

# **2021-22 NBA Season EDA**

# Yearbook Staff: Full House



**Meeyoung  
Park**



**David  
Mostacero**



**Peter  
Warren**

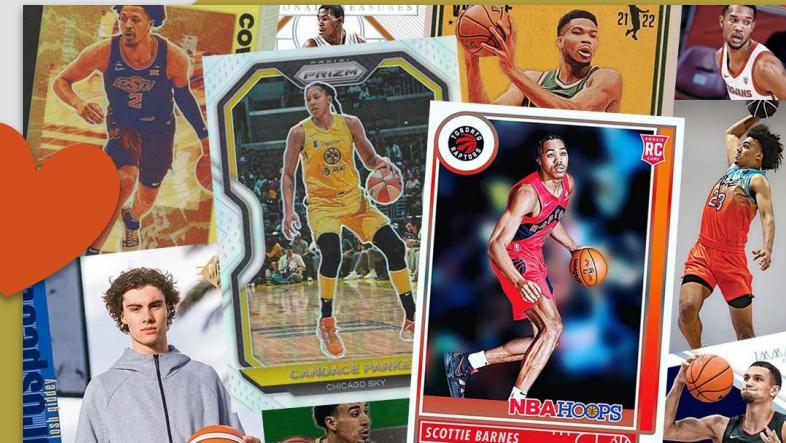


**Sun Choi**



**Alison  
Faulkner**

# Purpose: NBA 2021-22 EDA

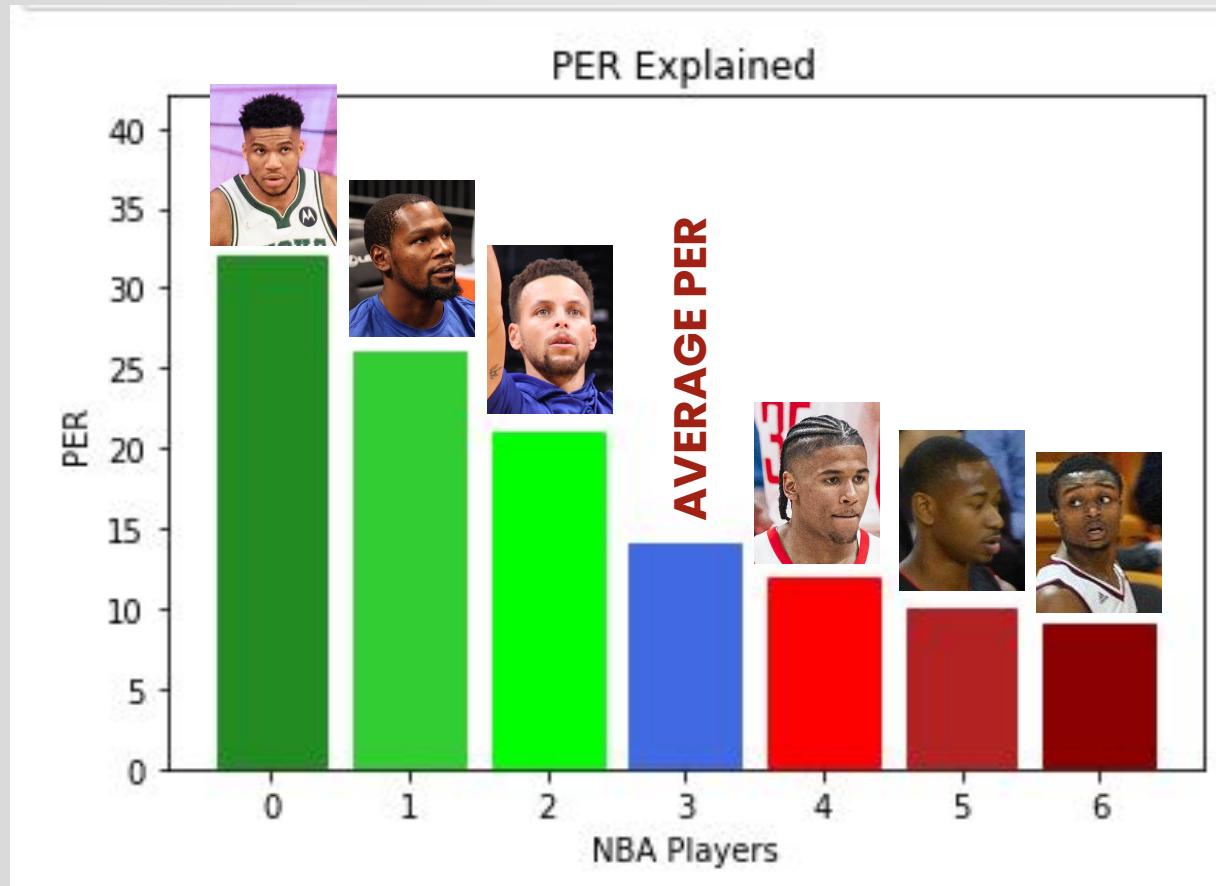


# PER, Explained

## Player Efficiency Rating

Metric used by NBA to distill a player's contributions into a single numerical statistic.

