A Yelp-Guided Tour of Las Vegas, Montreal, Phoenix, and Charlotte
By Lisa Barcelo, Melanie Costello & Priya Gupta

Summary

This analysis utilized four data sets from the Yelp Data Challenge, which provided data pertaining to user activity, reviews of businesses, check-ins at businesses, and some basic data about businesses in a select number of cities in the U.S. and abroad. For the purposes of this project, four cities were selected for analysis based on a combination of geographic diversity and the number of reviews posted in these locations.

The Data

All raw data and a compiled .csv file needed for this project can be found in the following folder submitted in conjunction with this project:
https://drive.google.com/file/d/0B7H1JlmYiCTtWG1Sajc2akFIRkU/view?usp=sharing

It was made available by Yelp as part of the Yelp Dataset Challenge and was originally downloaded from the following Yelp site: https://www.yelp.com/dataset_challenge/dataset

Raw data is all in .json format and specific files will be referenced in discussion as follows:

File Name	Referred to as
yelp_academic_dataset_business.json	Business data
yelp_academic_dataset_checkin.json	Check-in data
yelp_academic_dataset_review.json	Review data
yelp_academic_dataset_user.json	User data

Getting Started

The below table details an excerpt of the top cities sorted by the total number of reviews for businesses in each city (obtained from the review data):
To narrow the focus and provide more meaningful insight, this project looks specifically at information connected to Yelp data in Las Vegas, Phoenix, Charlotte and Montreal. The data was further narrowed to include only bars and restaurants in the select cities.

city	review_count
Las Vegas	957690
Phoenix	330437
Scottsdale	179521
Charlotte	131838
Tempe	95650
Pittsburgh	94942
Henderson	85463
Montréal	69973

Questions & Insights

1) Are new users more likely to post favorable (5 star) reviews? Are longtime users more likely to post less favorable reviews?

Before delving into each city's food scene, it became necessary to assess the demographics of the user types. It was hypothesized that there would be an inverse relationship between a user's longevity and the favorability of reviews given. A newer or less active user might give more high ratings, whereas a more seasoned or more active user might be more choosy with how he or she hand out stars.
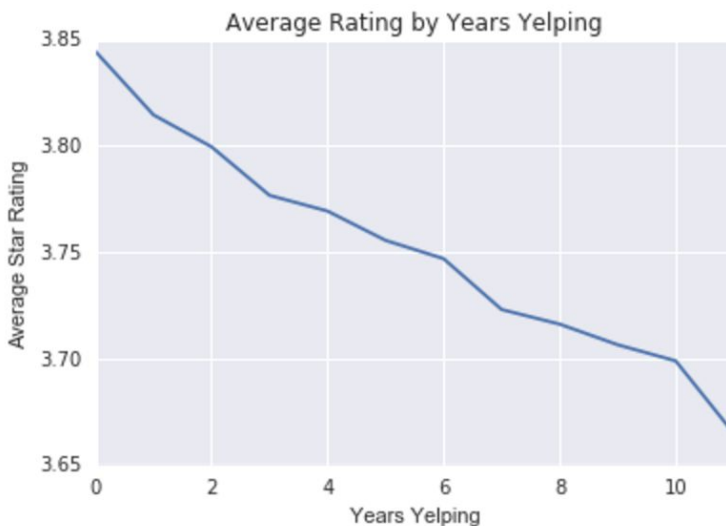
Through an analysis of the user data, we found that longer Yelping users give lower average star ratings to restaurants in comparison to users that have just started yelping. The graph at left shows the average stars given by a user steadily decreases the longer a user has been Yelping.

It's interesting to note that ratings are out of 5 stars, so the difference between a star rating of 3.8 and 3.65 may not seem significant but does demonstrate the changing behavior of users.

