

NBA Player Performance Analysis

Project Proposal by Ian Sellers & Ryan Brown

Introduction

This project aims to analyze NBA player performance data from 2012 to the present, focusing on field goal percentage, 3-point percentage, free throw percentage, assists and rebounds.

Project Goals

Deeper Analysis of Offensive Efficiency:

- Explore the **balance between scoring and playmaking** by analyzing **offensive contribution (PTS + AST)** for top players and how it changes with **minutes played** and **team dynamics**.
- Investigate whether players with high offensive contribution are also efficient in terms of shooting accuracy and turnovers.

Team Dynamics and Offensive Strategies:

- Focus on **high-scoring teams**. Analyze what specific **strategies or player roles** make these teams consistently strong in scoring, rebounds, and assists.
- Use **advanced metrics** like **pace**, **offensive rating**, and **defensive rating** to understand how team dynamics lead to success.

Shooting Accuracy and Play Style:

- Study players with high shooting accuracy in more depth.
- Analyze **three-point shooting specialists** to see how their scoring patterns differ from players with more traditional shot selection.

Consistency of Top Performers:

- Investigate players who appear across multiple metrics (efficiency, scoring, assists). Analyze how they maintain their performance across seasons and what metrics consistently predict high performance.

Team Evolution Over Time:

- Study **team synergy** by looking at the correlation between assists and points for top teams.

Impact of 3-Point Shooting in Modern Basketball:

- Given the strong correlation between 3-pointers and points, explore how the **rise of three-point shooting** is changing basketball. Look at trends in **3-point attempt rates** over the years and the impact on team success.
- **Methodology**

We will utilize the following Python libraries for data analysis and visualization:

- **Pandas:** For data manipulation and cleaning.
- **NumPy:** For numerical operations.
- **Matplotlib/Seaborn/ggPlot:** For data visualization.

Data Preparation and Exploration

To begin, we'll dive into the dataset by loading it into a Pandas DataFrame. This will provide a structured format for our analysis. We'll then conduct a thorough examination of the data, identifying and addressing any missing values that might skew our results. We'll calculate additional metrics such as win percentage, providing a more comprehensive view of team performance. With a clean and expanded dataset, we'll explore the data's characteristics using descriptive statistics, gaining insights into central tendencies and variability.

We will then focus on visualizing the distribution of shooting percentages, understanding how these values are spread across the league. We'll investigate the relationship between shooting accuracy and team success, exploring whether there's a correlation between these factors. To uncover trends and patterns, we'll examine how shooting performance has evolved over time. Finally, we'll dissect shooting performance based on player positions, searching for positional differences in shooting accuracy.

Scope

The project will focus on analyzing the provided dataset and deriving insights into the relationship between shooting accuracy and team success.

Expected Outcomes

- Visualizations that effectively communicate key findings.
- Potential identification of factors influencing shooting performance.
- A foundation for further exploration of NBA player performance metrics.

- Potentially predict player statistics for future seasons.- involves ML coding which is beyond the scope of the course.

Question 1: Group by teams and calculate average points & Top 10 players by points

Average Points by Team:

TEAM

ATL	357.038339
BKN	332.439628
BOS	332.835655
CHA	434.373913
CHI	390.211321
CLE	379.953488
DAL	352.360759
DEN	388.391304
DET	404.723849
GSW	389.580442
HOU	373.990291
IND	368.388889
LAC	351.793003
LAL	391.738516
MEM	314.759760
MIA	359.263804
MIL	336.054131
MIN	410.857708
NOH	459.647059
NOP	406.161157
NYK	396.583012
OKC	371.555911
ORL	395.247967
PHI	352.461783
PHX	420.568702
POR	362.545455
SAC	477.239819
SAS	371.836667
TOR	364.205047
UTA	360.555172
WAS	385.135714

Name: PTS, dtype: float64

Top 10 Players by Points for the last 10 years:

	PLAYER	PTS
4204	James Harden	2818
677	Kevin Durant	2593
2753	Russell Westbrook	2558
2062	James Harden	2376
2063	Stephen Curry	2375

8049	Luka Doncic	2370
2754	James Harden	2356
4946	James Harden	2335
0	Kevin Durant	2280
8050	Shai Gilgeous-Alexander	2254

Question 2

Identify the top scorers in different seasons and across all seasons.

