A2. Exploratory Data Analysis

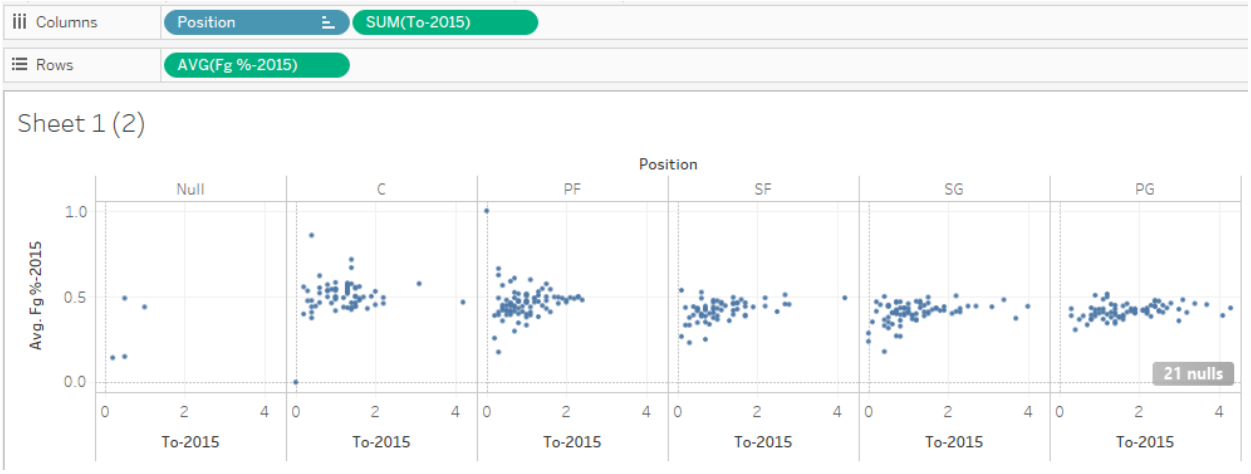
Krissy Gianforte

W209 - Section 1

Hypothesis 1: Players with high turnover statistics make up for that detriment in other ways. In particular, guards with high turnovers must also have high assist numbers, and lower-position players (centers and forwards) with high turnovers must also have high points-per-game.

What's informative about this view: This view shows the field goal shooting percentage of each player against his turnovers-per-game in 2015. It is a useful first glance at any correlation; if players truly made up for turnovers with higher shooting percentages, we'd expect to see the data slope upward to the right (more turnovers = higher percentage). It actually does not do that, it seems.

What could be improved about this view: More separation for features would be helpful to differentiate trends among different groups. For example, separating players of various positions.

