

The transnational migration of ingredients

Therese F. Martin, D.B.A.

Notebook

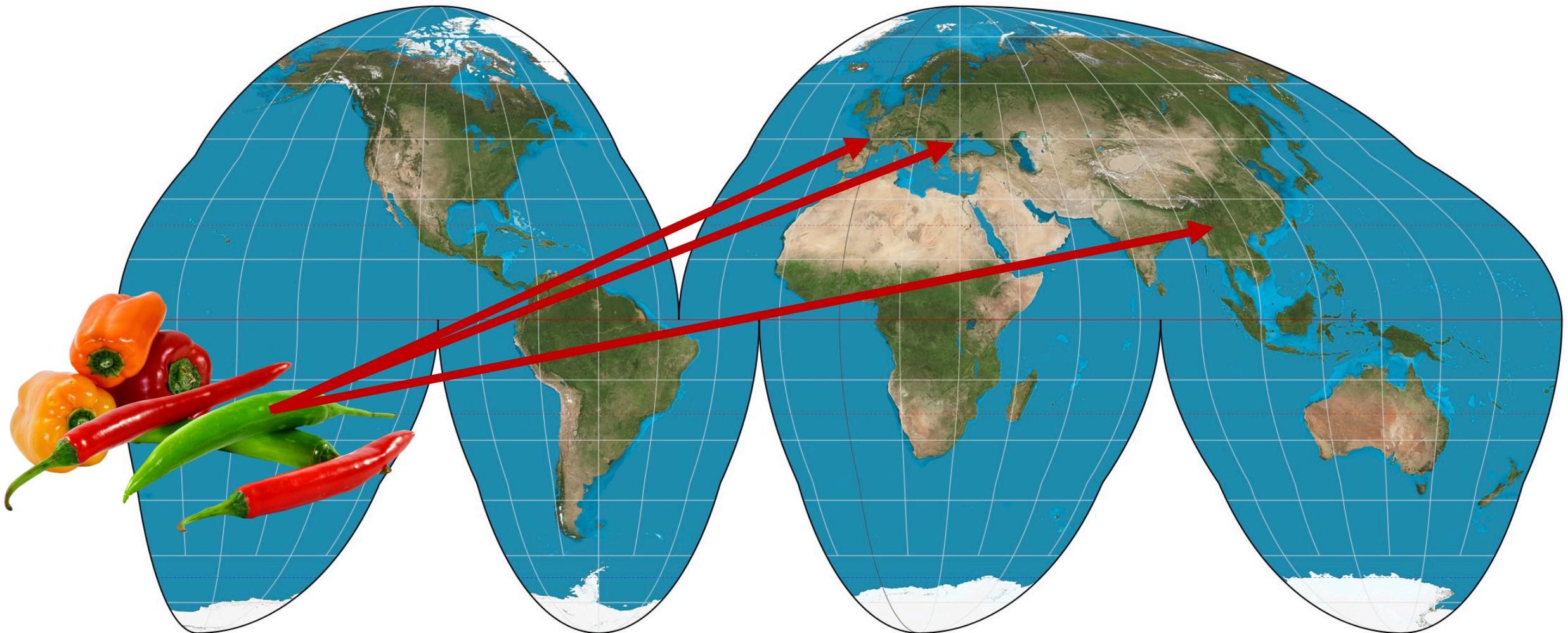
The Data Incubator Application

Ingredients that travel





Where's the pepper?





Data Source

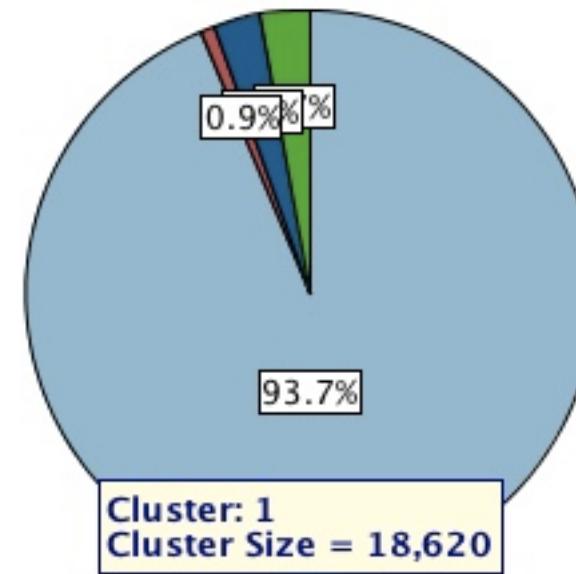
- Kaggle.com
- Hugo Darwood's Epicurious - Recipes with Rating and Nutrition
 - Over 20k recipes 55.3MB in CSV,
 - Metacolumns include garlic, red pepper, chili pepper, green pepper
- Next Steps
 - Expand ethnicity classification
 - Create variables for commonly used ingredients, such as the onion family
 - Rerun regressions
 - Rerun cluster analysis
 - Map clusters



