# CA675: Cloud Technologies
# Assignment 2 – Group 16

**Introduction:**

This project is a web application which allows users to browse and easily access cooking recipes. This application allows a user to search by recipe name, search by ingredient also displaying an interactive dashboard of charts that give insights into the data.

**Data -**

The data we have selected for this project is the "Food.com Recipes and Interactions" available on Kaggle. This dataset consists of 200K+ recipes and 900K+ recipe reviews covering 18 years of user interactions and uploads on Food.com. More on this dataset is available at -https://www.kaggle.com/shuyangli94/food-com-recipes-and-user-interactions

**Process:**

A cluster in Google cloud platform was used to build this project and spark environment was set up to do data pre-processing. Data cleaning was done in pyspark to remove rows in which any cell in the dataset is NA. The application was built using two datasets. The first one has information on recipes, i.e. its ingredients, steps to prepare and its attributes. The second one contains information on recipe reviews obtained from various users. We have stored the data in google cloud storage and loaded it in BigQuery.

Queries were written on big query-based SQL to provide data for creating visualizations on Data studio. These queries provide data for various trends that can be drawn from the dataset. These trends incude titles like, Recipes by Time Required to Cook,Recipe contributors,Ingredient Popularity(Top 50),Popular Recipes, Instant Recipes, Trending Tags,Monthly Healthy Posted Recipes Trend.

Application backend is written in Java using spring boot framework. It connects with BigQuery using a service account authentication key and subsequent JDBC calls. It dynamically generates queries for general recipes search, finding recipes by id, find interactions for recipes, search interactions by keywords etc. To speed up the search and minimize cost, frequently searched data, recipes and interactions are cached using java caching framework in memory. Backend REST API is documented using industry-standard Swagger v2 API documentation and is of Level 3 maturity with hypermedia controls.

A user can search the application based on ingredients or recipe or and get matching recipes. The user can then select a recipe and view the steps to prepare the recipe, time required, ingredients and any comments on the recipe. The frontend was developed using HTML, CSS and JavaScript. Data visualization charts and diagrams like the top healthy recipes, the tags of recipes that were trending during a given period.

**Code and scrits repo :**
https://gitlab.computing.dcu.ie/khandhh2/recipe_webapp

**Live Links:**

1. UI : https://newagent-11049.appspot.com/
2. Backend API Swagger UI: https://bigqueryproject-259721.appspot.com/swagger-ui.html

**Demo Link:**
https://youtu.be/VbiUCx9fQGU

**Related Work -**

Super cook and recipe land are similar applications to the project. In addition to it, this application also demonstrates a variety of graphical representations providing an insight into the data we use.

**Challenges and lessons learned –**

- Spark and Hadoop are jobs are difficult to integrate into online components like a REST API back end.
- It's cheaper and timesaving to use IaaS native solutions like BigQuery in GCP for the critical task as they are much reliable and performance-critical when compared to self-managed solutions when time and finance constrained.

**Individual responsibility and contributions-**

1. **Akhil:**
    a. **Role:** Lead Developer
    b. **Tasks assigned:**
        i. Develop backend application
        ii. Deploy and manage applications in GCP
        iii. Coordinate work efforts
    c. **Individual contributions:**
        i. Developed Java-based microservice application with caching of frequently used data to minimize big query usage.
        ii. Implemented industry standard swagger UI for RESt API documentation.
        iii. Deployed backend and frontend code in GCP
        iv. Coordinated work efforts

2. **Kirthy:**
    a. **Role:** UI Design and Implementation
    b. **Tasks assigned:**
        i. Design and create UI.
        ii. Connect the backend REST API points to display data in frontend.
        iii. Integrate visualizations in user interface.
    c. **Individual contributions:**
        i. Designed and created the UI using HTML, CSS, and JavaScript.
        ii. Connected backend to frontend to display data in the UI.
        iii. Integrated and displayed the visualizations in UI.

3. **Tejal:**
    a. **Role:** Backend Developer
    b. **Tasks assigned:**
        i. Develop backend scripts
        ii. Create visualizations of the data
        iii. Display them at frontend
    c. **Individual contributions:**
        i. Developing the SQL queries for visualization graphs in BigQuery
        ii. Creating the visualization in Data studio
        iii. Creating informative and descriptive visuals
        iv. Embedding the report in the frontend code to make it visible on UI

4. **Hithesh:**
    a. **Role : Developer**
    **b. Tasks assigned:**
        i. Set up an environment to run the application.
        ii. Data cleaning and loading.
    **c. Individual contributions:**
        i. The cluster was set up on GCP with required capabilities.
        ii. Deployed a pyspark script to clean the data and load to BigQuery for further processing.

5. **Barath:**
    a. **Role: UI Developer**
    b. **Tasks assigned:**
        i. Front end development.
    c. **Individual contributions:**
        i. Assisted with the development of User Interface development.

## Responsibility statement

| **Student Id** | **Name** | **Roles** | **Member Contribution** |
|---|---|---|---|
| **19210588** | Kirthy Francis | Frontend Development | Satisfactory |
| **19211063** | Barath Raja | Frontend Development | Satisfactory |
| **19210412** | Tejal Nijai | Backend scripts | Satisfactory |
| **19210864** | Hithesh Sai K B | Data cleaning and loading | Satisfactory |
| **19210912** | Akhil Alfons Kodiyan | Components Integration & Backend Development | Satisfactory |