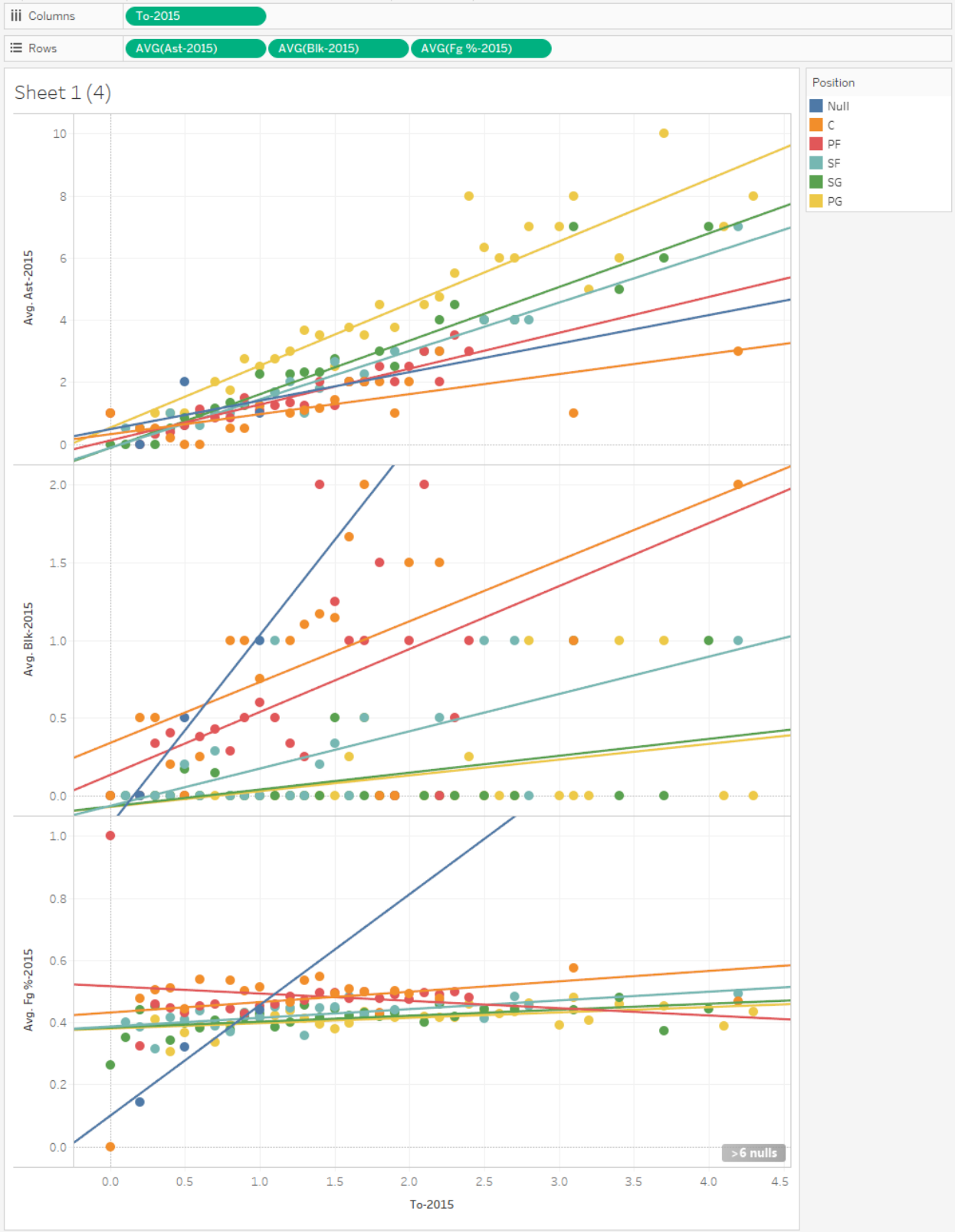


What's informative about this view: This view adds position as a filter, so that I can see the relationship between turnovers and shooting percentage for each of the five positions. I ordered the plots left-to-right from the “lower” to “higher” positions (Center, Power Forward, Small Forward, Shooting Guard, and Point Guard). As stated in the hypothesis, I think guards (PG) would make up for turnovers in different ways that centers (C).

What could be improved about this view: This view still doesn’t let me see all of the possible “make-up” statistics at once. I would like to see assists and blocks in the same view, so that I can see if different types of players bring different strengths. (I believe they do!)

What's informative about this view: This view adds plots for assists and blocks, still keeping things separated by position. This is helpful – I can see here that assists is actually most strongly correlated with turnovers. Those with higher turnovers tend to have higher assists, particularly with point guards (PG). That makes sense; to get many good assists through, you probably also throw some bad passes! I can also see that lower players (centers and power forwards) may have a correlation between assists and blocks. While those aren't necessarily related during actual play, perhaps the players are simply making up for turnovers by hustling on defense!

What could be improved about this view: Aggregating the data (rather than displaying by individual player) may help me confirm what I see in the plots.


Columns
To-2015
Rows
AVG(Ast-2015)
AVG(Blk-2015)
AVG(Fg %-2015)
Sheet 1 (4)
Position
Null
C
PF
SF
SG
PG
Avg. Ast-2015
Avg. Blk-2015
Avg. Fg %-2015
>6 nulls
To-2015

What's informative about this view: This view puts all positions' data on the same graph (still separating shooting percentage / assists / blocks) and displays trend lines for each position. From that, I can see trends in the data that are difficult to be sure of given the dots alone.

