Ammon Roberts 10793687

Connor Paulson 10684469

**Problem**: How to predict the team’s performance in the X amount of games, based on how they have performed in previous games, by looking at a data set of statistical categories ( for example: margin of victory 3-pt %, 2-pt %, etc.)

**Methods**:

* Carefully analyze and verify input data
* Create Python script to analyze input data
* Test script against data calculated by hand, to verify reliability
* Use script to analyze data

**Techniques**:

* Take a simple random sample of 5 games from 3 different teams (mean for total points, 3 point, 2 points, margin of victory)
* Determine the confidence interval of a few of the different statistical categories mentioned above
* From the confidence interval (determined in the above step), attempt to predict the next set of games performance when it comes to those statistical categories already mentioned
* Repeat, but sample a bigger dataset (for example: 10 games, 15 games, 30 games, a full season, etc.)

**Programming languages and packages**: Python, using numpy, scipy, and matplotlib libraries

**Results**: The result we hope to see is that we will be able to consistently predict a basketball team’s performance with a high rate of confidence

**Conclusion**: At the conclusion of the project, we will see the best sample size to use in order to predict the outcome of a game

**Timeline**:

* 5 hours - Find data samples and compute some samples by hand or another method (completed by 6/20)
* 5 hours - Plan how the Python script will read the sample data, calculate, and predict the final results (completed by 6/30)
* 10 hours - Create Python script to analyze data and make prediction (completed by 7/15)
* 10 hours - Test and verify data against samples previously calculated (completed by 7/31)
* 5 hours - Finalize any documentation and other tasks (completed by 8/14)

Total estimated time: 35 hours