Recipe and Ingredients Analysis using Networks

Group: 13

19Z213 - Gondi Rajeev
19Z223 - Koushik Balaji
19Z224 - Lohith Sowmiyan
19Z238 - Sairam Vaidya M
19Z257 - Venighalla Akhil

Problem Statement

The aim of the project is to analyze the different ingredients present in different types of recipes using three networks. Firstly, a complement network for ingredients that co-occur frequently, followed by a standard ingredient network for performing ingredient analysis, and finally a bipartite-graph to represent recipe-ingredient pairs for analysis. The project aims to identify the following,

1. Ingredient Complements
2. Community Detection
3. Niche Complementary Ingredients
4. Associate Rule Mining
5. Degree Analysis
6. Page-Rank Algorithm
7. Similarity between Recipes
8. Healthy Alternative Recipe Recommender

Dataset Description

The dataset has a total of 256 recipe names. Each recipe consists of,

- Recipe: Name of the recipe.
- Ingredients: It is one of the items of food you need to make a recipe.
- Diet: It specifies whether the recipe is vegetarian or non vegetarian.
- Preparation time:Time taken to prepare food.
- Cooking time: Time taken to cook food.
- Flavor profile: Flavor of the dish.

- Course:Specifies which type of recipe.
- State: state in which the recipe is used the most.
- Region: In which region of India the recipe used the most.

Link for dataset:

https://www.kaggle.com/datasets/nehaprabhavalkar/indian-food-101

Tools Used

- **Networkx**: It is used to create, manipulate, study the structure and functions of difficult graph networks.
- **Gephi**: It is an open source software for network visualization and analysis.
- **Pandas**: Python library for loading, viewing, exploring, and manipulating datasets.
- **Mlxtend**: A library of Python tools and extensions for data analysis.
- **Rapid API**: It is used to discover and connect to thousands of APIs.

Challenges Faced

- **Dataset**: The initial and major challenge we faced was in preparing the dataset. It posed various challenges as the information of recipes were ambiguous in almost all datasets and because of the ambiguity the analytics were not providing intended results.
- To counter this challenge we need to apply advanced natural language processing techniques to transform jargon into a common standard. Eg Extra hot chili needed to be converted to chili. This also required high computational power and time.
- **Health index:** To get the health associated parameters for a particular recipe we relied on a rapid api which contained nutritional information on 100000 recipes. But the dataset contained recipes that were not in the api, thereby limiting the scope of the functionality.

Contribution of Team Members

Roll No.	Name	Contribution
19Z213	Gondi Rajeev	Community Detection, Healthy Alternative Recipe Recommender
19Z223	Koushik Balaji P	Bipartite Graph Creation, Degree Analysis, Bipartite Matching
19Z224	Lohith Sowmiyan	Associate Rule Mining, Page-Rank Algorithm
19Z238	Sairam Vaidya M	Gephi Analysis - PMI Ingredient Network, Standard Ingredient Network, Recipe-Ingredient Bipartite Graph Analysis
19Z257	Venigalla Akhil	Dataset collection, Dataset preprocessing

