MIS 6309.502 Business Data Warehousing

Student Name: Sanyogeeta Patil

Instructor: Kevin R. Crook

Spring 2017

Data Visualization Project

NBA Analysis:

The data consist of NBA statistics from 1935 to 2011. It gives us information such as how many matches played, team statistics, coaches information as well as demographic information of teams. It is observed that NBA drafts players from all over the world and various universities as well. It is also observed that sometimes the best players comes from average universities and not from the top universities.

Data also gives us the information of salaries of the team from which we can see the highest paid team by NBA as well as it also gives the debts information of teams. It is observed that even of the team has good salary and still the team is in debt.

As the data contains information of various teams we can find which team performs best and which performs average and decide their salaries accordingly. The data also contains information about the where the match was played i.e. Home ground or Away from home. Sometimes win also depends on that.

Data also gives the information about the category of players such as best players, good players, average players, below average players. From this it can be seen that best players are not good in all aspects of game. So from this we can also find various statistics of categories of players.

The source of the data set was: http://opensourcesports.com/files/basketball/BasketballDB-20130121.zip.

- abbrev file gives us all the information about the abbreviations used in data set. Award coaches gives us information about coach associated with year.
- Draft files gives us information about which player was drafted from which university and in which year. It also tells us about draft rounds.
- Coaches files gives detail view about coaches, coach Id associated with year, team, win, loss and about previous wins or losses.
- Award Player gives us the information about the awards associated with players.
- Master files contains all details of the players in details including DOB, birth city, height, weight etc.
- HOF file i.e. Hall of Fame file gives information of legendary players as well coaches.
- Players files gives the team stats of a player such as how many assist he did, how many block, how many points he got etc.
- Teams file gives information about matches played, winning team, team rank, duration and location of the match.

Layout of the dataset: -

Variable	Description
Measure Names	Categories of players such as avg points, avg rebounds, avg assists, avg steals, avg blocks and avg three made.
Year	Contains range from 1935-2011
Player Name	Contains all the information about the player such as full name, id associated with player etc
College	Contains information about the university drafts made
Total Players	COUNTD([Bio ID]) - To get distinct count of no of players
Cluster Segmentation	IF [Cluster] = 3 then "Best Players" ELSEIF [Cluster] = 4 THEN "Good Players" ELSEIF [Cluster] = 2 Then "Average Players" ELSE "Below Average Players" END
Player Name	[First Name] +" "+ [Last Name] - To get Full name of player
3 Points Percentage	([O 3Pm]/[O 3Pa])*100 - To get percentage of 3 pointers scored

Defensive Points Saved	([D Pts]/[Games])*100 - To get percentage of defensive points saved
Field Goals Made Percentage	([O Fgm]/[O Fga])*100 - To get percentage of fields goals
Free Throws Percentage	([O Ftm]/[O Fta])*100 - To get percentage of fields throws
Points Per Game	[O Pts]/[Games] - To calculate points per game
Win Percentage	([Won]/[Games])*100 - To calculate win percentage

Data cleansing

As all the data was present in understandable format as views in spreadsheet no data cleansing was performed on data.

Analytical question at a high level that can be answered: -

Identify the top 5 team performances:

We can use small multiples to show the win and loss trends for the top 5 teams in the NBA between 1977 and 2008. The small multiple technique since it gave us an option to standardize the variables and allowed us to compare the trend line across four different parameters on a common axis.

Insight: -

"From the trend lines it can be seen that Los Angeles Lakers have been the most consistent team in terms of highest home and away wins and the least home and away losses."

How many teams have a One in Three chance of converting a three pointer in a game:

We can use text tables here to display the various team performance metrics since it offered transparency and depth of information to compare the teams, based on win rate since we wanted to display team efficiency. By adding multiple calculated columns namely: Win rate, points per game, field goals(in %), 3 points made (in %) and Free Throws Made. The table was sorted on the basis on Win Rate and the top 15 teams in the NBA are displayed.

Insight: -

"Top 15 Teams in the NBA have a One in Three chance of converting a three pointer in a game"

"San Antonio Spurs are almost twice as likely at winning games as compared to New York Knicks despite having lower average points per game"

Which type of players are best in all aspects:

We can use stack bar charts to visualize players to show comparisons across categories (Clusters Analysis).We did color encoding on player measures such as Assists Rebounds, Points, Steals, Blocks, Three Pointers made.

Classification of players can be done in 4 categories: 1) Best Players 2) Good Players 3) Average Players and 4) Below Average Players

Stacked bar chart was used to analyze players across different measures in a single stack for a rich and compelling visualization

Insight: -

"Best players are 3 times more likely to score points than the average NBA player"

Does all best players come from top universities?

We can use Bubble chart to visualize this information. This chart enables us to visually encode large amounts of data in a small space. We can do a color encoding on Colleges to better represent the information. We are comparing number of players drafted to the NBA from universities based on clusters.

Insight: -

"Most players drafted from UCLA between 2002-2008 fall in the Below Average Player"

"Best players originate from lesser known universities (Kansas and North Carolina)"

How diverse is NBA?

We can use filled maps to visualize this insight. Filled maps allowed us to show the geographic spread of foreign players drafted in the NBA. We applied color encoding on birth country. We also wanted to show the player diversity present in the NBA which was possible with this chart type.

Insight: -

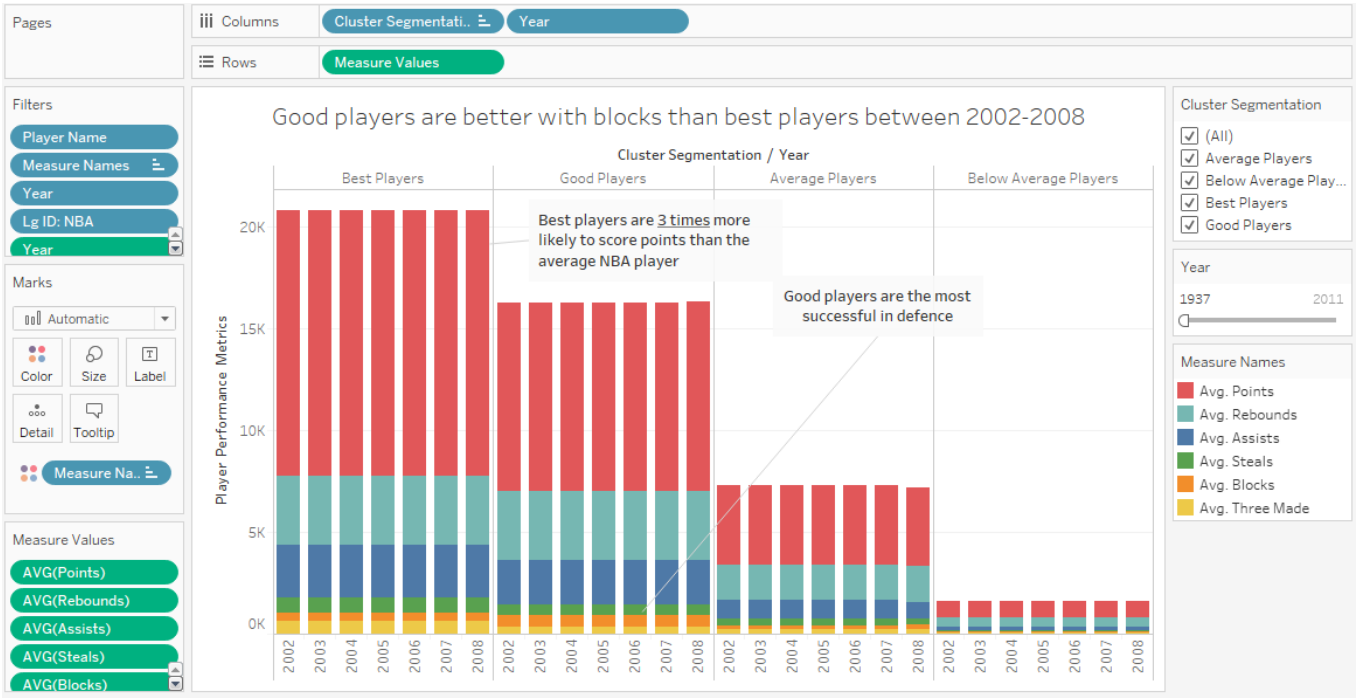
"Approximately 45% of all the foreign players drafted in the NBA are of European descent"

5 specific analytical questions that you want to answer: -

1. What impact team salaries have on debts?
2. Which team has good win rate and great points per game from 2002 – 2008.
3. Which are the top 5 teams of NBA and what is their winning and losing trends?
4. From where the max number of foreign players are drafted?
5. Does the best player is good in all aspects?

