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 juancsherrera

NYU

1 Introduction

- Recipes are one of the best sources of information on food pairings, taste, and food consumption
- Ingredients are paired together to form recipes that satisfy omnivore's need for variety
- Ingredient pairings reveal taste which is the basis of food consumption
- Network science studies of food recipes provide insight into food pairings by approaching recipes as complex interconnected systems
- These studies have acknowledged the relevance of considering longitudinal approaches but only a few have done so
- Longitudinal studies of food recipes can illuminate changes in food pairings
- As we rarely eat the same food for over a couple of days, network science studies of food recipes should account for the dynamics of food pairings

2 Research Questions

- **How do ingredient pairings in Colombia change over time?**
- Which are the advantages of using longitudinal datasets of recipes to study food recipes to understand food consumption?
- Is is possible to identify changes in taste using network science to study food recipes?

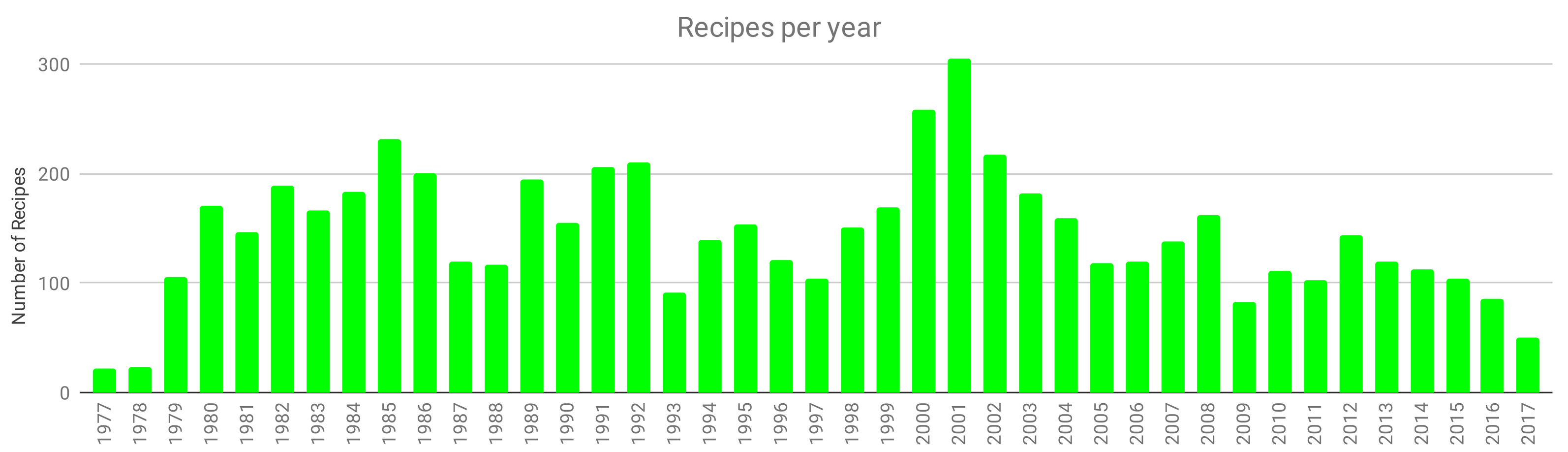
3 Methods

3.1 Network Analysis

- 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, using different sliding windows (i.e. 1 year, month to month, etc.)
- Two-mode to one-mode projections using different techniques: Euclidean, weighted, Jaccard, etc.
- Characterization of the network using backbone extraction and minimum thresholds
- Longitudinal network analysis. Example: Ingredients: Meats.
- Longitudinal network statistics. Example: Eigenvector Centrality

