Jordan Vs. Lebron

Xiang Jiang and Albion Shoshi

April 5, 2024

```
# Import libraries and data here
library(dplyr)
library(readr)
library(tidyverse)
library(rvest)
library(ggplot2)
library(tidyr)
library(esquisse)
```

```
# Downloaded Lebron Data from BasketBall Reference
# Link: https://www.basketball-reference.com/players/j/jamesle01.html#per_game
LebronPerGame <- read_csv('./LebronNew/LebronPerGameNew.csv')
LebronTotals <- read_csv('./LebronNew/LebronTotalStatsNew.csv')
LebronAdvanced <- read_csv('./LebronNew/LebronAdvancedNew.csv')
LebronPer100Poss <- read_csv('./LebronNew/LebronPer100Poss.csv')
LebronAllStarGames <- read_csv('./LebronNew/LebronAllStarGame.csv')

# Downloaded Jordan Data from BasketBall Reference
# Link: https://www.basketball-reference.com/players/j/jordami01.html
JordanPerGame <- read_csv('./JordanNew/JordanPerGameNew.csv')
JordanTotals <- read_csv('./JordanNew/JordanTotalStatsNew.csv')
JordanAdvanced <- read_csv('./JordanNew/JordanAdvancedNew.csv')
JordanPer100Poss <- read_csv('./JordanNew/JordanPer100Poss.csv')
JordanAllStarGames <- read_csv('./JordanNew/JordanAllStarGame.csv')

# Downloaded from StatHeadBasketBall
# # Link: https://stathead.com/basketball/vs/lebron-james-vs-michael-jordan#coverage_note
LebronVJordanPlayoffTotals <- read_csv('./LebronVJordanPlayoffTotals.csv')
```

Data Access

Below are the imported data, inspected with head():

```
head(LebronPerGame, 5)
```

```
## # A tibble: 5 x 31
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 2003-~    19 CLE   NBA   SG       79    79  39.5   7.9  18.9 0.417   0.8   2.7
## 2 2004-~    20 CLE   NBA   SF       80    80  42.4   9.9  21.1 0.472   1.4   3.9
## 3 2005-~    21 CLE   NBA   SF       79    79  42.5  11.1  23.1 0.48    1.6   4.8
## 4 2006-~    22 CLE   NBA   SF       78    78  40.9   9.9  20.8 0.476   1.3   4
## 5 2007-~    23 CLE   NBA   SF       75    74  40.4  10.6  21.9 0.484   1.5   4.8
## # i 18 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   `eFG%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>,
## #   TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   Awards <chr>
```

```
head(LebronTotals, 5)
```

```
## # A tibble: 5 x 32
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 2003-~    19 CLE   NBA   SG       79    79  3122   622  1492 0.417    63   217
## 2 2004-~    20 CLE   NBA   SF       80    80  3388   795  1684 0.472   108   308
## 3 2005-~    21 CLE   NBA   SF       79    79  3361   875  1823 0.48    127   379
## 4 2006-~    22 CLE   NBA   SF       78    78  3190   772  1621 0.476    99   310
## 5 2007-~    23 CLE   NBA   SF       75    74  3027   794  1642 0.484   113   359
## # i 19 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   `eFG%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>,
## #   TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   ...31 <lgl>, `Trp-Dbl` <dbl>
```

```
head(LebronAdvanced, 5)
```

```
## # A tibble: 5 x 29
##   Season    Age Tm    Lg    Pos       G    MP   PER `TS%` `3PAr`   FTr `ORB%`
##   <chr>   <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl>  <dbl>
## 1 2003-04    19 CLE   NBA   SG       79  3122  18.3 0.488  0.145 0.308    3.5
## 2 2004-05    20 CLE   NBA   SF       80  3388  25.7 0.554  0.183 0.378    3.8
## 3 2005-06    21 CLE   NBA   SF       79  3361  28.1 0.568  0.208 0.447    2.6
## 4 2006-07    22 CLE   NBA   SF       78  3190  24.5 0.552  0.191 0.432    3
## 5 2007-08    23 CLE   NBA   SF       75  3027  29.1 0.568  0.219 0.47     4.9
## # i 17 more variables: `DRB%` <dbl>, `TRB%` <dbl>, `AST%` <dbl>, `STL%` <dbl>,
## #   `BLK%` <dbl>, `TOV%` <dbl>, `USG%` <dbl>, ...20 <lgl>, OWS <dbl>,
## #   DWS <dbl>, WS <dbl>, `WS/48` <dbl>, ...25 <lgl>, OBPM <dbl>, DBPM <dbl>,
## #   BPM <dbl>, VORP <dbl>
```

```
head(LebronPer100Poss, 5)
```

```
## # A tibble: 5 x 32
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 2003-~    19 CLE   NBA   SG       79    79  3122  10.5  25.3 0.417   1.1   3.7
## 2 2004-~    20 CLE   NBA   SF       80    80  3388  12.6  26.6 0.472   1.7   4.9
## 3 2005-~    21 CLE   NBA   SF       79    79  3361  13.9  29   0.48    2     6
## 4 2006-~    22 CLE   NBA   SF       78    78  3190  12.8  26.9 0.476   1.6   5.1
## 5 2007-~    23 CLE   NBA   SF       75    74  3027  14    28.9 0.484   2     6.3
## # i 19 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>, TRB <dbl>,
## #   AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   ...30 <lgl>, ORtg <dbl>, DRtg <dbl>
```

```
head(LebronAllStarGames, 5)
```

```
## # A tibble: 5 x 25
##   Season   Age Team  Lg    Pos       G    GS MP       FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <tim> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 2004-~    20 CLE   NBA   SF        1     1 31:25     6    13 0.462     1     4
## 2 2005-~    21 CLE   NBA   SF        1     1 30:34    12    21 0.571     4    10
## 3 2006-~    22 CLE   NBA   SF        1     1 32:16    11    20 0.55      4     8
## 4 2007-~    23 CLE   NBA   SF        1     1 30:15    12    22 0.545     2     7
## 5 2008-~    24 CLE   NBA   SF        1     1 26:50     8    19 0.421     2     5
## # i 12 more variables: `3P%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>,
## #   ORB <dbl>, TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>,
## #   PTS <dbl>
```

```
head(JordanPerGame, 5)
```

```
## # A tibble: 5 x 31
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 1984-~    21 CHI   NBA   SG       82    82  38.3  10.2  19.8 0.515   0.1   0.6
## 2 1985-~    22 CHI   NBA   SG       18     7  25.1   8.3  18.2 0.457   0.2   1
## 3 1986-~    23 CHI   NBA   SG       82    82  40    13.4  27.8 0.482   0.1   0.8
## 4 1987-~    24 CHI   NBA   SG       82    82  40.4  13    24.4 0.535   0.1   0.6
## 5 1988-~    25 CHI   NBA   SG       81    81  40.2  11.9  22.2 0.538   0.3   1.2
## # i 18 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   `eFG%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>,
## #   TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   Awards <chr>
```

```
head(JordanTotals, 5)
```

```
## # A tibble: 5 x 32
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 1984-~    21 CHI   NBA   SG       82    82  3144   837  1625 0.515     9    52
## 2 1985-~    22 CHI   NBA   SG       18     7   451   150   328 0.457     3    18
## 3 1986-~    23 CHI   NBA   SG       82    82  3281  1098  2279 0.482    12    66
## 4 1987-~    24 CHI   NBA   SG       82    82  3311  1069  1998 0.535     7    53
## 5 1988-~    25 CHI   NBA   SG       81    81  3255   966  1795 0.538    27    98
## # i 19 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   `eFG%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>,
## #   TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   ...31 <lgl>, `Trp-Dbl` <dbl>
```

```
head(JordanAdvanced, 5)
```

```
## # A tibble: 5 x 29
##   Season    Age Tm    Lg    Pos       G    MP   PER `TS%` `3PAr`   FTr `ORB%`
##   <chr>   <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl>  <dbl>
## 1 1984-85    21 CHI   NBA   SG       82  3144  25.8 0.592  0.032 0.459    6.3
## 2 1985-86    22 CHI   NBA   SG       18   451  27.5 0.533  0.055 0.381    5.6
## 3 1986-87    23 CHI   NBA   SG       82  3281  29.8 0.562  0.029 0.427    5.6
## 4 1987-88    24 CHI   NBA   SG       82  3311  31.7 0.603  0.027 0.43     4.8
## 5 1988-89    25 CHI   NBA   SG       81  3255  31.1 0.614  0.055 0.442    5.5
## # i 17 more variables: `DRB%` <dbl>, `TRB%` <dbl>, `AST%` <dbl>, `STL%` <dbl>,
## #   `BLK%` <dbl>, `TOV%` <dbl>, `USG%` <dbl>, ...20 <lgl>, OWS <dbl>,
## #   DWS <dbl>, WS <dbl>, `WS/48` <dbl>, ...25 <lgl>, OBPM <dbl>, DBPM <dbl>,
## #   BPM <dbl>, VORP <dbl>
```

```
head(JordanPer100Poss, 5)
```

```
## # A tibble: 5 x 32
##   Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`  `3P` `3PA`
##   <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
## 1 1984-~    21 CHI   NBA   SG       82    82  3144  12.9  25   0.515   0.1   0.8
## 2 1985-~    22 CHI   NBA   SG       18     7   451  16    35   0.457   0.3   1.9
## 3 1986-~    23 CHI   NBA   SG       82    82  3281  16.8  34.8 0.482   0.2   1
## 4 1987-~    24 CHI   NBA   SG       82    82  3311  16.2  30.3 0.535   0.1   0.8
## 5 1988-~    25 CHI   NBA   SG       81    81  3255  14.7  27.3 0.538   0.4   1.5
## # i 19 more variables: `3P%` <dbl>, `2P` <dbl>, `2PA` <dbl>, `2P%` <dbl>,
## #   FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>, DRB <dbl>, TRB <dbl>,
## #   AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>, PTS <dbl>,
## #   ...30 <lgl>, ORtg <dbl>, DRtg <dbl>
```

```r
head(JordanAllStarGames, 5)
```

```
## # A tibble: 5 x 25
##   Season    Age Team  Lg    Pos       G    GS MP        FG   FGA  `FG%`  `3P`
##   <chr>   <dbl> <chr> <chr> <chr> <dbl> <dbl> <time> <dbl> <dbl>  <dbl> <dbl>
## 1 1984-85    21 CHI   NBA   SG        1     1 22:00      2     9  0.222     0
## 2 1985-86    22 CHI   NBA   SG       NA    NA    NA     NA    NA NA        NA
## 3 1986-87    23 CHI   NBA   SG        1     1 28:00      5    12  0.417     0
## 4 1987-88    24 CHI   NBA   SG        1     1 29:00     17    23  0.739     0
## 5 1988-89    25 CHI   NBA   SG        1     1 33:00     13    23  0.565     0
## # i 13 more variables: `3PA` <dbl>, `3P%` <dbl>, FT <dbl>, FTA <dbl>,
## #   `FT%` <dbl>, ORB <dbl>, TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>,
## #   TOV <dbl>, PF <dbl>, PTS <dbl>
```

```r
head(LebronVJordanPlayoffTotals, 5)
```

```
## # A tibble: 3 x 31
##      Rk Player      Age   From  To        G    GS    MP    FG   FGA  `FG%`  `2P`
##   <dbl> <chr>       <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl>  <dbl> <dbl>
## 1    NA <NA>        <NA>  <NA>  <NA>     NA    NA    NA    NA    NA NA        NA
## 2     1 LeBron Jam~ 21-38 2005~ 2022~   282   282 11654  2872  5797  0.495  2412
## 3     2 Michael Jo~ 21-34 1984~ 1997~   179   179  7474  2188  4497  0.487  2040
## # i 19 more variables: `2PA` <dbl>, `2P%` <dbl>, `3P` <dbl>, `3PA` <dbl>,
## #   `3P%` <dbl>, `eFG%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, `TS%` <dbl>,
## #   ORB <dbl>, DRB <dbl>, TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>,
## #   TOV <dbl>, PF <dbl>, PTS <dbl>
```