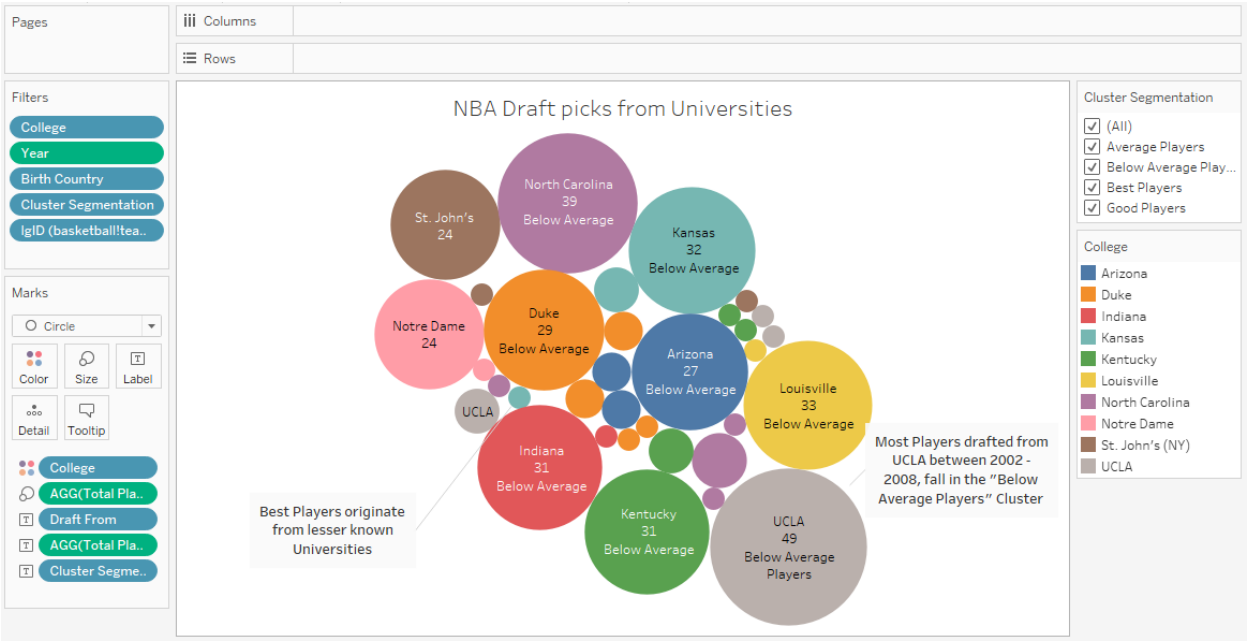
Data Visualization designed using Tableau

Viz 1. Does the best player is good in all aspects?



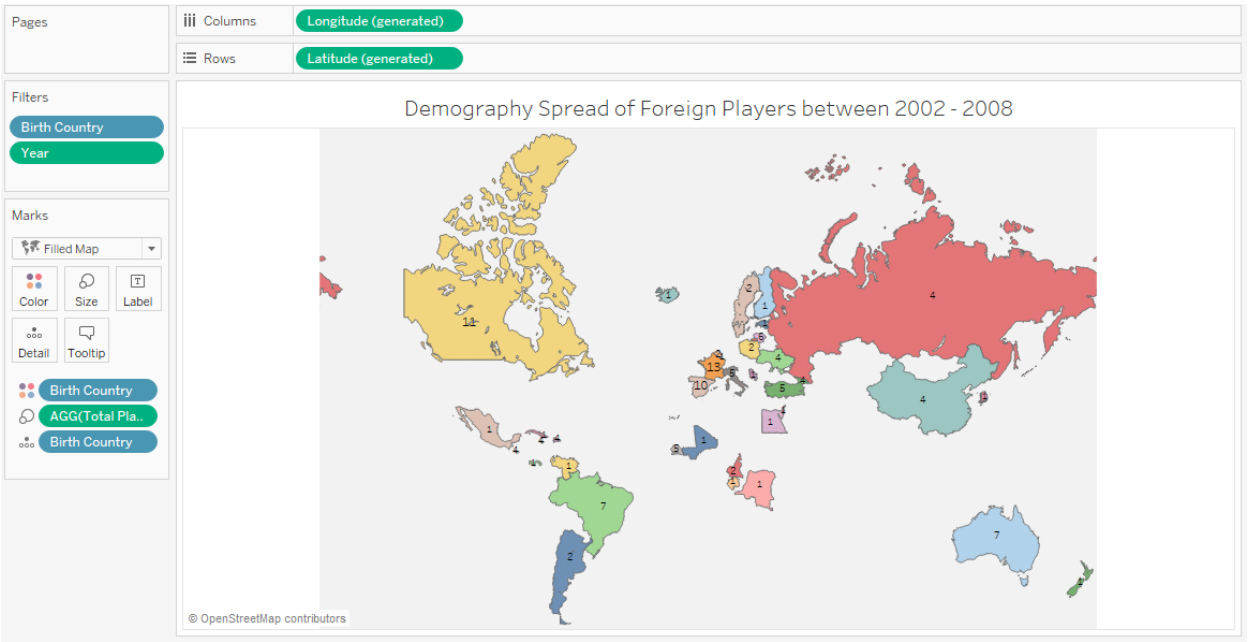
Above graph can be used to answer does the best player is good in all aspects? We can see that the best players are less successful in defence compared to good, average and below average players. Good players seem to be more successful in defence than best players.

Viz 2. Which university gives the best players?



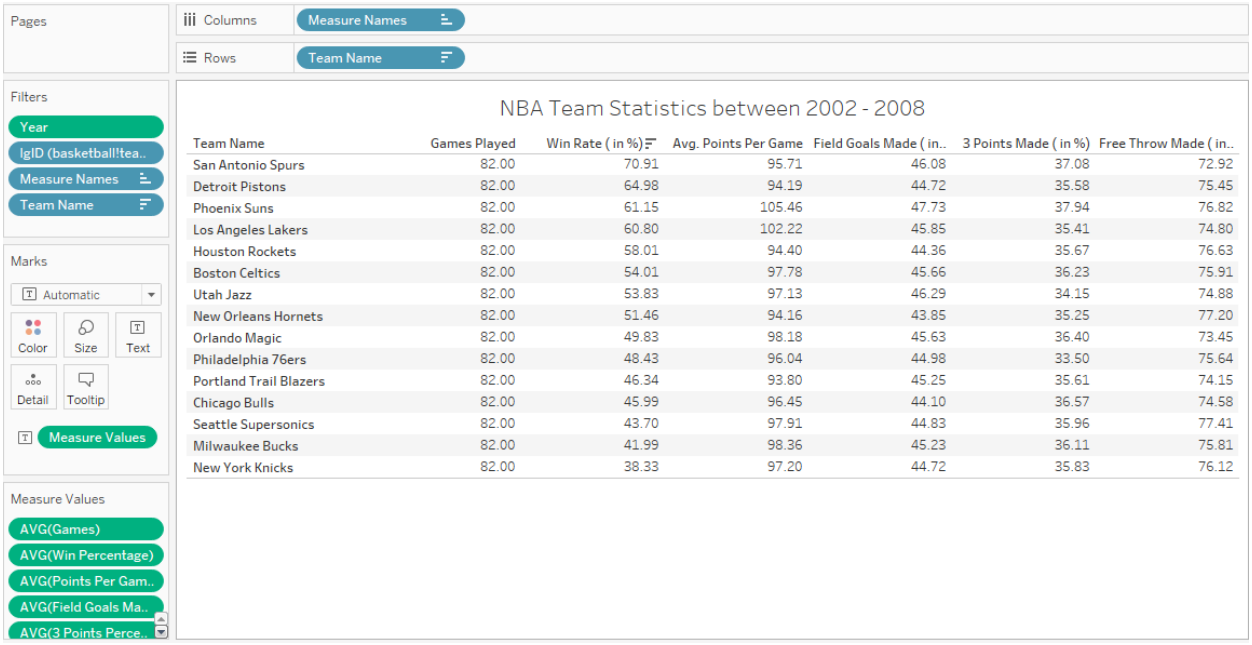
From the above bubble chart, we can see that the best players originate from lesser known universities compared to reputed universities. It is also observed that the players drafted from reputed universities are "Below Average Players".

Viz 3. How diverse is NBA?



We can see that players are drafted from all over the world. Number of players drafted are also known in above below.

Viz 4. Which team has good win rate and great points per game from 2002 – 2008.

