33	354	145.76	0	4	36	30
5	Quinton de Kock	16	16	2	503	78*	35.92	358	140.50	0	4	46	22
6	Suryakumar Yadav	16	15	3	480	79*	40.00	331	145.01	0	4	61	11
7	Devdutt Padikkal	15	15	0	473	74	31.53	379	124.80	0	5	51	8
8	Virat Kohli	15	15	4	466	90*	42.36	384	121.35	0	3	23	11
9	AB de Villiers	15	14	4	454	73*	45.40	286	158.74	0	5	33	23
10	Faf du Plessis	13	13	2	449	87*	40.81	319	140.75	0	4	42	14
11	Shubman Gill	14	14	1	440	70*	33.84	373	117.96	0	3	44	9
12	Manish Pandey	16	15	2	425	83*	32.69	333	127.62	0	3	35	18
13	Mayank Agarwal	11	11	0	424	106	38.54	271	156.45	1	2	44	15
14	Eoin Morgan	14	14	4	418	68*	41.80	302	138.41	0	1	32	24
15	Sanju Samson	14	14	1	375	85	28.84	236	158.89	0	3	21	26
16	Ambati Rayudu	12	11	2	359	71	39.88	282	127.30	0	1	30	12
17	Nicholas Pooran	14	14	4	353	77	35.30	208	169.71	0	2	23	25
18	Nitish Rana	14	14	0	352	87	25.14	254	138.58	0	3	43	12
19	Marcus Stoinis	17	17	3	352	65	25.14	237	148.52	0	3	31	16

In [66]:

```
#Q14. Bar chart for top 20 players with best strike rate in a season.
top_20_sr = ipl_data.nlargest(20 ,columns= 'SR')
#top_20_sr = ipl_data.sort_values(by='SR', ascending= False).head(20)
plt.figure(figsize=(10, 6))
plt.bar(top_20_sr['Player'], top_20_sr['SR'],color='blue')
plt.xticks(rotation=90)
plt.xlabel('Player')
plt.ylabel('Strik rate')
plt.show()
```



In [67]:

```
#Q15. Top 30 players with highest number of sixes.
top_30_six = ipl_data.nlargest(30, columns='6s')
top_30_six
```

Out[67]:

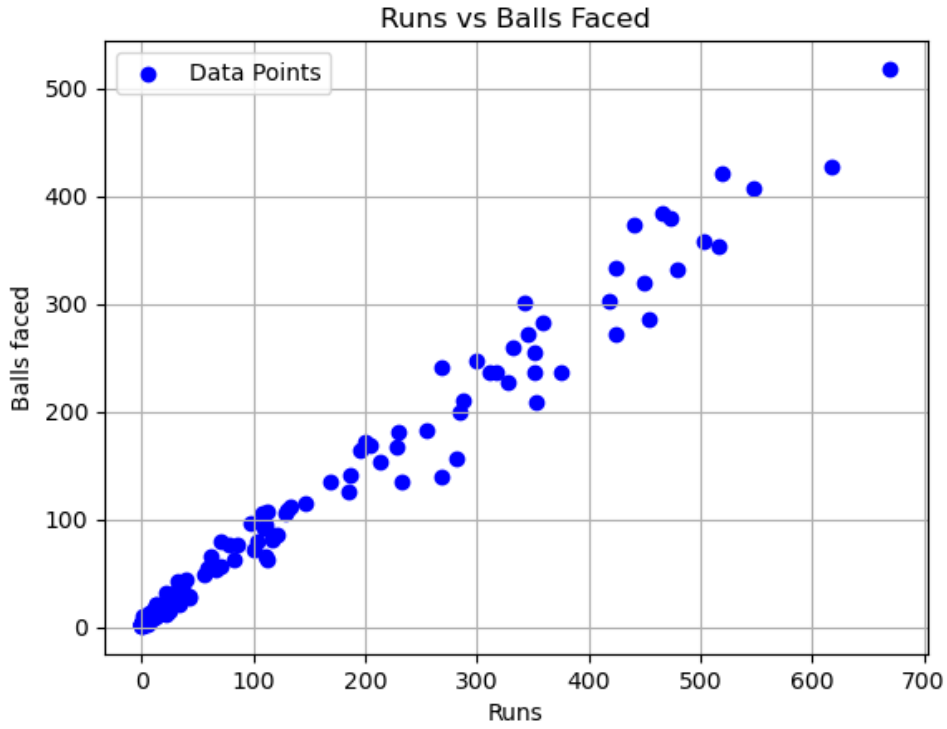
	Player	Mat	Inns	NO	Runs	HS	Avg	BF	SR	100	50	4s	6s
4	Ishan Kishan	14	13	4	516	99	57.33	354	145.76	0	4	36	30
15	Sanju Samson	14	14	1	375	85	28.84	236	158.89	0	3	21	26
17	Nicholas Pooran	14	14	4	353	77	35.30	208	169.71	0	2	23	25
29	Hardik Pandya	14	13	5	281	60*	35.12	157	178.98	0	1	14	25
14	Eoin Morgan	14	14	4	418	68*	41.80	302	138.41	0	1	32	24
0	KL Rahul	14	14	2	670	132*	55.83	518	129.34	1	5	58	23
9	AB de Villiers	15	14	4	454	73*	45.40	286	158.74	0	5	33	23
27	Chris Gayle	7	7	0	288	99	41.14	210	137.14	0	3	15	23
5	Quinton de Kock	16	16	2	503	78*	35.92	358	140.50	0	4	46	22
31	Kieron Pollard	16	12	7	268	60*	53.60	140	191.42	0	1	15	22
22	Rohit Sharma	12	12	0	332	80	27.66	260	127.69	0	3	27	19
12	Manish Pandey	16	15	2	425	83*	32.69	333	127.62	0	3	35	18
32	Rahul Tewatia	14	11	5	255	53	42.50	183	139.34	0	1	13	17
3	Shreyas Iyer	17	17	2	519	88*	34.60	421	123.27	0	3	40	16
19	Marcus Stoinis	17	17	3	352	65	25.14	237	148.52	0	3	31	16
23	Jos Buttler	13	12	2	328	70*	32.80	227	144.49	0	2	27	16
13	Mayank Agarwal	11	11	0	424	106	38.54	271	156.45	1	2	44	15
2	David Warner	16	16	2	548	85*	39.14	407	134.64	0	4	52	14
10	Faf du Plessis	13	13	2	449	87*	40.81	319	140.75	0	4	42	14
20	Jonny Bairstow	11	11	0	345	97	31.36	272	126.83	0	3	31	13
26	Shane Watson	11	11	1	299	83*	29.90	247	121.05	0	2	33	13
1	Shikhar Dhawan	17	17	3	618	106*	44.14	427	144.73	2	4	67	12
16	Ambati Rayudu	12	11	2	359	71	39.88	282	127.30	0	1	30	12
18	Nitish Rana	14	14	0	352	87	25.14	254	138.58	0	3	43	12
40	Sam Curran	14	11	3	186	52	23.25	141	131.91	0	1	12	12
41	Shimron Hetmyer	12	11	3	185	45	23.12	125	148.00	0	0	11	12
6	Suryakumar Yadav	16	15	3	480	79*	40.00	331	145.01	0	4	61	11
8	Virat Kohli	15	15	4	466	90*	42.36	384	121.35	0	3	23	11
33	Ravindra Jadeja	14	11	6	232	50	46.40	135	171.85	0	1	22	11
24	Kane Williamson	12	11	4	317	67	45.28	237	133.75	0	3	26	10

