# Janelle Goddard, Sarah Madden, Cole Pierie

# Washington University Data Analytics Boot Camp

# 02/09/20

# For this assignment, data was extracted from data.world via a compilation of NBA draft statistics in csv document form. Our team’s goal was to combine the data in order to toggle relevant NBA statistics for players drafted in 2014 through 2016; these variables include, but are not limited to, the player’s name, draft year, height with and without shoes, “wingspan”, reach height and step width, and agility. Analysts will be able to use our data to compare stats of new players either for personal use or in demonstrating the overall strength of a team based on the metrics of these newly added players. After inspection of the data, several null values were identified. Data cleaning originally included removing these data points, however due to technical issues the data was later changed to default values.

# It was determined to keep all data points relevant, However, troubleshooting included removing an entire column of inexplicably null statistics. No outlier values were determined and it is assumed all non-null data points are significant. Nearly all variables were determined to be double precision float values besides the “year” and “player”, defined in our schema as integer and variable character, respectively. Python and Pandas was used to establish connection from data frames we created from the NBA csv files to our database, as well as to clean and join the data. SQL was used to run database queries and visualize specific metrics. Some issues were encountered when establishing a connection to the tables generated in PgAdmin, however after further investigation we were able to rectify the logic used in connections between our code and data interface. Via pandas, we were able to join, inspect and clean the data completely in our script before importing into PgAdmin, thus we only had one table to generate information and streamline our data flow. We were thus able to generate a relational database.