

Buckets & Brackets: A Data-Driven Look at NBA Playoff Contenders

Introduction:

As a lifelong sports enthusiast and basketball fan, I knew I wanted my capstone project to focus on something that felt personally exciting, while still challenging me to apply the data analytics tools I've learned. I decided to explore what makes an NBA team a playoff contender, using team-level performance data over the past 25+ seasons.

Rather than tackling a highly complex machine learning problem, I focused on creating a clean, insightful, and flexible dataset that could support both descriptive analysis and potential future modeling. This project tested my ability to work with real-world data, clean and preprocess it, and communicate findings visually. I used Microsoft Excel for data cleaning and Tableau for data visualization.

Data Collection & Source:

To begin, I sourced a dataset from [basketball-reference.com](https://www.basketball-reference.com) that included team-level statistics from the 1996–97 NBA season through the 2022–23 season. This dataset covered regular season game-level stats for all NBA teams over that period.

I downloaded a ZIP file containing two CSV files:

- team_traditional.csv – containing team stats for every regular season and playoff game
- team_traditional_regular_season.csv – used later after filtering for regular season games only

I used Microsoft Excel to begin the cleaning process and Tableau for data visualization.

Business Problem:

NBA teams are constantly striving to make the playoffs, a benchmark that defines success for front offices, coaches, and players alike. With limited resources and high competition, it's essential for teams to understand which statistical factors most contribute to playoff qualification. Making the playoffs also brings increased revenue opportunities through ticket sales, merchandise, TV exposure, and sponsorships — making this not just a basketball goal, but a business one as well.

This project aims to uncover those factors by analyzing historical team performance data, helping identify what differentiates playoff teams from non-playoff teams.

Data Analytics Problem:

Using team-level performance data from the 1996–97 to 2022–23 NBA seasons, this project seeks to answer the question: What statistical patterns are associated with teams that make the playoffs?

I aim to structure, clean, and aggregate the dataset to compare playoff vs non-playoff teams, then visualize trends and insights using Tableau. The goal is to build a foundation for future predictive modeling and performance forecasting.

Data Preprocessing

I began by importing the dataset into Excel, focusing only on regular season games to ensure consistent comparisons across teams. From there, I went through several cleaning and transformation steps to prepare the data for analysis in Tableau. Each step was documented and captured with screenshots to demonstrate the progression of the dataset.

The image displays two Microsoft Excel spreadsheets. The top spreadsheet, titled 'team_traditional', shows a list of basketball games with columns for game ID, date, type, team, opponent, location, and various statistics. The bottom spreadsheet, titled 'team_traditional_regular_season', shows the same data but filtered to include only regular season games, with the 'type' column set to 'regular'.

Game ID	Date	Type	Team	Opponent	Location	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB
22301194	4/14/2024	regular	MIN	PHX		48	106	36	70	51.4	9	26	34.6	25	31	80.6	12	24
22301194	4/14/2024	regular	MIN	PHX		48	125	47	86	54.7	16	29	55.2	15	18	83.3	13	19
22301195	4/14/2024	regular	LAL	NOP		48	124	48	87	55.2	11	29	37.9	17	18	94.4	7	32
22301195	4/14/2024	regular	LAL	NOP		48	108	41	86	47.7	15	37	40.5	11	14	78.6	10	29
22301196	4/14/2024	regular	DAL	OKC		48	86	32	97	33	8	39	20.5	14	17	82.4	12	29
22301196	4/14/2024	regular	DAL	OKC		48	135	54	97	55.7	14	34	41.2	13	16	81.3	11	47
22301197	4/14/2024	regular	SAS	DET		48	95	36	92	39.1	8	33	24.2	15	20	75	11	29
22301197	4/14/2024	regular	SAS	DET		48	123	49	91	53.8	15	34	44.1	10	14	71.4	12	42
22301198	4/14/2024	regular	GSW	UTA		48	123	45	94	47.9	14	43	32.6	19	23	82.6	12	30
22301198	4/14/2024	regular	GSW	UTA		48	116	46	89	51.7	14	34	41.2	10	16	62.5	14	34
22301199	4/14/2024	regular	LAC	HOU		48	116	50	103	48.5	11	38	28.9	5	14	35.7	17	42
22301199	4/14/2024	regular	LAC	HOU		48	105	41	104	39.4	8	37	21.6	15	17	88.2	16	35
22301200	4/14/2024	regular	POR	SAC		48	82	31	98	31.6	7	45	15.6	13	17	76.5	23	31
22301200	4/14/2024	regular	POR	SAC		48	121	43	87	49.4	13	34	38.2	22	27	81.5	13	38
52300121	4/16/2024	playin	LAL	NOP		48	110	35	84	41.7	14	35	40	26	29	89.7	6	35
52300121	4/16/2024	playin	NOP	LAL		48	106	43	93	46.2	9	30	30	11	15	73.3	10	40
52300131	4/16/2024	playin	GSW	SAC		48	94	33	80	41.3	10	32	31.3	18	23	78.3	8	34
52300131	4/16/2024	playin	GSW	SAC		48	118	43	98	43.9	18	39	46.2	14	15	93.3	15	34
52300101	4/17/2024	playin	MIA	PHI		48	104	38	90	42.2	14	39	35.9	14	19	73.7	18	29
52300101	4/17/2024	playin	PHI	MIA		48	105	36	86	41.9	12	35	34.3	21	23	91.3	15	34
52300111	4/17/2024	playin	ATL	CHI		48	116	41	91	45.1	11	37	29.7	23	28	82.1	8	26
52300111	4/17/2024	playin	CHI	ATL		48	131	54	95	56.8	11	26	42.3	12	15	80	10	37
52300201	4/19/2024	playin	CHI	MIA		48	91	35	92	38	13	43	30.2	8	10	80	7	31
52300201	4/19/2024	playin	MIA	CHI		48	112	38	82	46.3	14	33	42.4	22	24	91.7	9	38
52300211	4/19/2024	playin	NOP	SAC		48	105	44	85	51.8	7	19	36.8	10	13	76.9	11	34
52300211	4/19/2024	playin	NOP	SAC		48	98	36	88	40.9	11	41	26.8	15	22	68.2	14	26
42300111	4/20/2024	playoff	NYK	PHI		48	111	36	91	39.6	16	35	45.7	23	28	82.1	23	32
42300111	4/20/2024	playoff	PHI	NYK		48	104	36	81	44.4	12	35	34.3	20	22	90.9	9	24
42300131	4/20/2024	playoff	CLE	ORL		48	97	36	81	44.4	8	30	26.7	17	21	81	10	44
42300131	4/20/2024	playoff	ORL	CLE		48	103	38	86	42.6	8	27	21.6	19	30	83.3	11	29