Top Scorers by Season:

	year	PLAYER	PTS
0	2012-13	Kevin Durant	2280
677	2013-14	Kevin Durant	2593
1362	2014-15	James Harden	2217
2062	2015-16	James Harden	2376
2753	2016-17	Russell Westbrook	2558
3454	2017-18	LeBron James	2251
4204	2018-19	James Harden	2818
4946	2019-20	James Harden	2335
5692	2020-21	Stephen Curry	2015
6471	2021-22	Trae Young	2155
7293	2022-23	Jayson Tatum	2225
8049	2023-24	Luka Doncic	2370

Top Scorer Across All Seasons:

year	2018-19
PLAYER	James Harden
PTS	2818

Name: 4204, dtype: object

Question 3: Teams with the Highest Average Stats

Teams with Highest Average Stats: 2012-13

	PTS	REB	AST
TEAM			
OKC	107.414634	44.829268	21.353659
DEN	106.146341	45.036585	24.414634
MIL	105.182927	45.829268	24.890244
MIA	102.756098	38.451220	23.036585
SAS	102.609756	41.134146	24.987805
SAC	102.353659	40.390244	20.378049
LAL	102.207317	44.804878	22.170732
LAC	101.768293	41.682927	24.121951
BOS	101.658537	39.780488	24.292683
GSW	99.975610	44.439024	22.195122

Teams with Highest Average Stats: 2013-14

	PTS	REB	AST
TEAM			
LAC	116.597561	47.268293	25.707317
GSW	111.439024	47.158537	28.073171
OKC	110.719512	46.658537	22.560976
IND	109.804878	51.634146	23.073171
HOU	107.182927	46.024390	20.682927
MIN	106.841463	44.817073	24.121951
POR	106.743902	46.439024	23.219512
CLE	106.134146	47.878049	23.829268
PHX	105.231707	43.036585	19.060976
DAL	104.841463	40.902439	23.597561

Teams with Highest Average Stats: 2014-15

	PTS	REB	AST
--	-----	-----	-----

TEAM			
OKC	116.512195	48.853659	22.109756
HOU	116.378049	49.512195	26.524390
POR	112.426829	47.097561	23.487805
CLE	111.987805	50.585366	24.121951
GSW	109.951220	44.719512	27.414634
LAC	108.280488	42.829268	25.512195
DAL	106.853659	44.804878	24.573171
TOR	104.170732	41.609756	20.841463
ATL	103.743902	41.256098	25.853659
BKN	102.512195	41.658537	21.621951

Teams with Highest Average Stats: 2015-16

	PTS	REB	AST
TEAM			
GSW	115.134146	46.646341	28.951220
WAS	114.451220	45.134146	26.317073
MIA	113.670732	47.987805	25.329268
OKC	112.219512	49.329268	23.524390
SAS	110.621951	45.390244	26.121951
LAC	107.707317	42.219512	22.817073
POR	107.231707	45.743902	22.182927
CLE	106.792683	45.329268	23.134146
SAC	106.585366	44.243902	24.500000
ATL	105.097561	43.987805	26.012195

Teams with Highest Average Stats: 2016-17

	PTS	REB	AST
TEAM			
HOU	125.402439	44.536585	26.500000
WAS	124.353659	46.902439	28.390244
GSW	120.780488	47.951220	32.146341
CLE	120.170732	47.829268	26.780488
DEN	117.036585	50.170732	27.426829
TOR	115.243902	50.402439	19.573171
MIL	112.841463	45.219512	25.451220
OKC	112.268293	49.780488	20.914634
ATL	111.451220	48.341463	24.780488
LAC	108.670732	43.024390	22.536585

Teams with Highest Average Stats: 2017-18

	PTS	REB	AST
TEAM			
PHI	121.121951	50.219512	28.646341
CLE	118.524390	43.560976	24.048780
LAC	116.463415	43.853659	22.329268
HOU	116.365854	45.487805	22.024390
NOP	115.548780	43.951220	26.817073
GSW	113.463415	43.512195	29.292683
TOR	112.646341	44.378049	24.475610
MIA	112.292683	46.573171	25.024390
OKC	110.268293	46.195122	21.841463
DEN	110.170732	44.402439	24.670732

