

# Who is the Best Scorer in the NBA?

David Bednarczyk, Luke Hoffner, Maurizio Hazoury

2023-12-02

«««< HEAD

===== ## Introduction

Who is the best scorer in the current NBA? Is it the person who scores the most points, or is it the most efficient scorer, or a combination of both? As a group, we wanted to determine what player is the best scorer in the NBA, so the easiest way to do this, was through analyzing data. Our group used data from 2022-23 NBA season, as the current season is incomplete at the time of creating this project. Our plan was to find data that included both scoring averages, and scoring totals of the season. We plan on graphing different aspects of the data using stats like points per game, true shooting percentage, scoring totals and more. After graphing the data we plan on being able to make a good estimation on the best scorer in the league based off our criteria. We also wanted observe the position of the top scorers and if that played a part in scoring. Our theory was that position and had scoring had no relationship.

## Our Data

For this project, we used data tables from two websites: ESPN and BasketballReference. We were able to successfully scrape both websites to obtain the data tables we needed, however the data wasn't tidy. There were several things wrong with the data, like player and team being written in once column, accents on letters in one table, but not in the other, unclear column names, and more. After tidying each data table, we combined them into our final table that we will be using to graph the players. Below is the final version of our data table.

```
library(tidyverse)
library(rvest)
```

```
## Warning: package 'rvest' was built under R version 4.3.2
```

```
##
```

```
## Attaching package: 'rvest'
```

```
## The following object is masked from 'package:readr':
```

```
##
```

```
##      guess_encoding
```

```
library(writexl)
```

```
## Warning: package 'writexl' was built under R version 4.3.2
```

```
library(data.table)
```

```
##  
## Attaching package: 'data.table'  
  
## The following objects are masked from 'package:lubridate':  
##  
##     hour, isoweek, mday, minute, month, quarter, second, wday, week,  
##     yday, year  
  
## The following object is masked from 'package:purrr':  
##  
##     transpose  
  
## The following objects are masked from 'package:dplyr':  
##  
##     between, first, last
```

```
library(stringi)
```

```
## Warning: package 'stringi' was built under R version 4.3.2
```

```
library(kableExtra)
```

```
##  
## Attaching package: 'kableExtra'  
  
## The following object is masked from 'package:dplyr':  
##  
##     group_rows
```

```
#Creating Totals table
```

```
#Scraping 22-23 season totals from basketball reference
```

```
nbaTopScorersTotals <- read_html(  
  x = "https://www.basketball-reference.com/leagues/NBA_2023_totals.html#totals_stats::pts"  
) %>%  
  html_elements(css = "table") %>%  
  html_table()  
nbaTopScorersTotals23 <- nbaTopScorersTotals[[1]]
```

```
#Eliminating duplicates from the table
```

```
tot_rows <- nbaTopScorersTotals23$Tm == "TOT"  
duplicates <- duplicated(nbaTopScorersTotals23$Player)  
removed_rows <- duplicates & !tot_rows  
nbaTopScorersTotals23 <- nbaTopScorersTotals23[!removed_rows, ]
```

```
#removing any accents from the players name
```

```
remove_accents <- function(x) {  
  stri_trans_general(x, "Latin-ASCII")  
}
```

```

}
nbaTopScorersTotals23$Player <- remove_accents(nbaTopScorersTotals23$Player)
remove_accents <- function(x) {
  stri_trans_general(x, "Latin-ASCII")
}

#Creating Averages table

#Scraping top scorer's by ppg data table from ESPN
nbaTopScorers <- read_html(
  x = "https://www.espn.com/nba/stats/player/_/season/2023/seasontype/2"
) %>%
  html_elements(css = "table") %>%
  html_table()
nbaTopScorers23 <- bind_cols(nbaTopScorers[[1]], nbaTopScorers[[2]])

#Removing unnecessary columns and rows
nbaTopScorers23 <- nbaTopScorers23[-c(26:50), -c(16:22)]

#Removing team abbreviation from end of player name
nbaTopScorers23$Name[nbaTopScorers23$Name == "Joel EmbiidPHI"] <- "Joel Embiid"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Luka DoncicDAL"] <- "Luka Doncic"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Damian LillardPOR"] <- "Damian Lillard"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Shai Gilgeous-AlexanderOKC"] <- "Shai Gilgeous-Alexander"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Giannis AntetokounmpoMIL"] <- "Giannis Antetokounmpo"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Jayson TatumBOS"] <- "Jayson Tatum"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Stephen CurryGS"] <- "Stephen Curry"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Kevin DurantBKN/PHX"] <- "Kevin Durant"
nbaTopScorers23$Name[nbaTopScorers23$Name == "LeBron JamesLAL"] <- "LeBron James"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Donovan MitchellCLE"] <- "Donovan Mitchell"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Devin BookerPHX"] <- "Devin Booker"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Kyrie IrvingBKN/DAL"] <- "Kyrie Irving"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Jaylen BrownBOS"] <- "Jaylen Brown"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Trae YoungATL"] <- "Trae Young"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Ja MorantMEM"] <- "Ja Morant"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Zion WilliamsonNO"] <- "Zion Williamson"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Anthony DavisLAL"] <- "Anthony Davis"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Lauri MarkkanenUTAH"] <- "Lauri Markkanen"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Julius RandleNY"] <- "Julius Randle"
nbaTopScorers23$Name[nbaTopScorers23$Name == "De'Aaron FoxSAC"] <- "De'Aaron Fox"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Zach LaVineCHI"] <- "Zach LaVine"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Brandon IngramNO"] <- "Brandon Ingram"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Anthony EdwardsMIN"] <- "Anthony Edwards"
nbaTopScorers23$Name[nbaTopScorers23$Name == "DeMar DeRozanCHI"] <- "DeMar DeRozan"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Nikola JokicDEN"] <- "Nikola Jokic"

#Changing first column name to Player
colnames(nbaTopScorers23)[2] <- "Player"

#Extracting name column of nbaTopScorers23 into a vector
top25 <- nbaTopScorers23$Player

#Creating a data frame of all players in the league

