COVID-19 Pandemic and the NBA Isolation Zone

Ethan Sheldon
Georgia Institute of Technology
esheldon6@gatech.edu

Brooks Slocum
Georgia Institute of Technology
bslocum3@gatech.edu

Justin Nielson
Georgia Institute of Technology
jnielson3@gatech.edu

Konrad Richter
Georgia Institute of Technology
krichter6@gatech.edu

1 INTRODUCTION

Statistical models to predict player and team outcomes is of great interest to sports teams, TV networks, advertisers, the sports betting industry, reporters, researchers, fans etc. Therefore it is a much studied area of data analysis [8]. Especially in popular sports like basketball, large amounts of data are collected and analyzed for team executives and coaches to make decisions on player contracts, trades, style of play, and game-time decisions [3].

In order to keep the game running during a global pandemic, the National Basketball Association (NBA) has implemented a special security concept, the "Bubble". The NBA, in essence, is quarantining itself from the rest of society. It was one of the first professional sports league to discontinue play in mid-March after a player tested positive for COVID-19 and then create a plan to return to play.

The Bubble plan instituted by the NBA was to have 22 teams out of the 30 teams that had a statistical chance to make the playoffs play games beginning in August. All teams would play at the ESPN Wide World of Sports Complex and stay at three different Disney hotels, located at Walt Disney World Resort in Florida. The 22 teams that were invited into the bubble would all play eight "seeding games" that would determine the 16 teams that would go to the playoff rounds. Players, coaches, reporters, and support staff are not allowed out of the bubble for a period of up to three months and family were not allowed in the bubble unless they quarantined beforehand.

These special circumstances changed the traditional everyday life of the players extremely on and off the court. In our project, we want to analyze how the factors that determine player and game outcome in the Bubble have changed without travel [7] or fans at games [1] compared to prior seasons and playoffs using established quantitative methods and metrics. These methods and metrics include but are not limited to basketball possessions, offensive ratings, defensive ratings, plays, per-minute statistics, pace adjustments, true shooting percentage, effective field goal percentage, rebound rates, Four Factors, plus/minus statistics, counterpart statistics, linear weights metrics, individual possession usage and individual efficiency, Pythagorean method and Bell Curve method [5].

2 PROPOSED METHOD

Hustle stats method of accurately classifying team wins and losses in the NBA playoffs and has that changed in the 2019-20 bubble season.

Hustle stats or defensive intangibles are a fairly new set of metrics which started to be tracked by the NBA in the 2015-16 season. Hustle stats include 2-point and 3-point contested shots, deflections, loose-ball-recoveries, screen assists and charges. Additional hustle metrics of offensive and defensive box outs were added in the 2017-18 season. By pulling team and player hustle box score metrics over the last five seasons, analyzing the correlations, and running multiple clustering and classification models such as Logistic Regression, KNN, SVM, and Random Forest, we will determine if hustle stats are good predictors on the two classes (win or loss).

3 DESIGN OF EXPERIMENTS AND EVALUATION

Along with the bubble environment, player movement and injuries made a huge impact on the 2019-20 playoffs. As a result of player injuries, the Golden State Warriors went from five straight finals appearances and three championships to missing the playoffs entirely. In addition, 2018-19 Finals MVP Kwahi Leonard who played for the Toronto Raptors championship team joined the

Los Angeles Clippers the following season. The challenge for this experiment is to try to isolate the impact of the bubble while accounting for these non-bubble factors.

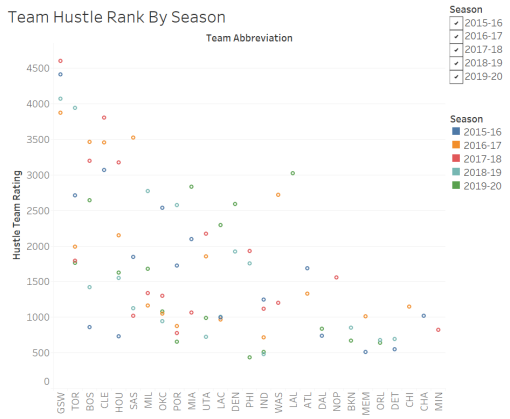

Figure 1: team hustle ranks

As you can see in figure 1, there is a good correlation of hustle stats ranking and the eventual NBA champion each season.

Total Points and Hustle Rating By Season

	Season				
	2015-16	2016-17	2017-18	2018-19	2019-20
PTS	189,716	184,560	181,086	190,064	186,155
Hustle Team Rating	26,733	31,275	30,857	26,488	24,258
Hustle Team Rating /PTS	0.14	0.17	0.17	0.14	0.13

PTS, Hustle Team Rating and Hustle Team Rating /PTS broken down by Season. The data is filtered on variance of Plus Minus and Team Name. The variance of Plus Minus filter includes everything. The Team Name filter has multiple members selected.

Figure 2: total points and hustle stats

However, when looking at hustle stats in figure 2, over the past five seasons it saw a sharp decline in the bubble playoffs of 2019-20 and prior 2018-19 playoffs as a percent of total points scored falling from 0.17 to 0.14 to 0.13. One explanation is the rise in 3-pt shot attempts and style of play that has moved to more quick pace that is offense focused over defense. The counter argument is championship teams in the last two playoffs the Toronto Raptors and the Los Angeles Lakers both ranked high in hustle stats. The NBA adage "defense wins championships" still holds true.