Teams with Highest Average Stats: 2018-19

	PTS	REB	AST
TEAM			
MIL	128.853659	54.609756	30.219512
PHI	128.390244	53.865854	28.085366
POR	125.573171	53.658537	24.390244
SAC	123.426829	47.841463	25.792683
OKC	118.890244	50.121951	24.182927
LAC	118.780488	47.475610	24.768293
LAL	117.878049	48.609756	27.012195
GSW	117.682927	46.182927	29.426829
TOR	117.646341	45.378049	28.524390
HOU	116.951220	43.060976	22.390244

Teams with Highest Average Stats: 2019-20

	PTS	REB	AST
TEAM			
DAL	112.475610	46.365854	23.914634
PHI	111.109756	44.219512	24.841463
LAC	109.512195	41.829268	21.012195
HOU	108.731707	37.792683	19.536585
MIL	108.682927	47.109756	23.426829
SAC	106.048780	40.475610	21.963415
MIA	105.439024	42.975610	24.914634
LAL	102.365854	41.268293	22.695122

NOP	101.719512	40.829268	23.560976
WAS	101.390244	37.548780	23.463415

Teams with Highest Average Stats: 2020-21

	PTS	REB	AST
TEAM			
MIL	109.268293	44.353659	22.926829
CHI	109.182927	44.536585	25.317073
BKN	107.365854	39.207317	24.658537
DEN	106.487805	41.573171	25.134146
ATL	104.756098	40.463415	21.780488
PHI	103.768293	40.621951	21.963415
LAL	103.487805	43.378049	22.695122
UTA	102.878049	42.512195	20.768293
POR	102.317073	38.646341	18.329268
SAC	101.914634	37.292683	23.256098

Teams with Highest Average Stats: 2021-22

	PTS	REB	AST
TEAM			
CHA	121.597561	47.682927	28.170732
PHX	119.792683	46.792683	28.951220
DEN	118.878049	45.939024	28.463415
UTA	118.731707	47.804878	22.487805
SAC	118.158537	48.000000	23.560976
MIL	116.817073	47.146341	24.292683
MIN	116.512195	44.634146	25.756098
MEM	115.609756	49.146341	25.987805
CHI	114.060976	44.146341	24.036585
TOR	113.768293	47.182927	23.170732

Teams with Highest Average Stats: 2022-23

	PTS	REB	AST
TEAM			
LAC	133.585366	51.853659	31.609756
ATL	129.573171	48.853659	26.097561
LAL	127.951220	48.573171	26.365854
NYK	124.695122	52.914634	26.317073
PHI	122.902439	44.780488	26.402439

TOR	122.780488	49.097561	26.695122
DEN	121.560976	46.109756	29.707317
BOS	121.243902	47.024390	27.134146
SAC	120.890244	42.621951	27.304878
MEM	120.146341	47.585366	26.463415

Teams with Highest Average Stats: 2023-24

	PTS	REB	AST
TEAM			
LAL	127.487805	45.743902	32.963415
OKC	125.621951	46.231707	28.195122
DAL	123.707317	46.548780	26.170732
IND	123.585366	40.792683	30.634146
BOS	123.292683	48.146341	27.707317
MIL	122.353659	46.390244	27.621951
PHI	121.512195	43.500000	27.341463
NYK	120.231707	47.646341	25.646341
PHX	119.829268	45.963415	28.878049
ATL	117.890244	44.475610	26.500000

Question 4: Identify the most efficient players (with a minimum of a certain number of minutes played to ensure relevance)Most Efficient Players Based on Simplified PER (Regular Season):

```

8049 2624 1.355945
8053 2737 1.354768