Game ID	Date	Type	Team	Opponent	Location	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB
29600001	11/1/1996	regular	BOS	CHI		48	98	38	84	45.2	8	20	40	14	23	60.9	14	22
29600001	11/1/1996	regular	CHI	BOS		48	107	42	74	56.8	1	13	7.7	22	32	68.8	8	29
29600002	11/1/1996	regular	CLE	NJN		48	90	34	72	47.2	5	10	50	17	20	85	12	23
29600002	11/1/1996	regular	NJN	CLE		48	77	23	58	39.7	9	18	50	22	33	66.7	11	24
29600003	11/1/1996	regular	MIL	PHI		48	111	38	84	45.2	4	9	44.4	31	44	70.5	19	31
29600003	11/1/1996	regular	PHI	MIL		48	103	37	83	44.6	2	10	20	27	36	75	14	26
29600004	11/1/1996	regular	ORL	WAS		48	92	35	91	38.5	5	23	21.7	17	20	85	22	26
29600004	11/1/1996	regular	WAS	ORL		48	96	35	74	47.3	4	13	30.8	22	30	73.3	11	26
29600005	11/1/1996	regular	ATL	MIA		48	81	25	65	38.5	4	16	25	27	32	84.4	10	30
29600005	11/1/1996	regular	MIA	ATL		48	94	35	78	44.9	10	21	47.6	14	19	73.7	13	27
29600006	11/1/1996	regular	NYK	TOR		48	107	33	68	48.5	4	11	36.4	37	51	72.5	11	33
29600006	11/1/1996	regular	TOR	NYK		48	99	34	78	43.6	12	25	48	19	32	59.4	11	26
29600007	11/1/1996	regular	DET	IND		48	95	32	68	47.1	6	20	30	25	35	71.4	11	28
29600007	11/1/1996	regular	IND	DET		48	89	33	74	44.6	6	21	28.6	17	23	73.9	10	26
29600008	11/1/1996	regular	MIN	SAS		48	82	31	74	41.9	3	8	37.5	17	23	73.9	9	27
29600008	11/1/1996	regular	SAS	MIN		48	78	31	72	43.1	4	13	30.8	12	18	66.7	9	31
29600009	11/1/1996	regular	HOU	SAC		48	96	34	87	39.1	8	25	32	20	31	64.5	17	28
29600009	11/1/1996	regular	SAC	HOU		47	85	32	72	44.4	4	12	33.3	17	23	73.9	8	32
29600010	11/1/1996	regular	DAL	DEN		48	92	35	88	39.8	4	18	22.2	18	26	69.2	20	33
29600010	11/1/1996	regular	DEN	DAL		48	91	28	78	35.9	3	8	37.5	32	43	74.4	18	29
29600011	11/1/1996	regular	SEA	UTA		48	91	30	75	40	5	18	27.8	26	34	76.5	12	34
29600011	11/1/1996	regular	UTA	SEA		48	99	34	78	43.6	8	18	44.4	23	33	69.7	11	32
29600012	11/1/1996	regular	LAL	PHX		48	96	31	63	49.2	6	18	33.3	28	40	70	11	40
29600012	11/1/1996	regular	PHX	LAL		48	82	32	91	35.2	3	13	23.1	15	21	71.4	15	21
29600013	11/1/1996	regular	GSW	LAC		48	85	27	75	36	7	18	38.9	24	29	82.8	12	27
29600013	11/1/1996	regular	LAC	GSW		48	97	41	86	47.7	4	8	50	11	15	73.3	16	33
29600014	11/1/1996	regular	POR	VAN		48	114	45	86	52.3	7	17	41.2	17	27	63	22	32
29600014	11/1/1996	regular	VAN	POR		48	85	34	77	44.2	5	14	35.7	12	19	63.2	9	18
29600015	11/2/1996	regular	CLE	WAS		53	98	37	85	43.5	2	14	14.3	22	26	84.6	17	35

