

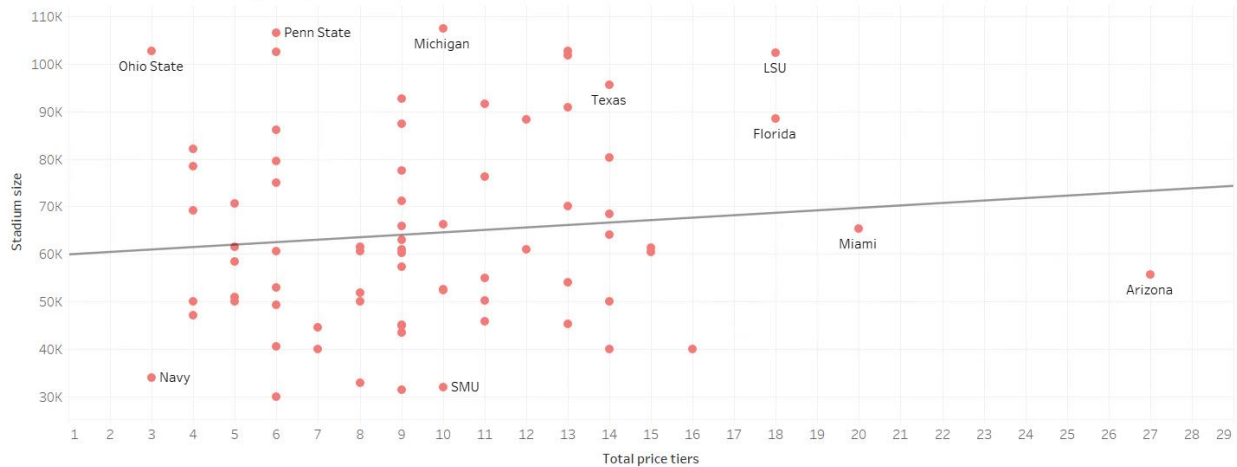
Analysis

For the purpose of this analysis, Notre Dame is included in the ACC conference. While we understand that they will not always be included in the ACC, it was important to include them in the place of best fit, rather than to leave them out altogether. Much of this analysis is based on correlations, so there is a short discussion at the very end to act as a primer on correlations.

The analysis is focused on 50 yd and Endzone prices. These sections were selected because they typically reflect the high and low end of pricing for a football stadium and given this commonality, they are better for comparing prices across a multitude of schools.

How does stadium capacity impact the number of pricing tiers?