```

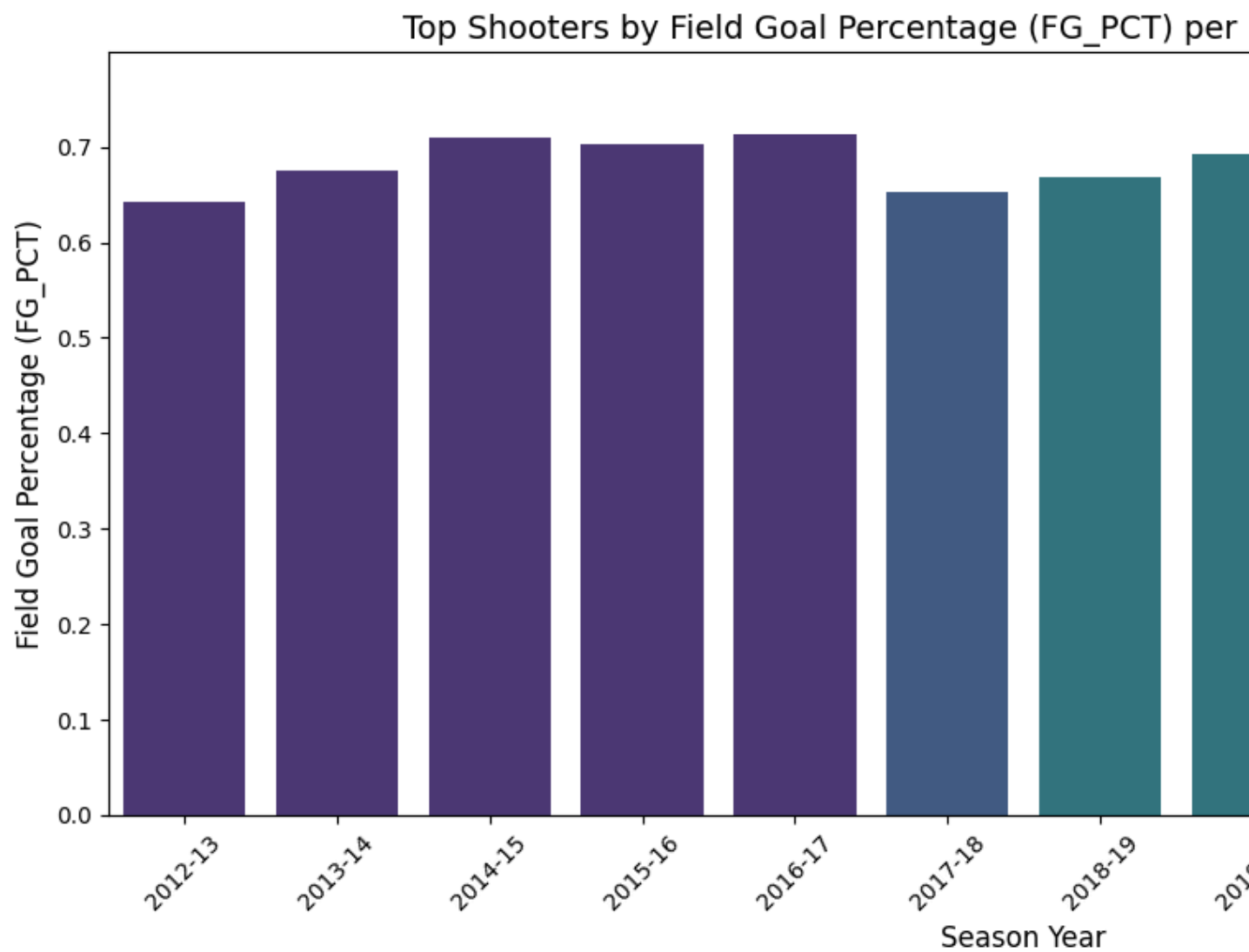
Question 5: Identify the most efficient player for each year

Most Efficient Players for Each Year (Regular%20Season):

	PLAYER	TEAM	year	PTS	REB	AST	STL	BLK	TOV
\									
2	LeBron James	MIA	2012-13	2036	610	551	129	67	226
680	Kevin Love	MIN	2013-14	2010	963	341	59	35	196
1364	Russell Westbrook	OKC	2014-15	1886	488	574	140	14	293
2063	Stephen Curry	GSW	2015-16	2375	430	527	169	15	262
2753	Russell Westbrook	OKC	2016-17	2558	864	840	132	31	438
3455	James Harden	HOU	2017-18	2191	389	630	126	50	315
4210	Giannis Antetokounmpo	MIL	2018-19	1994	898	424	92	110	268
4949	Giannis Antetokounmpo	MIL	2019-20	1857	856	354	61	66	230
5697	Giannis Antetokounmpo	MIL	2020-21	1717	671	357	72	73	207
6475	Nikola Jokic	DEN	2021-22	2004	1019	584	109	63	281
7297	Giannis Antetokounmpo	MIL	2022-23	1959	742	359	52	51	246
8096	Joel Embiid	PHI	2023-24	1353	430	219	46	66	150

