

Project1DAT301

Tim Schneider

2022-11-20

What determines wins in MCBB?

I used the cbb dataset from Kaggle that is described as

"Data from the 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020, and 2021 Division I college basketball seasons.

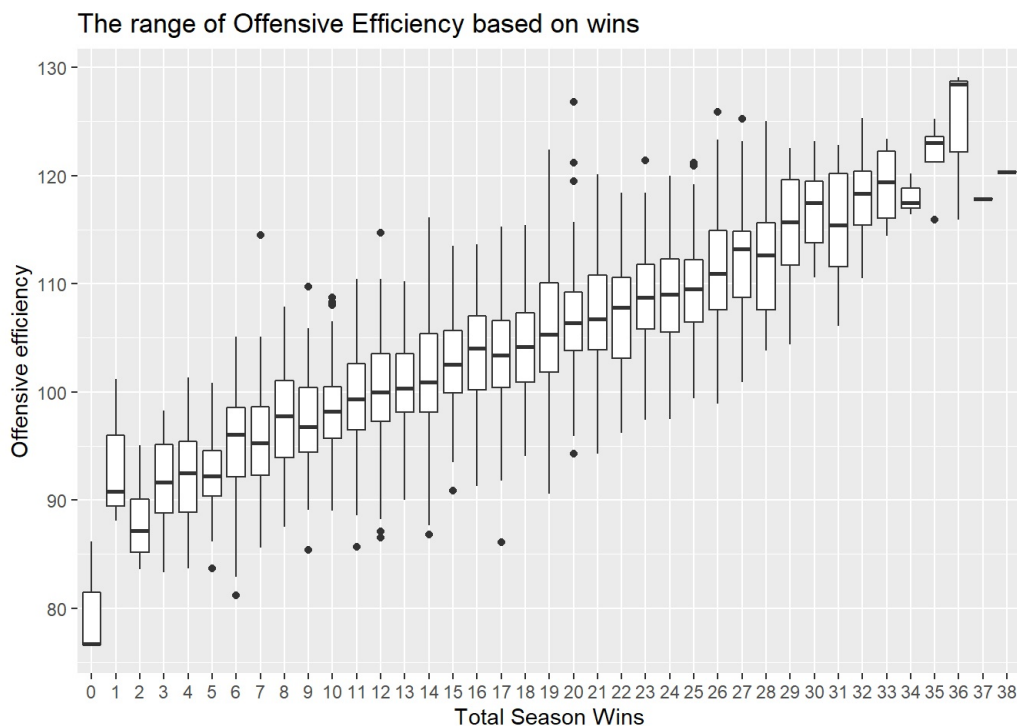
cbb.csv has seasons 2013-2019 combined."

This dataset allows us to look at trends in the efficiency of an offense and defense and where they placed in the years respective tournament.

This graph shows the Adjusted Offensive Efficiency which is the estimate of points that a team will score on an average division 1 basketball team per 100 possessions versus the number of wins that the team got that season.

```
WINX <- as.factor(cbb$W)
OFFWINY <- cbb$EFG_0
OFFWIN = ggplot(data = cbb, aes(WINX, ADJOE ) )+
  geom_boxplot()

OFFWIN +
  ggtitle("The range of Offensive Efficiency based on wins") +
  xlab("Total Season Wins") +
  ylab("Offensive efficiency")
```



