December 2020

Understanding Collegiate Football Ticket Prices

Ken Miller
kmiller2316@gmail.com

Kenny Miller and Dr. Wei Zheng
Melton Scholar’s Program Fall 2019 – Spring 2020
Understanding Collegiate Football Ticket Prices and Sales

Executive Summary

Problem

Deciding how to spend your money is one of the most important decisions anyone makes, and the purpose of this paper is to better understand the purchase behaviors of customers on the primary and secondary markets for college football ticket sales.

Data and Process

The data used in this study has been provided through a partnership with a college athletic department (AD). We have been given proprietary ticket sales data for the 2015 through 2019 seasons. The data is a combination of primary market (tickets sold directly by the AD) sales and sales on the secondary market for the 2017 through 2019 seasons. The secondary market sales data was provided through a partnership between the AD and VividSeats. There is limited previous research in this field, specifically among choice analysis of ticket sales for college football, with one other paper being the main guidance we used. This research replicates the previous research method but currently is unable to confirm the conclusions of the study due to limitations with the data provided.

Conclusions

- First, any ticket in the lower bowl of the stadium is more favorable than the best seat in the upper bowl.
- The east direction of the stadium is the least favorable section among the alumni, faculty, and donor season ticket segments.

- Team performance plays a major role in the difference between the resale price and original ticket price.
- Finally, the prices of the secondary market are driven by seat location (price zone) and how attractive the game is.

Next Steps

The next step of this research will be to confirm the selling prices of each price zone through the years with the AD, to ensure the models are accurate. Once this is completed, the next questions will be how a customer's choice varies in the primary and then the secondary resale market on a single game basis. Finally, how can we use this choice analysis to better price tickets for games.

Problem Statement

In developing this research study, we followed many frameworks to try and understand the choices made by consumers when they purchase football tickets, which is a “is this worth my time and money?” question. We identified and developed a choice model framework to understand customers’, in our case this was football fans, purchase behaviors on the primary and secondary markets. Specifically, we reviewed the purchase decisions made by distinct segments of season ticket holders, which is a major source of all college football revenue and looked to understand the differences in the decisions made by these segments.

Methodology

Using ticket sales data provided by a college athletic department (AD) and their ticketing partnerships, we analyzed purchase decisions and trends among their fans on both the primary and secondary markets. Our main resources are the primary market sales, tickets sold by the AD’s ticketing department, directly to fans for the 2015 through 2019 seasons and the secondary market sales from the AD’s partner VividSeats between fans for the 2017 through 2019 seasons. The datasets contain a total of 3.29 million tickets that were sold across the five seasons we reviewed. Both datasets contain the same variables which allowed for easy merging of the seasons and comparisons between the primary and secondary market prices for those seasons. A full description of the variables is in Appendix A. Ticket sales were marked with an order number and a few order specific variables but otherwise customer purchase profiles were anonymous for both markets. Beyond the data provided, other variables were generated from the data provided. These variables were used in our understanding of the primary market and secondary markets, both individually and merged.

For this analysis we focused mainly on the primary market season ticket sales, with the opportunities to continue this analysis to look at single game primary sales and resale decisions on the secondary market. With this focus in mind we utilized multiple packages in R and Tableau to review the data. We created a variable importance chart to identify the choice variable we would review, using the Boruta package. We utilized Tableau to generate visualizations to understand the nuances of each market. Finally, to answer the main question regarding the choice model of the season ticket holders across the 2015-2019 seasons, we developed a multinomial logit choice model using the mlogit package in R.

Analysis and Results

The first analysis of the ticket markets started by understanding the price distributions of the primary market. Table 1 contains the summary statistics of the price and resale price columns of the primary and secondary market data, respectively. The distributions are similar, but the secondary market has a large skew to the high-end prices, which will help us better understand how customer's value each game and how their willingness to pay fluctuates based on game and ticket attributes. The minimum of zero in the primary markets is due to some tickets that are given away by the athletic department for recruits and potentially their families, families of coaches, or any potential partners, which can be compensated by the AD or team at their discretion.

	Minimum	25% Percentile	Median	Mean	75% Percentile	Maximum
Primary Market	$0.00	$35.00	$48.00	$51.18	$70.00	$155.00
Secondary Market	$10.00	$35.00	$50.00	$55.26	$60.00	$155.00

Table 1: Summary Statistics of Price (Primary) and Resale Price (Secondary) Markets