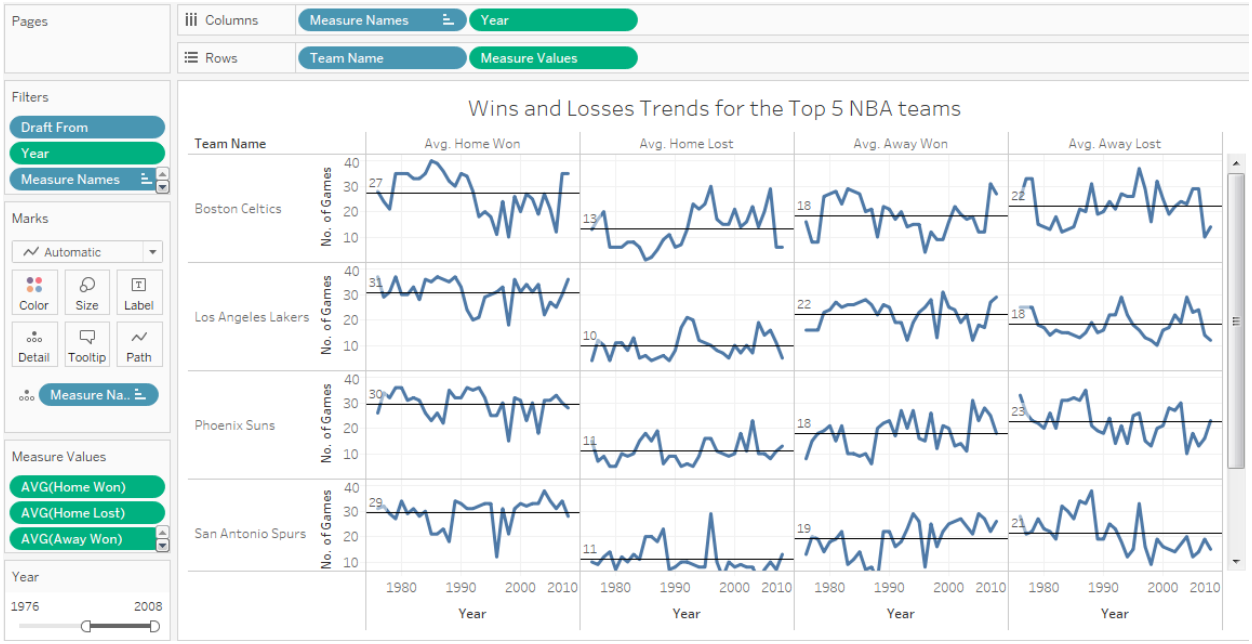


NBA Team Statistics between 2002 - 2008

Team Name	Games Played	Win Rate (in %)	Avg. Points Per Game	Field Goals Made (in..	3 Points Made (in %)	Free Throw Made (in..
San Antonio Spurs	82.00	70.91	95.71	46.08	37.08	72.92
Detroit Pistons	82.00	64.98	94.19	44.72	35.58	75.45
Phoenix Suns	82.00	61.15	105.46	47.73	37.94	76.82
Los Angeles Lakers	82.00	60.80	102.22	45.85	35.41	74.80
Houston Rockets	82.00	58.01	94.40	44.36	35.67	76.63
Boston Celtics	82.00	54.01	97.78	45.66	36.23	75.91
Utah Jazz	82.00	53.83	97.13	46.29	34.15	74.88
New Orleans Hornets	82.00	51.46	94.16	43.85	35.25	77.20
Orlando Magic	82.00	49.83	98.18	45.63	36.40	73.45
Philadelphia 76ers	82.00	48.43	96.04	44.98	33.50	75.64
Portland Trail Blazers	82.00	46.34	93.80	45.25	35.61	74.15
Chicago Bulls	82.00	45.99	96.45	44.10	36.57	74.58
Seattle Supersonics	82.00	43.70	97.91	44.83	35.96	77.41
Milwaukee Bucks	82.00	41.99	98.36	45.23	36.11	75.81
New York Knicks	82.00	38.33	97.20	44.72	35.83	76.12

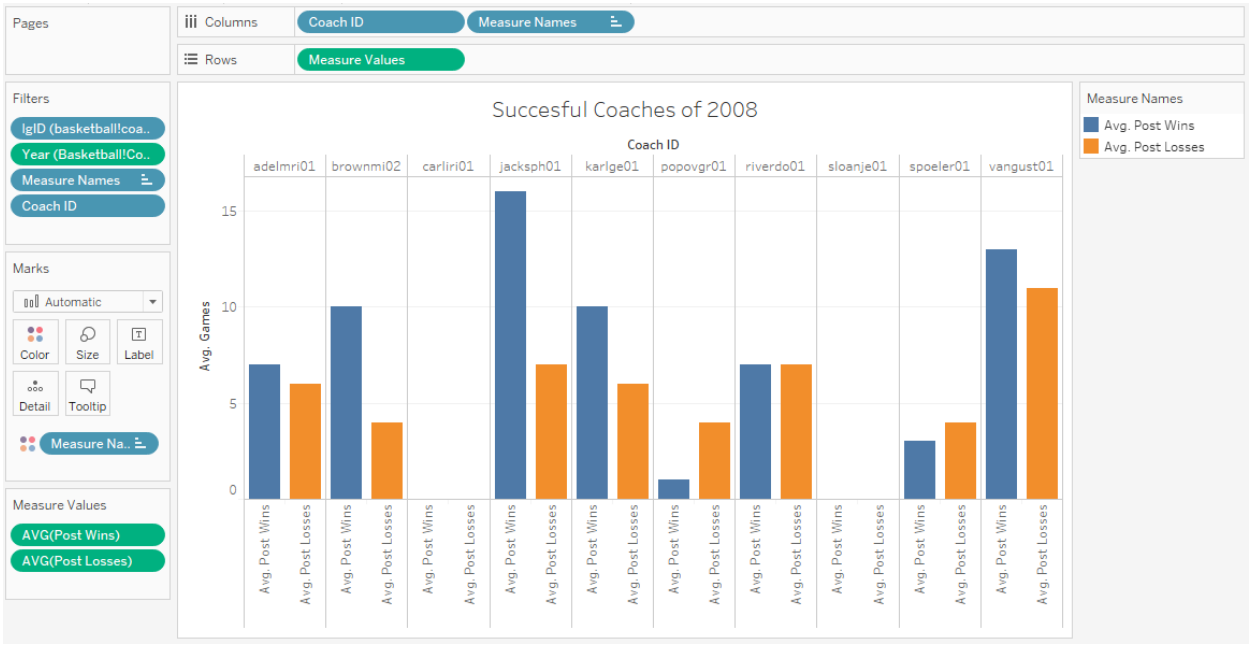
Following statistics shows which team played how many games? what was their win rate? Average points they earned per game and field goals made.

Viz 5. Which are the top 5 teams of NBA and what is their winning and losing trends?



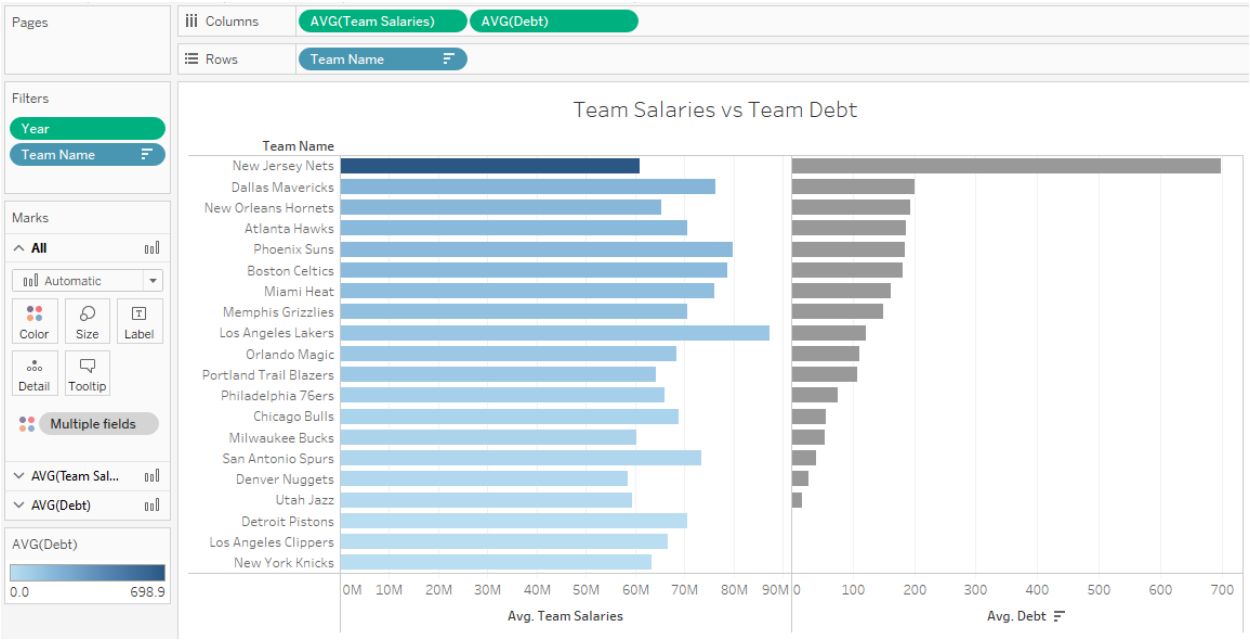
The following graph shows top 5 teams and their wins and losses with trend line.

Viz 6. Which all coaches are successful?



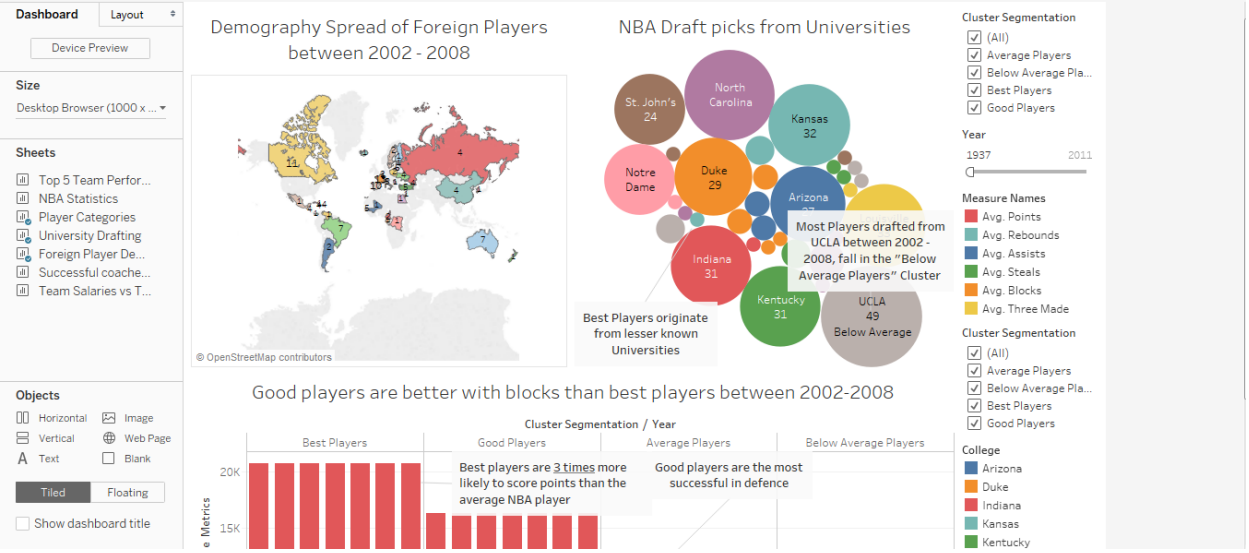
Above graph shows the IDs of all the successful coaches. Blanks shows that the particular coach did not coach a team in 2008.

