

Team Information			
Course	Statistical Methods in Text Analysis		
Term	Spring 2019	Course #	LING 583
Member	Glen Heungyong Im	Red ID	823395723
Primary Dataset	Yelp Open Dataset 2019		
Project Abstract			
Executive Summary	Identifying most frequent words in the review data by each restaurant category		
	Analyzing which words are most related with the star rate by each restaurant category.		
	Identifying frequently used words in tips and reviews by years if there is a change in language.		
	Predicting stars from reviews.		
Data Description	The dataset has 4 different data.		
	- Review: cumulated reviews throughout the years 2011 - 2018		
	- Tips: cumulated tips users tagged in each business		
	- Business: all kinds of business place with their essential information and geographical, multi-categorical, and star-rate data		
	- Check-in: the time stamp logging when a user posted that they checked in to a specific business place.		