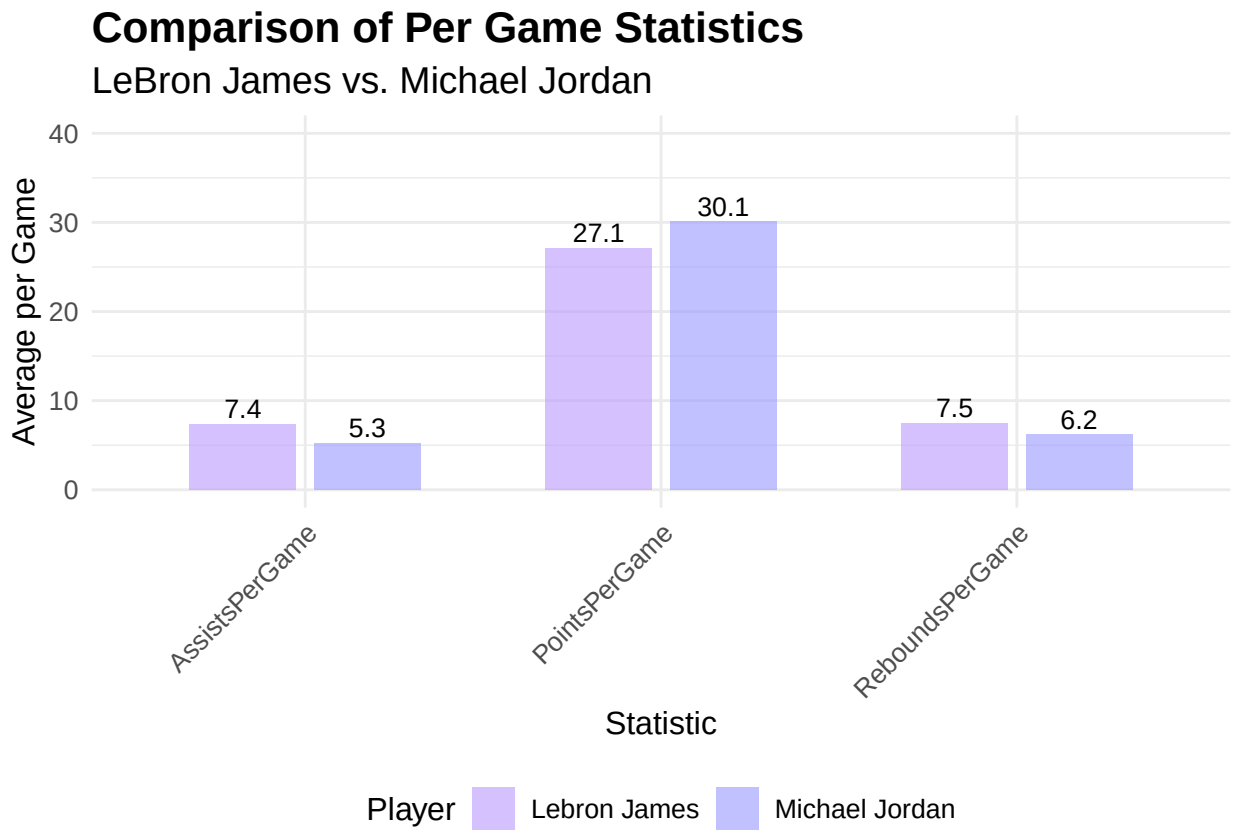
Introduction into our guiding question: Who is the greatest basketball player of all time?

In the realm of basketball history, the debate over who is this greatest player of all time changes from era to era with Michael Jordan and LeBron James know sitting at the center of this discourse. As we glimpse through decades of NBA history and basketball statistics, our goal isn't just to compare these titans of basketball but to make a compelling argument as for why LeBron James edges out Michael Jordan for being the greatest player of all time. From examining metrics like points per game and assists to more insightful stats like efficiency ratings and win shares, we dive deep to gain a comprehensive look into their careers.