Annexure I : Code

Colab Notebook

Annexure II : Snapshots of the Output

1. Gephi Analysis,

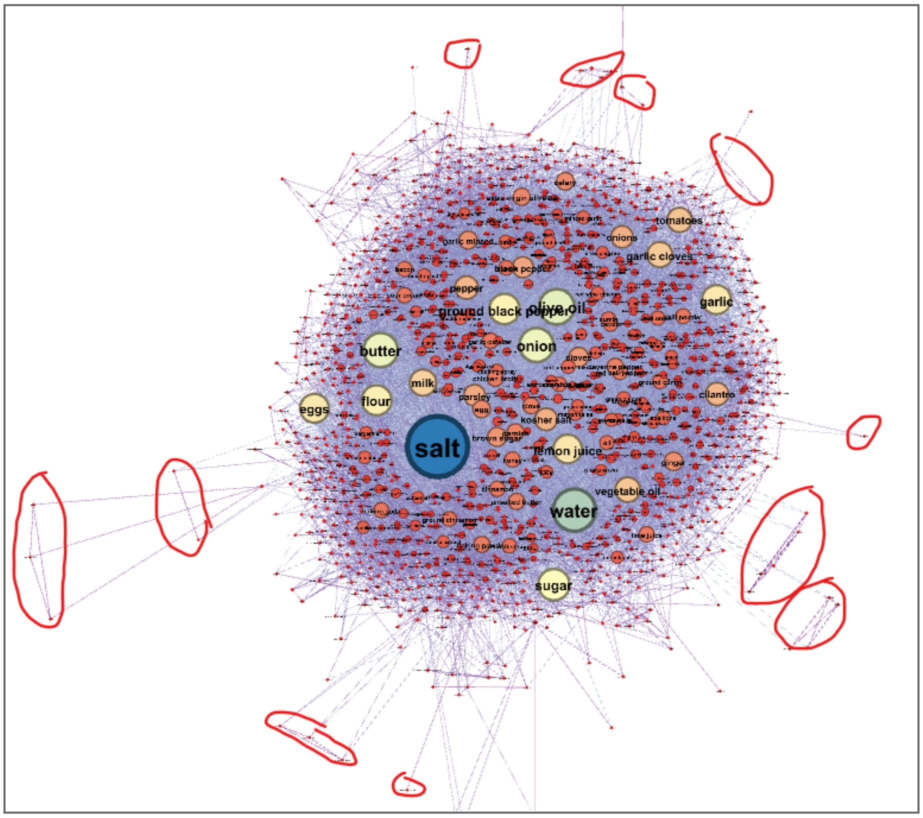

Fig 1. Identifying niche complementary ingredients using PMI based Ingredient Network

