1. Title

    - Generative Data Art using Restaurant Data Network Analysis
    - Generative Art Is Challenging What It Means to Be Human having taste-buds
    - Pick the right Restaurant for your cuisine through depicted Generative Art
    - Creative Taste buds through generative art amalgamation of Restaurants
    - Aesthetic and economic value by gastronomic guides using Generative Art
    - Generative Data Art created using Restaurants Data in Bangalore
    - Using Restaurants Data to Generative Data Art
    - Network Analysis and Generation of Data Art from Restaurants Data

2. Abstract

Tastemakers lay down the rules of what constitutes good or legitimate taste and influence the identifications and practices of producers. This paper examines the roles played by gastronomic critics/guides as tastemakers in the fine-dining restaurant industry of our Indian origins in Bangalore. A notional exploration—drawing on the turfs of both sociology of culture and cultural economy relating to singular cultural goods—states what cultural influence is, how it is achieved and what symbolic and material consequences it may have. The empirical section examines three gastronomic guides in the restaurant business industry based on Zomato data classification.

The paper employs qualitative research methods trusting on data sources: an analysis of Zomato data sets, and data archive of all Restaurants with star ratings of similar cuisines and palate taste-buds, further using Data Processing, extracting embeddings and similarity-based matching, use of network analysis and community discovery visualization and the '**Data Art**.'

3. Introduction

Exquisiteness is in the judgment of the beholder. At least that is a cliché spewed by wingchair philosophers as they grapple with ideas on what imbues value on any form of imaginative expression. Valuing pieces in the traditional physical art world has always vexed casual observers, collectors, and experts alike.

Contemporary art has high historical annual volatility of 25.8%, indicating the difficulty in objectively determining value. Additionally, approximately $200 billion is spent annually on physical art, of which $6 billion is estimated to be fraudulent or tainted by illicit activity. Generative art is a blanket term for any art form employing algorithmic or random process, as mentioned in Brian Eno's 77 Million Paintings, which composites Eno's visual art in such a way that there are 77 million possible combinations.

Other generative pieces are visualizations or realizations of abstract data sets or additional information. While this might sound a bit arcane, the production process (generative processes) can be a valuable addition to any designer's skill set, particularly for automating tedious tasks like creating "randomized" artwork or simulating realistic natural phenomena challenging to copy, like physical landscapes, intentionally. It can also be a prodigious way to rethink how you use your existing tools, like Actionscript or Illustrator.

Generative art can be a daunting topic — it seems like much math is involved, and art is tricky in itself. But it does not take to be intricate — you can build some cool things without a math or art degree.

This research project aims at creating '**Generative Data Art**' using the '**restaurants dataset**.' The dataset has been crawled from zomato.com. The data containing fields such as cuisines, menus and reviews were then converted to network graphs using natural language dispensation and an unsupervised machine learning approach.

The research further contains the community detection analysis for various insights in the created network graph, which has been visualized using beautiful generative data art.

4. Practice and Methodology

Both gustatory and metaphorical taste is about the closeness of pleasure or displeasure attending experience, and both concern matters of relative preference and standards for appealing discrimination. Haute gastronomy dishes invite aesthetic involvement by producers and consumers, and they demand judgments of both gustatory and metaphorical taste and may give rise to the emergence of tastemakers.

Tastemakers are highly influential individuals or social groups who, by laying down the rules of what establishes good or legitimate taste, may strongly influence aesthetic and economic identifications and practices among consumers and producers of cultural products. Tastemakers, by daunting a canon of rules and standards, establish an aesthetic trend and determine what is legitimate taste. Arbiters of taste may have far-reaching symbolic effects and material consequences for the whole cultural field.

This paper examines the roles played in the process of taste-making by gastronomic guides. It identifies their cultural and material resources and explores their influence on establishing and preserving a taste or culinary style in fine dining. The paper demonstrates that arbiters of taste exert symbolic and material power by attributing value to cultural goods.

It is assumed that the establishment of taste is accomplished interdependently with topmost chefs and diners/consumers. However, the impact of the latter two groups is not the subject of this paper, and the above claim is not empirically substantiated. This paper will concentrate mainly on gastronomic guides' crucial role as taste arbiters.

5. Key Findings of the theoretic outline

How is taste—that is, the size to be discerning and discriminating vis-à-vis a variety of items or experiences—developed and communicated, and what effects may be attributed to tastemakers in the social field of fine-dining restaurants? This section examines how philosophers view the role of gastronomic guides in taste-making.

6. Research method

The study is utilized to identify patterns of restaurant categorization, as well as to gain information on guides' ranking criteria and related practices.

- **Data Selection**

The data has been crawled from zomato.com, which contains 51,717 records and accounts for 8,792 full restaurants and branches across the city. Below are the top ten restaurants with the most units per the dataset.

Cafe Coffee Day	96
Onesta	85
Just Bake	73
Empire Restaurant	71
Five Star Chicken	70
Kanti Sweets	68
Petoo	66
Polar Bear	65
Baskin Robbins	64

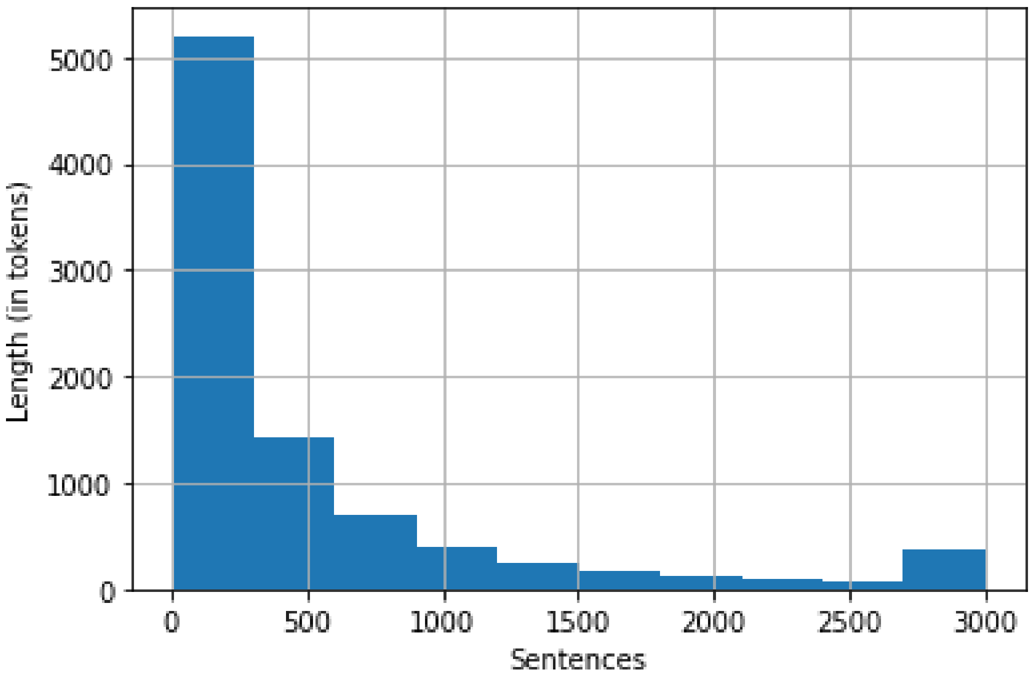
The data was then grouped by the names of restaurants aggregating over cuisines, featured dishes (dish_liked), menu items and reviews. Some of the fields have lots of missing values which have been compromised.

