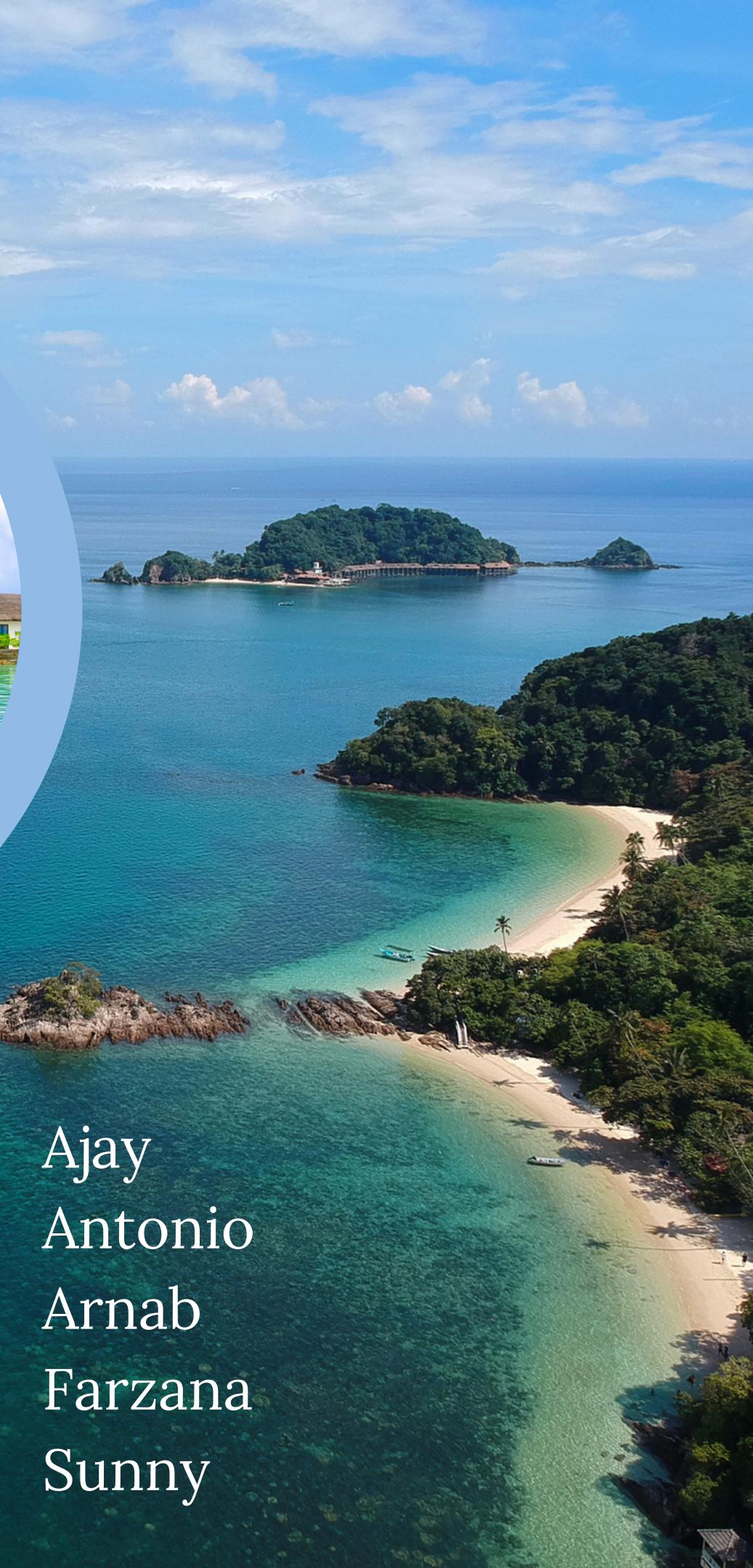


INNSIGHT

TRAVEL ASSITANT



GROUP 3
OSU AI-Bootcamp



Ajay
Antonio
Arnab
Farzana
Sunny

- 
- INTRODUCTION
 - DATA SOURCING/PROCESSING
 - FUNCTIONALITY
 - FEATURES/CAPABILITIES
 - TECHNOLOGIES
 - CHALLENGES
 - FUTURE INTERGRATION
 - USER EXP + DEMO

INTRODUCTION

The goal of our project is to create a virtual assistant to create a complete travel plan.