```

```

nbaTop25 <- data.frame(Player = nbaTopScorersTotals23$Player)

#Creating empty data frame
top25ScoringTotals <- data.frame()

#Creating a for loop to extract the players from the nbaTopScorers23 from the Totals table
for (string in top25) {
  #Taking the row from the totals table that has the same player as the nbaTopScorers23 table
  top25Rows <- nbaTopScorersTotals23$Player == string
  top25Scorers <- nbaTopScorersTotals23[top25Rows, ]
  #Combines selected columns into a new table
  if (nrow(top25Scorers) > 0) {
    top25ScoringTotals <- rbind(top25ScoringTotals, top25Scorers)
  }
}

#Changing column 5 to Team
colnames(top25ScoringTotals)[5] <- "Team"

#Changing a player's team names
top25ScoringTotals[8, 5] <- "BKN/PHX"
top25ScoringTotals[12, 5] <- "BKN/DAL"

#Removing unnecessary columns
top25ScoringTotals <- top25ScoringTotals[, -c(1, 3:7, 11, 14, 21:29)]

#Merging the data sets
nba23Data <- inner_join(nbaTopScorers23, top25ScoringTotals, by = "Player")

#Fixing column names
colnames(nba23Data)[6] <- "PPG"
colnames(nba23Data)[7] <- "FGM"
colnames(nba23Data)[8] <- "FGA"
colnames(nba23Data)[9] <- "FG%"
colnames(nba23Data)[11] <- "3PA"
colnames(nba23Data)[12] <- "3P%"
colnames(nba23Data)[14] <- "FTA"
colnames(nba23Data)[15] <- "FT%"
colnames(nba23Data)[17] <- "Total FGM"
colnames(nba23Data)[18] <- "Total FGA"
colnames(nba23Data)[19] <- "Total 3PM"
colnames(nba23Data)[20] <- "Total 3PA"
colnames(nba23Data)[21] <- "Total 2PM"
colnames(nba23Data)[22] <- "Total 2PA"
colnames(nba23Data)[23] <- "2P%"
colnames(nba23Data)[25] <- "Total FT"
colnames(nba23Data)[26] <- "Total FTA"
colnames(nba23Data)[27] <- "PTS"

#Turning Columns 4-27 into numeric values and adding True Shooting %
nba23Data[, 4:27] <- lapply(nba23Data[, 4:27], as.numeric)
nba23Data$`TS%` <- (nba23Data$PTS / (2 * (nba23Data$`Total FGA` + 0.44 * nba23Data$`Total FTA`))) * 100
nba23Data$`TS%` <- round(nba23Data$`TS%`, 2)