Merging the user data with the business and review data enabled a deeper dive to examine each of the four cities individually. The goal was to see if each city showed a similar trend of decreasing star ratings over time. The graph below details a city-by-city breakdown of average rating in relationship to the users' length of time using Yelp. Note that the graph reads from newest to oldest, meaning the newest users are represented on the left (e.g. 0-2 years Yelping) and the oldest users are represented on the right (e.g. 10+ years Yelping).
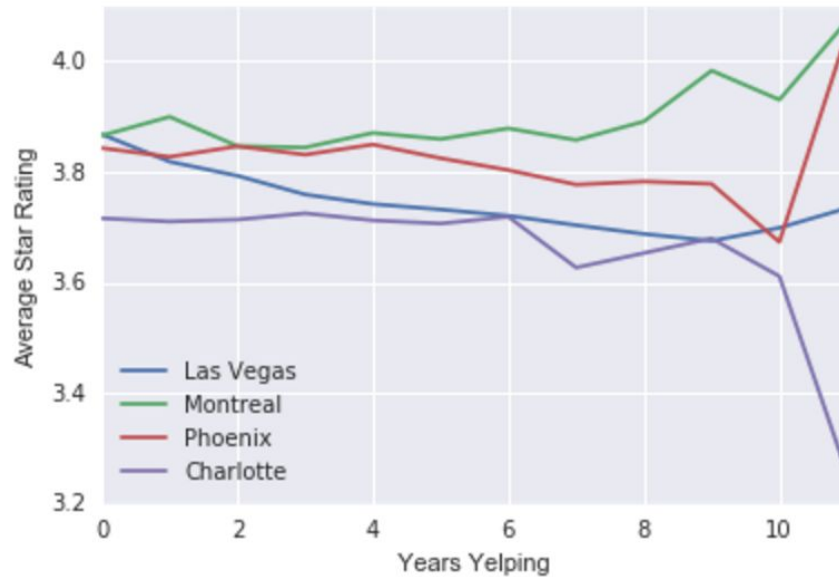
Overall, the data shows a somewhat similar downward trend in years 2-7 of Yelp use. However, the behavior of longtime users in each city differs greatly at approximately the 8 year mark. Notably, users in Montreal and Phoenix give much higher average ratings, while the average rating by users in Charlotte drops precipitously at the 10 year mark.

In an effort to explain these discrepancies, the merged user, business and review data was used to derive new insights. The variability in behavior in each city may be explained by the overall number of users for that time period. The graph below groups users based on the number of years they have been using Yelp.
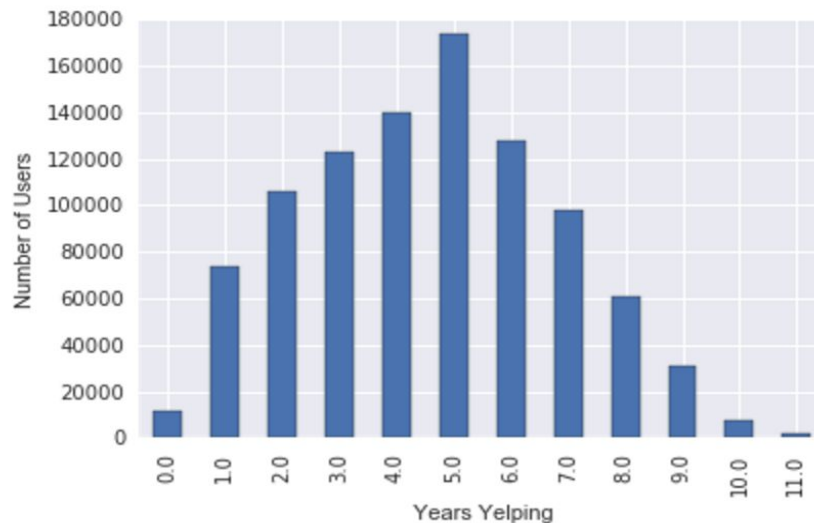
Yelp opened its doors in 2004, almost 12 years ago this October. There is a small population of Yelp users that have been yelping for 8 or more years. Therefore, with a smaller set of users, there is a chance for larger variation by city.

2) Who is behind the reviews?
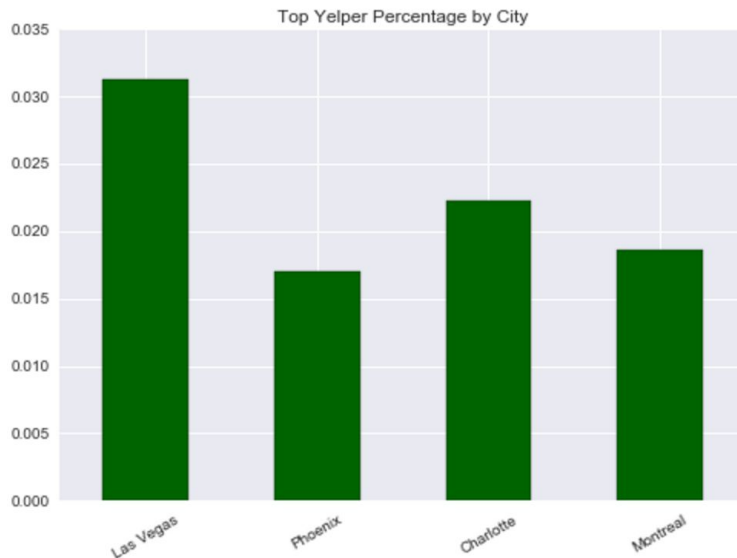
Not all reviews are created equal. For tourists in a new city, time is of the essence. Visitors generally have just a few days to find and visit the trendiest neighborhoods, and dine at the most popular eating spots. Yelp takes the guesswork out of it. By consulting with only the top Yelpers (users with an average rating of 4.0 stars or more), tourists can immediately find answers to their pressing questions.

An examination of the user data set joined with the business data set revealed that top-rated Yelpers made up around 3% of the Yelper demographic in Las Vegas, and between 1-2% for the other cities.

In delving a bit more into the user demographics, we learned that the most active users were those who signed up between 2008 and 2011, with over 90 million reviews combined. User activity follows a bell-shaped curve, with the least amount of reviews being offered up by both senior users (10+ years as Yelpers), and inchoate Yelpers (<2 years of members).
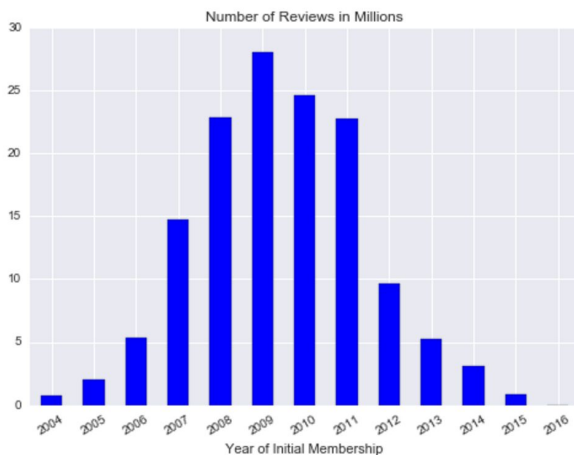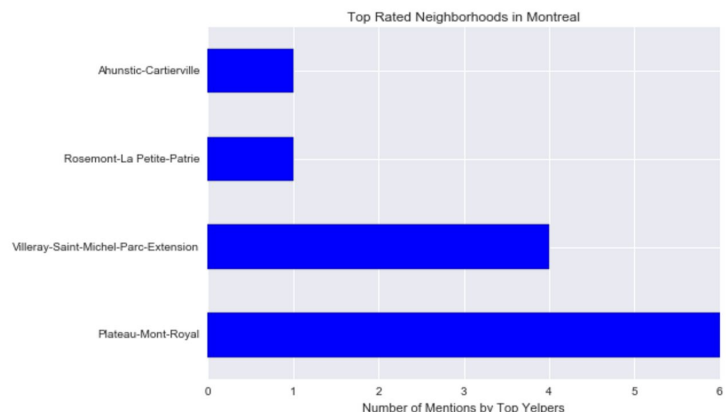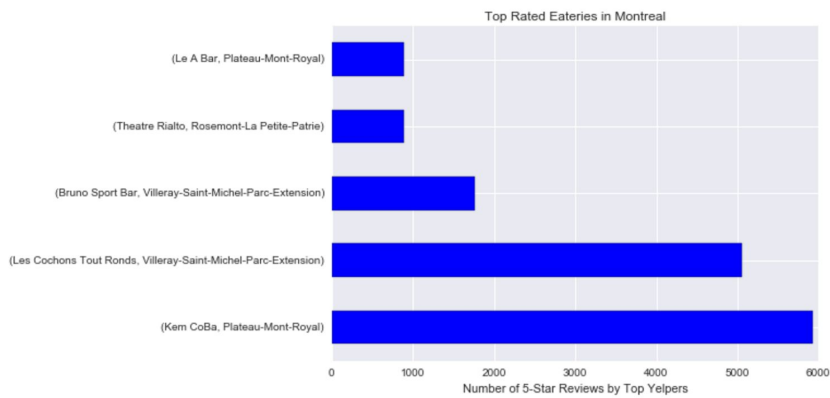

