Untitled

Ishaan Advani

19 December 2017

This is my Analysis MarkDown file. I will address the questions I mentioned in the readme file.

I will first create a GroupObject

minimum\_ovr <- 70  
maximum\_ovr <- 80  
minimum\_age <- 20  
maximum\_age <- 30  
source("FIFAConfig.R")

##   
## 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

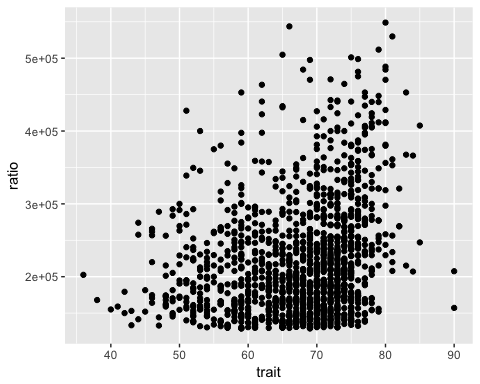
data <- load\_filtered\_dataset()  
GObj1 <- GroupObj(minimum\_ovr, maximum\_ovr, minimum\_age, maximum\_age, data)

Now that my Group Object has been created I will try and create an overvalued object that will help in identifying which players are valued at a premium.

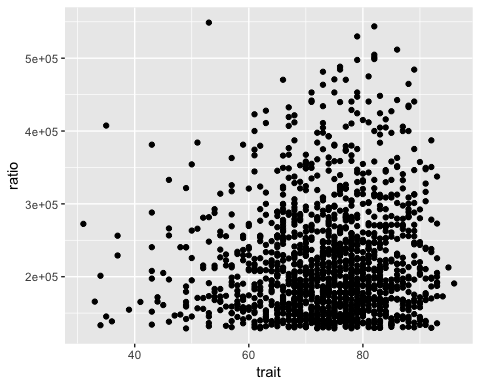
## [1] 4132

Now I will analyse 3 traits for the players valued at a premium and identify the relationship the traits have to the premium ratio. I will do this using regression and ggplot.

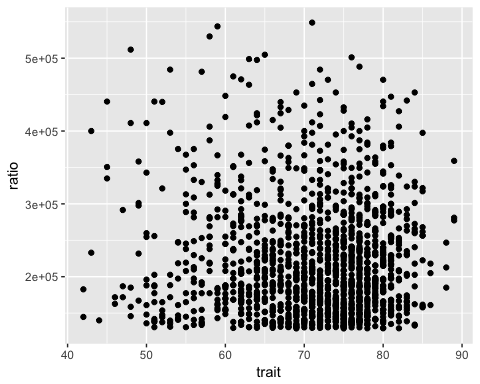
## [1] "Conducting a regression analysis for release to potential ratio and Passing"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 58619 2431   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Passing"



## [1] "Conducting a regression analysis for release to potential ratio and Pace"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 177006 618   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Pace"

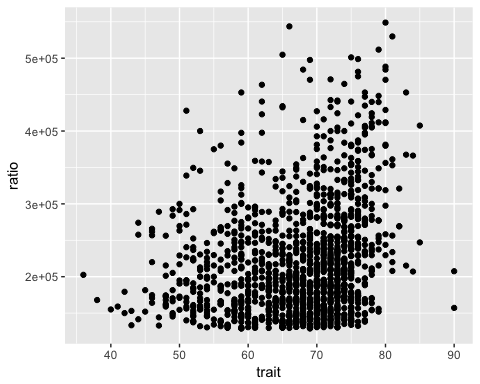


## [1] "Conducting a regression analysis for release to potential ratio and Physical"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 251151.4 -404.1   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Physical"

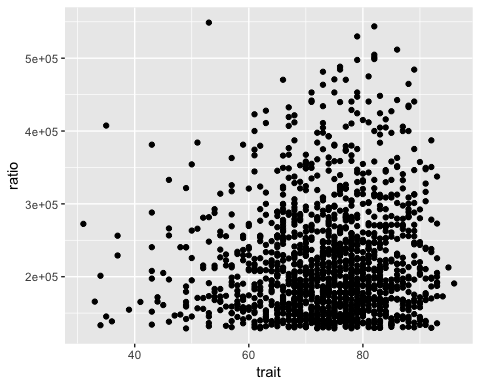


Now I will analyse 3 traits for the players valued at a premium and identify the relationship the traits have to the premium ratio. I will do this using regression and ggplot.

