Introduction/Business Problem

One interesting question facing a variety of businesses across markets involves the nature of major cities. In the current economy, many companies and businesses serve many major cities throughout the United States and even the world. For this reason, major U.S. businesses need to know about the major markets and cities. Further, as cities are vastly different based on region, location, size, and other demographic factors, it is important to know good information about each city beyond basic facts. A major company may be successful in some cities based on demographics and culture but may not be a good fit in another location even if that location is in close geographic proximity.

For this reason, my goal is to cluster major cities in the United States and Canada based on the types of businesses which exist in that city. Rather than simply a population or geographic approach, this provides another way to assess the similarity of cities. It will attempt to cluster cities which have similar composition of business types and demonstrate to companies which cities are similar based on the types of businesses and industries which flourish in those cities and metro areas.

I was able to gather 128 major cities in the U.S. and Canada from infoplease.com and based on population. Using foursquare data after cleaning up the latitudes and longitudes, I will be able to assess the types of businesses in these 128 metro areas and then come up with a similarity clustering to form another tool which businesses can use to assess the similarity of cities. Then, these businesses can examine which cities they are likely to thrive in, which they may not, and how to improve in the areas which they are not thriving if they'd like to enter new markets which may have different characteristics.