3) How can top Yelpers help people navigate a new city?

Tourists can easily get lost in the winding streets of a new city. Data from top-rated Yelpers can help them quickly find the neighborhoods with the most popular dining options! Merging the user and business data sets helped support a city-by-city breakdown of neighborhoods with the highest density of top-rated restaurants. Essentially, these neighborhoods are the best bets for a tourist looking to find a good bite to eat.

Montreal
Montreal is the largest city in Quebec, covering 141 square miles, about three times the size of San Francisco. One neighborhood, Plateau-Mont-Royal, clearly has a very high density of high-rated restaurants.

Top Rated Eateries in Montreal

(Le A Bar, Plateau-Mont-Royal)

(Theatre Rialto, Rosemont-La Petite-Patrie)

(Bruno Sport Bar, Villeray-Saint-Michel-Parc-Extension)

(Les Cochons Tout Ronds, Villeray-Saint-Michel-Parc-Extension)

(Kem CoBa, Plateau-Mont-Royal)
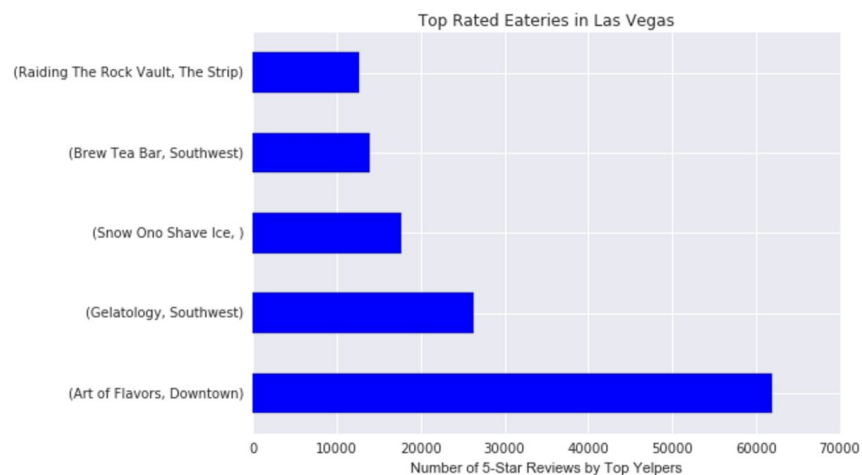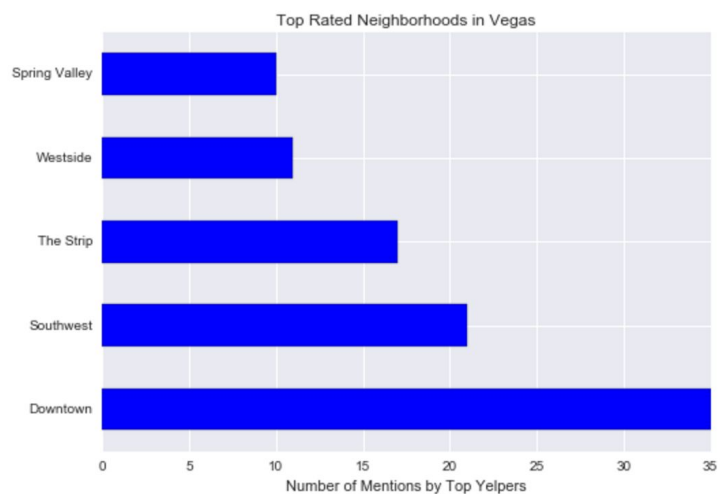
Number of 5-Star Reviews by Top Yelpers

A check of the top rated eateries in Montreal confirms the dominance of the Plateau-Mont-Royal neighborhood. The graph below shows the top-rated eateries in Montreal (determined by the number of 5-star reviews) and their neighborhoods. Two of the top five spots are in the top-rated neighborhood, Plateau-Mont-Royal.

