Money Ball Project

BACKGROUND

The 2002 Oakland A's

The Oakland Athletics' 2002 season was the team's 35th in Oakland, California. It was also the 102nd season in franchise history. The Athletics finished first in the American League West with a record of 103-59.

The Athletics' 2002 campaign ranks among the most famous in franchise history. Following the 2001 season, Oakland saw the departure of three key players (the lost boys). Billy Beane, the team's general manager, responded with a series of under-the-radar free agent signings. The new-look Athletics, despite a comparative lack of star power, surprised the baseball world by besting the 2001 team's regular season record. The team is most famous, however, for winning 20 consecutive games between August 13 and September 4, 2002.[1] The Athletics' season was the subject of Michael Lewis' 2003 book Moneyball: The Art of Winning an Unfair Game (as Lewis was given the opportunity to follow the team around throughout that season)

Source: Wikipedia

Moneyball Book

The central premise of book Moneyball is that the collective wisdom of baseball insiders (including players, managers, coaches, scouts, and the front office) over the past century is subjective and often flawed. Statistics such as stolen bases, runs batted in, and batting average, typically used to gauge players, are relics of a 19th-century view of the game and the statistics available at that time. The book argues that the Oakland A's' front office took advantage of more analytical gauges of player performance to field a team that could better compete against richer competitors in Major League Baseball (MLB).

Rigorous statistical analysis had demonstrated that on-base percentage and slugging percentage are better indicators of offensive success, and the A's became convinced that these qualities were cheaper to obtain on the open market than more historically valued qualities such as speed and contact. These observations often flew in the face of conventional baseball wisdom and the beliefs of many baseball scouts and executives.

By re-evaluating the strategies that produce wins on the field, the 2002 Athletics, with approximately US 44 million dollars in salary, were competitive with larger market teams such as the New York Yankees, who spent over US$125 million in payroll that same season.

Because of the team's smaller revenues, Oakland is forced to find players undervalued by the market, and their system for finding value in undervalued players has proven itself thus far. This approach brought the A's to the playoffs in 2002 and 2003.

In this project we will work with some data and with the goal of trying to find replacement players for the ones lost at the start of the off-season - During the 2001–02 offseason, the team lost three key free agents to larger market teams: 2000 AL MVP Jason Giambi to the New York Yankees, outfielder Johnny Damon to the Boston Red Sox, and closer Jason Isringhausen to the St. Louis Cardinals.

Data

We will be using data from Sean Lahaman's Website a very useful source for baseball statistics. The documentation for the csv files is in the readme2013.txt file. You may need to reference this to understand what acronyms stand for.

Importing Libraries

```
library(ggplot2)

library(dplyr)
```

Importing dataset

batting <- read.csv('C:\\Users\\Newton\\Desktop\\Batting.csv')

#Checking out the columns in the dataframe

head(batting)

```
   playerID yearID stint teamID lgID  G G_batting AB R H X2B X3B HR RBI SB CS BB SO IBB HBP SH SF GIDP G_old
1 aardsda01   2004     1    SFN   NL 11        11  0 0 0   0   0  0   0  0  0  0  0   0   0  0  0     0    11
2 aardsda01   2006     1    CHN   NL 45        43  2 0 0   0   0  0   0  0  0  0  0   0   0  1  0     0    45
3 aardsda01   2007     1    CHA   AL 25         2  0 0 0   0   0  0   0  0  0  0  0   0   0  0  0     0     2
4 aardsda01   2008     1    BOS   AL 47         5  1 0 0   0   0  0   0  0  0  0  1   0   0  0  0     0     5
5 aardsda01   2009     1    SEA   AL 73         3  0 0 0   0   0  0   0  0  0  0  0   0   0  0  0     0    NA
6 aardsda01   2010     1    SEA   AL 53         4  0 0 0   0   0  0   0  0  0  0  0   0   0  0  0     0    NA
```