Clusters

This preliminary examination begins to cluster the data into two groups. Expanding the ethnicity list & the pepperiness scale and including other variables, such as the onion family, will create more clusters.

Cluster Sizes



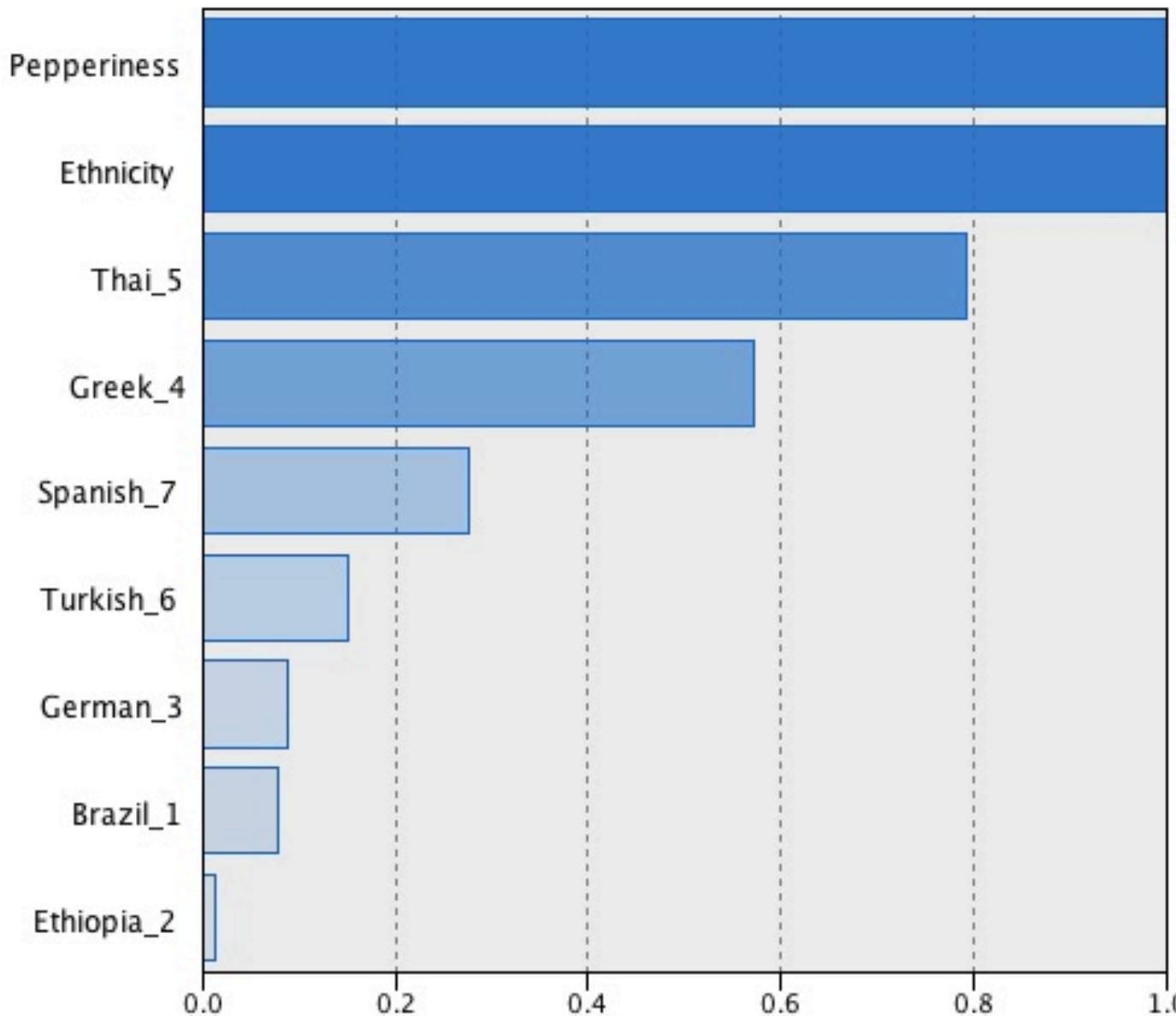
Cluster

- 1
- 2
- 3
- 4

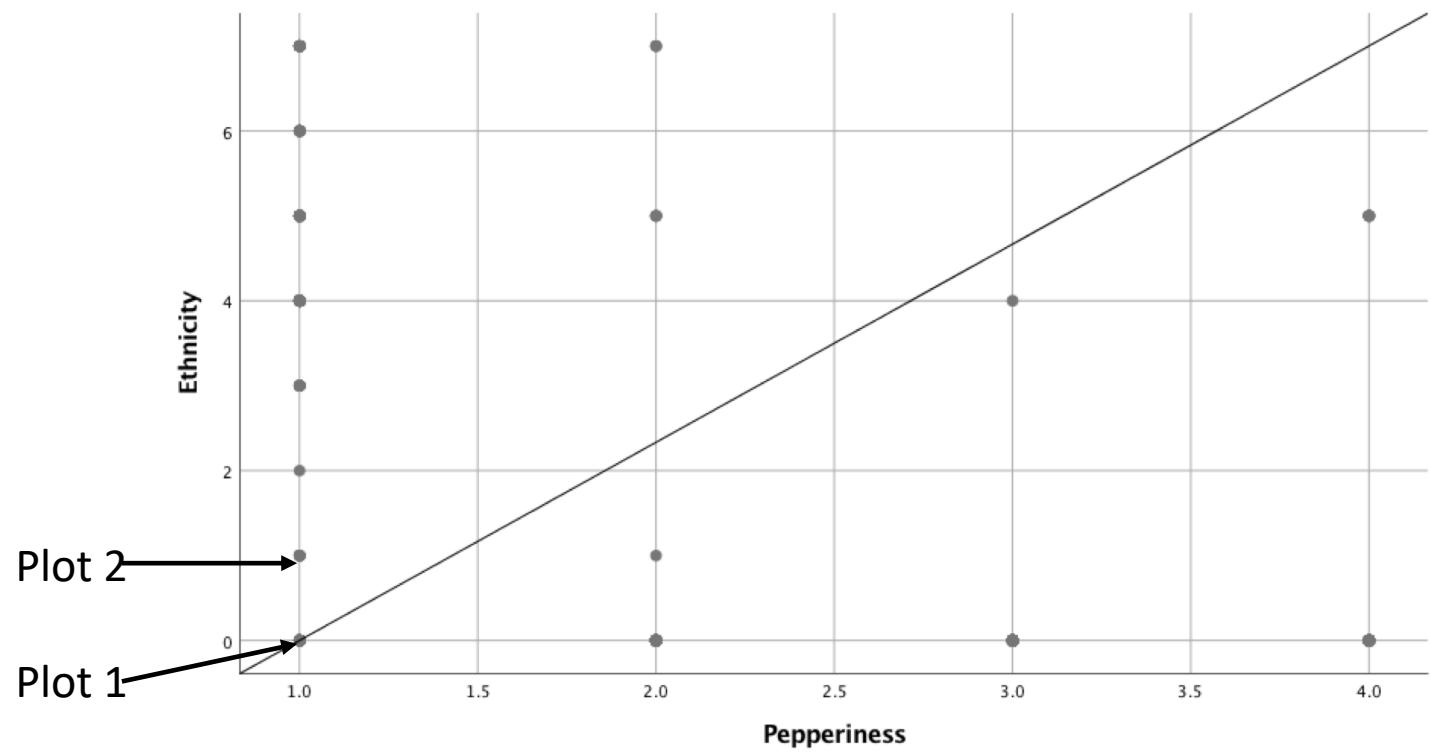
Size of Smallest Cluster	178 (0.9%)
Size of Largest Cluster	18620 (93.7%)
Ratio of Sizes: Largest Cluster to Smallest Cluster	104.61

Predictor Importance

Pepperiness and ethnicity have substantial influence over the model.



Scatter Plot





This preliminary examination suggests that pepperiness predicts the ethnicity of the food.

Regression Plots

Normal P-P Plot of Regression Standardized Residual