Las Vegas
"Sin City" is about the same size as Montreal, and there is no shortage of things to do under the bright lights. Downtown remains the best place to find top-rated places to eat, but you can't really go wrong.

Las Vegas's top rated eateries in the graph below don't connect as clearly to the popular neighborhoods as those in Montreal did. While some of the most popular neighborhoods are represented in this graph, it's also good to note that the third most popular restaurant, Snow Ono Shave Ice, does not have an associated neighborhood.

Top Rated Neighborhoods in Vegas

Spring Valley

Westside

The Strip

Southwest

Downtown

Number of Mentions by Top Yelpers

Top Rated Eateries in Las Vegas

(Raiding The Rock Vault, The Strip)

(Brew Tea Bar, Southwest)

(Snow Ono Shave Ice,)

(Gelatology, Southwest)

(Art of Flavors, Downtown)

Number of 5-Star Reviews by Top Yelpers

Charlotte

Top Rated Neighborhoods in Charlotte

Charlotte is the second-largest city geographically in this analysis, at just under 300 square miles. NoDa, Charlotte's famous historic art and entertainment district is the best place to have a great evening according to Yelpers, with Arboretum (the central shopping area) coming in a close second.

Examining the top-rated eateries in Charlotte, most of the top eateries reside in the NoDa neighborhood with a Chocolatier coming at the top of the list. It's interesting to note how distinctly named these restaurants are to a specific category such as spice, honey, and beer.

**Top Rated Eateries in Charlotte**
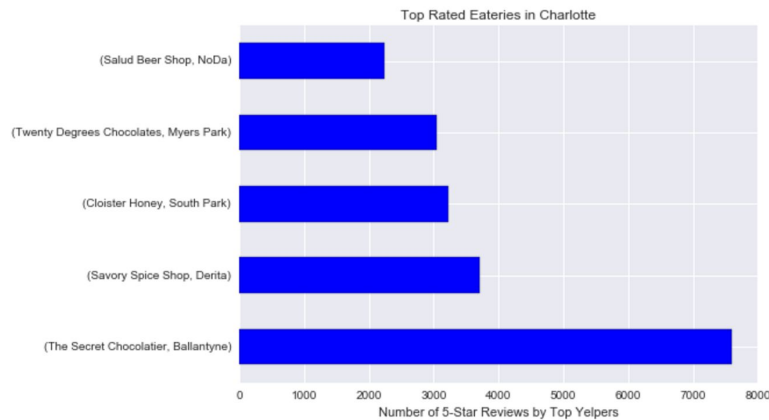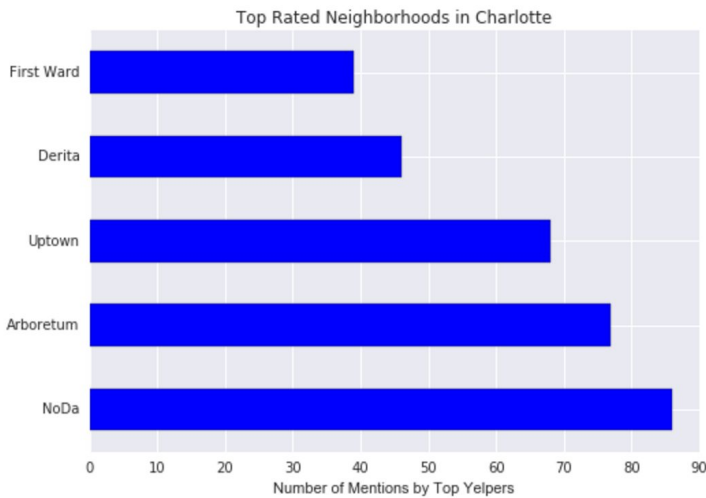
Phoenix

Phoenix is by far the largest city we surveyed, at over 500 square miles. In this particular data set, Yelp does not split Phoenix into neighborhoods, so there is no graphical representation of top neighborhoods. One possibility for this lack of distinction is that places in Phoenix that are considered neighborhoods, such as Scottsdale or Tempe, are technically separate cities.

