1 . Introduction/Business Problem

1.1 Background

A restaurant is looking to expand its business by opening a new chain in New York City. The restaurant serves food that is of a variety of cuisines. This has led them to figure out which categories of restaurants are available around Times Square. This is in order to view how to better create their menu to serve their business needs and to create competition. The chain has to be close to Times Square or in the next to it. And to figure out how it will be received.

1.2 Problem

To solve the problem for this business we would use Foursquare in order to view the best location where the business could start operating and where most of the work force are and view each area and the number of restaurants that location has. Having the ability to figure out the suitability of the location the business will have a clear lead gap compared to business that may not use the same methods.

1.3 Interest

The interest for this would be the business that would like to open up in the City of New York. This is to satisfy its business plan and projections. Other business who have similar interests may also find it interesting.

2. Data

2.1 Data Source

The data that will be used is from Foursquare that is used by people around the globe to add their reviews and locations of places of interest. We will be retrieving all the restaurant locations from this site. The data will also allow us to view the areas in which restaurants may not be as much to another place. And as well as places in which there appears to be a popularity amongst locals as well as tourists.

2.2 Data Cleaning

Before retrieving any data, I had to get the coordinates for Times Square. From here on out I accessed the Foursquare API and connected to it to retrieve all places with search query ‘restaurant’. Therefore, I retrieved all data associated with this request. I then cleaned by filtering it. This is by removing ‘location.’ before any word. I continued to delete columns that weren’t really necessary such as ‘neighborhood’.

2.3 Feature selection

After cleaning the data, the size of the data frame was around 135. I then continued to view the data frame for any errors that might be presented. I continued to remove any rows that had no values in them.

3. Exploratory Data Analysis

3.1

Due to the nature of the project not much was needed to be done for calculation. I continued on to display the types of restaurants that are around Times Square this is for that the requirements of the problem could be satisfied. Upon doing this a map was created and displayed and had all the categories

of food that were around the Times Square in order for the client to decide on the direction they would take.

Further upon this a chart was designed with all the details that has the retrieved categories of the restaurants that are around Times Square. A bar graph was plotted show the number of cuisines each category has in that area.