Eras of Baseball

The Fighting Franconas

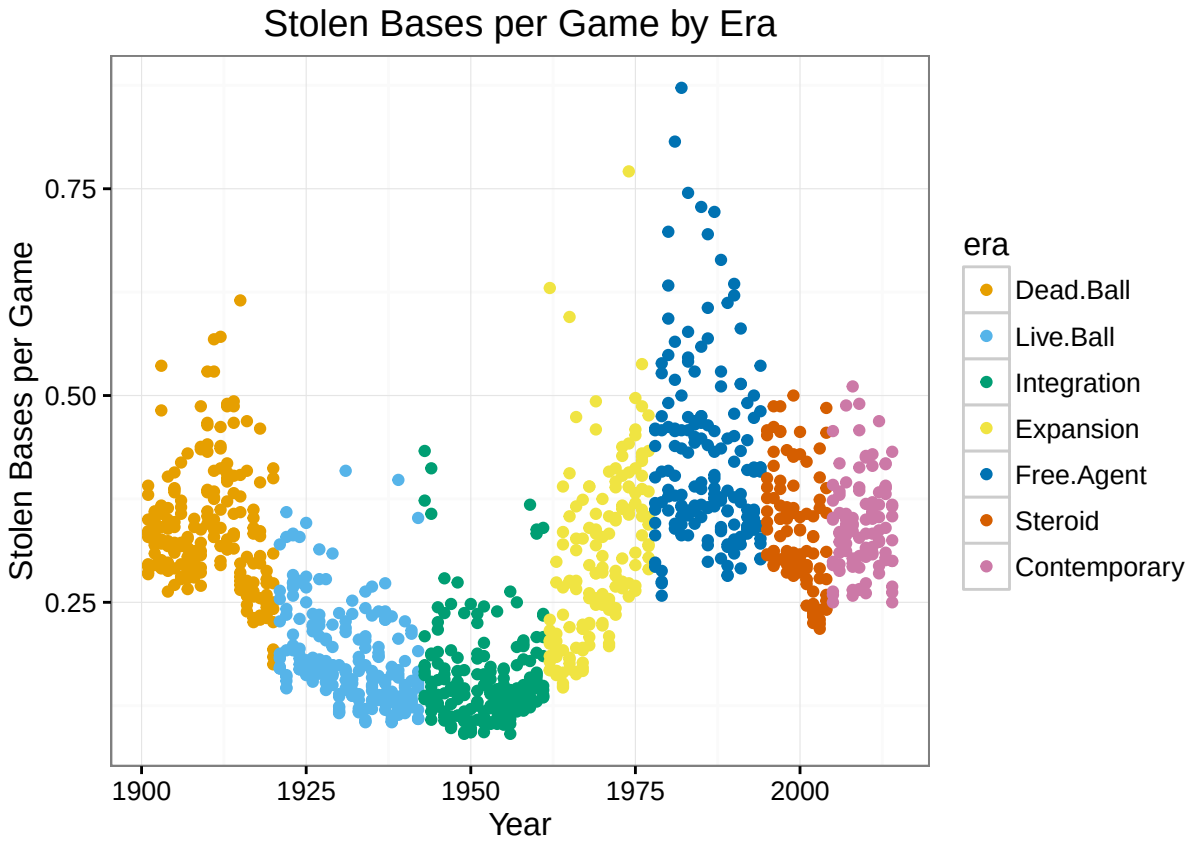
April 27, 2016

```
## Loading required package: Lahman

## Loading required package: dplyr

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

## Loading required package: ggplot2

## Warning: package 'ggplot2' was built under R version 3.2.4

## Loading required package: rmarkdown
```