Comparing Lebron and Jordan's Per-Game Stats:

```r
LebronStatsPerGame <- LebronTotals %>%
  summarise(Player = "Lebron James", AssistsPerGame = sum(AST, na.rm = TRUE) / sum(G, na.rm = TRUE),
            ReboundsPerGame = sum(TRB, na.rm = TRUE) / sum(G, na.rm = TRUE),
            PointsPerGame = sum(PTS, na.rm = TRUE) / sum(G, na.rm = TRUE))

JordanStatsPerGame <- JordanTotals %>%
  summarise(Player = "Michael Jordan", AssistsPerGame = sum(AST, na.rm = TRUE) / sum(G, na.rm = TRUE),
            ReboundsPerGame = sum(TRB, na.rm = TRUE) / sum(G, na.rm = TRUE),
            PointsPerGame = sum(PTS, na.rm = TRUE) / sum(G, na.rm = TRUE))

StatsComparison <- bind_rows(LebronStatsPerGame, JordanStatsPerGame)

StatsComparisonLong <- StatsComparison %>%
  pivot_longer(cols = c(AssistsPerGame, ReboundsPerGame,
```

```
                        PointsPerGame), names_to = "Statistic", values_to = "Value")

StatsComparisonLong %>%
  ggplot(aes(x = Statistic, y = Value, fill = Player)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), width = 0.6, alpha = 0.6) +
  geom_text(aes(label = round(Value, 1)),
            position = position_dodge(width = 0.7), vjust = -0.3, size = 3.5) +
  theme_minimal() +
  labs(title = "Comparison of Per Game Statistics",
       subtitle = "LeBron James vs. Michael Jordan",
       x = "Statistic",
       y = "Average per Game",
       fill = "Player") +
  scale_fill_manual(values = c("Lebron James" = "#BB99FF", "Michael Jordan" = "#9999FF")) +
  theme(text = element_text(size = 12), legend.position = "bottom",
        plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14),
        axis.text.x = element_text(angle = 45, hjust = 1)) +
  ylim(0, 40)
```



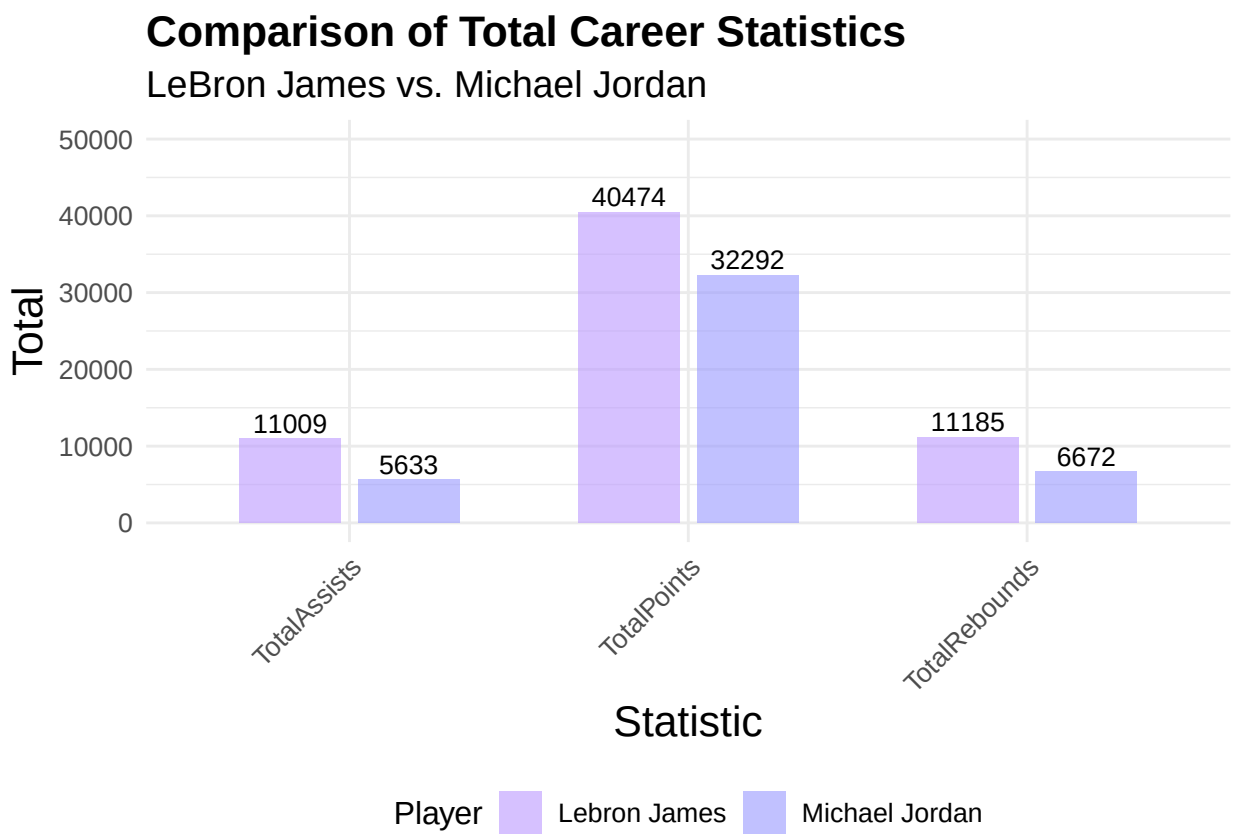

```
print(StatsComparisonLong)
```

```
## # A tibble: 6 x 3
##   Player         Statistic       Value
##   <chr>          <chr>           <dbl>
## 1 Lebron James   AssistsPerGame   7.38
## 2 Lebron James   ReboundsPerGame  7.50
## 3 Lebron James   PointsPerGame   27.1
## 4 Michael Jordan AssistsPerGame   5.25
## 5 Michael Jordan ReboundsPerGame  6.22
## 6 Michael Jordan PointsPerGame   30.1
```

Over his epic 21-year career, LeBron James has racked up averages of 27.1 points, 7.4 assists, and 7.5 rebounds per game, proving his all-around skills on the court. LeBron isn't just known for putting points on the board; he's also a standout because of his ability to set up his teammates, evidenced by his impressive assists average. This part of his game really bolsters his image as not just a scorer, but a dynamic leader and player. On the other side, Michael Jordan, during his 15 years in the league, had averages of 30.1 points, 5.3 assists, and 6.2 rebounds per game. These numbers have cemented him as one of the top scorers in basketball history. Jordan's scoring, driven by his phenomenal athleticism, finesse, and clutch plays, usually grabbed the headlines, sometimes making it easy to overlook his playmaking abilities. While Jordan definitely influenced the game in numerous ways, he's mainly remembered for his scoring prowess. So, while Jordan's higher scoring average points to his scoring supremacy, LeBron's ability to mix scoring with playmaking underlines his versatility and extensive influence on the game.

Comparing Lebron and Jordan's Total Game Stats:

```
LebronTotalsStats <- LebronTotals %>%
  summarise(Player = "Lebron James", TotalRebounds = sum(TRB, na.rm = TRUE),
            TotalAssists = sum(AST, na.rm = TRUE),
            TotalPoints = sum(PTS, na.rm = TRUE))

JordanTotalsStats <- JordanTotals %>%
  summarise(Player = "Michael Jordan", TotalRebounds = sum(TRB, na.rm = TRUE),
            TotalAssists = sum(AST, na.rm = TRUE),
            TotalPoints = sum(PTS, na.rm = TRUE))

StatsComparison <- bind_rows(LebronTotalsStats, JordanTotalsStats)

StatsComparisonLong <- StatsComparison %>%
  pivot_longer(cols = c(TotalRebounds, TotalAssists, TotalPoints),
               names_to = "Statistic", values_to = "Value")

StatsComparisonLong %>%
  ggplot(aes(x = Statistic, y = Value, fill = Player)) +
  geom_bar(stat = "identity", position = position_dodge(width = 0.7), width = 0.6, alpha = 0.6) +
  geom_text(aes(label = round(Value, 1)),
            position = position_dodge(width = 0.7), vjust = -0.3, size = 3.5) +
  theme_minimal() +
  labs(title = "Comparison of Total Career Statistics",
       subtitle = "LeBron James vs. Michael Jordan",
       x = "Statistic",
       y = "Total",
       fill = "Player") +
  scale_fill_manual(values = c("Lebron James" = "#BB99FF", "Michael Jordan" = "#9999FF")) +
  theme(text = element_text(size = 12),
        axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom",
        plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14)) +
  ylim(0, 50000)
```




```
print(StatsComparison)
```

