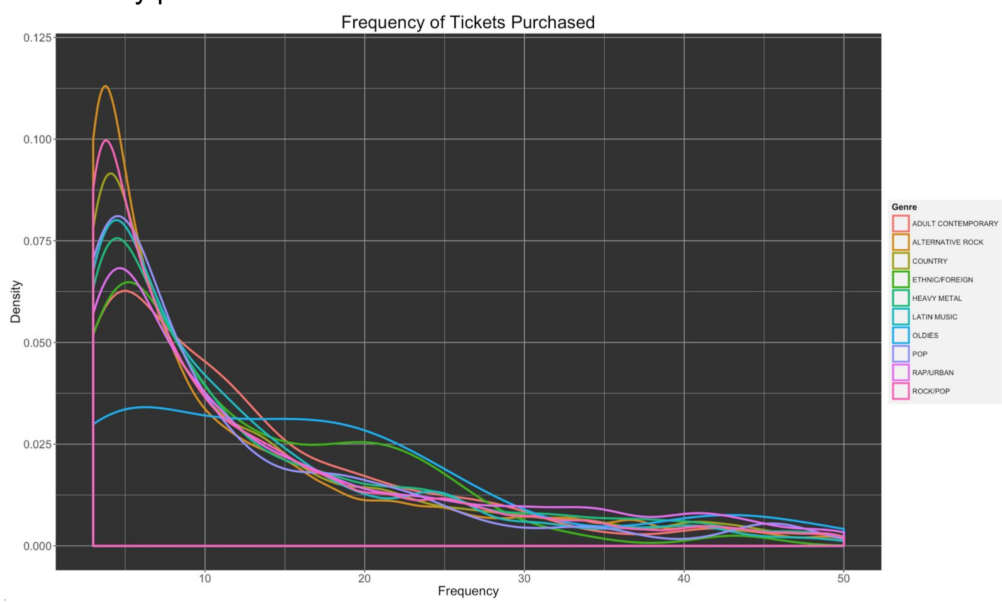
TicketMaster.com Visualization of Ad Suggestion Algorithm

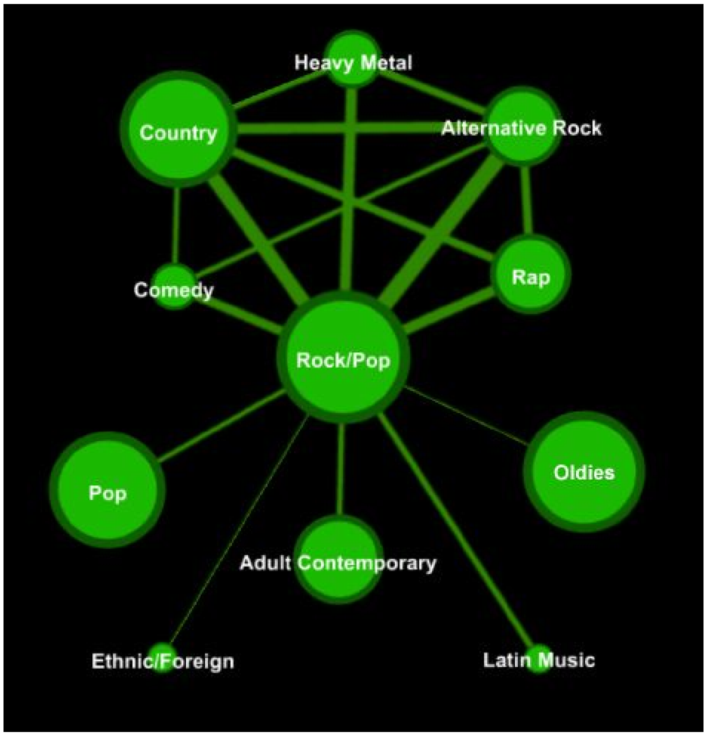
This project was completed from data supplied from TicketMaster.com as part of the 2016 DataFest. Using an adjacency matrix, we created concert genre suggestions for loyal users of the site, and using relative popularity of genre by location we create suggestions for new users based on the state or territory the are in.

Who are the different populations of users on TicketMaster.com, and how should we advertise to them? For this project, we have focused on the concert-going users of TicketMaster, and we are focused on finding out which genres of concerts we can suggest to them to optimize TicketMaster's ad campaigns. To determine which genres we should use, we brought in an outside data set from a survey on which music genres are the most popular in the US, the main country of residence for our ticketmaster sample[1]. Next, in order to screen out potentially malicious users of the website, we looked at all the Purchase Party I.D.'s that popped up more than 5 times for the exact same event. Interestingly enough, this list of Purchase Party IDs seemed to make up most of the population of users of TicketMaster that had 50+ total transactions on the site, leading us to believe these are "scalper" type users and should be ignored in ad suggestion.

