In [1]:
```
# Import necessary libraries
import pandas as pd # Pandas for data manipulation
import matplotlib.pyplot as plt # Matplotlib for basic plotting
import seaborn as sns # Seaborn for statistical data visualization
```

In [2]:
```
# Read the CSV file into a DataFrame
df = pd.read_csv("all_seasons.csv")
```

In [3]:
```
# Rename selected columns for better readability
df.rename(columns = {'player_name':'name', 'team_abbreviation':'team'
                     'pts':'avg_pts', 'reb':'avg_reb', 'ast':'avg_ast
```

In [4]:
```
# Display the first few rows of the DataFrame
df.head()
```

Out[4]:

ght	weight	college	country	draft_year	draft_round	...	avg_pts	avg_reb	avg_ast
.04	94.800728	Louisiana State	USA	1996	2	...	3.9	1.5	2.4
.50	86.182480	Northwestern Oklahoma	USA	1994	2	...	3.8	1.3	0.3
.20	103.418976	North Carolina	USA	1993	1	...	8.3	6.4	1.9
.20	102.058200	Florida State	USA	1989	1	...	10.2	2.8	1.7
.36	119.748288	UCLA	USA	1995	1	...	2.8	1.7	0.3

In [15]:
```
# Top Scorers
NumTopScorers = 10
top_scorers_names = df.groupby(['name'])['pts'].sum().sort_values(asc
top_scorers_names
```

Out[15]:
```
Index(['LeBron James', 'Kobe Bryant', 'Dirk Nowitzki', 'Carmelo Anth
ony',
       'Kevin Durant', 'Tim Duncan', 'Paul Pierce', 'Vince Carter',
       'Kevin Garnett', 'James Harden'],
      dtype='object', name='name')
```

In [25]:
```
# Display the first 10 rows of data
df.head(11)
```

Out[25]:

	Unnamed: 0	name	team	age	height	weight	college	country	draft_year
0	0	Randy Livingston	HOU	22.0	193.04	94.800728	Louisiana State	USA	1996
1	1	Gaylon Nickerson	WAS	28.0	190.50	86.182480	Northwestern Oklahoma	USA	1994
2	2	George Lynch	VAN	26.0	203.20	103.418976	North Carolina	USA	1993
3	3	George McCloud	LAL	30.0	203.20	102.058200	Florida State	USA	1989
4	4	George Zidek	DEN	23.0	213.36	119.748288	UCLA	USA	1995
5	5	Gerald Wilkins	ORL	33.0	198.12	102.058200	Tennessee-Chattanooga	USA	1985
6	6	Gheorghe Muresan	WAS	26.0	231.14	137.438376	None	USA	1993
7	7	Glen Rice	CHH	30.0	203.20	99.790240	Michigan	USA	1989
8	8	Glenn Robinson	MIL	24.0	200.66	106.594120	Purdue	USA	1994
9	9	Grant Hill	DET	24.0	203.20	102.058200	Duke	USA	1994
10	10	Gary Trent	POR	22.0	203.20	113.398000	Ohio	USA	1995

11 rows × 28 columns

In [39]:
```
# Display the last few rows of the data
df.tail()
```

Out[39]:

	Unnamed: 0	name	team	age	height	weight	college	country	draft_year
12839	12839	Joel Embiid	PHI	29.0	213.36	127.005760	Kansas	Cameroon	2014
12840	12840	John Butler Jr.	POR	20.0	213.36	86.182480	Florida State	USA	Undrafted
12841	12841	John Collins	ATL	25.0	205.74	102.511792	Wake Forest	USA	2017
12842	12842	Jericho Sims	NYK	24.0	208.28	113.398000	Texas	USA	2021
12843	12843	JaMychal Green	GSW	33.0	205.74	102.965384	Alabama	USA	Undrafted