#Using str() to check the structure of the dataset

str(batting)

```
'data.frame':	97889 obs. of  24 variables:
 $ playerID : Factor w/ 18107 levels "aardsda01","aaronha01",..: 1 1 1 1 1 1 1 2 2 2 ...
 $ yearID   : int  2004 2006 2007 2008 2009 2010 2012 1954 1955 1956 ...
 $ stint    : int  1 1 1 1 1 1 1 1 1 1 ...
 $ teamID   : Factor w/ 149 levels "ALT","ANA","ARI",..: 117 35 33 16 116 116 93 80 80 80 ...
 $ lgID     : Factor w/ 6 levels "AA","AL","FL",..: 4 4 2 2 2 2 2 4 4 4 ...
 $ G        : int  11 45 25 47 73 53 1 122 153 153 ...
 $ G_batting: int  11 43 2 5 3 4 NA 122 153 153 ...
 $ AB       : int  0 2 0 1 0 0 NA 468 602 609 ...
 $ R        : int  0 0 0 0 0 0 NA 58 105 106 ...
 $ H        : int  0 0 0 0 0 0 NA 131 189 200 ...
 $ X2B      : int  0 0 0 0 0 0 NA 27 37 34 ...
 $ X3B      : int  0 0 0 0 0 0 NA 6 9 14 ...
 $ HR       : int  0 0 0 0 0 0 NA 13 27 26 ...
 $ RBI      : int  0 0 0 0 0 0 NA 69 106 92 ...
 $ SB       : int  0 0 0 0 0 0 NA 2 3 2 ...
 $ CS       : int  0 0 0 0 0 0 NA 2 1 4 ...
 $ BB       : int  0 0 0 0 0 0 NA 28 49 37 ...
 $ SO       : int  0 0 0 1 0 0 NA 39 61 54 ...
 $ IBB      : int  0 0 0 0 0 0 NA NA 5 6 ...
 $ HBP      : int  0 0 0 0 0 0 NA 3 3 2 ...
 $ SH       : int  0 1 0 0 0 0 NA 6 7 5 ...
 $ SF       : int  0 0 0 0 0 0 NA 4 4 7 ...
 $ GIDP     : int  0 0 0 0 0 0 NA 13 20 21 ...
 $ G_old    : int  11 45 2 5 NA NA NA 122 153 153 ...
```

Feature Engineering

We need to add three more statistics that were used in Moneyball! These are:

- Batting Average
  - AVG = H/AB
- On Base Percentage
  - OBP = (H + BB + HBP) / (AB + BB + HBP + SF)
- Slugging Percentage
  - SLG = (s + 2d + 3t + 4hr)/ AB or SLG = (h + d + 2t + 3hr) / AB

```
# On Base Percentage
batting$OBP <- (batting$H + batting$BB + batting$HBP)/(batting$AB + batting$BB + batting$HBP + batting$SF)

# Creating X1B (Singles)
batting$X1B <- batting$H - batting$X2B - batting$X3B - batting$HR

# Creating Slugging Average (SLG)
```

batting$SLG <- ((1 * batting$X1B) + (2 * batting$X2B) + (3 * batting$X3B) + (4 * batting$HR)) / batting$AB

Merging Salary Data with Batting Data

We know we don't just want the best players, we want the most undervalued players, meaning we will also need to know current salary information! We have salary information in the csv file 'Salaries.csv'.

#Loading the Salaries.csv file into a data frame
sal <- read.csv('Salaries.csv')

Use summary to get a summary of the batting data frame and notice the minimum year in the yearID column. Our batting data goes back to 1871! Our salary data starts at 1985, meaning we need to remove the batting data that occurred before 1985.