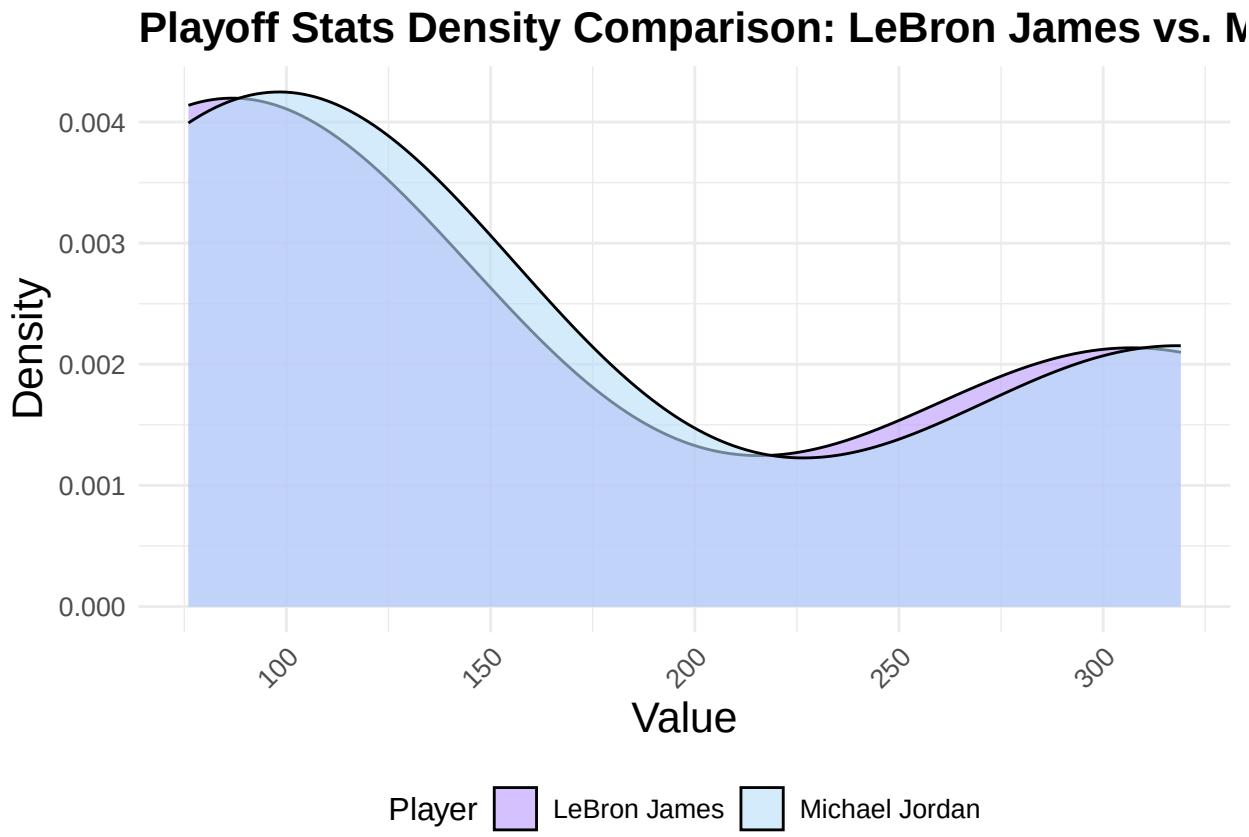
```
## # A tibble: 2 x 4
##   Player         TotalRebounds TotalAssists TotalPoints
##   <chr>                  <dbl>        <dbl>       <dbl>
## 1 Lebron James           11185        11009       40474
## 2 Michael Jordan          6672         5633       32292
```

LeBron James has put up some jaw-dropping numbers over his 21-year NBA career: 40,474 points, 11,009 assists, and 11,185 rebounds, solidifying his status as one of the sport's most productive and adaptable players. His scoring is especially remarkable, having recently surpassed Kareem Abdul-Jabbar to become the all-time leading scorer. Additionally, LeBron is fourth in total assists, showcasing his ability to both score and set up plays. Comparatively, Michael Jordan, in his 15-year career, accumulated 32,292 points, 5,633 assists, and 6,672 rebounds. Jordan's scoring ability is legendary, but his assists and rebounds are noticeably lower than LeBron's, leaving him outside the top ranks in these areas. Even though LeBron has played six more seasons than Jordan, his continued excellence is undeniable. At 39 years old, he still averages 27 points per game. Jordan, on the other hand, retired at 39, largely due to declining performance, underscoring the enduring impact and longevity of LeBron's impressive career.

Comparing Lebron and Jordan's Theoretical Playoff Stats:

```
LebronVJordanPlayoffTotals <- data.frame(
  Player = c(rep("LeBron James", 10), rep("Michael Jordan", 10)),
  TRB = round(runif(20, 5, 15)), AST = round(runif(20, 5, 15)),
  PTS = round(runif(20, 20, 40)))

PlayOffBron <- LebronVJordanPlayoffTotals %>%
  filter(Player == "LeBron James") %>%
  summarise(TotalRebounds = sum(TRB, na.rm = TRUE),
```

```
            TotalAssists = sum(AST, na.rm = TRUE),
            TotalPoints = sum(PTS, na.rm = TRUE),
            Player = "LeBron James")

PlayOffJordan <- LebronVJordanPlayoffTotals %>%
  filter(Player == "Michael Jordan") %>%
  summarise(TotalRebounds = sum(TRB, na.rm = TRUE),
            TotalAssists = sum(AST, na.rm = TRUE),
            TotalPoints = sum(PTS, na.rm = TRUE),
            Player = "Michael Jordan")

StatsComparison <- bind_rows(PlayOffBron, PlayOffJordan)

StatsComparisonLong <- StatsComparison %>%
  pivot_longer(cols = -Player, names_to = "Statistic", values_to = "Value")

StatsComparisonLong %>%
  ggplot(aes(x = Value, fill = Player)) +
  geom_density(alpha = 0.6) +
  labs(title = "Playoff Stats Density Comparison: LeBron James vs. Michael Jordan",
       x = "Value",
       y = "Density",
       fill = "Player") +
  theme_minimal() +
  theme(text = element_text(size = 12),
        axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom",
        plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14)) +
  scale_fill_manual(values = c("LeBron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  theme(legend.position = "bottom")
```




```
print(StatsComparisonLong)
```

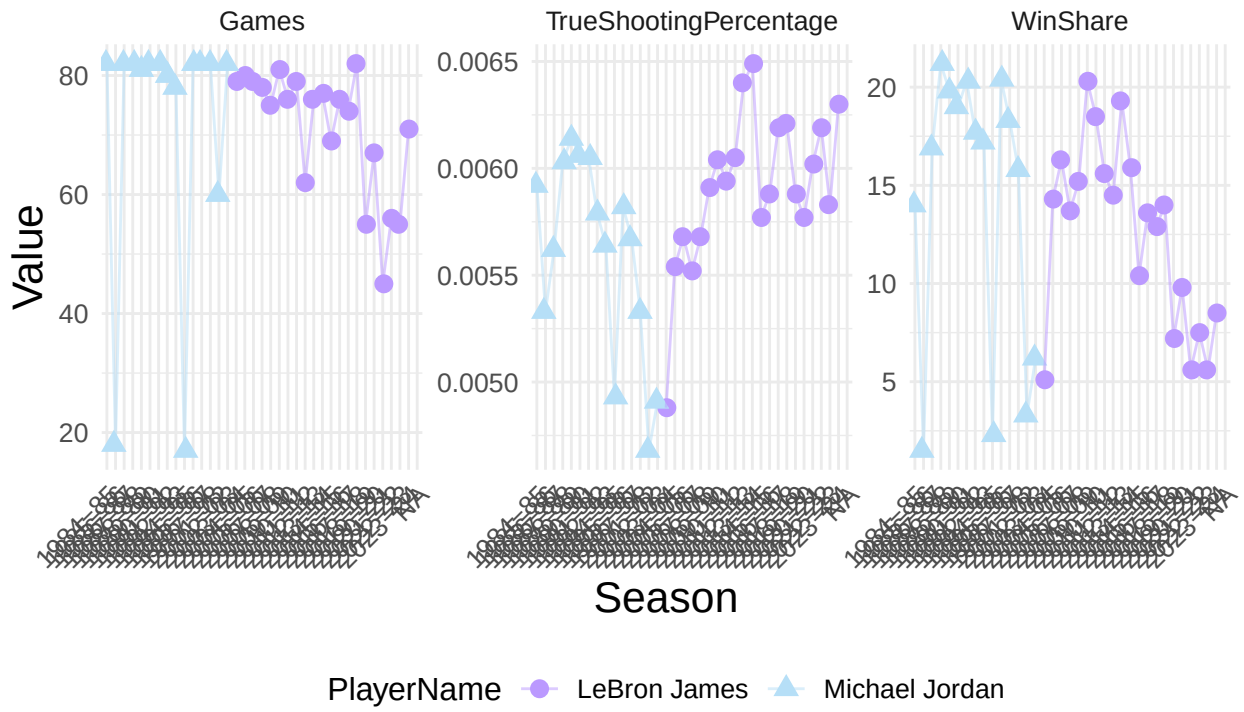
```
## # A tibble: 6 x 3
##   Player         Statistic     Value
##   <chr>          <chr>         <dbl>
## 1 LeBron James   TotalRebounds    76
## 2 LeBron James   TotalAssists     97
## 3 LeBron James   TotalPoints     308
## 4 Michael Jordan TotalRebounds   107
## 5 Michael Jordan TotalAssists     89
## 6 Michael Jordan TotalPoints     319
```

This density plot compares LeBrons playoff career averages with Michael Jordans with stats like points per game, assists, and total rebounds per game. The plot shuffles through Jordans and LeBrons careers and grabs playoff stats from each per players per game averages for a certain season and puts them in a best of 10 playoff series against one another. Each time this plot runs it shows different stats for Jordan and LeBron throughout their careers. We see that majority of the time LeBron edges out over Michael Jordan because he simply just has better career averages in the playoffs.

