IST 664 final project proposal_group3

1. **Group member's names and majors:**

   Mengdie Zhuang, Information management

   Chaoying Lyu, Applied data science

   Ruiwei Zhang, Applied data science

   Zeyi Luo, Business analytics

   Renjie Zhu, Applied data science

2. **Description about our project data set**

   Our group wants to use the public Boston Airbnb data set to do some analysis.

   The following Airbnb activity is included in this Boston dataset:

- Listings, including full descriptions and average review score
- Reviews, including unique id for each reviewer and detailed comments
- Calendar, including listing id and the price and availability for that day

3. **Major NLP techniques:**

   We searched relevant research resources and found that most researchers used sentiment analysis techniques to analyze the reviews. They did sentiment analysis on positive and negative reviews and caught the valuable information from them.

4. **The proposed task your group plan to complete:**

   Here is the business questions for our airbnb review analysis :

- How to improve the service quality based on our analysis?
- What are the major factors that affect customers' satisfaction?

   Firstly, just like most researchers did, we would use sentiment analysis as our main NLP technology.

Simply abstracting the positive and negative words using sentiment analysis is not doing a good job in the resources we found, so we would try to do some different things such as preprocessing the data or from different aspects to see if that would get better results.

Besides, we would also build an analysis model to predict the rating score based on customers' reviews to find something interesting relevant to our business questions since the rating score would reflect the customer satisfaction to some extent.

Then we combine the results of those two technologies to gain the final business insights.

5. A DRAFT week by week schedule of tasks for your group and who will complete them

Time	Task	Task distribution
Week8	Meet with group members, search information on the internet, decide what project topic to choose.	All together
Week9	Data cleaning and preprocessing	Ruiwei, Mengdie Zhuang
Week10	Positive and negative comment analysis	Ruiwei, Chaoying Lyu
Week11	Model building, model tuning and model evaluating	Renjie Zhu, Zeyi Luo
Week12	Compare the results and output our final report based on our weekly recording	All together