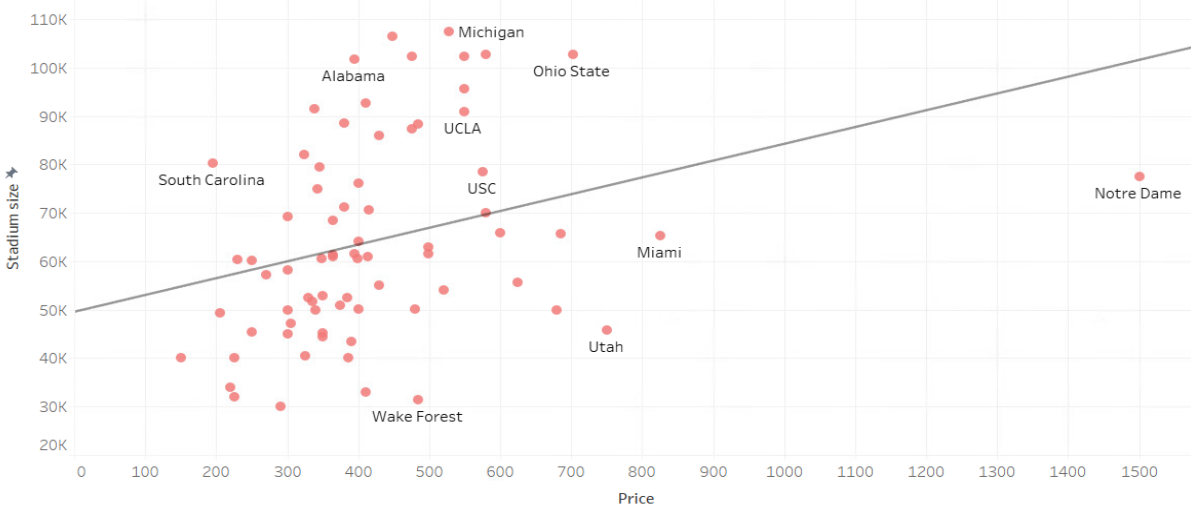
Price Tiers Vs. Stadium Capacity



- There is a weak relationship between pricing tiers and the size of the stadium. One might expect that with larger capacities more price tiers would be introduced, but this is not entirely the case.
- The visual does show that there is an opportunity for larger stadiums (Penn State, Michigan, Texas, etc.) to create more tiers if they chose.

What is the relationship between stadium capacity and 50-yard line tickets?

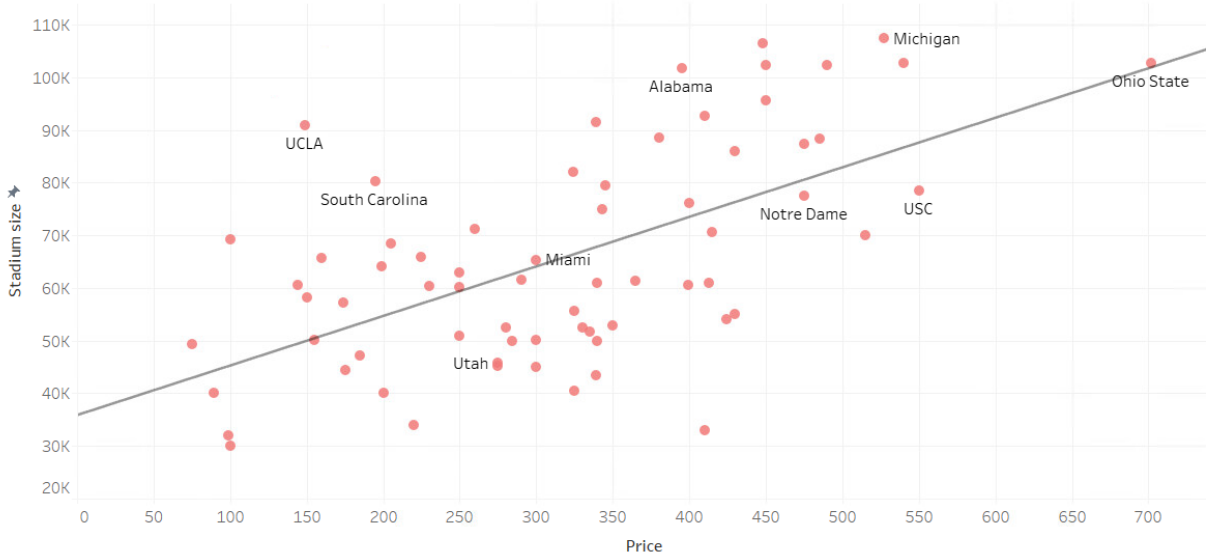
Low-50yd Cost Vs. Stadium Capacity



- There is a weak correlation between the two variables, even when accounting for outliers like Notre Dame.

What is the relationship between stadium capacity & endzone tickets?

Low-Endzone Cost Vs. Stadium Capacity



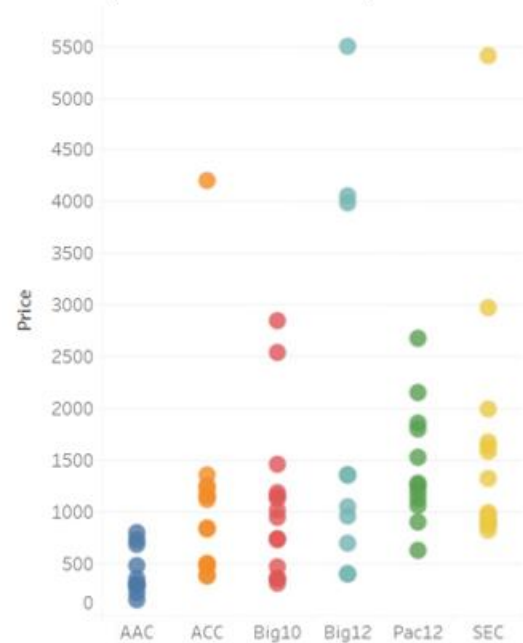
- There is a moderate correlation between these two variables. While the relationship between the 50 yd line seats and stadium capacity shown above is generally weak, there is a stronger correlation between Endzone seats and stadium capacity. This suggests that capacity plays a more direct role in setting prices for these endzone seats while there are other factors in play when pricing those desirable 50 yd line seats.