Primary Market

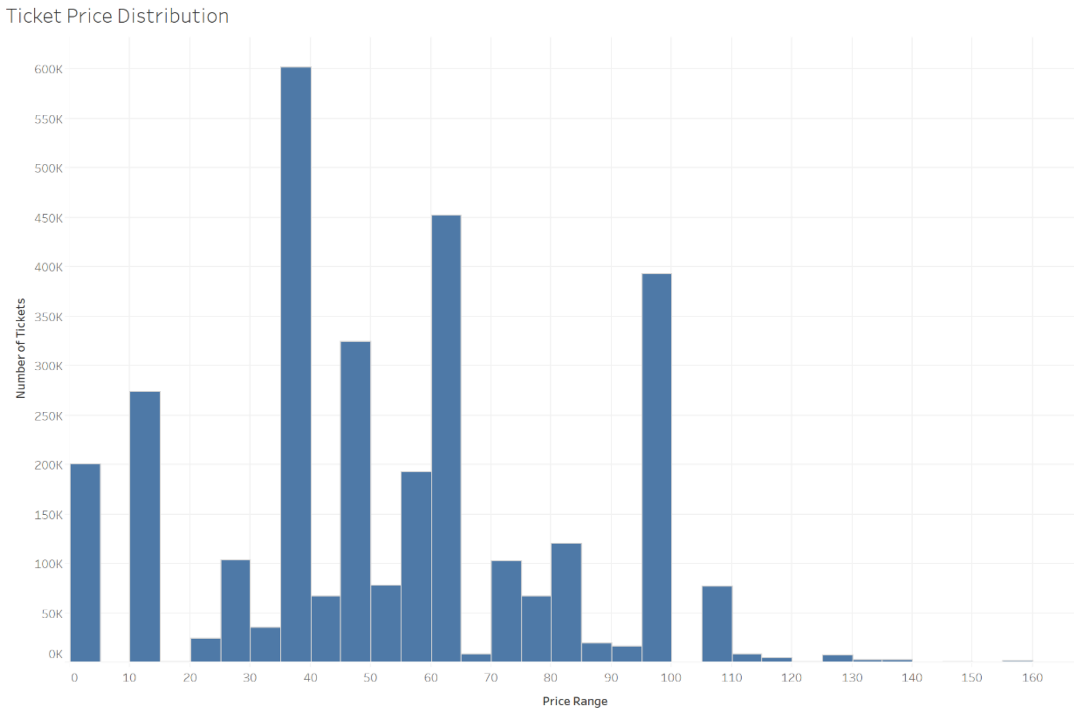

Figure 1: Histogram Distribution of Primary Market Ticket Prices

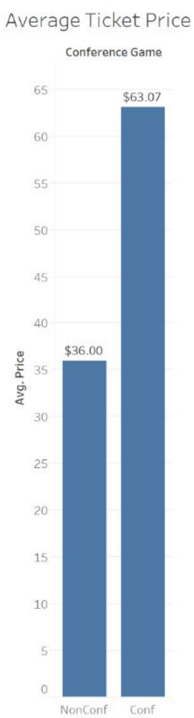

Figure 2: Average Price Conference Game vs Non-Conference

Our analysis of the primary market focused on understanding what causes prices to change from game to game and how prices fluctuate based on certain factors being looked at. In order to understand all these factors, we need to see the primary markets' numbers in action. Figure 1 shows the overall distribution of primary market ticket prices. This distribution is roughly normal, with a slight right-side skew. The large number of tickets in the $0-$5 range is due to tickets allotted to recruits, their families, and families of players and coaches.

Average Price by Home Team Record

	Wins							
Losses	0	1	2	3	4	5	6	7
0	$36.85	$54.90	$30.47	$80.11				
1	$42.66	$30.32	$56.83	$81.06		$79.58		
2	$30.39		$52.64	$49.02				
3		$81.92	$70.46	$84.54		$31.20	$51.72	$48.40
4		$52.05			$60.04	$29.75		$42.17
5			$50.42	$30.37	$50.70	$47.20	$49.50	
6					$83.90			
7					$53.51			

Figure 3: Average Price by Home Team Record

Figures 2 and 3 show how two important performance and game factors play a role in average price. Figure 2 shows average price by conference or non-conference game, whereas Figure 3 shows how price changes with the home team's record coming into the game, broken down via a matrix of wins and losses. Not all possible combinations of wins and losses are seen in this dataset which is why some boxes are empty. Figure 4 continues the trend of looking into factors that affect price, this time we see the distribution of price by game. The distribution is

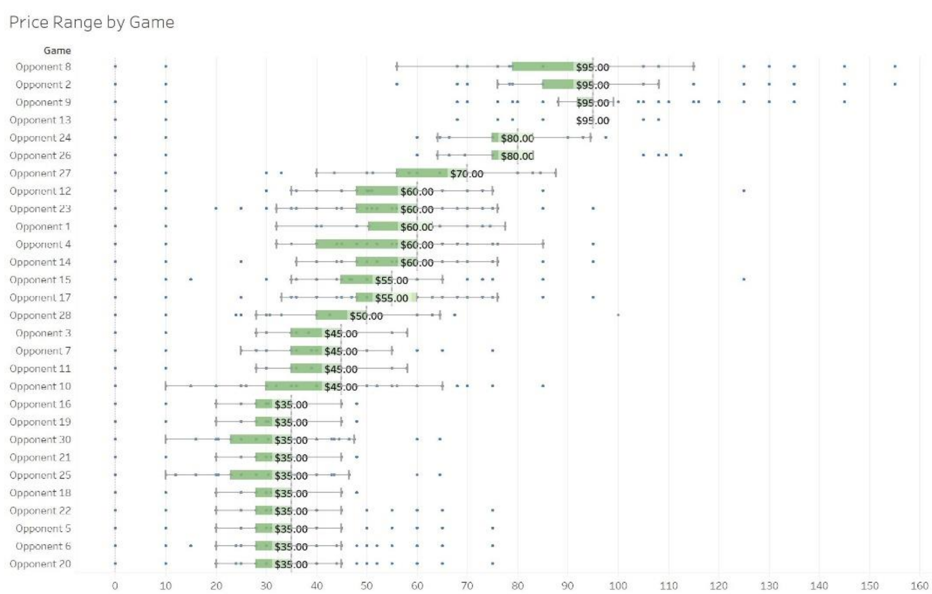

Figure 4: Price Boxplot Distribution by Game/Opponent

done by using boxplots with the median ticket price printed on the chart as well. In this figure we see how opponent affects the ticket price on average and the overall range. Also, we see that there is a variable pricing structure used for games based on the opponent. Opponent names are provided with aliases.

Secondary Market

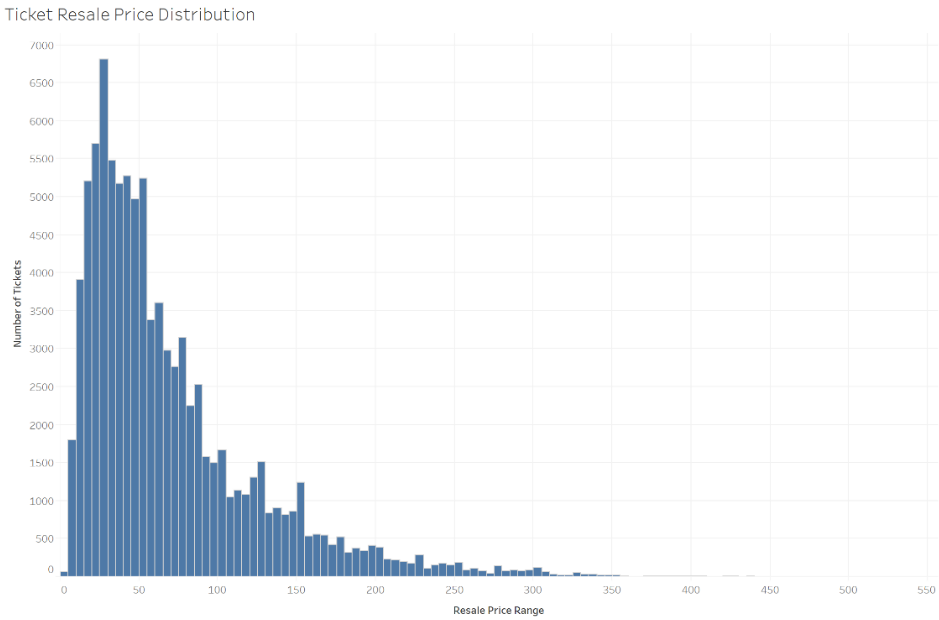

Figure 5: Histogram of Resale Prices on Secondary Market

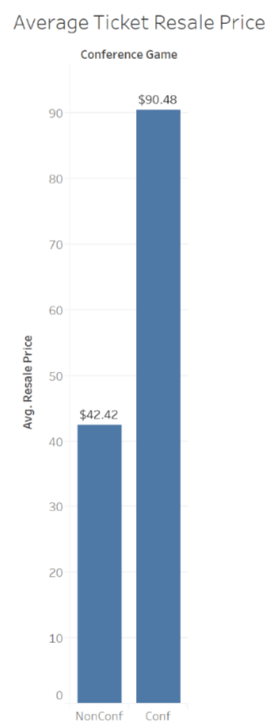

Figure 6: Average Resale Price Conference vs Non-Conference Game

For the secondary market and understanding fan spending on games, we took a similar approach as with the primary market. In our secondary market analysis, we can see and understand more of what customer’s value in games and what they are willing to pay for each game. This can be seen immediately through the right-skew of the histogram of resale prices of tickets sold on the secondary market, shown in Figure 5, that fans are willing to pay incredibly large sums of money if they deem the game to be worth it. Figures 6 and 7 also show this conclusion of paying large amounts in the game is “worth it”, where Figure 6 shows ticket price distributions by game, opponent aliases are matched with those provided in Figure 4, and Figure 7 shows average price for a game if it is a conference versus non-conference game. Figure 8 also shows average prices of tickets by the home team's

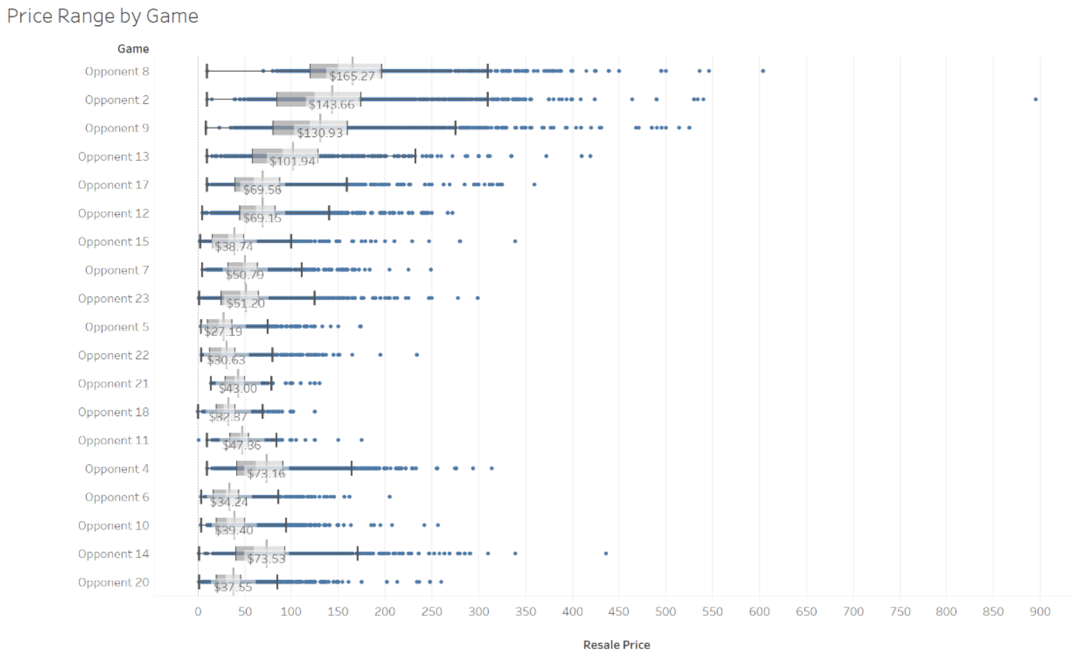

Figure 7: Boxplot Distribution of Resale Price by Game/Opponent

record, in this we see that as the home team's record improves the perceived customer value increases with prices as well.

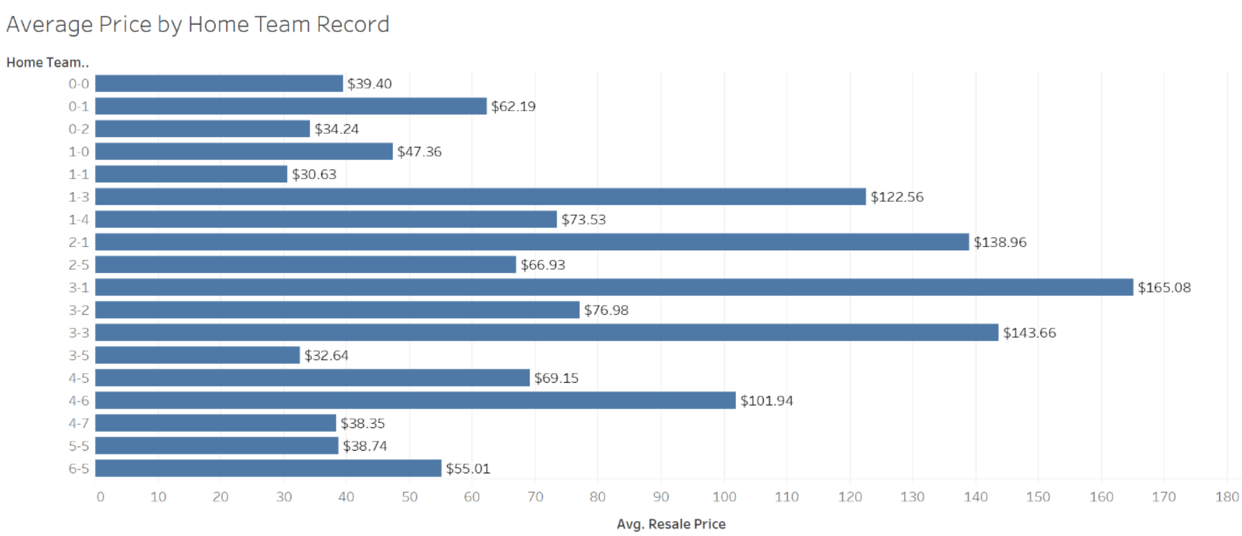

Figure 8: Average Resale Price by Home Team's Record

Merged Market Data

In this merged market data frame, it is a collection of all secondary market seats and their original primary purchase. That is if an individual bought a seat from the Athletic Department and then had to sell them on VividSeats because they couldn't go to the game, the seat and respective sales information would be included in this merged data.

The merged data shows similar factors as the primary and secondary data, but we can see this through comparisons quickly. Figures 9 and 10 show comparisons of the average price and average resale price by game and opponent, Figure 9, and by the home team's Record, Figure 10.