3.2 Dataset

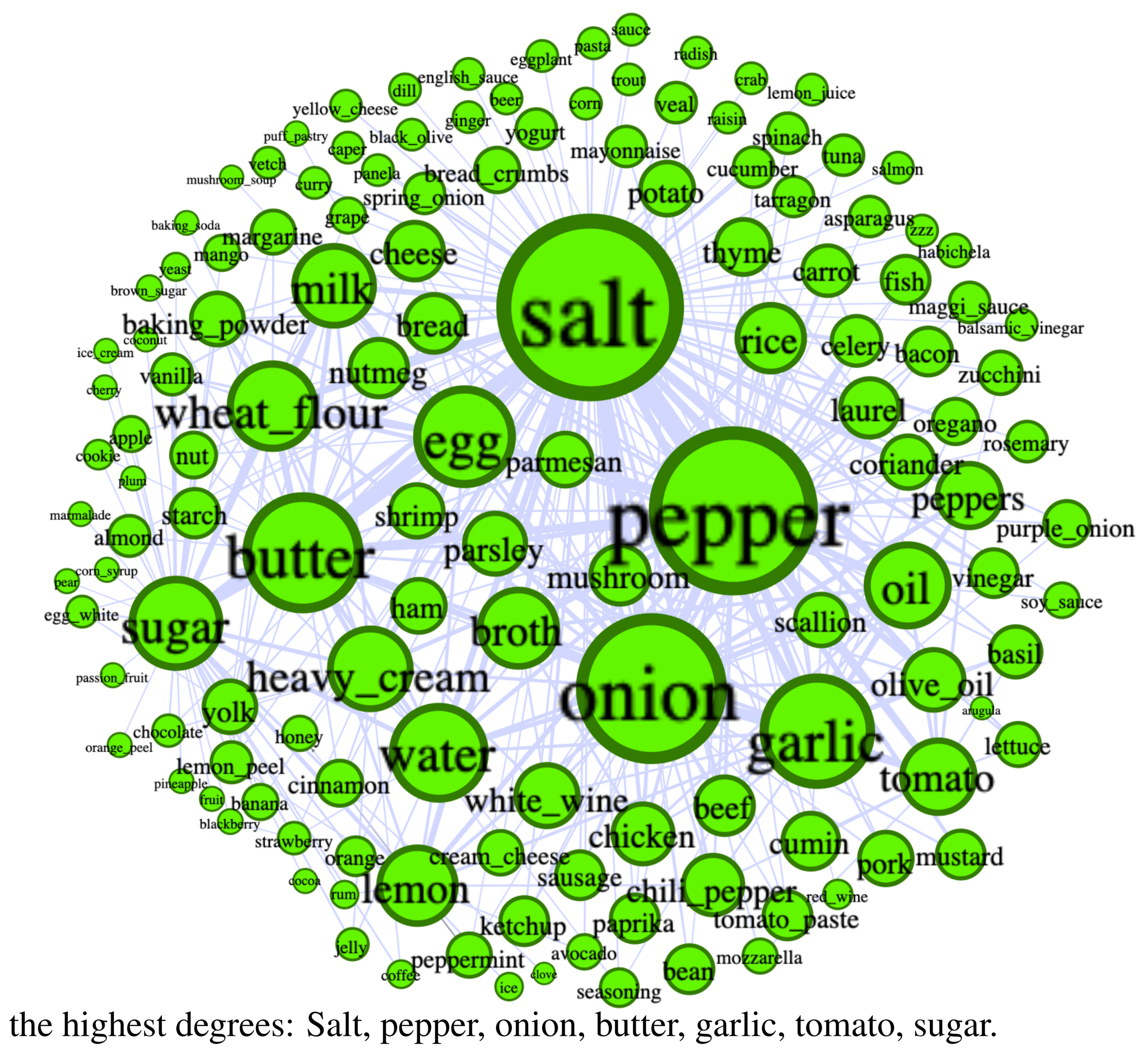
- Data Source: Recipes collected from *Revista Carrusel*
- Recipes published on this magazine come from a variety of sources: in-house editors, books, local restaurants, TV shows, sponsored recipes, etc.
- Data digitalization: Each recipe was digitalized, image dataset: 27.1GB
- Creation of the dataset: Image to dataset using manual transcription. The dataset contains: date, recipe name, ingredients
- Data cleaning: Data cleaned to remove typos, duplicates, etc.
- Ingredient standardization
- Translation: Translation done using Google translate and manual revision
- **Final Dataset**: 5,981 recipes, 559 ingredients, 41 years, 492 months