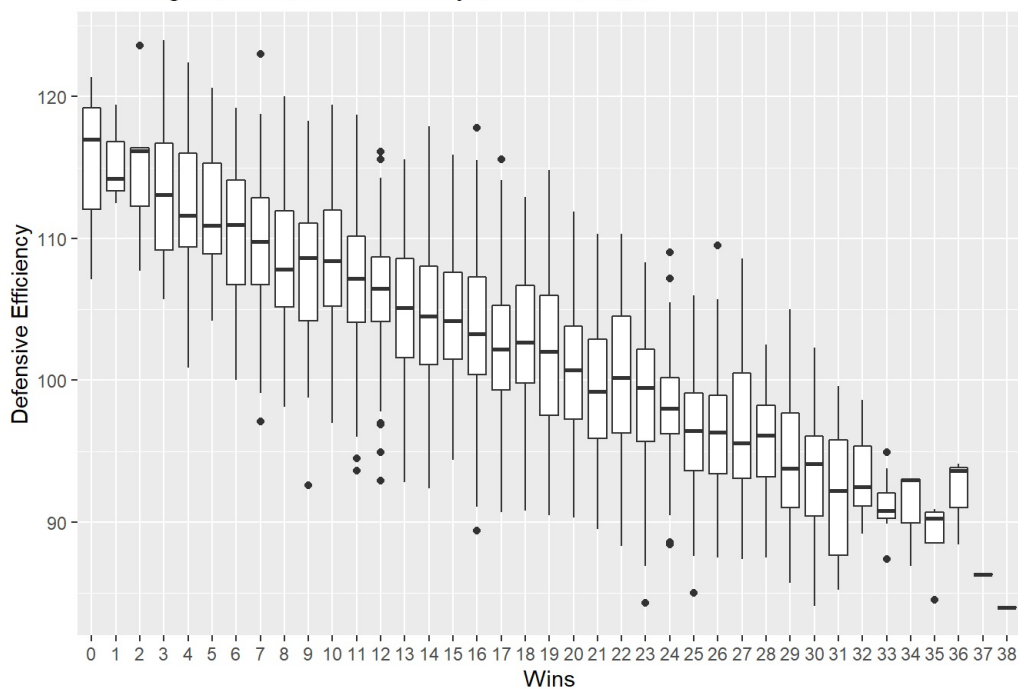
As the graph shows the trend is as offensive efficiency increases wins should also increase. This reaffirms that the more points a team scores helps them win games.

Defensive Efficiency

This graph shows the Adjusted Defensive Efficiency which is the estimate of points that a defense allows per 100 possessions against the average division 1 offense versus the number of wins that team got that year.

```
DEFWINY <- cbb$EFG_D
DEFWIN = ggplot(data = cbb, aes(WINX, ADJDE ) )+
  geom_boxplot()
DEFWIN +
  ggtitle("The range of Defensive Efficiency based on wins") +
  xlab("Wins") +
  ylab("Defensive Efficiency")
```

The range of Defensive Efficiency based on wins



The graph shows that the less points teams allows helps them win games but has a smaller slope than the offensive efficiency. So although it helps letting teams score less points on you, it matters a little bit more that you score.

