**Project Proposal**

Team: Data Explosion

Yiqing Liu

yiqing5@illinois.edu

**Project Summary**

Do a research about Yelp users’ tastes, purchasing habits and try to draw some interesting conclusions!

**Data**

A trove of reviews, businesses, users, tips and check-in data from *Yelp*

Size: 3GB

Source: Kaggle

**Project Description**

1. Description

Our project is going to be centered around Yelp. We will use the data from Kaggle to keep track of user tastes, purchasing habits and try to draw some interesting conclusions. (like draw a portrait of consumers, where are best restaurants )

1. Usefulness

“What to eat…” ,at least for me, it’s a big question.

Maybe I will know what to eat after the project

1. Realness

Our data is from Kaggle. And we can use Python to grab some other data as supplement.

1. Functionality

basic functions: use some basic visualization methods to represent data

advanced functions: more attractive way to present our conclusion

**Project Plan**

1. data collection **(completed)**

obtain the data, I collected the data from Kaggle. If it is needed for further research, I would use Python to grab some online data as supplement.

1. data cleaning

clean the data, including filtering (remove all but the data of interest), ordering the data…

1. data visualization stage1

after cleaning the data, we will use some visualization methods to take a look at the properties of our data. It is necessary to put the data into some structure, which help us understand them. And also we can get some inspirations for next from this!

1. Modeling and evaluation

Apply some methods or models from statistics or data mining to discern patterns, place the data in mathematical context.

1. conclusion (data visualization stage2)

after modeling work, we could have some visual representations, and I will focus on how the make it clearer and more visually engaging.

1. web-development*(optional)*

to make our conclusion more attractive, I am considering use some web skills to improve our work.