Price Ranges by Conference- 50 Yd Tickets & Total (Includes Donation Requirements)

Low-50yd by Conference



Low-50yd & Low-Don-50 by Conference



- When looking at the price range of tickets we find that the ACC (Notre Dame¹) and Pac 12 (Utah) have the highest average ticket price.
- When donation requirements are added in, the range of prices is a greater in the Big12 and SEC. The Big12 (Iowa St.) has the highest total cost for tickets at the 50-yard line.

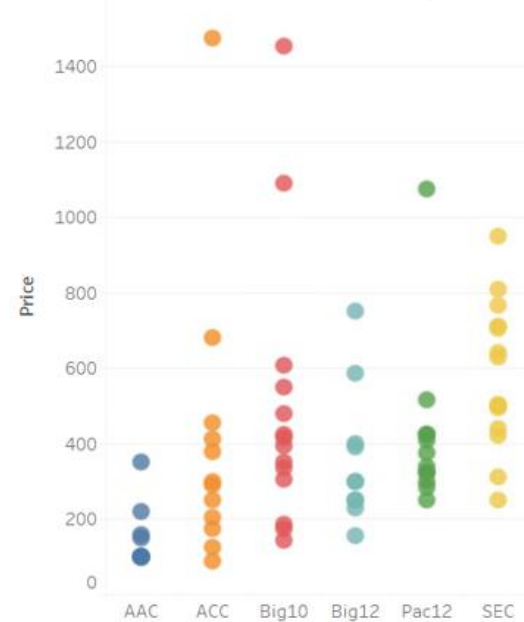
¹ It should be noted that Notre Dame does not actually appear as a point on the visual to the left, because leaving it in would skew the rest of the figure. The top point showing for the ACC there is Miami, but it should not be counted as having the highest ticket price, because for that school we were not able to separate out ticket and donation requirements.

Price Ranges by Conference- Endzone Tickets & Total (Includes Donation Requirements)