However, we were still able to tease out the top rated eateries, with Little Miss BBQ taking the number one spot by a landslide. Little Miss BBQ currently has a 5 star rating with over 4,000 reviews, and is renown in the area for it's Southern Style BBQ.

Top Rated Eateries in Phoenix
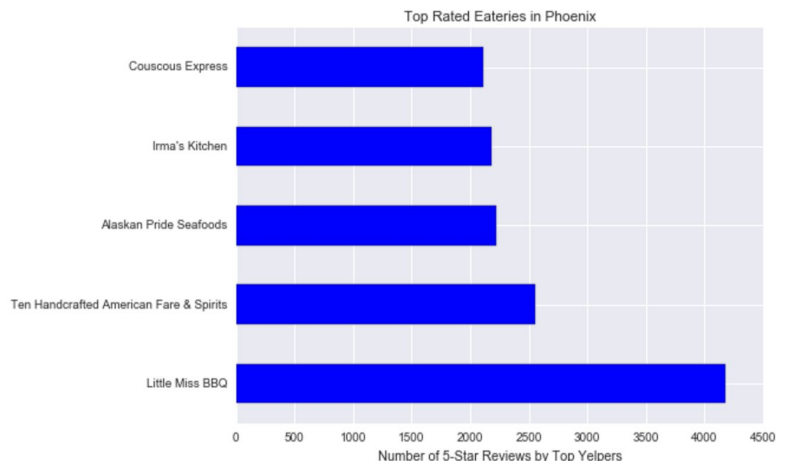
4) Which city has the best nightlife based on late-night checkins?

Las Vegas is known for its nightlife, and a basic query of the check-in data backs up that assumption. The graph below details a city-by-city breakdown of the total number of late-night check-ins at bars and restaurants. A check-in was deemed "late-night" if it happened between the hours of 11 p.m. and 4 a.m.
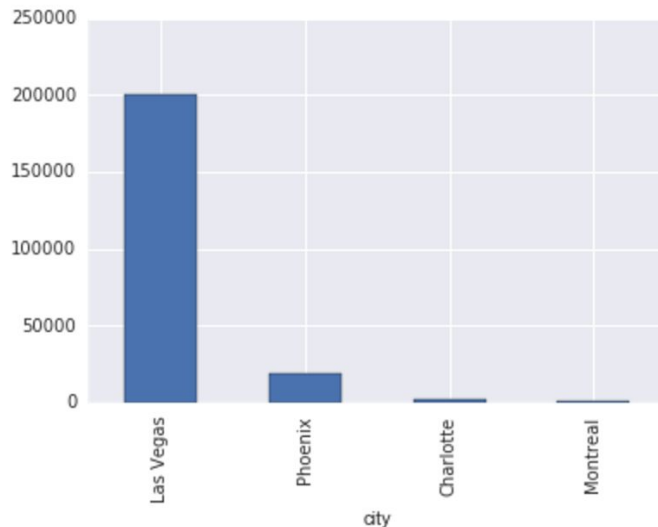
No surprise, Las Vegas has the most active nightlife, based on the total number of late-night check-ins on Yelp.

However, overall usage of the Yelp app may explain some of the discrepancies here. The app may not be as popular for users in the other cities, so looking at the data proportionally provides better insight.
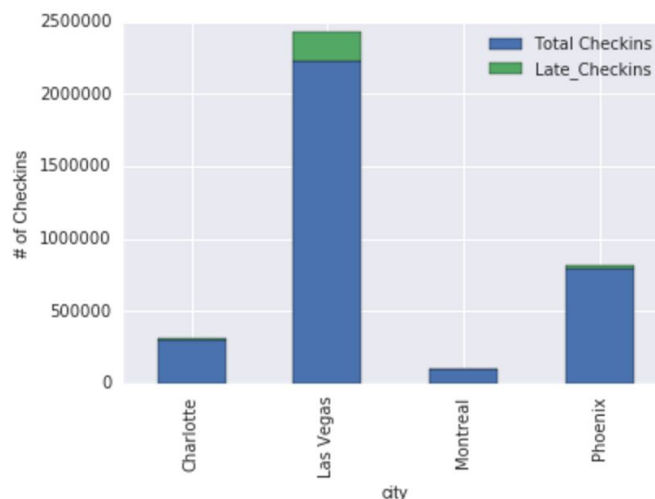
As evidenced in the graph above, Las Vegas maintains its late-night dominance, with the highest proportion of late night check-ins. Montreal and Charlotte had significantly smaller numbers of check-ins to begin with, and a very tiny proportion of late-night check-ins. In fact, the number of check-ins in Las Vegas almost equals the total number of check-ins (regardless of daypart) in Montreal, Phoenix, and Charlotte put together!