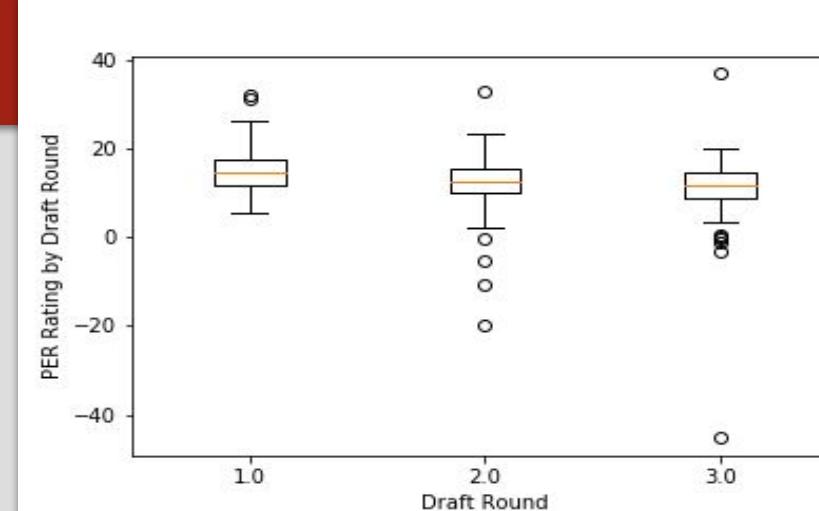
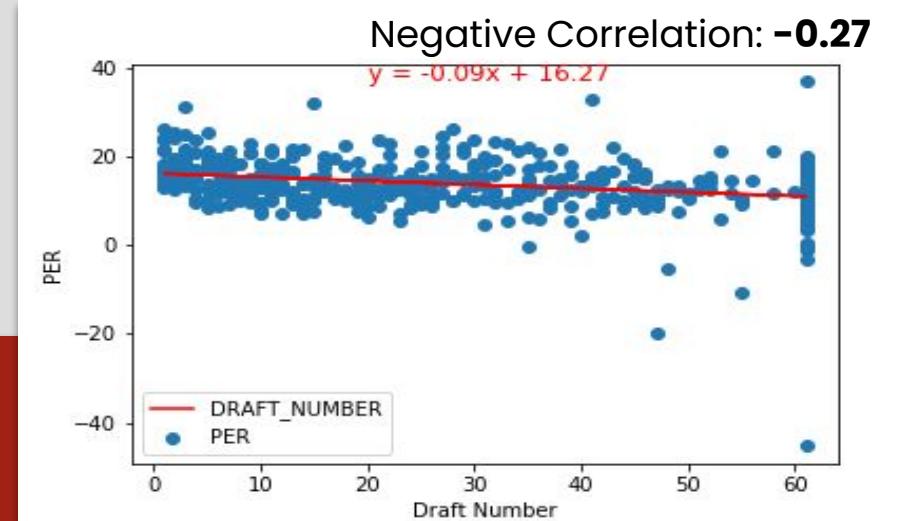
Average PER = 15



# Draft Picks and PER

DRAFT_NUMBER	Players in Draft Number	AVG PER in Draft Number
1.0	12	19.050000
3.0	13	18.915385
2.0	9	18.233333
27.0	8	17.675000
22.0	6	16.766667
43.0	3	16.600000
58.0	2	16.500000
5.0	11	16.309091
16.0	6	16.300000
41.0	6	16.250000
30.0	8	16.175000
29.0	6	16.016667

DRAFT_ROUND	Players in Draft Round	AVG PER in Draft Round
1.0	257	15.049027
2.0	119	12.601681
3.0	115	11.201739



