DATA VISUALIZATION ANALYSIS:

SCHOOL OF DANCE BOX OFFICE SALES

By

TALIA JENSEN HINTERMEISTER

A Thesis Submitted to The Honors College

In Partial Fulfillment of the Bachelor's degree
With Honors in

Management Information Systems

THE UNIVERSITY OF ARIZONA

M A Y 2 0 2 0

Approved by:

Dr. Lusi Yang
Department of Management Information Systems

ABSTRACT

The box office currently completes a variety of analyses on season ticket subscriber data, so this project aims to analyze different aspects of the box office sales from 2013 to 2020 to provide insights and findings to improve season ticket subscriptions and revenue for the University of Arizona School of Dance. The data comes from a dataset of over 34,000 records looking at accounts that purchased tickets to dance performances over the past seven years. There is information on ticket buyers; season ticket subscribers; number of seats sold; and the paid amount to the box office. I used several visualizations to display the data including bar charts, line graphs, pie charts, and maps. They all tell a story about the data, and I have included explanations, insights and rationale where necessary. This project evolved over time as I met with stakeholders, created project learning objectives, collected related data, analyzed and then organized it in Excel. The final step was to create visualizations in Tableau to review the data in a more digestible way. From there, I drew conclusions and made recommendations for the University of Arizona School of Dance to expand geographically and attract other ticket buyer groups among other improvements. I gained data analysis and visualization skills throughout this project and will continue to use these in my future career.

INTRODUCTION

Background

When asked which degree I wanted to earn honors under, I knew I wanted to do MIS. I wanted to dive deeper into what I have learned during my classes and internships, but I also wanted to combine it with my love for dance and the arts. Because I am a double degree student, I wanted to use both my majors to complete my thesis and do research in both areas. I thought it would be intriguing to use my MIS skills and interest in data visualization to analyze ticket sales to dance performances. This would benefit the School of Dance by providing them an analysis of who is buying tickets and how to target different geographical areas to improve ticket sales.

First, I met with Dr. Neumann, who is involved in both colleges (Dance and MIS) on campus. He is the "Introduction to MIS" professor and often attends dance performances, so I thought he would be a good starting point. He recommended some ideas on how to approach the project and ways to go about contacting the right people. He and Dr. Currim provided an introduction to a MIS professor that specializes in data visualization, Dr. Lusi Yang. He advised that I meet with Whitney Herr-Buchholz and see what the dance program is looking to improve and what questions they are asking about ticket sales, student attraction/recruitment, or demographics of current students.

After our discussion, it seemed most feasible and interesting to look at ticket sales for dance performances that run from 2013 – 2020, so we started looking at what the dance department's learning aspirations were and how I could use data visualizations to organize this information. We examined the data that would be interesting for the directors of the program, the dance advisory board, and the marketing/social media teams.

With my thesis, I hope to improve how the dance program analyzes the box office tickets purchased for UA Dance performances to enhance ticket sales. I wanted to gain experience in data visualization tools and box office analysis. This will be beneficial as I look into different career options especially jobs dealing with business intelligence. Data visualization has always been interesting to me, and this seemed like a great opportunity to take it a step further to dive into a passion of mine, dance, and benefit the program. I also wanted to gain experience using Tableau, and the data available from the box office made this the perfect chance. The box office analysis will be beneficial if I am dancing for a company and want to complete an analysis for them or make recommendations for future improvements.

Research Questions

During the early phases of my work, I brainstormed areas where I wanted to focus. From there, I used this list of questions to assess how I would obtain this information. My thesis seeks to answer the following questions among many others:

- How many seats are sold each year?
    - How many seats for annual shows?
- How much revenue does the box office make each year?
    - How much revenue for annual shows?
- When is the most popular time to buy tickets?

- What is the average number of tickets purchased per order?
- What age groups make up our ticket sales?
    - How can we target other age groups?
- Where do our ticket buyers come from?
    - How can we target other areas?
- What part of the week is the most common to attend a performance?
- Where do our season ticket subscribers come from?
- How many shows do ticket buyers attend before becoming a season ticket subscriber?
    - How can we attract new season ticket subscribers?

Research Process

I met with several professors to get background information before diving into my project. This helped me narrow down exactly what I wanted to find out and display for my thesis. In the fall, I met with Whitney, dance office business manager, to see what types of reports they already used and where there were gaps that I could fill. She noted that they already did a lot of work with season ticket subscriptions because they make up a good portion of their sales. I wasn't sure what would be possible in terms of the scale of data the box office has and is able to release. We brainstormed the above research questions, so I could bring them to Charles, the box office manager, to see what would be possible.

Charles pulled up the box office data in Ticketmaster, and we looked at the different reports that were automatically created. Ticketmaster holds an abundance of information about the people who purchase tickets through the College of Fine Arts Box Office. We met a couple times to brainstorm exactly what it was I wanted and what would be the easiest way to obtain the data while following privacy protocols. I was most interested in the dataset with over 34,000 records that had a variety of attributes from number of seats purchased to the purchaser's zip code and whether or not they were a season ticket subscriber. Following my meetings with Charles, I met with Whitney to find out the best method for getting the data in my hands and how we might proceed going forward.

METHODOLOGY

Data Collection

Charles gave me access to the Ticketmaster data using a computer located near his office and the ability to download the information that I needed. I went into the account and decided to pull the dataset with over 34,000 records and 41 attributes that described the purchaser and the dance seats purchased from 2013 - 2020. Whitney decided that it would be best if the dataset was emailed to her from Charles, so that it would follow privacy protocols. We eliminated the columns containing data like name and address, so purchaser data remained anonymous. We were able to keep zip code, so I could use it for location and map data visualizations. From there, Whitney transferred the data to my computer where I was able to analyze and organize it in ways that would be helpful before importing it to Tableau. Once I am finished with this project, I will delete the Excel data file from my computer and USB drive to ensure compliance with University privacy requirements.

Data Preparation

There was quite a bit of data cleanup required due to the way the data was stored in Ticketmaster. The first column combined the year, arts genre, performance, and show number into one cell. A sample cell is: 15DJAZZ3 meaning that this was during 2014-2015 school year, in the dance genre, the Jazz in AZ performance and show number three. I had to break these out into four separate columns where Tableau could digest it more easily. Since the data came from shows as early as 2013, there has been variation in the way the data was recorded. For example, ticket type had over 20 options when it could be summarized into five ticket buyer options. Most of these were named "On stage student" for a specific show rather than just "student." I cleaned these up so that the ticket buyer charts would demonstrate the five main groups rather than specifics over the years.

