Table of Contents

Difference between cookies, pastries and pizzas

Source of data

Food Network (http://www.foodnetwork.com/), compiled by the user *everest4ever* on reddit.com

Description of the dataset

From the online description: “I scraped 1931 recipes from the Food Network that contain the keywords cookies (my group of interest), pastry, or pizza (two control groups). Next I extracted the ingredient list and pooled similar ingredients together (e.g. salt, seasalt, Kosher salt), coming up with a total of 133 unique ingredients. I ended up with a 1931×133 matrix, where each row is one recipe, and each column is whether this recipe contains a certain ingredient (0 or 1).”

Ingredients contain 133 items, from almonds, anchovies, anise and apples to tomatoes, tortillas, vinegar, wine or zucchini.

Locality

Global (perhaps, given the variety of cooking recipes the Food Network website contains)

Environmental variables

Name of variable	Description
type_of_food	A factor with three levels: `Cookies`, `Pastries` vs `Pizzas`

Data for download

File name	File type	Description
cookies-pastry-pizza dataset (everest4ever).xlsx	Excel file	Contains Recipes × ingredients matrix, assignment of recipes to food type, and metadata
cookie_dataset_everest4ever_composition.txt	tab-delimited txt format	Recipes × ingredients matrix (1931 recipes in rows, 133 ingredients in columns)
cookie_dataset_everest4ever_type.txt	tab-delimited txt format	Type of food (a single column with 1931 rows, values: `Cookies`/`Pastries`'/Pizzas'')

Compositional and environmental data (for all species)

```
recipes.ingr <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/cookie_
dataset_everest4ever_composition.txt', row.names = 1)
recipes.type <- read.delim
('https://raw.githubusercontent.com/zdealveindy/anadat-r/master/data/cookie_
dataset_everest4ever_type.txt', row.names = 1)
```

References

- *everest4ever* posted on www.reddit.com on 2017/12/18. Original data are stored here.