Final Project Proposal – IST 652

**ANALYZING AIRBNB USER DATA**

DATA SET USED

The Kaggle challenge provides a list of users along with their demographics, web session records, and some summary statistics.

Data Source: <https://www.kaggle.com/c/airbnb-recruiting-new-user-bookings>

TOPIC OF INVESTIGATION

**Where do new users book their first destination?**

This topic of investigation will be used to address the following:

1. **Exploratory Analysis**

An analysis of the types of users and their preferred destinations will give us insight into who AirBnb’s users are and their tendencies.

1. **Predictive Modeling**

The question will help us find which countries new travelers or users of Airbnb will first book. Can we predict which country they will choose just by a few initial pieces of information a user provides?

1. **Text Mining / Sentiment analysis / Natural Language Processing**

I will attempt to pull comments from AirBnb destination countries to see if the sentiment toward the country matches the ratings that users provide for their initial stay in their first AirBnb.

PROJECT SCOPE

This project will include charts and maps to show where users are staying. If successful with the webscraping, I believe 80 mb of data will be used.