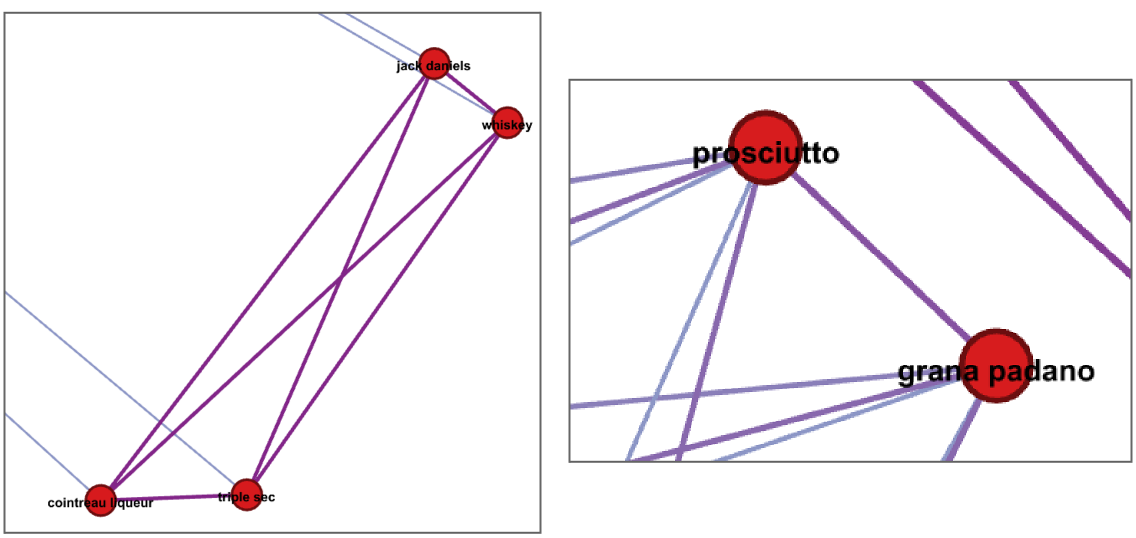

Fig 2. Sample complementary ingredients

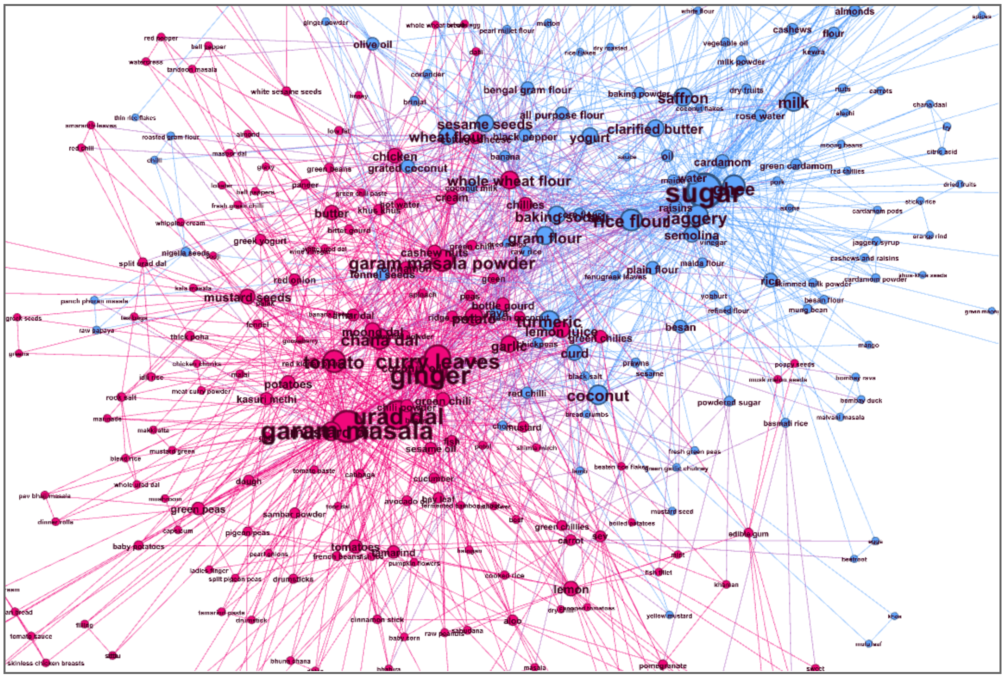

Fig 3. Two communities