Viz 7. What impact team salaries have on debts?



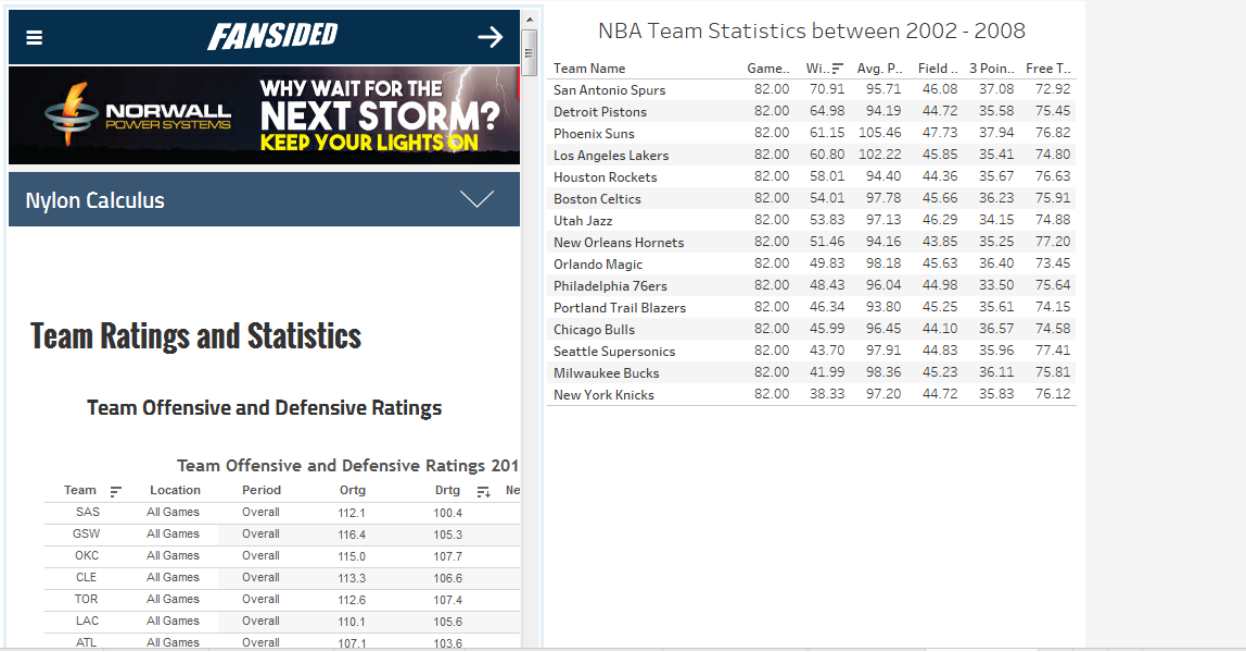
From the above graph we can see that New Jersey Nets have average salary but highest in debts. We can also see that Los Angeles Lakers have highest salary and low on average debts.

Dashboard (without Mashup):



In this dashboard, we can see that the players were selected from all over the world as well as university drafts made. We can control the categories of players from cluster segmentation between average players, good players etc. We can also select university's to see the category of players drafted from that university.

Dashboard (with Mashup):

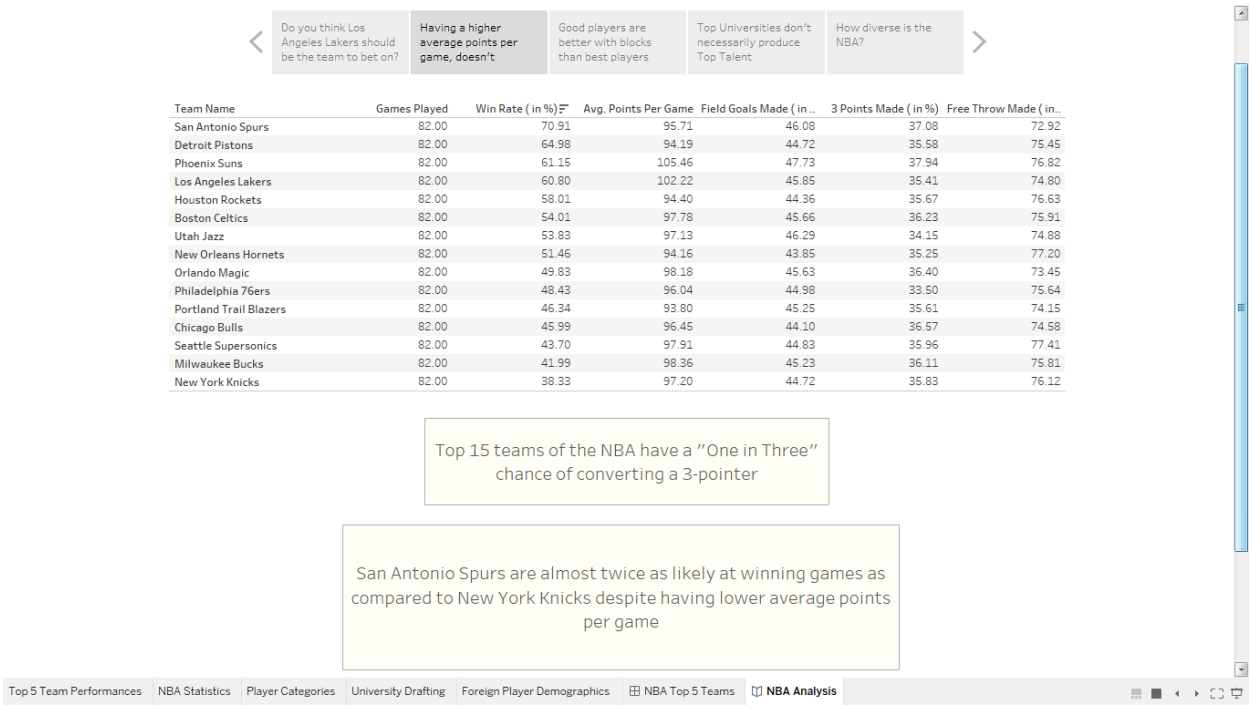


In this above mashup, we can see the team statistics we got from the data set and the team statistics from NBA website.