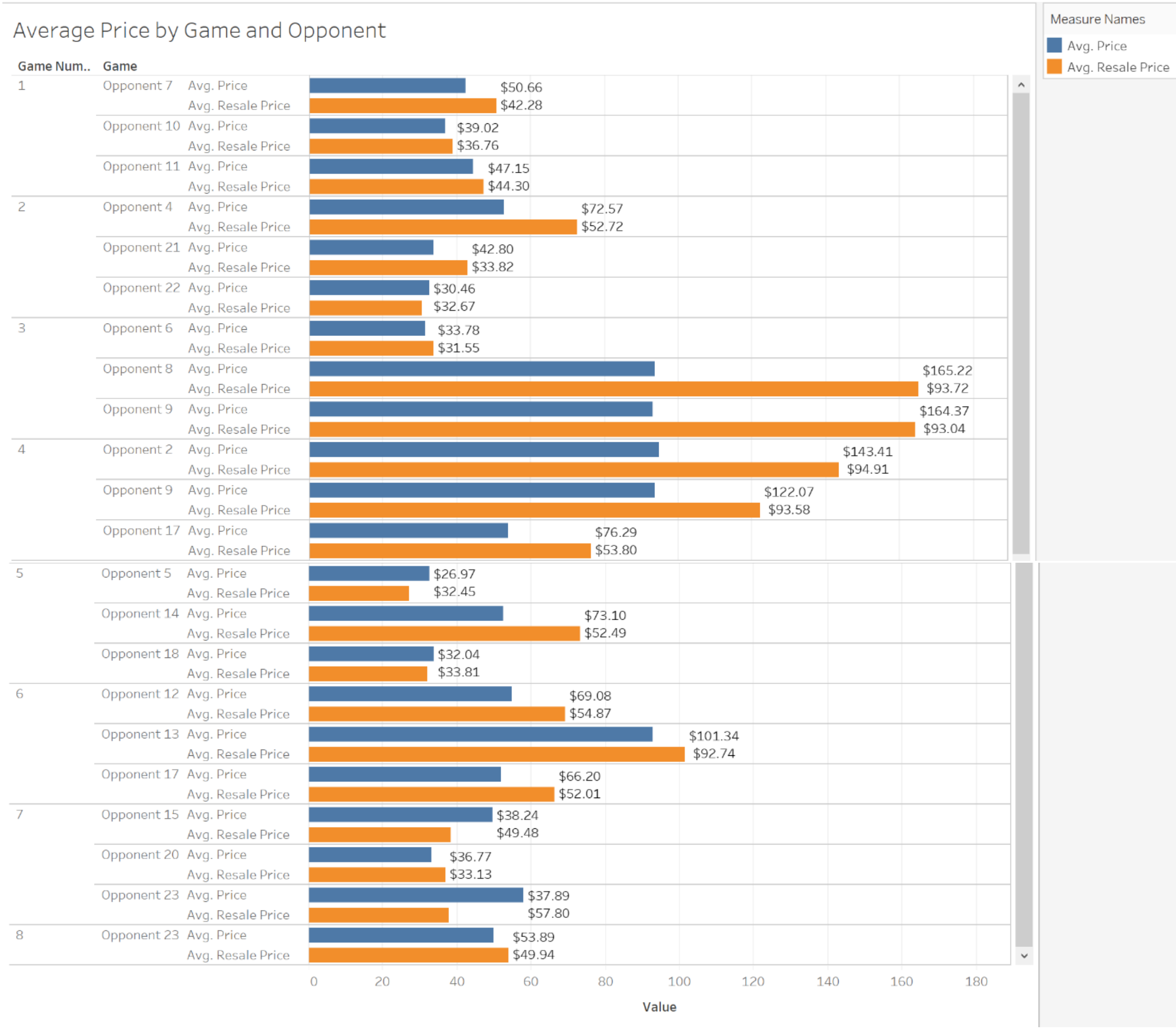

Figure 9: Merged Price Comparison by Game and Opponent

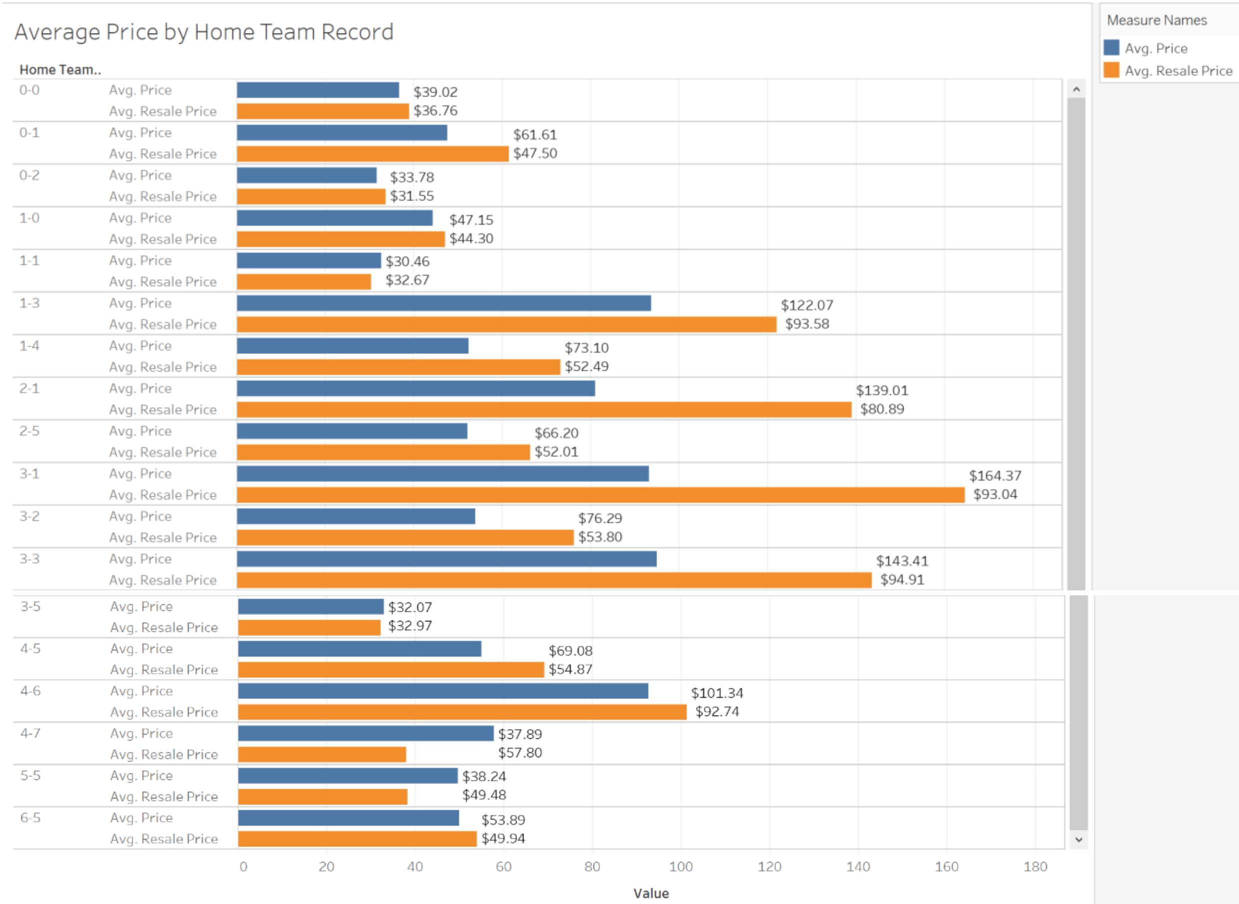

Figure 10: Merged Comparison of Prices by Home Team's Record

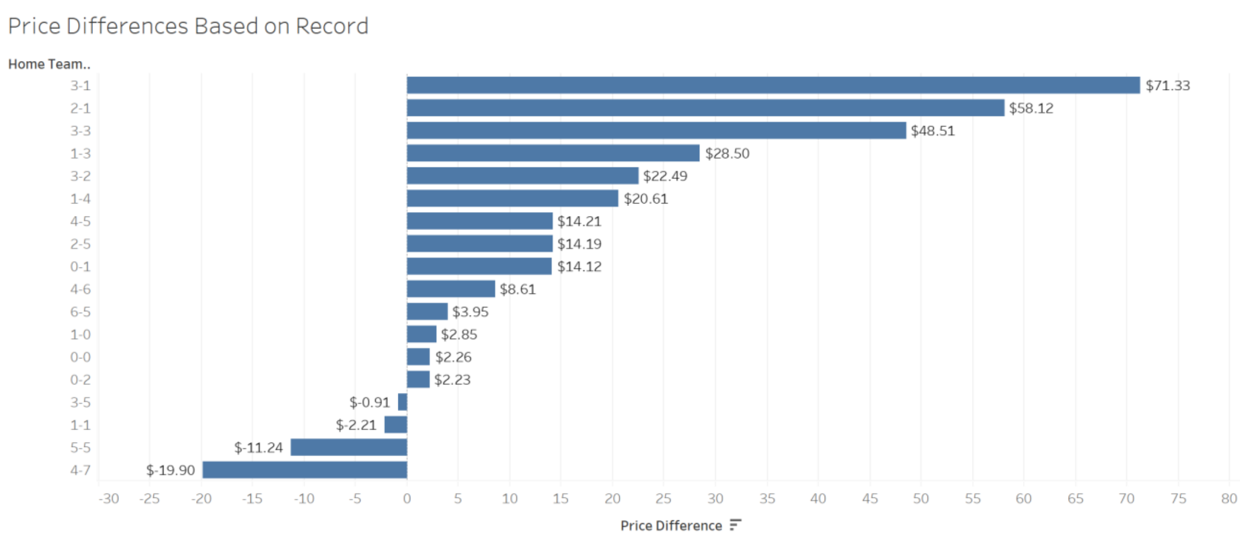

Figure 11: Resale and Primary Price Differences based on Home Team's Record

Figure 11 continues to show the relationship between the home team's success and price differences by showing the average resale price minus the average price by record.