- **Network Creation**

Out of the total of 8,792 data points, here is the list of some missing values in the required fields:

Field	Missing Values	Missing value %
menu_item	5,561	63.25%
cuisines	7	0.0007%
dish_liked	5,597	63.66%
reviews	1,753	19.93%

The above four rows have then been concatenated with a newline separator to create a single text field. The text field is then trimmed to have a maximum of 3,000 tokens (space-separated strings). Below is the brief about the length of text fields in space-separated tokens.



- **Term frequency-inverse document frequency (TF-IDF)** is a text vectorizer that transforms the text into a usable vector. It combines two concepts, Term Frequency (TF) and Document Frequency (DF). The tenure frequency is the number of occurrences of a specific term in a document.

  Thus, the TF-IDF vectorizer has been used to extract embeddings from the text fields. After we have the TF-IDF embeddings, we get the distances and indices of the 20 closest neighbours to each text embedding.

  - A quick experiment was done to decide the number of neighbours. A higher number would yield an extensive network and unnecessary connections, while a lower number would fail to capture more similar restaurants

  - In the Initial approaches, we used pre-trained transformer embeddings instead of TF-IDF vectors, and it did not yield a meaningful result

  - A high similarity score was assigned to each data point. Then we decided on TF-IDF which works by increasing proportionally to the number of times a word appears in a record but is offset by the number of records that contain the word, which makes more sense here as the language is not much diverse here in terms of context and word choices and it is too simple a use-case for transformers

  After we have the nearest 20 neighbours for each Restaurant, a quantile-based scoring is done to provide weight to the network, e.g. the close-connected nodes are assigned a weight of 4, while nodes connected the furthest are assigned a weight of 1. The network here is undirected.

- **Network Stats and Interpretations**

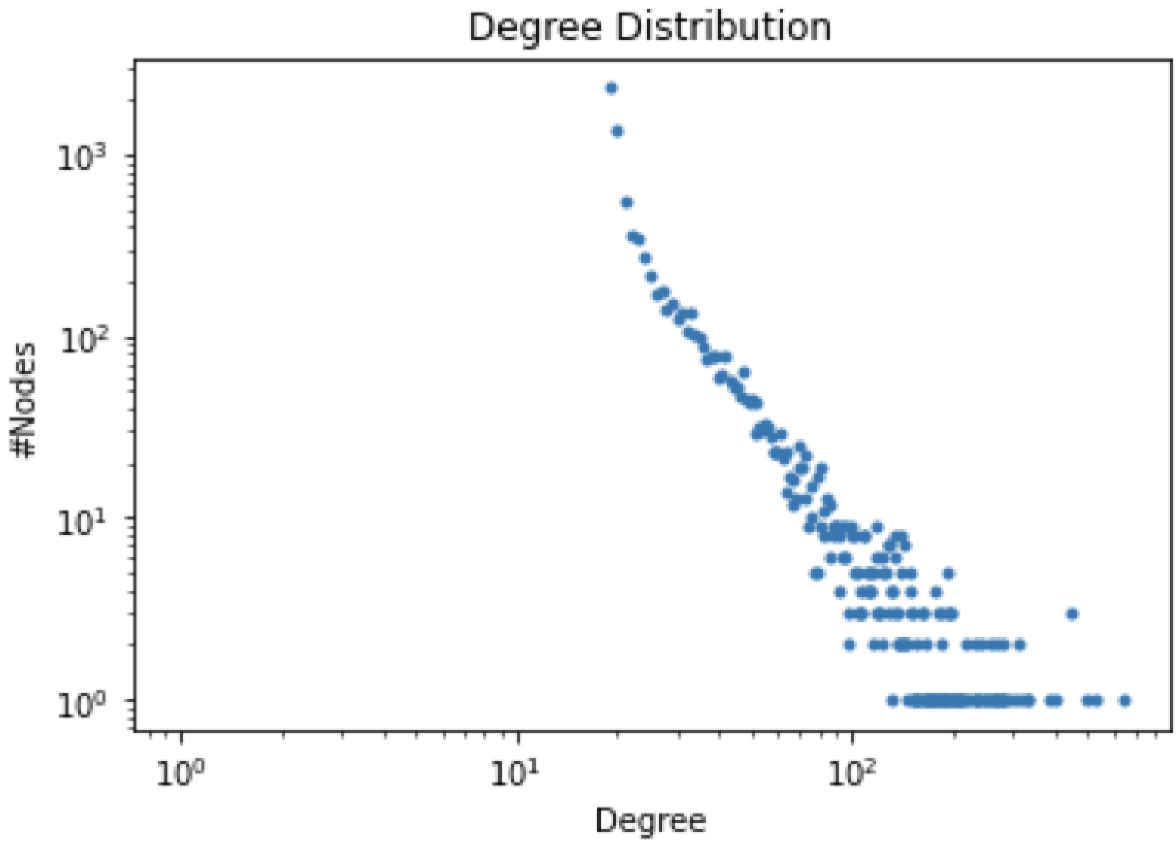
  The created network is a connected undirected graph with 8,792 total nodes and 1,47,424 total unique edges.

  In an undirected graph, the degree is the number of nodes connected to a particular node. It tells us in our work how many restaurants are similar to one Restaurant, e.g. Swad has the highest degree of 638.

  When we did further data analysis, Swad's menu and reviews contained diverse foods, from pizza, sweets, and drinks to foods from various parts of India, making it an alternative to many other restaurants. Alternatively, it gives its consumers an option to try at 638 different restaurants depending upon the dish they want to try.

  The network has an average degree of 33.536 and a weighted degree of 92.482.

Given below is the degree distribution of the graph. It can be pragmatic that the degree of distribution follows the power law, which verifies it to be a scale-free network that incorporates two important concepts: growth and preferential attachment. Many real-world networks such as Social Media, Internet, or World Wide Web are scale-free.



The network diameter is the shortest distance between the two most distant nodes in a network. Our network has a diameter of 8.

Betweenness Centrality is sensing the sum of influence a node has over the flow of data in a graph. It is used to find nodes that serve as conduits from one part of a graph to another. The algorithm calculates unweighted shortest paths between all pairs of nodes in a graph. It tells us which nodes/restaurants are 'bridges' between nodes in a network. Swad, Pallavi Restaurant, and Tamarind are some of the nodes with the highest betweenness centrality hence the key players influencing the whole system.

Closeness centrality is the average shortest distance from each vertex to another vertex. It scores each node based on its closeness to all other nodes in the network. Swad, Chef-In, Melange, and Food Point are the restaurants with the highest closeness centrality. In networks, closeness centrality is used to find the 'broadcasters'. In the context of our work, it can be inferred that these restaurants' menus or cuisine are the most appropriate to influence the entire network most quickly.

- **EigenCentrality**

  Measures a node's influence based on the number of links it has to other nodes in the network. But it is different from degree centrality.

