Guy Dotan

Stats 404 – Ikukuyeva

1/15/19

**Homework #1**

**Part One - GitHub username**

guy-dotan

**Part Two – Final project data set**

<https://www.kaggle.com/noahgift/social-power-nba>

nba\_2016\_2017\_100.csv (100 x 63)

nba\_2017\_att\_val.csv (30 x 7)

nba\_2017\_att\_val\_elo.csv (30 x 9)

nba\_2017\_att\_val\_elo\_with\_cluster.csv (30 x 10)

nba\_2017\_attendance.csv (30 x 5)

nba\_2017\_br.csv (486 x 30)

nba\_2017\_elo.csv (30 x 3)

nba\_2017\_endorsements.csv (10 x 4)

nba\_2017\_nba\_players\_with\_salary.csv (342 x 39)

nba\_2017\_pie.csv (486 x 22)

nba\_2017\_player\_wikipedia.csv (163k x 5)

nba\_2017\_players\_stats\_combined.csv (446 x 38)

nba\_2017\_players\_with\_salary\_wiki\_twitter.csv (239 x 42)

nba\_2017\_real\_plus\_minus.csv (8 columns)

nba\_2017\_salary.csv (449 x 4)

nba\_2017\_team\_valuations.csv (30 x 2)

nba\_2017\_twitter\_players.csv (329 x 3)

I will be using a dataset from Kaggle that compiles a list of current NBA players and their basketball statistics, financial metrics, and social media presence. In addition to this data being interesting to me as an avid basketball fan, the topic is also quite relevant to my work.

I am employed at GumGum Sports which is a computer vision sports marketing company. We digest television broadcasts and social media data to determine the sponsorship valuations for players and teams. I think I could gather some extremely valuable insights from this dataset regarding which types of players are more active and social media and how it correlates with their on-the-court success, salaries, etc. Hopefully at the end of this project I might have some interesting results to share with my team.