

## XML vs. SQL Bake-off - November 7, 2011

Welcome to the XML vs. SQL bake-off, where you and your group will race other groups in answering a set of queries!

The data set is given to you in JSON, and you will have to choose to convert it to either XML or SQL for querying. (Hint: look at the queries before deciding whether to use XML or SQL!) We've provided starter code to get you most of the way to either creating an XML file or creating a CSV file to load into SQLite. You can run the starter code and queries on your own machine if you have python, plus SQLite (for SQL) or Saxon or Kernow (for XML). You may also choose to use the Corn cluster.

Data set: `/usr/class/cs145/bakeoff/bakeoff_cookbook.json`

Starter code: `/usr/class/cs145/bakeoff/bakeoff_starter.py`

### Queries:

1. Find the titles of all recipes in the breakfast food section.
2. Find the number of recipes with more than 5 ingredients.
3. Find the average number of ingredients needed for a recipe.
4. Find the number of different ingredients in this recipe book.
5. Find the recipe with the fewest number of ingredients.
6. Find the average number of ingredients needed for recipes requiring butter.
7. For recipes with more than 4 ingredients, find the 4th ingredient.
8. Find recipes requiring at least  $\frac{1}{2}$  tsp of two or more ingredients. (Don't worry about converting measurements, just look at ingredients measured in tsp.)

Once you're finished, bring your solutions up to the TAs to have them checked!