  EigenCentrality goes a step further by considering how well connected a node is, how many links its contacts have, and so on throughout the network. In our network, the restaurants with high closeness centrality mostly correspond to high EigenCentrality.

- **Community Discovery**

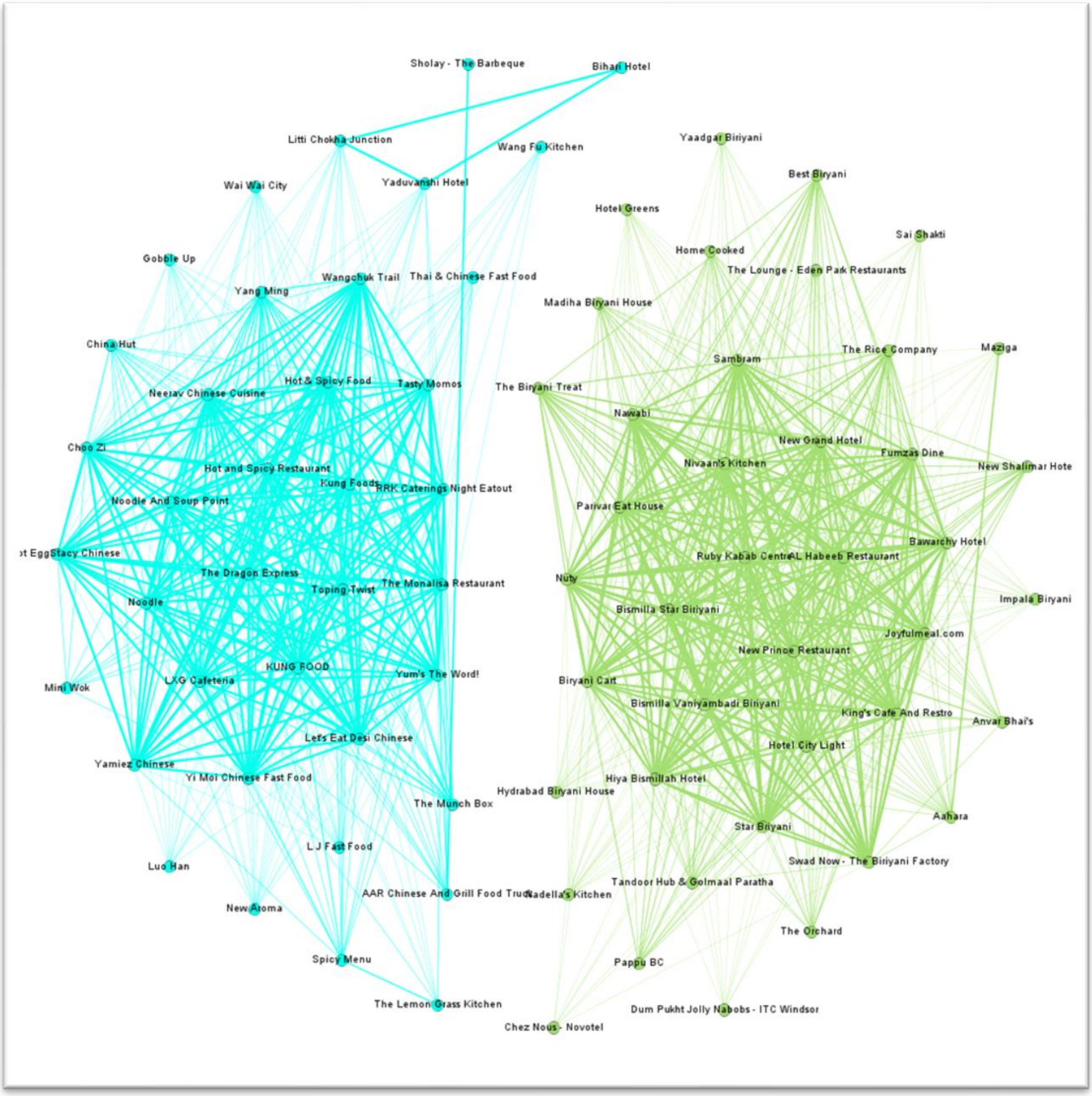
  Community Discovery automatically clusters similar kinds of restaurants based on the connections made.

- **Label Propagation**
  https://arxiv.org/abs/1103.4550

  The label propagation algorithm is fast for finding communities in a graph. It detects communities by propagating the labels throughout the network and forming communities based on this label propagation process.

  Label Propagation for our network resulted in a total of 71 discovered communities. The top five cultures with the highest number of nodes were [1,119, 590, 522, 478, 378], and the top five with the lowest number of nodes were [10, 14, 18, 18, 25].

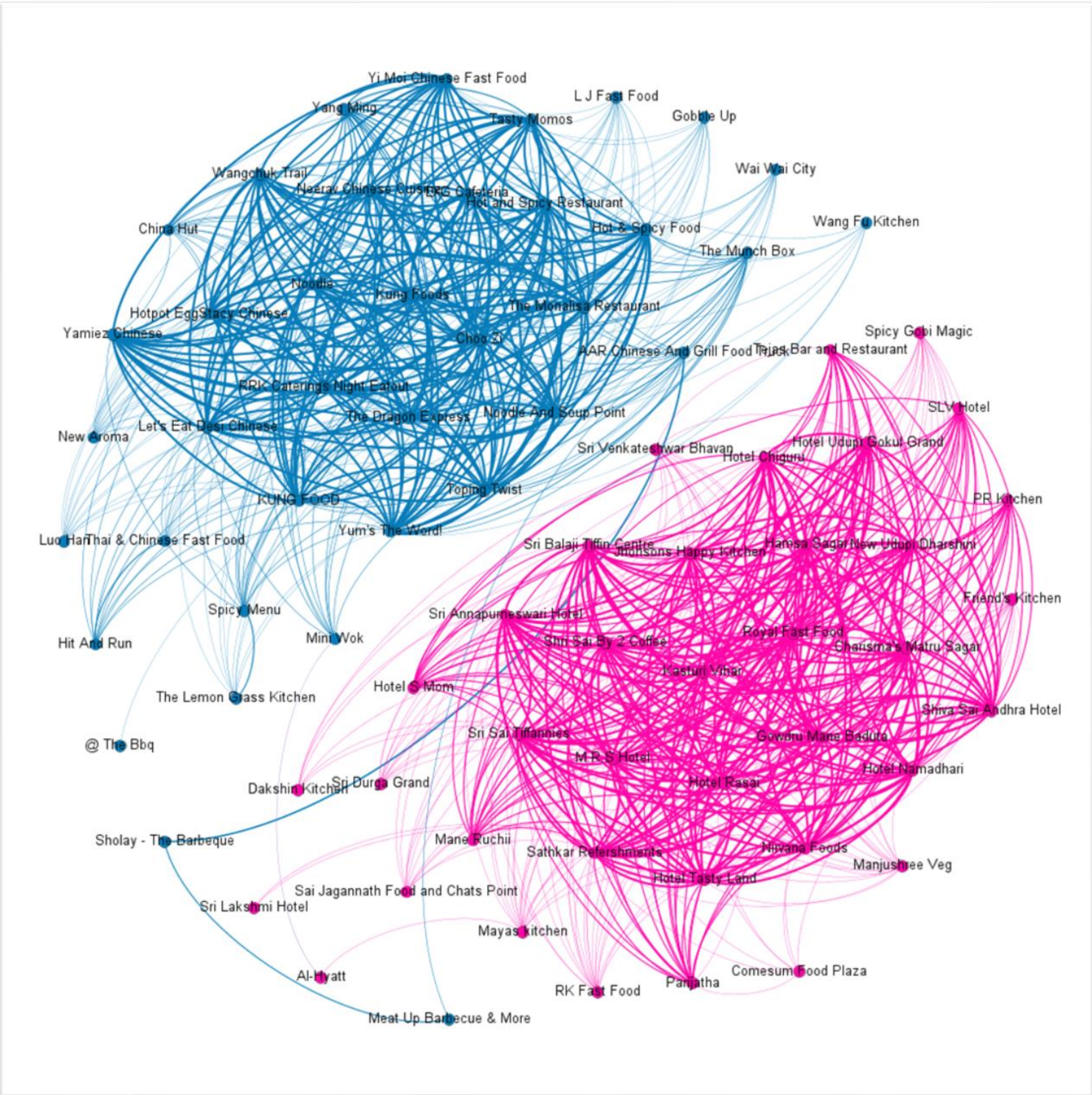
  The example below shows two communities where the blue contains primarily Chinese food Restaurants, and the green one includes mostly Mughlai-style restaurants.