Ignoring the "null" position values (the blue lines), I can see that point guards have the strongest positive correlation between assists and turnovers; centers and power forwards have the strongest correlation between blocks and turnovers (though centers have simply more blocks than PFs – the orange line is shifted up from the pink one). Most players have a slight positive relationship between field goal percentage and turnovers, though power forwards actually have worse shooting given more assists.

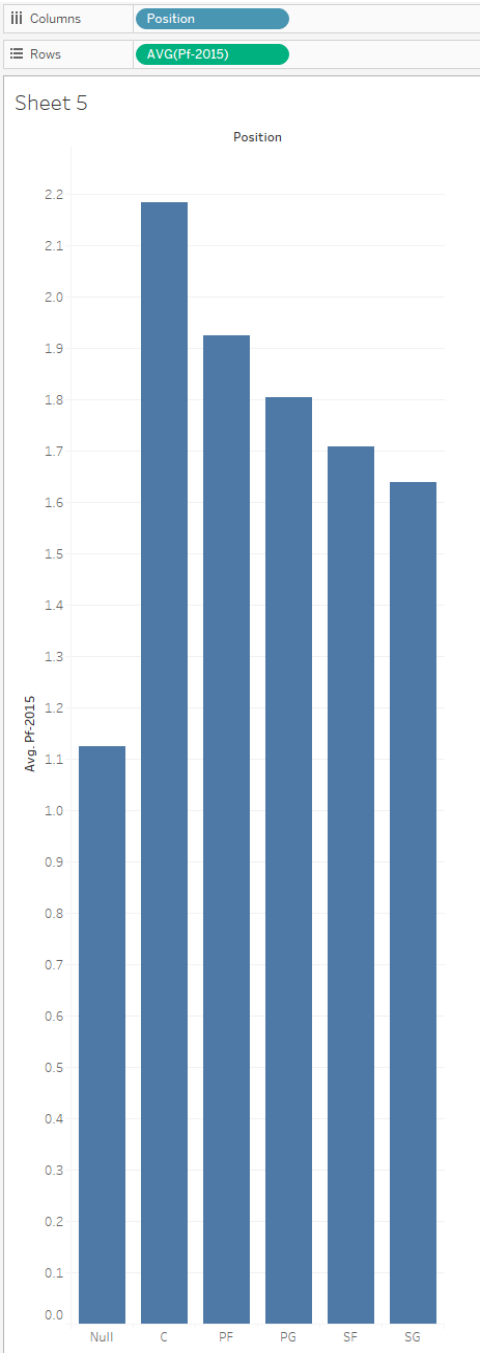
What could be improved about this view: The view is simply a bit busy! I could do more work to reduce some clutter and make the trend lines the focus. Additionally, I could change the color palate to use more pastels and less harsh neon colors.

Conclusion (do the data appear to support the hypothesis, or not?): These data views suggest that players *do* make up for turnovers by increasing their "positive" player statistics. Guards tend to increase assists, whereas lower players like centers and power forwards tend to increase their blocking.

