**Sentiment Analysis and topic extraction Research.**

The process of determining the emotion or attitude expressed in a piece of text, particularly a review, is known as sentiment analysis. It can determine whether the text is positive, negative, or neutral. This is useful for hotel reviews because it helps to determine how guests feel about the hotel overall.

The process of automatically identifying the main topics or themes present in a piece of text, is known as topic extraction. Topic extraction in the context of hotel reviews can assist in identifying common issues or areas of interest mentioned by guests. The hotel can then use this information to improve its services or facilities and customers can use the dashboard to see what hotel to choose according to their preferences and priorities.

**Data Extraction—**

There are some APIs for hotel reviews:

The Yelp API allows us to access over 50 million reviews, photos, and business information. We can use the API to get information about hotels and their reviews.

The TripAdvisor API allows us to access traveller reviews and ratings, as well as photos and other information about hotels and other travel-related businesses.

Booking.com API: The Booking.com API allows us to access information about hotels and other lodging options, such as guest reviews and ratings.

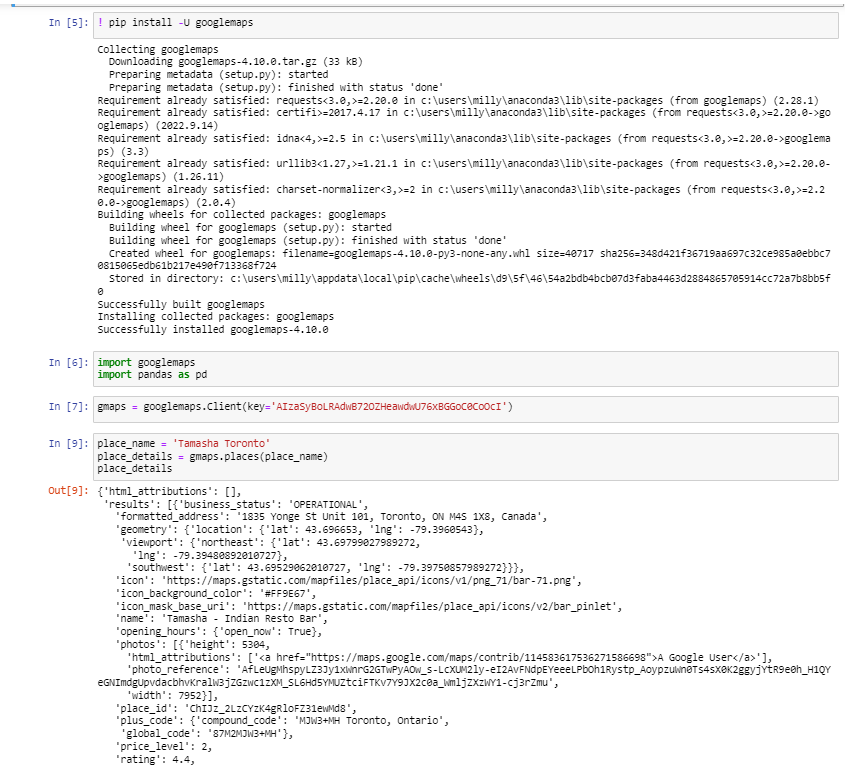
Expedia API: The Expedia API allows us to access a large database of travel information, such as hotel reviews, prices, and availability.