- **Louvain Algorithm**
  https://arxiv.org/pdf/1410.1237.pdf

  In the Louvain method, first, small communities are found by optimizing modularity locally on all nodes, then, each small community is grouped into one node, and the first step is repeated. A total of 45 districts were found using Louvain Algorithm on our network with Constant Potts Model [Narrow scope for resolution-limit-free community detection (arxiv.org)].

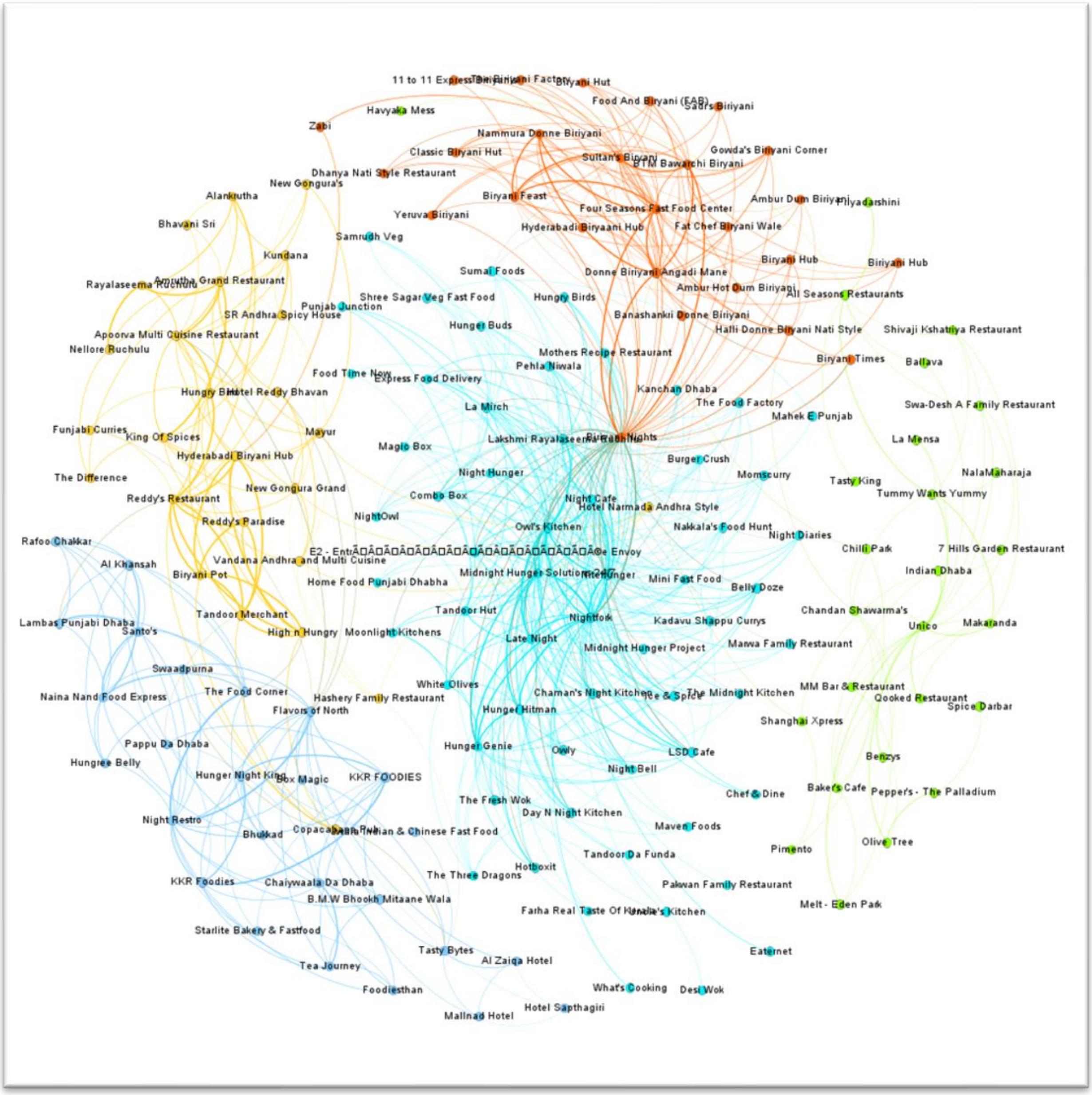
  The figure below shows two communities detected by using Louvain Algorithm. The community in blue has primarily Chinese restaurants, while in pink, we have public food messes.



- **Statistical Inference**
  [2006.14493] Statistical inference of assortative community structures (arxiv.org)]

  This method is used to infer assortative communities in networks based on a nonparametric Bayesian formulation of the planted partition mode. A total of 177 communities were discovered using this approach on our network.

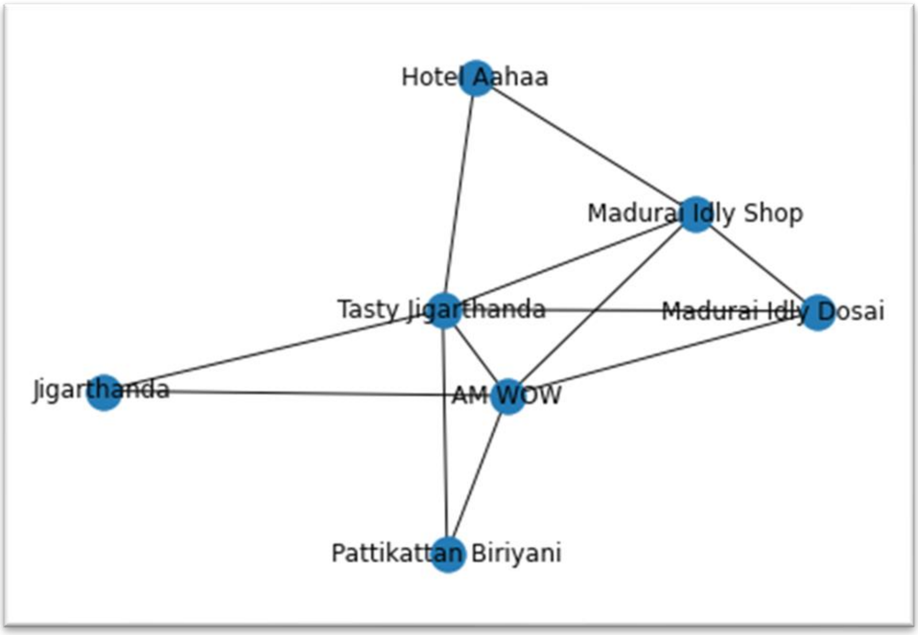
  The figure below shows some of the smaller communities detected in different colours.