```

Table 1: Top 25 NBA Players by Points per Game in the 2022-2023 Season

RK	Player	POS	MIN	PPG	FGM	FGA	FG%	3PM	3PA	3P%
1	Joel Embiid	C	34.6	33.1	11.0	20.1	54.8	1.0	3.0	33.0
2	Luka Doncic	PG	36.2	32.4	10.9	22.0	49.6	2.8	8.2	34.2
3	Damian Lillard	PG	36.3	32.2	9.6	20.7	46.3	4.2	11.3	37.1
4	Shai Gilgeous-Alexander	PG	35.5	31.4	10.4	20.3	51.0	0.9	2.5	34.5
5	Giannis Antetokounmpo	PF	32.1	31.1	11.2	20.3	55.3	0.7	2.7	27.5
6	Jayson Tatum	SF	36.9	30.1	9.8	21.1	46.6	3.2	9.3	35.0
7	Stephen Curry	PG	34.7	29.4	10.0	20.2	49.3	4.9	11.4	42.7
8	Kevin Durant	PF	35.6	29.1	10.3	18.3	56.0	2.0	4.9	40.4
9	LeBron James	SF	35.5	28.9	11.1	22.2	50.0	2.2	6.9	32.1
10	Donovan Mitchell	SG	35.8	28.3	10.0	20.6	48.4	3.6	9.3	38.6
11	Devin Booker	SG	34.6	27.8	9.9	20.1	49.4	2.1	6.0	35.1
12	Kyrie Irving	PG	37.4	27.1	9.9	20.1	49.4	3.1	8.3	37.9
13	Jaylen Brown	SG	35.9	26.6	10.1	20.6	49.1	2.4	7.3	33.5
14	Trae Young	PG	34.8	26.2	8.2	19.0	43.0	2.1	6.3	33.5
15	Ja Morant	PG	31.9	26.2	9.3	19.9	46.6	1.5	4.9	30.7
16	Zion Williamson	PF	33.0	26.0	9.8	16.2	60.8	0.2	0.7	36.8
17	Anthony Davis	PF	34.0	25.9	9.7	17.2	56.3	0.3	1.3	25.7
18	Lauri Markkanen	PF	34.4	25.6	8.7	17.3	49.9	3.0	7.7	39.2
19	Julius Randle	PF	35.5	25.1	8.5	18.6	46.0	2.8	8.3	34.3
20	De'Aaron Fox	PG	33.4	25.0	9.3	18.2	51.2	1.6	5.0	32.4
21	Zach LaVine	SG	35.9	24.8	8.7	18.0	48.5	2.6	7.1	37.5
22	Brandon Ingram	SF	34.2	24.7	9.0	18.6	48.4	1.4	3.6	39.0
23	Anthony Edwards	SG	36.0	24.6	8.9	19.5	45.9	2.7	7.3	36.9
24	DeMar DeRozan	SF	36.2	24.5	8.9	17.6	50.4	0.6	1.9	32.4
25	Nikola Jokic	C	33.7	24.5	9.4	14.8	63.2	0.8	2.2	38.3

```

#Subsetting the data for naming purposes on the scatterplot
subsetNba23Data <- nba23Data[nba23Data$PPG * nba23Data$`TS%` > 1900, ]
Jokic <- nba23Data[nba23Data$`TS%` > 70, ]

firstHalfData <- nba23Data[, 1:15]
secondHalfData <- nba23Data[, c(1:3, 16:28)]
firstHalfData <- firstHalfData[, -c(4, 13:15)]

firstHalfDataTable <- firstHalfData %>%
  kable(
    caption = "Top 25 NBA Players by Points per Game in the 2022-2023 Season",
    booktabs = TRUE,
    align = c("l", rep("c", 6))
  ) %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "condensed"),
    font_size = 6
  )

secondHalfDataTable <- secondHalfData %>%
  kable(
    booktabs = TRUE,
    align = c("l", rep("c", 6))
  ) %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "condensed"),
    font_size = 6
  )
firstHalfDataTable