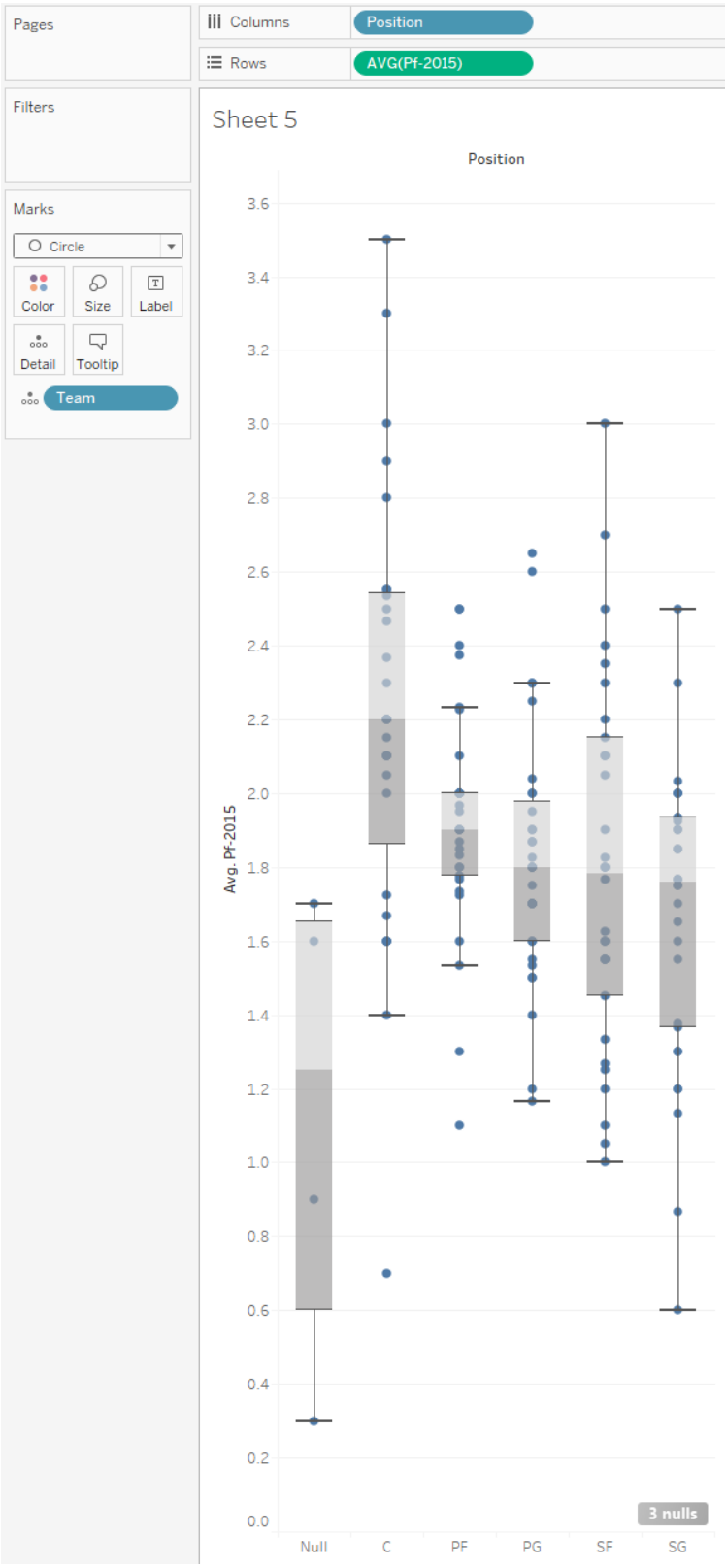
PLEASE NOTE THAT THE TEMPLATE IS NEEDED FOR ONLY THE FIRST HYPOTHESIS. THE OTHER TWO ONLY NEED TO INCLUDE THE HYPOTHESIS AND THE FIRST AND LAST VIEWS CREATED (NO ADDITIONAL TEXT NEEDED).

Hypothesis 2: Players in the Center position hade more personal fouls on average in 2015 than players in higher positions (shooting guards and point guards).

FIRST:



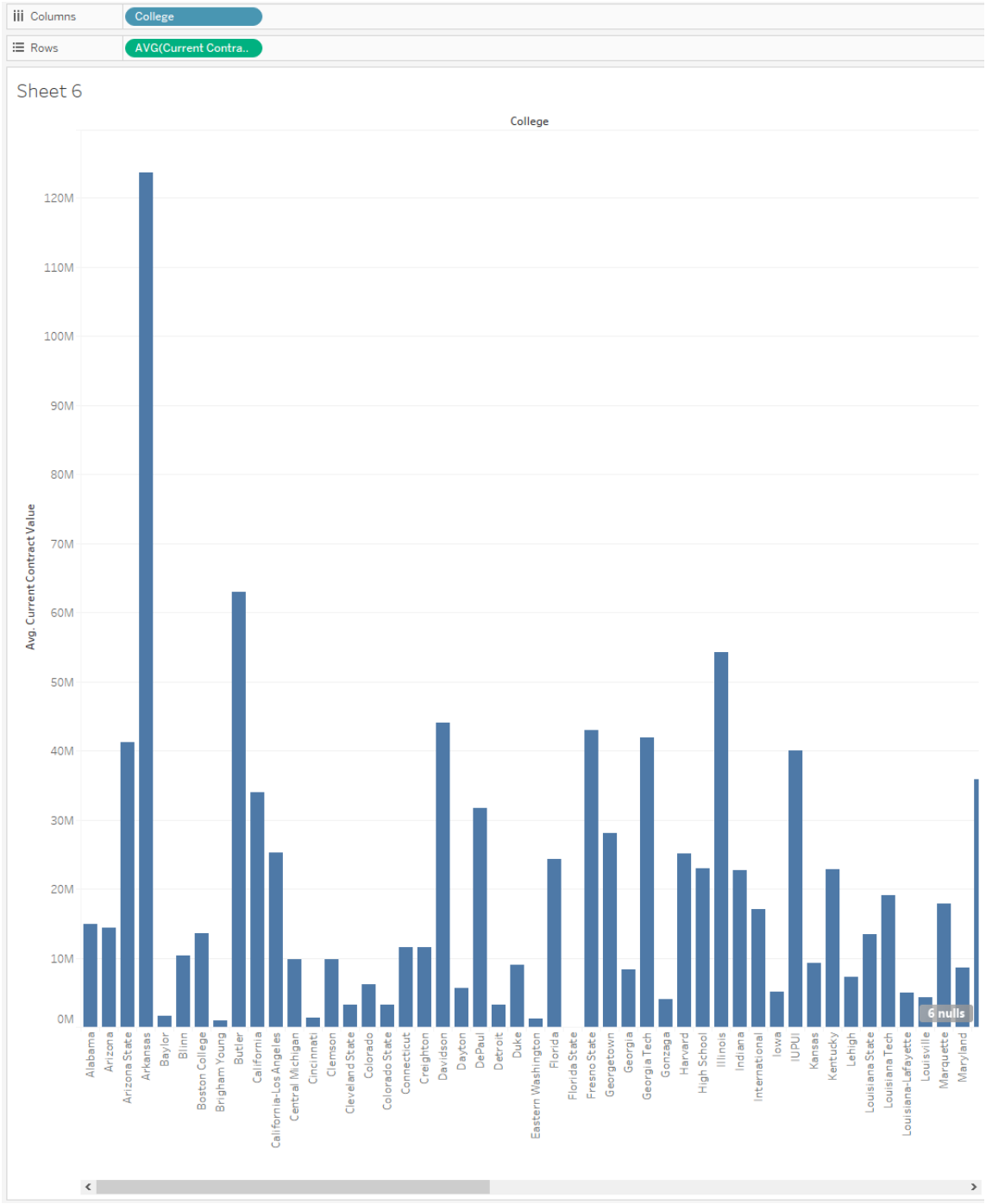
FINAL:


Pages
Columns
Position
Rows
AVG(Pf-2015)
Filters
Marks
Circle
Color
Size
Label
Detail
Tooltip
Team
Sheet 5
Position
Avg. Pf-2015
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
3 nulls
Null
C
PF
PG
SF
SG