5 rows × 28 columns

In [6]:
```
# Calculate cumulative sums for points, rebounds, and assists
df['total_pts'] = df.groupby('name')['pts'].cumsum()
df['total_reb'] = df.groupby('name')['reb'].cumsum()
df['total_ast'] = df.groupby('name')['ast'].cumsum()
```

In [7]:
```
# Display the first few rows of the DataFrame for Kyrie Irving
df[df['name']=='Kyrie Irving'].head()
```

Out[7]:

e	team	age	height	weight	college	country	draft_year	draft_round	...	usg_pct	ts_pct
e g	CLE	20.0	190.5	86.636072	Duke	Australia	2011	1	...	0.281	0.566
e g	CLE	21.0	190.5	86.636072	Duke	Australia	2011	1	...	0.298	0.553
e g	CLE	22.0	190.5	87.543256	Duke	Australia	2011	1	...	0.280	0.533
e g	CLE	23.0	190.5	87.543256	Duke	Australia	2011	1	...	0.260	0.583
e g	CLE	24.0	190.5	87.543256	Duke	Australia	2011	1	...	0.293	0.540

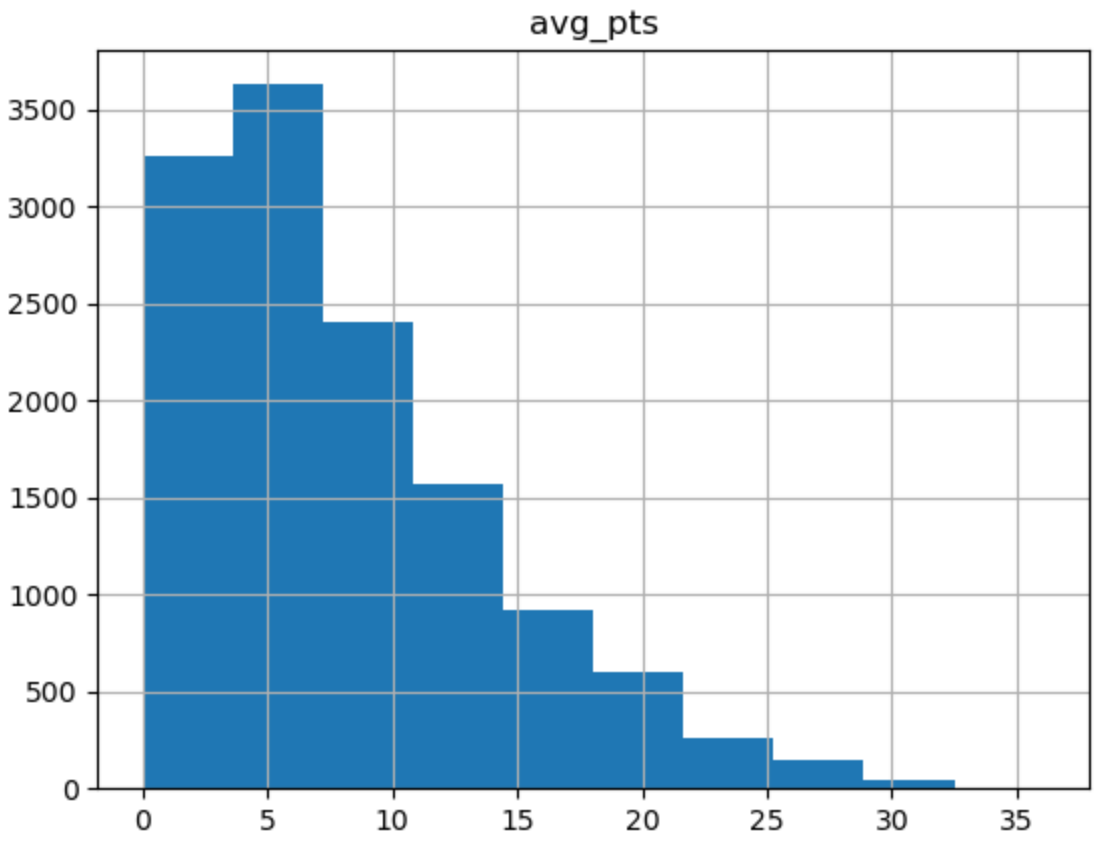
In [30]:
```
# Calculate Total Points, Rebounds, and Assists
df['pts'] = (df['gp']*df['avg_pts']).astype(int)
df['reb'] = (df['gp']*df['avg_reb']).astype(int)
df['ast'] = (df['gp']*df['avg_ast']).astype(int)
```