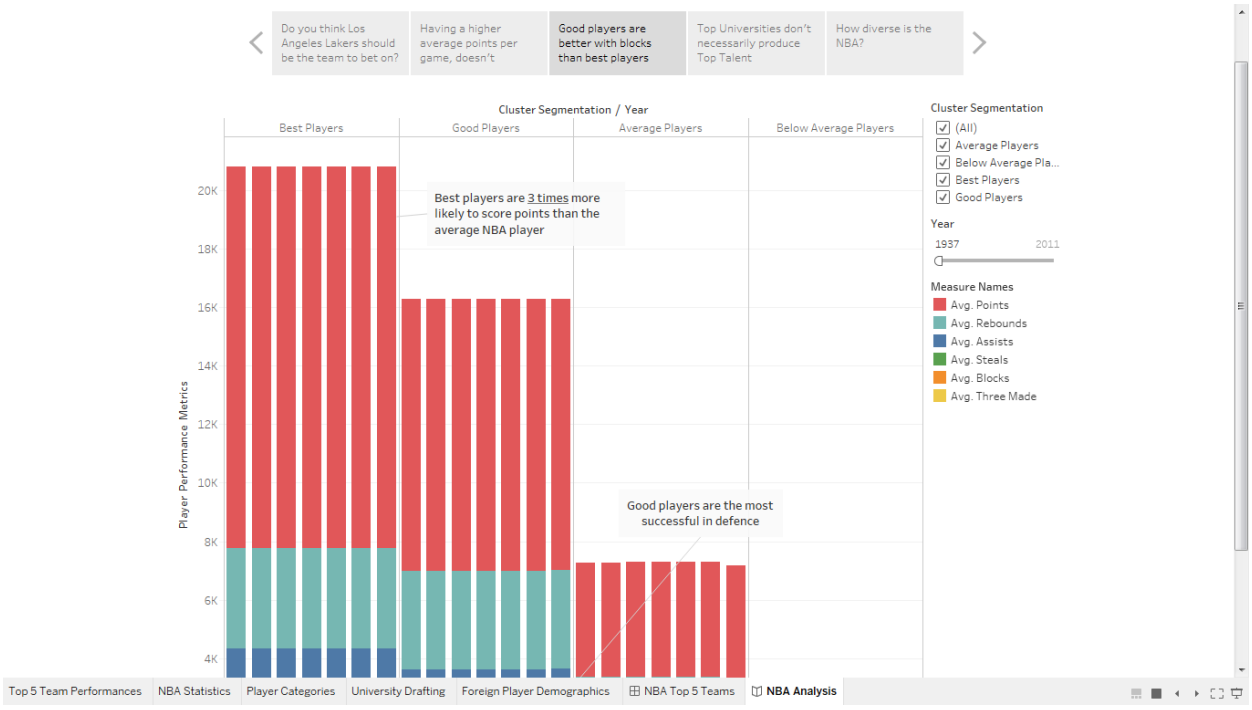
Story Line: -

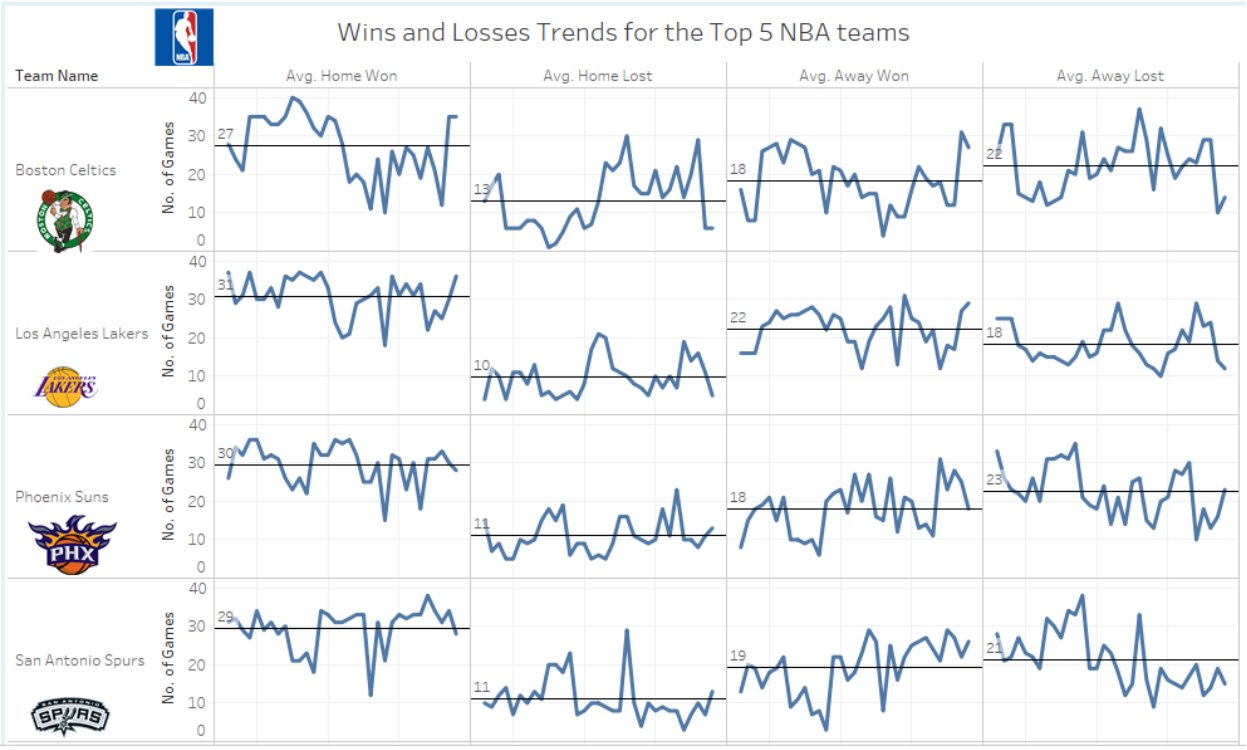
The primary actors/characters in this storyline are the basketball teams and players.

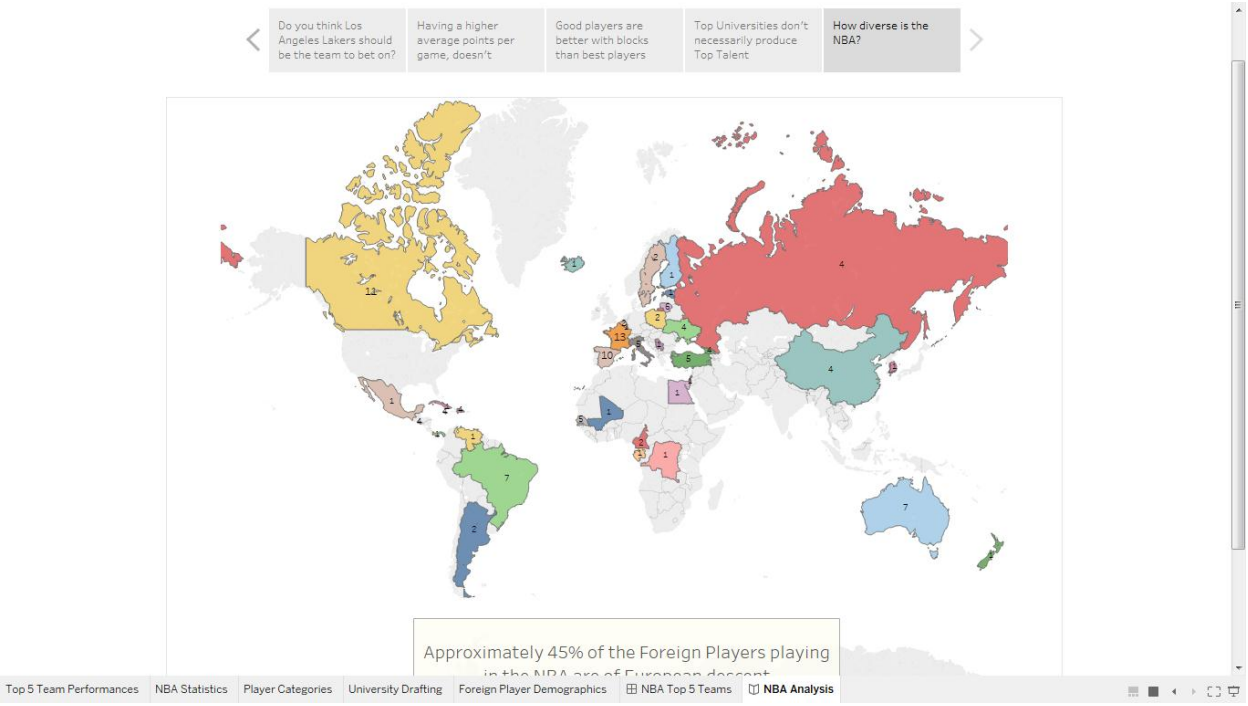
- We have woven a story around NBA analysis wherein we start off by showing trend analysis of the top 5 teams over a 30 year period.
- Next we move on by comparing the win rate between the top 15 most recent teams and try to capture the audience's attention towards metrics which may influence the teams standing and efficiency in the league.
- We show the most and the least impactful players by plugging in cluster analysis to get an idea of their differences.
- We then move onto showcase universities from where the NBA drafts its new players from. We have added player categories as a filter for exploratory analysis for the audiences.
- Lastly we sign off by showing the diversity amongst the players in the NBA and give an insight to our audience that NBA is a truly global sport.