Introduction

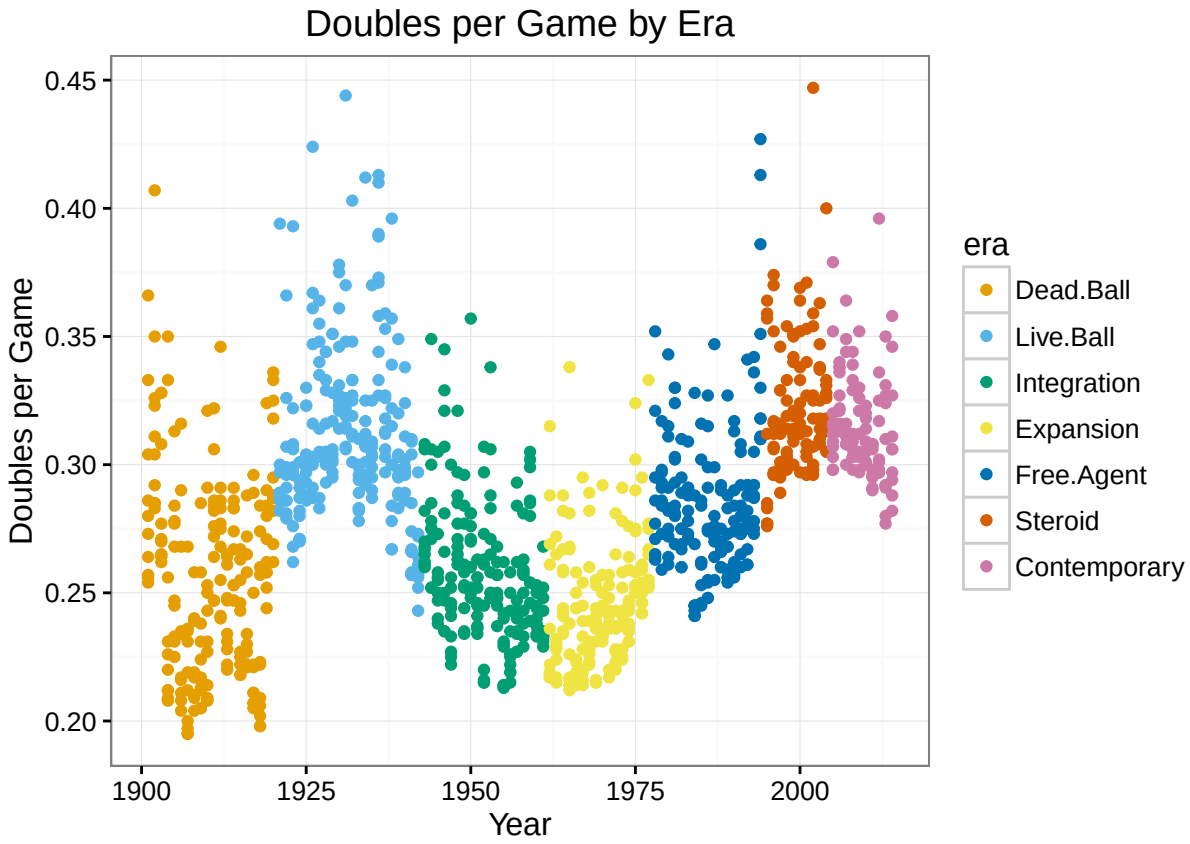
As an examination of the eras of baseball, our group has chosen to look at stolen bases per game, doubles per game, and triples per game. These era-by-era comparisons reveal fascinating aspects about the evolution of baseball, and each show a different change in the game.

```
data("Master")
cbPalette <- c("#E69F00", "#56B4E9", "#009E73", "#F0E442",
               "#0072B2", "#D55E00", "#CC79A7", "#999999")
Batting.Eras <- left_join(batting, Master, by = "playerID")
Batting.Eras %>%
  mutate(name=(paste(nameFirst,nameLast,sep=" "))) %>%
  select(yearID,G,AB,R,H,X2B,X3B,HR,RBI,SB,CS,BB,SO,IBB,HBP,SH,SF,GIDP,era,name) %>%
  filter(AB>=130) %>%
  group_by(yearID) %>%
  mutate(SBRate = round((SB)/(G),3)) %>%
  top_n(10,SBRate) %>%
  ggplot(., aes(yearID, SBRate)) + geom_point(aes(col=era), pch=19) + xlab("Year") + ylab("Stolen Bases
```