The data held the ticket buyers' zip code information, but it did not include the city or state that they were located because we got rid of the address attribute. In Tableau, the tool uses zip code to populate the data onto a map; but unfortunately, it cannot summarize data on the city and state level based off of the zip code. After reading and researching the Tableau help page, I found a site that listed states with their corresponding minimum and maximum zip codes. I pasted this data into a sheet of my Excel Workbook and used lookup functions to reveal the state that contained each zip code. I attached the zip code/state data that I used to add state to the original box office dataset as Figure 13 in my Appendix. In this figure, you can see one state name is "kc96 DataMO" that I renamed to be Missouri after verifying that those zip codes corresponded to those listed online. This enabled me to use the state of each ticket buyer in Tableau and create a map with number of seats sold grouped by state.

Whitney had mentioned that she wanted to know how many shows people attended before becoming a season ticket subscriber. This was one of the more challenging analyses completed in this whole process. I discovered that it would be easiest to organize the data in Excel using functions and formulas first rather than attempting to organize it in Tableau. I created a column to reveal if the person buying tickets was a season ticket subscriber or not. Then, I made another column that would tell us what number occurrence that the account ID showed up. For example, if it was the first time that account ID purchased dance tickets, a one would appear in that cell, and if it was the third time this account ID purchased dance tickets, a three would appear and so on. This would be helpful in knowing how many times this purchaser has been to a show and also whether or not they were a season ticket subscriber each time. From there, I copied and pasted the values, so there would be no formulas and it would be solely text. This enabled me to sort and filter without having the cells change values. Next, I filtered the season ticket subscriber column so that only people with subscriptions showed. Finally, I removed duplicates based on the account ID, so it would show the first time the account ID appeared in the system with a subscription. The occurrence column became valuable because now I could see how many times this purchaser had went to School of Dance shows before becoming a season ticket subscriber.

RESULTS

After experimenting with Tableau display functionalities, the following results emerged. The outcomes revealed a variety of information and findings including, but not limited to, ticket buyer patterns; number of tickets sold of each performance and each year; amount of money made by the box office; and the amount of times people came to a show before buying a season ticket subscription.

Seats Sold and Amount Paid to Box Office

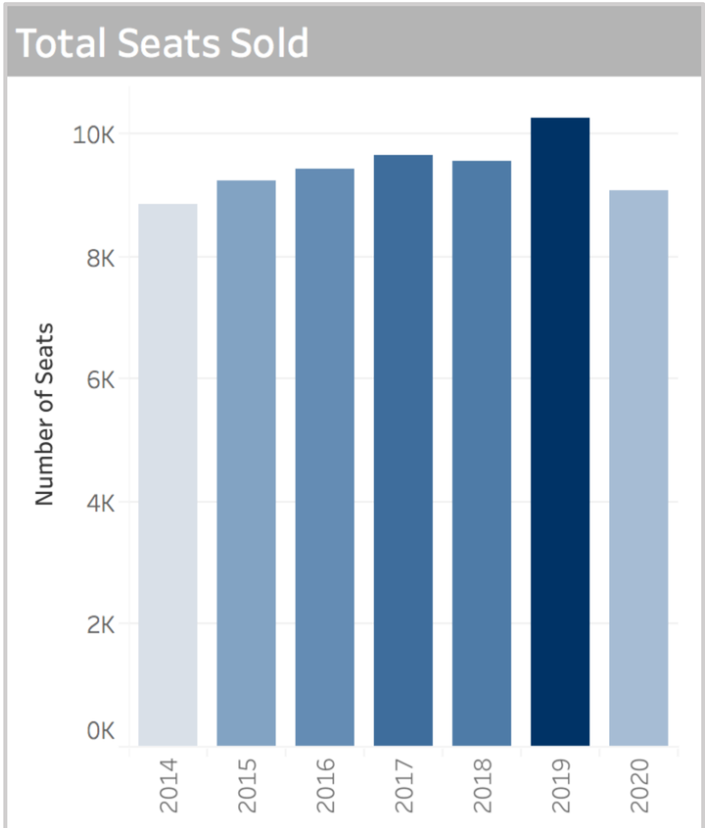


Figure 1. Total seats sold each year

Figure 1 displayed above shows the difference between seats sold each year. Seats refer to the tickets that people purchased and what seat they sat in while watching the dance performance. The 2018 - 2019 season (labeled as 2019) had the highest number of seats sold due to its bar exceeding the 10K line. It is important to note that 2019 was higher than 2020 due to the 2020 season being shortened due to COVID-19 and the University closing in March and moving to online learning. As mentioned before, two shows were cancelled, and therefore, not included, resulting in the lower value in 2020. The data reveals an increasing trend over the years. However, there was a slight dip in 2018 where it is unknown what the cause of decline was since there are many potential qualitative factors that may have contributed to the slight decline. I imagine that 2020 would continue to follow an upward trend if the season hadn't ended early. Let's take a look at what causes these increases over the years: How has the number of seats increased over the years? The theater has not gotten bigger (in terms of space), but there are a number of other ways that have contributed to this increasing trend.

First, it is predicted that word has gotten out about the talent of the dancers and the high quality of the performances. People talk about it or they might see something on social media as that has continued to emerge, and this influences other people to come. Another potential factor is the quality of the choreography and performances of popular master works. Additionally, the increase could be related to the program increasing in size and more families attending the performances. Another reason for the increase in seats sold each year is the added performances. In recent years, the school of dance has added a Sunday night performance to the shows that are especially popular. This adds an entire theater of seats to the number of seats sold for that year. These factors partially explain the increase in seats sold over the years.