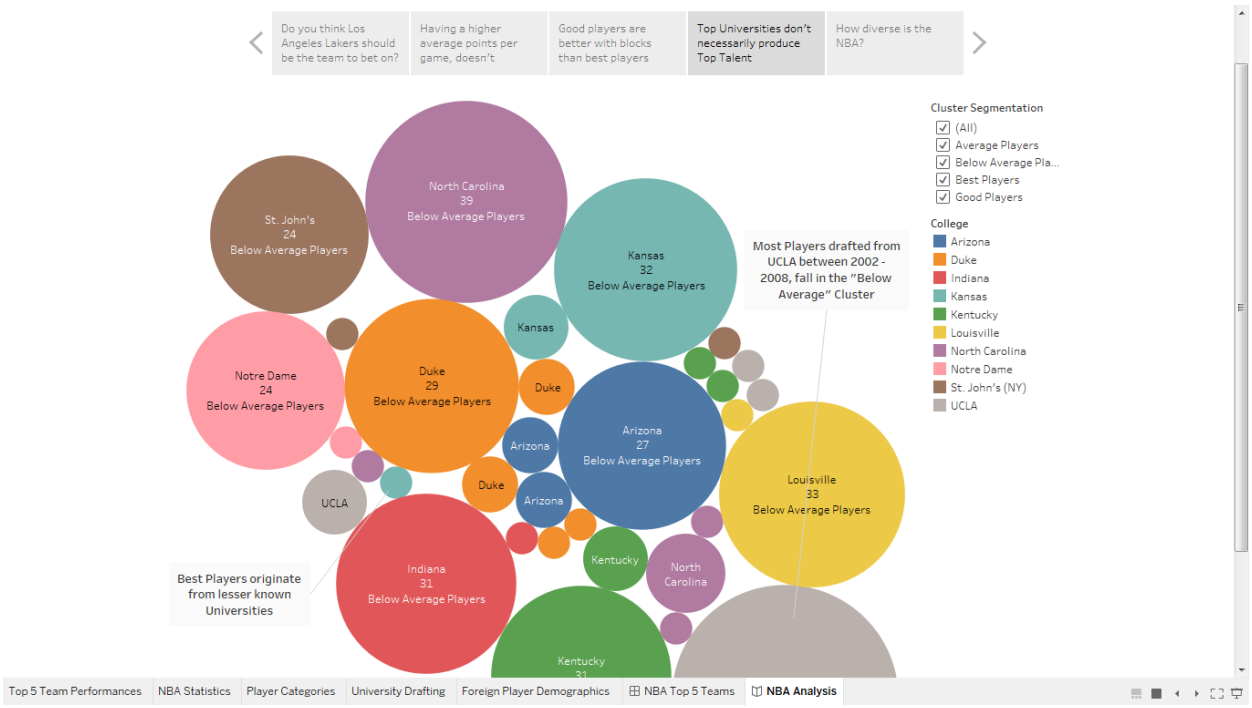


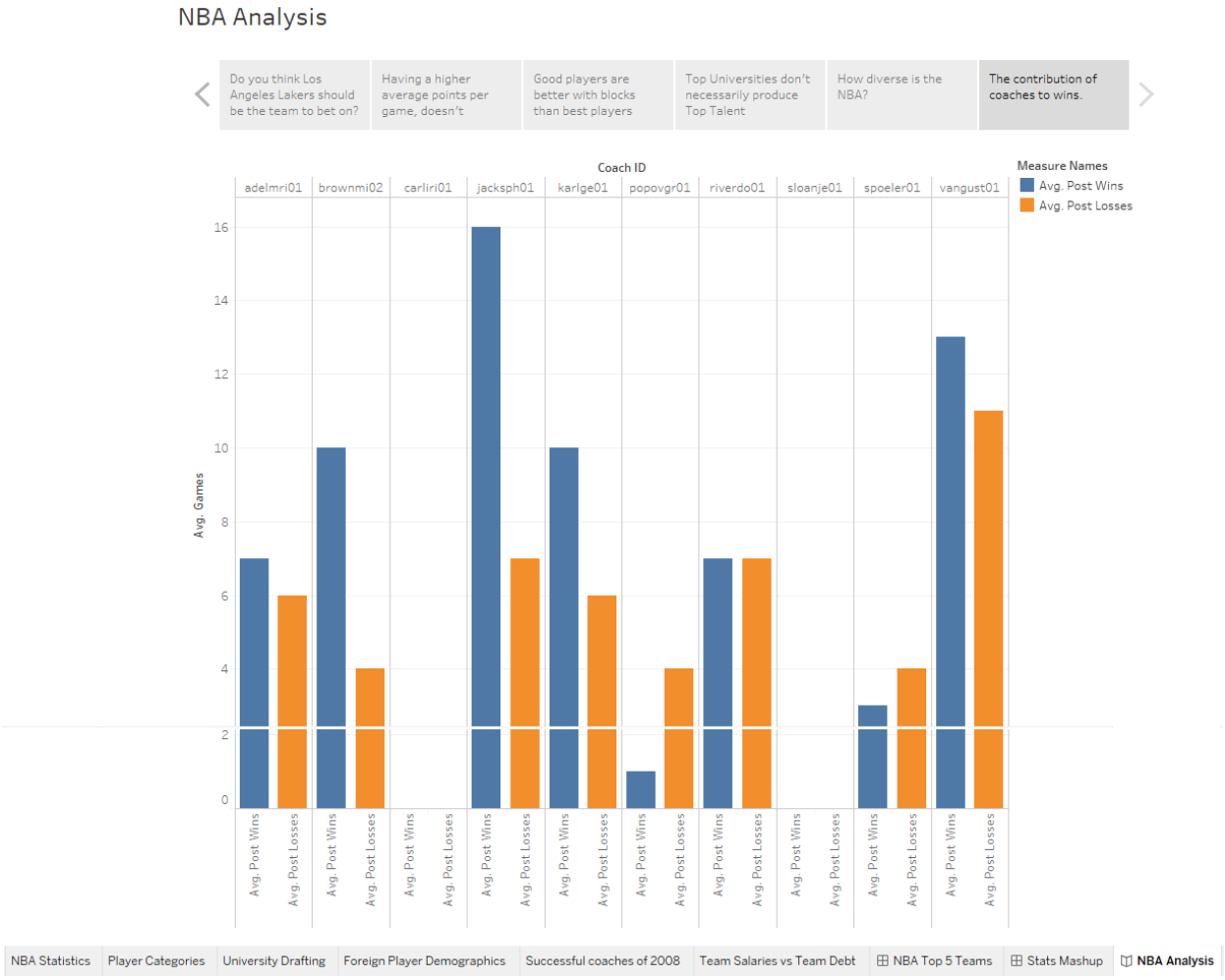
Team Name	Games Played	Win Rate (in %)	Avg. Points Per Game	Field Goals Made (in ..	3 Points Made (in %)	Free Throw Made (in..
San Antonio Spurs	82.00	70.91	95.71	46.08	37.08	72.92
Detroit Pistons	82.00	64.98	94.19	44.72	35.58	75.45
Phoenix Suns	82.00	61.15	105.46	47.73	37.94	76.82
Los Angeles Lakers	82.00	60.80	102.22	45.85	35.41	74.80
Houston Rockets	82.00	58.01	94.40	44.36	35.67	76.63
Boston Celtics	82.00	54.01	97.78	45.66	36.23	75.91
Utah Jazz	82.00	53.83	97.13	46.29	34.15	74.88
New Orleans Hornets	82.00	51.46	94.16	43.85	35.25	77.20
Orlando Magic	82.00	49.83	98.18	45.63	36.40	73.45
Philadelphia 76ers	82.00	48.43	96.04	44.98	33.50	75.64
Portland Trail Blazers	82.00	46.34	93.80	45.25	35.61	74.15
Chicago Bulls	82.00	45.99	96.45	44.10	36.57	74.58
Seattle Supersonics	82.00	43.70	97.91	44.83	35.96	77.41
Milwaukee Bucks	82.00	41.99	98.36	45.23	36.11	75.81
New York Knicks	82.00	38.33	97.20	44.72	35.83	76.12




Wins and Losses Trends for the Top 5 NBA teams
Team Name
Avg. Home Won
Avg. Home Lost
Avg. Away Won
Avg. Away Lost
Boston Celtics
Los Angeles Lakers
Phoenix Suns
San Antonio Spurs
No. of Games
27
13
18
22
31
10
22
18
30
11
18
23
29
11
19
21