# 2021-22 Superlatives based on Draft Round and PER

```
#figuring out who is the veteran (drafted earlier than 2020) MVP
vet_sup = cor_player_draft_pick.loc[(cor_player_draft_pick['DRAFT_YEAR']<2021)
                                     & (cor_player_draft_pick['G'] >= cor_player_draft_pick['G'].mean())]
vet_sup = vet_sup[vet_sup['PER']==vet_sup['PER'].max()][['Player Name']]
print(f"the veteran superlative goes to {vet_sup}")

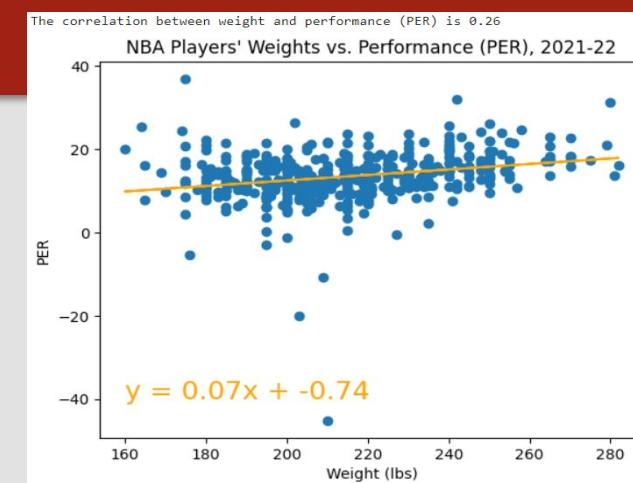
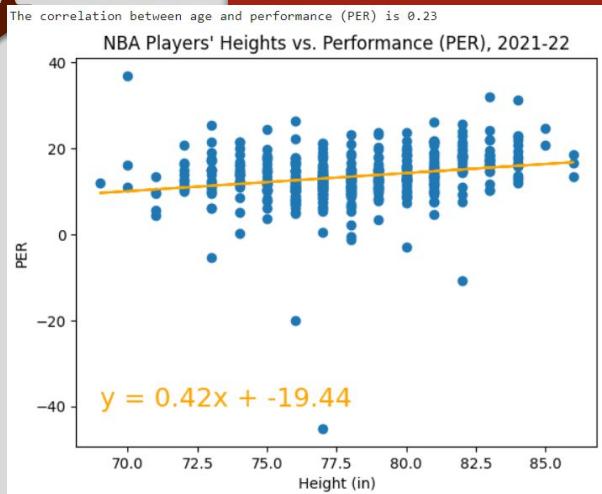
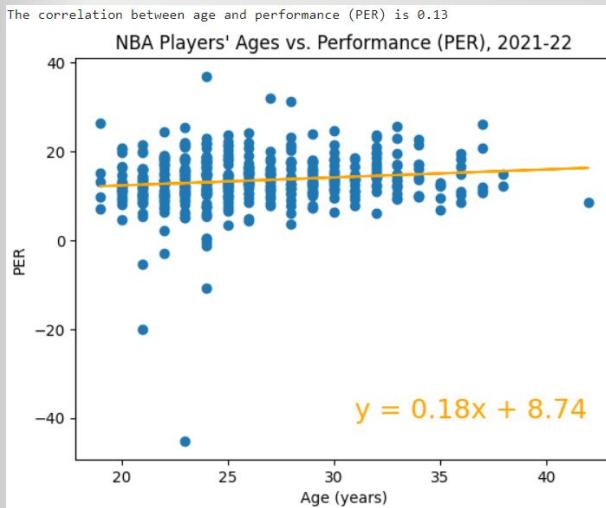
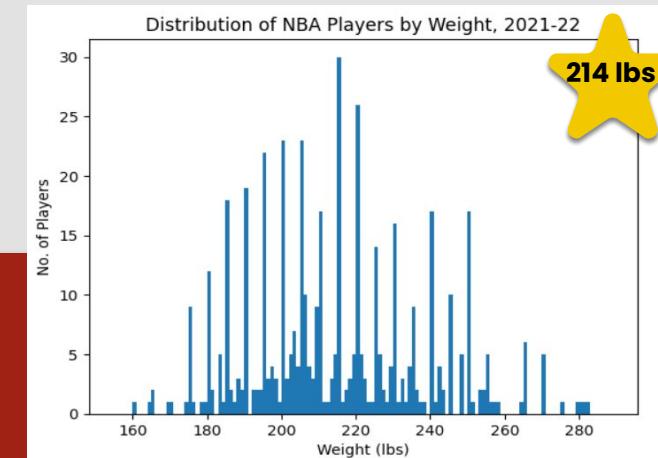
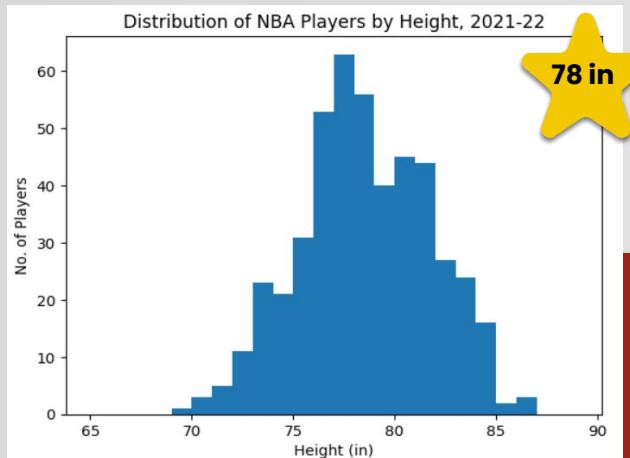
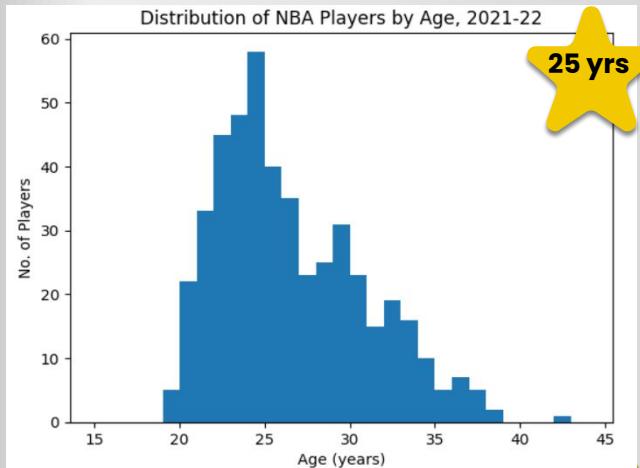
#figuring out who is the rookie (drafted in 2021) MVP
rook_sup = cor_player_draft_pick.loc[(cor_player_draft_pick['DRAFT_YEAR']==2021)
                                      & (cor_player_draft_pick['G'] >= cor_player_draft_pick['G'].mean())]
rook_sup = rook_sup[rook_sup['PER']==rook_sup['PER'].max()][['Player Name']]
print(f"the rookie superlative goes to {rook_sup}")

#figuring out who the undrafted MVP is
undraft_sup = cor_player_draft_pick.loc[(cor_player_draft_pick['DRAFT_YEAR']==3000)
                                         & (cor_player_draft_pick['G'] >= cor_player_draft_pick['G'].mean())]
undraft_sup = undraft_sup[undraft_sup['PER']==undraft_sup['PER'].max()][['Player Name']]
print(f"the undrafted superlative goes to {undraft_sup}")

veteran superlative goes to 112    Nikola Jokic
: Player Name, dtype: object
rookie superlative goes to 256    Scottie Barnes
: Player Name, dtype: object
undrafted superlative goes to 152   Christian Wood
: Player Name, dtype: object
```