```

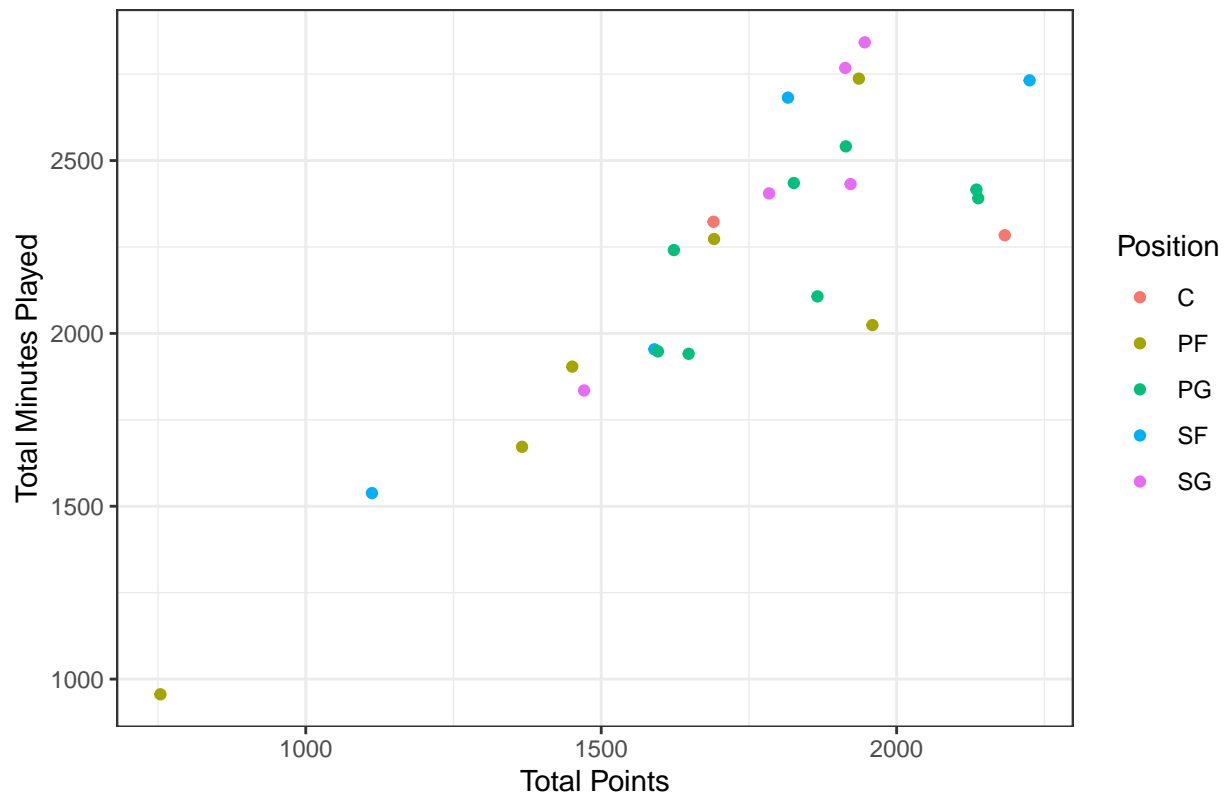
RK	Player	POS	MP	Total FGM	Total FGA	Total 3PM	Total 3PA	Total 2PM	Total 2PA	2P%	eFG%	Total FT
1	Joel Embiid	C	2284	728	1328	66	200	662	1128	0.587	0.573	661
2	Luka Doncic	PG	2391	719	1449	185	541	534	908	0.588	0.560	515
3	Damian Lillard	PG	2107	556	1202	244	658	312	544	0.574	0.564	510
4	Shai Gilgeous-Alexander	PG	2416	704	1381	58	168	646	1213	0.533	0.531	669
5	Giannis Antetokounmpo	PF	2024	707	1278	47	171	660	1107	0.596	0.572	498
6	Jayson Tatum	SF	2732	727	1559	240	686	487	873	0.558	0.543	531
7	Stephen Curry	PG	1941	559	1133	273	639	286	494	0.579	0.614	257
8	Kevin Durant	PF	1672	483	862	93	230	390	632	0.617	0.614	307
9	LeBron James	SF	1954	609	1219	121	377	488	842	0.580	0.549	251
10	Donovan Mitchell	SG	2432	679	1402	245	635	434	767	0.566	0.572	319
11	Devin Booker	SG	1835	527	1067	111	316	416	751	0.554	0.546	306
12	Kyrie Irving	PG	2241	594	1203	188	496	406	707	0.574	0.572	247
13	Jaylen Brown	SG	2405	679	1383	163	487	516	896	0.576	0.550	263
14	Trae Young	PG	2541	597	1390	154	460	443	930	0.476	0.485	566
15	Ja Morant	PG	1948	566	1214	92	300	474	914	0.519	0.504	372
16	Zion Williamson	PF	956	285	469	7	19	278	450	0.618	0.615	177
17	Anthony Davis	PF	1904	542	962	19	74	523	888	0.589	0.573	348
18	Lauri Markkanen	PF	2273	571	1145	200	511	371	634	0.585	0.586	349
19	Julius Randle	PF	2737	658	1432	218	636	440	796	0.553	0.536	402
20	De'Aaron Fox	PG	2435	682	1331	119	367	563	964	0.584	0.557	343
21	Zach LaVine	SG	2768	673	1388	204	544	469	844	0.556	0.558	363
22	Brandon Ingram	SF	1538	404	835	64	164	340	671	0.507	0.522	240
23	Anthony Edwards	SG	2842	707	1541	213	578	494	963	0.513	0.528	319
24	DeMar DeRozan	SF	2682	657	1303	46	142	611	1161	0.526	0.522	456
25	Nikola Jokic	C	2323	646	1022	57	149	589	873	0.675	0.660	341

```
secondHalfDataTable
```