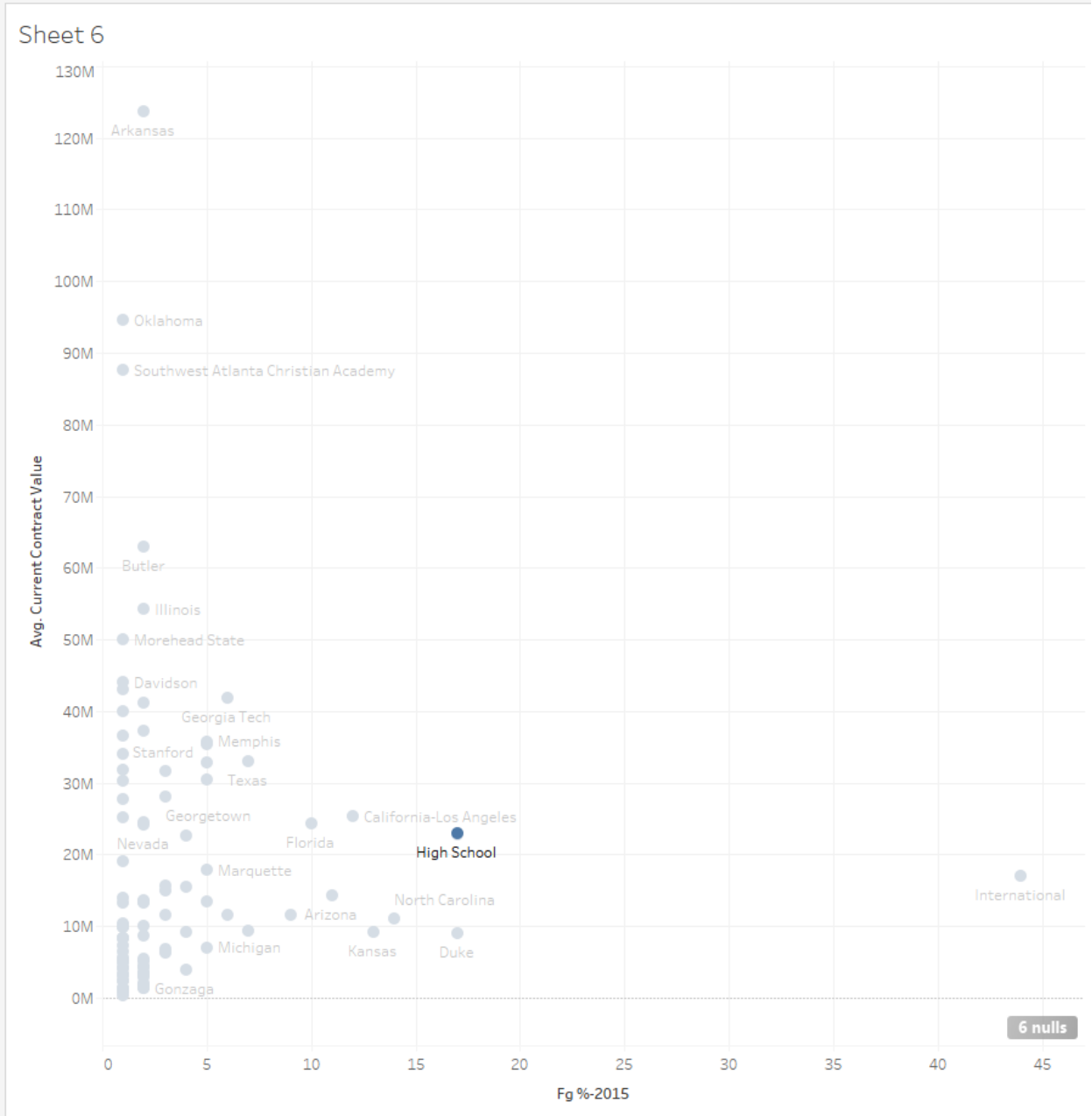
Hypothesis 3: Players with only a high school education (where “College” = “high school”) have higher shooting percentage but lower salaries / contract value.

(I imagine they must be good players to be recruited straight from high school, but I think their limited alternate career options would make them accept lower contract values.)

FIRST:



FINAL:


Sheet 6
Avg. Current Contract Value
130M
120M
110M
100M
90M
80M
70M
60M
50M
40M
30M
20M
10M
0M
Arkansas
Oklahoma
Southwest Atlanta Christian Academy
Butler
Illinois
Morehead State
Davidson
Georgia Tech
Memphis
Stanford
Texas
Georgetown
California-Los Angeles
Nevada
Florida
High School
Marquette
North Carolina
International
Arizona
Kansas
Duke
Michigan
Gonzaga
6 nulls
0
5
10
15
20
25
30
35
40
45
Fg %-2015