4 Results

4.1 1977-2017 Ingredient Network

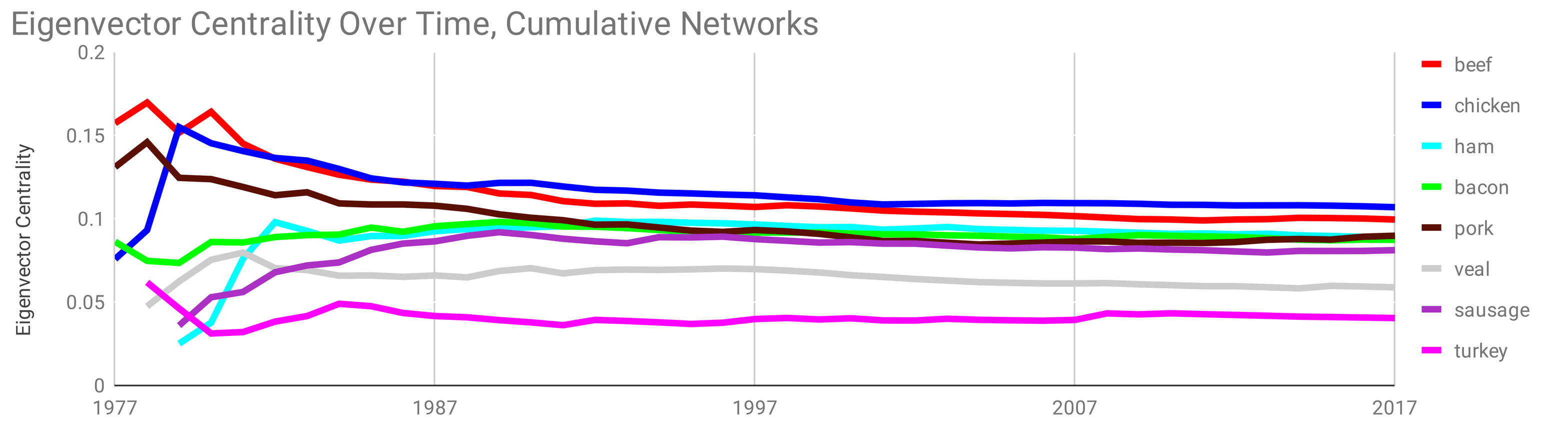
- Backbone of the ingredient network, each node represents an ingredient. Size of the node indicates its degree. Size of the edges are relative to their weights.