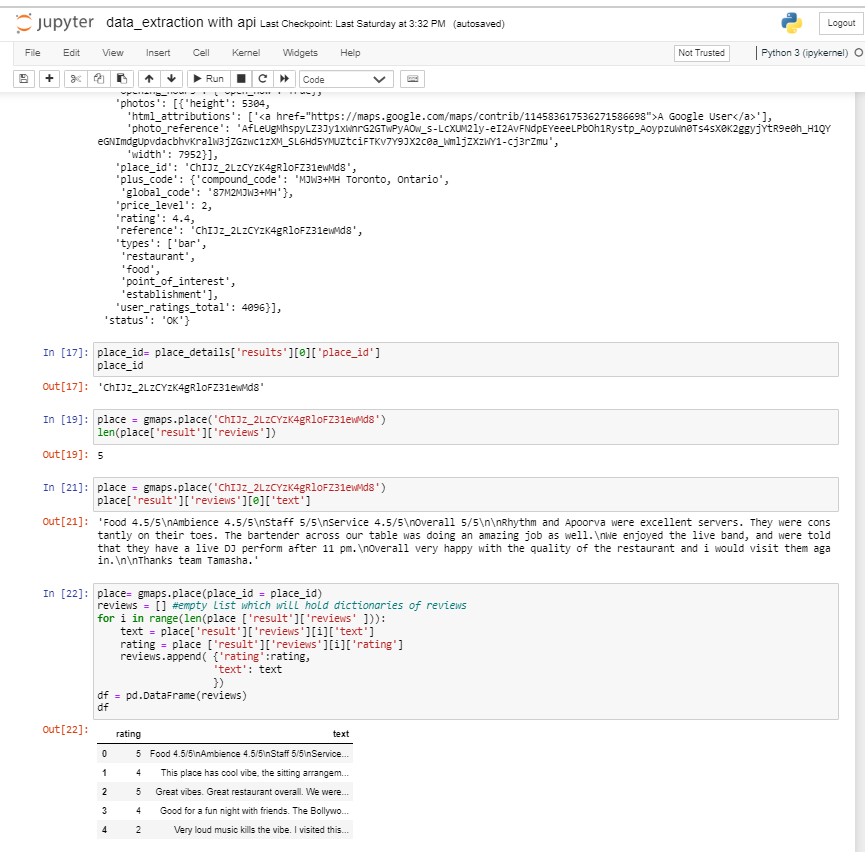
Hotels.com API: The Hotels.com API provides access to hotel and other lodging information, including guest reviews and ratings.

Google Maps API: The Google Maps API provides information about local businesses, including hotels. The API provides access to business information as well as user reviews, which can be used to retrieve hotel reviews.

I have worked on getting data from Google Places Api, and followed the following steps :--

* To get hotel reviews data from Google Places API, sent an API request with the hotel's Place ID. The Place ID is a unique identifier for a location in the Google Places database that can be obtained by performing a Place Search request.
* Obtained a Google API key: To make requests to the Google Places API, first obtained a Google API key. Got a key by going to the Google Cloud Console.
* Made a Place Search request.
* Requested Place Details: After having the Place ID, requested Place Details to get more information about the location, including reviews. In the Place Details request, included the Place ID as well as API key.
* Parsed the response: The API response was in JSON format, containing information about the location, including reviews. parsed the JSON data and extract the necessary review information.





Was only getting 5 reviews at a time and couldn’t pull the other reviews of the place with free API key, so had to drop google places api.

Decided on the whole machine learning process for the Sentiment Analyser :-

* Data Preparation: Will Gather and clean the reviews that will be used to train the sentiment analyzer.
* Text Pre-processing: Pre-process the text data by removing stop words and punctuation and converting it to numerical representations. USING( from nltk.corpus import stopwords, from nltk.stem import WordNetLemmatizer) ( from nltk.corpus import stopwords, from nltk.stem import WordNetLemmatizer)
* Feature Engineering: From the pre-processed data, Will create features to be used in training the sentiment analyzer.
* Modeling and Training: Using the pre-processed data and extracted topics, Will choose an appropriate ML algorithm (e.g., Naive Bayes, SVM, etc.) and train the sentiment analyzer.
* Model Evaluation: Test the sentiment analyzer on a separate validation set to assess its performance.
* Topic Extraction: Will be Extracting topics from reviews using spaCy's Classy- classification model.

Deployment: Finally, will be putting the trained sentiment analyzer into action.

References : - <https://developers.google.com/maps/documentation/places/web-service/overview>

<https://rapidapi.com/collection/hotels-apis>

<https://www.dataquest.io/blog/tutorial-text-classification-in-python-using-spacy/>

<https://www.analyticsvidhya.com/blog/2021/06/nlp-sentiment-analysis/>