- Please list out changes in the directions of your project if the final project is different from your original proposal (based on your stage 1 proposal submission).

In general, we change our ER diagram. In the sage 1 proposal, inside the Nutrition table, the primary key is ingredientID, but later we changed it into recipeID.
In the user table, we previously only had userID(as the primary key), userName, and userEmail. Later, we added a password attribute.
In our stage 1 proposal submission, we have the FavoriteRecipe table, with userID and favoriteRecipeID as the primary keys, and recipeID and dataAdded as other keys. However, we remove the FavoriteRecipe table and change it to the Markets table. The Markets table has marketID, marketName, and marketLocation, with marketID as the primary key.

- Discuss what you think your application achieved or failed to achieve regarding its usefulness.

Our application successfully fulfilled the core functionality, which is recipe search. Right now, our application supports two different kinds of recipe search (search by name and search by ingredient). However, there are still some functions that are not fulfilled like filter search results.

- Discuss if you change the schema or source of the data for your application

The main source of the data remained unchanged, which is still https://www.kaggle.com/datasets/shuyangli94/food-com-recipes-and-user-interactions/data?select=PP_recipes.csv
This dataset contains our recipe data, ingredient data, and nutrition data. Note that due to the huge original size of the recipe data (about 250k recipes) that causes disastrous upload time to the GCP, we have reduced the recipe size to about 1500.

- Discuss what you change to your ER diagram and/or your table implementations. What are some differences between the original design and the final design? Why? What do you think is a more suitable design?

First, We have added the Market Table, RecipeLikeLog, and RecipeRateLog in our databases. Now the user can check out the locations to buy the ingredients. The Log table will give us insight of when and what user like/rate . Next, we have deleted the table FavortiteRecipe, instead, we made it a relation between the recipes and the user table. I think our current design is suitable.

- Discuss what functionalities you added or removed. Why?

Until now, our application supports the following functionalities:

- Search recipe by recipe name
- Search recipe by recipe ingredients
- User Registration and Login
- User setting changes (change username, change password)
- Like, remove like, and show like for logged-in users
- Rate recipes for logged-in users

- Explain how you think your advanced database programs complement your application.

- Each team member should describe one technical challenge that the team encountered. This should be sufficiently detailed such that another future team could use this as helpful advice if they were to start a similar project or where to maintain your project.

Jinheng Li: GCP database backup will drop all databases, so disable it when using the feature. Also, I am having trouble host the application on GCP, due to their confusing document. If you are using expressJS, make sure you include an app.yaml file in the project directory that specifies the runtime env you want. After that use “gcloud app deploy” to deploy your app on the google app engine.

Siran Xianyu: When I was doing data cleaning, it was important to be clear how to extract relevant information from a complex original dataset. I first extracted the dataset from main tables and then populated other information. In connection to databases from IDE, I encountered the problem of invalid IP addresses. Additionally, I figured out how to define a transaction in javascript to provide a valid transaction to ensure my queries are processed correctly and use an appropriate isolation level.

Yixuan Li: After importing SQL tables to GCP, we encountered a problem with inserting data via CSV file. Finally, I found out that it is the column sequence does not match the definition in the SQL table (So the table recognized user email as username, and username as user email). Also, it seems like when GCP reads CSV files, it does not require a title row, it will parse the columns and start inserting right from the first row.
Also, there is a challenge that I encountered, which is that when users are signed in, any windows popped cannot be closed properly, which I am still not able to fix.

Qiyang Liu: I encountered the challenge when I was doing indexing in stage 3 part 2 of the project. As a result, one of the requirements for this part is that we cannot index primary keys. At the same time, we have to have at least 3 indexing designs. And most keys in our written advanced queries are primary keys. In this case, I need to go back and make my advanced query much more complicated so that I can have at least 3 non-primary keys to do indexing. At the same time, when I first do the indexing, the time after indexing does not change. Later, I noticed that after running the data each time, I have to turn it off, restart, and then re-run so that it does not perform the same indexing.

- Are there other things that changed comparing the final application with the original proposal?

We have nothing changed in our final application compare to the original proposal.

- Describe future work that you think, other than the interface, that the application can improve on

We have developed user logging in, reading recipes based on keywords of ingredient name and recipe name, rating, and liking recipes with a valid user account.

Firstly, As we only used 1500 recipes in our current development, the types of recipes are significantly reduced, so we will upload new data into our database. Secondly, we will develop

the functionality of getting nutrition information from each recipe. Thirdly, we will provide another search interface that tells users where a specific ingredient can be bought. Additionally, we will display the average rating and comments of a recipe when the user searches for the recipe.

We would also need to implement a user upload recipe function. Also, currently, though we have implemented a rating function, users cannot check the rating of the recipes yet. In the future, we should allow users to check the comments on the recipes.

- Describe the final division of labor and how well you managed teamwork.

Siran Xianyu: Data cleaning, backend development
Yixuan Li: GCP SQL setup, data import, frontend development
Jinheng Li: Host website on Google App Engine, Plan B application development
Qiyang Liu: Data cleaning, query optimization