

Wendy Phillips

May 6, 2019

1. INTRODUCTION

The American Pet Products Association (APPA) calculated that the total pet industry expenditures in the United States in the year 2018 was over 72 billion dollars¹. As of 2011, almost $\frac{3}{4}$ of U.S. households included a pet. What is more, pet expenditures are more resilient to economic downturns than many other household expenditures², making the pet industry a business market with high potential.

A company with a limited budget that plans to launch a new product, begin a marketing campaign, or open a new business location might would benefit from determining a narrow range of locations best suited for that purpose.

The cities of the United States are a diverse mix and may have distinct characteristics in their cultural relationship to their pets. One way to study these characteristics is to look at the number and relative abundance of different types of pet businesses in different communities. Another measure of interest would be the overall density of pet related businesses. Yet another measure might be the relative frequency of participation in internet reviews of pet related businesses.

2. DATA

To address these questions, I used recent statistics on U.S. cities such as population and land area³. I focused on cities with populations ranging from approximately 400,000 to 1,000,000 people with the idea such data will be of interest to companies focusing on mid-level markets. To address questions of pet business abundances and consumer participation in internet ratings, I queried data on pet businesses from both Foursquare and Yelp.

3. METHODOLOGY

I had two dataframes as a source for analysis: one with statistics on the 38 U.S. cities with populations ranging from ~400,000 to ~1,000,000 and a second with venues related to pets in those cities.

The venues could be listed in more than one city and under more than one category. To simplify analysis, I reduced the data set to only a single entry per venue. I first randomized the order of the rows. Then I kept only the first entry row for any given venue so that venues that the city and/or category were randomly retained for those with multiple entries.

¹ https://www.americanpetproducts.org/press_industrytrends.asp, accessed May 6, 2019

² <https://www.bls.gov/opub/btn/volume-2/spending-on-pets.htm>, accessed May 6, 2019

³ https://en.wikipedia.org/wiki/List_of_United_States_cities_by_population, accessed May 6, 2019

I did exploratory analysis of the frequency of different venue types in the cities, how many venues are listed under more than one category, what the most frequent venue type is for each city, and how many total venues each city has.

I performed three clustering methods, all of which produced similar results. I then mapped the clusters by color onto the map of the U.S. using folium.

4. RESULTS

There was a very large range in the total number of venues per city, with Mesa having only 4.5% the number of venues as Long Beach (Figure 1). The number of pet venues were not highly correlated with population size of cities (Figure 2).

