Taste Does Not Endure: The Dynamics of Ingredient Pairings, Colombia, 1977-2017

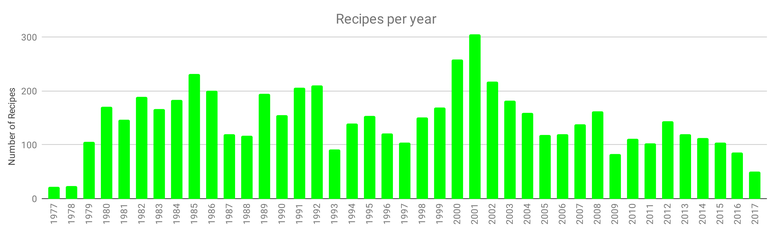
Juan C. S. Herrera, PhD Candidate, Food Studies, New York University, jsh501@nyu.edu

1 Introduction

- Food constitutes the sustenance of humanity. Omnivores need to pair different ingredients in order to survive
- Recipes are one of the best sources of information on food consumption and food pairings
- Network science studies of food recipes provide insight into food culture and food preferences. Network science studies of food recipes have acknowledged the relevance of considering longitudinal approaches but only a few studies have done so
- **How do food recipes change over time ?**
- The findings evidence that taste, as a set of ingredient combinations, does change over time.

2 Data Collection

- Longitudinal study of food recipes using self-collected data from a Colombian magazine spanning 1977-2017
- **Data identification**: The recipes were identified by manually examining magazines through in situ archival research at Biblioteca Luis Angel Arango and Biblioteca Nacional de Colombia. The most recent figures, for 2012 and 2013, estimate that 795,862 copies of this magazine were published nationally every other Thursday (Prensanet 2012). That is equivalent to roughly 1.5% of the Colombian population, or 6% assuming four readers per magazine.
- **Data digitalization**: Each recipe was digitalized by taking a photograph using an iPhone 6 camera. Sometimes two or more photos were taken in order to have a backup and guarantee readability. The image dataset with a filesize of 27.1GB was securely stored.
- **Creation of the dataset**: Initially, several OCR methods were tested, nevertheless recipes appearing on the gutter, glossy paper, text in different orientations, text juxtaposed with images, and recipe text juxtaposed with text from text that didnt belong to the recipe made OCR processing cumbersome. For this reason, this process was completed manually by transcribing the variables from the image dataset. The dataset contains: date, recipe name, ingredients
- **Data cleaning**: The dataset was cleaned to remove typographical errors, a result of the manual transcription. Other considerations for the data cleaning process were: removing blank spaces, use lowercase, removing punctuation, removing duplicates, separating ingredients (oftentimes recipes call for salt and pepper), in order to ensure a consistent dataset. Furthermore, ingredients were standardized.
- **Translation**: Since the recipes were originally published in Spanish, each ingredient was translated into English. The translation was initially done using Google translate. Afterwards it was revised. In the cases were there was no translation available the original name in Spanish was used. Recipes names were not translated.
- **Final Dataset**: 5,981 recipes, 559 unique ingredients, 41 years, 492 months



3 Methods

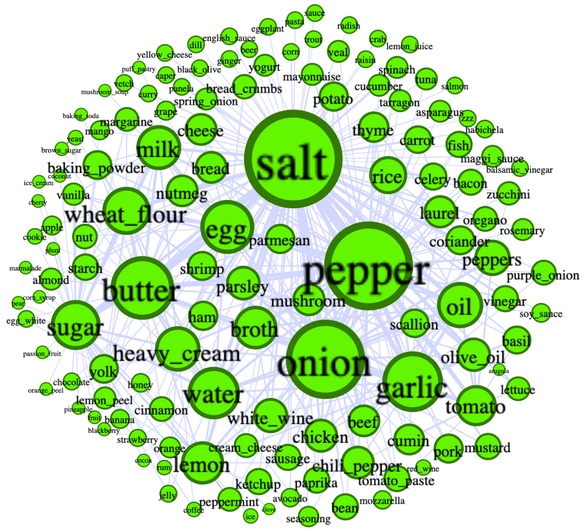
- 1. Creation of different bipartite recipe - Ingredient networks:
  - 1977 - 2017 Network
  - Cumulative Networks for different years and months
  - non-Cumulative Networks for different years and months
- Two-mode to one-mode projections using different techniques: Euclidean, weighted, Jaccard, etc
- Characterization of the network using backbone extraction and minimum threshold techniques
- Longitudinal network analysis. Example: Ingredients; Meats.

conference tip: If you eat *arepa de huevo* in Cartagena, *you should!* season it with plenty of "hogao" and "suero costeño".