identified in the ingredient graph - savory and sweet

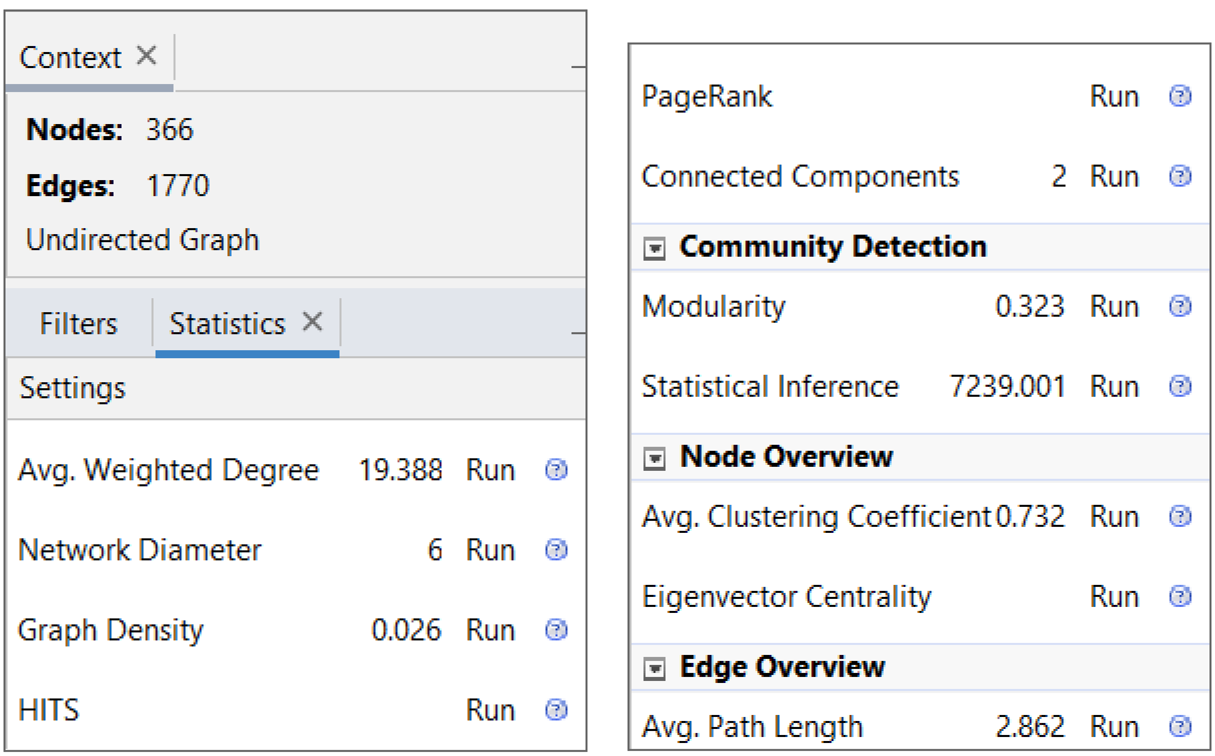

Fig 4. Statistics for the Ingredients Graph