summary(batting)

```
     playerID         yearID         stint           teamID        lgID       
 mcguide01:   31   Min.   :1871   Min.   :1.000   CHN    : 4720   AA  : 1890  
 henderi01:   29   1st Qu.:1931   1st Qu.:1.000   PHI    : 4621   AL  :44369  
 newsobo01:   29   Median :1970   Median :1.000   PIT    : 4575   FL  :  470  
 johnto01 :   28   Mean   :1962   Mean   :1.077   SLN    : 4535   NL  :49944  
 kaatji01 :   28   3rd Qu.:1995   3rd Qu.:1.000   CIN    : 4393   PL  :  147  
 ansonca01:   27   Max.   :2013   Max.   :5.000   CLE    : 4318   UA  :  332  
 (Other)  :97717                                  (Other):70727   NA's:  737  
      G            G_batting            AB              R         
 Min.   :  1.00   Min.   :  0.00   Min.   :  0.0   Min.   :  0.00  
 1st Qu.: 13.00   1st Qu.:  7.00   1st Qu.:  9.0   1st Qu.:  0.00  
 Median : 35.00   Median : 32.00   Median : 61.0   Median :  5.00  
 Mean   : 51.65   Mean   : 49.13   Mean   :154.1   Mean   : 20.47  
 3rd Qu.: 81.00   3rd Qu.: 81.00   3rd Qu.:260.0   3rd Qu.: 31.00  
 Max.   :165.00   Max.   :165.00   Max.   :716.0   Max.   :192.00  
                  NA's   :1406     NA's   :6413    NA's   :6413    
      H              X2B             X3B              HR        
 Min.   :  0.00   Min.   : 0.0   Min.   : 0.000   Min.   : 0.000  
 1st Qu.:  1.00   1st Qu.: 0.0   1st Qu.: 0.000   1st Qu.: 0.000  
 Median : 12.00   Median : 2.0   Median : 0.000   Median : 0.000  
 Mean   : 40.37   Mean   : 6.8   Mean   : 1.424   Mean   : 3.002  
 3rd Qu.: 66.00   3rd Qu.:10.0   3rd Qu.: 2.000   3rd Qu.: 3.000  
 Max.   :262.00   Max.   :67.0   Max.   :36.000   Max.   :73.000  
 NA's   :6413     NA's   :6413   NA's   :6413     NA's   :6413    
     RBI              SB               CS              BB        
 Min.   :  0.00   Min.   :  0.000   Min.   : 0.000   Min.   :  0.00  
 1st Qu.:  0.00   1st Qu.:  0.000   1st Qu.: 0.000   1st Qu.:  0.00  
 Median :  5.00   Median :  0.000   Median : 0.000   Median :  4.00  
 Mean   : 18.47   Mean   :  3.265   Mean   : 1.385   Mean   : 14.21  
 3rd Qu.: 28.00   3rd Qu.:  2.000   3rd Qu.: 1.000   3rd Qu.: 21.00  
 Max.   :191.00   Max.   :138.000   Max.   :42.000   Max.   :232.00  
 NA's   :6837     NA's   :7713      NA's   :29867    NA's   :6413    
      SO              IBB              HBP              SH        
 Min.   :  0.00   Min.   :  0.00   Min.   : 0.000   Min.   : 0.000  
 1st Qu.:  2.00   1st Qu.:  0.00   1st Qu.: 0.000   1st Qu.: 0.000  
 Median : 11.00   Median :  0.00   Median : 0.000   Median : 1.000  
 Mean   : 21.95   Mean   :  1.28   Mean   : 1.136   Mean   : 2.564  
 3rd Qu.: 31.00   3rd Qu.:  1.00   3rd Qu.: 1.000   3rd Qu.: 3.000  
 Max.   :223.00   Max.   :120.00   Max.   :51.000   Max.   :67.000  
 NA's   :14251    NA's   :42977    NA's   :9233     NA's   :12751   
      SF              GIDP            G_old             BA        
 Min.   : 0.0    Min.   : 0.00   Min.   :  0.00   Min.   :0.000  
 1st Qu.: 0.0    1st Qu.: 0.00   1st Qu.: 11.00   1st Qu.:0.148  
 Median : 0.0    Median : 1.00   Median : 34.00   Median :0.231  
 Mean   : 1.2    Mean   : 3.33   Mean   : 50.99   Mean   :0.209  
 3rd Qu.: 2.0    3rd Qu.: 5.00   3rd Qu.: 82.00   3rd Qu.:0.275  
 Max.   :19.0    Max.   :36.00   Max.   :165.00   Max.   :1.000  
 NA's   :42446   NA's   :32521   NA's   :5189     NA's   :13520  
      OBP              X1B              SLG        
 Min.   :0.00    Min.   :  0.00   Min.   :0.000  
 1st Qu.:0.19    1st Qu.:  1.00   1st Qu.:0.179  
 Median :0.29    Median :  9.00   Median :0.309  
 Mean   :0.26    Mean   : 29.14   Mean   :0.291  
 3rd Qu.:0.34    3rd Qu.: 48.00   3rd Qu.:0.397  
```

```
 Max.   :1.00    Max.   :225.00   Max.   :4.000
 NA's   :49115   NA's   :6413     NA's   :13520
```

#Using subset() to reassign batting to only contain data from 1985 and onwards.

batting <- subset(batting,yearID >= 1985)