	MIN	Efficiency
2	2877	1.100799
680	2797	1.148373
1364	2302	1.220243
2063	2700	1.205185
2753	2802	1.422912
3455	2551	1.203842
4210	2358	1.378287
4949	1917	1.546166
5697	2013	1.332837
6475	2476	1.412763
7297	2024	1.441206
8096	1309	1.500382

Question 6: Identify players with the highest shooting accuracy in each season (normalized for minutes played)



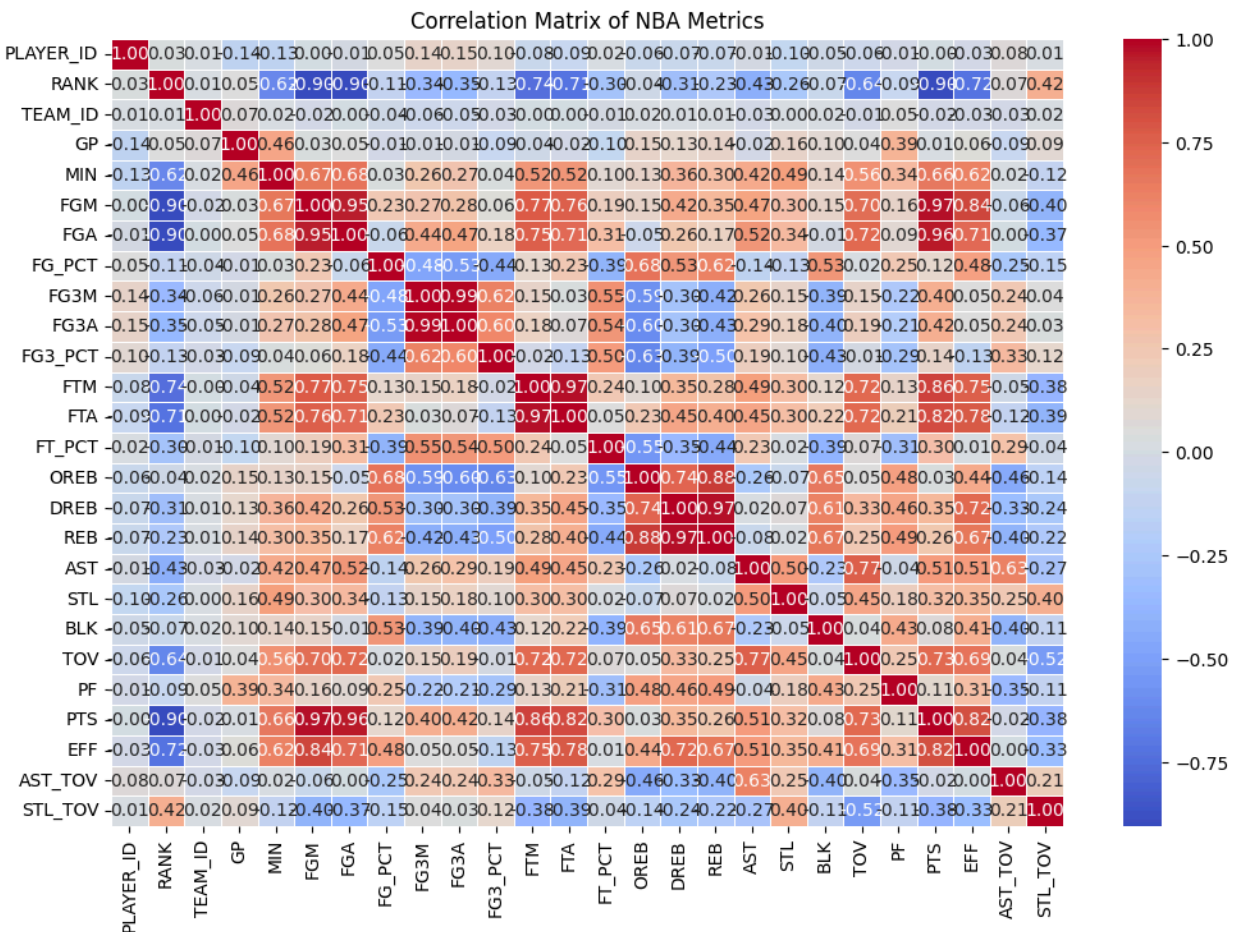
Players with the Highest Shooting Accuracy by Season (Field Goal Percentage):