Filtered the dataset to include only regular season games, removing playoff games to ensure a consistent basis for team comparison. (Shown above)

Excel screenshot showing the 'Format Cells' dialog box for the 'date' column. The 'Date' category is selected, and the 'Sample' shows 'date'. The 'General format cells have no specific number format.' message is displayed.

date	team	PTS	FGM	FGA	FG%	3PM
11/1/1996	BOS	98	38	84	45.2	8
11/1/1996	CHI	107	42	74	56.8	1
11/1/1996	CLE	90	34	72	47.2	5
11/1/1996	NIN	77	23	58	39.7	9
11/1/1996	MIL	111	38	84	45.2	4
11/1/1996	PHI	103	37	83	44.6	2
11/1/1996	ORL	92	35	91	38.5	5
11/1/1996	WAS	96	35	74	47.3	4
11/1/1996	ATL	81	25	65	38.5	4
11/1/1996	MIA	94	35	78	44.9	10
11/1/1996	NYK	107	33	68	48.5	4
11/1/1996	TOR	99	34	78	43.6	12
11/1/1996	DET	95	32	68	47.1	6
11/1/1996	IND	89	33	74	44.6	6
11/1/1996	MIN	82	31	74	41.9	3
11/1/1996	SAS	78	31	72	43.1	4
11/1/1996	HOU	96	34	87	39.1	8
11/1/1996	SAC	85	32	72	44.4	4
11/1/1996	DAL	92	35	88	39.8	4
11/1/1996	DEN	91	28	78	35.9	3
11/1/1996	SEA	91	30	75	40	5
11/1/1996	UTA	99	34	78	43.6	8
11/1/1996	LAL	96	31	63	49.2	6
11/1/1996	PHX	82	32	91	35.2	3
11/1/1996	GSW	85	27	75	36	7
11/1/1996	LAC	97	41	86	47.7	4
11/1/1996	POR	114	45	86	52.3	7
11/1/1996	VAN	85	34	77	44.2	5
11/2/1996	CHI	98	37	85	43.5	2

Excel screenshot showing the 'Format Cells' dialog box for the 'date' column. The 'Date' category is selected, and the 'Sample' shows 'date'. The 'General format cells have no specific number format.' message is displayed.

