Noah Beller and Iris Zhao

A PDF file containing a written description of your project. There are no specific page or word limits. This document should contain:

* + **a)  A description of the data. Report where you got the data. Describe the variables. If you had to reformat the data or filter it in any way, provide enough details that someone could repeat your results. If you combined multiple datasets, specify how you integrated them. Mention any additional data that you used, such as shape files for maps. Editing is important! You are not required to use every part of the dataset. Selectively choosing a subset can improve usability. Describe any criteria you used for data selection.**
  + We got data from Kaggle. We looked at two datasets, one on NBA player statistics (https://www.kaggle.com/drgilermo/nba-players-stats) which has detailed NBA statistics dating back to the 1950s. And one on the NBA draft, which detailed every draft pick since 1980.  (https://www.kaggle.com/pmp5kh/nba-draft-19802017?fbclid=IwAR18smgIdWyLAzMxq7BF9DQX5-bytohCyPZiGdx-ROAaiQY5AB4Lt\_IS4oA)

There were a lot of formatting differences between the two datasets. One data set was organized “First Name Last Name”. The other dataset used “Last Name, First Name”. One dataset also used accents on letters while the other didn’t. We had a few edge cases we had to work with like a one dataset referring to Michael Sweetney as Mike Sweetney, or in another case omitting a player’s first name entirely (Poeltl). Additionally, players who missed their entire rookie year (of which there are 7), had no cell for that year. So we had to make cells with empty data for those players.

The NBA player statistics dataset also uses 200Y to refer to the 200X-200Y season, while the draft picks uses 200X. Lastly, we wanted to look at points per game, but that statistic was not in the dataset, so we had to manually create it by dividing total points by games played.

We decided to look at the draft classes from 2003-2016. We started with the year 2003 because it is one of the most notably talented draft classes, and many view it as the start of the current era of basketball, and it is one of the latest years in which many of the players are still in the league. Looking at years earlier than this is looking at a league with a completely different context, and therefore the stats have to be understood in a different context.

* + **b)  An overview of your visual design rationale. A good rule of thumb to follow is “every pixel must be justified.” Instead of a 100,000-element breakdown, give us an overview of the design decisions you made and the trade-offs inherent in how you displayed the data. This part ought to include a description of the mapping from data to visual elements. Describe marks and channels you employ such as position, color, or shape. Mention any transformations you performed, such as log scales.**

For the overview graph, we wanted to make the graph longer horizontally because there would 20 bars on the graph at any given time and we wanted enough space so that each bar would be clear and each draft pick would be different from the other. We also wanted the bars to be somewhat thicker so that it would be easier for users to discover the mouseover functionality. The colors blue and yellow were chosen for the two different groups because they have distinctive contrast and red was chosen as the mouseover color to highlight which bar is being hovered over and it contrasts well with both blue and yellow.

For the charts on points per game, player efficiency rating, and effective field goal percentage, we wanted to use a similar format to the overview graph for easy comparison. We again used contrasting colors. The vertical bar chart works well because it becomes easy to see the trend of players being worse, the lower down in the lottery they are.

* + **c)  An overview of your interactive elements and their design rationale. Give us an outline of the design decisions that went into the interaction affordances you added to your visualization. What process did you use to choose the interactions you developed? How did you make them discoverable, usable, and interesting?**

For the overview graph we chose to have two different interactive elements. First, we chose to do a dropdown menu so the user can change between statistics to view different statistics for comparison. We chose a drop-down rather than a side by side comparison since each statistic is calculated differently and have different values it is unnecessary and actually more confusing to have the graphs appear at the same time. We chose to have a mouseover for users to view the player information for the highest value / “best” individual of that draft pick and statistic. We kept it as a mouseover because we want the main focus of the graph to be on comparing between the numbers and the statistic across draft pick order. If the user chooses to learn more, this will be an option. We chose a mouseover because it is easy to be discovered by the user and we made the bars larger to make this functionality more discoverable.

For the charts on points per game, player efficiency rating, and effective field goal percentage, we wanted to draw clear comparisons between the draft classes so we incorporated a slider, allowing the viewer to look through the years. We liked a slider because it has a very clear affordance, and is hard for the user to miss.

We wanted allowed for the user to toggle between each draft classes’ rookie year stats and career stats. We had the rookie year stats show first since this is a fairer playing field since the most recent draft class in our dataset (2016), only has 1 year of statistics, while the oldest class (2003) has as many as 14 years worth of statistics, with time in the league to improve those stats.

* + **d)  The story. What does your visualization tell us? What was surprising about it? What insights do you want to convey to the viewer of your visualization?**

We looked at 3 main stats for this analysis. Points per game, Player efficiency rating, and effective field goal percentage. Points per game is a standard statistic and metric used for a players talent. This can be understood by many viewers

PER is an incredibly complicated statistic

**Team Contributions:**

Noah:

* Joined the data
* Created the three graphs for individual player statistics by year - toggles between rookie year and career statistics

Iris:

* Created the graphs for cumulative averages compared to best player for each draft pick number
* Completed CSS and formatting.