OEFG vs DEFG

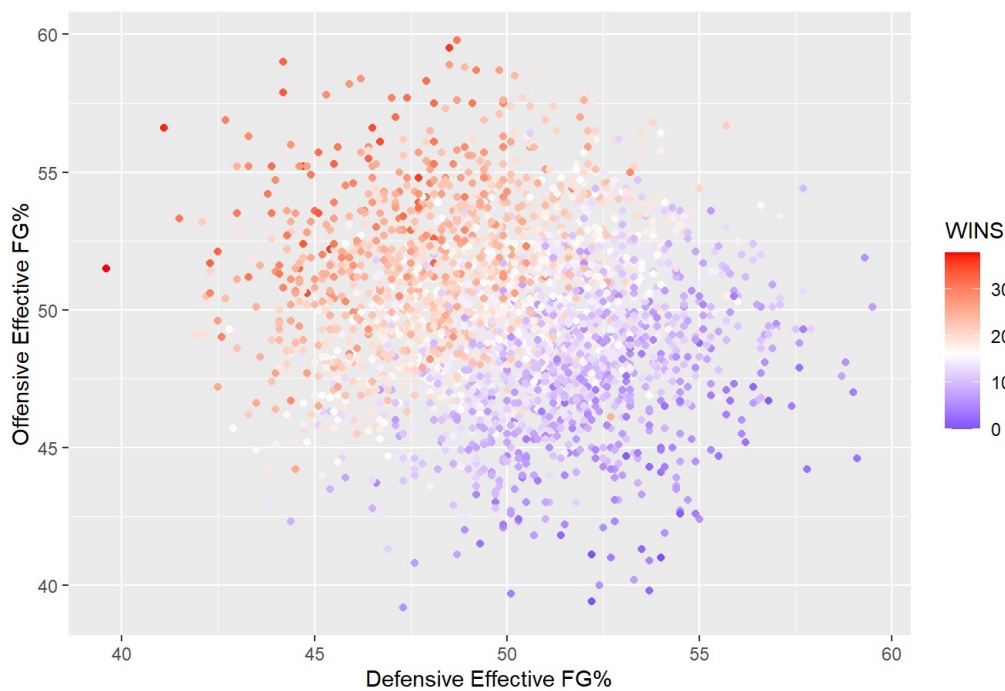
This graph takes key indicators of how well an offense and defense run and plots them with the color indicating the amount of wins. Using Offensive Effective Field goal percentage (O_eFG%) it accounts for the different types of shots offenses take. It gets calculated by $\frac{2ptFGM + 1.5 * 3ptFGM}{FGA}$ Which normalizes points scored, allowing us to take the field goal percentage of each point scored. The other key indicator is the Defensive Effective Field goal percentage (D_eFG%). Its the same statistic as O_eFG% but on the defensive side, so it normalizes the field goal percentage that a team allows.

```
WINS <- as.numeric(as.character(WINX))
OFFDEFWIN <- ggplot(data = cbb,aes(x = DEFWINY, y = OFFWINY, color = WINS)) +
  geom_point()+
  guides(color = guide_colorbar(),
  col=guide_legend("Wins") )

mid<-mean(WINS)

OFFDEFWIN + scale_color_gradient2(midpoint=mid, low="blue", mid="white",high="red", space ="Lab") +
  ggtitle("Offensive Effective FG% vs Defensive Effective FG%") +
  xlab("Defensive Effective FG%") +
  ylab("Offensive Effective FG%")
```

Offensive Effective FG% vs Defensive Effective FG%



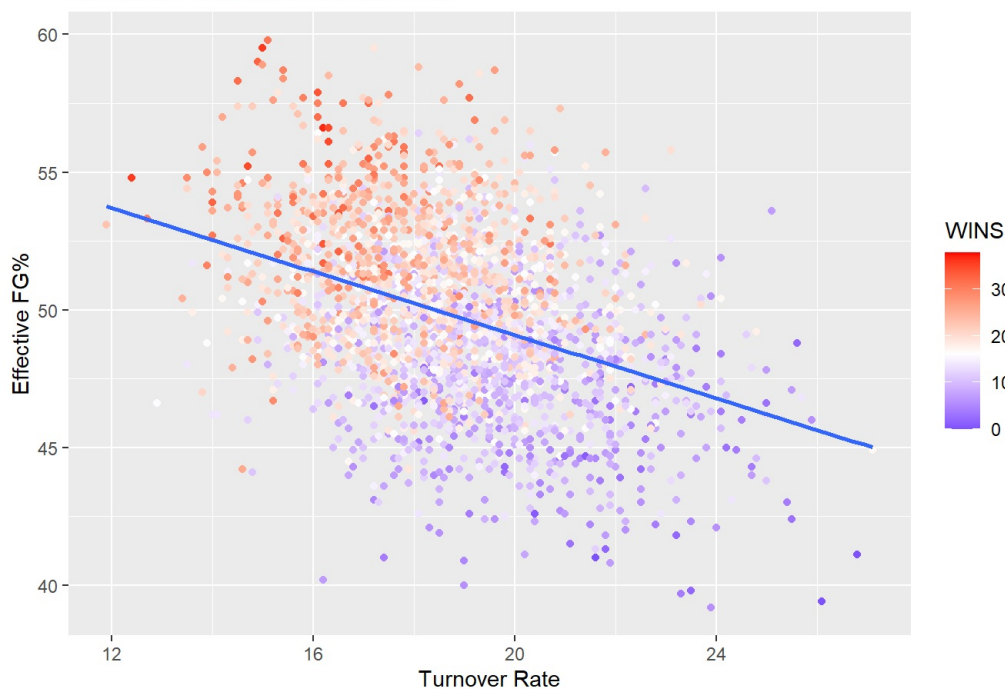
The graph shows that teams with higher O_eFG% and the lower the D_eFG% the more wins a team has.

