Predicting NBA Season Scores Based on Past Performance and Statistics

UCB Extension: Data Analytics
Ilana Portugal, Eric Ralston, Philip Ring, James Fang, Nassim Ataii

Project Topic & Background

Our topic is focused on NBA player statistics and their ability to predict performance in the 2019 season. This topic was selected to determine if there are stats of NBA players' career that are indicators of NBA career success, and if they can be used to predict future metrics. We are building a model that will be able to predict future NBA player statistics. The season statistics we hope to predict are: rebounds, points and assists.

Description of Data

- Position
- Age
- Height
- Weight
- Draft number
- Draft round
- Average rebounds
- Average points
- Average assists
- Massive injuries that will serve as outliers

The sources for our data are the following:

- To compare player career statistics between the NBA and NCAA to give an overall view of a player's career trajectory: https://data.world/bgp12/nbancaacomparisons
- Measurements for NBA draft combine participants from DraftExpress.com: https://data.world/achou/nba-draft-combine-measurements
- NBA Player Data Per Season From 1978-2016: https://data.world/jgrosz99/nba-player-data-1978-2016

Questions To Be Answered

We aim to answer the following questions:

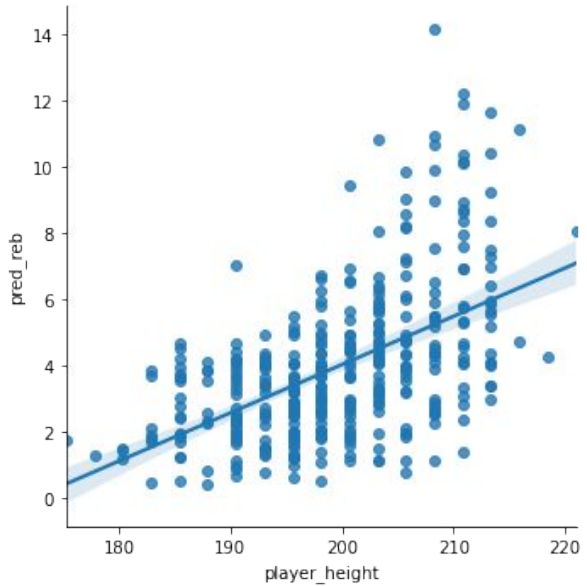
- Is there any correlation between NBA player past and future performance?
- Which metrics strongly correlate with player performance?
- Can we predict an NBA players season scores (points, rebounds and assists)?

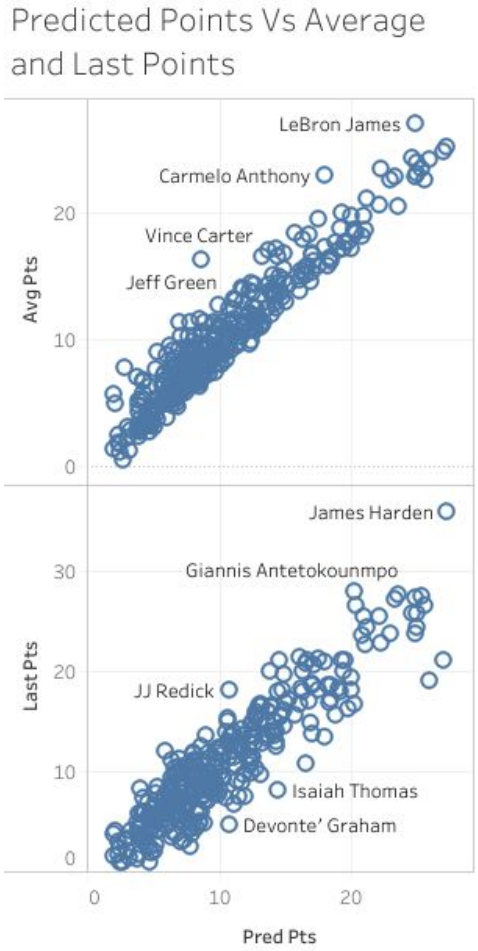
Description of Data Exploration Completed

A machine learning model was created by the following:

- Determine the number of unique values in each column for potential binning
- Split preprocessed data into our features and target arrays
- Define the model and layers
- Compile, Train, and Evaluate the model using the test data

Description of Analysis





Technologies, Tools, Language, Algorithms

The following were used to complete the exploration and analysis of our data:

- Tableau - data visualization.
- SQL - clean, organize, store data.
- Machine Learning - create model using sklearn, pandas to import data, establish model, and evaluate its accuracy.