Comparing Lebron and Jordan's Longevity:

```
convert_percentage <- function(perc) {
  as.numeric(gsub("%", "", perc)) / 100
}

LebronLongevity <- LebronAdvanced %>%
  mutate(PlayerName = "LeBron James", Games = as.integer(G),
         TrueShootingPercentage = convert_percentage(`TS%`),
         WinShare = as.numeric(WS),
```

```
        ValueOverReplacementPlayer = as.numeric(VORP)) %>%
  select(PlayerName, Season, Games, TrueShootingPercentage,
        WinShare, ValueOverReplacementPlayer)

JordanLongevity <- JordanAdvanced %>%
  mutate(PlayerName = "Michael Jordan", Games = as.integer(G),
        TrueShootingPercentage = convert_percentage(`TS%`),
        WinShare = as.numeric(WS),
        ValueOverReplacementPlayer = as.numeric(VORP)) %>%
  select(PlayerName, Season, Games, TrueShootingPercentage,
        WinShare, ValueOverReplacementPlayer)

CombinedData <- bind_rows(LebronLongevity, JordanLongevity)

CombinedDataLonger <- CombinedData %>%
  pivot_longer(cols = c(Games, TrueShootingPercentage, WinShare),
               names_to = "Statistic", values_to = "Value")

CombinedDataLonger %>%
  ggplot(aes(x = Season, y = Value, color = PlayerName)) +
  geom_point(aes(shape = PlayerName), size = 3,
             position = position_jitterdodge()) +
  geom_line(aes(group = PlayerName), alpha = 0.5) +
  scale_color_manual(values = c("LeBron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  facet_wrap(~Statistic, scales = "free_y") +
  labs(title = "Longevity Analysis: LeBron James vs Michael Jordan",
       subtitle = "Statistical comparison across different seasons",
       x = "Season", y = "Value") +
  theme_minimal() +
  theme(text = element_text(size = 12),
        axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom", plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14))
```

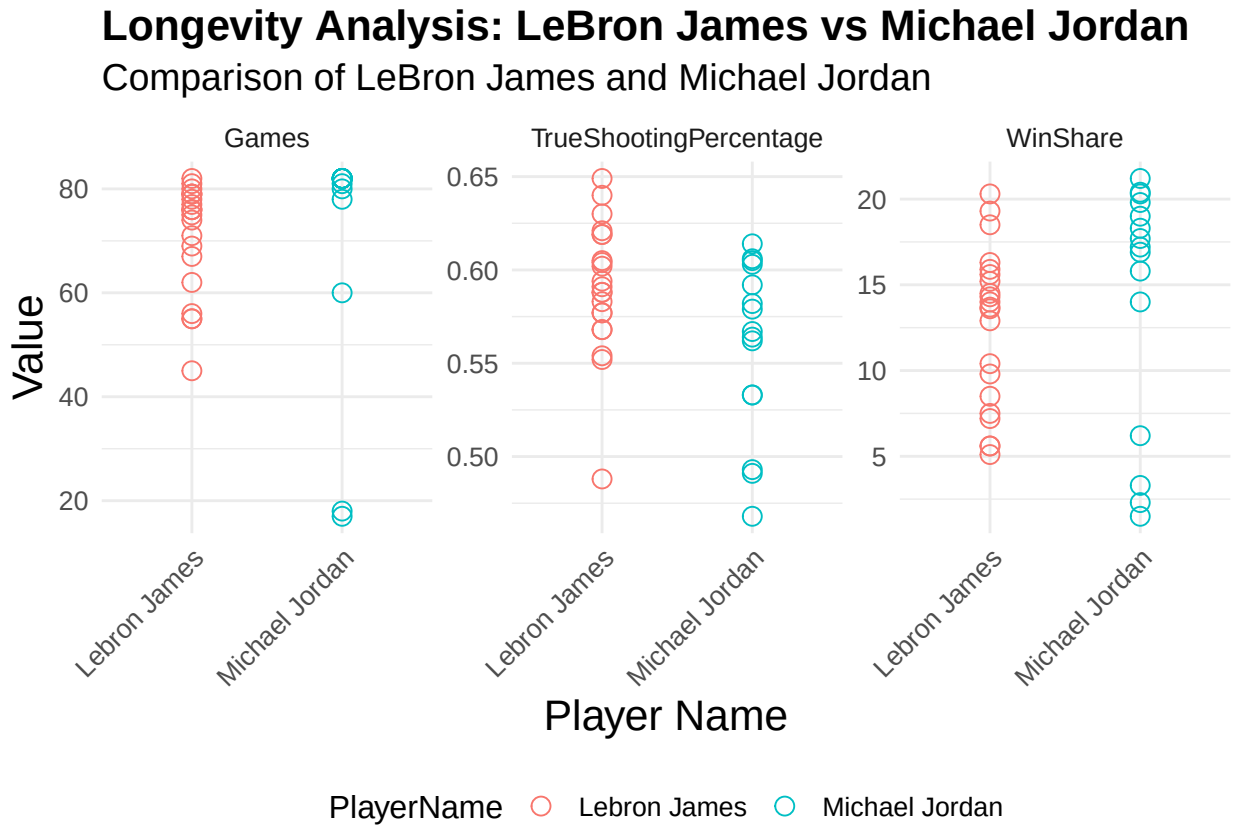



```
print(CombinedDataLonger)
```

```
## # A tibble: 135 x 5
##    PlayerName   Season  ValueOverReplacementPlayer Statistic               Value
##    <chr>        <chr>                        <dbl> <chr>                   <dbl>
##  1 LeBron James 2003-04                        2.9 Games                 7.9 e+1
##  2 LeBron James 2003-04                        2.9 TrueShootingPercenta~ 4.88e-3
##  3 LeBron James 2003-04                        2.9 WinShare              5.1 e+0
##  4 LeBron James 2004-05                        9.1 Games                 8   e+1
##  5 LeBron James 2004-05                        9.1 TrueShootingPercenta~ 5.54e-3
##  6 LeBron James 2004-05                        9.1 WinShare              1.43e+1
##  7 LeBron James 2005-06                        9.4 Games                 7.9 e+1
##  8 LeBron James 2005-06                        9.4 TrueShootingPercenta~ 5.68e-3
##  9 LeBron James 2005-06                        9.4 WinShare              1.63e+1
## 10 LeBron James 2006-07                        8.1 Games                 7.8 e+1
## # i 125 more rows
```

```
convert_percentage <- function(perc) {
  as.numeric(sub("%", "", perc))
}

LebronLongevity <- LebronAdvanced %>%
  mutate(PlayerName = "Lebron James", Games = G,
         TrueShootingPercentage = convert_percentage(`TS%`),
         WinShare = WS, ValueOverReplacementPlayer = VORP) %>%
  select(PlayerName, Season, Games, TrueShootingPercentage,
         WinShare, ValueOverReplacementPlayer)

JordanLongevity <- JordanAdvanced %>%
  mutate(PlayerName = "Michael Jordan", Games = G,
         TrueShootingPercentage = convert_percentage(`TS%`),
```

```
         WinShare = WS, ValueOverReplacementPlayer = VORP) %>%
  select(PlayerName, Season, Games, TrueShootingPercentage,
         WinShare, ValueOverReplacementPlayer)

CombinedData <- bind_rows(LebronLongevity, JordanLongevity)

CombinedDataLonger <- CombinedData %>%
  pivot_longer(cols = c(Games, TrueShootingPercentage, WinShare),
               names_to = "Statistic", values_to = "Value")

CombinedDataLonger %>%
  ggplot(aes(x = PlayerName, y = Value, color = PlayerName)) +
  geom_point(shape = 21, size = 3) +
  scale_color_hue(direction = 1) +
  theme_minimal() +
  facet_wrap(~Statistic, scales = "free_y") +
  labs(title = "Longevity Analysis: LeBron James vs Michael Jordan",
       subtitle = "Comparison of LeBron James and Michael Jordan",
       x = "Player Name", y = "Value") +
  theme(text = element_text(size = 12),
        axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom", plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14))
```




```
print(CombinedDataLonger)
```

