|  |  |  |
| --- | --- | --- |
|  | **Document** | **Annotation** |
|  | **Root** |  |
| 0.0 | Draft.docx | Report |
| 1.1 | 2015-16\_playoffs\_refs\_complete.xlsx  2015-16\_season\_refs\_complete.xlsx | All extracted information  (GameID:int, Date:date, Hometeam:varchar(3), Awayteam:varchar(3), Homescore:int, Awayscore:int, HFTA:int, AFTA:int, Ref:string, Ref:string, Ref:string) |
| 1.2 | 2015-16\_playoffs\_refs\_complete.csv  2015-16\_season\_refs\_complete.csv | All extracted information  Generated by: 4.1  (GameID:int, Date:date, Hometeam:varchar(3), Awayteam:varchar(3), Homescore:int, Awayscore:int, HFTA:int, AFTA:int, Ref:string, Ref:string, Ref:string) |
| 1.3 | 2014-15\_playoffs\_refs\_1hot.csv  2014-15\_season\_refs\_1hot.csv  2015-16\_playoffs\_refs\_1hot.csv  2015-16\_season\_refs\_1hot.csv  2016-17\_playoffs\_refs\_1hot.csv  2016-17\_season\_refs\_1hot.csv | All extracted information, ref names are 1-hot encoded  Generated by: 2.4  (GameID:int, Date:date, Hometeam:varchar(3), Awayteam:varchar(3), Homescore:int, Awayscore:int, HFTA:int, AFTA:int, ‘Bennie Adams’:int, ‘Bill Kennedy’:int,…, Zach Zarba:int) |
| 1.4 | 2015-16\_season\_fta\_refs.csv | For each ref: avg FTAs per game and no. of games reffed  Generated by: 2.1  (pk:int, ‘FTA/Game’:double, Ref:string, ‘Refd Games’:int) |
| 1.5 | 2015-16\_season\_fta\_total.csv | For each team: total FTAs at home, total FTAs away and grand total FTAs  Generated by: 2.1  (team:varchar(3), HFTA:int, AFTA:int, Total:int) |
| 2.1 | FTAs.py | Reads: 1.3  Generates: 1.4, 1.5, 3.1, 3.2 |
| 2.15 | FTA\_Refs\_v\_Teams.py | Reads: 1.3, 1.4, 1.5  Generates: 3.25 |
| 2.2 | PointDifferentials.py | Reads: 1.3  Generates: 3.3 |
| 2.25 | PointDifferentialSampler.py | Reads: 1.3  Generates: 3.0 |
| 2.3 | RefDistributions.py | Incomplete?  Reads: 1.3 |
| 2.4 | Transform.py | Reads: 1.2  Generates: 1.3 |
|  | **Charts** |  |
| 3.0 | 2016\_season\_epd\_gsw\_scottfoster.png | Simulated mean point differentials by team and by ref’s proportion of home/away games  Generated by: 2.25 |
| 3.1 | 2016\_season\_fta\_corr.png | Correlation between total FTA @ home and total FTA away, for all teams  Generated by: 2.1 |
| 3.2 | 2016\_season\_fta\_refs.png | Box plot for average FTAs per game by refs  Generated by: 2.1 |
| 3.25 | 2016\_season\_fta\_refvteam.png | Heat map for avg FTA taken by teams by refs (ordered by rank desc)  Generated by: 2.15 |
| 3.3 | 2016\_season\_pd01\_BOS-MIL.png  2016\_season\_pd02\_ATL-UTA.png  2016\_season\_pd03\_GSW-SAS.png  2016\_season\_pdCLE.png  2016\_season\_pdGSW.png  2016\_season\_pdSAS.png | Histogram of point differentials for home and away for each team  Generated by: 2.2 |
|  | **Crawler** |  |
| 4.1 | basketballCrawlerNew.py | Crawls BasketballReference.com for ref data and FTAs  Imports: 4.3  Reads: 4.5  Generates: 1.2 |
| 4.2 | bballReferenceCleaner.py | Utility for cleaning up the table of games off BasketballReference.com (run season\_data\_clean in command line)  Reads: 4.4  Genenates: 4.5 |
| 4.3 | soup\_utils.py | Imported by: 4.1 |
|  | Crawler/Data |  |
| 4.4 | 2014-15\_season\_raw.csv  2014-15\_playoffs\_raw.csv  2015-16\_season\_raw.csv  2015-16\_playoffs\_raw.csv  2016-17\_season\_raw.csv  2016-17\_playoffs\_raw.csv | Copy and pasted table of all season’s games from BasketballReference.com  (Date:string, Start(ET):string, “Box Score”, Visitor/Neutral: string, PTS:int, Home/Neutral: string, PTS:int, OT?:string, Notes:string) |
| 4.5 | 2014-15\_season\_clean.csv  2014-15\_playoffs\_clean.csv  2015-16\_season\_clean.csv  2015-16\_playoffs\_clean.csv  2016-17\_season\_clean.csv  2016-17\_playoffs\_clean.csv | Cleaned up 4.4 for crawler  Generated by: 4.2  (GameID:int, Date:date, Hometeam:varchar(3), Awayteam:varchar(3), Homescore:int, Awayscore:int) |
| 1.2 | 2014-15\_season\_refs\_complete.csv  2014-15\_playoffs\_refs\_complete.csv  2015-16\_season\_refs\_complete.csv  2015-16\_playoffs\_refs\_complete.csv  2016-17\_season\_refs\_complete.csv  2016-17\_playoffs\_refs\_complete.csv | All extracted information  (GameID:int, Date:date, Hometeam:varchar(3), Awayteam:varchar(3), Homescore:int, Awayscore:int, HFTA:int, AFTA:int, Ref:string, Ref:string, Ref:string) |
|  | **Archives** |  |
|  |  |  |