

1. Why was a python script used?

I chose to use a python script to generate thousands of recipes, ingredients, and user information from a public API available instead of manually writing out everything or making things up. The script also saved time instead of manually inserting the data, it generates a SQL file with all the insert statements.

The python script gave me control over:

- Primary key uniqueness
- Data volume requirements
- Randomization
- All constraints

2. How the script works?

The script pulls real data from OpenFoodFacts API and generates artificial fake users and recipes based off our schema.

Core features:

- Duplication filtering

The API contains foods that are the same just different accents on letters from different countries, so the script has a normalization function that removes accents and all other extra things that can make it seem like they're different words. This guaranteed no duplications.

- Joinability

I generated different "userId" values from (1-3000) then used those IDs when creating mealplans, to give each plan to a user.

3. Size Requirements

- Ingredients – 2000 rows
- Recipe – 3000 rows
- Site_user – 300 rows
- All others are in the range of 10-50 rows

4. Samples

- Ingredients

```
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Coca-Cola', 'Bevande e preparati per bevande', '100 grams', 'g', 42, 0, 11, 0);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Creamy', 'Produits laitiers', '100 grams', 'g', 109, 4, 17, 3);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Maruja', 'Snacks', '100 grams', 'g', 539, 7, 53, 33);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Cream cheese', 'Dairies', '100 grams', 'g', 257, 10, 3, 18);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Spreadable Slightly Salted', 'Dairies', '100 grams', 'g', 706, 0, 1, 78);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('Organic 70% Dark Chocolate Bar', 'Snacks', '100 grams', 'g', 580, 0, 36, 42);
INSERT INTO ingredients (ingredientName, category, measurementUnit, unitShorthand, calories, protein, carbs, fats) VALUES ('jibal', 'Chees', '100 grams', 'g', 138, 2, 6, 12);
```

- Recipe

```
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (2, 'perly', 'Prepared meal based on perly.', 49, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (3, 'Lait', 'Prepared meal based on Lait.', 25, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (4, 'Fromage blanc nature', 'Prepared meal based on Fromage blanc nature.', 21, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (5, 'uht jaouda 1L', 'Prepared meal based on uht jaouda 1L.', 12, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (6, 'Jben', 'Prepared meal based on Jben.', 7, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (7, 'Marmite Yeast Extract', 'Prepared meal based on Marmite Yeast Extract.', 55, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (8, 'Jaouda', 'Prepared meal based on Jaouda.', 30, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (9, 'Lait Frais', 'Prepared meal based on Lait Frais.', 20, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (10, 'lillia', 'Prepared meal based on lillia.', 21, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (11, 'Prince Goût Chocolat', 'Prepared meal based on Prince Goût Chocolat.', 46, 'api_meal');
INSERT INTO recipe (recipeId, name, recipecol, cookTimeMin, recipecol1) VALUES (12, 'Excellence Noir Prodigieux 90% Cacao', 'Prepared meal based on Excellence Noir Prodigieux 90% Cacao.', 53, 'api_meal');
```

- Users

```
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (1, 'Emma', 'Kim', 'user1@mealapp.com', '2025-04-25', NULL, 1811, 62, 369, 77);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (2, 'Aisha', 'Ali', 'user2@mealapp.com', '2025-04-27', NULL, 3480, 103, 262, 80);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (3, 'Noah', 'Wang', 'user3@mealapp.com', '2025-03-09', NULL, 3252, 108, 173, 95);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (4, 'Chen', 'Hammad', 'user4@mealapp.com', '2025-04-24', NULL, 2437, 140, 140, 117);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (5, 'Olivia', 'Garcia', 'user5@mealapp.com', '2025-04-24', NULL, 2942, 140, 289, 69);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (6, 'Ali', 'Patel', 'user6@mealapp.com', '2025-03-22', NULL, 1826, 124, 393, 93);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (7, 'James', 'Hammad', 'user7@mealapp.com', '2025-03-24', NULL, 2067, 104, 332, 52);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (8, 'Hassan', 'Lopez', 'user8@mealapp.com', '2025-03-05', NULL, 2016, 102, 366, 102);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (9, 'Sara', 'Santos', 'user9@mealapp.com', '2025-04-21', NULL, 3459, 112, 227, 42);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (10, 'Youssef', 'Chen', 'user10@mealapp.com', '2025-03-09', NULL, 1731, 169, 280, 115);
INSERT INTO site_user (userId, firstName, lastName, email, startDate, endDate, dailyCalories, dailyProtein, dailyCarbs, dailyFat) VALUES (11, 'James', 'Ivanov', 'user11@mealapp.com', '2025-03-09', NULL, 1731, 169, 280, 115);
```