With limited time as first phase, we have implemented an app to assist with hotel recommendations based on reviews and ratings and we call our first phase “Inn”Sight(○○).

- Search hotels based on natural language input.
- Recommend hotels based on rating.
- Analyze reviews and provide recommendations based on score.
- Suggest hotels based on amenities and key words found in reviews.
- Top attractions to visit in a city.

DATA SOURCES

Airbnb Rating Dataset (Kaggle)

Containing over 200k samples globally

Datafinity Hotel Reviews

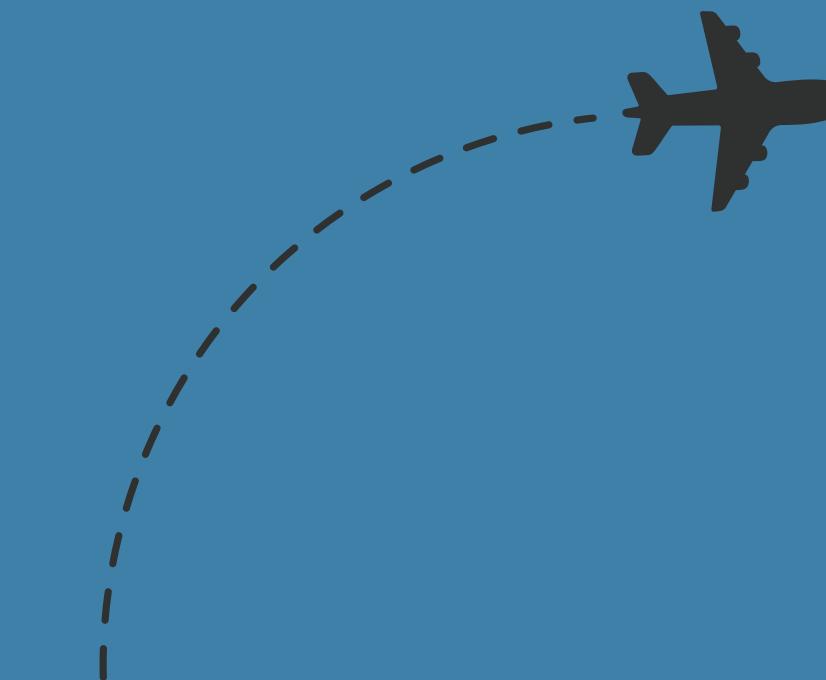
Containing over 20K Hotel Reviews

Google Places (API)

Get realtime reviews for hotels.

Open Trip Map (API)

Get top attractions by location.



DATA PROCESSING

- Hotel review data had 27 features but we focused on rating and review text to analyze and use in our models.
- All ids, date values, categories was removed along with all null values.

	<u>id</u>	<u>dateAdded</u>	<u>dateUpdated</u>	<u>address</u>	<u>categories</u>	<u>primaryCategories</u>	<u>city</u>	<u>country</u>
0	AWE2FvX5RxPSIh2RscTK	2018-01-18T18:43:12Z	2019-05-20T23:55:47Z	5620 Calle Real	Hotels,Hotels and motels,Hotel and motel mgmt....	Accommodation & Food Services	Goleta	United States
1	AVwcj_OhkufWRAb5wi9T	2016-11-06T20:21:05Z	2019-05-20T23:31:56Z	5th And San Carlos PO Box 3574	Hotels,Lodging,Hotel	Accommodation & Food Services	Carmel by the Sea	United States
2	AVwcj_OhkufWRAb5wi9T	2016-11-06T20:21:05Z	2019-05-20T23:31:56Z	5th And San Carlos PO Box 3574	Hotels,Lodging,Hotel	Accommodation & Food Services	Carmel by the Sea	United States

DATA PROCESSING

	dateAdded	address	city	country	latitude	longitude	name	postalCode	province	reviews.rating	reviews.text
0	2018-01-18	5620 Calle Real	Goleta	US	34.44178	-119.81979	Best Western Plus South Coast Inn	93117	CA	3.0	This hotel was nice and quiet. Did not know, t...
1	2016-11-06	5th And San Carlos PO Box 3574	Carmel by the Sea	US	36.55722	-121.92194	Best Western Carmel's Town House Lodge	93921	CA	4.0	We stayed in the king suite with the separatio...