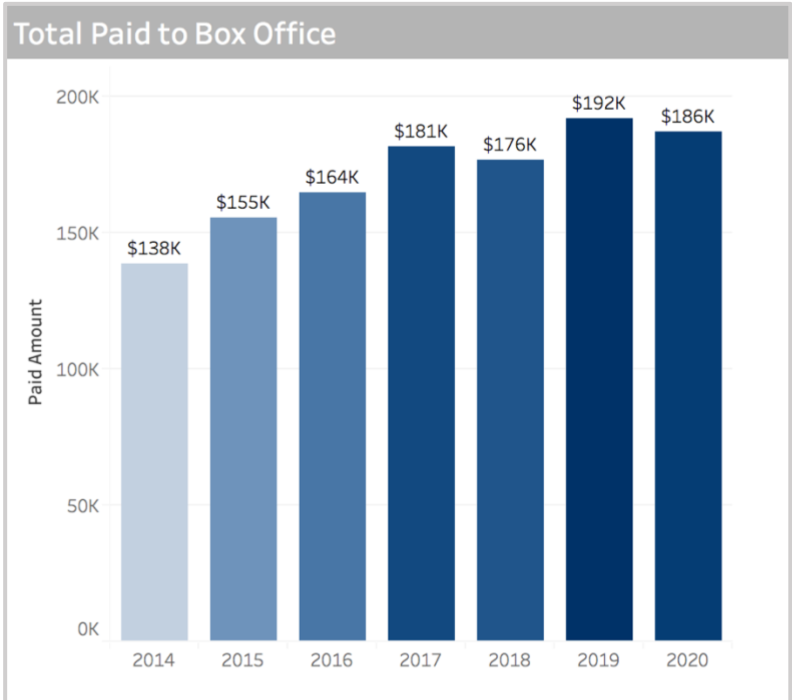
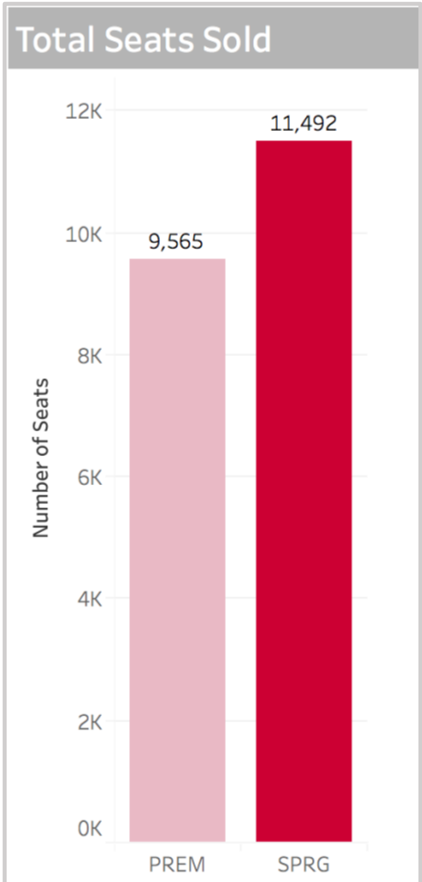


Figure 2. Total amounts paid to box office each year

As shown in Figure 2, the total amount paid to the box office each year is often reflective of the number of seats sold. Both bar charts have similar trends where the values dip in 2018 and are the highest in 2019. It is impressive to see the value in 2020 so high when we remember that the season got cut short. The season, 2019-2020, has the second highest value whereas it had the second to least value for number of seats sold. This difference could be due to prices of seats increasing this past year, so they made more money even though they didn't sell more seats than past years. There are certain discounts for age groups, so there could also be more people that attended shows in 2020 that didn't receive a discount. It is assumed that if sales continued through the rest of the 2020 dance season, the values would show an upward trend following 2018 and 2019.



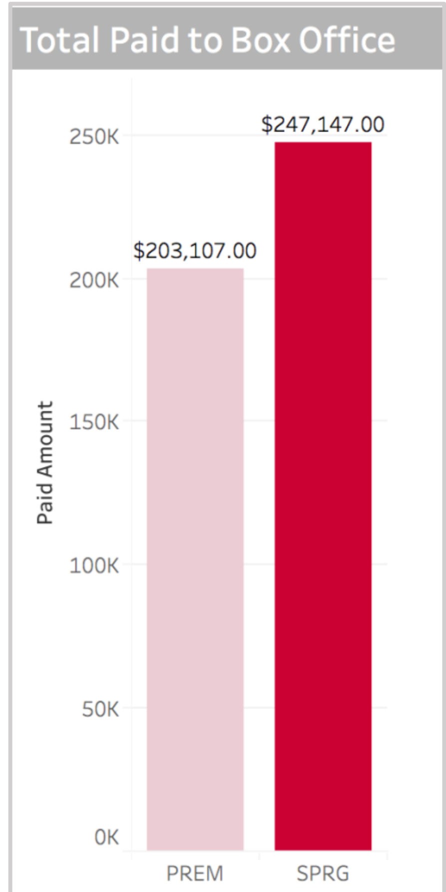


Figure 3 & 4. Comparing Premium Blend and Spring Collection's paid amount and number of seats 2013 - 2019

After looking at total number of seats sold and the total amount paid to the box office, I decided to narrow down the data set and analyze two specific shows and their corresponding values in Figures 3 and 4. This analysis includes all data recorded in Ticketmaster over the years with the exception of 2020. The 2020 shows were omitted from the analysis since Spring Collection did not occur. The three shows that occur each year and have the same name each year are Jazz in AZ, Premium Blend and Spring Collection. The student shows and the February show happen annually as well, but they have different names each year which made it hard to compare year to year since it wasn't consistent in the dataset. Premium Blend is abbreviated as "PREM" and Spring Collection is abbreviated as "SPRG" as seen above.

Jazz in AZ is not included in this visualization because it is hard to compare it to the other annual shows. Jazz in AZ is slightly different because half of the tickets to these performances are sold as a package to attend the weekend workshop with master classes. The Friday performances are reserved for Jazz in AZ participants that are groups of dancers from studios across the country. Therefore, the money made from the Friday performances do not go directly to the box office and are not recorded the same as Premium Blend and Spring Collection.

Looking at the data displayed, it is evident that Spring Collection sold more seats and made more money in the box office. This could be due to a variety of reasons. Premium Blend happens in

November whereas Spring Collection typically occurs in late April. Since Premium Blend occurs earlier in the dance season, there may not be as many people that have subscribed or heard about the dance program for that year. In addition, Spring Collection may have more performances because it is spread across two weekends, whereas Premium Blend is just one weekend. This could also argue that since it is spread across more time, more people are available because of the increase in options.