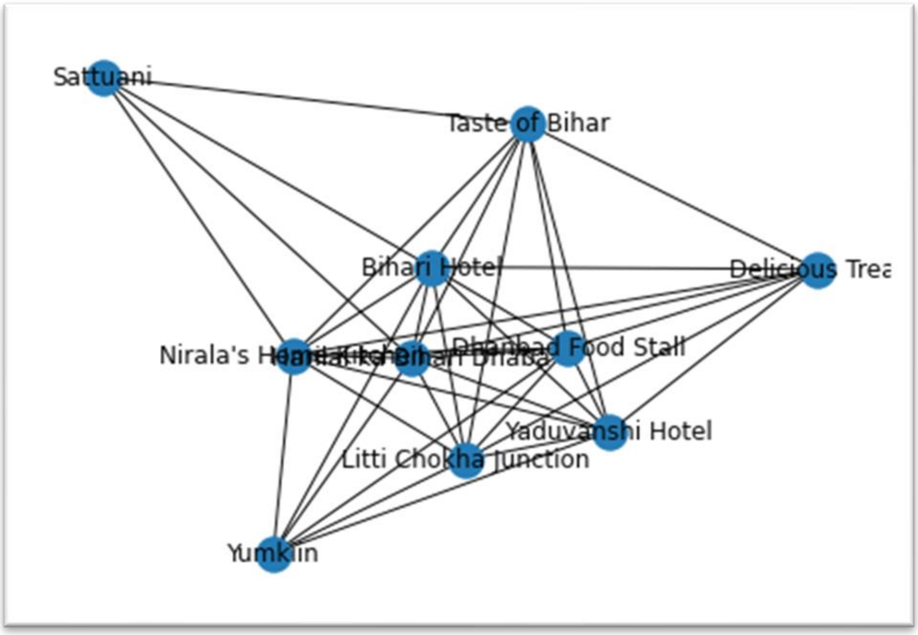


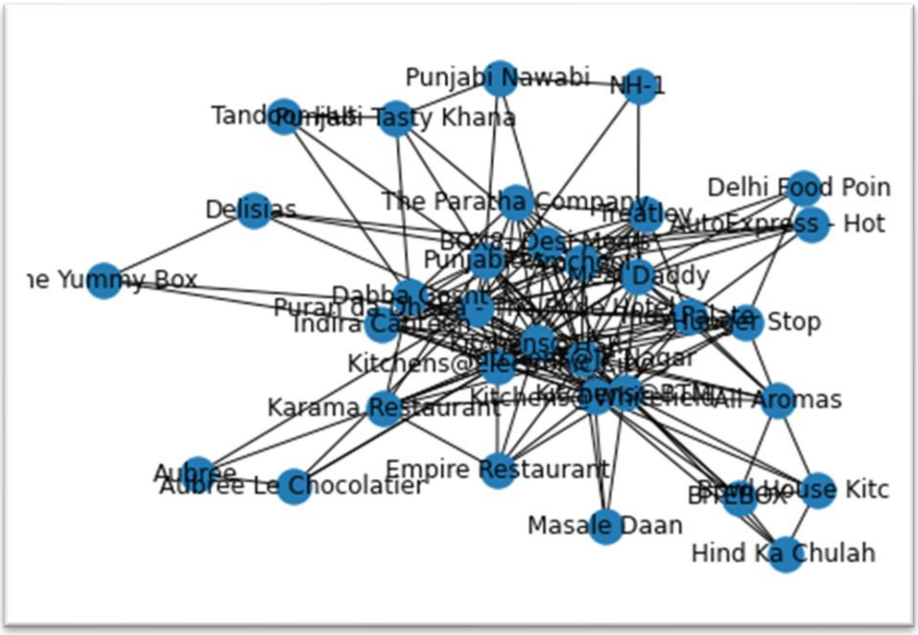
- **Demon [https://arxiv.org/abs/1206.0629]**

  Many large networks often lack a particular community organization at a global level. In these cases, traditional graph partitioning algorithms fail to let the latent knowledge embedded in modular structure emerge because they impose a top-down global view of a network. Demon is a simple local-first approach to community discovery, able to unveil the modular organization of real complex networks. This is attained by constitutionally letting each node vote for the communities it sees adjacent to it in its limited sight of the global system, i.e. its ego neighbourhood, using a label propagation algorithm; finally, the local communities are merged into a worldwide collection.

  We have tested the Demon method on our network, and it yielded the most intuitive results of all others. Below are some examples where it was able to classify restaurants from different parts of India without having any direct information about the location. The first figure contains some of the restaurants from Tamil Nadu. The second has grouped some Bihari restaurants, while the third figure shows communities formed with restaurants from Delhi, Punjab and Haryana regions.




Sattuani
Taste of Bihar
Bihari Hotel
Delicious Trea
Yaduvanshi Hotel
Litti Chokha Junction
Yumklin

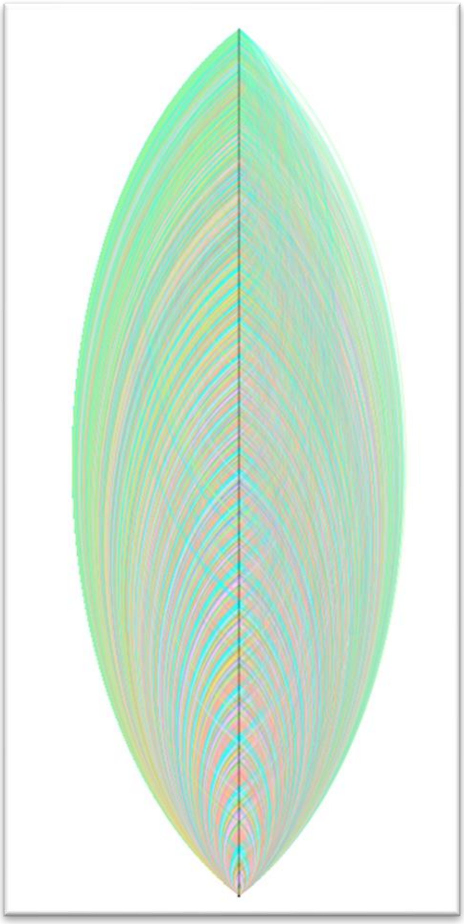

Punjabi Nawabi
NH-1
The Paratha Company
Delisias
Delhi Food Poin
AutoExpress - Hot
Indira Canteen
Karama Restaurant
Empire Restaurant
Aubree
Aubree Le Chocolatier
Masale Daan
Hind Ka Chulah