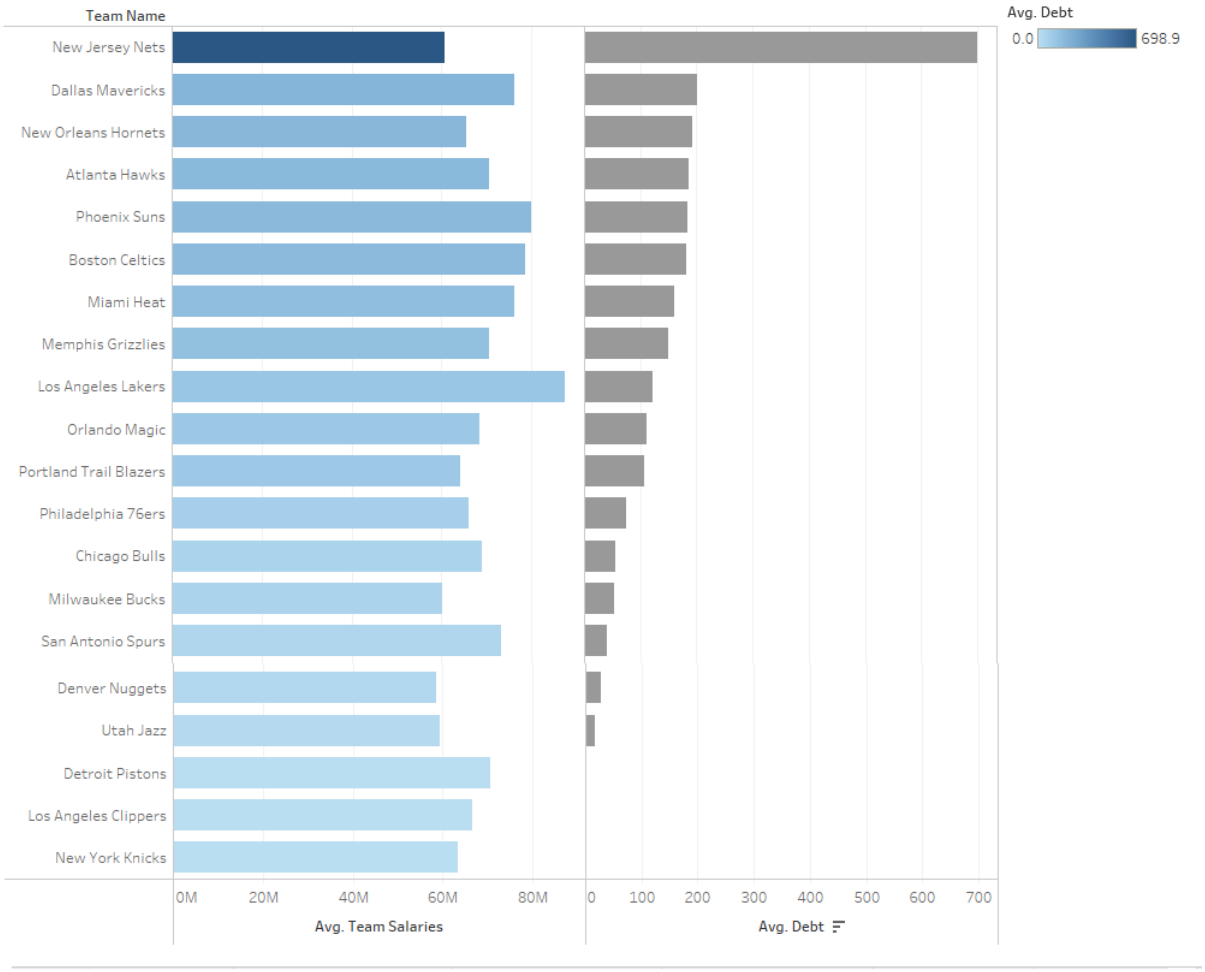

Do you think Los Angeles Lakers should be the team to bet on?
Having a higher average points per game, doesn't
Good players are better with blocks than best players
Top Universities don't necessarily produce Top Talent
How diverse is the NBA?
Approximately 45% of the Foreign Players playing
Top 5 Team Performances
NBA Statistics
Player Categories
University Drafting
Foreign Player Demographics
NBA Top 5 Teams
NBA Analysis


Do you think Los Angeles Lakers should be the team to bet on?
Having a higher average points per game, doesn't
Good players are better with blocks than best players
Top Universities don't necessarily produce Top Talent
How diverse is the NBA?
Cluster Segmentation
(All)
Average Players
Below Average Pla...
Best Players
Good Players
College
Arizona
Duke
Indiana
Kansas
Kentucky
Louisville
North Carolina
Notre Dame
St. John's (NY)
UCLA
North Carolina 39 Below Average Players
St. John's 24 Below Average Players
Kansas 32 Below Average Players
Most Players drafted from UCLA between 2002 - 2008, fall in the "Below Average" Cluster
Notre Dame 24 Below Average Players
Duke 29 Below Average Players
Kansas
Duke
Arizona 27 Below Average Players
Arizona
Duke
Arizona
UCLA
Louisville 33 Below Average Players
Best Players originate from lesser known Universities
Indiana 31 Below Average Players
Kentucky
North Carolina
Kentucky 31
Top 5 Team Performances
NBA Statistics
Player Categories
University Drafting
Foreign Player Demographics
NBA Top 5 Teams
NBA Analysis


NBA Analysis
Do you think Los Angeles Lakers should be the team to bet on?
Having a higher average points per game, doesn't
Good players are better with blocks than best players
Top Universities don't necessarily produce Top Talent
How diverse is the NBA?
The contribution of coaches to wins.
Coach ID
adelmri01
brownmi02
carliri01
jacksph01
karlge01
popovgr01
riverdo01
sloanje01
spoeler01
vangust01
Measure Names
Avg. Post Wins
Avg. Post Losses
Avg. Games
16
14
12
10
8
6
4
2
0
Avg. Post Wins
Avg. Post Losses
NBA Statistics
Player Categories
University Drafting
Foreign Player Demographics
Successful coaches of 2008
Team Salaries vs Team Debt
NBA Top 5 Teams
Stats Mashup
NBA Analysis

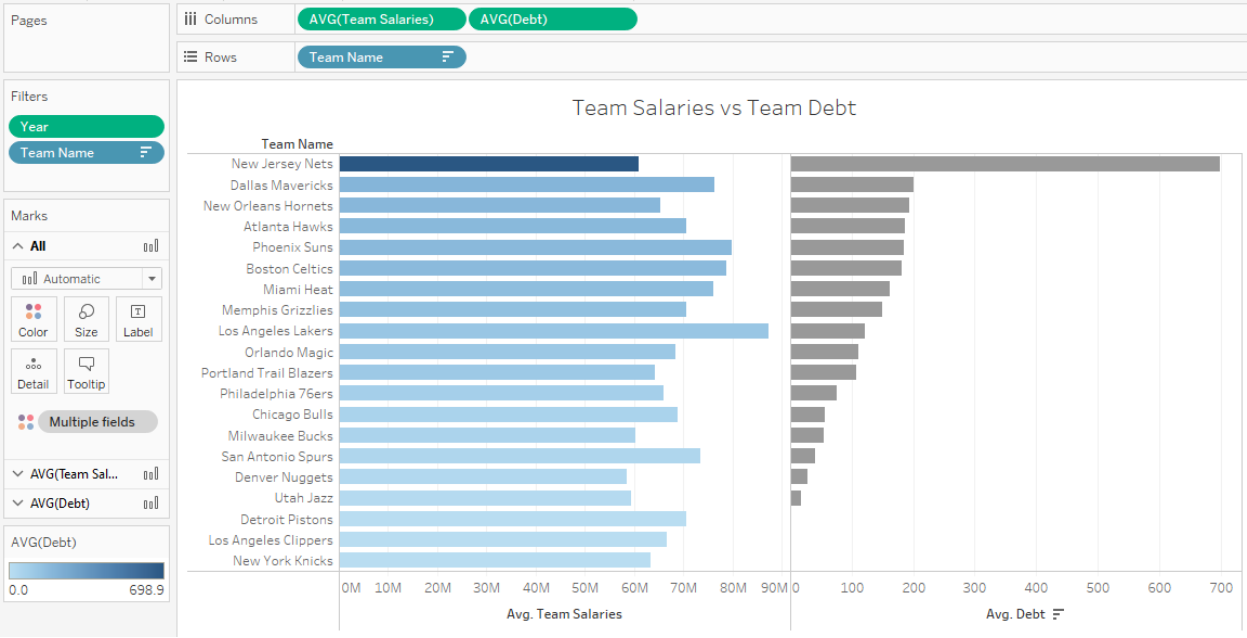
NBA Analysis

Having a higher average points per game, doesn't | Good players are better with blocks than best players | Top Universities don't necessarily produce Top Talent | How diverse is the NBA? | The contribution of coaches to wins. | Does the team salaries have effect on debts?



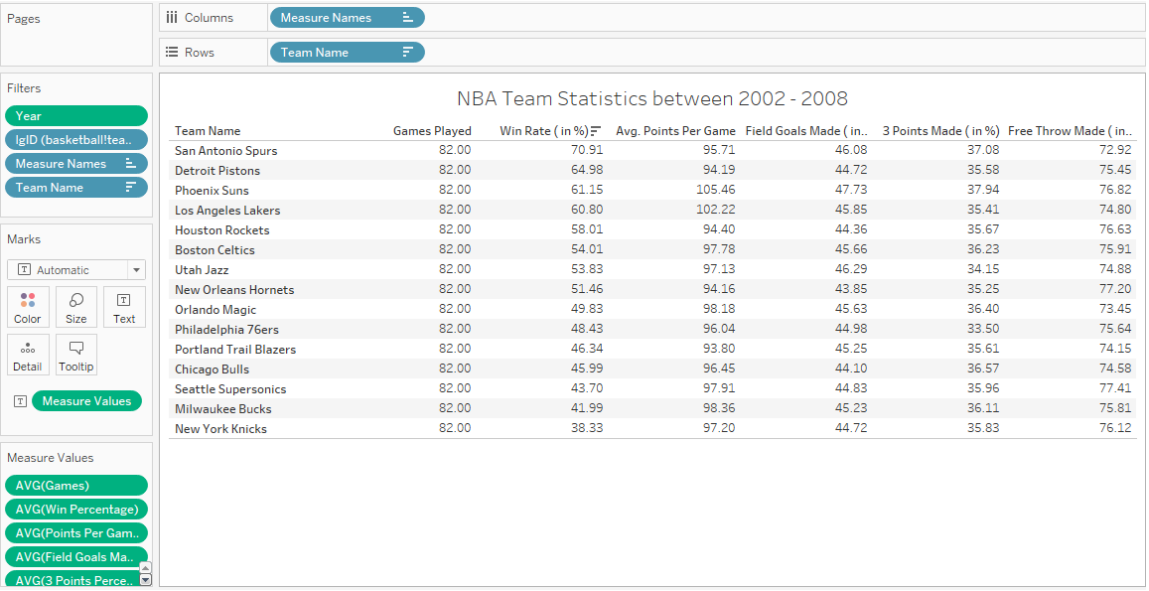
Answers for 5 specific analytical questions: -

1. What impact team salaries have on debts?



It is observed that team with highest salary is very less in debt but is also observed that team with average salary is highest in debts.