Using summary again to make sure the subset reassignment worked

summary(batting)

```
      playerID         yearID         stint          teamID      lgID
 moyerja01:   27   Min.   :1985   Min.   :1.00   SDN    : 1313   AA:    0
 mulhote01:   26   1st Qu.:1993   1st Qu.:1.00   CLE    : 1306   AL:17226
 weathda01:   26   Median :2000   Median :1.00   PIT    : 1299   FL:    0
 maddugr01:   25   Mean   :2000   Mean   :1.08   NYN    : 1297   NL:18426
 sierrru01:   25   3rd Qu.:2007   3rd Qu.:1.00   BOS    : 1279   PL:    0
 thomeji01:   25   Max.   :2013   Max.   :4.00   CIN    : 1279   UA:    0
 (Other)  :35498                                 (Other):27879
       G           G_batting           AB              R
 Min.   :  1.0   Min.   :  0.00   Min.   :  0.0   Min.   :  0.00
 1st Qu.: 14.0   1st Qu.:  4.00   1st Qu.:  3.0   1st Qu.:  0.00
 Median : 34.0   Median : 27.00   Median : 47.0   Median :  4.00
 Mean   : 51.7   Mean   : 46.28   Mean   :144.7   Mean   : 19.44
 3rd Qu.: 77.0   3rd Qu.: 77.00   3rd Qu.:241.0   3rd Qu.: 30.00
 Max.   :163.0   Max.   :163.00   Max.   :716.0   Max.   :152.00
                 NA's   :1406     NA's   :4377    NA's   :4377
       H              X2B              X3B              HR
 Min.   :  0.00   Min.   : 0.000   Min.   : 0.000   Min.   : 0.000
 1st Qu.:  0.00   1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.: 0.000
 Median :  8.00   Median : 1.000   Median : 0.000   Median : 0.000
 Mean   : 37.95   Mean   : 7.293   Mean   : 0.824   Mean   : 4.169
 3rd Qu.: 61.00   3rd Qu.:11.000   3rd Qu.: 1.000   3rd Qu.: 5.000
 Max.   :262.00   Max.   :59.000   Max.   :23.000   Max.   :73.000
 NA's   :4377     NA's   :4377     NA's   :4377     NA's   :4377
      RBI              SB               CS              BB
 Min.   :  0.00   Min.   :  0.000   Min.   : 0.000   Min.   :  0.00
 1st Qu.:  0.00   1st Qu.:  0.000   1st Qu.: 0.000   1st Qu.:  0.00
 Median :  3.00   Median :  0.000   Median : 0.000   Median :  3.00
 Mean   : 18.41   Mean   :  2.811   Mean   : 1.219   Mean   : 14.06
 3rd Qu.: 27.00   3rd Qu.:  2.000   3rd Qu.: 1.000   3rd Qu.: 21.00
 Max.   :165.00   Max.   :110.000   Max.   :29.000   Max.   :232.00
 NA's   :4377     NA's   :4377      NA's   :4377     NA's   :4377
       SO              IBB              HBP              SH
 Min.   :  0.00   Min.   :  0.000   Min.   : 0.000   Min.   : 0.000
 1st Qu.:  1.00   1st Qu.:  0.000   1st Qu.: 0.000   1st Qu.: 0.000
 Median : 12.00   Median :  0.000   Median : 0.000   Median : 0.000
 Mean   : 27.03   Mean   :  1.171   Mean   : 1.273   Mean   : 1.465
 3rd Qu.: 42.00   3rd Qu.:  1.000   3rd Qu.: 1.000   3rd Qu.: 2.000
 Max.   :223.00   Max.   :120.000   Max.   :35.000   Max.   :39.000
 NA's   :4377     NA's   :4378      NA's   :4387     NA's   :4377
       SF              GIDP           G_old             BA
 Min.   : 0.000   Min.   : 0.00   Min.   :  0.0   Min.   :0.000
 1st Qu.: 0.000   1st Qu.: 0.00   1st Qu.: 11.0   1st Qu.:0.136
 Median : 0.000   Median : 1.00   Median : 32.0   Median :0.233
 Mean   : 1.212   Mean   : 3.25   Mean   : 49.7   Mean   :0.205
 3rd Qu.: 2.000   3rd Qu.: 5.00   3rd Qu.: 77.0   3rd Qu.:0.274
 Max.   :17.000   Max.   :35.00   Max.   :163.0   Max.   :1.000
 NA's   :4378     NA's   :4377    NA's   :5189    NA's   :8905
      OBP              X1B              SLG
 Min.   :0.000   Min.   :  0.00   Min.   :0.000
 1st Qu.:0.188   1st Qu.:  0.00   1st Qu.:0.167
 Median :0.296   Median :  6.00   Median :0.333
 Mean   :0.262   Mean   : 25.66   Mean   :0.304
 3rd Qu.:0.342   3rd Qu.: 42.00   3rd Qu.:0.423
 Max.   :1.000   Max.   :225.00   Max.   :4.000
 NA's   :8821    NA's   :4377     NA's   :8905
```