Ticket Buyer Patterns

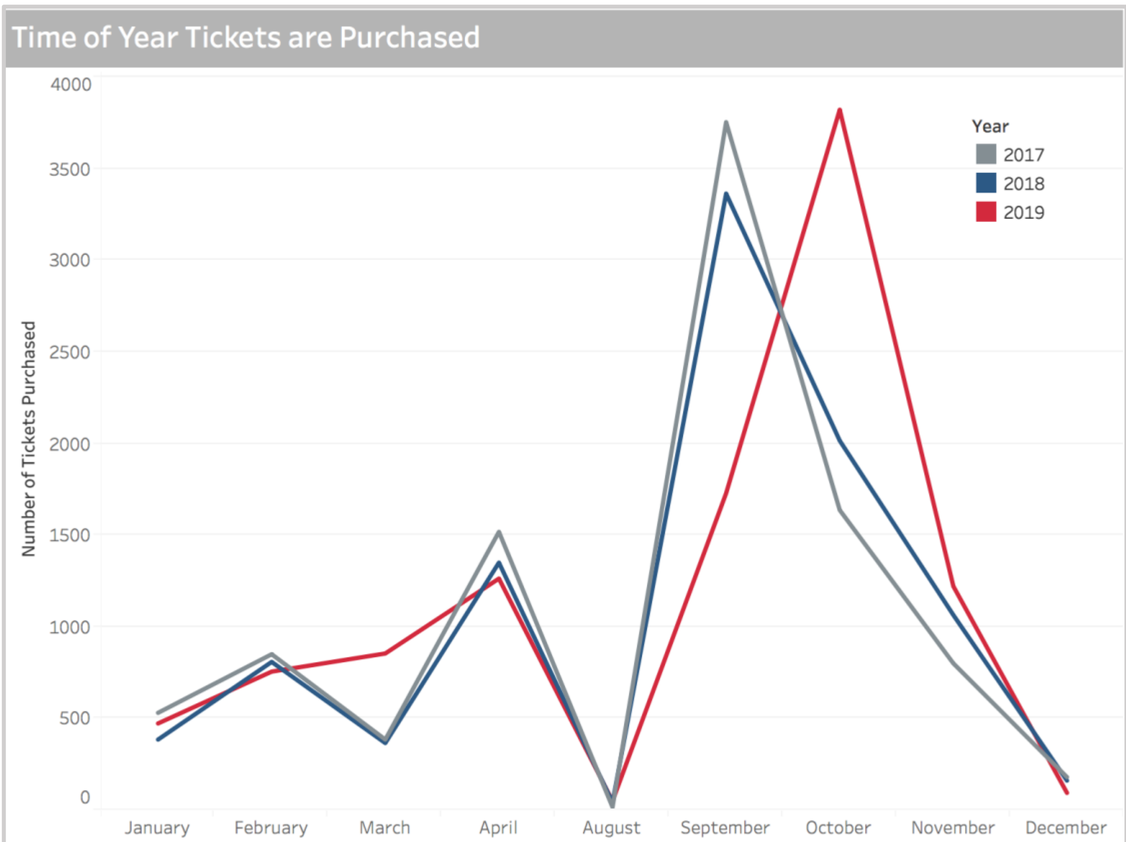


Figure 5. Number of tickets bought throughout the year

In Figure 5, the time of year tickets are purchased is shown; it varies since there are shows scattered unevenly throughout the year. As seen above, tickets are purchased almost year-round with the exception being May, June, and July. School usually ends around mid-May, so it makes sense that tickets wouldn't be purchased during summer because there are no shows. The last show that is bought through the box office is Spring Collection and the end of the year student show. Most of the time, these are over by the end of April, so there aren't any tickets to purchase after that.

There are similar patterns that emerged for the month people purchased tickets even when looking at three different years. The data for 2019 indicated that more tickets were purchased in October when compared to 2017 and 2018. Although, these two years have higher amounts in September than 2019. I suspect that this is due to Jazz in AZ being later in October of the 2019 season compared to being in September in previous years. It is easy to tell what time of year shows are because there are clear peaks in the line graph. It is interesting that August was included, so there must be a small number of tickets that were bought for Jazz in AZ extra early.

Another thing to note is that the September/October amounts are much higher than any other time of the year, but this could be due to season ticket subscriptions being purchased near the beginning of the school year. Season ticket subscriptions typically include tickets for all the main shows for the dance season, so it might make sense that this is when the box office records the tickets even for later shows. More tickets were purchased in March 2019 than previously, and this could be due to releasing more information about the Spring shows during this time.

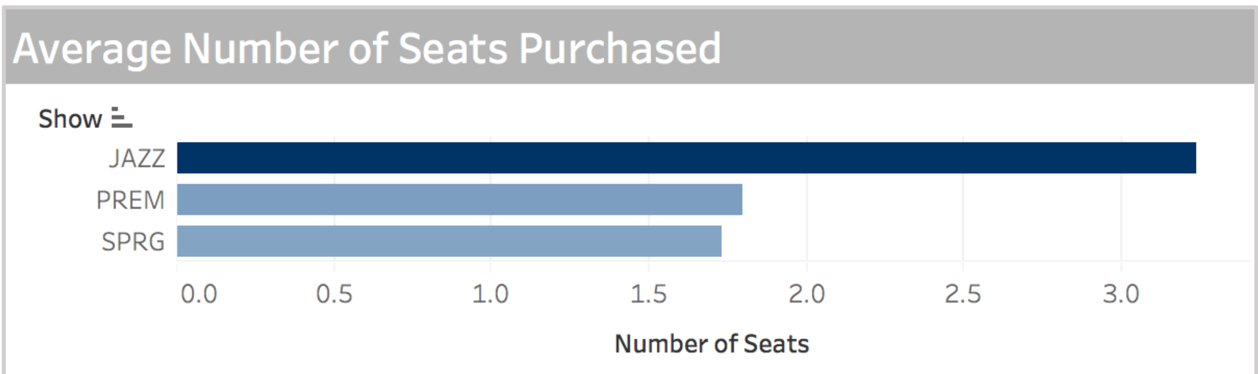


Figure 6. Average number of seats purchased per order

The box office tracks the number of seats that are purchased per order as seen in Figure 6. This data was intriguing and answers the question of how many tickets on average were purchased together to tell us how many people usually attend shows together. I wondered if people went alone, couples went together, or if families attended more commonly and how this varied depending on the show. Jazz in AZ is abbreviated as "JAZZ" in Ticketmaster, and we also looked at the two other shows that occur annually. It is clear that people usually attend Jazz in AZ with more people (purchasing multiple seats under one account) when compared to the lower values for Premium Blend and Spring Collection. Jazz in AZ is targeted for younger audiences, so it makes sense that people might bring their kids or other family members and purchase tickets under one account.

Premium Blend's and Spring Collection's average number of seats purchased per order are closer to each other which led to a hypothesis that more season ticket subscribers may attend these shows. They would often have one or two people under an account and would therefore have a smaller number of seats purchased in a transaction. People may attend the show with others but only purchase their ticket under their account. This seems more common since I have observed people who come to the performances with a friend where they both would have purchased their seats on their own account.

We cannot rely on this to explain how many people attended the show together; neither can we assume that only one person paid for the tickets purchased. With the amount of electronic payment options, it could be possible that people purchase tickets with an account but are paid by their friend or family member via Venmo or Paypal. There is also the possibility that someone may not be able to attend with the rest of the group, and the account holder allows someone else to buy the ticket from them. This happens a lot with family members of dancers in the show that might buy extra tickets to make sure they have enough, but they end up having leftover tickets. The tickets cannot be sold back to the box office, but they could sell it for the purchase price to another person needing an extra ticket.