Figure 1. Number of pet related venues in 38 U.S. cities

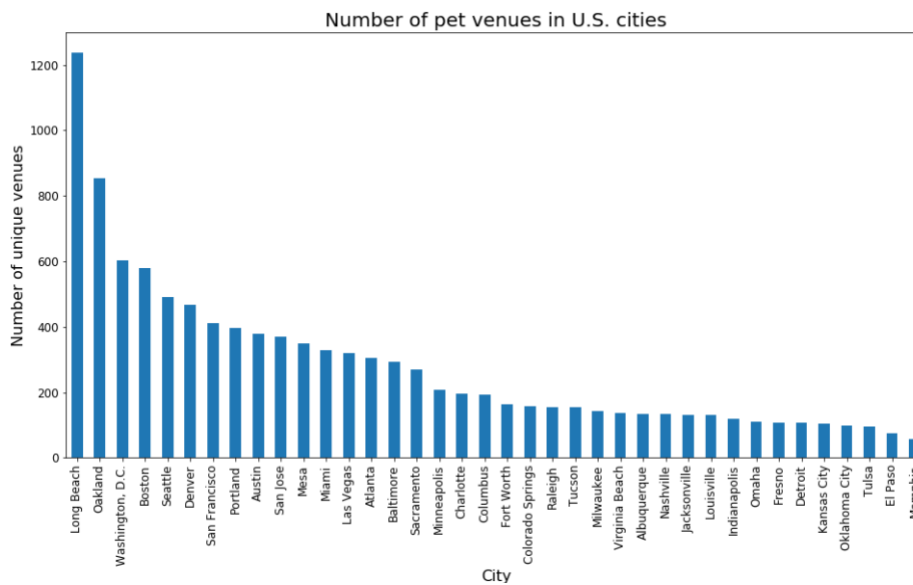
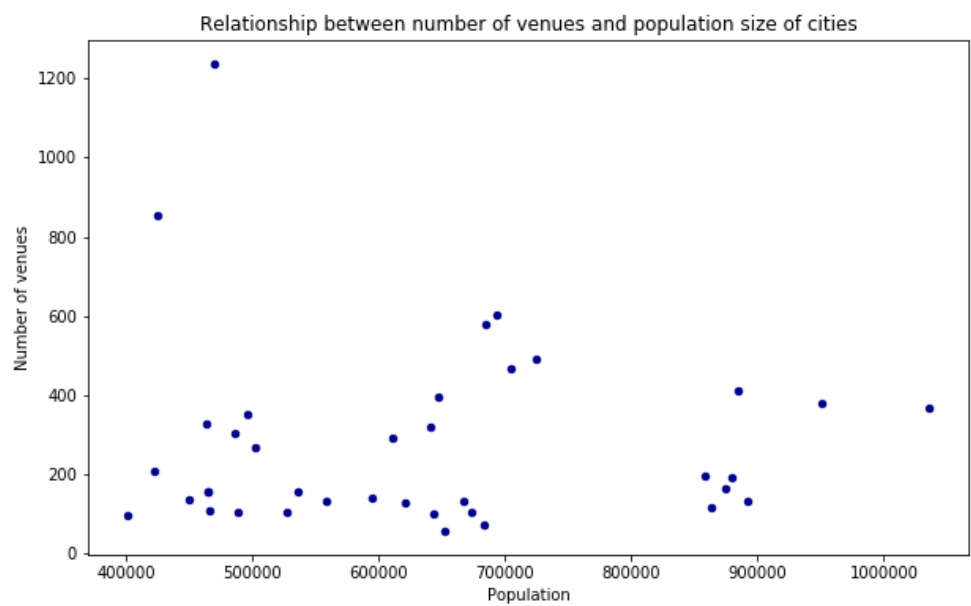


Figure 2. The number of venues vs. the population of 38 U.S. cities



The pet venue related activity can be measured by factors such as cities with the most venues per person and the most reviews per person and per city. Another measure of pet culture is the cities with the highest rated venues. A few cities are consistently at the top in most of those rankings: Long Beach, Oakland, Seattle, and Washington, D.C. (Table 1).

Table 1. City rankings by different calculations of pet venue related measures

Rank	Population density	Reviews per city	Mean venue rating per city	Venues per person	Venues per square mile	Reviews per person	Reviews per venue
1	San Francisco	Long Beach	San Francisco	Long Beach	Long Beach	Long Beach	San Jose
2	Boston	Oakland	Oakland	Oakland	Oakland	Oakland	Long Beach
3	Miami	San Jose	Seattle	Washington, D.C.	Boston	Washington, D.C.	Las Vegas
4	Washington, D.C.	San Francisco	Long Beach	Boston	Washington, D.C.	Las Vegas	Oakland
5	Long Beach	Washington, D.C.	San Jose	Miami	Miami	Sacramento	San Francisco
6	Seattle	Las Vegas	Austin	Mesa	San Francisco	San Jose	Sacramento
7	Minneapolis	Seattle	Mesa	Seattle	Seattle	Seattle	Seattle
8	Oakland	Boston	Portland	Denver	Minneapolis	San Francisco	Austin
9	Baltimore	Sacramento	Sacramento	Atlanta	Baltimore	Mesa	Washington, D.C.
10	Milwaukee	Austin	Denver	Portland	Denver	Boston	Portland

Clustering of the cities based on relative pet venue frequencies was performed by three different methods, all producing similar results. Clustering using k-means with 4 clusters resulted in groups with distinct patterns (Table 2). Cluster 0 had the most evenly distributed venue type frequencies. Cluster 2 had a much higher percentage of groomers than any other cluster. Clusters 1 and 3 both had a high frequency of pet sitters, with cluster 3 also having the highest number of dog walkers (Table 2, Figure 3).

Mapping the clusters onto a map of the United States revealed observable patterns (Figure 4). The cities with the most pet sitters, clusters 1 and 3, were all in coastal states. Cities with the most groomers, cluster 0, were largely in the southern half of the U.S., with the exceptions of Lincoln and Milwaukee.

Table 2. Mean frequencies of venue types by cluster assignment

Venues	Cluster0	Cluster1	Cluster2	Cluster3
Groomer	23.8%	18.0%	31.1%	12.1%
Pet sitting	21.7%	29.8%	19.0%	27.9%
Pet store	15.3%	11.8%	19.6%	9.2%
Pet training	12.5%	12.7%	9.1%	12.5%
Dog walkers	8.5%	14.2%	4.9%	24.1%
Pet photography	3.4%	2.7%	2.6%	3.6%
Dog Run	3.3%	4.0%	4.6%	2.4%
Pet cremation	1.9%	1.6%	1.5%	0.8%
Pet transport	1.5%	0.7%	1.4%	0.8%
Pet waste removal	1.5%	1.3%	1.0%	0.6%
Animal holistic	1.2%	0.7%	0.0%	0.6%
Pet hospice	0.9%	0.5%	0.9%	0.7%
Animal physical therapy	0.4%	0.3%	0.0%	0.8%

Figure 3. Stacked bar graph of the number of pet venue types of the most abundant categories in 38 U.S. cities. Bars are color coded by venue categories. The green, red, blue, and purple horizontal lines between the bars and the city names indicate clusters (k-mean) 0, 1, 2, 3, respectively.

