

Airbnb San Francisco Reviews NLP Proposal

By Asteway Kebede

- What is the question behind your analysis or model and what practical impact will your work have? How will that client benefits from exploring this question or building this model/system?

For this project, customer reviews posted on Airbnb in San Francisco will be analyzed using Natural Language Processing methods such as sentiment analysis and topic modeling. The goal is to help hosts better understand and learn their customers' needs to better serve their customers within their city.

- What dataset(s) do you plan to use, and how will you obtain the data?

Downloading Airbnb customer reviews from [Insideairbnb.com](#) 280K reviews - [San Francisco Airbnb Reviews](#)

- What is an individual sample/unit of analysis in this project? In other words, what does one row or observation of the data represent?

listing_id	id	date	reviewer_id	reviewer_name	comments	year
53258020	501567230864641900	2021-11-22	104345330	Paige	We had a lovely stay at Craigs place. The room...	2021

- What characteristics/features do you expect to work with? In other words, what are your columns of interest?

'comments' column will be used as the text reviews

- How do you intend to meet the tools requirement of the project?

Python for EDA, Modeling and Visualization as well as NLP algorithms

- Are you planning in advance to need or use additional tools beyond those required?

With any new skills learned from this class

- What would a [minimum viable product \(MVP\)](#) look like for this project?

The MVP will contain results from the initial sentiment analysis and topic modeling.