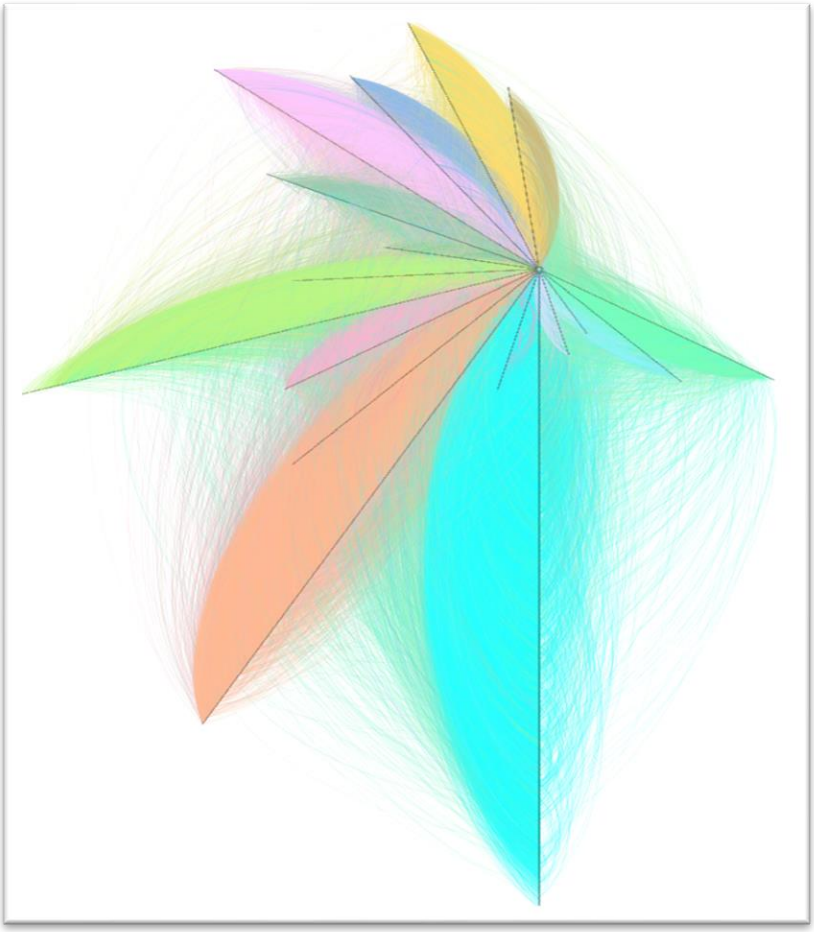
- **Visualization and Creation of Data Art**

Once we had created the network, we tried various approaches to generate the data art. The overall goal was to create something beautiful which would also make sense.

Art 1. Paper-Spinner:

We started aligning all the nodes on a single line, ordering by their modularity class/community id. Then we made connections as per the edges in randomized colours, which denote the different partitions in the graph. In the figure below, the vertical line contains the nodes ordered in decreasing order of community sizes from top to bottom. Within this hierarchy, the nodes are ordered in decreasing order of their degree. Eg. The bottommost node is the most significant degree node of the biggest community.



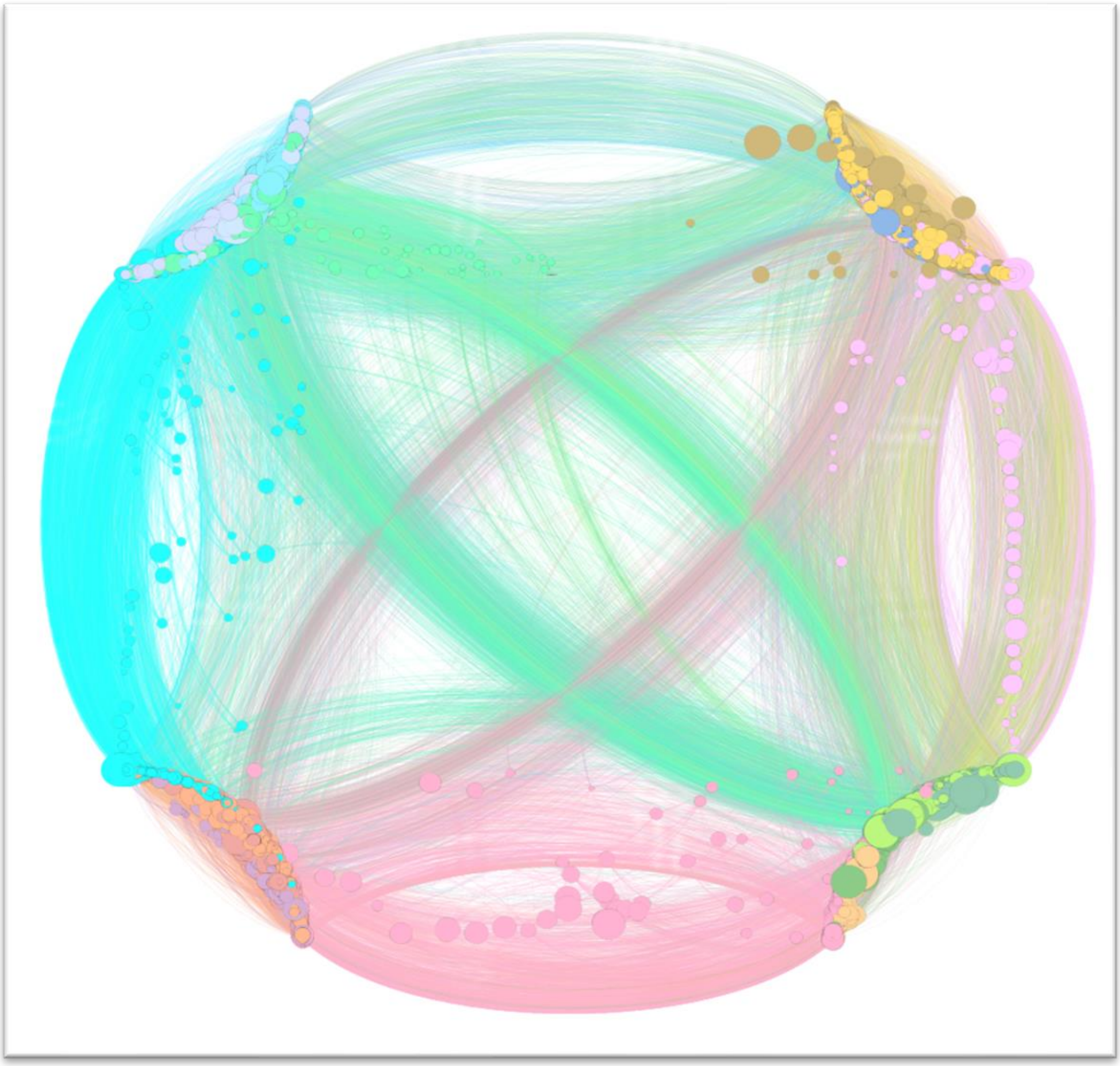


While the above approach creates beautiful graphics, zooming in does not give a very partitive way of identifying different types of restaurants and their connections. Therefore, we experimented with a number of axes/spars to find a visualization that would give us a sense of the dataset while also preserving the beauty of the generated art.

