# Project Proposal: Study Yelp

An application to help students find the best study spaces in their surrounding area based on their preferences.

Meerza Ahmed, mhahmed@usc.edu

Melissa Perkins, perkinsm@usc.edu

Alan Trinh, alantrin@usc.edu

DSCI 551 - Foundations of Data Management
Prof. Wensheng Wu
02/02/24

# Team Member Background & Skills

**Meerza Ahmed** - B.S. in Computer Programming & Information Systems from SUNY Farmingdale. I mainly used C++, and Java during my bachelor's, and got to work extensively with database systems. I have previously used SQL, firebase, HTML/CSS, JavaScript, and data modeling to build a web application. I have also worked with PostGIS, to map spatial data.

**Melissa Perkins** - Undergraduate degree in Cognitive Science from UC Berkeley with some Python, R and SQL experience during that time. I started the MS program in Applied Data Science in 2021 as a part-time student while working full-time on the HSC campus. I have since gained more experience with Python and various libraries, I was also introduced to Java and CSS. As a project for DSCI 554, I created a dashboard using a Javascript framework- Vue.

**Alan Trinh** - Minored in Data Science from UC Berkeley with experience in R, SAS, SQL, and Python. I started the MS program in Applied Data Science last fall and have taken DSCI 550. Extensive knowledge in building predictive models, classification models, data analysis, and visualization. Have some experience with Python programming and am looking forward to learning more about full-stack development. Currently working in product development and have lots of experience working with software engineers, data scientists, and UI/UX designers.

# Project requirements

We believe the project requirements are to build a web application that users can visit to interact with a data source. The web app should ideally have a distributed database, with a minimum of two tables, and two UI interfaces. The first one should be front-end heavy and be designed to display our dataset in a meaningful way so that users can interact with it, the other UI should be reserved for data manipulation by a data manager; this one can be a CLI. It is also important to note that the data manager needs to be able to add, delete, and modify any specific tuple in any of our database tables. Our implementation needs to include a standout feature or element, to gain full credit on the project.

# Planned implementation

1. **Dataset & Data Cleaning**
   - Download the Yelp dataset and perform data cleaning, handling null values, features selection, etc
2. **Database Setup**
   - Setup Firebase databases (2) will separate the data by Eastern and Western states.
3. **Backend Development**
   - Create a hash function to split the data by East and West
   - Write Python functions for data management (add, delete, modify) study spots
   - Express.js framework
4. **Develop CLI interface**
   - Write Python functions to develop a CLI menu for data managers for add, delete, and modifying
5. **Frontend Development**
   - Deciding between Javascript or Python framework
6. **Testing & Debugging**
   - Unit tests
   - Documentation
   - User guide and code documentation

# Team Responsibilities

| | |
|---|---|
| Meerza | Data collection, Data cleaning, Management, Data Modeling (ER / EER), CLI Implementation |
| Melissa | Data collection, Data Analysis, Data Visualization, UI Design, Website Component Implementation, CLI implementation |
| Alan | Data collection, Data Analysis, UI Design, Data Prediction Modeling, Python queries, CLI Implementation |

# Timeline

**Weekly Meetings: Mon & Thur (Started on 1/25)**

| February - 02 | Submit project proposal |
|---|---|
| February - 09 | Finish Firebase skeleton and upload Yelp dataset into dedicated tables |
| February - 23 | Build a working front-end skeleton |
| March - 01 | Midterm Report |
| March - 08 | Program and test data queries; data to be displayed on the frontend |
| March - 15 | Build the function calls for the CLI part of the project |
| April - 06 | Finish front-end display working data, and test user interactions |
| April - 12 | Testing / Debugging |
| April - 17 | Implementation & Demo |
| May - 03 | Final Report |