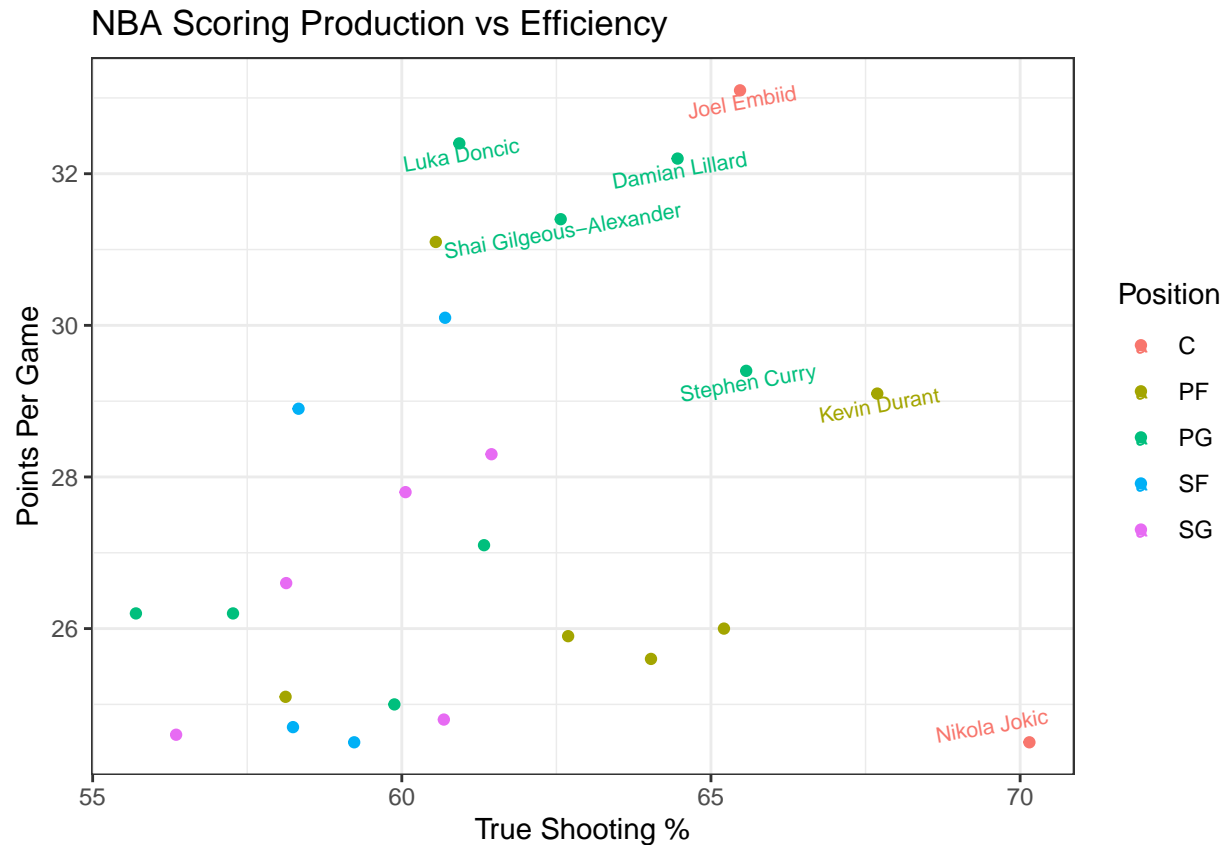
## Graphing the Data

```
nba23Data %>% ggplot(aes(x = PTS,
                        y = MP,
                        color = POS))+
  geom_point()+
  labs(x = "Total Points",
       y = "Total Minutes Played",
       color = "Position",
       title = "Points vs Minutes Played for the NBA's top 25 PPG Scorers")+
  theme_bw()
```

