**PER Calculation**

1.Introduction

2.Formula

3.Practice

1.Introduction

The Player Efficiency Rating (PER) is a per-minute rating developed by ESPN.com columnist John Hollinger. In John's words, "The PER sums up all a player's positive accomplishments, subtracts the negative accomplishments, and returns a per-minute rating of a player's performance." It appears from his books that John's database only goes back to the 1988-89 season. I decided to expand on John's work and calculate PER for all players since minutes played were first recorded (1951-52).

Some pros of PER

1)Can give you a straightforward idea of how good a player is

2)Useful in comparing seasons

3)A universal recognized valuable metric to evaluate a player using a “Single number”

Some cons of PER:

1)Doesn’t give enough credits to a Player’s defensive value

2)Overrate the Rebounds a little bit.

3)Doesn’t value FT enough

4)Overvalue the “volume” players

5)Undervalues the MPG

2.Formula

All calculations begin with what we call an unadjusted PER (uPER). The formula is:

uPER = (1 / MP) \*

[ 3P

+ (2/3) \* AST

+ (2 - factor \* (team\_AST / team\_FG)) \* FG

+ (FT \*0.5 \* (1 + (1 - (team\_AST / team\_FG)) + (2/3) \* (team\_AST / team\_FG)))

- VOP \* TOV

- VOP \* DRB% \* (FGA - FG)

- VOP \* 0.44 \* (0.44 + (0.56 \* DRB%)) \* (FTA - FT)

+ VOP \* (1 - DRB%) \* (TRB - ORB)

+ VOP \* DRB% \* ORB

+ VOP \* STL

+ VOP \* DRB% \* BLK

- PF \* ((lg\_FT / lg\_PF) - 0.44 \* (lg\_FTA / lg\_PF) \* VOP) ]

Most of the terms in the formula above should be clear, some of the factors are defined by Mr. Hollinger:

factor = (2 / 3) - (0.5 \* (lg\_AST / lg\_FG)) / (2 \* (lg\_FG / lg\_FT))

VOP = lg\_PTS / (lg\_FGA - lg\_ORB + lg\_TOV + 0.44 \* lg\_FTA)

DRB% = (lg\_TRB - lg\_ORB) / lg\_TRB

**Note here that “lg” is not log10 but means “league”**

After uPER is calculated, an adjustment must be made for the team's [pace](https://www.basketball-reference.com/about/glossary.html#pace). The pace adjustment is:

pace adjustment = lg\_Pace / team\_Pace

**Pace**  
Pace Factor (available since the 1973-74 season in the NBA); the formula is 48 \* ((Tm Poss + Opp Poss) / (2 \* (Tm MP / 5))). Pace factor is an estimate of the number of possessions per 48 minutes by a team. (Note: 40 minutes is used in the calculation for the College Basketball.)

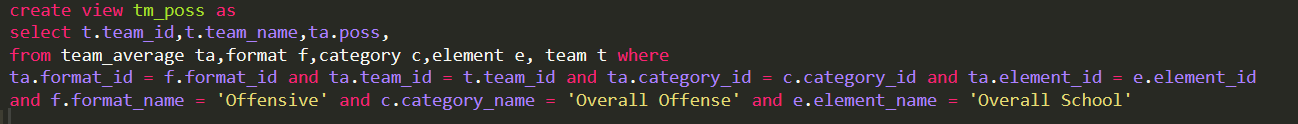
3.Practice

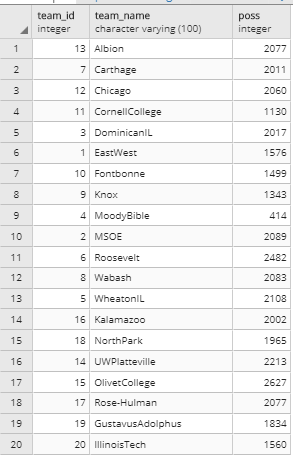
In this section I will show you how did I implement each step in the formula using Postgresql Database ”Basketball” I created and Python module “Pandas”.

One thing to note is that since the formula includes the “league” information and IIT doesn’t have a formal kind of league, so I just created a so-called league which includes the teams that IIT played against in the last season.