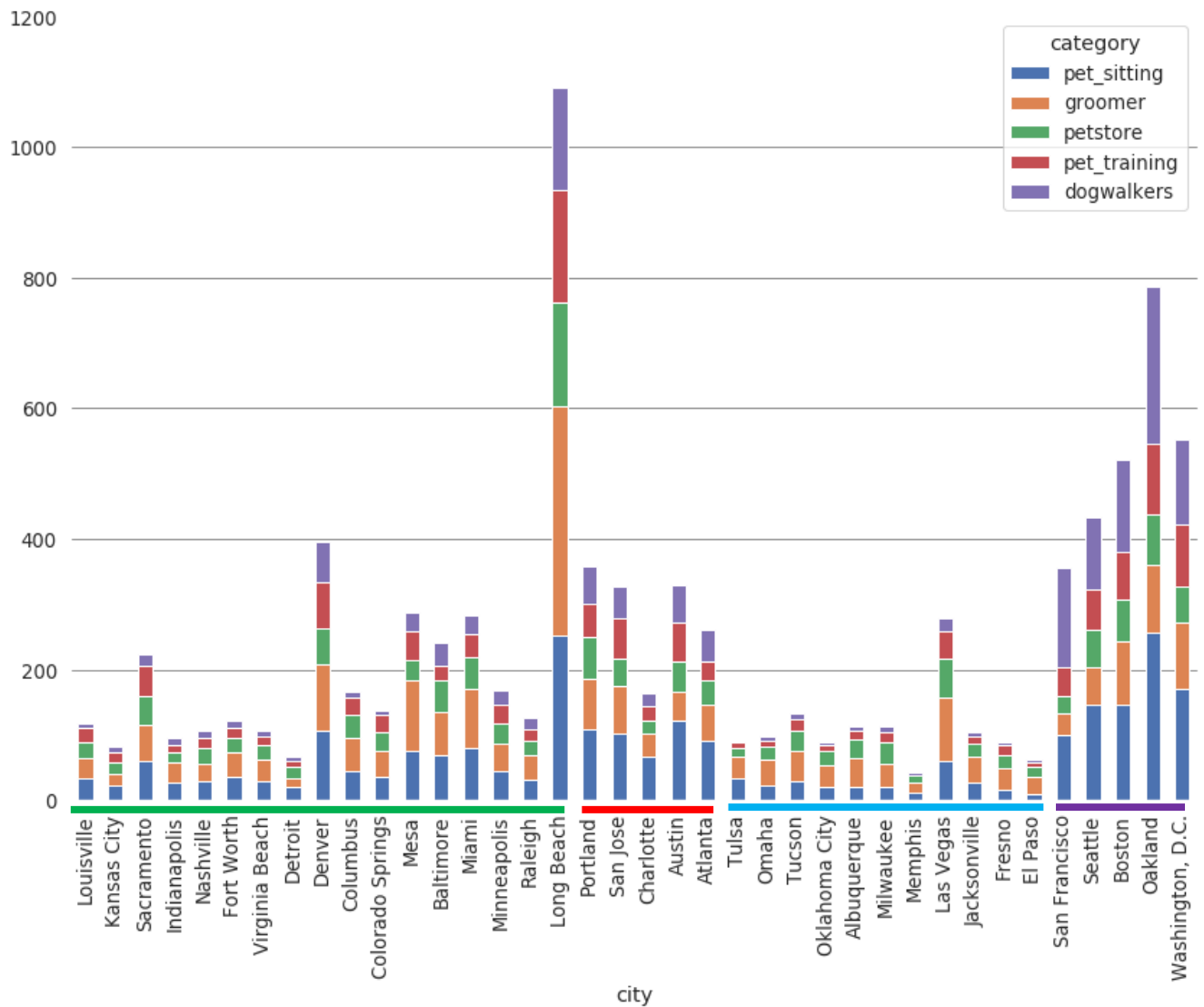
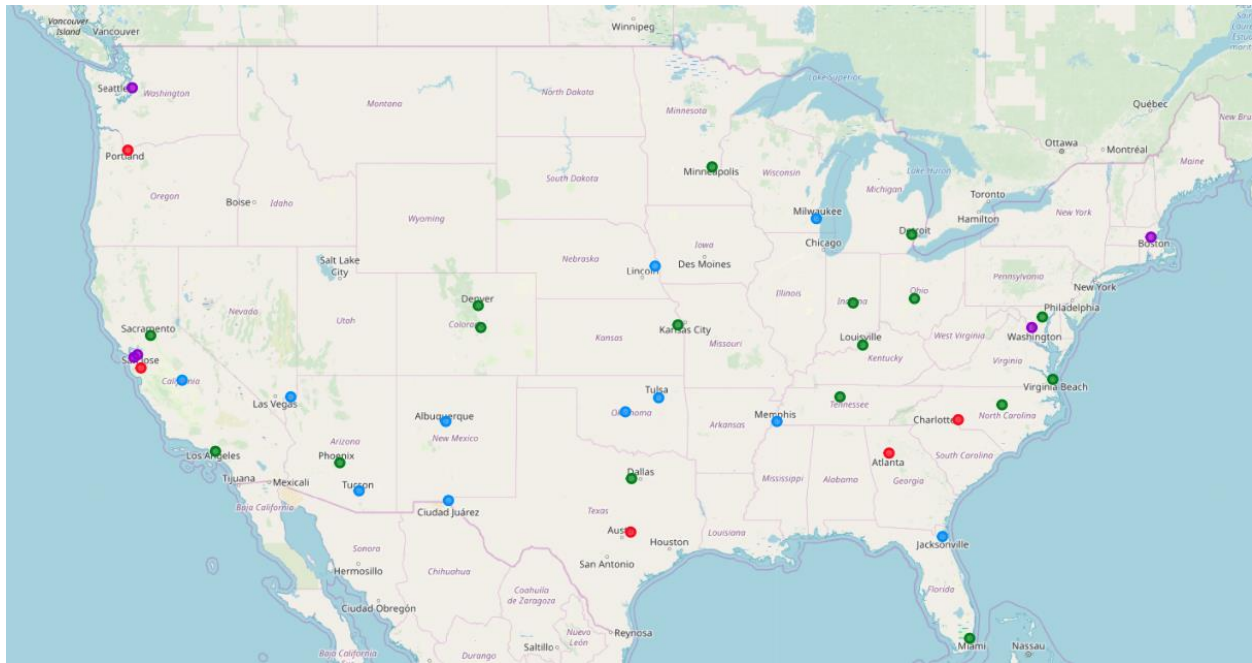