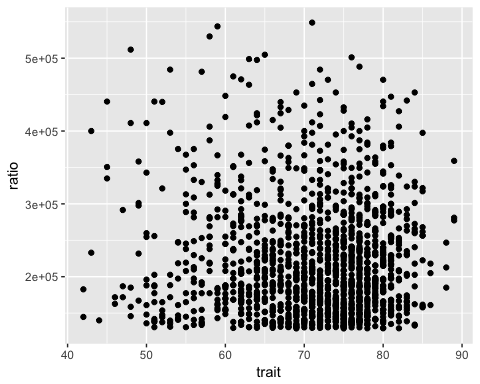
## [1] "Conducting a regression analysis for release to potential ratio and Passing"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 58619 2431   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Passing"



## [1] "Conducting a regression analysis for release to potential ratio and Pace"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 177006 618   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Pace"



## [1] "Conducting a regression analysis for release to potential ratio and Physical"  
##   
## Call:  
## lm(formula = ratio ~ t)  
##   
## Coefficients:  
## (Intercept) t   
## 251151.4 -404.1   
##   
## [1] "Creating a plot to show the relationship between overprice ratio and Physical"



I will now move on to analysis from the perspective of different countries/leagues. Each country has a typical playing style that requires more of some attributes and less of other. For example, players in Spain tend to have higher passing abilities than physical abilities in general. I will first create a country object based on a list of countries/leagues I extracted from the group object.

## [1] "The country I'm going to be analysing is Italian Serie A"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.3727948  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.4851828  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.3189259  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.3994921  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.4101669  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.6219327  
## [1] "The country I'm going to be analysing is French Ligue 1"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.4543739  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.6267577  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.2372175  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.2776999  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.3296345  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.6018398  
## [1] "The country I'm going to be analysing is Spanish Primera Divisi\303\263n"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.411425  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.6862161  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.2239987  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.2356987  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.2161193  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.5787765  
## [1] "The country I'm going to be analysing is German Bundesliga"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.4894731  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.6689654  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.2466104  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.324875  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.3468916  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.6418276  
## [1] "The country I'm going to be analysing is Portuguese Primeira Liga"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.3568979  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.5136723  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.1990076  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.2646671  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.1765049  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.3635901  
## [1] "The country I'm going to be analysing is Ukrainian Premier League"  
## [1] "The correlation between the wage paid and a players Passing is:"

## Warning in cor(country\_salary, cur\_skill): the standard deviation is zero

## [1] NA  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.7481882  
## [1] "The correlation between the wage paid and a players Pace is:"

## Warning in cor(country\_salary, cur\_skill): the standard deviation is zero

## [1] NA  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.5315494  
## [1] "The correlation between the wage paid and a players Physical is:"

## Warning in cor(country\_salary, cur\_skill): the standard deviation is zero

## [1] NA  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.3234133  
## [1] "The country I'm going to be analysing is English Premier League"  
## [1] "The correlation between the wage paid and a players Passing is:"  
## [1] 0.6098966  
## [1] "The correlation between the overall and Passing is:"  
## [1] 0.7215576  
## [1] "The correlation between the wage paid and a players Pace is:"  
## [1] 0.2579243  
## [1] "The correlation between the overall and Pace is:"  
## [1] 0.2741436  
## [1] "The correlation between the wage paid and a players Physical is:"  
## [1] 0.5031323  
## [1] "The correlation between the overall and Physical is:"  
## [1] 0.7172529

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.