Using the Boruta package in R we were also able to identify and rank the variable importance for our model development. In Figure 12 you will find the plot output from Boruta showcasing each variable's importance in predicting the ratio of resale price over price, using the following equation:

$$\mathit{log\ (resale\_price/original\_price)} = \beta_0 + \beta_1\mathit{section} + \beta_2\mathit{row} + \beta_3\mathit{seat} + \beta_4\mathit{order\_city} + \beta_5\mathit{order\_state} + \beta_6\mathit{order\_zip} + \beta_7\mathit{type} + \beta_8\mathit{game\_date\_time} + \beta_9\mathit{opponent\_rank} + \beta_{10}\mathit{conference\_game} + \beta_{11}\mathit{home\_record} + \beta_{12}\mathit{tv} + \beta_{13}\mathit{game\_number} + \beta_{14}\mathit{game\_type} + \beta_{15}\mathit{home\_team\_wins} + \beta_{16}\mathit{home\_team\_losses} + \beta_{17}\mathit{home\_winning{+}percentage} + \beta_{18}\mathit{home\_foregiveness\_wp} + \beta_{19}\mathit{opponent} + \beta_{20}\mathit{ticket\_in\_hand} + \beta_{21}\mathit{streak} + \beta_{22}\mathit{bs}$$

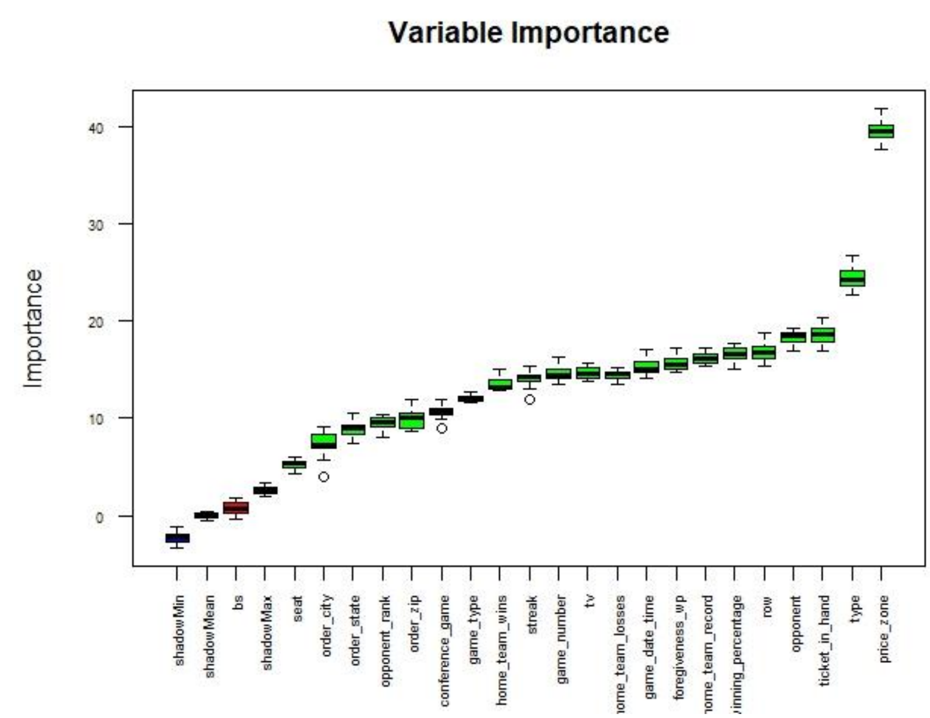

Figure 12: Ranked Variable Importance, using Boruta package, for predicting logged resale price and price relationship

The Boruta package and ranking allowed us to identify the main choice fans make when they purchase is dependent on the “price zone”, how sections are grouped by the AD, they would be sitting in. This provided us with the main choice our customers make and would be the basis for the logit model we develop. The stadium layout and available price zones is shown in Figure 13.