year	PLAYER TEAM	FG_PCT
2012-13	DeAndre Jordan LAC	0.643
2013-14	DeAndre Jordan LAC	0.676
2014-15	DeAndre Jordan LAC	0.710
2015-16	DeAndre Jordan LAC	0.703
2016-17	DeAndre Jordan LAC	0.714
2017-18	Clint Capela HOU	0.652
2018-19	Rudy Gobert UTA	0.669

2019-20	Rudy Gobert	UTA	0.693
2020-21	Rudy Gobert	UTA	0.675
2021-22	Mitchell Robinson	NYK	0.761
2022-23	Walker Kessler	UTA	0.720
2023-24	Daniel Gafford	DAL	0.725

Question 8:

Calculate the correlation matrix for the numeric columns



The heatmap displays the correlation matrix for the NBA metrics. Here's an analysis of the key observations:

Strong Positive Correlations (Values close to 1):

FGM (Field Goals Made) vs. PTS (Points): Correlation ~ 0.99 . This is expected, as scoring more field goals directly leads to more points. **FGA (Field Goals Attempted) vs. PTS (Points):** Correlation ~ 0.99 . Players who attempt more field goals tend to score more points.

REB (Total Rebounds) vs. OREB/DREB (Offensive/Defensive Rebounds): Correlation ~ 0.99 . Total rebounds are almost perfectly correlated with offensive and defensive rebounds combined.

MIN (Minutes Played) vs. various stats (PTS, FGM, etc.): Players who spend more time on the court tend to accumulate more stats across different categories, such as points, field goals, and assists.

Strong Negative Correlations (Values close to -1):

There aren't many significant negative correlations, though some metrics are inversely related, but with weak correlation values.

Notable Medium-High Correlations (0.7 - 0.9):

AST (Assists) vs. PTS (Points): A relatively strong positive correlation (~ 0.79). Players with more assists are often involved in more scoring plays.

OREB/DREB (Offensive/Defensive Rebounds) vs. REB (Total Rebounds): As expected, these individual components contribute directly to the total rebounds.

Weak or No Correlations (Values close to 0):

TEAM_ID vs. all other stats: As expected, there's no significant correlation between team identifiers and player statistics.

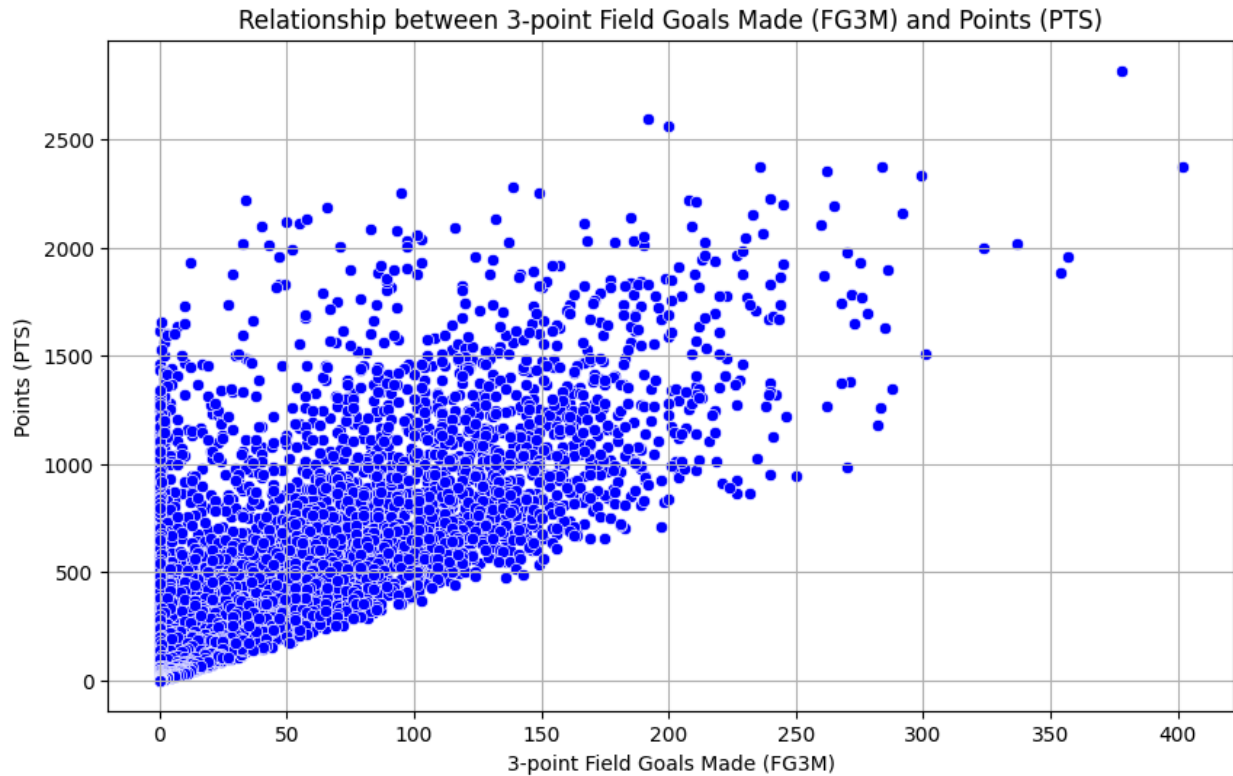
PLAYER_ID vs. all other stats: This is a unique identifier and doesn't correlate with performance metrics.