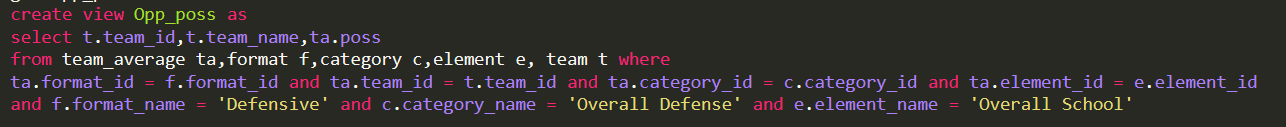
***3.1Get Pace factors***

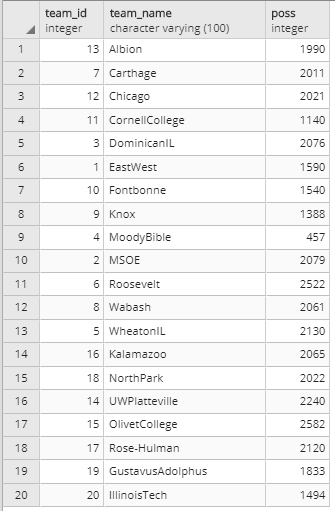
1)Get Tm\_Poss



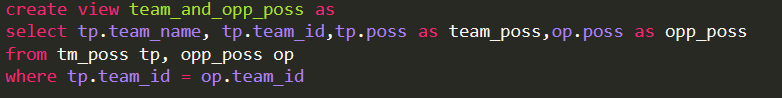


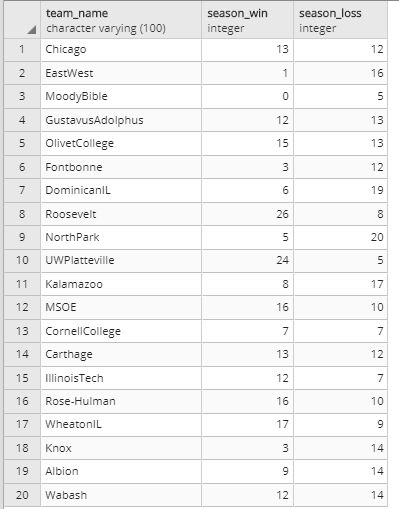
2)Get Opp\_Poss



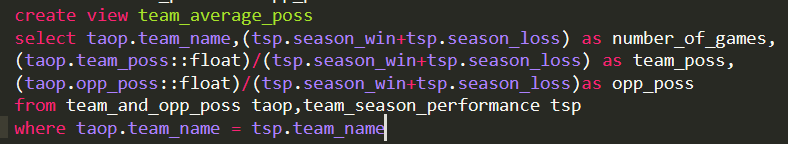


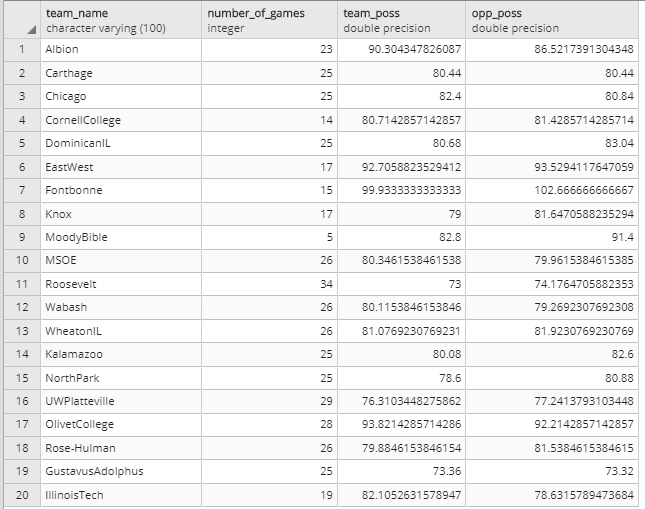
3) Combine 1) and 2) get team and opp poss





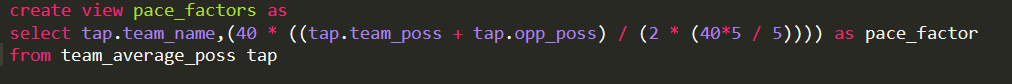
4) Calculate average team\_poss and average opp\_poss





5)calculate Pace factors:

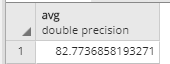
The original formula is 48 \* ((Tm\_Poss + Opp\_Poss) / (2 \* (Tm\_MP / 5))).Since PER was invented based on NBA stats, but now we are analyzing the stats of College basketball, so we replace “48” in the formula with 40 and Tm\_MP = 5\*40 (instead of 5\*48)





6)get league average Pace\_Factor: lg\_Pace





***3.2 Get player\_stats and team\_stats used in the formula.***

According to the formula of the ‘uPER’,

the stats needed for players are :

PTS,3P,AST,turnover,FG,FT,FGA,STL,PF,FTA,TRB,ORB,DRB,BLK.