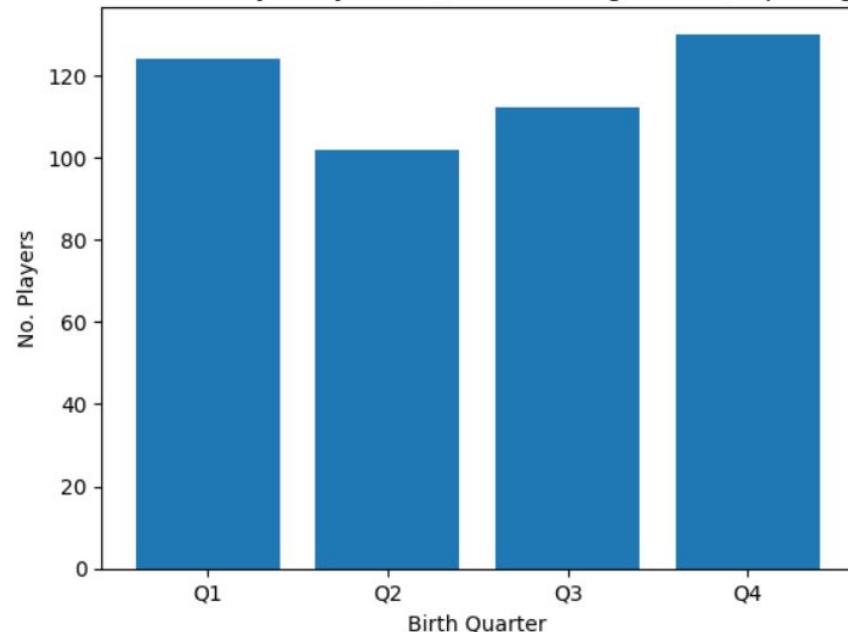
# Biometrics



# Birth Month by Quarter



Distribution of NBA Players by Birth Quarter of League Year (Sep - Aug), 2021-22



## NBA League Year: August – September

- Q1 (Sep – Nov)
- Q2 (Dec – Feb)
- Q3 (Mar – May)
- Q4 (Jun – Aug)



$H_1$ : Due to cumulative advantages of being the oldest players in respective youth teams, players born in the 1st quarter of each league year would have higher representation in the NBA.

$H_0$ : There was no advantage in representation between players born in the first and the other quarters of the NBA league year.

# Analysis by Origin

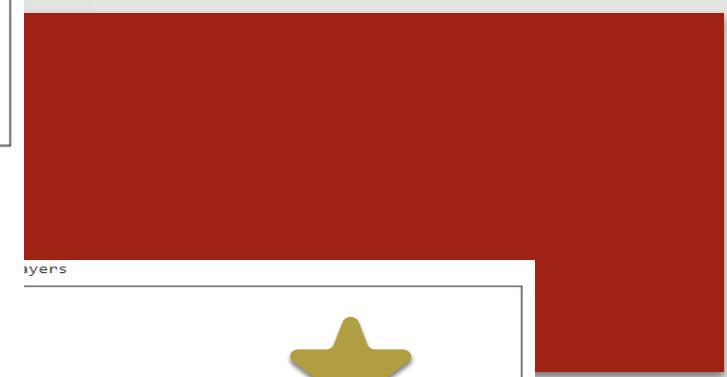
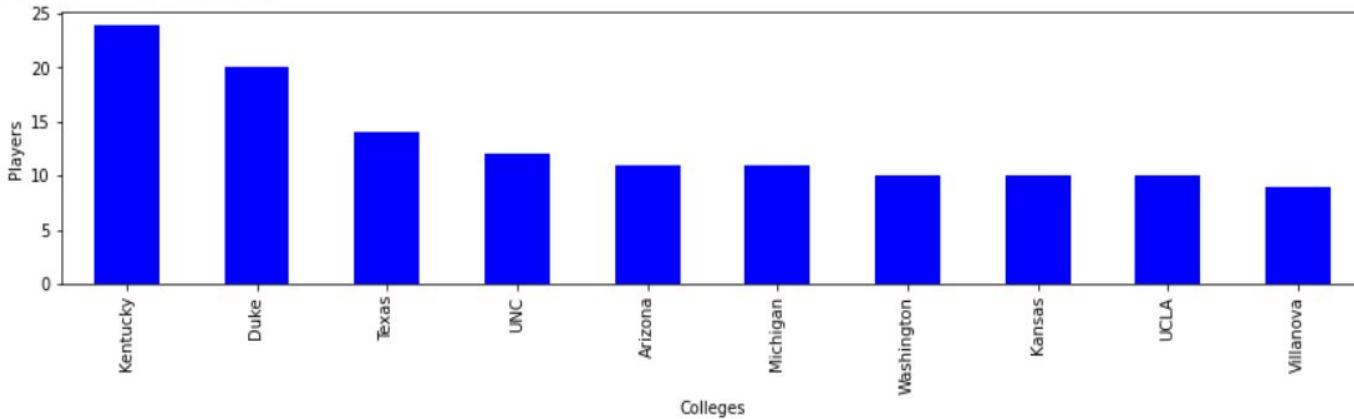