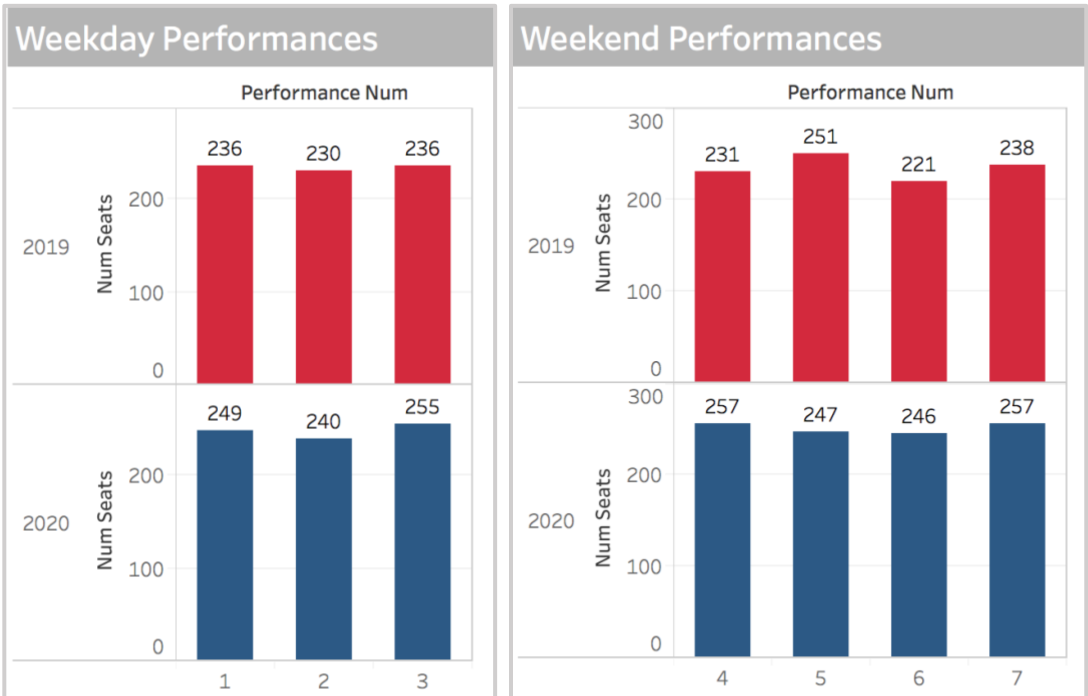


Figure 7 & 8. Number of seats bought for weekday vs. weekend performances of Premium Blend

Figure 7 and 8 look at number of seats sold on weekdays compared to weekends. Weekday is considered performances Wednesday through Friday, whereas weekend is considered performances on Saturday or Sunday (matinee and evening both days). Wednesday night's performance is labeled as "1" and goes sequentially through the week where Saturday matinee is labeled "4" and Sunday night is "7". I used Premium Blend for this analysis, so we could compare a show using the year 2020. Premium Blend in 2019 and 2020 included the exact amount of performances, so it is easy compare the two.

The most common performance in 2019 was the Saturday night; whereas in 2020, the Saturday matinee and Sunday night performances were most popular. It is surprising to me that Sunday night was one of the most attended performances; one hypothesis is that Sunday evening may have been a "last chance" opportunity and driven attendance. Opening night typically has a bigger audience because it is the premiere of all the works for that series. We can see that in both 2019 and 2020, it was the third most popular performance when looking at the number of seats sold.

As you can see, there isn't a lot of variation between different days of the week and weekend. Most shows and performances are sold out, but for the ones that aren't, there are a variety of tactics to attract more ticket buyers. We could use this information to offer discounts and advertise about the less attended shows as mentioned above. The social media platforms could release special deals for tickets to be bought that same day through the box office for shows like the Thursday night (performance after opening night) and the Sunday morning performances. This would be a great way to increase the number of seats sold, especially for performances that typically have less people.

Ticket Buyer Information

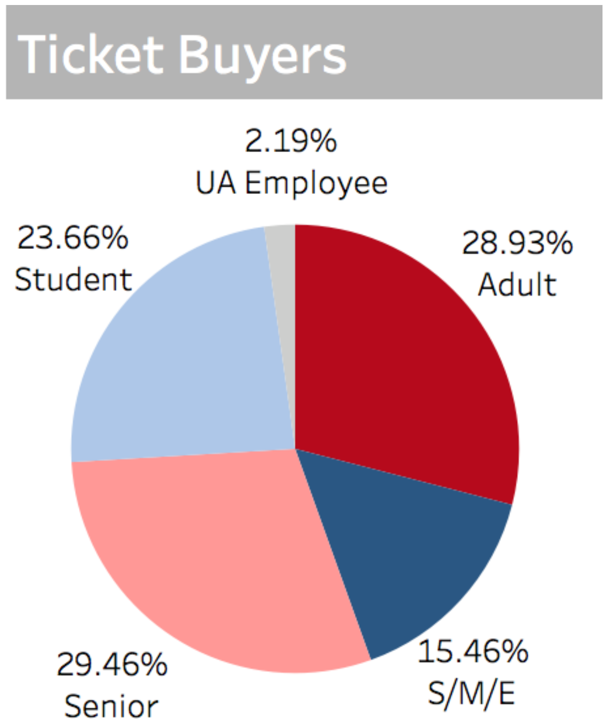


Figure 9. Groups of ticket buyers

As seen in Figure 9, there are a variety of different types of attendees that go to dance shows over the years (2014-2020), both in terms of age demographic and title (senior/military/employee). This data took a bit more cleaning than the other pieces because there was over 20 options in this Excel column as mentioned in the data preparation section. I could summarize them into the groups displayed in the pie chart that would effectively cover five primary groups.

In the limitations section below, I describe the ticket buyer type, "S/M/E", which describes Seniors, Military and Employee. The box office typically recorded someone as type, S/M/E, if they fell under two of the three categories. This is important to note because Senior and UA Employee are also ticket buyer types listed. It is unclear whether or not these categories could also be counted with the S/M/E or if they are solely a Senior or UA Employee and do not fall under another category explained.

The biggest ticket buyer type is Senior with the next largest being Adult. Arizona attracts many older groups of people because of the attractive temperatures and "snowbird" destination during the winter months. The dance season is during the school year which is when many Seniors come to Arizona to stay during the winter. Dance shows provide this audience entertainment, and many of them are connected through season ticket subscriptions and are also members of the board.