Points vs Minutes Played for the NBA's top 25 PPG Scorers



```
nba23Data %>%
  ggplot(aes(x = `TS%`,
             y = PPG,
             color = POS)) +
  geom_point() +
  labs(x = "True Shooting %",
       y = "Points Per Game",
       color = "Position",
       title = "NBA Scoring Production vs Efficiency") +
  theme_bw() +
  geom_text(data = subsetNba23Data, aes(label = Player), vjust = 1.25, size = 2.75, angle = 10) +
  geom_text(data = Jokic, aes(label = Player), vjust = -1, hjust = .8, size = 2.75, angle = 10)
```



«««< HEAD

## Analyzing the Data

This graph compares player production and efficiency. The “True Shooting %” metric is an advanced basketball statistic that evaluates a player’s shooting efficiency by taking free throws, 2 point shots, and 3 point shots into account. It is widely considered the best way to measure a player’s efficiency. To represent the scoring production of the players, points per game is used. The players in the top right quadrant are the most elite scorers in the NBA, with both high PPG and TS%. This graph also shows us that 2x MVP center Nikola Jokic is the most efficient player in the top 25 PPG scorers. It is still difficult to name one player as the best scorer because while one team may favor slightly higher production, other teams may favor the player with less production and higher efficiency.



## Code Appendix

```
library(dplyr)
library(tidyverse)
library(tidyverse)
library(rvest)
library(writexl)
library(data.table)
library(stringi)
library(kableExtra)

#Creating Totals table

#Scraping 22-23 season totals from basketball reference
nbaTopScorersTotals <- read_html(
  x = "https://www.basketball-reference.com/leagues/NBA_2023_totals.html#totals_stats:pts"
) %>%
  html_elements(css = "table") %>%
  html_table()
nbaTopScorersTotals23 <- nbaTopScorersTotals[[1]]

#Eliminating duplicates from the table
tot_rows <- nbaTopScorersTotals23$Tm == "TOT"
duplicates <- duplicated(nbaTopScorersTotals23$Player)
removed_rows <- duplicates & !tot_rows
nbaTopScorersTotals23 <- nbaTopScorersTotals23[!removed_rows, ]

#removing any accents from the players name
remove_accents <- function(x) {
  stri_trans_general(x, "Latin-ASCII")
}
nbaTopScorersTotals23$Player <- remove_accents(nbaTopScorersTotals23$Player)
remove_accents <- function(x) {
  stri_trans_general(x, "Latin-ASCII")
}

#Creating Averages table

#Scraping top scorer's by ppg data table from ESPN
nbaTopScorers <- read_html(
  x = "https://www.espn.com/nba/stats/player/_/season/2023/seasontype/2"
) %>%
  html_elements(css = "table") %>%
  html_table()
nbaTopScorers23 <- bind_cols(nbaTopScorers[[1]], nbaTopScorers[[2]])

#Removing unnecessary columns and rows
nbaTopScorers23 <- nbaTopScorers23[-c(26:50), -c(16:22)]

#Removing team abbreviation from end of player name
nbaTopScorers23$Name[nbaTopScorers23$Name == "Joel EmbiidPHI"] <- "Joel Embiid"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Luka DoncicDAL"] <- "Luka Doncic"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Damian LillardPOR"] <- "Damian Lillard"