# College

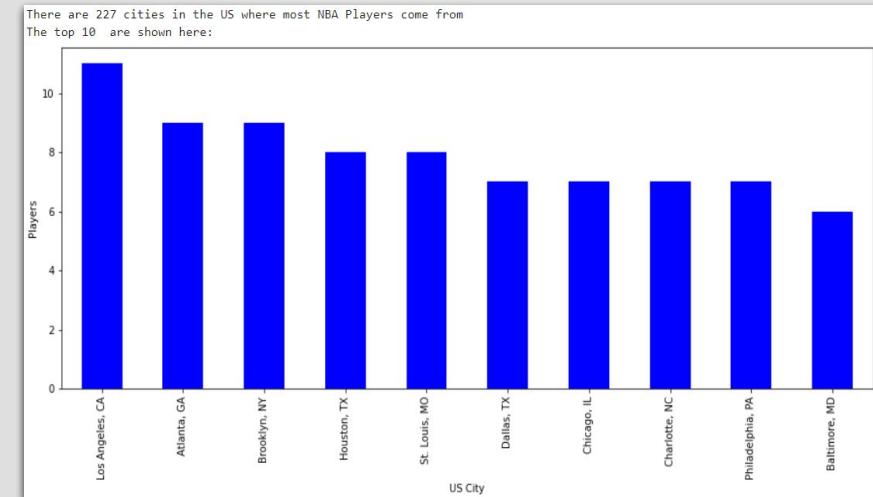
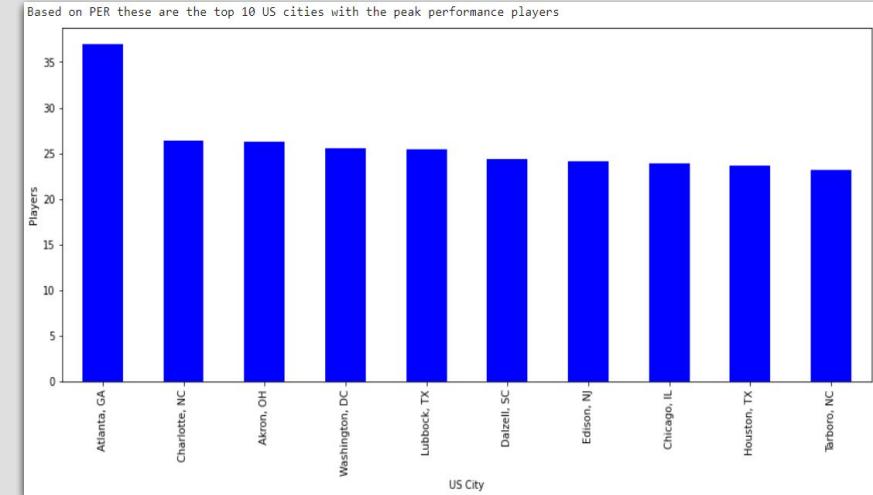
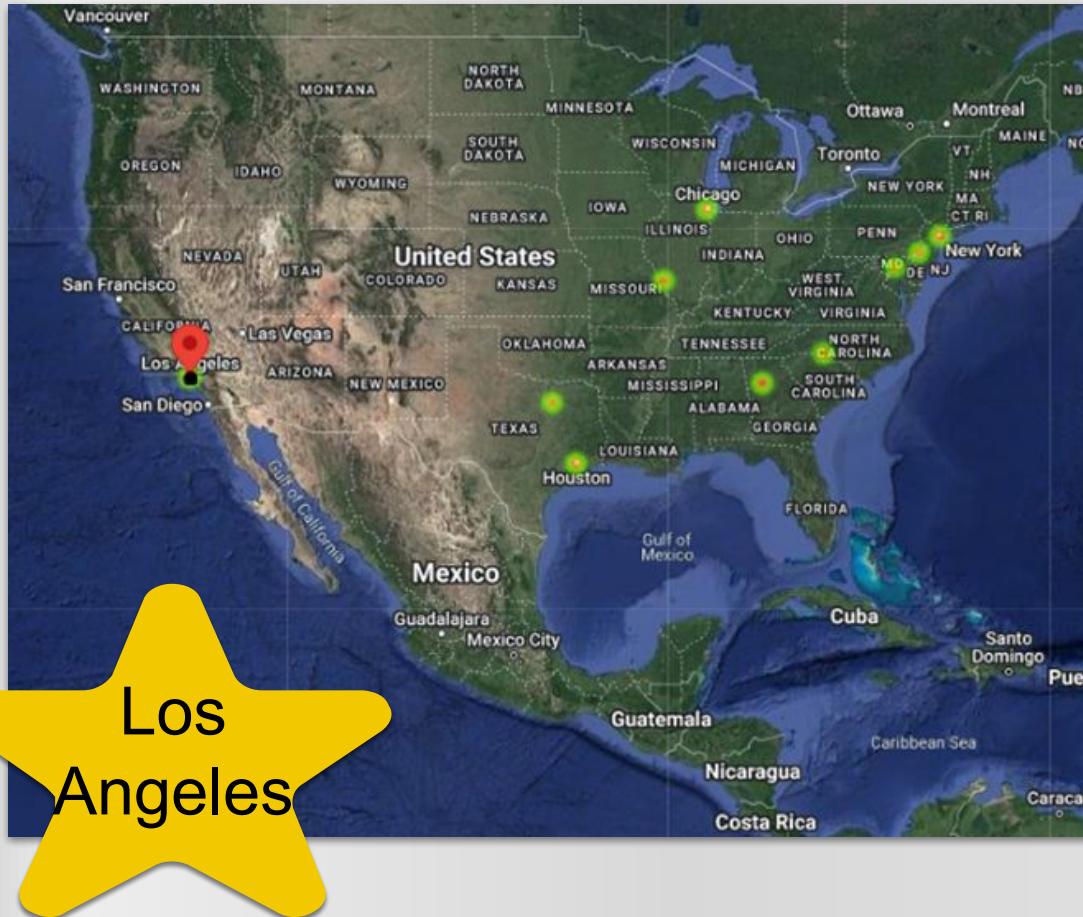


It's  
finger lickin'  
good

There are 173 College/University where most NBA Players come from.  
The top 10 are shown here:



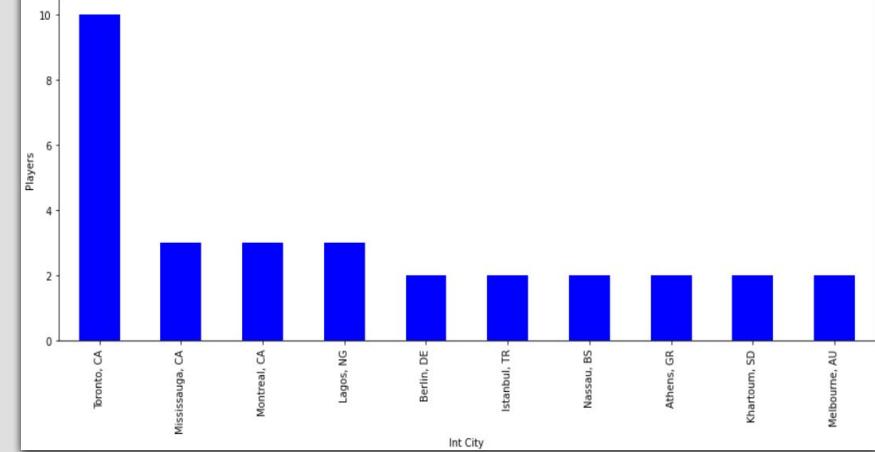
# US City



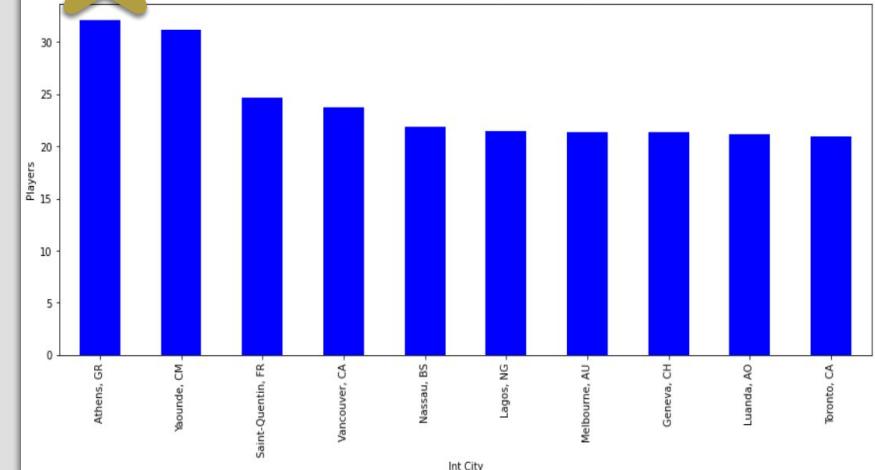
# Int. City



There are 69 cities where NBA player come from around the world  
The top 10 are shown here:



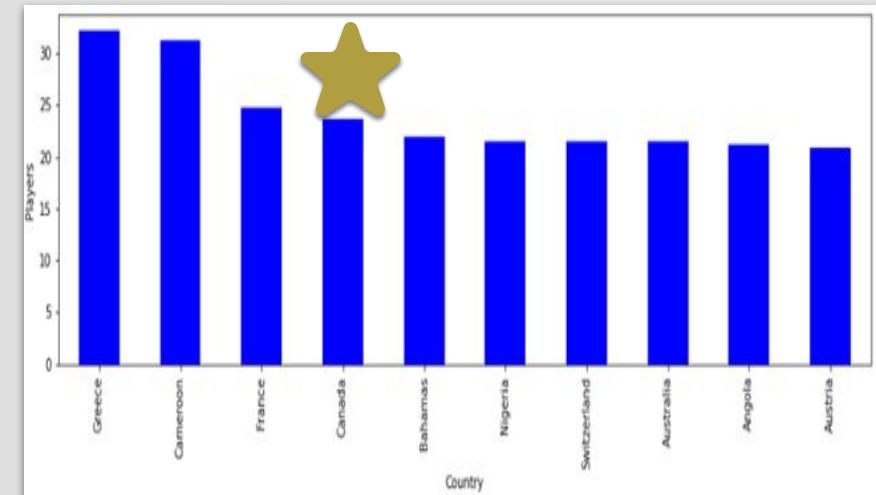
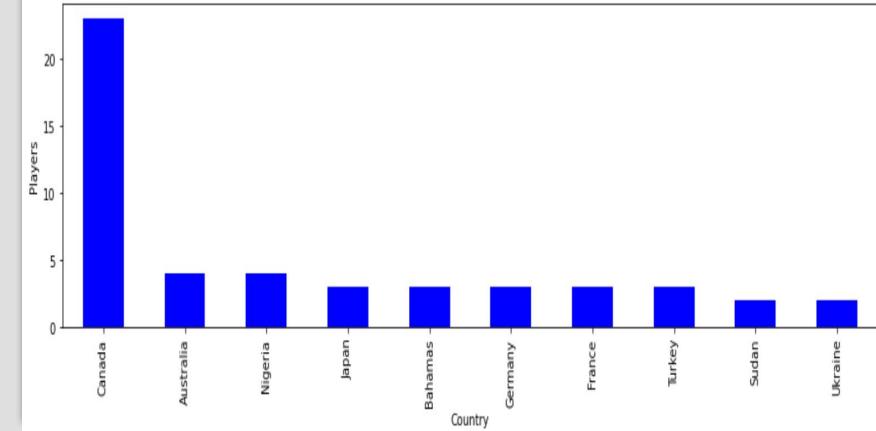
Based on the chart above are the top 10 International cities with the peak performance players