The smaller groups of ticket buyers include UA Employees and Students (if we exclude the S/M/E group since it is unspecific). We could target these ticket buyer groups by offering discounts to them for shows that aren't as sold out. In Figures 9 and 10, we examine weekday and weekend performances. We could look at which performances don't have as many tickets sold and offer discounts to UA Employees and Students as it gets closer to that performance. These groups already have certain discounts, but it may be helpful to also advertise this outside the box office application. This could be sent out in a University-wide email one week highlighting the talent in the dance program. It could also be advertised around campus in the Union, buildings where classes are held, or dorms, so people can see them as they walk past.

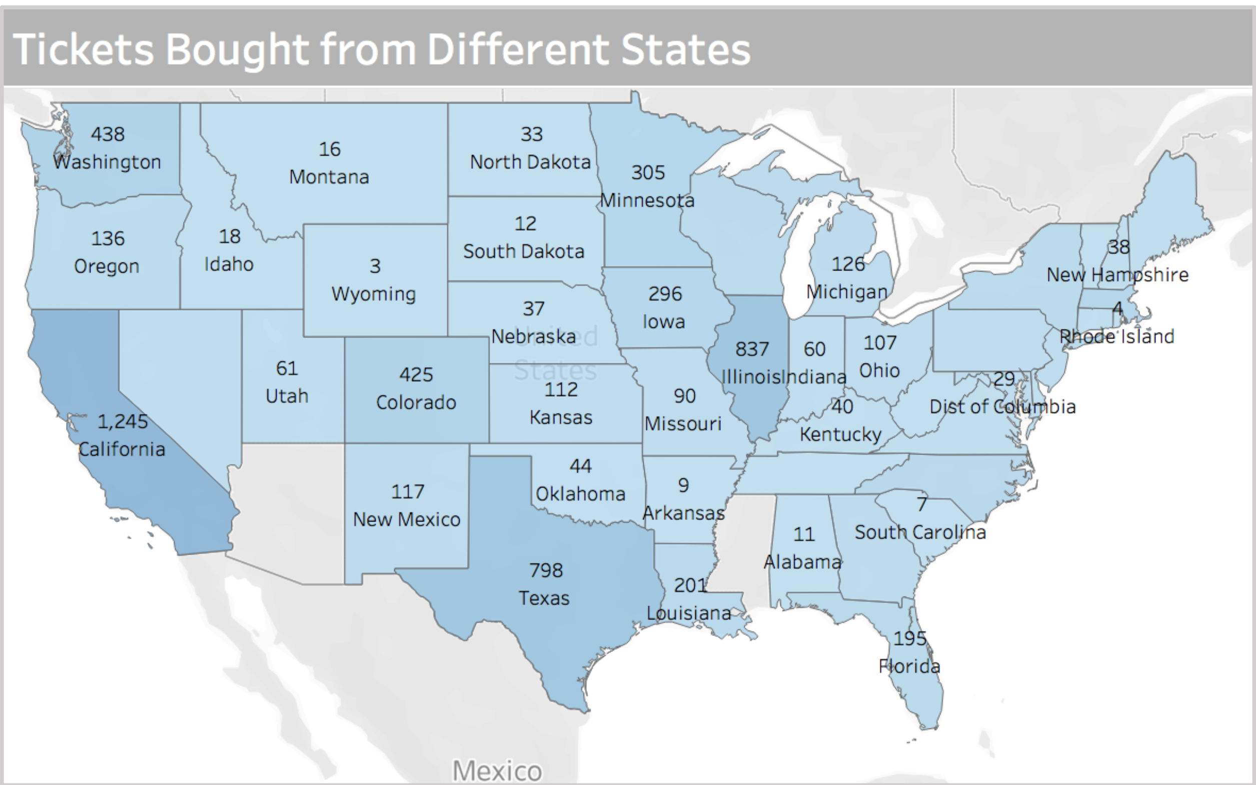


Figure 10. Number of ticket buyers from each state

Because I had ticket buyers' zip code and state, I could use this data to visualize the number of seats purchased by state as seen in Figure 10. It is obvious that Arizona has the most tickets each year with a total of 55,128 tickets bought over the years because it is the host state. I chose to eliminate Arizona from this map visualization so that we could easily see what other states had ticket buyers that purchased tickets to shows. This data comes from 2014 and is inclusive of all shows each year following that.

It is evident that California is the next state with the largest number of seats bought. It is interesting to see that Illinois has had 837 tickets bought over the years, and they are not located in close proximity to Arizona as the other states with higher number of seats. This might be due to the fact that there is a high number of students from Illinois with family attending shows, and the fact that there may be seniors that spend winter in Arizona. On the other hand, there is one state, Mississippi, that doesn't have any ticket buyers from the past seven years. This is surprising especially considering that all other states have been covered. Ticket buyers from across the country could be people that move to Arizona during the winter months and attend dance shows as a form of entertainment or it could be people that are related to dancers in the shows.

One way we could expand ticket buyers across the country and the world is by increasing social media coverage and promotion of upcoming shows. With the abundance of technology, it has exposed everyone to more information without having to be in the same physical location. The University of Arizona Dance Instagram and Facebook are already releasing previews and video clips from rehearsals for shows, but they could advertise to people who aren't following them as well. I know with Instagram ads you can have ads appear within a certain location radius. It could be beneficial to advertise to a radius surrounding the U of A campus to target more students and UA Employees. With Facebook, you can advertise using users' interests, so the dance social media account could target people with interests in dance, choreography, and theater, for example.