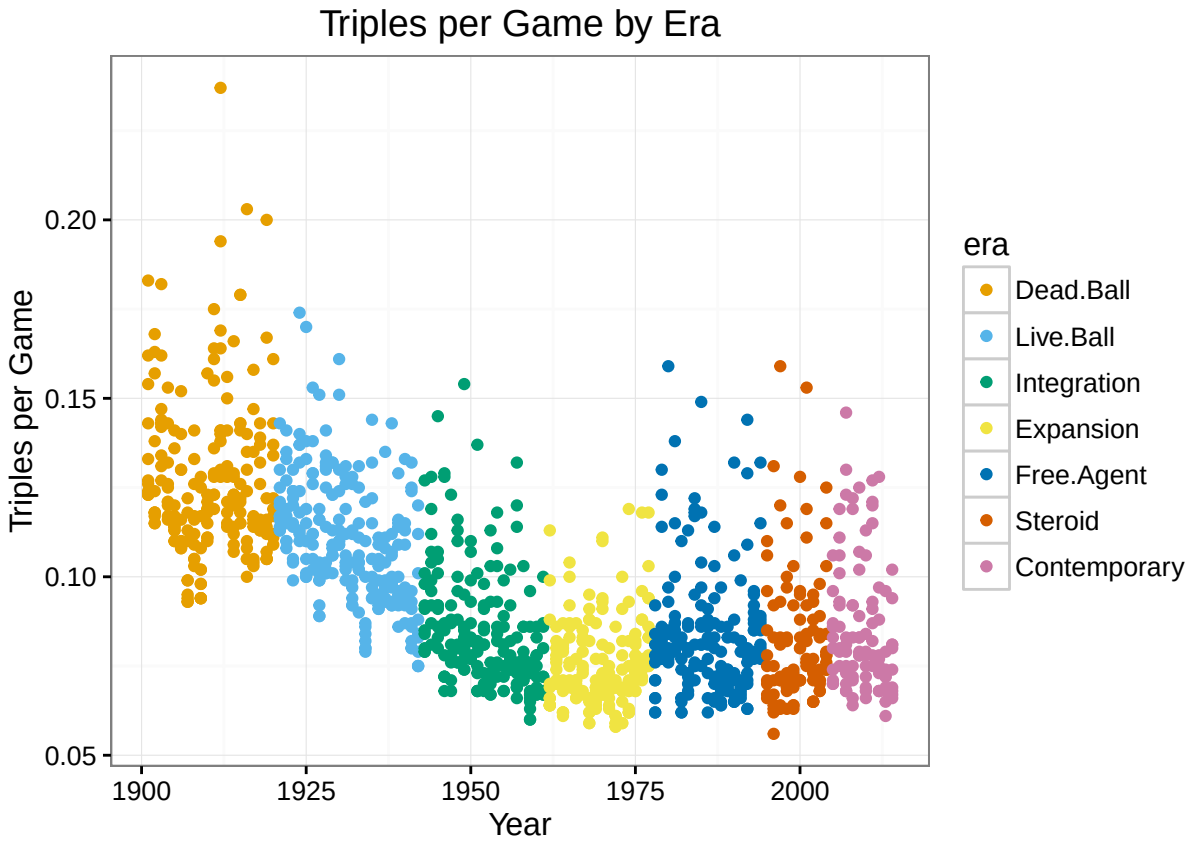
Stolen bases per game by era shows a uniquely high output of steals in the Free Agent Era(all top ten players in this category in this era had more than .25 steals per game), and subsequently high numbers (although this stat appears to be on the wane in recent years) in the Steroid and Contemporary Eras. These results may be indicative of the increased use of amphetamines (also known as "greenies") and cocaine that proliferated in the Major Leagues in the 1980s, although the somewhat higher than average steals per game outputs of the last few years may be better explained by a more aggressive stealing strategy and more value being placed in speedy players.

```
Batting.Eras %>%
  mutate(name=(paste(nameFirst,nameLast,sep=" "))) %>%
  select(yearID,G,AB,R,H,X2B,X3B,HR,RBI,SB,CS,BB,SO,IBB,HBP,SH,SF,GIDP,era,name) %>%
  filter(AB>=130) %>%
  group_by(yearID) %>%
  mutate(DoubleRate = round((X2B)/(G),3)) %>%
  top_n(10,DoubleRate) %>%
  ggplot(., aes(yearID, DoubleRate)) + geom_point(aes(col=era), pch=19) + xlab("Year") + ylab("Doubles
```



Doubles per game has a unique up-and-down appearance, wherein the Live Ball Era (all but one player in the top ten yearly in this era was over .25 doubles per game) and the recent Steroid and Contemporary Eras are the leading eras. This is best explained by the newfound power of batters in the Live Ball Era, and in the more recent Steroid Era, as more balls find there way deep into the outfield and drop in for extra-base hits.

```
Batting.Eras %>%
  mutate(name=(paste(nameFirst,nameLast,sep=" "))) %>%
  select(yearID,G,AB,R,H,X2B,X3B,HR,RBI,SB,CS,BB,SO,IBB,HBP,SH,SF,GIDP,era,name) %>%
  filter(AB>=130) %>%
  group_by(yearID) %>%
  mutate(TripleRate = round((X3B)/(G),3)) %>%
  top_n(10,TripleRate) %>%
  ggplot(., aes(yearID, TripleRate)) + geom_point(aes(col=era), pch=19) + xlab("Year") + ylab("Triples
```



Triples per game has a more consistent level throughout the last 50 years, hovering mostly below .1 triples per game. The greatest exceptions to this rule are the Dead Ball Era (where most of the top ten players in this statistic were over .1 triples per game) and the Live Ball Era (when the level begins to dip below .1 triples per game). The high output of triples per game is indicative of the lack of power in the game in those early years coupled with poor fielding and speedy lineups (note that the steal rate in the Dead Ball Era was also similar to that seen in the Steroid and Contemporary Eras).