In [31]:
```
# Display first few rows
df.head()
```

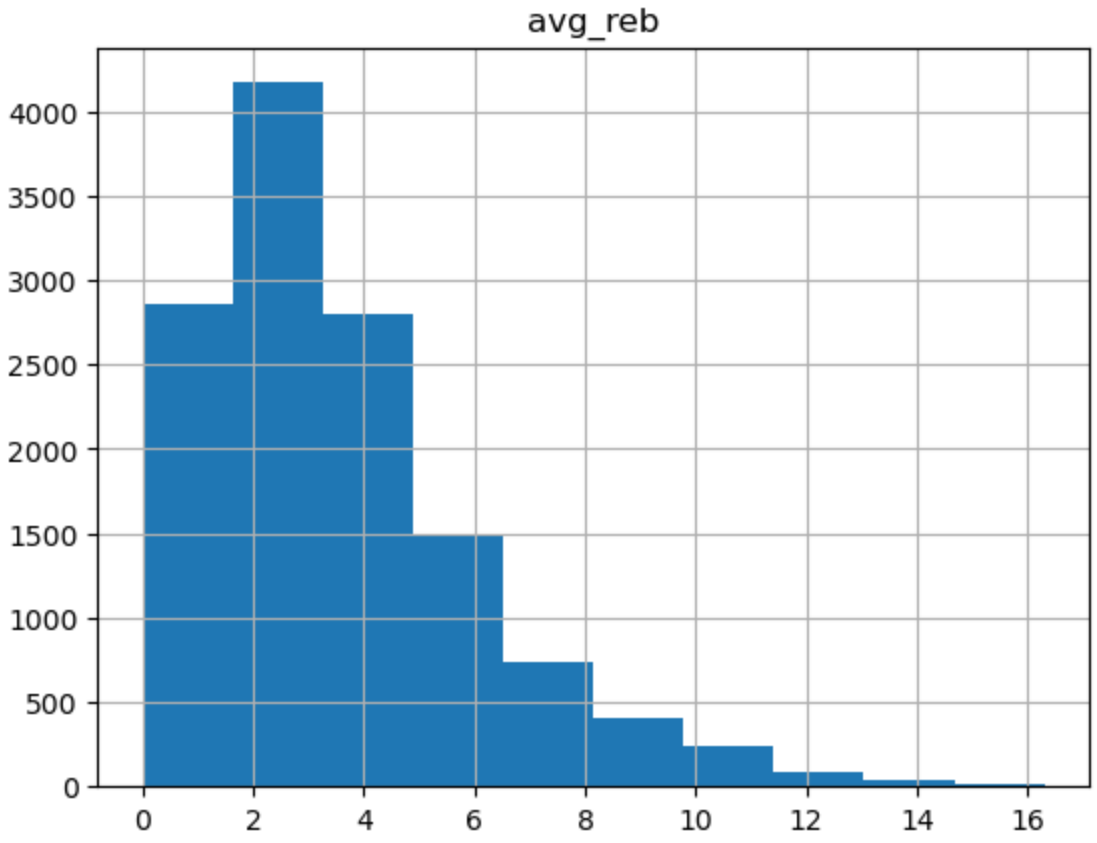
Out[31]:

am	age	height	weight	college	country	draft_year	draft_round	...	usg_pct	ts_pc
OU	22.0	193.04	94.800728	Louisiana State	USA	1996	2	...	0.169	0.48
AS	28.0	190.50	86.182480	Northwestern Oklahoma	USA	1994	2	...	0.174	0.49
AN	26.0	203.20	103.418976	North Carolina	USA	1993	1	...	0.175	0.51
AL	30.0	203.20	102.058200	Florida State	USA	1989	1	...	0.206	0.52
EN	23.0	213.36	119.748288	UCLA	USA	1995	1	...	0.195	0.50