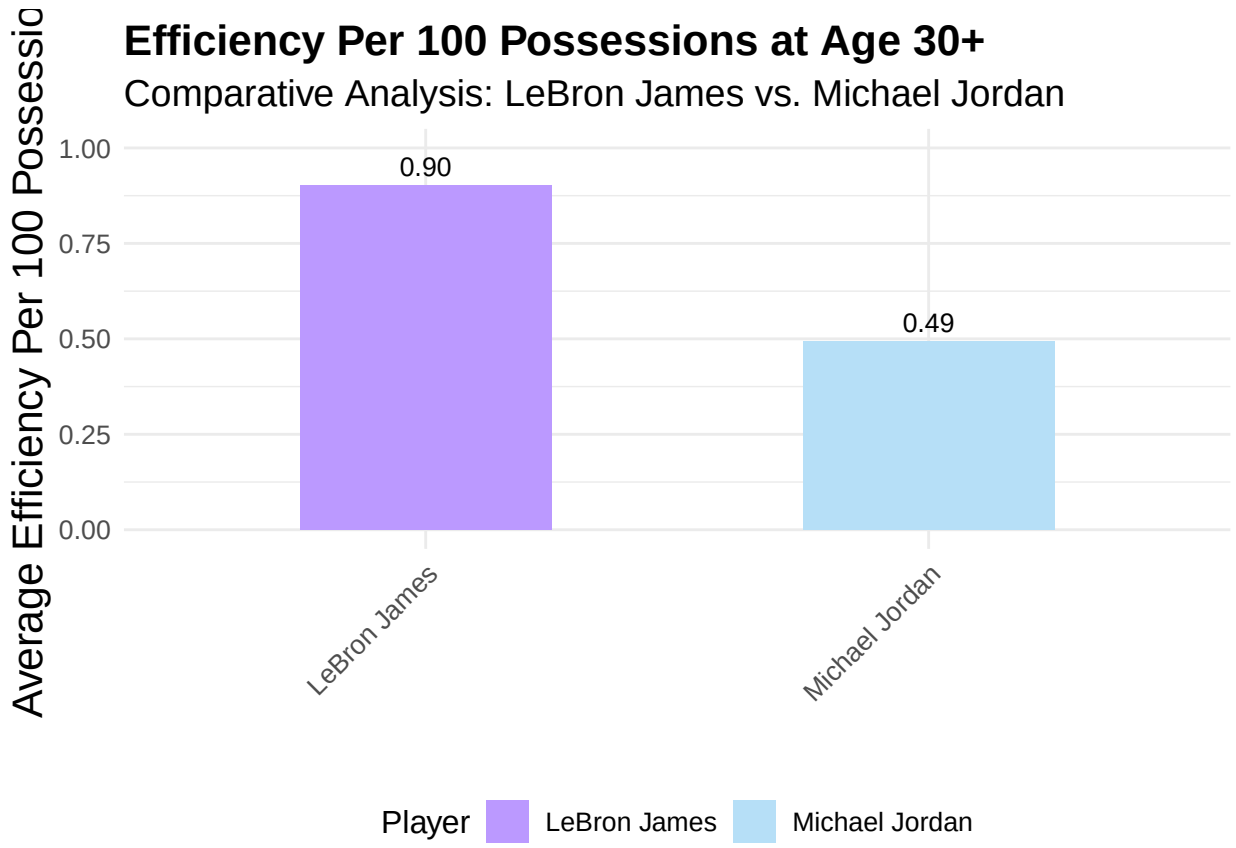
```
## # A tibble: 135 x 5
##    PlayerName   Season  ValueOverReplacementPlayer Statistic              Value
##    <chr>        <chr>                        <dbl> <chr>                  <dbl>
```

```
##  1 Lebron James 2003-04                    2.9 Games                   79
##  2 Lebron James 2003-04                    2.9 TrueShootingPercentage  0.488
##  3 Lebron James 2003-04                    2.9 WinShare                5.1
##  4 Lebron James 2004-05                    9.1 Games                   80
##  5 Lebron James 2004-05                    9.1 TrueShootingPercentage  0.554
##  6 Lebron James 2004-05                    9.1 WinShare                14.3
##  7 Lebron James 2005-06                    9.4 Games                   79
##  8 Lebron James 2005-06                    9.4 TrueShootingPercentage  0.568
##  9 Lebron James 2005-06                    9.4 WinShare                16.3
## 10 Lebron James 2006-07                    8.1 Games                   78
## # i 125 more rows
```

For this two plots, we are comparing advanced stats for LeBron James and Michael Jordan. Win shares and true shooting percentage are some very insightful statistics not many basketball fans pay attention to as much. Win shares represents the number of wins contributed by that player and it reveals to us how valuable that player truly is to the team. True shooting percentage provides a more accurate measure of a players shooting efficiency. The stat incorporates all forms scoring that contributes to a players total points. When it comes to LeBron and Jordan, we see LeBron being the more efficient shooter.

Comparing Lebron and Jordan's Personal Efficiency Rating At Ages 30+:

```
OldBron <- LebronPer100Poss %>%
  filter(Season >= "2014-15") %>%
  mutate(PER = (PTS + TRB + AST + STL + BLK - ((FGA - FG) + (FTA - FT) + TOV)) / G,
         Player = "LeBron James")

OldJordan <- JordanPer100Poss %>%
  filter(Season >= "1995-96") %>%
  mutate(PER = (PTS + TRB + AST + STL + BLK - ((FGA - FG) + (FTA - FT) + TOV)) / G,
         Player = "Michael Jordan")

OldComp <- bind_rows(OldBron, OldJordan) %>%
  select(Player, everything())

maxPER <- OldComp %>%
  group_by(Player) %>%
  summarize(MaxPER = max(PER))

OldComp %>%
  ggplot(aes(x = Player, y = PER, fill = Player)) +
  geom_bar(stat = "identity", position = position_dodge(), width = 0.5) +
  geom_text(data = maxPER, aes(label = sprintf("%.2f", MaxPER), y = MaxPER),
            position = position_dodge(width = 0.5), vjust = -0.5, size = 3.5) +
  theme_minimal() +
  labs(title = "Efficiency Per 100 Possessions at Age 30+",
       subtitle = "Comparative Analysis: LeBron James vs. Michael Jordan",
       x = "", y = "Average Efficiency Per 100 Possessions", fill = "Player") +
  scale_fill_manual(values = c("LeBron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  theme(text = element_text(size = 12), axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom", plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14)) +
  ylim(0, 1)
```

