**Must have features**

**Feature Selection Methodology:**

In our project, one of the most important steps in the data visualization process is feature exploration and selection where we can select inputs or features to help the end user understand and explore the NBA player’s information better.

Based on the original NBA players’ dataset, we don’t need to show every single data. All we need to do is to find the major features which could give best data understanding. For this project, we have to define the features that are potentially useful, and then a subset is chosen from the original pool of NBA Players’ features.

All features available in our dataset are listed below: ["player\_id","season\_id","league\_id","team\_id","team\_abbreviation","player\_age"

"gp","gs","min","position","fgm","fga","fg\_pct","fg3m","fg3a","fg3\_pct","ftm","fta","ft\_pct","oreb","dreb","reb","ast","stl","blk","tov","pf","pts","gp","gs","rank\_pg\_min","rank\_pg\_fgm","rank\_pg\_fga","rank\_fg\_pct","rank\_pg\_fg3m","rank\_pg\_fg3a","rank\_fg3\_pct","rank\_pg\_ftm","rank\_pg\_fta","rank\_ft\_pct","rank\_pg\_oreb","rank\_pg\_dreb","rank\_pg\_reb","rank\_pg\_ast","rank\_pg\_stl","rank\_pg\_blk","rank\_pg\_tov","rank\_pg\_pts","rank\_pg\_eff"]

From all of these features in our dataset, first, we divide the features into two group. One of the group of features represents the basic personal information of each player. Another group of features represents the performance information of each player. Second, based on those two feature groups, we choose the must features in each group to show the visualization elements and design different functions.

Group1 (player’s personal information): ["player\_id","season\_id","league\_id","team\_id","team\_abbreviation","player\_age"]

Group2 (player’s performance information):

["gp","gs","min","position","fgm","fga","fg\_pct","fg3m","fg3a","fg3\_pct","ftm","fta","ft\_pct","oreb","dreb","reb","ast","stl","blk","tov","pf","pts","gp","gs","rank\_pg\_min","rank\_pg\_fgm","rank\_pg\_fga","rank\_fg\_pct","rank\_pg\_fg3m","rank\_pg\_fg3a","rank\_fg3\_pct","rank\_pg\_ftm","rank\_pg\_fta","rank\_ft\_pct","rank\_pg\_oreb","rank\_pg\_dreb","rank\_pg\_reb","rank\_pg\_ast","rank\_pg\_stl","rank\_pg\_blk","rank\_pg\_tov","rank\_pg\_pts","rank\_pg\_eff"]

Must have features that we choose to tell the interesting story for the audiences are:

Group1 (player’s personal information): ["player\_id","season\_id","league\_id","team\_id","team\_abbreviation","player\_age"]

Group2 (player’s performance information)

["gp","gs","min","position","fgm","fga","fg\_pct","fg3m","fg3a","fg3\_pct","ftm","fta","ft\_pct","oreb","dreb","reb","ast","stl","blk","tov","pf","pts","gp","gs","rank\_pg\_fgm","rank\_pg\_fga","rank\_fg\_pct"]

**Optional features**

The features which are considered to be nice to have, but not critical are listed as follows:

Group2 (player’s performance information):

["rank\_pg\_fg3m","rank\_pg\_fg3a","rank\_fg3\_pct","rank\_pg\_ftm","rank\_pg\_fta","rank\_ft\_pct","rank\_pg\_oreb","rank\_pg\_dreb","rank\_pg\_reb","rank\_pg\_ast","rank\_pg\_stl","rank\_pg\_blk","rank\_pg\_tov","rank\_pg\_pts","rank\_pg\_eff"]

According to our visualization task, the first two important questions we want to solve are:

How does a player behave in different locations of the court?

What is a player’s overall performance?

Therefore, the ranking details of each play are optional features since those features are extra information. We only need to show the major rankings of each player in certain season such as the shooting score ranking or total performance ranking. In this way, people can see the major performance of each player and understand who are the top players based on basketball capabilities that people care.

**Time Table and Task Schedule**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| October | | | | | | | | | | | | | | | | | | | |
| Sunday | | | Monday | | | Tuesday | | | Wednesday | | | | Thursday | | | Friday | | Saturday | |
| 23 |  | 24 | | Proposal Due  Group Meeting | 25 | |  | 26 | | |  | 27 | | Group Meeting | 28 | | Obtain the data | 29 | Provide data structure |
|  |  | |  | |  | | |  | |  | |  |
| 30 | Choose basic visual model | 31 | |  |  | |  |  | | |  |  | |  |  | |  |  |  |
|  |  | |  | |  | | |  | |  | |  |
|  |  | |  | |  | | |  | |  | |  |
| November-December | | | | | | | | | | | | | | | | | | | |
| Sunday | | | Monday | | | Tuesday | | | Wednesday | | | | Thursday | | | Friday | | Saturday | |
|  |  |  | |  | 1 | | Group Meeting | 2 | | | Choose must visual model | 3 | | Group Meeting | 4 | | Project Check | 5 |  |
|  |  | |  | |  | | |  | |  | |  |
| 6 | Improve visual model | 7 | | Project Check | 8 | | Group Meeting | 9 | | Add methods manipulating data | | 10 | | Group Meeting | 11 | | Project Milestone due | 12 |  |
|  |  | |  | |  | |  | |  | |  |
| 13 | Improve visual model | 14 | | adds interaction | 15 | | Group Meeting | 16 | | | further clarify representation | 17 | | Group Meeting | 18 | | Project Check | 19 |  |
|  |  | |  | |  | | |  | |  | |  |
| 20 |  | 21 | | Iteration and Combination | 22 | | Group Meeting | 23 | | |  | 24 | | Group Meeting | 25 | | Project Check | 26 |  |
|  |  | |  | |  | | |  | |  | |  |
| 27 |  | 28 | | Summary | 29 | | Group Meeting | 30 | | |  | 1 | | Group Meeting | 2 | | Final Project Due |  |  |
|  |  | |  | |  | | |  | |  | |  |