**Yelp API and Google Places API Restaurant Reviews**

Prepared by Dinnara Hitt and Alberto Molina

**Introduction and Summary:**

An Orange county client is interested in opening a new restaurant and hired our team to provide customer review data for market research.

We’ve identified our datasets from Yelp and Google APIs as our data sources for this project. The goal is to perform an accurate and efficient Extract, Transfer, and Load (ETL) process before moving to a production database. The ETL

**ETL Process:**

**Tools**

* Yelp and Google API
* Jupyter Notebook
* MongoDb

**Extract**

* Retrieved the data sets from Yelp Fusion and Google API websites and submitted multiple API queries. The scope of the data extraction was based on the inputted parameters in the API calls.

Text

Description automatically generated

Figure 1

**Transform**

* The transformation process started into the staging area for data cleansing and analysis.
* The first step is to convert the data into JSON formats.
* The Data columns are dropped and renamed to fit the tables created in the database.
* The Data is filtered, checked, and dropped duplicates, and copied the columns into a new DataFrame.
* The Data is QC and tested numerous times.

Graphical user interface, text, application

Description automatically generated

**Load**

* Created a connection to the database using MongoDB and confirmed the successful connections.
* Migrated Yelp and Google reviews and transport the data into Mongo DB . **Graphical user interface, text, application, email

  Description automatically generated**
* Confirmed the tables had been created by performing numerous queries in the database.

**Text

Description automatically generated**

**Conclusion:**

The purpose of the project is to gather data for restaurant reviews using the ETL process. We successfully applied the process that involves extracting, transforming, and loading the data. It involves other components, which includes authentication, data handling, and exception. The final output is to load data into the database.

We have though some limitations in combining Yelp and Google APIs due to their varying parameters. Also, by default Google API could only obtain 60 results per query to avoid data scraping.

**Reference:**

**<**[**https://www.yelp.com/developers**](https://www.yelp.com/developers)**>**

**<**[**https://maps.googleapis.com/maps/api/place/nearbysearch**](https://maps.googleapis.com/maps/api/place/nearbysearch)**>**