2. Which team has good win rate and great points per game from 2002 – 2008.

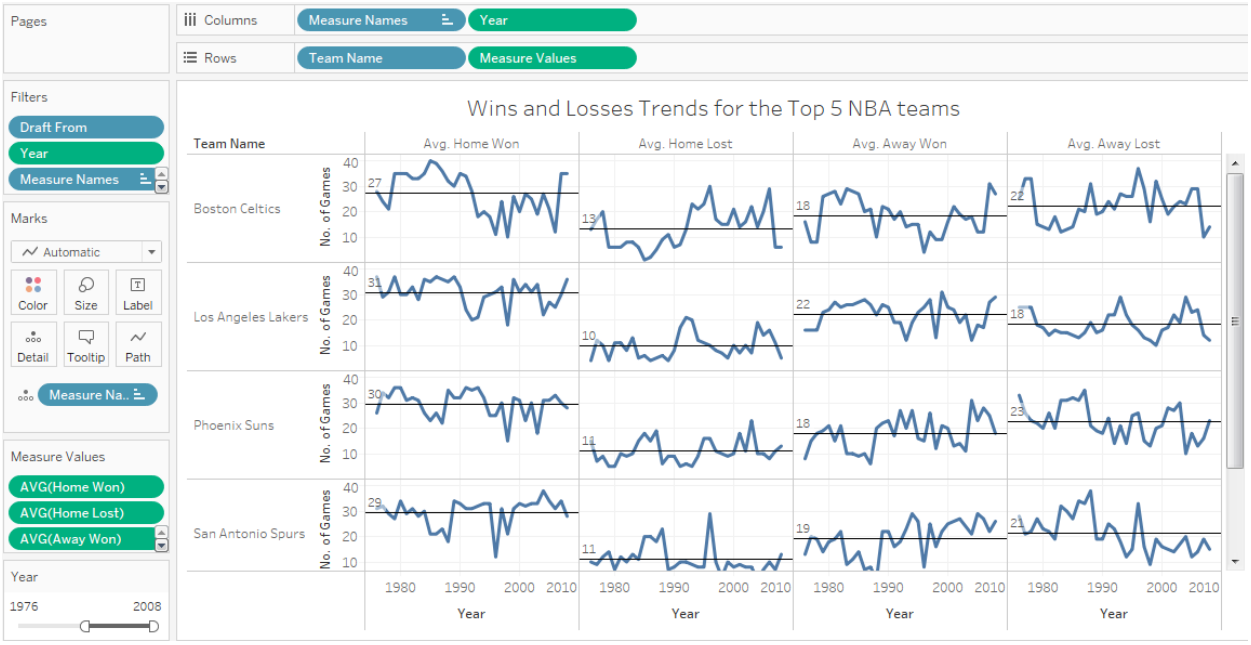


NBA Team Statistics between 2002 - 2008

Team Name	Games Played	Win Rate (in %)	Avg. Points Per Game	Field Goals Made (in..	3 Points Made (in %)	Free Throw Made (in..
San Antonio Spurs	82.00	70.91	95.71	46.08	37.08	72.92
Detroit Pistons	82.00	64.98	94.19	44.72	35.58	75.45
Phoenix Suns	82.00	61.15	105.46	47.73	37.94	76.82
Los Angeles Lakers	82.00	60.80	102.22	45.85	35.41	74.80
Houston Rockets	82.00	58.01	94.40	44.36	35.67	76.63
Boston Celtics	82.00	54.01	97.78	45.66	36.23	75.91
Utah Jazz	82.00	53.83	97.13	46.29	34.15	74.88
New Orleans Hornets	82.00	51.46	94.16	43.85	35.25	77.20
Orlando Magic	82.00	49.83	98.18	45.63	36.40	73.45
Philadelphia 76ers	82.00	48.43	96.04	44.98	33.50	75.64
Portland Trail Blazers	82.00	46.34	93.80	45.25	35.61	74.15
Chicago Bulls	82.00	45.99	96.45	44.10	36.57	74.58
Seattle Supersonics	82.00	43.70	97.91	44.83	35.96	77.41
Milwaukee Bucks	82.00	41.99	98.36	45.23	36.11	75.81
New York Knicks	82.00	38.33	97.20	44.72	35.83	76.12

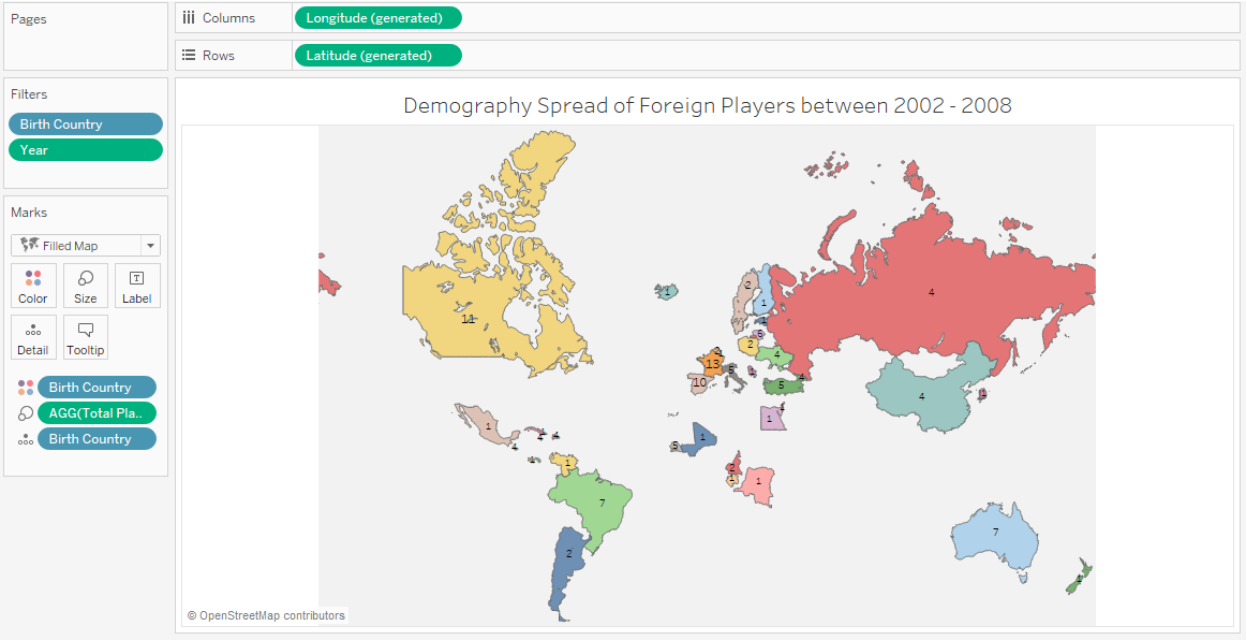
San Antonio Spurs are almost twice as likely at winning games as compared to New York Knicks despite having lower average points per game.

3. Which are the top 5 teams of NBA and what is their winning and losing trends?



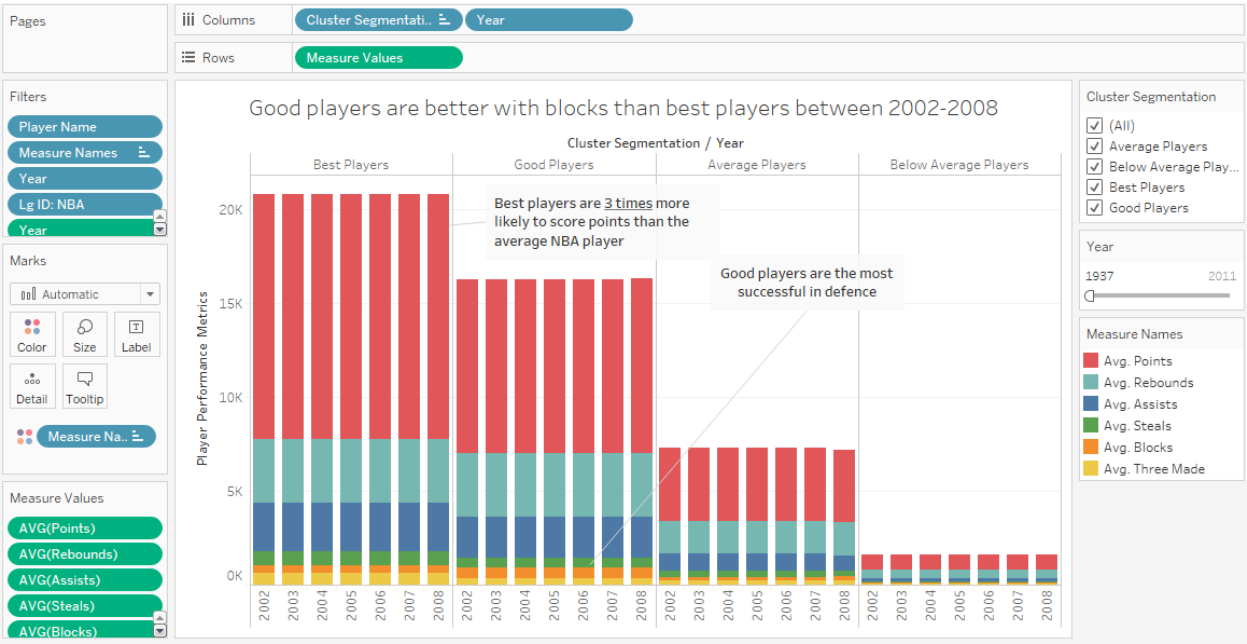
We can see that Boston Celtics is top team from the trend line.

4. From where the max number of foreign players are drafted?



NBA has players from all over the world. Approximately 45% of all the foreign players drafted in the NBA are of European descent

5. Does the best player is good in all aspects?



Best players are 3 times more likely to score points than the average NBA player but good players are the most successful in defence.