The stats needed for teams are:

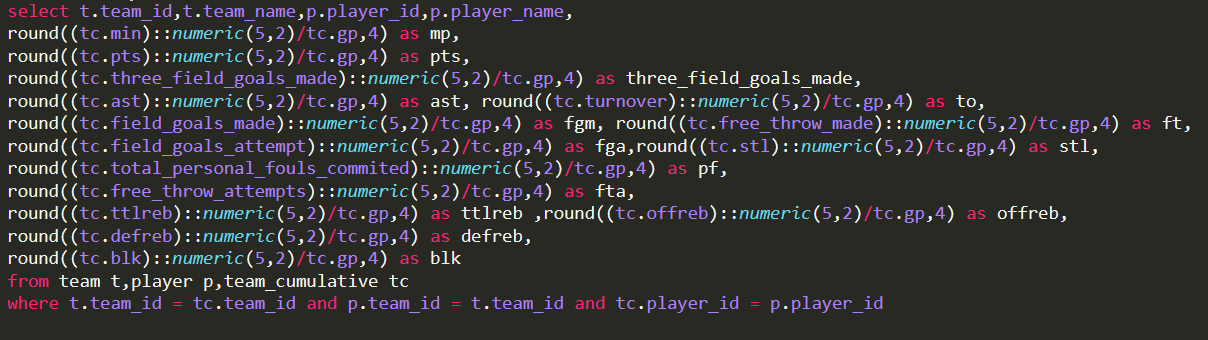
team\_AST,team\_FG

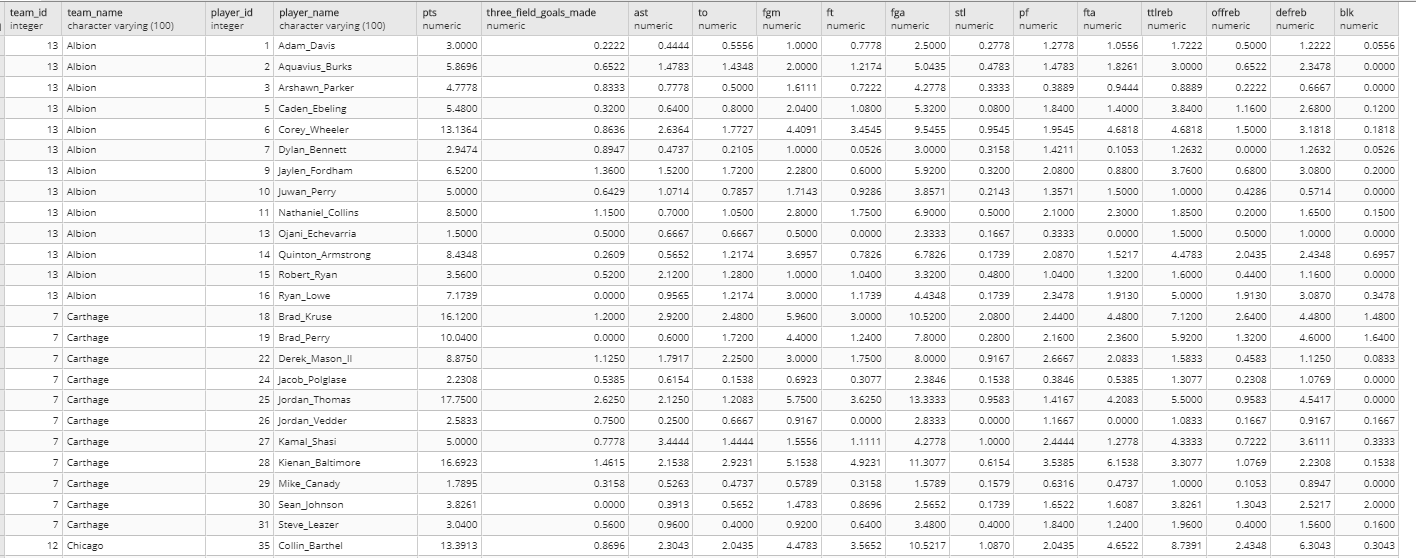
Since in our database”basketball”, we don’t have “average stats” of these metrics, we need to calculate those using queries.