Low-Endzone by Conference



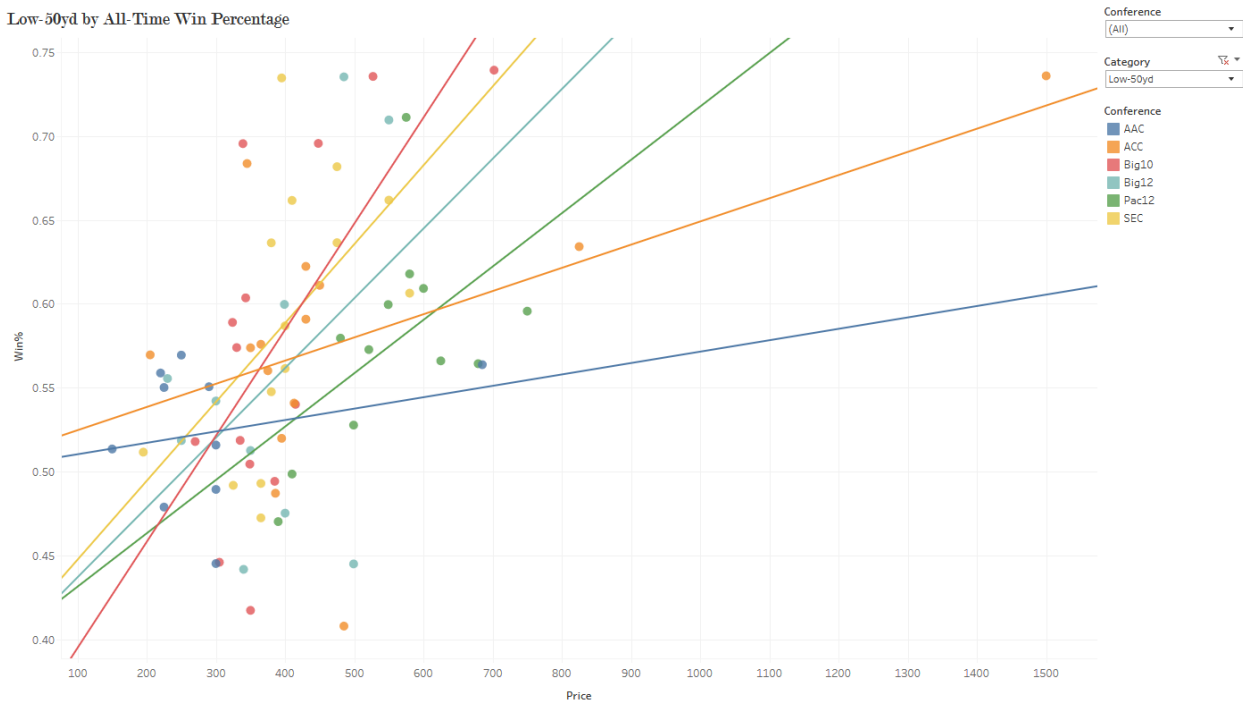
Low-Don-EZ & Low-Endzone by Conference



- When looking at the price range of endzone tickets we find that the Big10 has the highest average ticket price, as well as the widest range of prices
- With donation requirements taken into account, the ACC has the highest total cost for Endzone seating

How does 50-yard line prices relate to historic win percentage?

