Data Science Project
Stage 1

Objective -

We want to analyze the two structured data sets that contain information about various restaurants and also the restaurant review text documents to find out the influence of the various attributes on the rating of the restaurants.

Data Sources –

Zomato.com
Yelp.com

Data Extraction –

Used scrapy web crawling library in python to crawl the web pages and extract the information of various attributes.

Text Documents -

Text documents consist of restaurant reviews and will be used for sentiment analysis.

Open Source tools –

scrapy