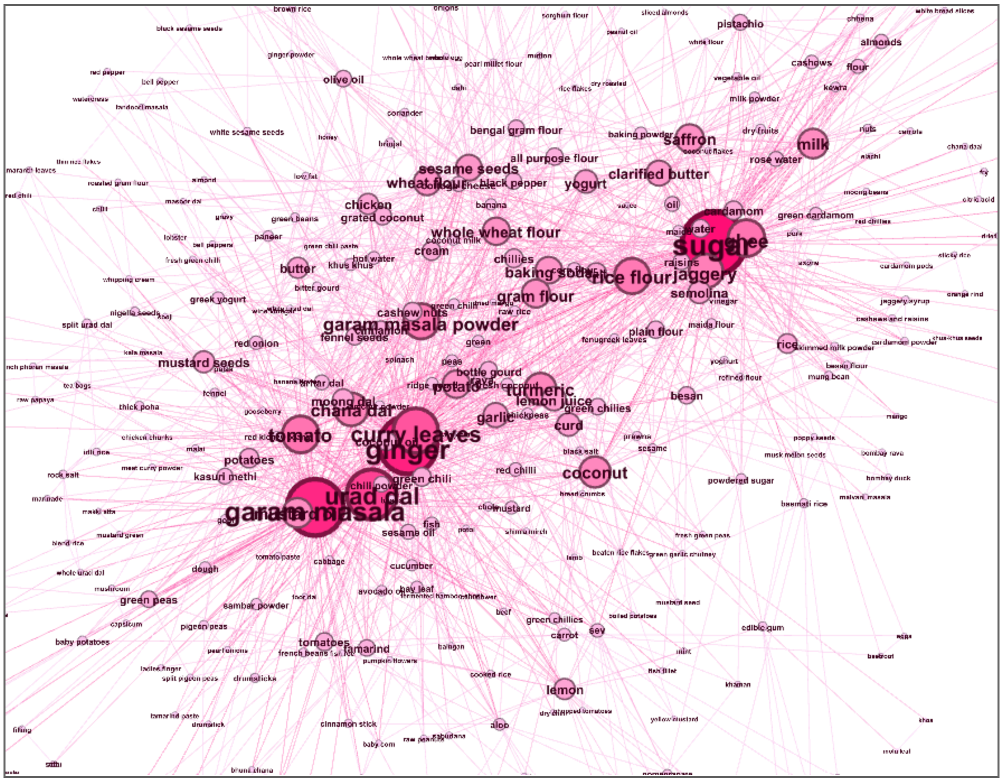

Fig 5. Degree Centrality in Ingredient Network

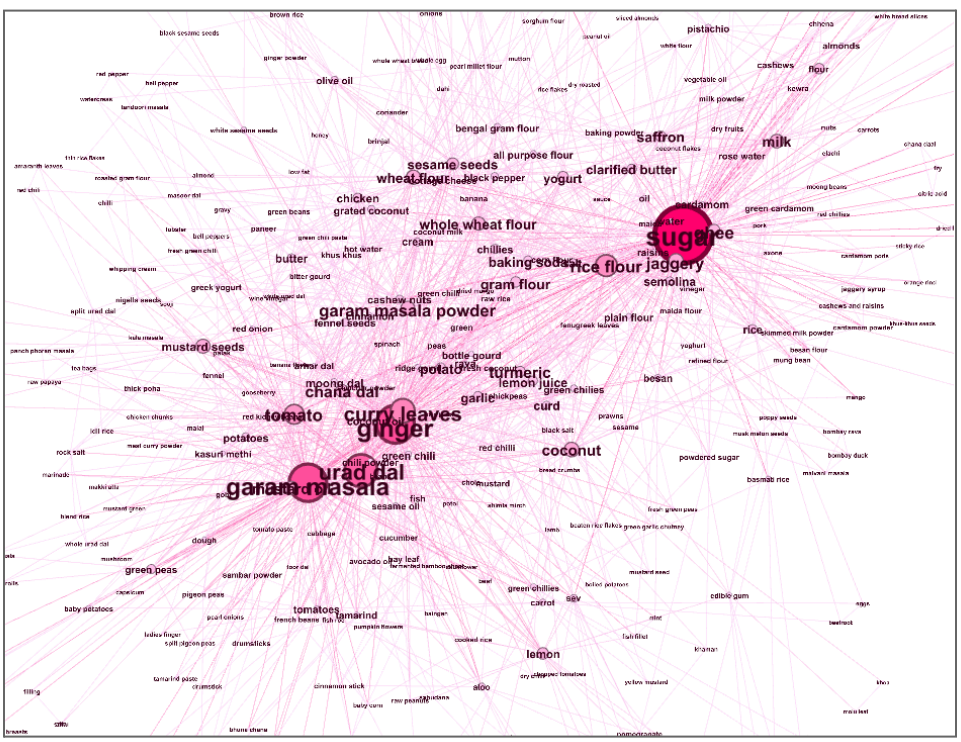

Fig 6. Betweenness Centrality in Ingredient Network