O_eFG% vs Turnover Rate

The graph takes the same O_eFG% and plots it against Turnover rate, showing not only how well a team can score the basketball but also how well they do to keep it in their hands.

```
TORPLOT <- ggplot(data=cbb, aes(x = TOR, y = EFG_0,color = WINS)) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  guides(color = guide_colorbar(),
         col=guide_legend("Wins"))
TORPLOT + scale_color_gradient2(midpoint=mid, low="blue", mid="white",
                                high="red", space = "Lab") +
  ggtitle("Effective FG% vs. Turnover Rate") +
  xlab("Turnover Rate") +
  ylab("Effective FG%")
```

Effective FG% vs. Turnover Rate

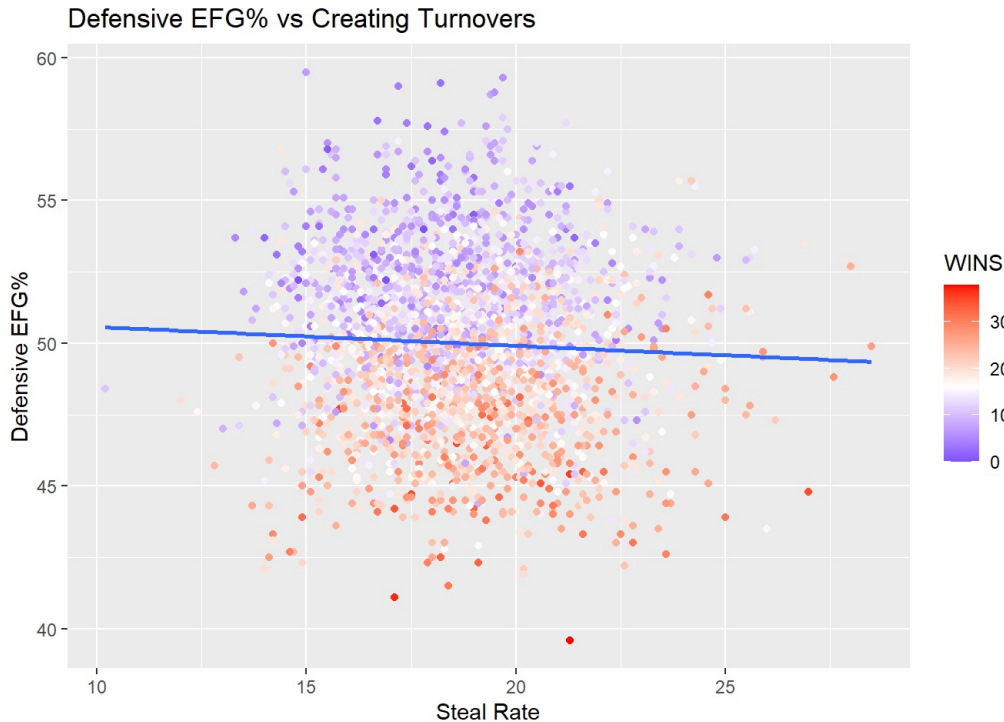


Teams who tend to make more shots also on average tend to not turn the ball over and end up winning more games because of it. A little surprising considering some turnovers don't relate to how well the opposing defense is playing.

D_eFG% vs. Steal Rate

This graph shows the opposite of the last, showing the D_eFG% and how well the defense can create turnovers by steal rate. This is a great measure of a defense as steals end up creating extra possessions for the offense.

```
STLPLLOT <- ggplot(data=cbb, aes(x = TORD, y = EFG_D,color = WINS))+
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  guides(color = guide_colorbar(),
         col=guide_legend("Wins"))
STLPLLOT + scale_color_gradient2(midpoint=mid, low="blue", mid="white",
                                high="red", space = "Lab") +
  ggtitle("Defensive EFG% vs Creating Turnovers") +
  xlab("Steal Rate") +
  ylab("Defensive EFG%")
```