4 COMPLETED PLAN OF ACTIVITIES

In our project, we want to use a proven quantitative prediction approach [6] and incorporate the latest data from the Bubble. We want to find out how the game-decisive factors have changed with the particular game situation. This analysis is new, as there has never been a situation like the Bubble before.

The NBA bubble provides a natural experiment that controls for many factors such as travel [4], home-team preferences, and many more.

The first milestone is acquiring all of the necessary data and making it uniform using big data querying and wrangling tools such as SQLite, R, Python and Spark. The second key milestone is identifying trends through data analysis and version control using Git, Jupyter, and Google Colab. The mid-point would be applying statistical methods and machine learning to predict player and game outcomes in the Bubble. The final analysis and report would include data visualizations in Tableau on key metrics, predictions and methodology.

5 REVISED PLAN OF ACTIVITIES

All team members have contributed similar amount of effort. Below is a list of completed activities by team member and work in progress.

Justin Nielson: 1. Completed data pull of NBA hustle stats in Python for the last five playoffs from the api, cleaned data sets, merged team and player box scores on GAME ID and TEAM ID. 2. Created Hustle Rating column based on the NBA Hustle Stat composite formula outlined in the July, 2015 Grantland article[2]. 3. Exploratory Data Analysis on hustle stats by team, player, and playoff season and correlation with win-loss records (in progress). 4. Fit classification and clustering models to explore if hustle stats are a good way of predicting wins and losses and if that has changed in the 2019-20 bubble season playoffs (in progress). 5. Explain and visualize results of each model (in progress).

6 INNOVATIONS 2-4

Innovations here...

* One innovation is using hustle stats as a gauge of player value to the team in winning the NBA championship.

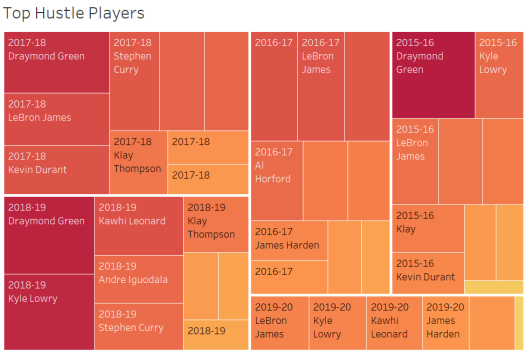

Figure 3: player hustle ranks

As you can see in figure 3, the top hustle players each playoffs have generally been members of the winning NBA championship team. Having multiple players on the hustle stats ranking is a good indicator of teams winning playoff games and the championship over the past five seasons.

For example, Draymond Green, Klay Thompson and Stephen Curry, who all play for the Golden State Warriors, were the top hustle players in the 2017-18 playoffs and ended up with them winning the championship. However, in the 2018-19 playoffs, while the same three GSW players ranked high in hustle stats two players Kyle Lowry and Kwahi Leonard from the eventual championship team the Toronto Raptors where also near the top of the hustle rankings. Injuries in the NBA Finals to Stephen Curry and Klay Thompson played a critical role in the Warriors not repeating as champion.

More recently, the Los Angeles Lakers won the championship in the 2019-20 bubble playoffs behind perennial All-Star and four-time league MVP Lebron James. In past seasons, Lebron James ranked high in hustle stats with his prior team the Cleveland Cavaliers.

REFERENCES

[1] Alex Bryson, Peter Dolton, J. James Reade, Dominik Schreyer, and Carl Singleton. 2020. *Experimental effects of an absent crowd on performances and refereeing decisions during Covid-19.* Economics Discussion Papers em-dp2020-18. Department of Economics, Reading University. https://ideas.repec.org/p/rdg/emxxdp/em-dp2020-18.html

[2] Jason Concepcion. 2015. How Do You Measure Hustle in the NBA? There's a Stat for That. https://grantland.com/the-triangle/how-do-you-measure-hustle-in-the-nba-theres-a-stat-for-that/

[3] Thomas H Davenport. 2014. Analytics in sports: The new science of winning. *International Institute for Analytics* 2 (2014), 1–28.

[4] Thomas Huyghe, Aaron T. Scanlan, Vincent J. Dalbo, and Julio Calleja-González. 2018. The Negative Influence of Air Travel on Health and Performance in the National Basketball Association: A Narrative Review. *Sports* 6, 3 (2018). https://doi.org/10.3390/sports6030089

[5] Justin Kubatko, Dean Oliver, Kevin Pelton, and Dan Rosenbaum. 2007. A Starting Point for Analyzing Basketball Statistics. *Journal of Quantitative Analysis in Sports* 3 (02 2007), 1–1. https://doi.org/10.2202/1559-0410.1070

[6] Kai Song, Qingrong Zou, and Jian Shi. 2018. Modelling the scores and performance statistics of NBA basketball games. *Communications in Statistics - Simulation and Computation* 0, 0 (2018), 1–13. https://doi.org/10.1080/03610918.2018.1520878 arXiv:https://doi.org/10.1080/03610918.2018.1520878

[7] Kyle Steenland and James A. Deddens. 1997. Effect of Travel and Rest on Performance of Professional Basketball Players. *Sleep* 20, 5 (05 1997), 366–369. https://doi.org/10.1093/sleep/20.5.366 arXiv:https://academic.oup.com/sleep/article-pdf/20/5/366/8596564/sleep-20-5-366.pdf

[8] Fabian Wunderlich and Daniel Memmert. 2020. Forecasting the outcomes of sports events: A review. *European Journal of Sport Science* 0, 0 (2020), 1–14. https://doi.org/10.1080/17461391.2020.1793002 PMID: 32628066.

List of Figures