- Featured engineering by adding columns to with derived values from review text.

tokenized_reviews_text
[this, hotel, was, nice, and, quiet, did, not, ...]
[we, stayed, in, the, king, suite, with, the, ...]

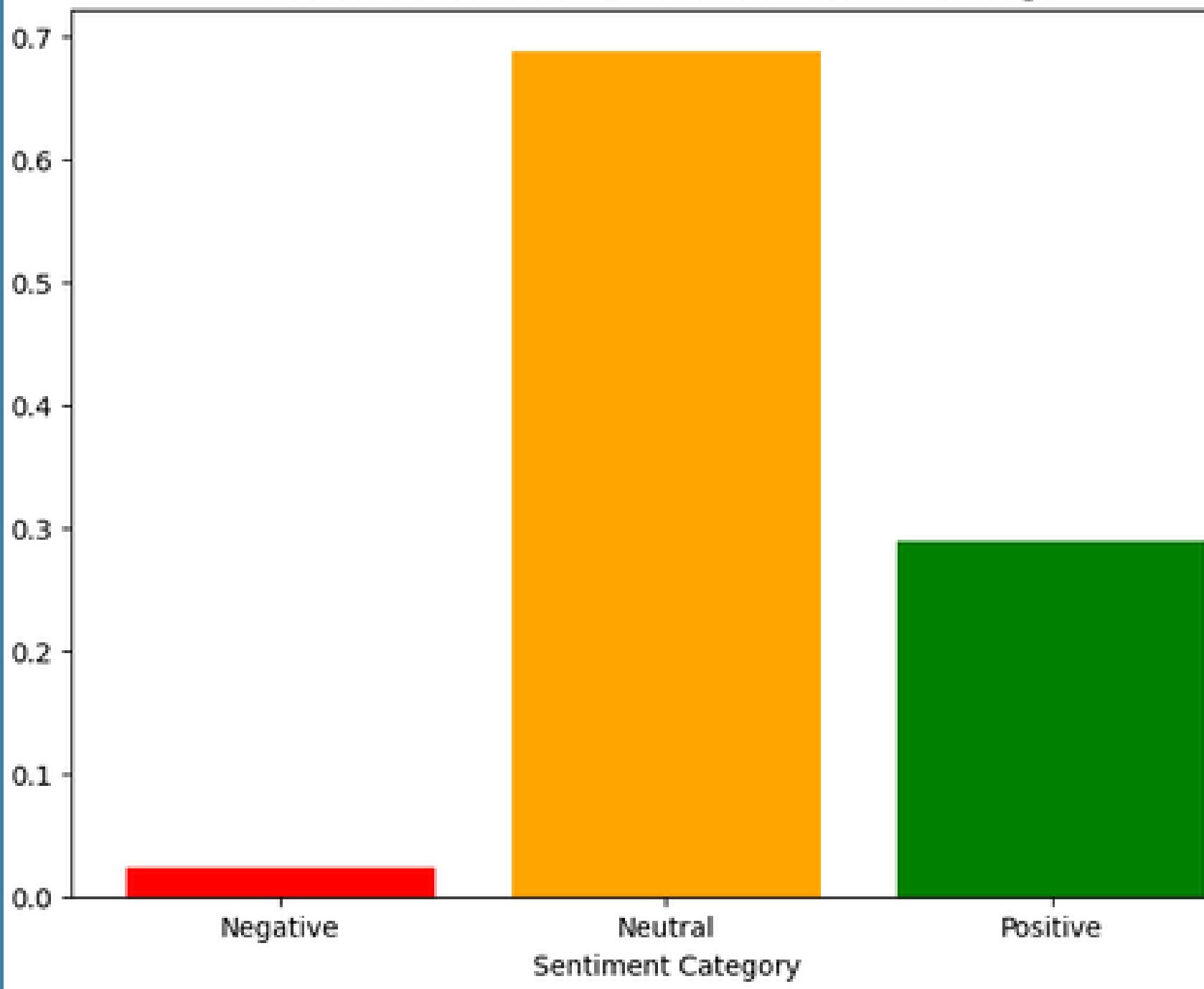
DATA ANALYSIS

- Extract top hotels by city, state using the ratings.
- The review text was used to analyze sentiment and provide a score and suggest based on scoring.
- Identify key words from review text and extract amenities.