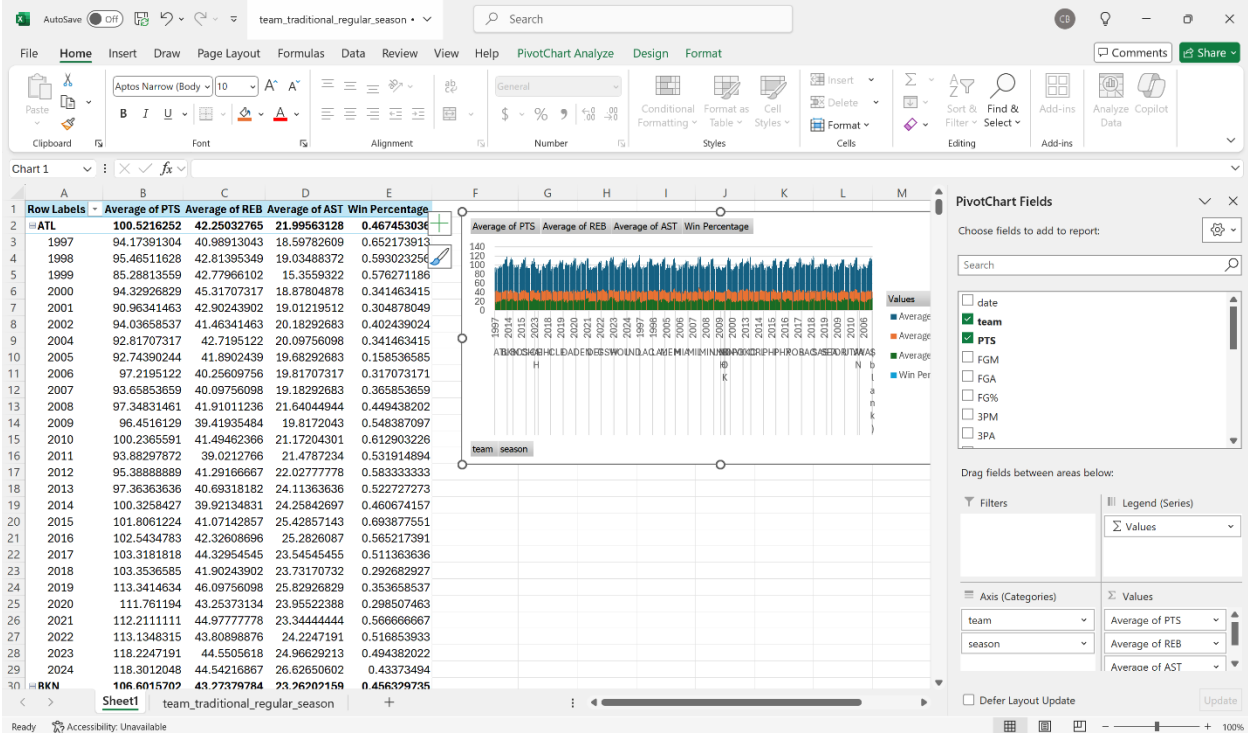
date	team	PTS	FGM	FGA	FG%	3PM
11/1/1996	BOS	98	38	84	45.2	8
11/1/1996	CHI	107	42	74	56.8	1
11/1/1996	CLE	90	34	72	47.2	5
11/1/1996	NIN	77	23	58	39.7	9
11/1/1996	MIL	111	38	84	45.2	4
11/1/1996	PHI	103	37	83	44.6	2
11/1/1996	ORL	92	35	91	38.5	5
11/1/1996	WAS	96	35	74	47.3	4
11/1/1996	ATL	81	25	65	38.5	4
11/1/1996	MIA	94	35	78	44.9	10
11/1/1996	NYK	107	33	68	48.5	4
11/1/1996	TOR	99	34	78	43.6	12
11/1/1996	DET	95	32	68	47.1	6
11/1/1996	IND	89	33	74	44.6	6
11/1/1996	MIN	82	31	74	41.9	3
11/1/1996	SAS	78	31	72	43.1	4
11/1/1996	HOU	96	34	87	39.1	8
11/1/1996	SAC	85	32	72	44.4	4
11/1/1996	DAL	92	35	88	39.8	4
11/1/1996	DEN	91	28	78	35.9	3
11/1/1996	SEA	91	30	75	40	5
11/1/1996	UTA	99	34	78	43.6	8
11/1/1996	LAL	96	31	63	49.2	6
11/1/1996	PHX	82	32	91	35.2	3
11/1/1996	GSW	85	27	75	36	7
11/1/1996	LAC	97	41	86	47.7	4
11/1/1996	POR	114	45	86	52.3	7
11/1/1996	VAN	85	34	77	44.2	5
11/2/1996	CHI	98	37	85	43.5	2

Converted the date column to proper datetime format to enable seasonal grouping and chronological analysis. (Shown above)