The image below is the result of using 17 axes. Here, from the second-largest line to the smallest line, the community with the most significant size up to the smallest 16th size are denoted. All the other smaller communities are then aligned on the most central line, and connections are made to the respective nodes. This approach generated a beautiful paper-spin-looking art which, when zoomed-in, can also be used to find similar restaurants following the connections made in another colour than the node itself. The distance here does not denote any similarity/meaning in the dataset.

Art 2: Circular Layouts

In the arts below, we have experimented with the circular layout. The circular design is an O(n) complexity algorithm which aligns the nodes in a circle, ordered by the parameter specified. We started by using a circular layout with an ordered hierarchy of Partition (modularity class), degree and



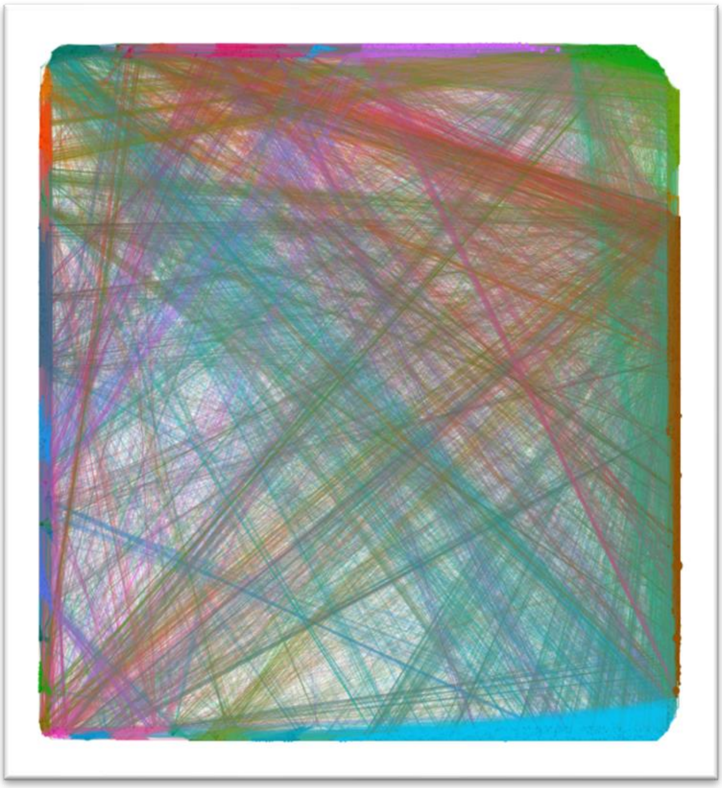
betweenness centrality. We experimented with other algorithms, such as Fructerman Reingold (Graph drawing by force-directed placement - Fruchterman - 1991 - Software: Practice and Experience - Wiley Online Library)

Force Atlas 2

[https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0098679] and Radial Axis Layout over the circular layout image. A setting of Area, Speed and Gravity was varied continuously to create the data art by applying Fructerman Reingold over Circular Layout for various minutes. Below is an example of such art. Here the degree information has been added to the node's size, i.e., the node with the highest overall degree is the largest. This type of art is essential to quickly identify the high-size-nodes-density regions as they might contain the group of big restaurants in the city with diverse but similar tastes. The distance here denotes the difference in taste/variety of the food they offer. E.g. Two distant clusters of big-size nodes indicate the group of restaurants in the city with diverse and different tastes. The other kinds of arts generated by mixing the algorithms and varying the time are shown in the appendix section.

Art 3: Rectangular Layouts:

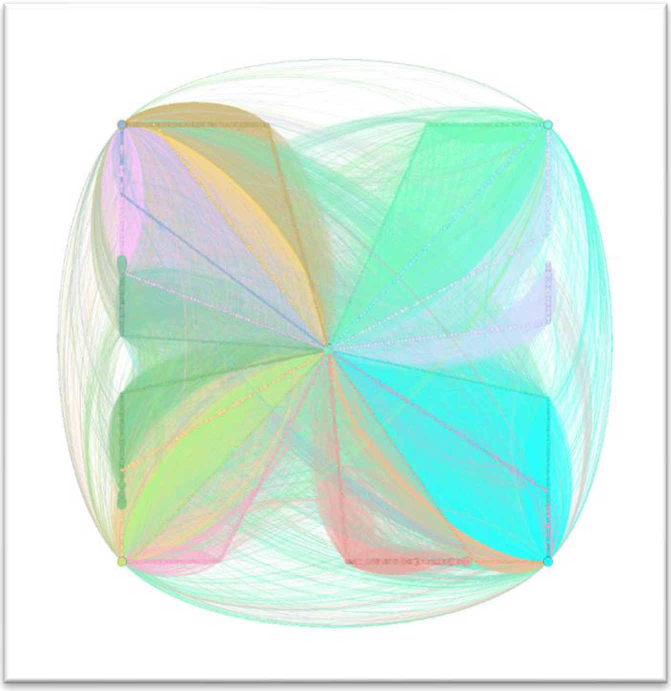
The art below is generated by applying a rectangular layout algorithm, grouping nodes by their modularity class and ordering by their degree. Finally, connections are made with the other nodes in a straight line. This kind of art gives a sense of the size of the community and hence could help the viewer look for a restaurant from a different size group. Since the smaller communities indirectly denote the rarity in terms of taste and cuisine, the nodes from respective communities could be searched from a chain of big restaurants which usually belong in the larger communities as they offer a wide and more common variety of food. Usually, the restaurants which offer less variety of food specialize in specific cuisines.