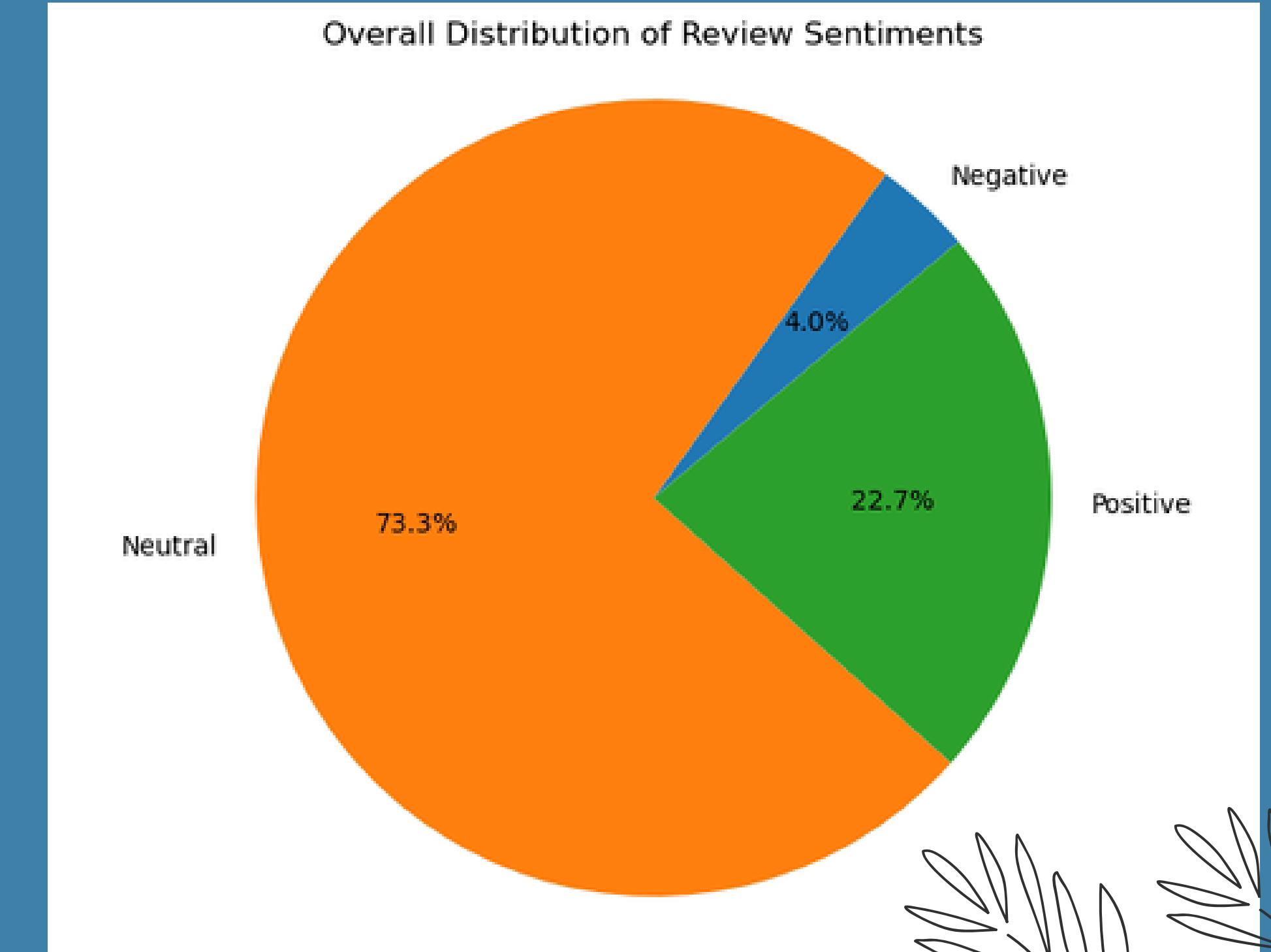
Modules used: **NLTK, SKLearn, Numpy, Textblob**