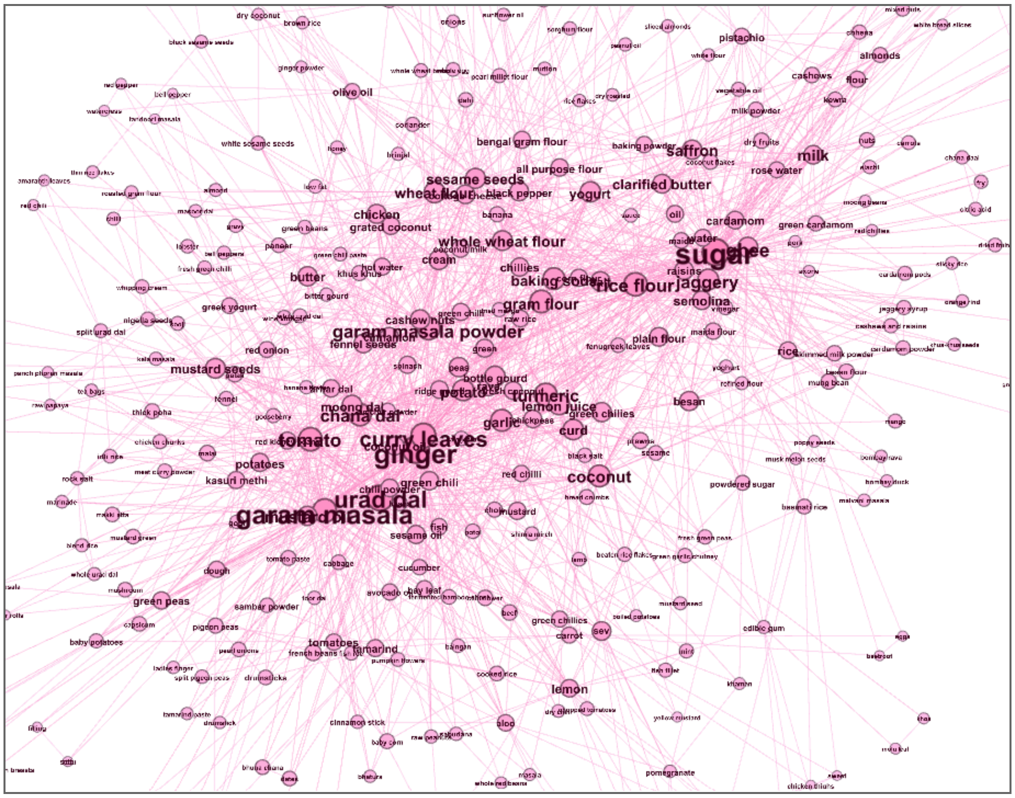

Fig 7. Closeness Centrality in Ingredient Network

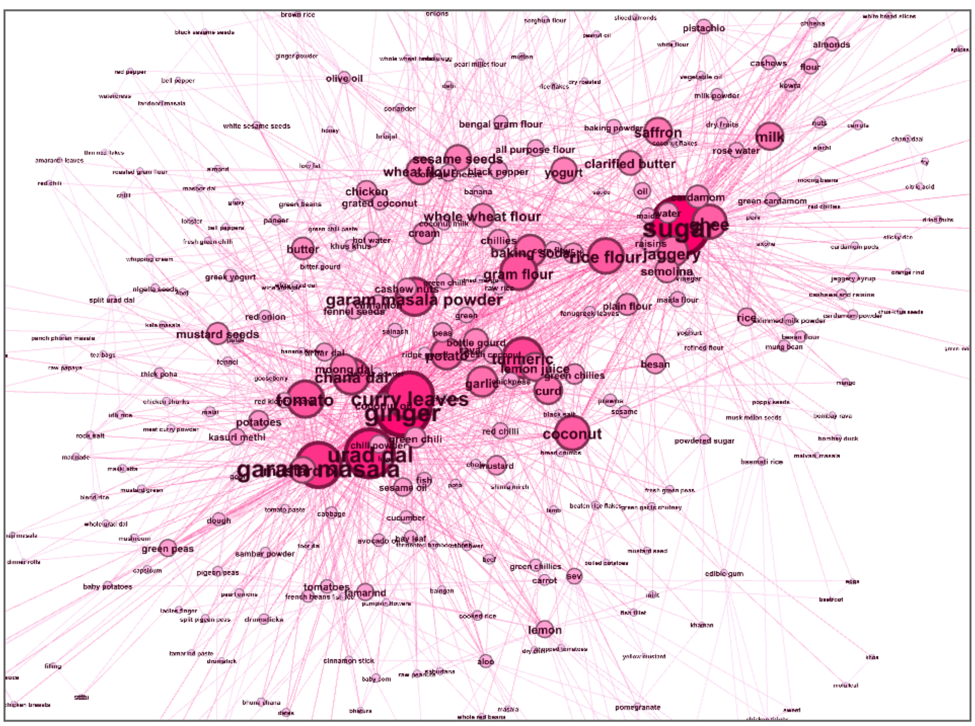

