Food: We are What We Eat

Group 3: Meina Bian, Norma Espinosa, Dan Kananen, Henol Sososa, Angele Yazbec

1. Extract
   1. Each group member is going to obtain data from a variety of sources (e.g. Kaggle, Spoonacular API, USDA API).
   2. The data will contain the following information:
      1. Recipe Name “str”
      2. Ingredients in “str”
      3. Fat Grams “float”
      4. Carb Grams “Float”
      5. Protein Grams “Float”
      6. Total calories “float”
      7. Total servings “int”
      8. Calories per serving “float”
      9. Cuisine type “string”
      10. Recipe\_instructions “string” (url)
      11. Image\_url “string” (url)
      12. Food aversion
          1. Gluten Free (boolean)
          2. Vegan (boolean)
          3. Vegetarian (boolean)
          4. Ketogenic (boolean)
2. Transform
   1. Each group member will clean their acquired data set using Pandas.
   2. Part of this process will include standardizing the column names to ease the joining process later on in Postgres.
   3. Loading the cleaned dataframes into Postgres.
3. Load
   1. Create a relational database that combines these datasets.
   2. Outline the schema, including primary keys and foreign keys.
   3. Join tables, as they make sense to create larger tables.
   4. Create sample queries to respond to prompts such as:

“I’m on a diet that only allows x calories per meal and must have “a” grams of protein, “b” grams of fats, “c” grams of carbohydrates, and may or may not include ingredients such as “gluten”, “peanuts”, “green produce”. And I’m cooking for “y” amount of people.”