

# CLUSTERING THE NEIGHBORHOODS OF DATA SCIENCE COMPANIES IN US

## Introduction- Business Problem:

As I am interested in data science jobs, I would like to choose **clustering the neighborhoods of Data Science companies in US** as my Applied Capstone Project. The basic idea is using the dataset of all the companies in US and location dataset provided by **FourSquare API**, we can **cluster the location of different domain companies in US**. We can also cluster the neighborhoods of top venues around all the data science companies in US using FourSquare API.

The main problem faced by all the jobseekers is **relocation**. Since we don't have proper dataset of the facilities in the new location, we will always feel uncomfortable and sometimes jobseekers will be frustrated by searching for too many venues around the new place. We should love the environment where we are going to work, so that we will do the work with pleasure.

Our main objective of this project is to find out the locations of all the IT companies in the US and cluster them based on the domain of the company. Then based on the input provided by the jobseeker, we can filter the neighborhood top venues of the selected domain companies in US. In my project, I am going to use Data Science companies as my desired input and I will focus more on clustering the top venues around the selected data science companies. We can also create a **map** to visualize the locations of the different domain companies and the neighborhood top venues of the selected domain companies in US.