In [68]:

```
#Q16. Scatter plot of runs scored by a player v/s balls faced in a season.
plt.scatter(ipl_data['Runs'], ipl_data['BF'],color='blue', marker='o', label='Data Points')
plt.xlabel('Runs')
plt.ylabel('Balls faced')
plt.title('Runs vs Balls Faced')

plt.legend()
plt.grid(True)

plt.show()
```



In [69]:

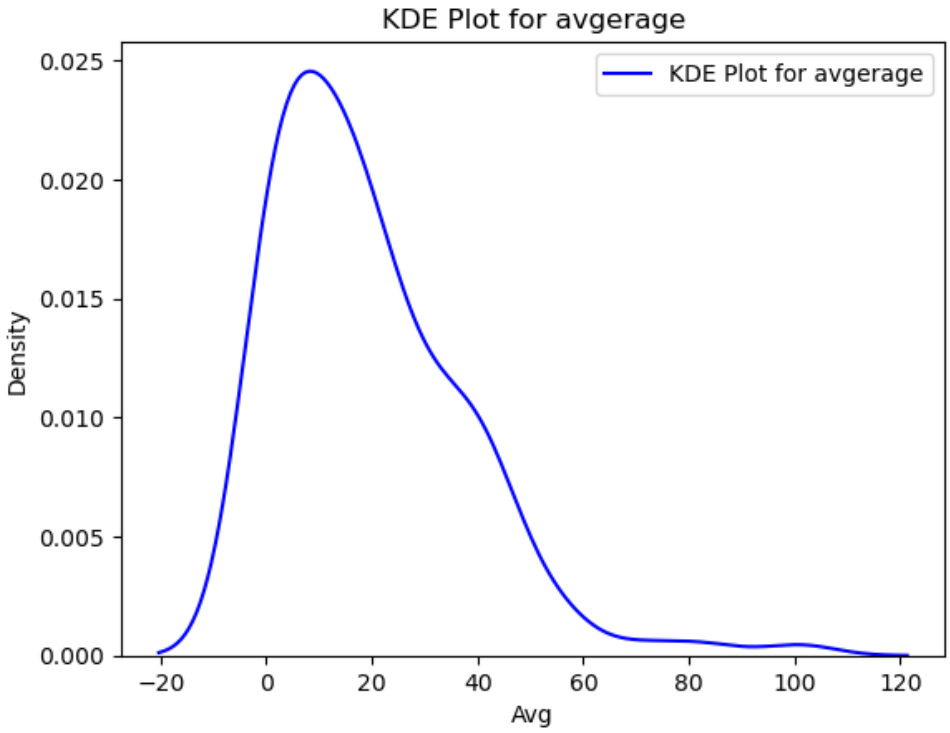
```
#Q17. KDE plot for averages on a season.
import seaborn as sns
```

In [70]:

```
sns.kdeplot(ipl_data['Avg'],color='blue', label='KDE Plot for avgerage')
plt.xlabel('Avg')
plt.ylabel('Density')
plt.title('KDE Plot for avgerage')

plt.legend()  # Show legend if you added a label

plt.show()
```



In [ ]: