**Introduction/Business Problem:**

I am an employee of Siemens Limited which has 17 offices in India. Each office is either situated in an exclusive Siemens building or in a business center (which means that this building houses the offices of many other organizations besides Siemens). Naturally, each office is situated in a location wherein a huge amount of employees work. I want to identify and analyze 100 venues which are nearest to each office. The analysis will be followed by a k means clustering to identify various clusters. Based on the clusters obtained, I will try to classify each and over office on its popularity and the kind of venues situated around it.

**Data**:

The 1st piece of data is the location of each of Siemens’ office in India. This will be uploaded in CSV format as shown below.

|  |  |
| --- | --- |
| **Office City** | **Office Location** |
| Bangalore | St Marks Road, Bangalore |
| Hyderabad | Saifabad, Hyderabad |
| Visakhapatnam | Allipuram |
| Chennai | Anna Salai |
| Coimbatore | P N Palayam |
| Mumbai | Airoli |
| Ahmedabad | Navrangpura |
| Vadodara | Maneja |
| Pune | ICC Trade Tower, Pune |
| Nagpur | Ramdaspeth, Nagpur |
| Delhi | Sector 18, Gurgaon |
| Chandigarh | Sector 17c, Chandigarh |
| Jaipur | Ashok Nagar Jaipur |
| Lucknow | Gomti Nagar |
| Kolkata | Acropolis Mall, Kolkata |
| Bhubaneswar | Patia, Bhubaneswar |
| Jamshedpur | Bistupur |

Using geocoder, geographical coordinates of all the 17 places will be found out. Then using these geographical coordinates and Foursquare data, I will identify the venues located in and around these office locations.