Microsoft Excel window showing a dataset of basketball game statistics for the 1995-1996 season. The dataset is titled "team_traditional_regular_season" and contains 30 rows of data. A "No cells were found" error message is displayed in the center of the spreadsheet.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	
1	date	team	PTS	FGM	FGA	FG%	3PM	3PA	3P%	FTM	FTA	FT%	OREB	DREB	REB	AST	TOV	STL	BLK	PF	+/-
2	11/1/1996	BOS	98	38	84	45.2	8	20	40	14	23	60.9	14	22	36	20	18	10	2	33	
3	11/1/1996	CHI	107	42	74	56.8	1	13	7.7	22	32	68.8	8	29	37	28	19	7	8	23	
4	11/1/1996	CLE	90	34	72	47.2	5	10	50	17	20	85	12	23	35	16	15	11	1	24	
5	11/1/1996	NJN	77	23	58	39.7	9	18	50	22	33	66.7	11	24	35	13	22	7	7	19	
6	11/1/1996	MIL	111	38	84	45.2	4	9	44.4	31	44	70.5	19	31	50	21	15	9	7	30	
7	11/1/1996	PHI	103	37	83	44.6	2	10	20	27	36	75	14	26	40	25	14	6	4	29	
8	11/1/1996	ORL	92	35	91	38.5	5	23	21.7	17	20	85	22	26	48	17	17	7	3	22	
9	11/1/1996	WAS	96	35	74	47.3	4	13	30.8	22	33	66.7	11	26	37	25	16	10	10	24	
10	11/1/1996	ATL	81	25	65	38.5	4	16	25	25	33	75.8	10	30	40	13	24	10	7	20	
11	11/1/1996	MIA	94	35	78	44.9	10	21	47.6	25	33	75.8	10	30	40	13	24	10	7	20	
12	11/1/1996	NYK	107	33	68	48.5	4	11	36.4	25	33	75.8	10	30	40	13	24	10	7	20	
13	11/1/1996	TOR	99	34	78	43.6	12	25	48	25	33	75.8	10	30	40	13	24	10	7	20	
14	11/1/1996	DET	95	32	68	47.1	6	20	30	25	33	75.8	10	30	40	13	24	10	7	20	
15	11/1/1996	IND	89	33	74	44.6	6	21	28.6	25	33	75.8	10	30	40	13	24	10	7	20	
16	11/1/1996	MIN	82	31	74	41.9	3	8	37.5	17	23	73.9	9	27	36	18	18	6	12	13	
17	11/1/1996	SAS	78	31	72	43.1	4	13	30.8	12	18	66.7	9	31	40	21	18	10	4	23	
18	11/1/1996	HOU	96	34	87	39.1	8	25	32	20	31	64.5	17	28	45	25	23	20	1	19	
19	11/1/1996	SAC	85	32	72	44.4	4	12	33.3	17	23	73.9	8	32	40	18	26.2	11	8	24	
20	11/1/1996	DAL	92	35	88	39.8	4	18	22.2	18	26	69.2	20	33	53	21	22	13	7	30	
21	11/1/1996	DEN	91	28	78	35.9	3	8	37.5	32	43	74.4	18	29	47	13	21	9	10	21	
22	11/1/1996	SEA	91	30	75	40	5	18	27.8	26	34	76.5	12	34	46	16	12	6	6	26	
23	11/1/1996	UTA	99	34	78	43.6	8	18	44.4	23	33	69.7	11	32	43	26	9	7	6	27	
24	11/1/1996	LAL	96	31	63	49.2	6	18	33.3	28	40	70	11	40	51	26	23	6	8	26	
25	11/1/1996	PHX	82	32	91	35.2	3	13	23.1	15	21	71.4	15	21	36	21	12	11	4	25	
26	11/1/1996	GSW	85	27	75	36	7	18	38.9	24	29	82.8	12	27	39	20	20	8	3	17	
27	11/1/1996	LAC	97	41	86	47.7	4	8	50	11	15	73.3	16	33	49	21	20	7	2	27	
28	11/1/1996	POR	114	45	86	52.3	7	17	41.2	17	27	63	22	32	54	30	17	12	7	20	
29	11/1/1996	VAN	85	34	77	44.2	5	14	35.7	12	19	63.2	9	18	27	23	16	7	5	21	
30	11/2/1996	CLE	88	37	85	43.5	2	14	14.3	22	26	84.6	17	35	52	13	18	7	2	33	

Checked the dataset for missing values and duplicates. No data quality issues were found. (shown above)



Grouped the dataset by team and season to calculate average statistics per season, transforming the data into a team-season level format. (shown above).

Also manually added a ‘made playoffs’ column for each team-season using historical playoff data from Basketball Reference. This served as the target variable for later analysis.

Data Analysis & Visualizations

With the dataset cleaned and labeled, I transitioned to Tableau for visual exploration and analysis. I created three key visualizations that helped uncover patterns in team performance and how those patterns relate to playoff qualification. Each chart focused on a different aspect of performance — scoring, team play, and efficiency — and compared playoff vs non-playoff teams across seasons.

Tableau - Book1

File Data Server Window Help

Final_team_season_dataset

Connections

Final_team_season_dataset

Files

Use Data Interpreter

Data Interpreter might be able to clean your Text file workbook.

Final_team_s...dataset.csv

team_traditional.csv

team_traditio...ar_season.csv

traditional.csv

New Union

New Table Extension

Final_team_season_data...

Connection

Live Extract

Filters

0 Add

Final_team_season_data...

Final_team_season_data... 23 fields 803 rows

100 rows

Name	Final_team_season_dataset.csv	Final_team_season_dataset.csv	Final_team_season_dataset.csv	Final_team_season_dataset.csv	Final_team_season_dataset.csv
Team	Season	PTS	FGM	FGA	FTG
ATL	1.997	94.8049	34.2927		
ATL	1.998	95.8537	35.2073		
ATL	1.999	86.3000	30.7800		
ATL	2.000	94.3293	36.5854		
ATL	2.001	90.9634	35.0732		
ATL	2.002	94.0366	35.3780		
ATL	2.004	92.8171	34.5000		

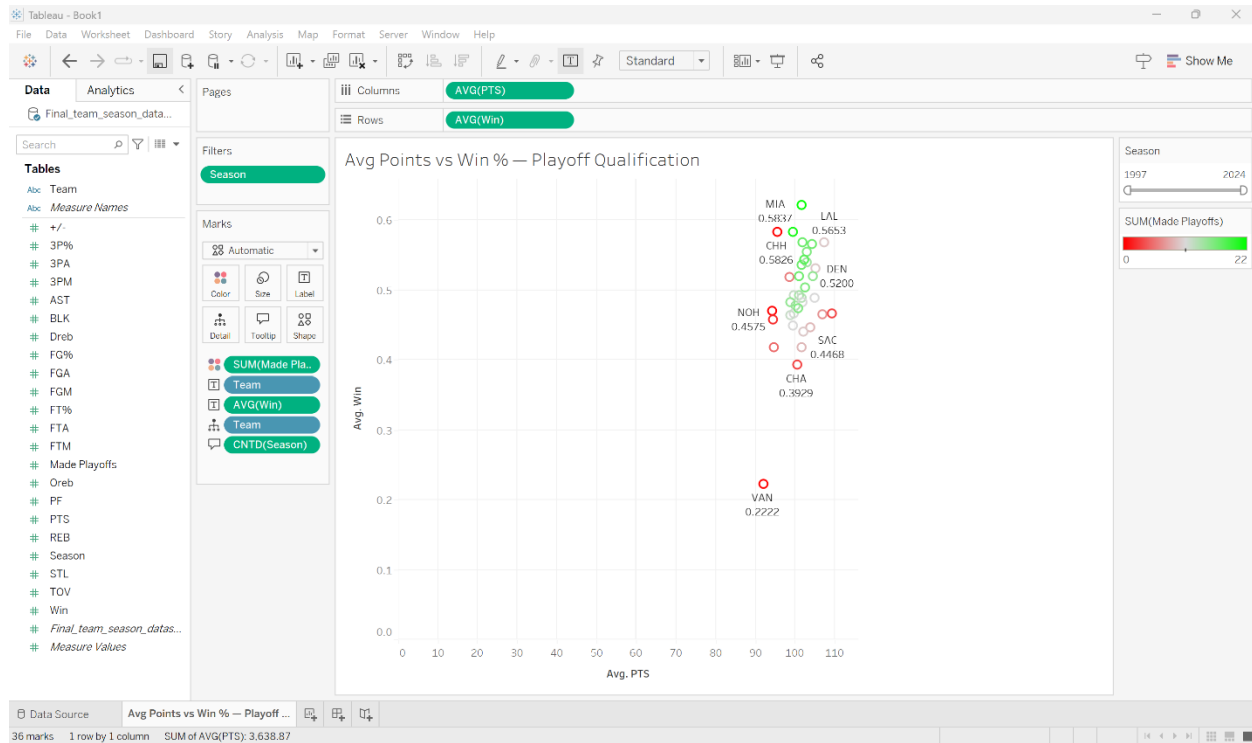
Fields

Type	Field Name	Physical Table	Rem...
ABC	Team	Final_team_season_dataset...	team
#	Season	Final_team_season_dataset...	season
#	PTS	Final_team_season_dataset...	PTS

Data Source

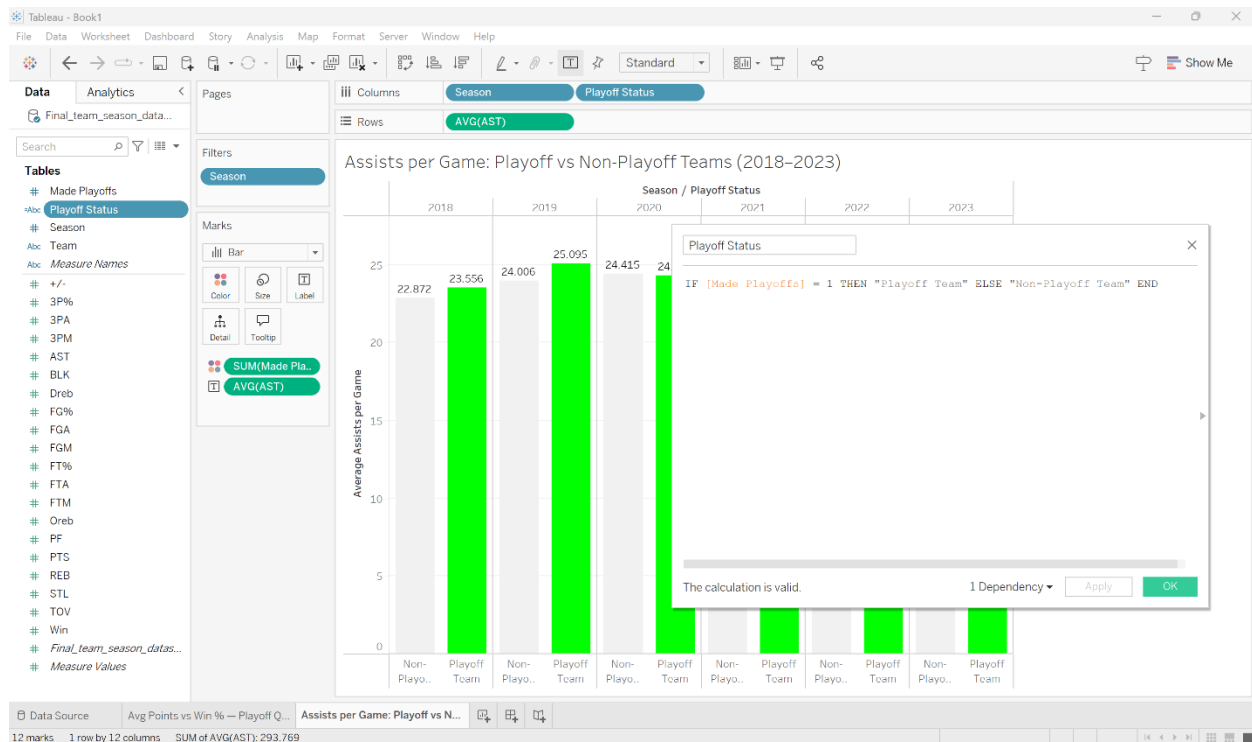
Sheet 1

Loading the dataset into Tableau. Once loaded, I began building visualizations to explore key statistical relationships between team performance and playoff success. (Shown above)

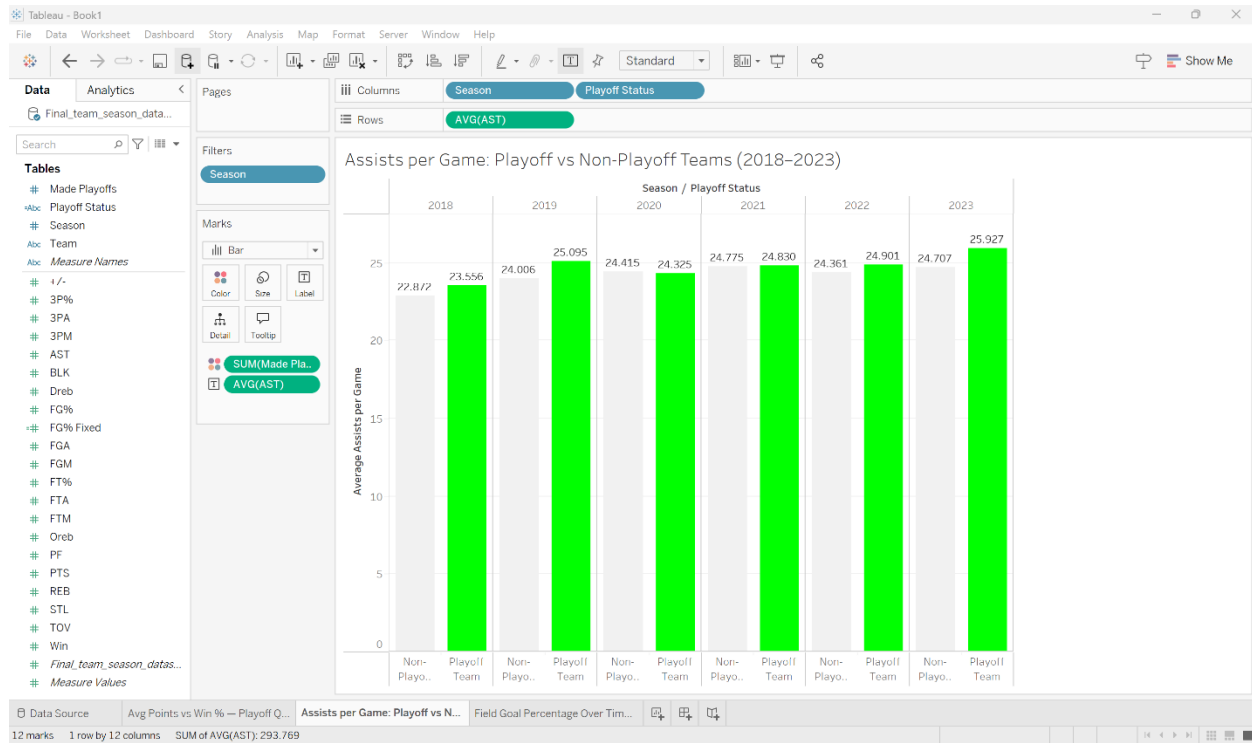


Visualization 1: Avg Points vs Win Percentage (Scatterplot)