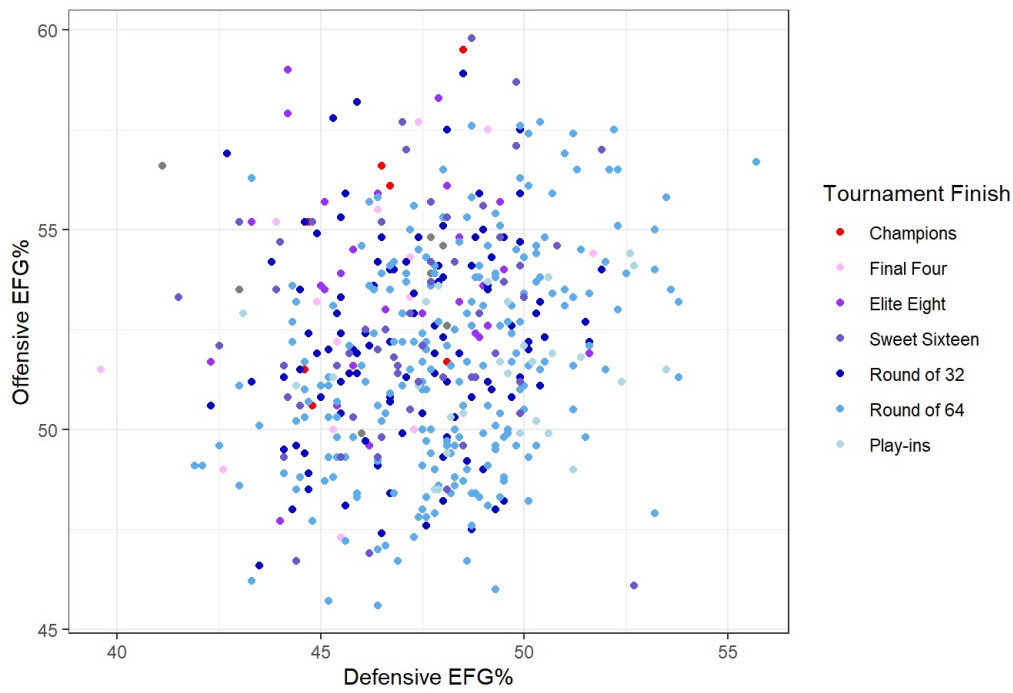
The graph shows that although teams get more wins by allowing a less defensive eFG% that steal rate doesn't play into the likelihood of getting wins. With the data not really having any other trends.

O_eFG% vs D_eFG% of the top 68 teams

This graph took only the teams that made it to the March Madness tournament, so only the better of the teams that season and plots the O_eFG% vs D_eFG% in respect to how they finished in the tournament.

```
POST <- na.omit(cbb)
POSTDEF <- POST$EFG_D
POSTOFF <- POST$EFG_O
POSTCONTWIN <- as.numeric(as.character(POST$W))
TOURNOFFDEFWIN <- ggplot(data = POST, aes(x = POSTDEF, y = POSTOFF, color = POST$POSTSEASON)) +
  geom_point()
TOURNOFFDEFWIN + scale_color_manual(values = c("red","purple4","plum1","purple1","slateblue","mediumblue","steelblue2","lightblue"),
  name = "Tournament Finish",
  breaks=c("Champions","2nd","F4","E8","S16","R32","R64","R68"),
  labels = c("Champions","Second Place","Final Four","Elite Eight","Sweet Sixteen","Round of 32","Round of 64","Play-ins")) +
  theme_bw() +
  ggtitle("Offensive EFG% vs Defensive EFG% Of the Teams that made March Madness") +
  xlab("Defensive EFG%") +
  ylab("Offensive EFG%")
```

Offensive EFG% vs Defensive EFG% Of the Teams that made March Madness



The graph shows that everybody who made the final four with the exception of one team had a D_eFG% > 50% where as the champion also had a D_eFG% of < 50%

Loading MainJax/jax/output/HTML-CSS/jax.js