```

```

nbaTopScorers23$Name[nbaTopScorers23$Name == "Shai Gilgeous-AlexanderOKC"] <- "Shai Gilgeous-Alexander"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Giannis AntetokounmpoMIL"] <- "Giannis Antetokounmpo"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Jayson TatumBOS"] <- "Jayson Tatum"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Stephen CurryGS"] <- "Stephen Curry"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Kevin DurantBKN/PHX"] <- "Kevin Durant"
nbaTopScorers23$Name[nbaTopScorers23$Name == "LeBron JamesLAL"] <- "LeBron James"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Donovan MitchellCLE"] <- "Donovan Mitchell"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Devin BookerPHX"] <- "Devin Booker"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Kyrie IrvingBKN/DAL"] <- "Kyrie Irving"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Jaylen BrownBOS"] <- "Jaylen Brown"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Trae YoungATL"] <- "Trae Young"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Ja MorantMEM"] <- "Ja Morant"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Zion WilliamsonNO"] <- "Zion Williamson"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Anthony DavisLAL"] <- "Anthony Davis"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Lauri MarkkanenUTAH"] <- "Lauri Markkanen"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Julius RandleNY"] <- "Julius Randle"
nbaTopScorers23$Name[nbaTopScorers23$Name == "De'Aaron FoxSAC"] <- "De'Aaron Fox"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Zach LaVineCHI"] <- "Zach LaVine"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Brandon IngramNO"] <- "Brandon Ingram"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Anthony EdwardsMIN"] <- "Anthony Edwards"
nbaTopScorers23$Name[nbaTopScorers23$Name == "DeMar DeRozanCHI"] <- "DeMar DeRozan"
nbaTopScorers23$Name[nbaTopScorers23$Name == "Nikola JokicDEN"] <- "Nikola Jokic"

#Changing first column name to Player
colnames(nbaTopScorers23)[2] <- "Player"

#Extracting name column of nbaTopScorers23 into a vector
top25 <- nbaTopScorers23$Player

#Creating a data frame of all players in the league
nbaTop25 <- data.frame(Player = nbaTopScorersTotals23$Player)

#Creating empty data frame
top25ScoringTotals <- data.frame()

#Creating a for loop to extract the players from the nbaTopScorers23 from the Totals table
for (string in top25) {
  #Taking the row from the totals table that has the same player as the nbaTopScorers23 table
  top25Rows <- nbaTopScorersTotals23$Player == string
  top25Scorers <- nbaTopScorersTotals23[top25Rows, ]
  #Combines selected columns into a new table
  if (nrow(top25Scorers) > 0) {
    top25ScoringTotals <- rbind(top25ScoringTotals, top25Scorers)
  }
}

#Changing column 5 to Team
colnames(top25ScoringTotals)[5] <- "Team"

#Changing a player's team names
top25ScoringTotals[8, 5] <- "BKN/PHX"
top25ScoringTotals[12, 5] <- "BKN/DAL"