Scatterplot comparing average points per game to win percentage, colored by playoff qualification. While higher points often align with more wins, not all high-scoring teams made the playoffs — indicating that scoring alone is not a guaranteed predictor. (shown above)

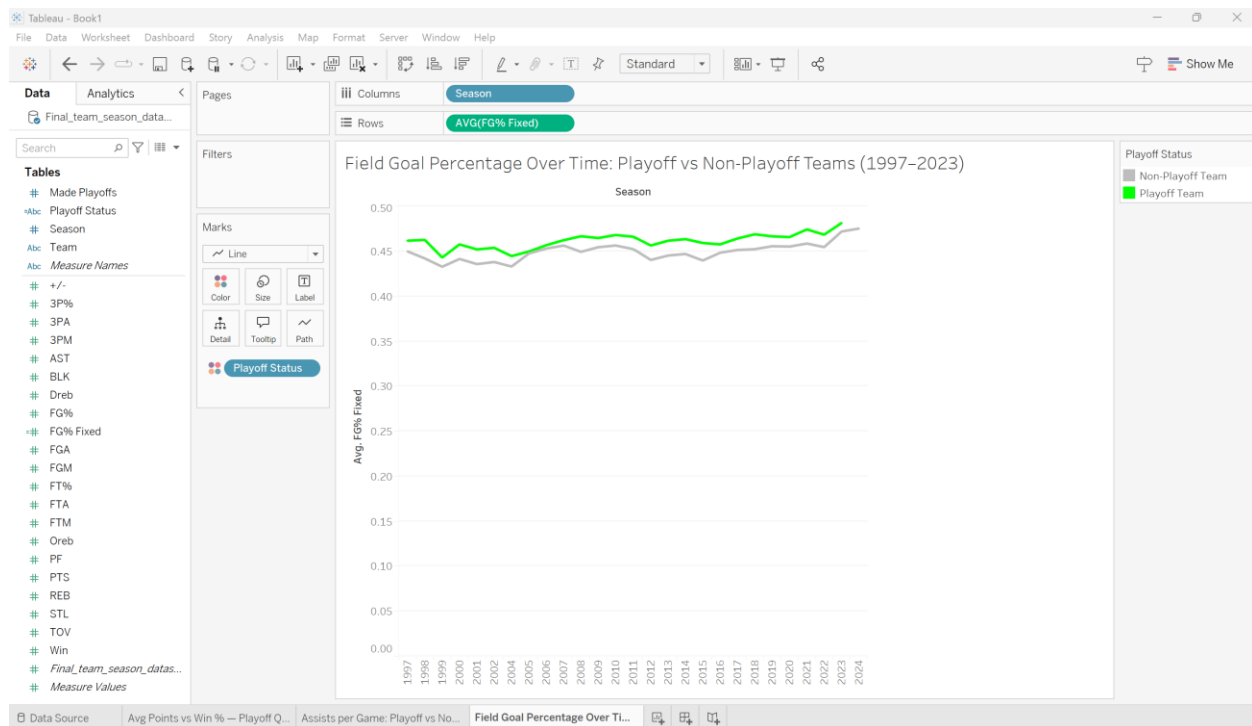


Created a calculated field determining if a team made the playoffs (using the column Made Playoffs from earlier) (shown above)



Visualization 2: Assists per Game — Playoff vs Non-Playoff Teams (Bar Chart)

Bar chart showing average assists per game from 2018 to 2023, split by playoff status. Playoff teams consistently average more assists, highlighting the role of ball movement and team coordination in postseason success. (shown above)



Visualization 3: Field Goal Percentage Over Time (Line Chart)

Line chart showing average field goal percentage for playoff vs non-playoff teams from 1997 to 2023. Playoff teams have maintained a higher and more consistent shooting efficiency across seasons. (Shown above)

Insights & Conclusion

This project highlighted several key trends that help differentiate playoff teams from non-playoff teams. While scoring is important, it isn't the sole factor. Playoff teams consistently show higher assist rates and better field goal percentages — pointing to the importance of team play and offensive efficiency. These findings align with the modern NBA's emphasis on ball movement and shot selection. “Space and pace”, as they like to call it now. Gone are the days of using up most of the clock to get a good shot but rather the ball movement and improvements in player talent and shooting ability has seen percentages increase over the years.

In the future, I'd like to expand this project by incorporating more advanced metrics (like offensive/defensive ratings or pace), player-level data, and even predictive models to forecast playoff outcomes. This version served as a foundation for building a repeatable analysis pipeline and refining my skills in data cleaning, visualization, and storytelling with data.

As a basketball fan and aspiring analyst, this project was a great opportunity to combine my passion with practical data skills — and it's just the beginning.