SENTIMENT ANALYSIS

Review Sentiments - Downtown Hotels in Chicago



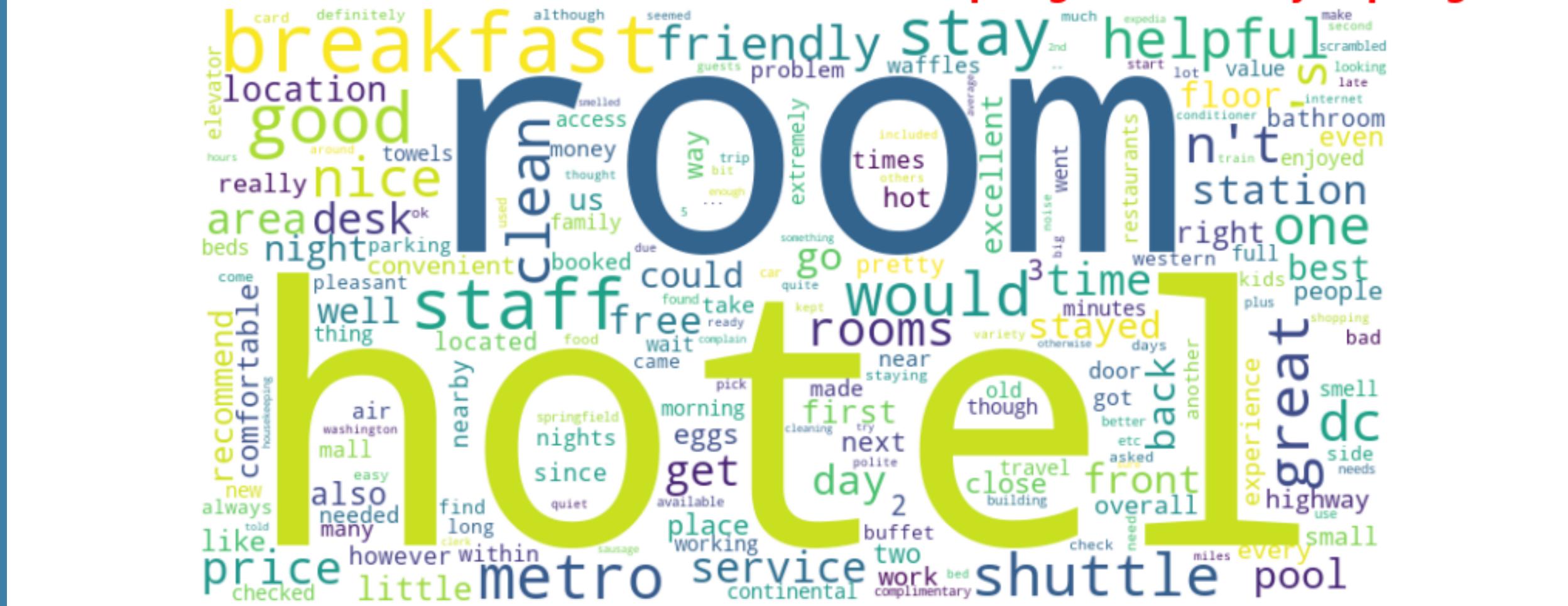
Overall Distribution of Review Sentiments



Modules used: Matplotlib

WORDS ANALYSIS

Word Cloud for Reviews of Hotel "Best Western Springfield" in City "Springfield"