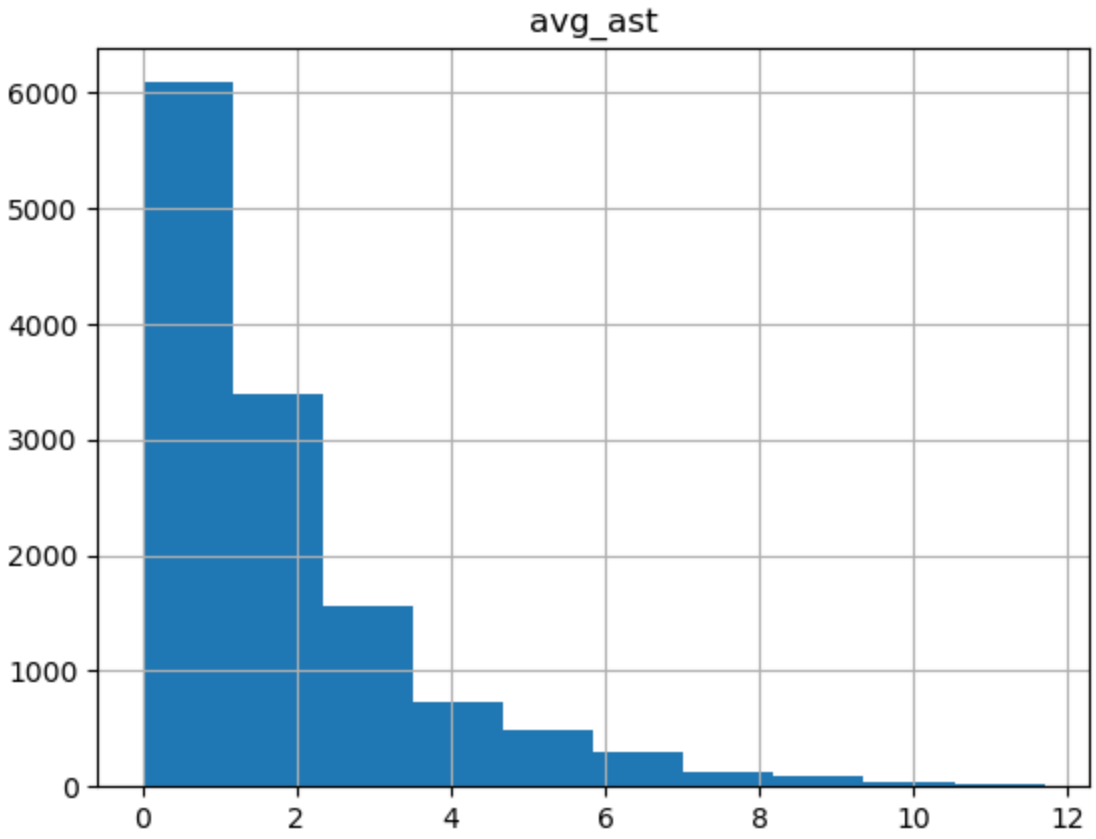
In [35]:
```
df.hist(['avg_pts']); # Histograms for Average Points
```