Season ticket subscriber Information

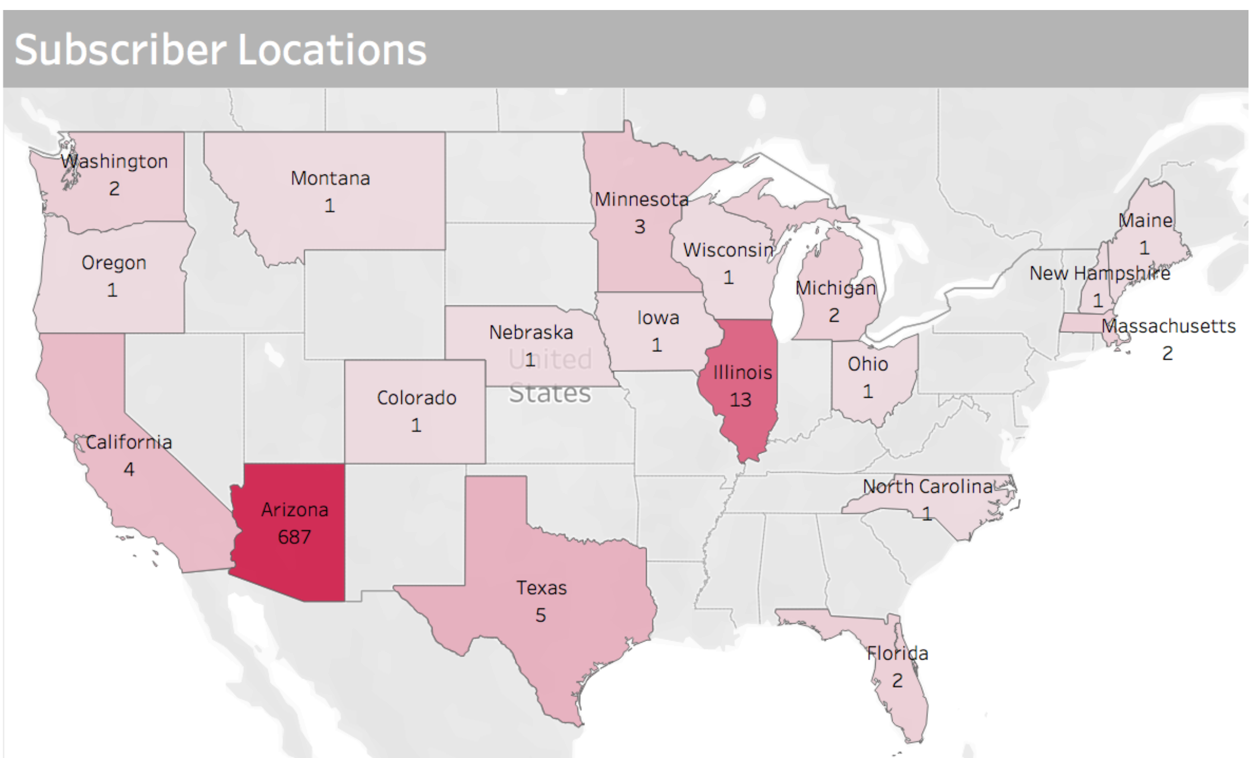


Figure 11. Season ticket season ticket subscribers' locations based on zip code attached to their account ID

Most season ticket subscribers come from Arizona due to the close proximity to the University and being there year-round. As shown in Figure 11, there are season ticket subscribers that are scattered across the country, but it is not as common. This visualization shows season ticket subscribers over the years that have purchased a subscription one year or maybe multiple years. This data was obtained from the zip code that is listed on their account. The season ticket subscribers that are considered out of state may live in Arizona for part of the year or some may just come to Tucson for each series of shows.

It is interesting that Illinois has the second highest number of season ticket subscribers, and it is not located very close to Arizona. This is a highly populated state with its major city of Chicago which has a recognized dance scene. There are also a lot of current students and alumni that are located in Chicago, so one hypothesis is that this could be correlated. As mentioned before, people could use their address from Illinois but live in Arizona for the entire dance season. If I were to study further, I would contact season ticket subscribers with a survey to see how long they live in Arizona and where else they reside.

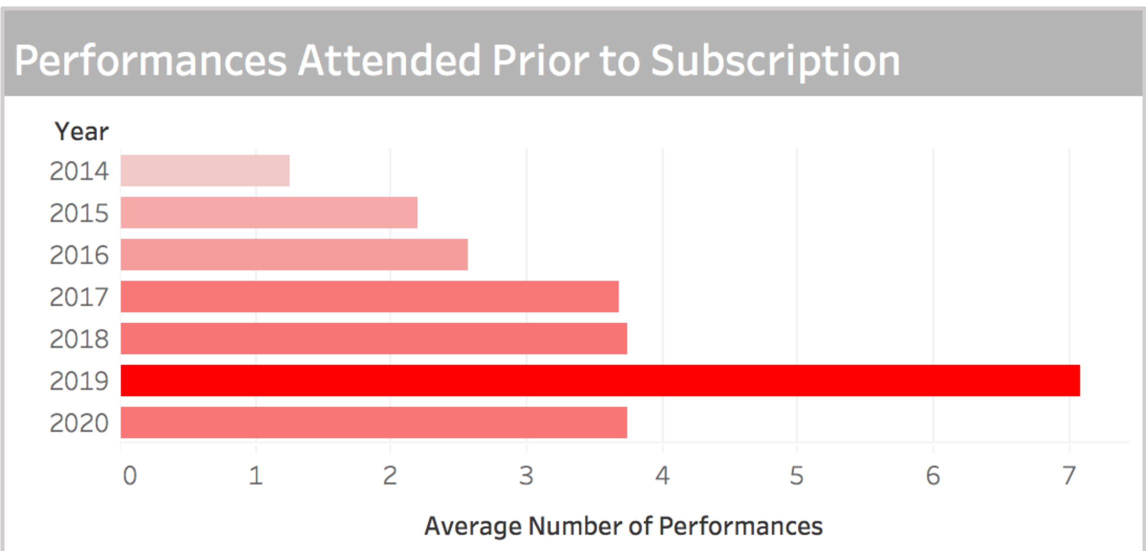


Figure 12. Average number of performances attended prior to getting a subscription

When meeting with Whitney at the beginning of the year, this was one of the specific charts she had requested. She explained that the box office does an analysis on season ticket subscribers since they support a large portion of the dance program. Whitney asked if it would be possible to see how many performances ticket buyers attend before buying a subscription. I didn't know at the time if this would be possible, but as I described my process in the data preparation section, I ended up with this result. This visualization shows the average number of performances attended before purchasing a subscription over the years from 2014 to 2020.

We can see an increase in the average number of performances people attended before buying a subscription through 2019. In 2019, the average number almost doubled from previous years which means people attended more performances before purchasing a subscription during that year. I don't predict that the drop in average number of performances attended prior to getting a subscription during 2020 can be linked to the season being shortened. It seems more likely that the number in 2019 is extra high and that the average would stay between three and four. The overall average is 3.048 which indicates that on average people attend roughly three performances before purchasing a subscription.

This information is helpful for the dance program and board members because they are looking for ways to increase the number of people that have season ticket subscriptions. Since we know the average number of performances attended before purchasing a subscription, board members and/or Whitney's team could reach out to people who have attended three performances. This is when they could target people, especially those that have an Arizona zip code, and promote the subscription to them. They could also survey people at this point to see if they are considering getting a subscription, and if they aren't, ask them their rationale.

