

STAT231 Group C Revised Project Proposal

Kenny Chen, Jessica Yu

October 18, 2019

date?

Contents

Group Letter: C	1
Group Members:	1
Title:	1
Purpose:	1
Data:	1
Variables:	1
End Product:	1

Group Letter: C

Group Members:

Kenny Chen and Jessica Yu

Title:

Examining healthcode violation trends and similarities between NYC restaurants

Purpose:

We want to use spatial data to map restaurants in NYC to examine if certain types of healthcode violations occur in a specific area/district more frequently than others. We want to combine that information with sentiment analysis of Yelp reviews to also show "similar" restaurants.

Data:

We will be using the dataset provided by Yelp that contains information about restaurants in NYC along with reviews as well as NYC Health Department data on healthcode violations.

Variables:

We will be looking at longitude, latitude, postal code, type of cuisine, price range, number and type of healthcode violations, and conducting sentiment analysis on Yelp reviews.

End Product:

Ultimately

this is a good start but I need more detail.
can you please submit a revised proposal
by the end of the day on 11/12 with
sample data for 2 restaurants from both
datasets?

how many?