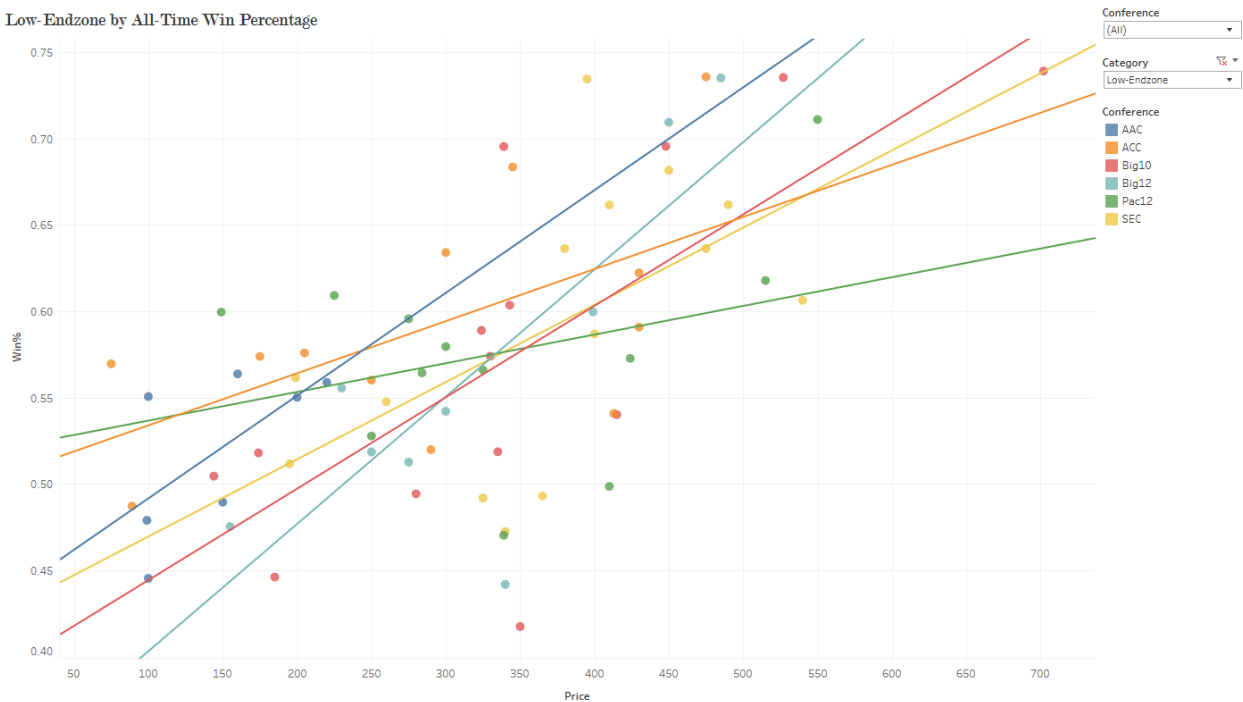
Low-50yd by All-Time Win Percentage



- A positive correlation between win percentage and the price of 50 yd line seats is evident for all conferences
- This would indicate that winning does affect the prices for that area of the stadium.
- While the correlation is positive, it is important to note that the correlation is generally weak. For example, the largest correlation (Big10) is only .45, or moderate.

How do endzone ticket prices relate to historic win percentage?

Low-Endzone by All-Time Win Percentage



- The results for Endzone seating are similar to those for the 50 yd line, all conferences have a positive correlation.
- It is important to note that the correlations are higher in this case, with the largest one being a strong correlation, (0.63 for the Big 12) and pricing at the end zone.
- This would mean that win percentages, much like stadium capacity (mentioned above), play a larger role in determining pricing for the End zone sections of a football stadium, than they do for the 50 yd line seats.

Interactive Visuals

All visuals shown above can be found by visiting the Tableau Dashboard: [HERE](#)

These visuals are fully interactive. Feel free to apply filters on the right-hand side of the screen to create different visuals or to conduct further analysis as needed. The tabs from left to right are:

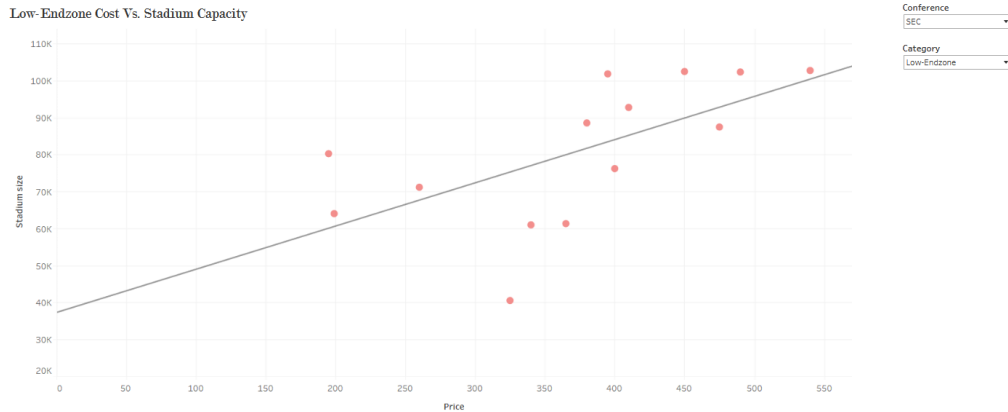
- Price Tiers Vs. Cap
- Price Vs. Stadium Cap
- Ticket Price Ranges
- Price by Win Percentage
- Price by Win Percentage (No TL) – same a previous tab, but with no trend lines fit to the data
- Stadium Capacity Range
- Price Tiers

When using the filter keep these definitions in mind:

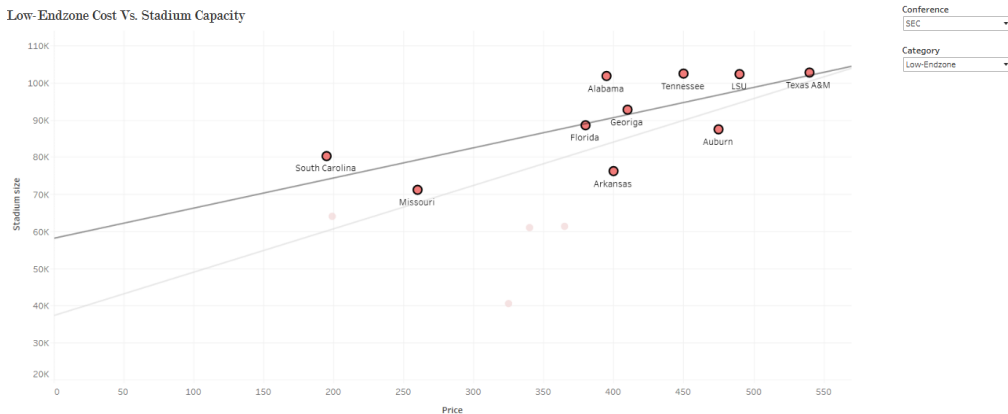
- If a Category filter has “Don” in the name, that means it represents the donation data for the price level. For example: Low-Don-50 should be read as Lower level, donation requirement for 50 yd line seats.
 - If “Don” is not mentioned, that is referencing the ticket cost alone. For example, Low-50yd, is the price for just the tickets at 50 yd line.
- If you click multiple filters, the cost of both will be added up and that is what the visual will show.
 - For example if you are on ‘Price by Win Percentage’ tab, and you select ‘Big10’ as the Conference filter, and if you click on Low-Don-50 ,and Low-50yd you will see the total ticket cost (tickets + per seat donation) for the 50 yd seat for the Big 10 conference schools (plotted on the horizontal axis) vs the school’s all-time win percentage (plotted the vertical axis)
 - If you hover over the Red line you can see the linear equation for this ‘best fit line’ and the R-squared and p-value of this fit

For some of that data, you can edit the trend lines by holding and dragging the mouse over a set of data points and that will change the trendline.

- Starting with the Price vs. Stadium Cap tab, and setting the Conference setting to 'SEC', and the setting the Category setting to Low-Endzone
- You should see:



- Now if you only want to examine the trend for stadium's larger than 60k, you can click and drag the mouse over those data points, and you should see this:



- Where you can see the schools that have been highlighted, and a new trend line, above the old one

If you have any questions about the data presented or the using the analysis, please reach out to Andrew Gillingham (Data Analytics Manager) at: andrew.gillingham@imglearfield.com

Primer: What is correlation?

A correlation is way to talk about whether two variables are related. Let's start with some real-life examples. For one, let's look at the number of carbs you have before a marathon and whether you finish. We would expect that there is some positive relationship between these two things, that is, if you carb-load before a marathon you are more likely to finish. Now, what about carb-loading, getting a good-night's sleep, and training for a few months, we would think that this combination of that physical condition would be strongly related to whether someone finishes the marathon. How about the relationship between how tightly the marathoner tied their shoes to whether they finished? I know it seems pretty random, right? That's sort of the point. You would not expect there to be a strong, or any relationship between those two things, and so we would say they are weakly correlated.

Now in those examples, we were only talking about positive correlations, one thing seems to help the other (either strongly or weakly). There is also negative correlation. To keep on this contrived analogy, let's think about the relationship between the number of "adult beverages" ingested the night before the marathon and whether someone finishes. In this example, as the number of beverages goes up it is fair to assume that the runner will not finish the race. Therefore, these two things are negatively (weakly or strongly) correlated. One is hurting, not helping the other.

This marathon example hopefully makes intuitive sense, but it is helpful to add a quantitative measure to these weak and strong labels for correlation. It is possible to mathematically quantify how related two variables are to each other by defining something called the correlation coefficient. The value of what constitutes a strong correlation coefficient can vary by industry but for the analysis presented above:

- The variables have NO correlation if the value of the coefficient is 0
- The variables are totally correlated if the value of the coefficient is 1
- The variables are weakly correlated if the value of the coefficient is less than 0.2
- The variables are strongly correlated if the value of the coefficient is greater than 0.6
- The variables are moderately correlated if the value of the coefficient is between 0.2 and 0.6

Correlation is NOT causation

Something that you've surely heard, is that "correlation is not causation". This is an important, if not overused, warning about what is safe to assume with correlated data. It is a reminder that just because you have two variables that have strong correlation, it does NOT mean that one variable is causing the other. From the marathon example above, you can't say that eating carbs is going to CAUSE you to win marathons. If it did, I'd be in the Olympics.