#Using the merge() function to merge the batting and sal data frames by c('playerID','yearID').

combo <- merge(batting,sal,by=c('playerID','yearID'))

#Use summary to check the data

summary(combo)

```
     playerID         yearID         stint          teamID.x     lgID.x
 moyerja01:   27   Min.   :1985   Min.   :1.000   LAN    :  940   AA:    0
 thomeji01:   25   1st Qu.:1993   1st Qu.:1.000   PHI    :  937   AL:12292
 weathda01:   25   Median :1999   Median :1.000   BOS    :  935   FL:    0
 vizquom01:   24   Mean   :1999   Mean   :1.098   NYA    :  928   NL:13105
 gaettga01:   23   3rd Qu.:2006   3rd Qu.:1.000   CLE    :  920   PL:    0
 griffke02:   23   Max.   :2013   Max.   :4.000   SDN    :  914   UA:    0
 (Other)  :25250                                  (Other):19823
      G            G_batting           AB              R
 Min.   :  1.00   Min.   :  0.00   Min.   :  0.0   Min.   :  0.00
 1st Qu.: 26.00   1st Qu.:  8.00   1st Qu.:  5.0   1st Qu.:  0.00
 Median : 50.00   Median : 42.00   Median : 85.0   Median :  9.00
 Mean   : 64.06   Mean   : 57.58   Mean   :182.4   Mean   : 24.71
 3rd Qu.:101.00   3rd Qu.:101.00   3rd Qu.:336.0   3rd Qu.: 43.00
 Max.   :163.00   Max.   :163.00   Max.   :716.0   Max.   :152.00
                  NA's   :906      NA's   :2661    NA's   :2661
      H              X2B             X3B              HR
 Min.   :  0.00   Min.   : 0.000   Min.   : 0.000   Min.   : 0.000
 1st Qu.:  1.00   1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.: 0.000
 Median : 19.00   Median : 3.000   Median : 0.000   Median : 1.000
 Mean   : 48.18   Mean   : 9.276   Mean   : 1.033   Mean   : 5.369
 3rd Qu.: 87.25   3rd Qu.:16.000   3rd Qu.: 1.000   3rd Qu.: 7.000
 Max.   :262.00   Max.   :59.000   Max.   :23.000   Max.   :73.000
 NA's   :2661     NA's   :2661     NA's   :2661     NA's   :2661
     RBI              SB              CS              BB
 Min.   :  0.00   Min.   :  0.000   Min.   : 0.00   Min.   :  0.00
 1st Qu.:  0.00   1st Qu.:  0.000   1st Qu.: 0.00   1st Qu.:  0.00
 Median :  8.00   Median :  0.000   Median : 0.00   Median :  6.00
 Mean   : 23.56   Mean   :  3.568   Mean   : 1.54   Mean   : 17.98
 3rd Qu.: 39.00   3rd Qu.:  3.000   3rd Qu.: 2.00   3rd Qu.: 29.00
 Max.   :165.00   Max.   :110.000   Max.   :29.00   Max.   :232.00
 NA's   :2661     NA's   :2661      NA's   :2661    NA's   :2661
      SO              IBB             HBP              SH
 Min.   :  0.00   Min.   :  0.000   Min.   : 0.000   Min.   : 0.000
 1st Qu.:  2.00   1st Qu.:  0.000   1st Qu.: 0.000   1st Qu.: 0.000
 Median : 20.00   Median :  0.000   Median : 0.000   Median : 0.000
 Mean   : 33.52   Mean   :  1.533   Mean   : 1.614   Mean   : 1.786
 3rd Qu.: 55.00   3rd Qu.:  2.000   3rd Qu.: 2.000   3rd Qu.: 2.000
 Max.   :223.00   Max.   :120.000   Max.   :35.000   Max.   :39.000
 NA's   :2661     NA's   :2662      NA's   :2670     NA's   :2661
      SF              GIDP            G_old             BA
 Min.   : 0.000   Min.   : 0.000   Min.   :  0.00   Min.   :0.000
 1st Qu.: 0.000   1st Qu.: 0.000   1st Qu.: 20.00   1st Qu.:0.160
 Median : 0.000   Median : 2.000   Median : 47.00   Median :0.242
 Mean   : 1.554   Mean   : 4.127   Mean   : 61.43   Mean   :0.212
 3rd Qu.: 2.000   3rd Qu.: 7.000   3rd Qu.:101.00   3rd Qu.:0.276
 Max.   :17.000   Max.   :35.000   Max.   :163.00   Max.   :1.000
 NA's   :2662     NA's   :2661     NA's   :3414     NA's   :5618
     OBP              X1B             SLG           teamID.y     lgID.y
 Min.   :0.000   Min.   :  0.0   Min.   :0.000   CLE    :  935   AL:12304
 1st Qu.:0.208   1st Qu.:  0.0   1st Qu.:0.200   PIT    :  932   NL:13093
 Median :0.305   Median : 13.0   Median :0.351   PHI    :  931
 Mean   :0.270   Mean   : 32.5   Mean   :0.317   SDN    :  923
 3rd Qu.:0.346   3rd Qu.: 59.0   3rd Qu.:0.432   LAN    :  921
 Max.   :1.000   Max.   :225.0   Max.   :4.000   CIN    :  912
 NA's   :5562    NA's   :2661    NA's   :5618    (Other):19843
    salary
 Min.   :       0
 1st Qu.:  255000
 Median :  550000
 Mean   : 1879256
 3rd Qu.: 2150000
 Max.   :33000000
```