Conclusion:

The most strongly correlated metrics are field goals attempted/made with points, as well as minutes played with performance metrics like points and assists. Rebounds also show strong internal correlation between offensive, defensive, and total rebounds. These correlations make sense in the context of basketball performance, where key performance indicators tend to cluster.

Question9: Calculate the correlation between 3-point field goals made and total points

Correlation between 3-point Field Goals Made (FG3M) and Points (PTS): 0.77



Top 10 Scorers and Their 3-Point Field Goals Made:

	PLAYER	PTS	FG3M
4204	James Harden	2818	378
677	Kevin Durant	2593	192
2753	Russell Westbrook	2558	200
2062	James Harden	2376	236
2063	Stephen Curry	2375	402
8049	Luka Doncic	2370	284
2754	James Harden	2356	262
4946	James Harden	2335	299
0	Kevin Durant	2280	139
8050	Shai Gilgeous-Alexander	2254	95

Scatter Plot Results

The scatter plot shows a clear positive trend: as the number of 3-point field goals made (FG3M) increases, the total points (PTS) also tend to increase. However, the spread of points indicates that some players score a significant number of points even without relying heavily on 3-pointers, while others who make a lot of 3-pointers also score a lot.

Top 10 Scorers:

James Harden appears multiple times in the top 10, making him one of the top scorers and also one of the top 3-point shooters, especially with 378 and 299 made 3-pointers in different seasons. Stephen Curry is a notable example with 402 made 3-pointers, showing that a significant portion of his points come from 3-point shooting. Other top scorers, like Shai Gilgeous-Alexander and Kevin Durant in one of the entries, score high points but have relatively fewer 3-point field goals made compared to others (95 and 139).

Conclusion:

There is a clear correlation between making 3-point field goals and scoring high points, but not all top scorers are necessarily the best 3-point shooters. Players like Shai Gilgeous-Alexander rely less on 3-pointers, while players like Stephen Curry and James Harden significantly contribute to their points through 3-pointers.

Question 10

To find out which players contribute the most to their team's offense by combining their points (PTS) and assists (AST), we can create a new metric called "Offensive Contribution" (which is the sum of points and assists) and then identify the top players.

Top 10 Players with the Highest Offensive Contribution (Points + Assists):

	PLAYER	TEAM	PTS	AST	Offensive_Contribution
4204	James Harden	HOU	2818	586	3404
2753	Russell Westbrook	OKC	2558	840	3398
2754	James Harden	HOU	2356	907	3263
8049	Luka Doncic	DAL	2370	686	3056
677	Kevin Durant	OKC	2593	445	3038
3454	LeBron James	CLE	2251	747	2998
2062	James Harden	HOU	2376	612	2988
2063	Stephen Curry	GSW	2375	527	2902
6471	Trae Young	ATL	2155	737	2892
3457	Russell Westbrook	OKC	2028	820	2848