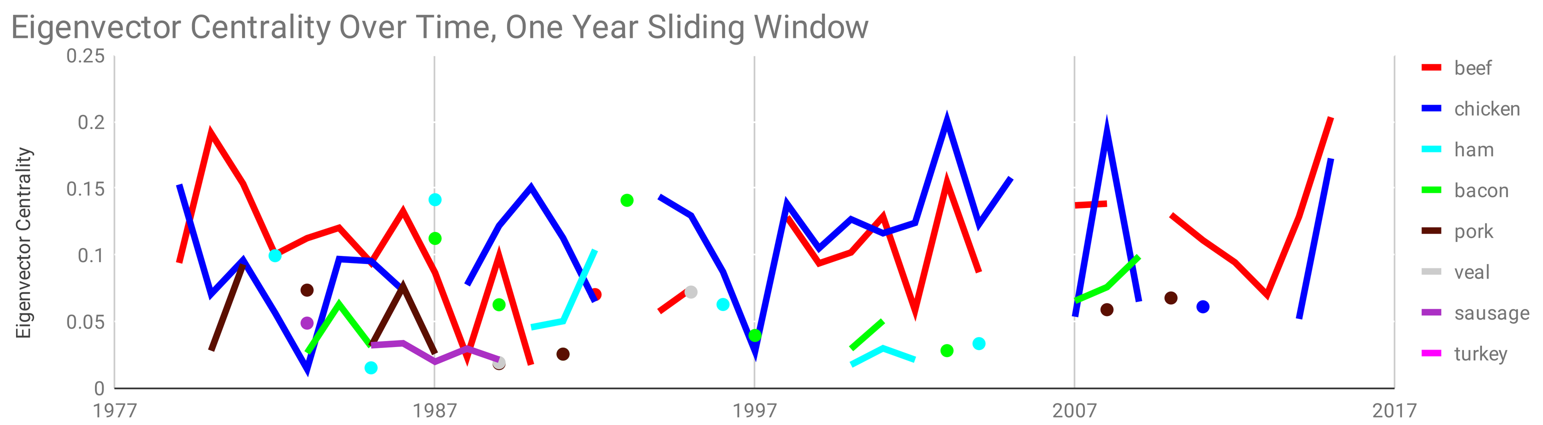


- Nodes with the highest degrees: Salt, pepper, onion, butter, garlic, tomato, sugar.
- Note the ingredients for *guiso* or *hogao*: garlic, onion, oil, pepper, salt, tomato, olive oil, scallion (optional). This is an ubiquitous sauce that is used to cook dishes as well as to accompany them

4.2 Recipes as an approximation to food consumption

- Approximating food consumption using network analysis of food recipes: Example: Meats
- The following graphs show the eigenvector centrality over time for the eight most central meats
- The first figure is computed on an aggregated network The second is computed using a one year sliding window, that is one network per year