Analyzing the Lost Players

As previously mentioned, the Oakland A's lost 3 key players during the off-season. We will want to get their stats to see what we must replace. The players lost were: first baseman 2000 AL MVP Jason Giambi (giambja01) to the New York Yankees, outfielder Johnny Damon (damonjo01) to the Boston Red Sox and infielder Rainer Gustavo "Ray" Olmedo ('saenzol01').

#Using the subset() function to get a data frame called lost players from the combo data frame consisting of those 3 players.

```
lost_players <- subset(combo,playerID %in% c('giambja01','damonjo01','saenzol01') )
head(lost_players)
```

```
      playerID   H   X2B X3B     HR       OBP       SLG       BA  AB
5141  damonjo01 165  34   4   9    0.3235294 0.5093168 0.2562112   644
7878  giambja01 178  47   2  38    0.4769001 0.9942308 0.3423077   520
20114 saenzol01  67  21   1   9    0.2911765 0.5868852 0.2196721   305
```

#Using the subset again to only grab the rows where the yearID was 2001 for lost players.

```
lost_players <- subset(lost_players,yearID == 2001)
Reduce the lost_players data frame to the following columns: playerID,H,X2B,X3B,HR,OBP,SLG,BA,AB

lost_players <-lost_players[,c('playerID','H','X2B','X3B','HR','OBP','SLG','BA','AB')]
head(lost_players)
```

Replacement Players

Now we have all the information we need! Here is your final task - Find Replacement Players for the key three players we lost! However, you have three constraints:

1. The total combined salary of the three players cannot exceed 15 million dollars.
2. Their combined number of At Bats (AB) needs to be equal to or greater than the lost players.
3. Their mean OBP had to equal to or greater than the mean OBP of the lost players.

#filter year =2002 year after players left

```
avail.players <- filter(combo, yearID == '2002')
head(avail.players)
```

#filtering out columns needed for analysis

```
avail.players <-avail.players[c("playerID","yearID","AB","salary","OBP")]
```

#filtering out unnecessary values

```
avail.players <-filter(avail.players, OBP > 0 )
avail.players <-filter(avail.players,salary < 8000000)
avail.players <-filter(avail.players, AB > 200)
```

#Top three players I would choose as replacements with

```
replacements <-head(avail.players[order(avail.players$AB, decreasing = TRUE),],3)
replacements
```

```
    playerID  yearID  AB  salary      OBP
70  hattesc01   2002 492  900000 0.3738977
96    leeca01   2002 492 2700000 0.3593750
158 spiezsc01   2002 491 2275000 0.3714789
```