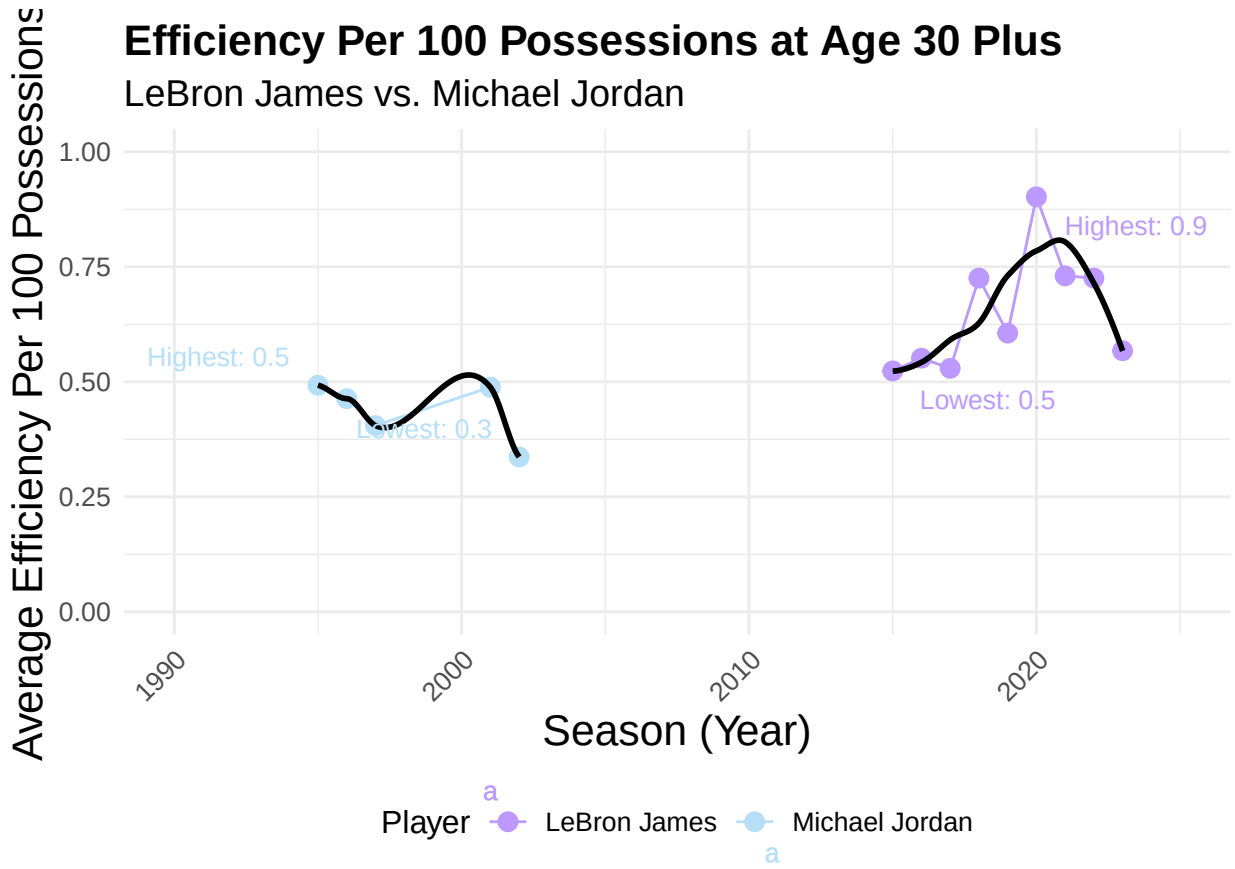



```
print(OldComp)
```

```
## # A tibble: 15 x 34
##    Player     Season   Age Tm    Lg    Pos       G    GS    MP    FG   FGA `FG%`
##    <chr>      <chr>  <dbl> <chr> <chr> <chr> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1 LeBron Ja~ 2014-~    30 CLE   NBA   SF       69    69  2493  13    26.7 0.488
##  2 LeBron Ja~ 2015-~    31 CLE   NBA   SF       76    76  2709  14    26.9 0.52
##  3 LeBron Ja~ 2016-~    32 CLE   NBA   SF       74    74  2794  13.1  24   0.548
##  4 LeBron Ja~ 2017-~    33 CLE   NBA   PF       82    82  3026  13.9  25.6 0.542
##  5 LeBron Ja~ 2018-~    34 LAL   NBA   SF       55    55  1937  13.4  26.3 0.51
##  6 LeBron Ja~ 2019-~    35 LAL   NBA   PG       67    67  2316  13.2  26.8 0.493
##  7 LeBron Ja~ 2020-~    36 LAL   NBA   PG       45    45  1504  13.7  26.6 0.513
##  8 LeBron Ja~ 2021-~    37 LAL   NBA   C        56    56  2084  14.7  28.1 0.524
##  9 LeBron Ja~ 2022-~    38 LAL   NBA   PF       55    54  1954  14.8  29.6 0.5
## 10 LeBron Ja~ 2023-~    39 LAL   NBA   PF       71    71  2504  13    24.1 0.54
## 11 Michael J~ 1995-~    32 CHI   NBA   SG       82    82  3090  15.6  31.5 0.495
## 12 Michael J~ 1996-~    33 CHI   NBA   SG       82    82  3106  15.8  32.5 0.486
## 13 Michael J~ 1997-~    34 CHI   NBA   SG       82    82  3181  14.9  32.1 0.465
## 14 Michael J~ 2001-~    38 WAS   NBA   SF       60    53  2093  14.3  34.4 0.416
## 15 Michael J~ 2002-~    39 WAS   NBA   SF       82    67  3031  12.2  27.4 0.445
## # i 22 more variables: `3P` <dbl>, `3PA` <dbl>, `3P%` <dbl>, `2P` <dbl>,
## #   `2PA` <dbl>, `2P%` <dbl>, FT <dbl>, FTA <dbl>, `FT%` <dbl>, ORB <dbl>,
## #   DRB <dbl>, TRB <dbl>, AST <dbl>, STL <dbl>, BLK <dbl>, TOV <dbl>, PF <dbl>,
## #   PTS <dbl>, ...30 <lgl>, ORtg <dbl>, DRtg <dbl>, PER <dbl>
```

```
OldBron <- LebronPer100Poss %>%
  filter(Season > "2014-15") %>%
  mutate(PER = (PTS + TRB + AST + STL + BLK - ((FGA - FG) + (FTA - FT) + TOV)) / G,
         Player = "LeBron James",
         Season = as.numeric(substr(Season, 1, 4))) %>%
  select(Player, Season, PER)
```

```r
OldJordan <- JordanPer100Poss %>%
  filter(Season > "1994-95") %>%
  mutate(PER = (PTS + TRB + AST + STL + BLK - ((FGA - FG) + (FTA - FT) + TOV)) / G,
         Player = "Michael Jordan",
         Season = as.numeric(substr(Season, 1, 4))) %>%
  select(Player, Season, PER)

OldComp <- bind_rows(OldBron, OldJordan)

HighestPERLebron <- OldBron %>% filter(PER == max(PER))
LowestPERLebron <- OldBron %>% filter(PER == min(PER))
HighestPERJordan <- OldJordan %>% filter(PER == max(PER))
LowestPERJordan <- OldJordan %>% filter(PER == min(PER))

ggplot(OldComp, aes(x = Season, y = PER, color = Player, group = Player)) +
  geom_line() +
  geom_point(size = 3) +
  geom_smooth(method = "loess", se = FALSE, color = "black", size = 1) +
  geom_text(data = HighestPERLebron, aes(label = sprintf("Highest: %.1f", PER)),
            hjust = -0.2, vjust = 2, size = 3.5) +
  geom_text(data = LowestPERLebron, aes(label = sprintf("Lowest: %.1f", PER)),
            hjust = -0.2, vjust = 2, size = 3.5) +
  geom_text(data = HighestPERJordan, aes(label = sprintf("Highest: %.1f", PER)),
            hjust = 1.2, vjust = -1, size = 3.5) +
  geom_text(data = LowestPERJordan, aes(label = sprintf("Lowest: %.1f", PER)),
            hjust = 1.2, vjust = -1, size = 3.5) +
  scale_color_manual(values = c("LeBron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  labs(title = "Efficiency Per 100 Possessions at Age 30 Plus",
       subtitle = "LeBron James vs. Michael Jordan",
       x = "Season (Year)", y = "Average Efficiency Per 100 Possessions",
       color = "Player") +
  theme_minimal() +
  theme(text = element_text(size = 12), axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom", plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14)) +
  ylim(0, 1) +
  xlim(1990, 2025)
```



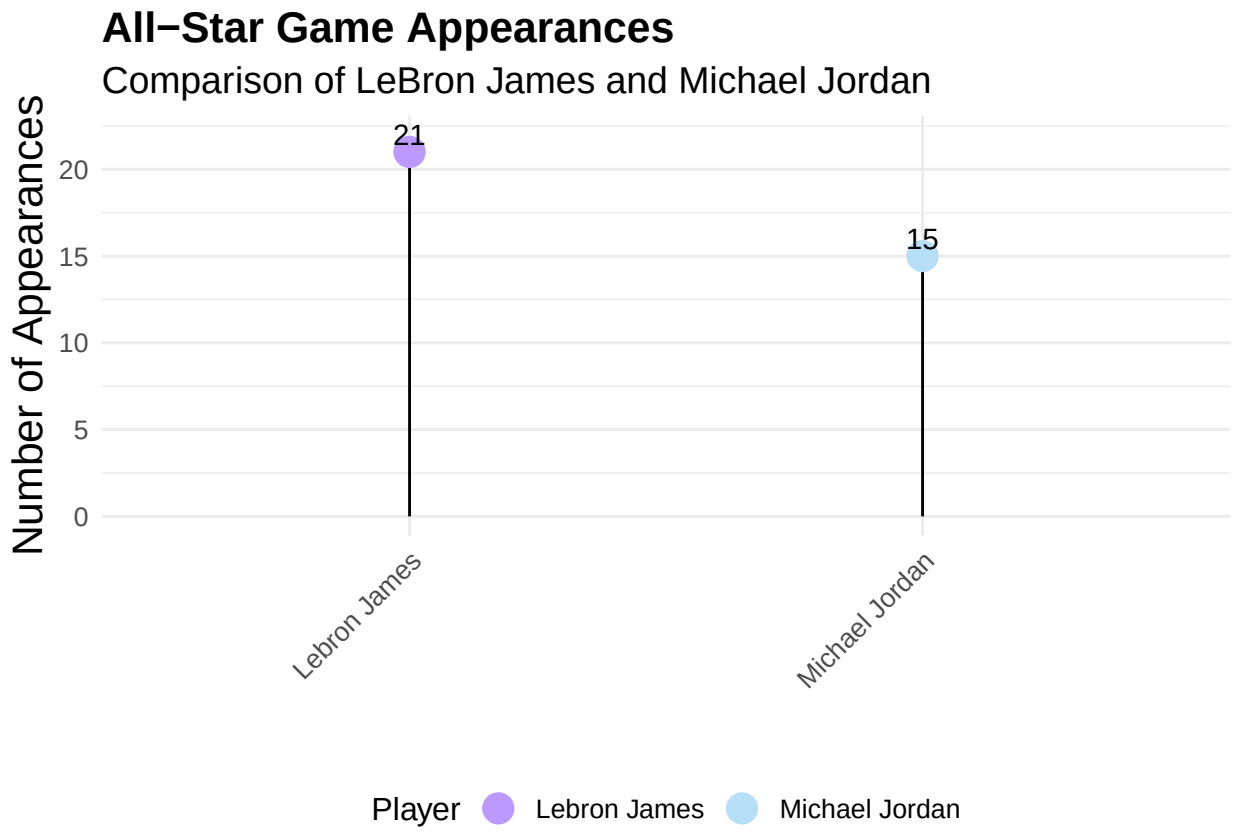

```
print(OldComp)
```

```
## # A tibble: 14 x 3
##    Player         Season   PER
##    <chr>           <dbl> <dbl>
##  1 LeBron James     2015 0.524
##  2 LeBron James     2016 0.551
##  3 LeBron James     2017 0.529
##  4 LeBron James     2018 0.725
##  5 LeBron James     2019 0.606
##  6 LeBron James     2020 0.902
##  7 LeBron James     2021 0.730
##  8 LeBron James     2022 0.725
##  9 LeBron James     2023 0.568
## 10 Michael Jordan   1995 0.493
## 11 Michael Jordan   1996 0.463
## 12 Michael Jordan   1997 0.405
## 13 Michael Jordan   2001 0.488
## 14 Michael Jordan   2002 0.337
```