One thing to notice is that in our table ”team\_cumulative”, we can calculate those metrics needed by dividing the existing metrics by the “game\_played” column.

Another thing we need to notice is that since initially we created the fields types as “integers” for the most of our columns, we need to change it to float when we do averages.

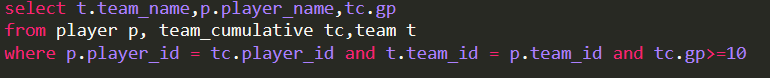
Here is the view I created for player\_raw\_stats needed for calculating PER:





***3.3 Set a filter for the players included***

I calculated the PER once and found some “end of bench” players can get an unexpected high PER since they only played 1 or 2 games and happened to player well in their limited minutes. So I think it is necessary to set a “filter” to rule those players out. I set the minimum number of game a player played as 10.Out of 321 players, I finally got 235 players.

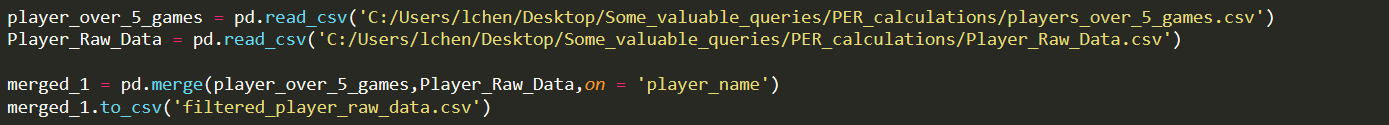


***3.4 Merge “Filtered\_Player\_Raw\_Stats” with”Player\_Raw\_Stats”and “Team\_Pace\_Factors”***

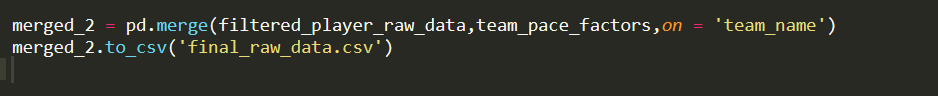
Taking the results of 3.1, 3.2 and 3.3, now we need to “merge” those 3 so that it will facilitate us to calculate the PERs

I used pandas module in Python to implement this step

First, merge “Player\_Raw\_Data” and “Player\_Over\_5\_Games”