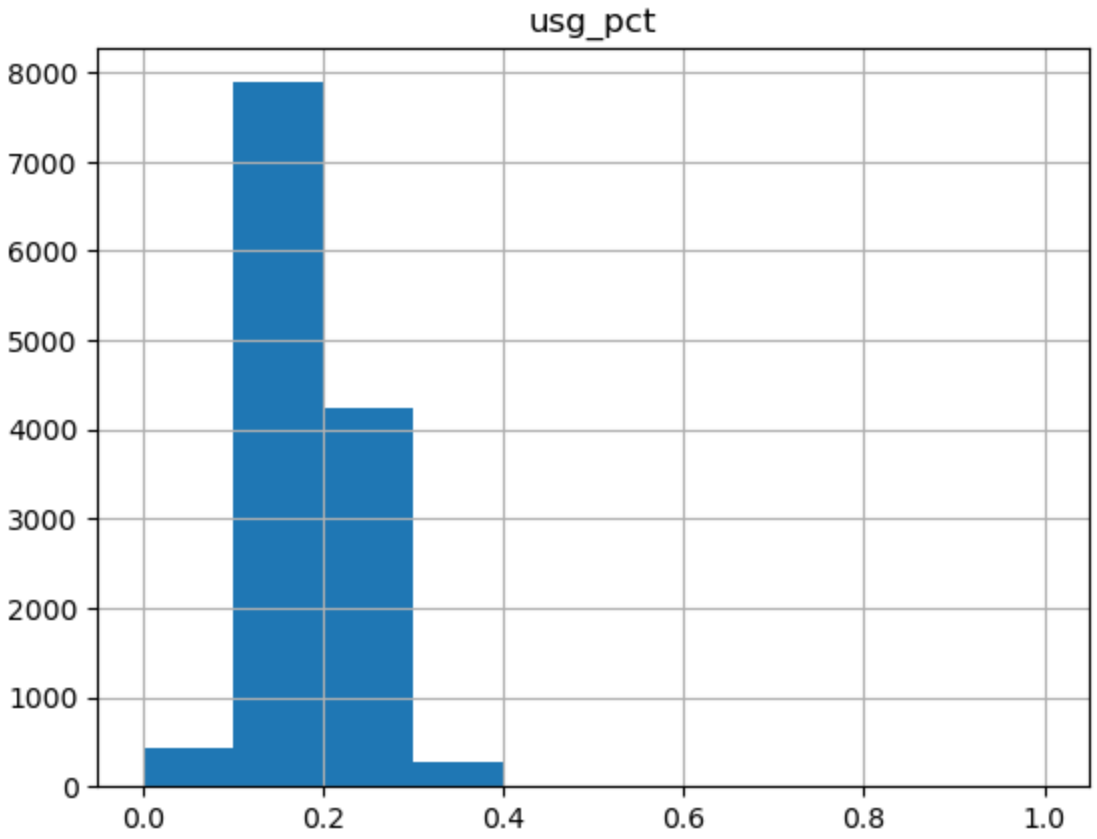
In [32]:
```
# Average Rebounds
df.hist(['avg_reb']);
```



