Names:

Christopher Miklasevich

Maggie Greco

Bryan Veres

Project: NBA Survival

The data used for this project will be a combination of two NBA datasets to see the survivability of the NBA players who were drafted from 2007 and beyond. Note: there is a lot of data that is being left truncated, however we believe that this is not a problem because this would be the start of “Super Team” era (3 “superstar” players on a team). This started as the 2007 Boston Celtics got Kevin Garrett and Ray Allen with Paul Pierce, and this would continue as Miami Heat, Golden State Warriors, and Oklahoma City Thunder would follow this trend.

The variable that we will be using are: Years Position, Points, Assists, Rebounds, Blocks, Steals, Personal Fouls, Games played, and Games Started.

Years is the amount of years in the NBA the player has been playing for. This is the Time-to-Event data which is measured in years and starts in 2007 and ends in 2018.

Position is just what position you are. The positions usually hold some type of strength the player has. For example, shooting guard (SG) are usually get long range shooters, while Centers (C) are usually more defensive anchors, rebounding machines, screen setters, and dunkers.

Points is the average amount of points per game that the player scores over from 2007 to 2018

Assists is the average amount of assists per game that the player does over from 2007 to 2018

Rebounds is the average amount of rebounds per game over from 2007 to 2018

Blocks is the average amount of blocks per game over from 2007 to 2018

Steals is the average amount of steals per game over from 2007 to 2018

Personal Fouls (PF) is the average amount of fouls committed per game over from 2007 to 2018

Games played is the total amount of games played from 2007 to 2018

Games started is the total amount of games started playing from 2007 to 2018

**NOTE: There are a lot more statistics to describe the basketball player’s effect on the court. The list given is the most basic ones to be used**, **and perhaps the easiest to interpret.**

The data is a combination of Kaggle datasets. We are using the player\_data.csv and the Seasons\_Stats.csv from this link: <https://www.kaggle.com/drgilermo/nba-players-stats>

The original data is from Basketball Reference: <https://www.basketball-reference.com/>