**Yelp Review Project**

Data for this project was collected: <https://www.yelp.com/dataset_challenge/dataset>

We have splitted the project into 4 files for better execution:

Yelp\_Project\_1:

1. We start with getting a regression between the sentiment score and the Business rating
2. We then create a network graph on the sentiment of the reviews

Yelp\_Project\_2:

1. Taking up a general notion of Location influencing the star ratings and testing it. Finding any correlation between distance of a business from a popular location.
2. Once we find the effect of a location on a business, we try to find average rating in different zones of the city. The plot explains the average rating along with the density of businesses in that zone.

Yelp\_Project\_3:

1. Finding Influencer of a Business for targeted marketing.
2. Recommendation of business based on user’s past star ratings given to business. [We have filled NA’s with mean and if it has only one review we have replaced ]
3. Finding the correlation between stars given and length of the review given.

Yelp\_Project\_4:

1. What are the most common type of restaurants in the state
2. Does the length of review determine its usefulness

Shiny App: Server.R, ui.R, app.R

**Note: All the files submitted were running on minimum data inorder to knit them, So, outputs might not align with the presentation. Please run them based on hardware capacity.**