**Introduction**

**Background**

With the rising trend in science fiction and the increased knowledge in science itself, it is only a matter of time until a zombie breakout occurs. When it comes to disaster scenarios, many people have leanings into three categories. First group are the survivalists who prefer making a bunker to live out their days. Second are the runners, those who hop from place to place when an area gets too crazy. Third category holds those that accept the doom and want a spectacular end.

**Problem**

Since the zombie apocalypse is the inevitable outcome, deciding what city to live in can be crucial to survival or preferred ending. In order to make an informed decision for the individual, a comparison of different cities and what they have to offer is important. There are many cities and each have different distributions of resources. These resources need to be categorized in a meaningful way so an individual can choose which city will suit their needs at the end.

**Data**

### Data Source

The Foursquare database will be utilized to look at different venues in the five largest cities in the United States.

### Data Selection

Several venue features will be important to categorizing each city into Survivability, Easy Escape and Spectacular cities. For these three city types, an idea of the key venue features is listed below.

1. Survivability
   1. Pharmacy
   2. Gym
   3. Shopping Mall
2. Easy Escape
   1. Gas Stations
   2. Train Stations
   3. Car Rentals
3. Spectacular
   1. Bars
   2. Theater
   3. Bookstore

**Methodology**

**Exploratory Data Analysis**

In order to retrieve venue information from the five largest cities in the United States, a data frame with those cities and coordinates needs to be made. These features can be found on a Wikipedia page and manually entering the values is less code than scraping the table therefore the values were manually entered. Initially one dataframe was made with all the cities location data. This later became a problem when trying to find individual venues associated with a given city so each city had its own dataframe.

Next the venue data was acquired from Foursquare and iterated through each city dataframe. Initially the Foursquare call returned fewer venues than expected. It was determined that the city coordinates for the city correspond to the city center therefore the radius passed in the beginning was not large enough to encompass the city.

Once the venue data is stored in the city dataframes, the information needed to be separated based on venue category. A function was created to break out the venue categories into a new dataframe for each city. The venue categories where then categorized based on city into the following categories: Food, Transportation, Survival and Good Time.

Food is essential for survival regardless of the individual’s apocalypse strategy. The venues that include food of any sort are included in this category. The food category is the largest for each city and was not used as a direct measurement for the three main categories rather a supplement for deciding between cities with similar resource distributions.

Transportation is the main driver for those in the Escape category. Venues dealing with transportation such as gas stations, rental car places and train stations were sorted into the Transportation category for each city.

Survival includes the venues best suited towards surviving in the city. The venues that include gyms lots of different stores and places suitable for conversion into a bunker were sorted into the survival category.

The Good Time category is for those who wish a spectacular end and are willing to risk turning into a zombie as long as they were entertained first. The category was populated with museums, galleries, bars and other similar type venues.

Once each city venues were categorized, a single dataframe was created to house the raw counts for each category in each city. From this final dataframe, the data was analyzed and visualized in order to determine which city to live in depending on apocalypse preferences.

**Results**

|  |
| --- |
|  |
| Figure 1. Dataframe of the total number of venue types for the different categories. |

|  |
| --- |
|  |
| Figure 2. Bar graph of the total venues in each category for each city. |

|  |
| --- |
|  |
| Figure 3. The food category will be supplemental information for the three main categories considering food is needed for survival regardless of apocalypse approach. |

**Survivalists**

The Survivalist specific venues display Chicago has the most venues geared towards survival during the zombie apocalypse.

|  |
| --- |
|  |
| Figure 4. The bar graph above displays the number of venues in each city that are useful for surviving in the city. |

**Transportation**

There is a three way time for transportation venues in Chicago, Houston and Phoenix. The food category will be used to decide which city has the easiest escape survivability.

|  |
| --- |
|  |
| Figure 5. The bar graph above displays the number of venues that are useful for escaping the city. |

**Good Times**

New York has the most venues with an entertainment focus making it best suited for those looking for a good last time.

|  |
| --- |
|  |
| Figure 6. The bar graph above displays the venues that are useful for having a good time before being eaten by zombies. |

**Discussion**

For the survivalists, Chicago has the greatest number of venues geared towards survival. Phoenix has the second most venues for survival but lacks food venues which would impact the overall ability to survive. Chicago is the clear choice for survival.

For the runners, there is a three way tie between Chicago, Houston and Phoenix for transportation venues. Houston has the greatest number of food venues of the three making it more preferable than the other two. In addition, Houston has more entertainment options so it is possible there will be higher traffic to Houston allowing for additional escape options. Houston is the ideal city for a quick escape from zombies.

For those that want a good time before the end, New York has the most venues for entertainment. New York also has the most food venues making the city best suited for those searching for a spectacular end.

Phoenix has the most diverse selection of venues that could accommodate any of the three categories. Being the smallest of the cities leave Phoenix with a lack of venues but it does have the most diverse venues. Los Angeles is similar to New York and has mainly food and entertainment venues. Los Angeles also has better survival venues than New York. The city would be suited for those wanting a good time and deciding later they want to survive.

**Conclusion**

The three people groups in the zombie apocalypse are survivalists, runners and good timers. The survivors should find a home in Chicago to increase their chances of survival based on the venues in Chicago. Runners should start their journey in Houston from where they have food readily available and easy access out of the city. The good timers should park it in New York and enjoy the many options for entertainment until the end comes. Los Angeles and Phoenix are good for mixed camps of thinking. The program could be applied to several different cities for greater statistics and characterization of zombie apocalypse friendly cities.