The graph above makes clear the significantly different behavior in each city, with the presence of late-night checkins almost invisible for Charlotte and Montreal. The table below detailing the proportion of late-night check-ins better illustrates the behavior happening in those cities.

City	Late Night Proportion
Charlotte	0.87%
Las Vegas	8.9%
Montreal	1.33%
Phoenix	2.37%

While predicting that Las Vegas would have the most active nightlife was an obvious choice, perhaps the most useful insight in this analysis is that Charlotte is simply not a place where people party all night long and make it known on Yelp.

5) Which city has the most kid-friendly restaurants?

Due to significantly differing volumes of businesses per city in the Yelp dataset, the comparison of kid-friendly places is done proportionally as well, examining the number of kid-friendly restaurants in relation to total number of restaurants.

It was hypothesized that Charlotte or Montreal might offer the most kid-friendly options for dining. Exploration of the business data revealed a different answer.
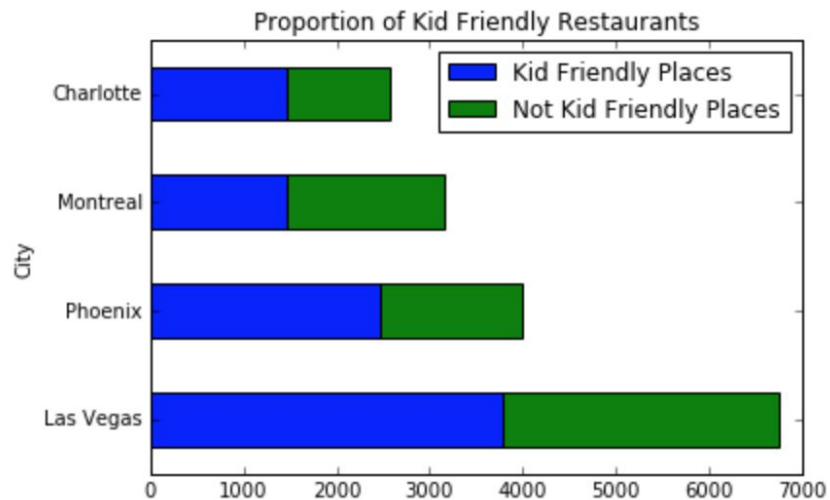
Phoenix wins the award for the most "kid-friendly" city, with 62% of its restaurants deemed kid-friendly on Yelp. Even more surprising, rounding to a tenth of a percent, Charlotte is only slightly more kid-friendly than Las Vegas (57% vs 56%) and Montreal appears to be the least kid friendly place at 46%.

The table below details the specific percentages represented in the graph above.

City	Proportion Kid-friendly
Las Vegas	56.2%
Phoenix	61.9%
Montreal	46.1%
Charlotte	57.2%

While the results defied the initial hypothesis, they did provide valuable insights.   Most intriguing were the results from Las Vegas, where 56% of restaurants on Yelp were deemed 'kid-friendly.' This is a surprising number, as the initial look into Las Vegas Yelp data revealed a vibrant nightlife that didn't seem suited to kids.

Conclusion

The analysis of the four selected cities in the Yelp dataset confirmed some logical assumptions, such as Las Vegas's active nightlife, and revealed some less predictable elements, including Phoenix's abundance of kid-friendly restaurant options. Hopefully the findings in this project help travelers to Las Vegas, Phoenix, Montreal and Charlotte make plans based on deeper insights than they would find solely through the Yelp app or website.