Modules used: Matplotlib, WordCloud

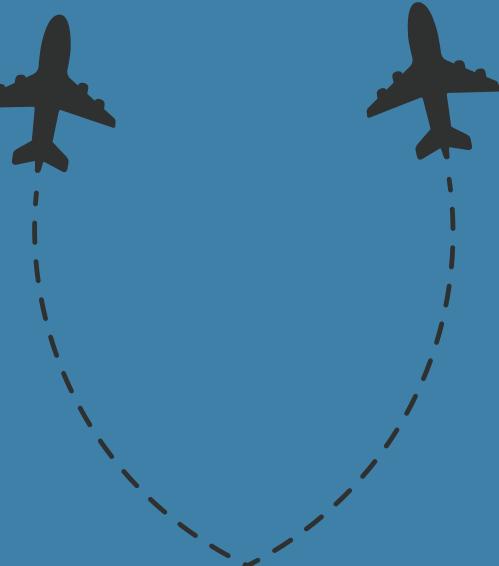


LOCATION



HOTEL
RECOMMENDATIONS

FEATURES

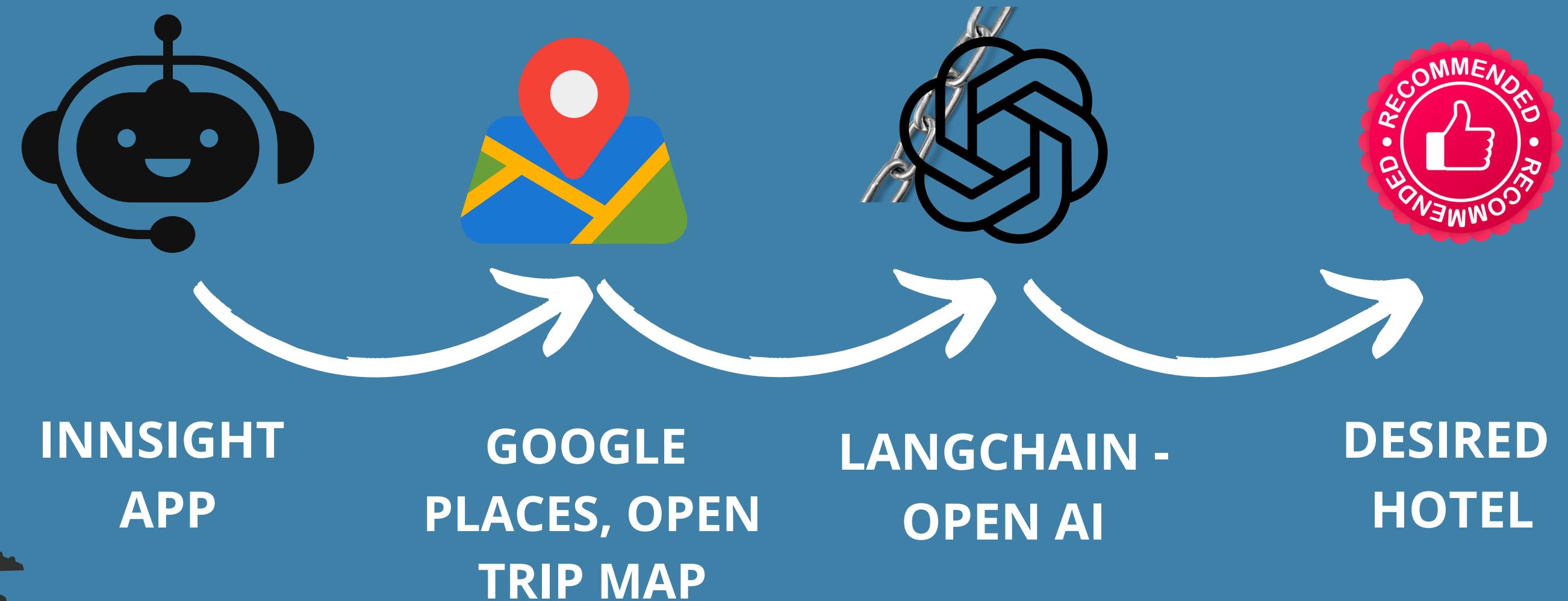


TOP ATTRACTIONS



AMENITIES

WEB APP



Modules used: Flask, Langchain, OpenAI

CHALLENGES

- Data collection - Tokenization limits when trying to convert travel pdfs, consume large static data, paid apis....
- Conversation Chat Interface - Amazon Lex UI & Set-up
- Complete travel assistant involves many categories, so had to breakdown on feature to start with.
- Scheduling that works with all team members schedule

FUTURE ENHANCEMENTS

- Continue with web interface to provide more content images, live weather, travel restrictions and so on.
- Update our LLM to make it more conversational inclusive and finally join all the pieces to provide a final full travel itenary.
- New Features - more details on top attractions, events, visual map itinerary, planning calendar.....

DEMO



QUESTIONS/COMMENTS