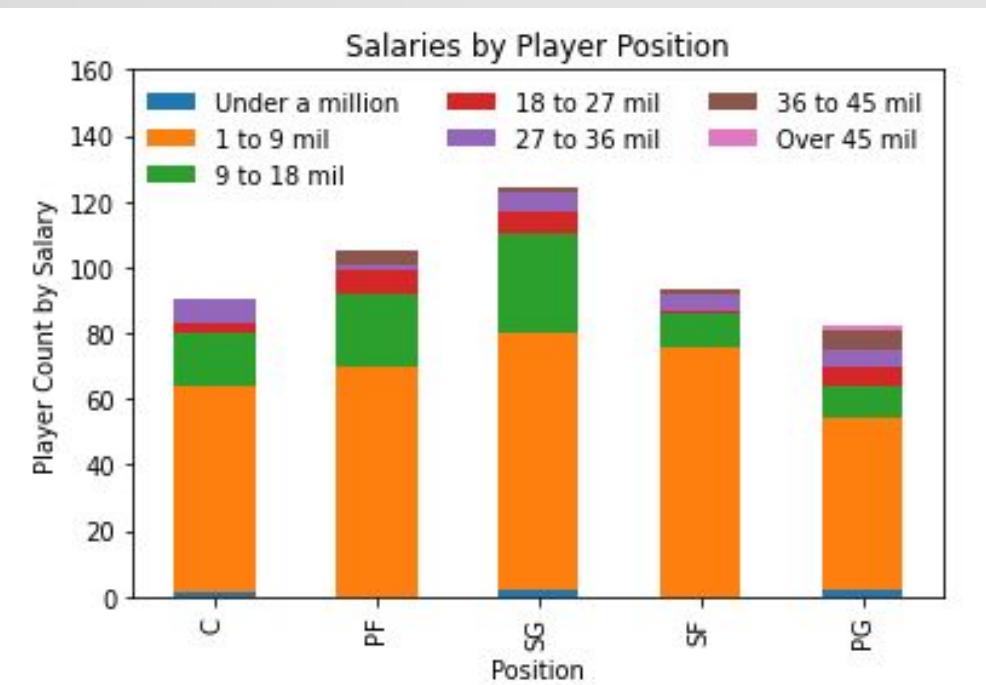
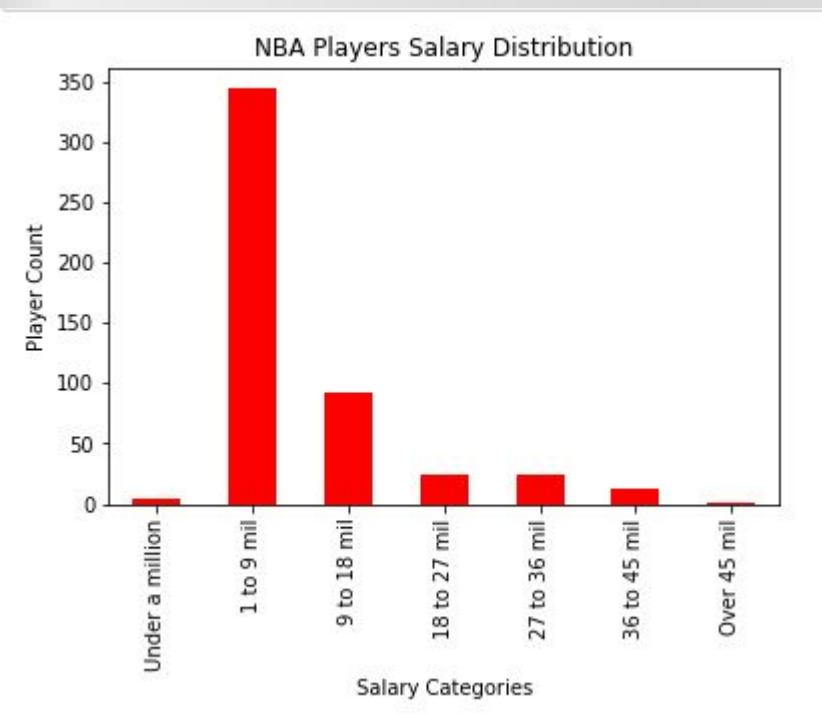
# Country



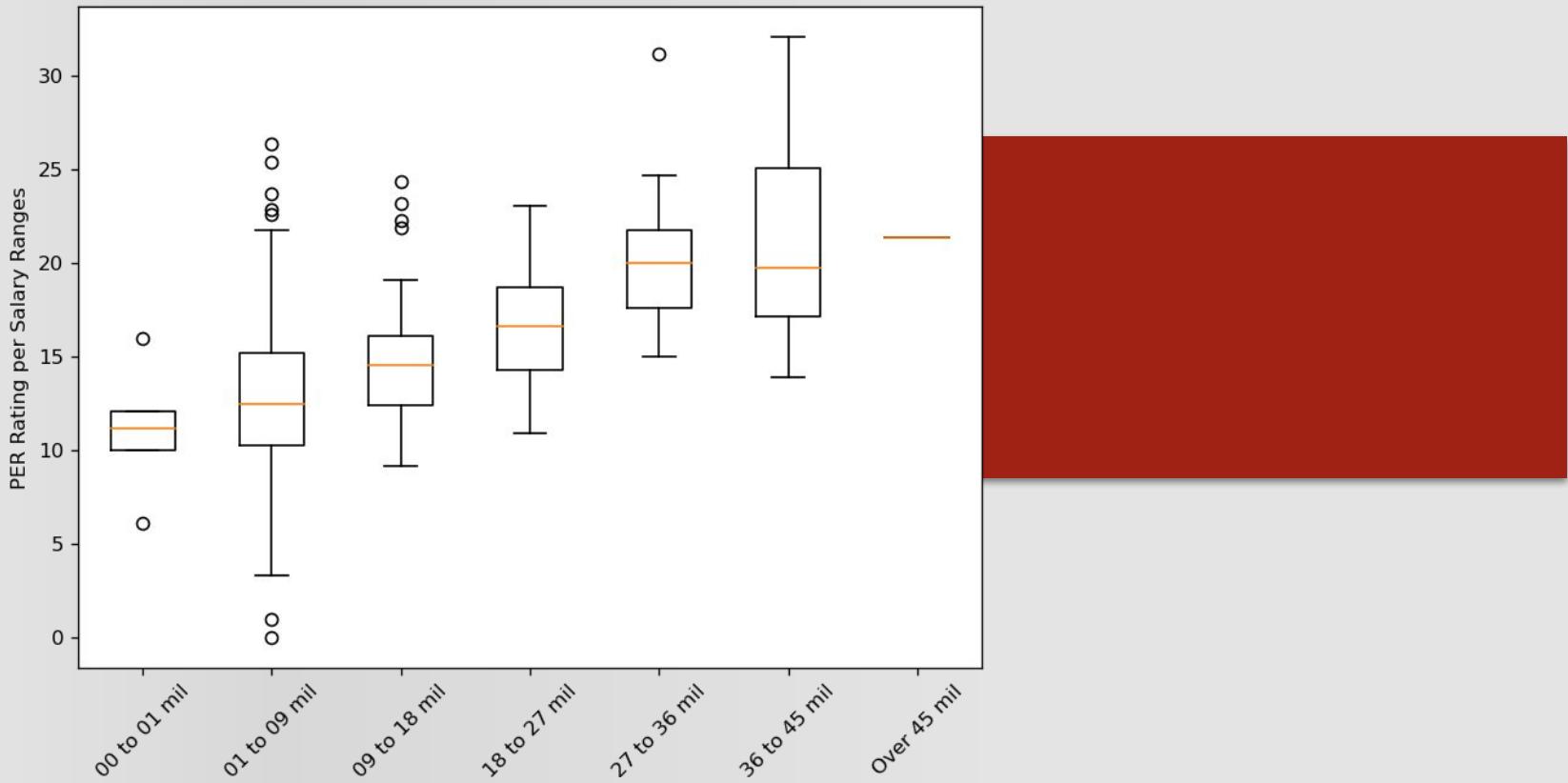
There are 39 countries where most players in the NBA that were not born in the US come from  
The top 10 are shown here:



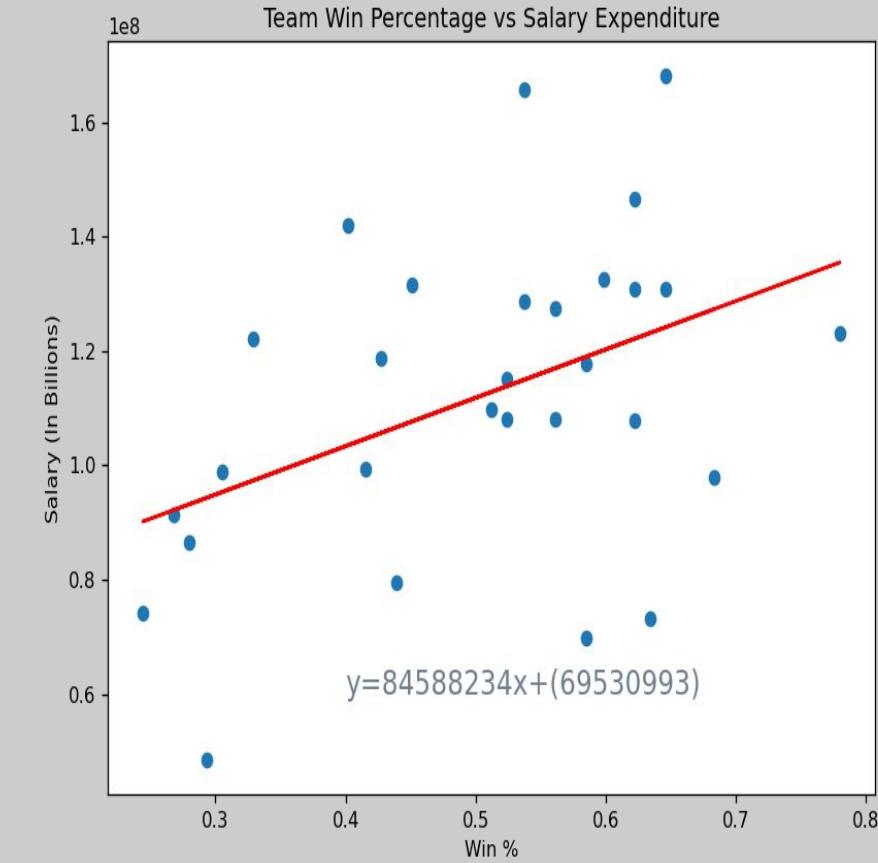
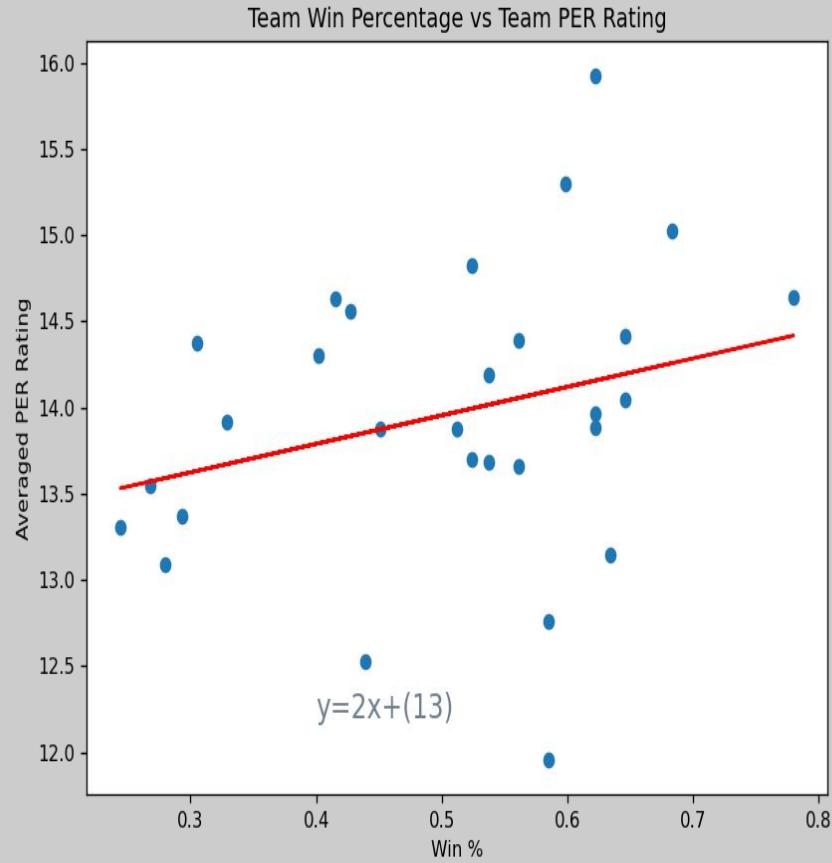
# Salary Distributions



# PER vs. Salary



# Does PER Work?



# Closing Statements

- Exploration of 2021–22 NBA Active Player Data
  - Biometric
  - Draft Pick
  - Origin
  - Salary
- Limitations
- Findings

**Stephen Curry**