In [36]:
```
# Average Assists
df.hist(['avg_ast']);
```



In [38]:
```
# Usage Percentage
df.hist(['usg_pct']);
```



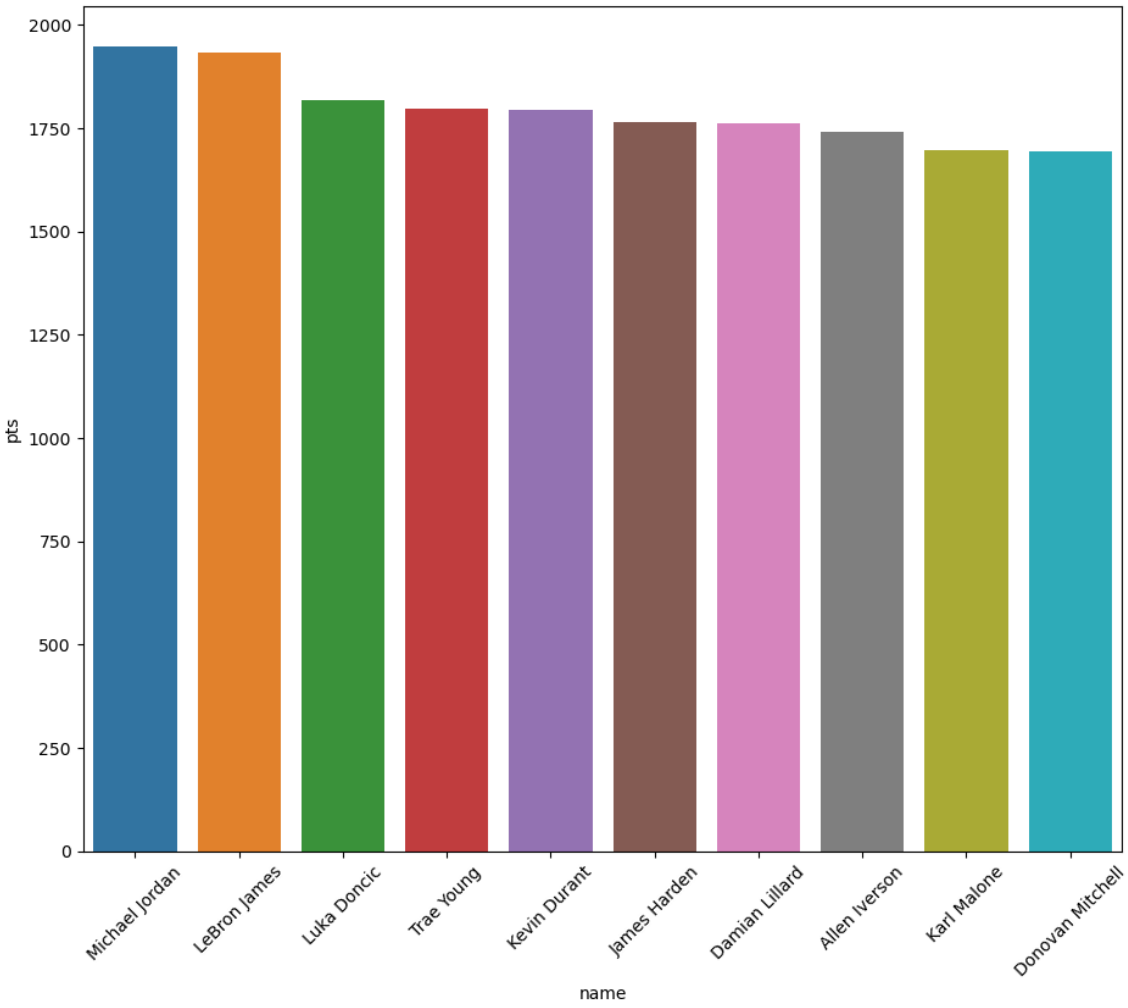
In [8]:
```
# Create a DataFrame containing mean points for the top 10 players
pts_data = pd.DataFrame(df.groupby('name')['pts'].mean().sort_values(
```

In [9]:
```
# Display the resulting DataFrame
pts_data
```

Out[9]:

	name	pts
0	Michael Jordan	1948.500000
1	LeBron James	1932.050000
2	Luka Doncic	1818.800000
3	Trae Young	1797.200000
4	Kevin Durant	1792.400000
5	James Harden	1763.285714
6	Damian Lillard	1762.090909
7	Allen Iverson	1741.142857
8	Karl Malone	1697.875000
9	Donovan Mitchell	1694.166667

In [11]:
```
# Create a bar plot for the top 10 players with the highest mean poin
plt.figure(figsize = (11,9) , dpi = 100)
sns.barplot(data = pts_data , x = 'name' , y = 'pts')
plt.xticks(rotation = 45);
```



In [ ]:
```
```