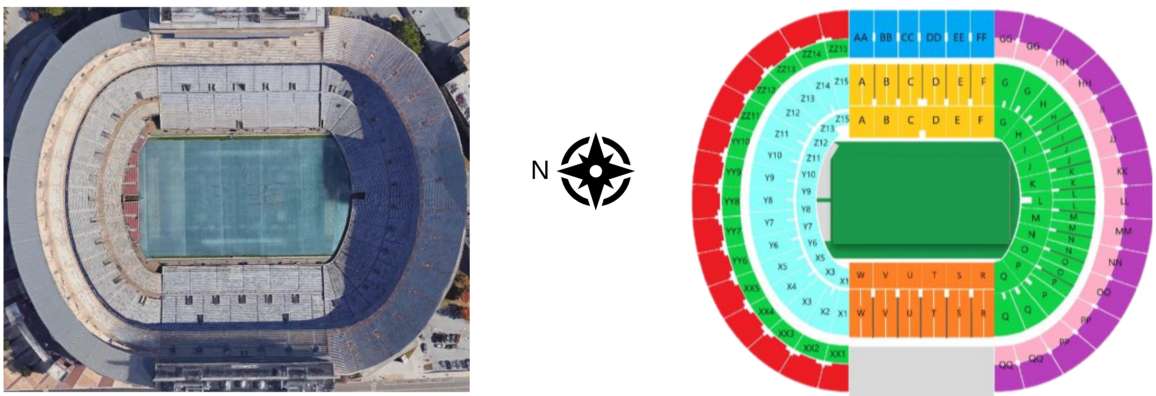

Figure 13: Stadium Price Zone Layouts for the Partnered AD, Included with Compass Rose to Showcase Seat Directions

Logit Model Development

Using the mlogit package in R, we were able develop a multinomial logit choice model to better understand the choice decisions of the main season ticket holder segments. A multinomial logit model is used when the person of interest, in our case a fan, has multiple options in front of them, different price zones, when and they can only choose one. It is best utilized to describe how the different variables interact to affect an individual’s choice in a decision. To run our logit model, we needed to develop a formula, in out case *price zone* = *direction* + *distance* + *price* + *0* and correct the data frame to be in long form and not wide. When a data frame is in long form, it means everyone has an opportunity for each choice and a true/false is used to show which decision was made. After these adjustments we can run the logit model and see how each variable impacts the choice made by a buyer.

We determined the main season ticket holder segments to be alumni, donor, faculty, and other. This decision was made via a table and understanding which segments are most likely to purchase tickets for an entire season. With price zone as our choice for these segments, we analyzed the decision using direction (sitting in the North, South, East, or West sections of the stadium), distance (lower versus upper bowl), and the average purchase price for the price zone. The average price was calculated using by price zone and year, with any compensated tickets for that zone excluded in the calculation. Table 2 shows the results of the multinomial logit models across the four segments for season ticket purchase decisions across the 2015 – 2019 seasons.

2015 – 2019 Parameter Estimates for Season Ticket Segments					
Segment	Distance	DirectionN	DirectionS	DirectionW	Price
Alumni	-2.22 (0.03) **	0.95 (0.03) **	0.43 (0.03) **	0.38 (0.03) **	0.07 (0.01) **
Donor	-1.43 (0.01) **	0.72 (0.02) **	0.02 (0.02)	0.46 (0.02) **	0.09 (0.00) **
Faculty	-2.16 (0.03) **	1.51 (0.05) **	1.53 (0.05) **	0.77 (0.05) **	0.01 (0.01)
Other	-1.00 (0.01) **	0.90 (0.01) **	0.82 (0.01) **	-0.23 (0.02) **	-0.01 (0.00) **

Table 2: Parameter Estimates and Standard Errors for Logit Models. ** show the estimate is significant at the 99.7% confidence level

In conclusion, the results of our model show that any seat in the lower bowl, regardless of segment, is more favorable than the best seat in the upper bowl. The Alumni, Donor, and Faculty segments all view the East direction seats as the least favorable whereas the other segment views West as the least favorable. Price is an interesting case we see that demand and favorability increases as the price increase. There are two possible cases for this: 1) when the price is higher the seat is placed in a more favorable area (price zone, distance, direction) and the fans are more likely to make this purchase regardless of the price increase, or 2) demand has decreased in recent years due to games on TV, poor team performance and this has caused the AD to lower prices from 2015 – 2019, which has happened.

Additional Recommendations

As discussed above our logit model only looks at season ticket purchase decisions across the 2015 through 2019 seasons. While these models were insightful and aided in better understanding the choice decisions of our customers, it would worth while to review and develop models on single-game ticket purchases and develop the same model on the secondary market to see the sensitivity of the decisions being made. Once these models have been developed and analyzed to see trends in the single-game purchase decision a pricing model should be developed to better price each game and maximize stadium capacity, therefore maximizing revenues for the Athletic Department.

Appendix A

Variable Descriptions

Variable Name	*Description*
Id	Unique ticket purchase identifier
Order Number	Unique identifier for entire ticket order
Game	Opponent for game
Performance code	Game identifier
Price/Resale Price	Ticket Purchase Price
Section	Section ticket is in
Row	Ticket row number
Seat	Ticket seat number
Order city	Billing city of customer
Order state	Billing state of customer
Order zip	Billing zip of customer
Type	How ticket was purchased
Order date time	When ticket order was placed
Game date time	Game kickoff time
Opponent city	City opponent is from
Opponent state	State opponent is from
Opponent rank	AP and Coaches poll ranking of opponent week of game
Conference game	Is the game in conference
Home Team record	AD’s team record at start of game
TV	Channel the game is broadcast on