Final Project

Baseball Statistics

[Fall 2015] DataVisualizations, Cannata

12/7/15

Brian Bargas, German Martinez, Chris Pena

Data:

Our Data was obtained from the Sean Lahman Online Baseball Database. It contains complete batting and pitching statistics from 1871 to 2014, plus fielding statistics, standings, team stats, managerial records, post-season data, and more. Specifically we took the master, hall of fame, and xxx cvs.

Once we had this csv chosen we reformatted them using the ETL.R scripts then imported them into our oracle database.

The csv we used were;

Batting

Hall of Fame

Master

Salary

Team

TeamFrancise

summary(df1)

%head(df)

summary(df2)

%head(df)

summary(df3)

%head(df)

summary(df4)

%head(df)

summary(df5)

%head(df)

summary(df6)

%head(df)

Visualization:

![](./GamesToHOFVotes.jpg)

![](./HallOfFameGeo.jpg)

![](./HallOfFameVotes.jpg)

![](./HomerunAVGatBat.jpg)

![](./SalaryToWins.jpg)

![](./TeamBattingAvg.jpg)

![](./TeamRecord.jpg)

Shiny: