1. A description of the problem and a discussion of the background.

With broader application of data science arising from new capabilities and ideas led to existing businesses facing another level of threats. Start-ups are keener to leverage new technologies and innovative business models consistently looking for ways to impact customers in such a way that they influence customers on the products they are choosing and how their chosen product better impact their lives.

To remain competitive existing businesses have embraced use of location data to identify consumer patterns, tastes, and trends. They are readily available these days from several social media platforms including Google maps, Facebook check-in service, Yelp.ca platform, location review on sites like TripAdvisor and Booking, and Foursquare among other location data providers.

Working as a data scientist, one of customers have asked for consulting as to their business strategy involving data science technologies to come up with business models to keep them competitive as a start-up restaurant. Location is one most important factor when considering to opening a new restaurant and it is strongly believed that data services using data science technology will give them advantages.

1. A description of the data and how it will be used to solve the problem.

The start-up intends to collect location data from Foursquare and apply data science techniques and tools. The data collected will involve comparison of two locations to determine which is the best location to recommend to a customer.

1. Clustering regions of distinct post code by the number of restaurants within specific radius will be done by using K-means clustering to group regions into “high number of restaurants with high review ratings”, “high number of restaurants with low average review ratings”, “low number of restaurants with high review ratings”, and “low number of restaurants with low review ratings”.
2. Population data for each region with distinct post code will be used to in conjunction with the number of restaurants within each region of specific radius to give priorities of each region.
3. Reviews from each sub region of the city will be evaluated from data collected from Foursquare and determine their ratings and number of potential customers estimated from the size of reviews.
4. Using all of analysis done above, customer will be advised of best type of restaurants for each region, with number of competitors around the region so ultimately give them insights on their choice of restaurants, its best location as a final answer.