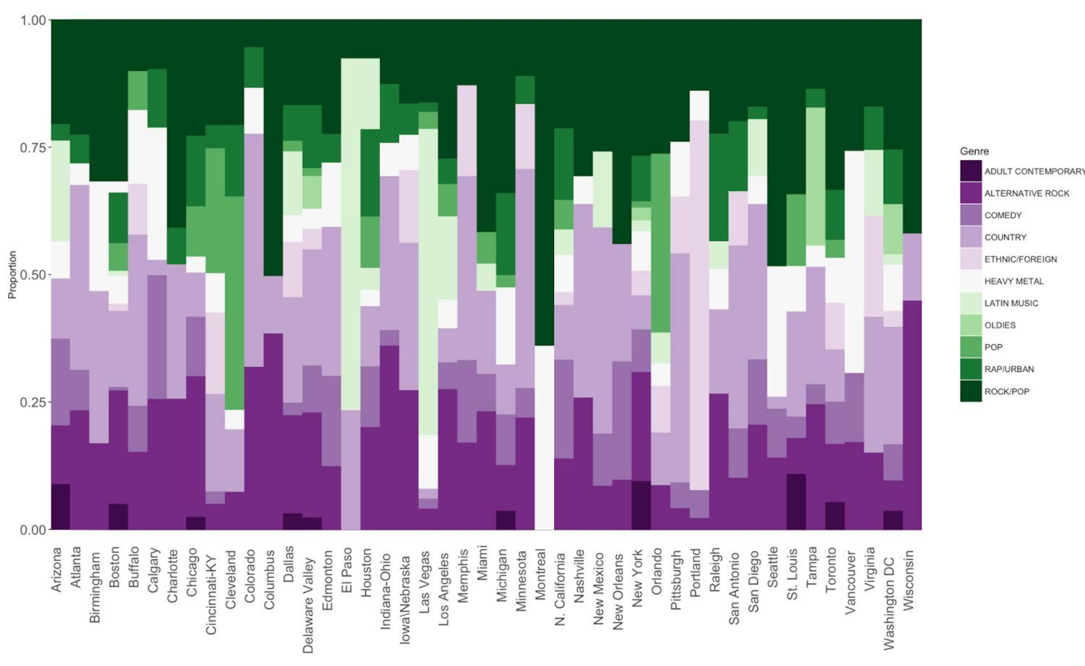
Next, to distinguish between "one-time" or "non-loyal" and "loyal" users of the site, we evaluated a density plot of transactions on the site separated by our most popular genre categories. We saw two clear trends: users that tend to buy 1-4 tickets and never visit the site again and users that are frequent customers and have visited the site around 5-50 times. We see large amounts of these "loyal" type users buy tickets for Ethnic Music and Oldies concerts. Here is our density plot:



Having identified loyal users, we then wanted to study what related genres TicketMaster should suggest to these returning users. To do this, we calculated the correlations between each genre for this segment of the population using an adjacency and the past concert history of these "loyal" users. Thicker edges thus correspond to a stronger correlation, and the size of each node corresponds to the popularity of each genre (from the statista preference poll, not TicketMaster's data, as to not confound the actual popularity). In making suggestions, a big node with few edges represents a genre whose recommendations would only include itself, but a small node with many strong edges represents a genre whose recommendations would include many other genres. We may use these data, combined with a returning user's previous purchases, to suggest new (or keep the same) genres for the ads targeted at them.



After targeting the loyal users, we focused on advertising to new customers, of whom all we may know is their location. Taking data from the one-time user demographic, we calculated the relative popularity of each genre in each location based off of TicketMaster's data. We then created a type of heatmap that visually shows the popularity of each genre in each region, where colors represent different genres and the size of the bar represents the percent of new users in that region who buy tickets to concerts of that genre. Now given a new user's location, TicketMaster can suggest ads for concerts in these popular genres for the area. Since new users do not have a large concert history like other users do, this type of suggestion will make relevant suggestions instead of extrapolating from a small concert history.



Appendix

1. Statista user preference poll:
http://www.statista.com/statistics/442354/music-genres-preferred-consumers-usa/

2. ASA DataFest website:
https://www.amstat.org/education/datafest/

3. This project uses the TicketMaster purchase dataset given to participants in 2016 DataFest.

4. We used RStudio and Eclipse as our programming tools for this project.