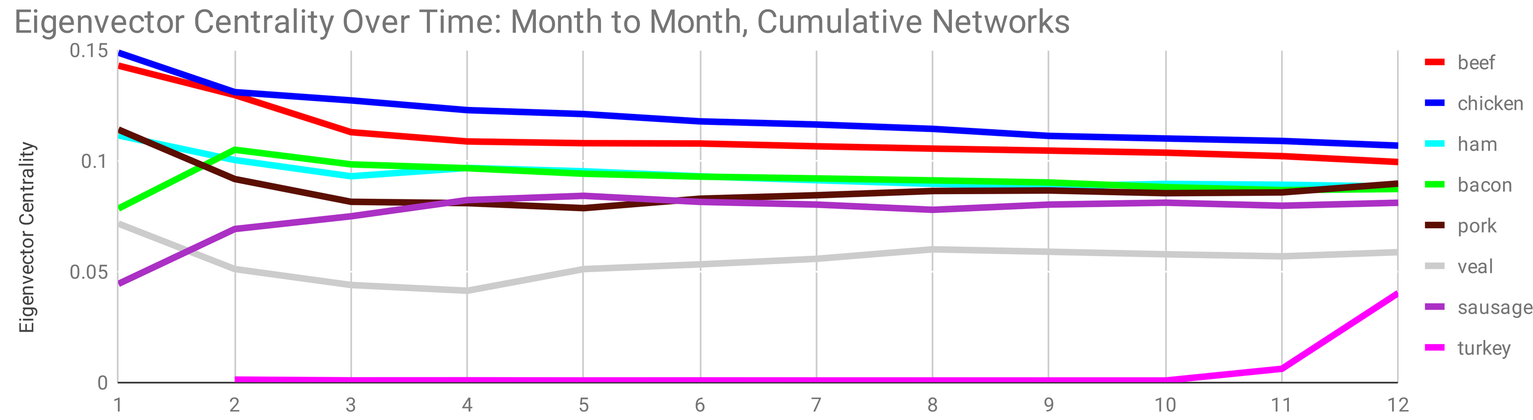


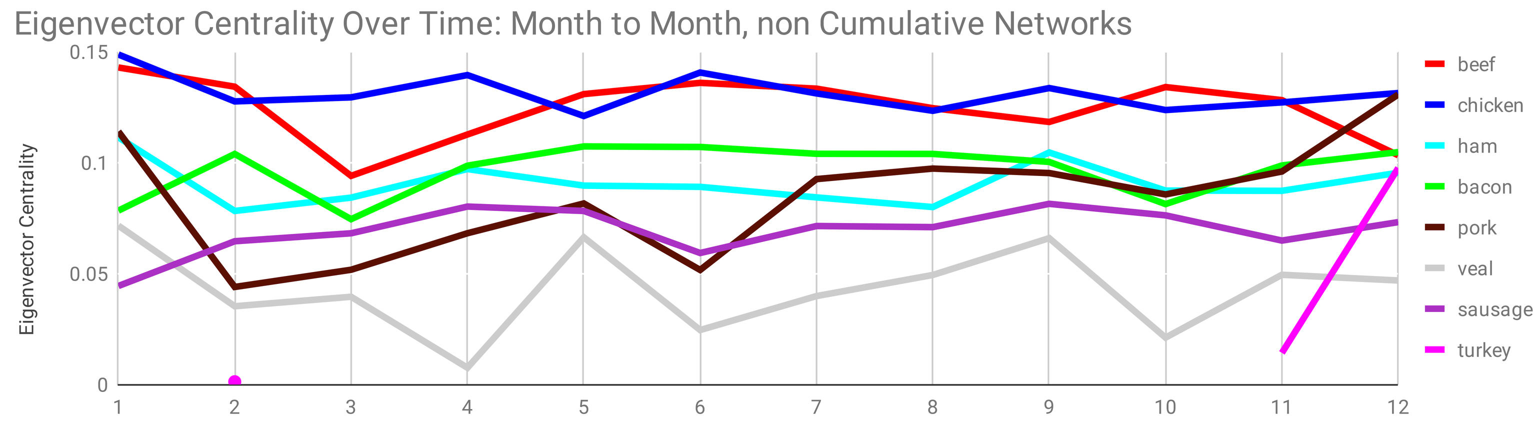


- Chicken is the most central meat, followed by beef and pork (particularly if pork and bacon were added together)
- This hierarchy reflects the statistics on consumption of meats according to Fedegan

4.3 Seasonal Ingredients and Trends

- Long term centrality measures capture general food consumption
- Seasonality appears after aggregating the networks by month





- Turkey's centrality spikes around Christmas
- Beef declines between February and April, during Lent
- These results are surprising as Colombia does not have traditional seasons, it is located close to the Equator
- Seasonality should play a larger role in countries further away from the Equator

5 Limitations

- Dataset is comprised of recipes which appear in a paid publication thus they are biased
- Dataset contains recipes that encompass multiple cuisines
- Food recipes may be leading or lagging in culinary innovation; leading chef's social media accounts could provide a better source of information to study popular ingredient trends

6 Conclusion

- Ingredient pairings in Colombia change over time by providing seasonality in a country that is located on the Equator. This may have it's roots in omnivore's need for variety and innovation
- Network analysis is a method to systematically analyze food recipes and thus gain an understanding of food pairings, taste, and food consumption and their development over time
- Network statistics can identify changes in ingredient pairings. This can be done using different approaches such as aggregated networks and sliding windows
- Network analysis studies of food recipes should consider changes in taste
- AI and Machine learning applications that use network analysis such as computational creativity systems aiming to develop food recipes should use longitudinal analysis to produce recipes that adapt to omnivore's desire for variety by using seasonality and accounting for changes in taste
- Deviations from longitudinal patterns could help identify fads and trending topics such as ingredients that are popular in a given moment

7 Future Work

- The examples using meat and centrality show part of larger analysis that entails more ingredients
- These examples centered around single nodes, I am working on dyads, triads, and community detection
- Modify the existing copy mutate model for recipe evolution to account for seasonality and ingredient trends
- Design an ingredient recommendation system that accounts for seasonality and trends

THANK YOU