Python Files:

1. Run data\_analysis.py
   1. This will create the data frames needed to run all regression, random forest, and decision tree models, as well as analyze much of the data.
   2. This will also populate the **Statistics** folder’s sub folders **All Players** and **Individual** **Players.** 
      1. Within the **All Players** sub folder: **Player Averages**, **Player Totals**, **Points VS FGA**, **Points VS Games**, and **Points VS Minutes** sub folders will be populated with the applicable charts.

*nba\_project > Statistics > All Players*

Text

Description automatically generated

* + 1. Within the **Individual Players** sub folder: **FGA VS Season**, **Games VS Season**, **Minutes VS Season**, and **Points VS Season** sub folders will be populated with the applicable charts.

*nba\_project > Statistics > Individual Players*

Text

Description automatically generated

* 1. This will also populate the **Heat Maps** folder’s sub folders **All Players** and **Individual** **Players** with the applicable charts.

*nba\_project > Heat Maps*

Text

Description automatically generated with medium confidence

1. Run regression.py
   1. This will print the regression outputs for both linear and quadratic regression.
   2. This will also populate the **Regression** folder, along with its subfolders: **Regression Models** and **Regression Predictions**.
      1. Within the **Regression Models** sub folder: **Points VS FGA**, **Points VS Games**, and **Points VS Minutes** sub folders will be populated with the applicable charts.

*nba\_project > Regression > Regression Models*

Text

Description automatically generated

* + 1. Within the **Regression Predictions** sub folder: **Points VS FGA**, **Points VS Games**, and **Points VS Minutes** sub folders will be populated with the applicable charts.

*nba\_project > Regression > Regression Predictions*

Text

Description automatically generated

1. Run lebron\_james\_forest\_tree.py
   1. This will print the random forest and decision predictions for LeBron James.
   2. This will also populate the **Random Forest** subfolder **LeBron James** with applicable charts.

*nba\_project > Random Forest*



1. Run james\_hard\_forest\_tree.py
   1. This will print the random forest and decision predictions for James Harden.
   2. This will also populate the **Random Forest** subfolder **James Harden** with applicable charts.

*nba\_project > Random Forest*



1. Run demar\_derozan\_forest\_tree.py
   1. This will print the random forest and decision predictions for DeMar Derozan
   2. This will also populate the **Random Forest** subfolder **DeMar Derozan** with applicable charts.

*nba\_project > Random Forest*



1. Feel free to explore all charts and analyze all system printed results. Additionally, citations for the project can be viewed in the **Citations.txt** file. Resourced data is stored in the following CSV files: **2012\_2013.csv**, **2013\_2014.csv**, **2014\_2015.csv**, **2015\_2016.csv**, **2016\_2017.csv**, **2017\_2018.csv**, **2018\_2019.csv**, **2019\_2020.csv**, **2020\_2021.csv**, **2021\_2022.csv**

**Folder Structure to Store Data**

* + Folder: Heat Maps
    - Subfolder: All Players
    - Subfolder: Individual Players
  + Folder: Random Forest
    - Subfolder: DeMar Derozan
    - Subfolder: James Harden
    - Subfolder: LeBron James
  + Folder: Regression
    - Subfolder: Regression Models
      * Subfolder: Points VS FGA
      * Subfolder: Points VS Games
      * Subfolder: Points VS Minutes
    - Subfolder: Regression Predictions
      * Subfolder: Points VS FGA
      * Subfolder: Points VS Games
      * Subfolder: Points VS Minutes
  + Folder: Statistics
    - Subfolder: All Players
      * Subfolder: Player Averages
      * Subfolder: Player Totals
      * Subfolder: Points VS FGA
      * Subfolder: Points VS Games
      * Subfolder: Points VS Minutes
    - Subfolder: Individual Players
      * Subfolder: FGA VS Season
      * Subfolder: Games VS Season
      * Subfolder: Minutes VS Season
      * Subfolder: Points VS Season