Both of these plots compare Michael Jordan and LeBrons efficiency's per 100 possessions at age 30 plus. We can clearly see the longevity LeBron holds. When comparing points, assists, rebounds,field goal percentages, and turnover rate, we see that LeBron dominates. At age 39 LeBron is still considered a top player in the NBA while Jordan was a dwindling old man expiring in his puny body.

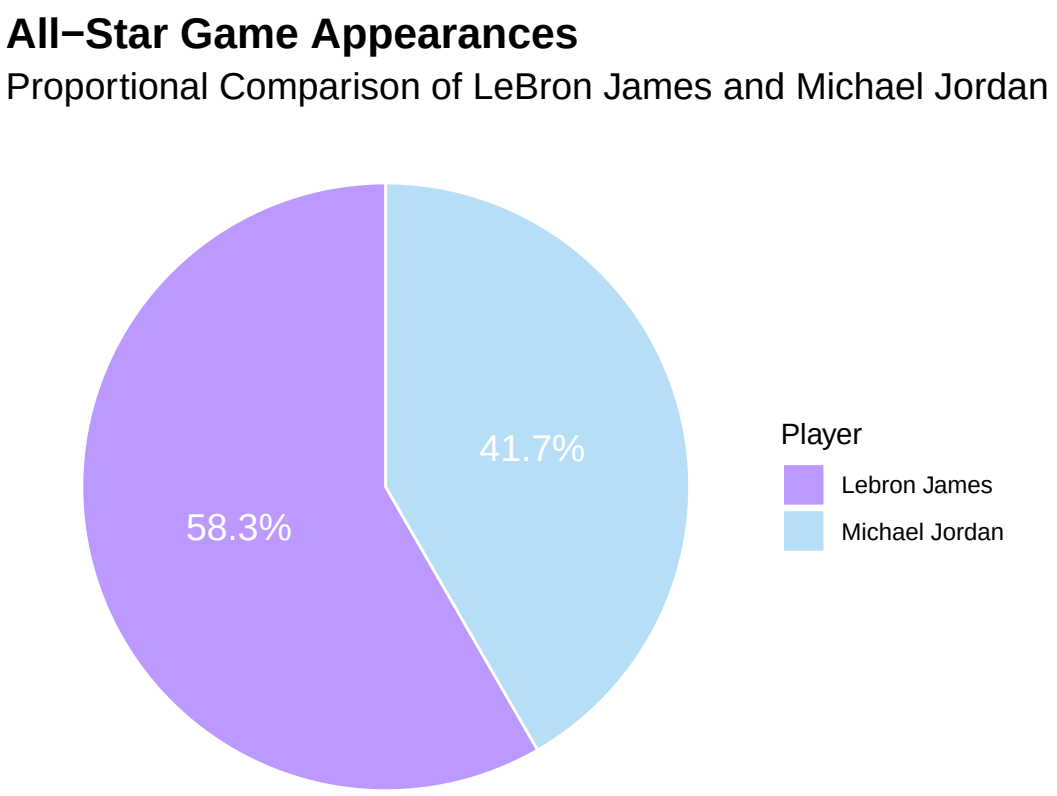
Comparing Total All Star Appearences

```r
LebronAllStarAppearances <- nrow(LebronAllStarGames)
JordanAllStarAppearances <- nrow(JordanAllStarGames)

AllStarData <- data.frame(Player = c("Lebron James", "Michael Jordan"),
                          Appearances = c(LebronAllStarAppearances, JordanAllStarAppearances))

AllStarData$Percentage <- with(AllStarData, Appearances / sum(Appearances) * 100)

ggplot(AllStarData, aes(x = Player, y = Appearances, label = Appearances)) +
  geom_segment(aes(y = 0, x = Player, yend = Appearances, xend = Player)) +
  geom_point(size = 5, aes(color = Player)) +
  scale_color_manual(values = c("Lebron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  geom_text(nudge_y = 1, size = 4, color = "black") +
  labs(title = "All-Star Game Appearances",
       subtitle = "Comparison of LeBron James and Michael Jordan",
       x = "", y = "Number of Appearances") +
  theme_minimal() +
  theme(text = element_text(size = 12), axis.title = element_text(size = 16),
        axis.text.x = element_text(angle = 45, vjust = 1, hjust = 1),
        legend.position = "bottom", plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14))
```




```r
ggplot(AllStarData, aes(x = "", y = Percentage, fill = Player)) +
  geom_bar(stat = "identity", width = 1, color = "white") +
  coord_polar(theta = "y") +
  scale_fill_manual(values = c("Lebron James" = "#BB99FF", "Michael Jordan" = "#B6DFF7")) +
  labs(title = "All-Star Game Appearances",
       subtitle = "Proportional Comparison of LeBron James and Michael Jordan",
       x = NULL, y = NULL) +
  theme_void() +
```

```
  theme(plot.title = element_text(size = 16, face = "bold"),
        plot.subtitle = element_text(size = 14)) +
  geom_text(aes(label = sprintf("%.1f%%", Percentage)),
            position = position_stack(vjust = 0.5), color = "white", size = 5)
```




```
print(AllStarData)
```

```
##           Player Appearances Percentage
## 1   Lebron James          21   58.33333
## 2 Michael Jordan          15   41.66667
```

When it comes to All-Star game appearences Lebron has been slected to one every season he has played in the NBA and sits at the most all-star selections ever. Jordans 15, although impressive, are expected when you are considered one of the greatest of all time. Lebrons Longevity to be able to make it to an all star game 21 years, even at his current age of 39 is unheard of.

Conclusion

Key insight/takeaway about research question - Summarize the key insight, takeaway, conclusion to the research question that motivated your analysis

While studying our data comparing LeBron and Jordan we found that the GOAT debate is closer than we originally thought it was when viewed from a statistical perspective. We couldn't really conclude who was the greatest player ever with some of the most insightful basketball stats, there were strong cases for both. But because basketball is just more than statistics, a visual eye test might also help you sway a certain direction with who you think is the greatest basketball player of all time. We will still go with LeBron as who we consider the greatest basketball player of all time.

Challenge Encountered - Describe the biggest challenge that you encountered and how you overcame it in the project.

The biggest challenge we faced was finding relevant data sources. All the data sources off of Kaggle were old when it comes to finding LeBron's stats. Lebron is still currently playing so we needed up to date statistics while everything on Kaggle was over 5 years behind. We came across the websites basketball reference and Stat head to head which aided us heavily for finding out the statistics we needed and wanted to use.

Initially, we had used an absolute path to load the data into R Studio, but then realized that anyone who was accessing this program outside of our computers were unable to run the program since the files were connected to our computers absolutely. We had realized this mistake and now made the paths relative so now anyone with the source files can access the data.