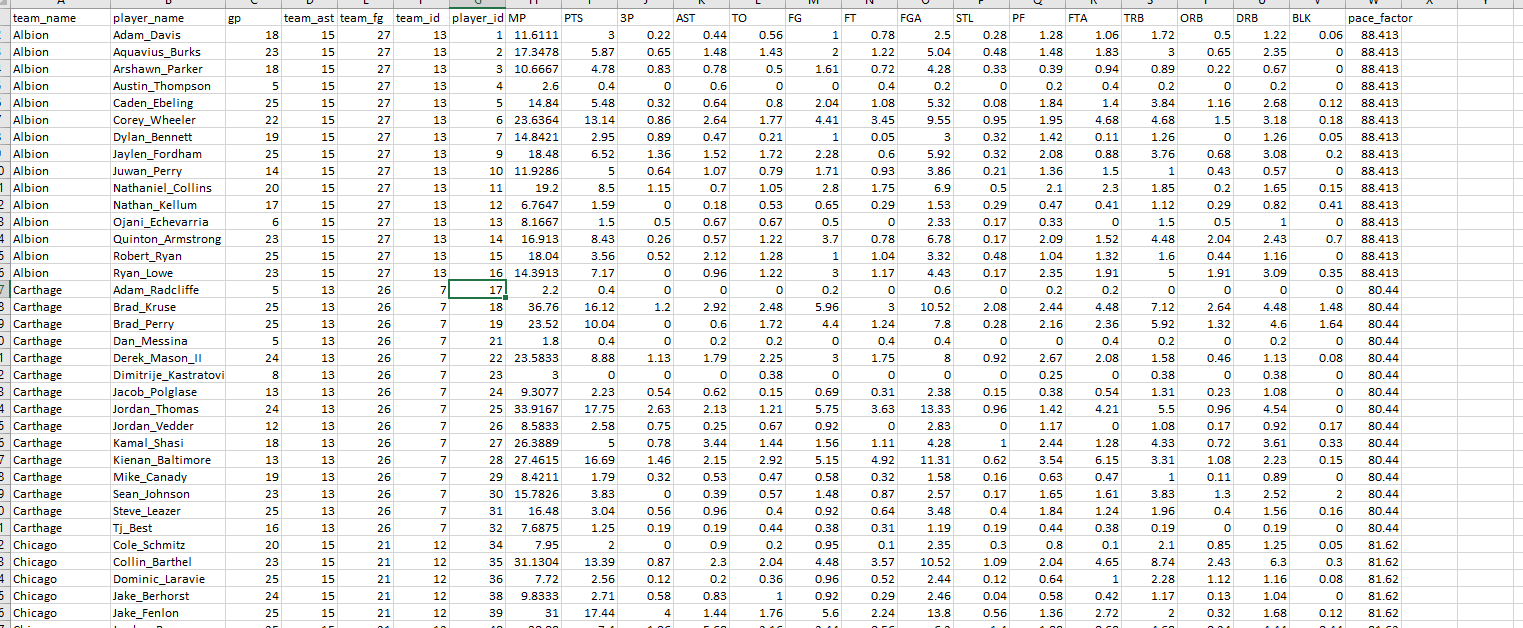


Then, merge”filtered\_player\_raw\_data” and “team\_pace\_factors”



Now, we are set to move to next step: Calculating PER

“final raw data”

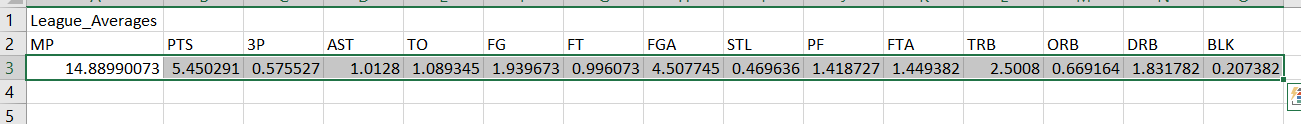


***3.5 Calculating PER***

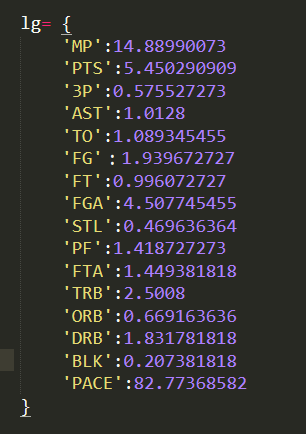
Now we have all the stats in a table called “final\_raw\_data”, now we can do our long-anticipated step: calculating the PER

In this session, we continued to use the “Pandas” as tool to calculate the PER for each player.

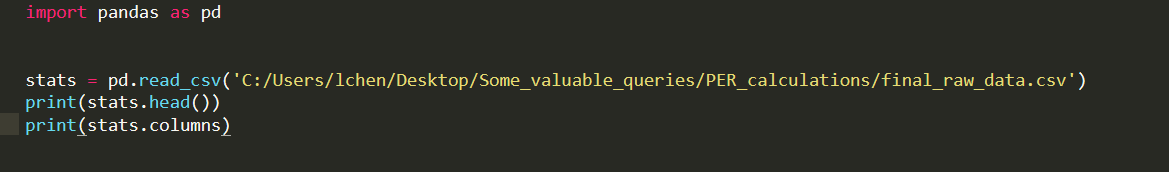
First, in EXCEL sheet, we can get the average stats of the so-called “League”(includes 275 players).

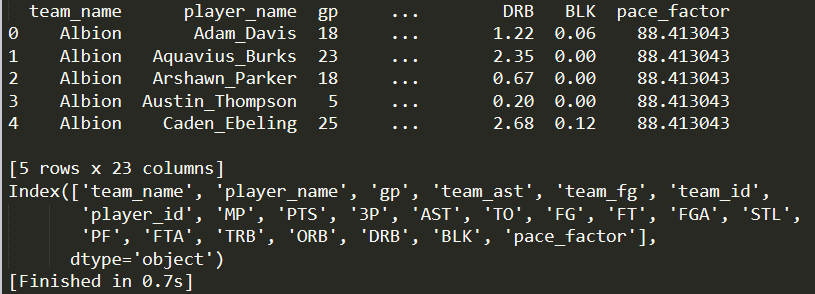


Based on this, we create a dictionary called “lg” to store those average metrics.