Fig 8. Eigenvector Centrality in Ingredient Network

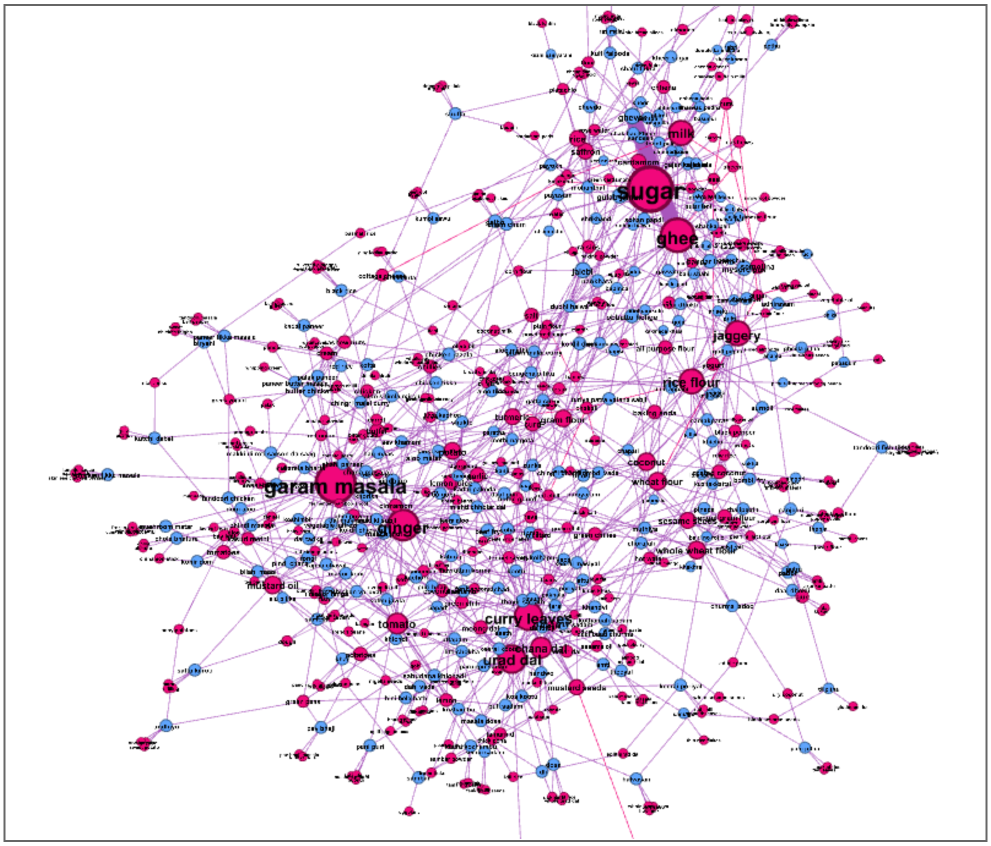

Fig 9. Bipartite Graph for Recipe (Blue) - Ingredients (Red)

