

## Data Analysis Project Proposal

- **Project Title:**
  - Exploring the Relationship between Restaurant Diversity and City Demographics
- **Goal:**
  - By leveraging the Yelp Business Search API for restaurant data and public demographic data from government sources, we will analyze the distribution of different types of restaurants and explore how it relates to the racial composition of the city's population.
- **Data Source:**
  - **Yelp API:** We will use the Yelp API to collect data on various restaurants, including their cuisine types, ratings, and locations.
  - **City Demographics Data:** Publicly available demographic data for the selected city, including race and ethnicity distribution, will be obtained from reliable sources such as government databases or official city websites.
- **Data Analysis Tools:**
  - The analysis will be conducted using libraries including Pandas, NumPy, and Matplotlib.
- **Analysis approach**
  - Data Collection:
    - Retrieve restaurant data from the Yelp API, focusing on different cuisine types, ratings, and locations. Simultaneously, gather demographic data for the selected city, specifically focusing on race and ethnicity distribution.
  - Data Cleaning:
    - Clean and preprocess the collected data to ensure consistency and accuracy, handling missing values and outliers as needed.
  - Data Integration:
    - Merge the restaurant data from Yelp API with the demographic data based on the city location, creating a unified dataset for analysis.
  - Data Visualization:
    - Heat map/scatter plot that visualize the correlation matrix between different cuisine types and racial demographics.
  - Correlation Analysis:
    - Utilize appropriate statistical methods to determine the correlation between the types of cuisine offered by restaurants and the racial composition of the city's population.