MA678

Midterm Project Proposal

Daniel Xu

November 4th, 2021

### **Personal Statement**:

My career goal is to work as a data scientist in a technology company or consulting company that has a department related to sports analytics that I can get in touch with some projects about on-field sports analytics. I know that will be hard but I still want to mention that my dream job at present is to make into the basketball data science team of Los Angeles Lakers. As a basketball data scientist, I can travel with the team all season to do data analysis game by game, play by play to help improve the team's performance on the field and monitor players' physical stats to help them prevent injuries and prolong their careers. My project is made up of two data sets, one is about NBA(National Basketball Association) players’ stats from the season 1996-1997 to the season 2019-2020. In this data set, it contains every player’s biographic information, every season average stats. From this data set, I can have an obvious view of NBA players’ career data trajectory. Another is about NBA players’ injuries from the season 2010-2011 to the season 2019-2020. Obviously, I chose data sets that fit well with my current career plan. Thus, in this project I will have an initial contact with this kind of data which may help me get work under way more quickly if I have a job in the future.

### **Question**:

1. Can we find any evidence of the change of basketball game style from the changing trend of NBA player data?
2. Whether the types of injuries suffered by players have changed as the style of basketball has changed?
3. Whether the career-ending injuries type change or not?

### **data source**:

1. <https://www.kaggle.com/justinas/nba-players-data>
2. <https://www.kaggle.com/ghopkins/nba-injuries-2010-2018>

### **Proposed Timeline of work**:

* Proposal: 11/4
* EDA: 11/9-11/14
* Data Processing: 11/15-11/18
* Modeling and Validation: 11/19-11/25
* Write up: 11/26-12/1
* Submission: 12/2