**Data**

In order to find the most suitable city for Joy, a list of cities with their corresponding zipcodes was downloaded from http://www.downloadexcelfiles.com/us\_en/download-list-zip-codes-washington-state in csv format.

Venue queries will then be made by neighborhood name/cities using FourSquare Api. The resulting data regarding venue category will be used to find similarities in the corresponding cities. With a proper list of top venue categories for each state, a cluster can be made grouping them together for easier understanding.

These are the data needed for the success of this project:

* Cities and Zip codes in Washington State
* Coordinates of the cities
* Top Venues for each city
* Categories of Venue

**Methodology**

As a result of the nature of the problem at hand, we will make use of k-means clustering to properly group the cities based on their most in-demand venue.

A one-hot encoding will be the done on the data and then grouped by city. The encoding will return venue categories as a column per city which will then be further grouped. The encoded dataframe will further be filtered into the top venues and then the k-means algorithm will be passed over it. This will return a cluster of cities. Then an adequate recommendation can be given to Joy.