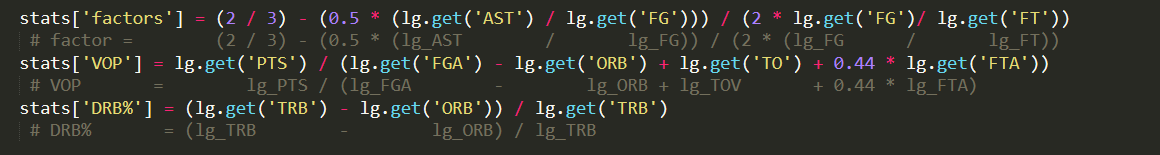


The we import the “final\_raw\_data” :

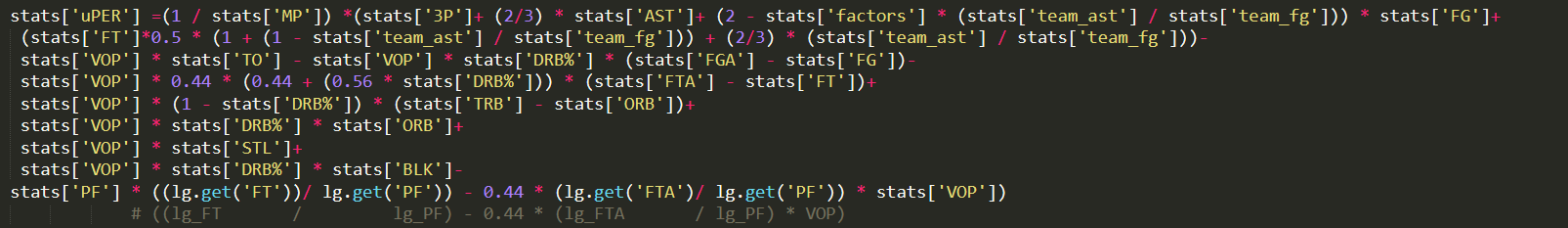




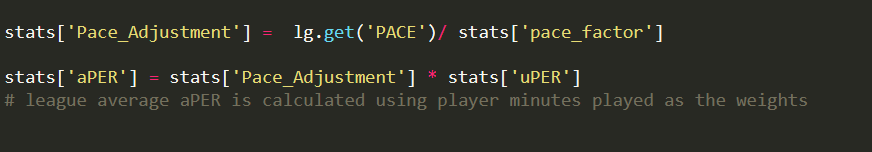
Corresponding codes for 3 factors author introduced



Unadjusted PER:

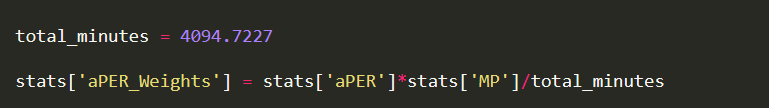


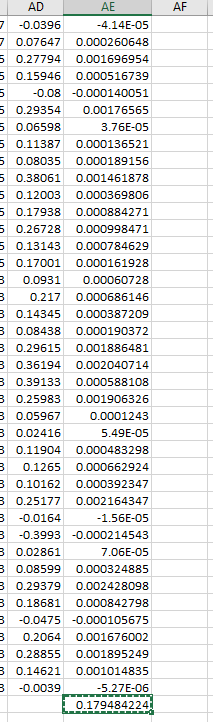
Pace adjustment:



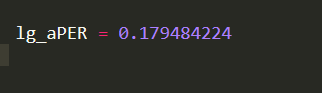
The final step is to calculate the lg\_aPER which use player minutes played as the weights.

To do that I use this equation to get each players “contribution” to the lg\_aPER and then sum them up in EXCEL sheet.

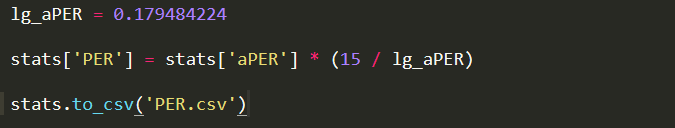




So we get lg\_aPER is



Finally, we get what we want!



Now we can manipulate the results in Tableau!