Figure 4. Mapping of clusters onto U.S. cities. Clusters are represented by four colors: green, red, blue, and purple for clusters numbered 0,1,2,3, respectively



5. DISCUSSION

The most abundant venue types for all cities but one are either groomers or pet sitters. The exception is San Francisco, where dog walker is the most abundant venue.

There is an extraordinary more than 20-fold difference between the least amount of venues per city, at 56 in Mesa, and the most, at 1238 in Long Beach. One factor to note is that Yelp does not provide all venues but only those with at least one review. Therefore, it is possible and even likely, that the numbers of venues are not entirely accurate. However, it seems unlikely that the large variation in venues per city could be explained completely by differences in review submission rates by city in Yelp. Given much of the data is meant to measure pet related activity, this information is still useful.

When considering where to focus business efforts, in some cases it may make sense to look for cities with the most venues by land area or population, either total or in particular categories of interest. In other cases, it may be more relevant to look for cities where a particular venue type is more abundant regardless of overall venue numbers.

Cities could be clustered into 4 groups, with primary differences in their frequencies of groomers, pet stores, pet sitters, pet trainers, and dog walkers. Interestingly, five of the seven cities with the highest total venue number clustered together in one group when clustering was done solely on relative frequency of venue types, not total numbers. Those five cities (Boston, Oakland, San Francisco, Seattle, and Washington, D.C.) had the highest frequency of dog walkers by almost double the next highest

cluster and also had a high frequency of pet sitters. Overall, pet sitters and dog walkers made up more than 50% of the pet venues in those cities. In contrast, cluster 2, which contained mostly cities in the southern half of the U.S. (with the exceptions Lincoln and Milwaukee), had groomers and pet stores making up more than 50% of their venues.

One confounding factor is that ~40% of the venues were listed under more than one category or in more than one city. This was addressed with the simplistic method of randomizing entries and saving only the first occurrence of a venue for subsequent analysis. A more thorough approach for eliminating duplications in different cities would have been to separate by zip code or latitude and longitude. Venue types were frequently linked, in that a venue listed as pet sitter was more likely to also be listed as a dog walker. Likewise, venues were often categorized as both pet stores and groomers. Further analysis could be done to characterize more fully the depth of those correlations. The randomized selecting of only one entry per venue meant that the category assigned to a venue was randomized if it had multiple categories.

6. CONCLUSION

The cities of the United States vary a great deal in their pet related cultural characteristics. Determining the primary characteristics of different cities could be helpful for business entrepreneurs. Depending on their business objectives, they may want to focus their efforts on a particular subgroup of cities to stay within budgetary limitations.

I chose to focus on the 38 U.S. cities with populations between approximately 400,000 and 1,000,000 people where a business could focus on a moderate market size. Using queries of pet related venues from both Foursquare and Yelp, various measures were used to classify cities.

Several cities were regularly at the top of the lists of most active pet communities judged by rankings such as venues per person and reviews per venue. Those cities included Oakland, San Francisco, Long Beach, Seattle, and Washington, D.C.

Cities could be clustered into 4 groups, with primary differences in their frequencies of groomers, pet stores, pet sitters, pet trainers, and dog walkers. Altogether, this information could be used to choose the best group of cities for the business or product one wants to launch.