<https://www.kaggle.com/shuyangli94/food-com-recipes-and-user-interactions>

Data for recommender system:

* **ingredients**
* tags(cuisine category)
* allergy (For Sprint 3)
* ratings,
* nutrition(calories, sugar level, fat), ,
* time to prepare + number of ingredients + number of steps

Data to display to the users:

* Steps + all the data

Data from user:

ingredients

when ingredients expire: fresh, mature, ASAP

Get / Filter recipes based on current ingredients

Use other categories for prediction algorithm:

Naive Bayes, 95% confidence intervals based on current data,

Format:

Keys as pickle files

ingredients.pkl

cuisines.pkl

recipe\_id\_to\_column\_index.pkl

Ingredients:

Binary bag of words...?

**store indices only**

**list of unique ingredients**

Tags**:**

**Cuisine: Others, Mexican, American, Chinese, Italian, Thai, French, Japanese,**

**Korean, Indian, Arabic**

**Assign indexes to different cuisines and have mapping in memory**

**maybe just key word search**

***Tasks:***

process data: Christian Rodriguez

get rid of useless columns V

convert ingredients to more useful format V

convert tags to useful numbers V

generate pickles

ingredients.pkl V

cuisines.pkl V

recipe\_id\_to\_column\_index.pkl

Using ingredients to filter out elements：james, Dong

Threshold to like 70%

Use rest of data to do "cheap" prediction:

User input: Ingredients available; expiration status; personal preference

How to init user profiles (For Future)...TODO?