1. **Introduction**

An entrepreneur is interested in opening a pet store in Toronto. They would like to determine what is the best neighborhood for the pet store. The entrepreneur does not want to simply open up the pet store by looking at neighborhoods that currently have a high number of existing pet stores as the supply in those existing neighborhoods may exceed current demands. Therefore, the goal is to identify neighborhoods that have characteristics that are suitable for pet stores but currently is undersupplied. They believe that the best indicators for identifying the location are the type of venues that are situated in a particular neighborhood. In particular, it is assumed that the types of venues in a particular neighborhood act as a proxy to represent the local demographic, including their needs and interests. In this project, we will determine which type of venues are often established in conjunction with pet stores. Once we understand what the top type of venues are, we will proceed to locate neighborhoods that have these venues as potential targets for the new pet store.

1. **Data** 
   1. We will use data from Wikipedia to define the neighborhoods in Toronto.
   2. We will use foursquare data to build a list of the venues in each neighborhood in Toronto.
2. **Methodology**
   1. In the first stage we will attempt to determine the relationship between Pet Stores and other types of venues. This is accomplished by first counting the number of venues for each type of venue in each neighborhood in Toronto. We will then calculate the correlation between each type of venue with Pet Stores. The top 5 venues with the highest correlation to the instances of Pet Stores will therefore be identified as the top most related venues, in other words, these types of venues represent the demographic who are likely suitable customers of our new pet store.
   2. In the second stage, we will identify the neighborhoods that fit the characteristics above best. This is accomplished by the counting the instances of the top venue types as determined above, and applying a weighting to that count. The weighting is the correlation calculated above. The weighted count for relevant venues is stored as the “weighted score” for each neighborhood in Toronto. The weighted score will be sorted to identify the neighborhoods that best fits our targeted characteristics.
   3. In the final stage, we will examine the top neighborhoods and eliminate the neighborhoods with a high number of existing pet stores. The remaining neighborhoods will then be the suitable neighborhoods for the new pet store.
3. **Results**

W

1. **Discussion**
2. **Conclusion**