CONCLUSION

Limitations

There are some data discrepancies and limitations that kept the analysis from being completely accurate. If there was a cell that was left blank, it was not included in the visualization; and therefore, did not impact the conclusions drawn. The year that is included on multiple charts' axes is the spring of that school year. For example, 2014 means it was during the 2013-2014 school year. There is a ticket buyer group labeled "S/M/E" which stands for Senior, Military, UA Employee. The data revealed that this category was selected if the ticket buyer qualified for two of the three categories. This made for some confusion in my data because I also had separate categories for Senior and UA Employee. Another thing to note is that COVID-19 ended the school year early for data with the year 2020. There were two shows, Spring Collection and the student choreographed show, that didn't occur which could lead to lower numbers during the 2019-2020 season and should be noted when reviewing the findings.

Some other discrepancies include complimentary tickets which are indicated on the Excel sheet under the "comp_name" column. Occasionally, this impacted the visualizations unintentionally, like the number of tickets sold, comp tickets would be included but not in the paid amount. People purchasing tickets to the shows may live in Arizona only during the winter months which happens to be the dance season. The zip code is whatever is listed on their account which may reflect their Arizona address or another address that they live for the other part of the year. Finally, the student shows and show in February is called something different each year, so we couldn't compare those as easily. Instead, I tried to compare shows that happened annually and were abbreviated as the same in Excel.

Future Direction

Through my analysis, I obtained a variety of information on ticket buyers; season ticket subscribers; revenue the box office calculated; and seats sold. I studied several of the attributes listed in the dataset and combined new attributes that were used to create the visualizations in Tableau. These visualizations were used to display the summarized data in a way that is easily understood. From there, I could make conclusions for why the data appeared this way and make recommendations for how the School of Dance could improve. It was useful to have data from seven dance seasons and recognize the changes that have occurred since the box office started recording this data. I am hopeful that this information will be useful for the School of Dance and making decisions in the future.

This research is beneficial for the School of Dance to make improvements as they move forward in future seasons. They can use the visualizations and data analysis to make changes in how they advertise and target certain groups for their performances. The social media teams can use this information to see which shows should be previewed more on the Instagram or Facebook. The School of Dance might be considering adding another performance to a series of shows, and this information would help them decide which day of the week would be best. It is unknown how COVID-19 will impact gatherings in the next couple of years, but the dance program could use these findings to decide how they make changes to how many people can come to performances and what age groups are included.

If this project were to proceed further, I would explore other topics to improve outcomes for the School of Dance. The dataset of 34,000 records contained ample information for this and additional research. I might study what seats are purchased earliest or look at the student shows and compare fall semester to spring semester. I could also look at international ticket buyers that did not include a zip code since that is not applicable. There were also a few discrepancies with the data, so I might try to do more data cleanup to increase consistency and accuracy.

I have created a business report with visualizations that make recommendations on the business decisions that can be made by the box office and the dance program. I enjoyed using my dance major to find useful data that I could translate into visualizations and findings to wrap up my undergraduate experience. I look forward to expanding upon this gained knowledge and exploring areas of analysis that I am not familiar with in the future!

APPENDIX

Figure 13. Zip codes by state used to do map visualizations

Zip codes by State

USPS Zip Code Locator

ID	State Name	ST	Zip Min	Zip Max
6	Alaska	AK	99501	99950
5	Alabama	AL	35004	36925
69	Arkansas	AR	71601	72959
8	Arkansas (Texarkana)	AR	75502	75502
7	Arizona	AZ	85001	86556
9	California	CA	90001	96162
10	Colorado	CO	80001	81658
11	Connecticut	CT	6001	6389
73	Connecticut	CT	6401	6928
13	Dist of Columbia	DC	20001	20039
74	Dist of Columbia	DC	20042	20599
63	Dist of Columbia	DC	20799	20799
12	Delaware	DE	19701	19980
14	Florida	FL	32004	34997
15	Georgia	GA	30001	31999
65	Georga (Atlanta)	GA	39901	39901
16	Hawaii	HI	96701	96898
20	Iowa	IA	50001	52809
66	Iowa (OMAHA)	IA	68119	68120
17	Idaho	ID	83201	83876
18	Illinois	IL	60001	62999
19	Indiana	IN	46001	47997
21	Kansas	KS	66002	67954
22	Kentucky	KY	40003	42788
68	Louisiana	LA	70001	71232
23	Louisiana	LA	71234	71497
26	Massachusetts	MA	1001	2791
58	Massachusetts (Andover)	MA	5501	5544
77	Maryland	MD	20331	20331
78	Maryland	MD	20335	20797
25	Maryland	MD	20812	21930
24	Maine	ME	3901	4992
27	Michigan	MI	48001	49971
28	Minnesota	MN	55001	56763
57	kc96 DataMO	MO	63001	65899
29	Mississippi	MS	38601	39776
59	Mississippi(Warren)	MS	71233	71233
31	Montana	MT	59001	59937
38	North Carolina	NC	27006	28909
39	North Dakota	ND	58001	58856
67	Nebraska	NE	68001	68118
32	Nebraska	NE	68122	69367
34	New Hampshire	NH	3031	3897
35	New Jersey	NJ	7001	8989
36	New Mexico	NM	87001	88441
33	Nevada	NV	88901	89883
60	New York (Fishers Is)	NY	6390	6390
37	New York	NY	10001	14975
40	Ohio	OH	43001	45999
70	Oklahoma	OK	73001	73199
41	Oklahoma	OK	73401	74966
42	Oregon	OR	97001	97920
43	Pennsylvania	PA	15001	19640
44	Puerto Rico	PR	0	0
45	Rhode Island	RI	2801	2940
46	South Carolina	SC	29001	29948
47	South Dakota	SD	57001	57799
48	Tennessee	TN	37010	38589
49	Texas (Austin)	TX	73301	73301
71	Texas	TX	75001	75501
72	Texas	TX	75503	79999
61	Texas (El Paso)	TX	88510	88589
50	Utah	UT	84001	84784
64	Virginia	VA	20040	20041
75	Virginia	VA	20040	20167
76	Virginia	VA	20042	20042
52	Virginia	VA	22001	24658
51	Vermont	VT	5001	5495
62	Vermont	VT	5601	5907
53	Washington	WA	98001	99403
55	Wisconsin	WI	53001	54990
54	West Virginia	WV	24701	26886
56	Wyoming	WY	82001	83128

Found from following resources:

https://www.extendoffice.com/documents/excel/5877-excel-convert-zip-code-to-text.html
http://www.structnet.com/instructions/zip_min_max_by_state.html