4 Results

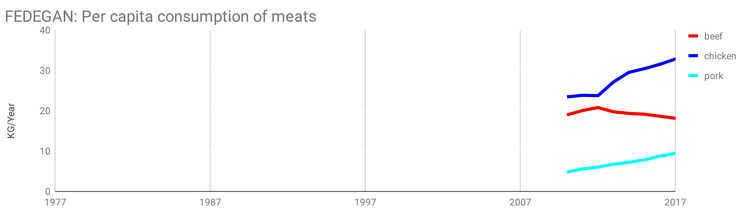
4.1 1977-2017 Ingredient Network

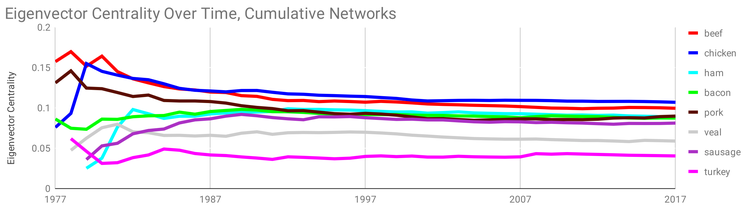
- 1977-2017 aggregated Ingredient-Ingredient Network.
- The nodes with the highest degrees are: Salt, pepper, onion, butter, garlic, tomato, sugar, are some of the nodes with the highest degrees.
- Note (in the lower right corner) the ingredients for *guiso* or *hogao*: garlic, onion, oil, pepper, tomato, olive oil, scallion (optional). This is an ubiquitous sauce that is used to cook dishes as well as to accompany them[1] This visualization presents the backbone of the ingredient network, each node represents an ingredient. The size of the nodes represents the degree. The size of the edges is relative to its weight.



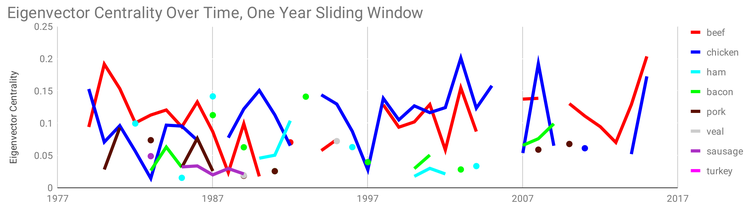
4.2 Recipes as an approximation to food consumption

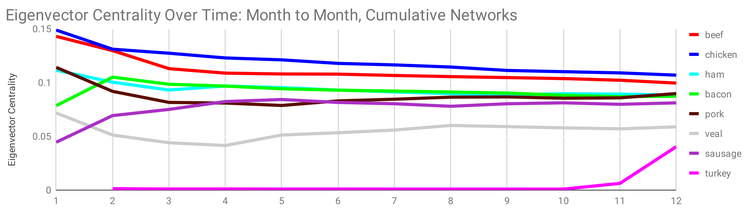
- Recipes can be used as source of information to understand food consumption

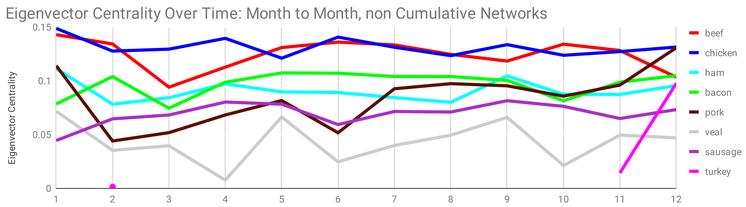




4.3 Taste Does not Endure: Seasonal Ingredients and Trends







5 Limitations

6 Conclusion

-

THANK YOU
Contact Information:
jsh501@nyu.edu


Universidad de los Andes
Colombia
Facultad de Ingeniería
LA NET
Social and Health Complexity Center
EpiAndes
Grupo de Epidemiología de la Universidad de los Andes
Universidad de Ibagué
LATIN AMERICAN CONFERENCE
2.0 ON COMPLEX NETWORKS
CARTAGENA - COLOMBIA
AGOSTO 05-09
INSCRÍBETE
HTTPS://LANET2019.UNIANDES.EDU.CO/
DÍA 1
-Escuela Internacional
DÍA 2
-Escuela Internacional
-Bienvenida a la conferencia
-Conferencia Día 1
DÍA 3
-Conferencia Día 2
-Recepción y sesión de posters
DÍA 4
-Conferencia Día 3
-Sesiones de minisimposios
-Cena del congreso
DÍA 5
-Conferencia Día 4
-Cierre y ceremonia de premiación
-Tour por Cartagena
]