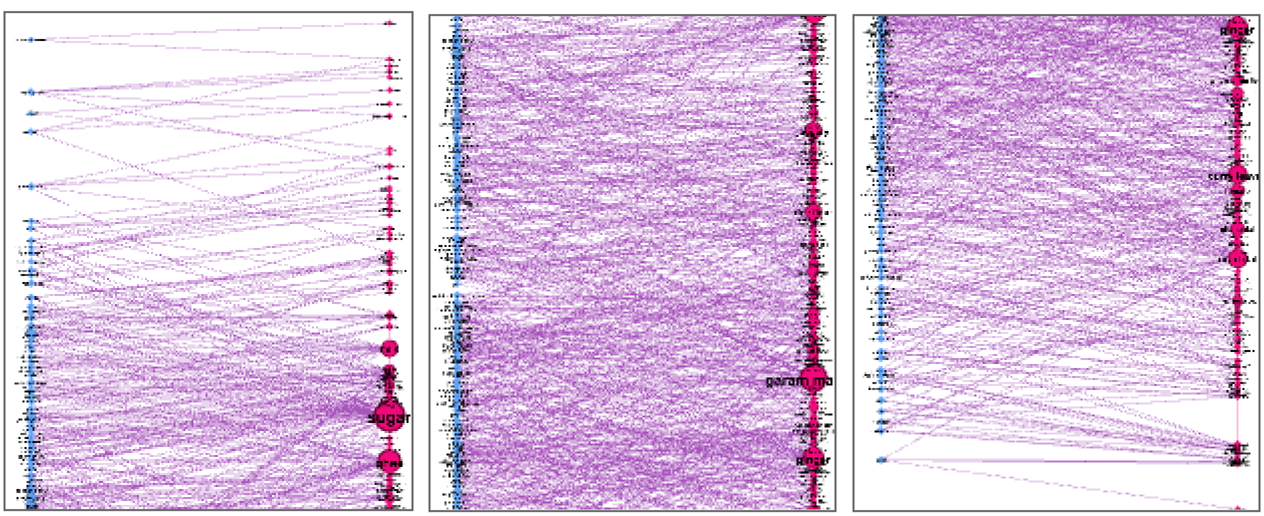

Fig 10. Ordered Bipartite Graph using Customized Event Graph Layout

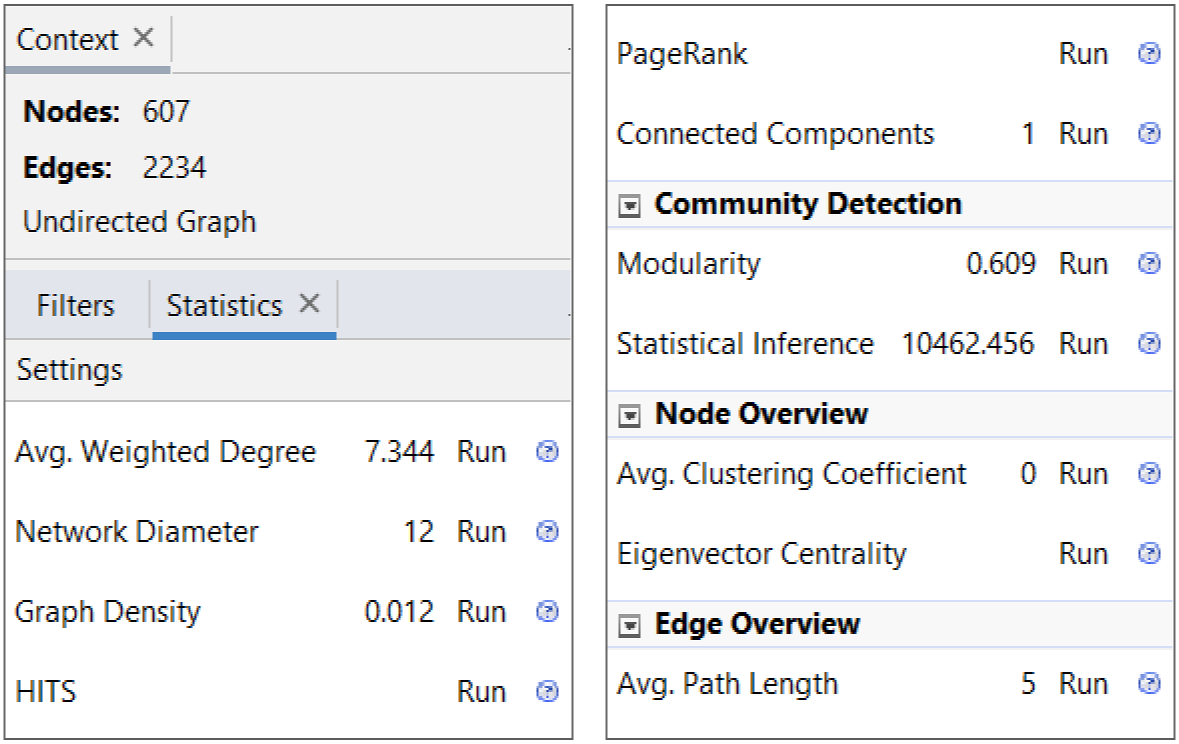

Fig 11. Statistics for the Bipartite Recipe-Ingredient Graph

References

1. https://networkx.org/
2. https://link.springer.com/article/10.1007/s10994-020-05898-0
3. https://gephi.org/
4. https://focaalvarez.medium.com/shake-it-up-network-analytics-of-cocktail-ingredients-2be5c3a8d3d2
5. https://www.askpython.com/python/examples/bipartite-graph-in-python