```

```

#Removing unnecessary columns
top25ScoringTotals <- top25ScoringTotals[, -c(1, 3:7, 11, 14, 21:29)]

#Merging the data sets
nba23Data <- inner_join(nbaTopScorers23, top25ScoringTotals, by = "Player")

#Fixing column names
colnames(nba23Data)[6] <- "PPG"
colnames(nba23Data)[7] <- "FGM"
colnames(nba23Data)[8] <- "FGA"
colnames(nba23Data)[9] <- "FG%"
colnames(nba23Data)[11] <- "3PA"
colnames(nba23Data)[12] <- "3P%"
colnames(nba23Data)[14] <- "FTA"
colnames(nba23Data)[15] <- "FT%"
colnames(nba23Data)[17] <- "Total FGM"
colnames(nba23Data)[18] <- "Total FGA"
colnames(nba23Data)[19] <- "Total 3PM"
colnames(nba23Data)[20] <- "Total 3PA"
colnames(nba23Data)[21] <- "Total 2PM"
colnames(nba23Data)[22] <- "Total 2PA"
colnames(nba23Data)[23] <- "2P%"
colnames(nba23Data)[25] <- "Total FT"
colnames(nba23Data)[26] <- "Total FTA"
colnames(nba23Data)[27] <- "PTS"

#Turning Columns 4-27 into numeric values and adding True Shooting %
nba23Data[, 4:27] <- lapply(nba23Data[, 4:27], as.numeric)
nba23Data$`TS%` <- (nba23Data$PTS / (2 * (nba23Data$`Total FGA` + 0.44 * nba23Data$`Total FTA`))) * 100
nba23Data$`TS%` <- round(nba23Data$`TS%`, 2)

#Subsetting the data for naming purposes on the scatterplot
subsetNba23Data <- nba23Data[nba23Data$PPG * nba23Data$`TS%` > 1900, ]
Jokic <- nba23Data[nba23Data$`TS%` > 70, ]

firstHalfData <- nba23Data[, 1:15]
secondHalfData <- nba23Data[, c(1:3, 16:28)]
firstHalfData <- firstHalfData[, -c(4, 13:15)]

firstHalfDataTable <- firstHalfData %>%
  kable(
    caption = "Top 25 NBA Players by Points per Game in the 2022-2023 Season",
    booktabs = TRUE,
    align = c("l", rep("c", 6))
  ) %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "condensed"),
    font_size = 6
  )

secondHalfDataTable <- secondHalfData %>%
  kable(
    booktabs = TRUE,

```

```

    align = c("l", rep("c", 6))
  ) %>%
  kableExtra::kable_styling(
    bootstrap_options = c("striped", "condensed"),
    font_size = 6
  )
firstHalfDataTable
secondHalfDataTable
nba23Data %>% ggplot(aes(x = PTS,
                        y = MP,
                        color = POS))+

  geom_point()+
  labs(x = "Total Points",
       y = "Total Minutes Played",
       color = "Position",
       title = "Points vs Minutes Played for the NBA's top 25 PPG Scorers")+
  theme_bw()
nba23Data %>%
  ggplot(aes(x = `TS%`,
             y = PPG,
             color = POS)) +
  geom_point()+
  labs(x = "True Shooting %",
       y = "Points Per Game",
       color = "Position",
       title = "NBA Scoring Production vs Efficiency")+
  theme_bw()+
  geom_text(data = subsetNba23Data, aes(label = Player), vjust = 1.25, size = 2.75, angle = 10)+
  geom_text(data = Jokic, aes(label = Player), vjust = -1, hjust = .8, size = 2.75, angle = 10)

```