DS-GA 1007 Final Project User Guide
NBA Players' Performance: Real Life and Fantasy Sports Competition
Chuan-Ya Hsu

PROJECT DESCRIPTION

This program provides users the performance of NBA players within late 2015 to early 2016.

The dataset includes 36294 observations and 31 features which consist of the basic information of the players, the performances of the players, and the fantasy sports competition information from two different websites. This project allows users to examine the data in different aspects:

- For each player,
    - Learn the player's information,
    - Examine the real life scoring history, and
    - Examine the fantasy competition performance

- For each team,
    - Learn the information about the team,
    - Examine the points difference between the team and the opposite team over time, and
    - Examine on average, how the players in the team perform in both real life and fantasy competition

The program also allows users to add more data into the dataset. The data file which be allowed to read is csv (.csv) or excel (.xlsx) file.

ABOUT DATA

The dataset which be used in this project be found at:
https://www.dropbox.com/s/s52281dyd33lk8g/Players_w_spreads_2015.xls?dl=0

The original usage for the dataset is for examining the relationship between players' minutes in games and score spreads.
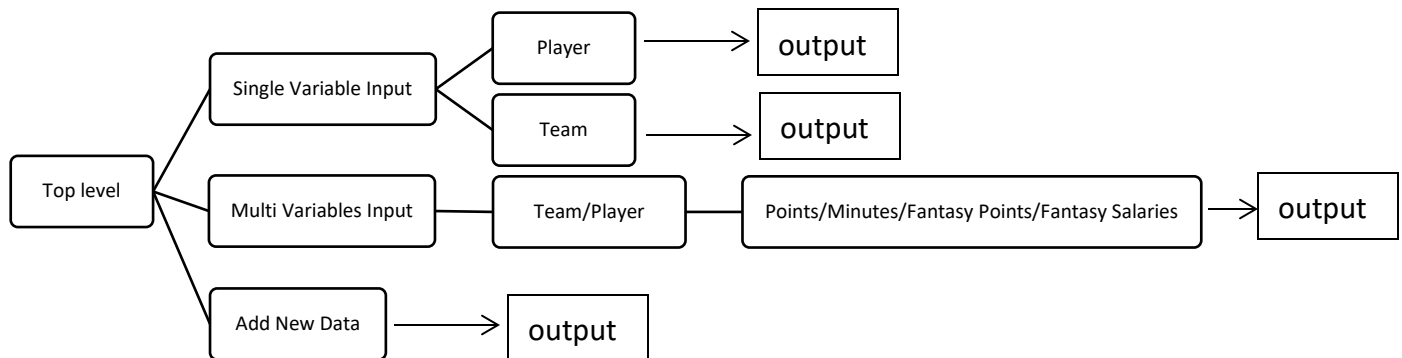
INSTALL

- Python 3.5 or above installed is needed
- Required python packages: Pandas, Numpy, matplotlib
    - ➔ All these packages can be installed with pip in the shell/terminal on Windows/Mac:
            pip install [package name]

HOW TO RUN

Use the shell/terminal, change the directory to where the main file located and enter

python main.py   -or-   python3 main.py

PROGRAM STRUCTURE

PROGRAM INPUT

The program allows users to enter S, M, A, and Q at the top level, string inputs for the following steps. Notice that the string inputs can only be the given options or names which include in the dataset.

The followings are the input explanations:
- S: Choosing single variable analysis
- M: Choosing two variables and examining the relationship between them
- A: Choosing to add new data.

    Notice that the new data must have the player id (GID) variable.
- Q: Quit the program
- [Player name]:   player's last name without '-' or other symbols
- [Team name]:   3 words abbreviation of the team
- pt/Minutes/FanPt/FanSal   Please enter these inputs as what the program given like

OPERATING PROCEDURE

The program consists three main functions: analysis single variable, examines the relationship between two variables, and adds more data. The following guide will discuss these three functions respectively.

In the top level, the program will ask users to enter S, M, A, or Q. After input the options, the program will separate into 4 conditions:

1. **Type in S:**
   ⇨ The program will ask users to enter the name of the player or the name of the team, or input Q to go back to the top level.
      ⇨ **Type in [player name]:**
         ⇨ Output the information of the player,
         ⇨ Output the simple statistics of the player's score, minutes in games, fantasy competition scores, and fantasy salaries,
         ⇨ Output the scoring history graph, and
         ⇨ Ouput the time series graph for fantasy salaries.
      ⇨ **Type in [team name]:**
         ⇨ Output the information of the team,
         ⇨ Output the simple statistics of the player's average score, minutes in games, fantasy competition scores, and fantasy salaries in the team,
         ⇨ Output the bar chart of the time change of the team points versus the opposite team, and
         ⇨ Output the curve of the spread (the points difference between two teams) of the team.

2. **Type in M:**
   ⇨ The program will ask users to click enter to keep go on, or input Q to go back to the top level.
      ⇨ **Type in [player name/team name]:**
         ⇨ **Type in pt/Minutes/FanPt/FanSal:**
            ⇨ Output the graph of the selected information for the player/team by time.

3. **Type in A:**
   ⇨ The program will ask users to input the path of the new data, or input Q to go back to the top level.
   Example:
      C:/~/another_players_spread.csv
   ⇨ Save the merged data, which called *'players_new.xlsx'*

4. **Type in Q:**
   ⇨ Quite the program.