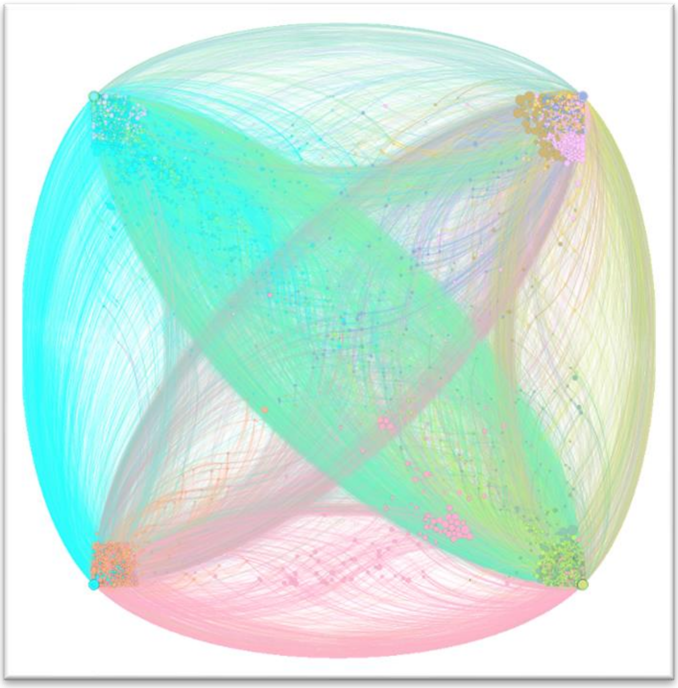


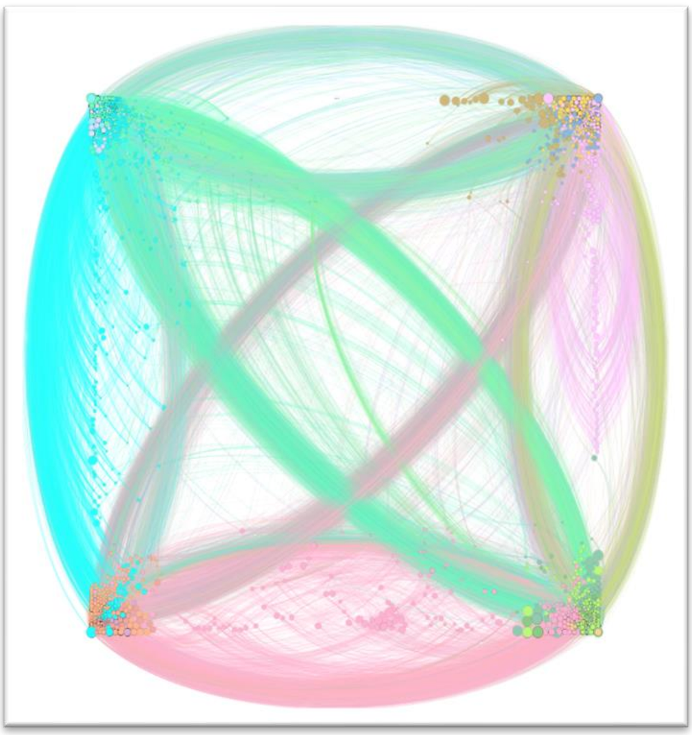
Therefore, this art could open a possibility for a user to discover restaurants that specialize in specific rare cuisines and for a restaurant to attract visitors without being famous.

- **Appendix:**

  The two arts below were generated by applying Force Atlas 2 over a rectangular layout arrangement. The setting, such as speed, approximation, approximate repulsion, and overlap, varied with variable time and the connections were made in a curved fashion to obtain art that gives a flower-like design. The size of the nodes here denotes the degree within the same community.



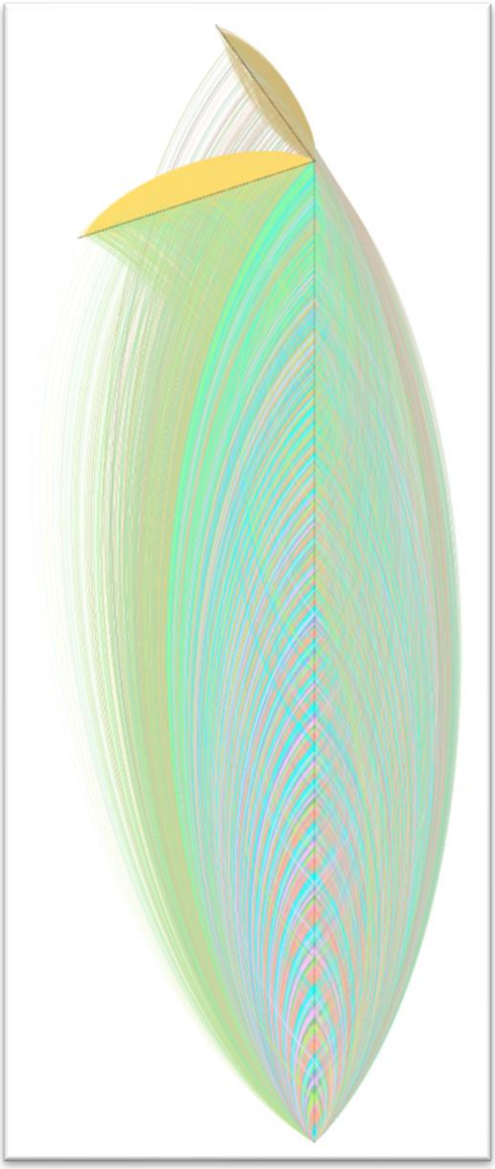




Below are some more arts generated using Radial Axis Layout by experimenting with the number of axes/spars and varying hierarchy of nodes.

Figure 1:

Radial axis layout with just two spars and curved line connections



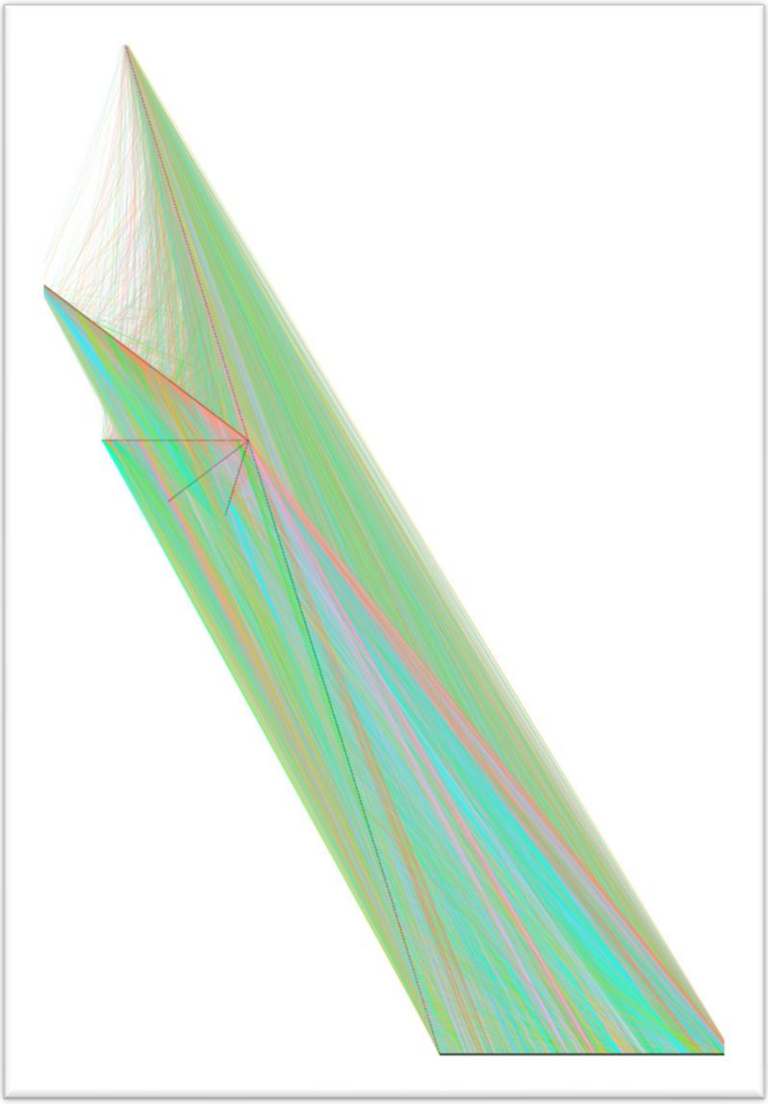
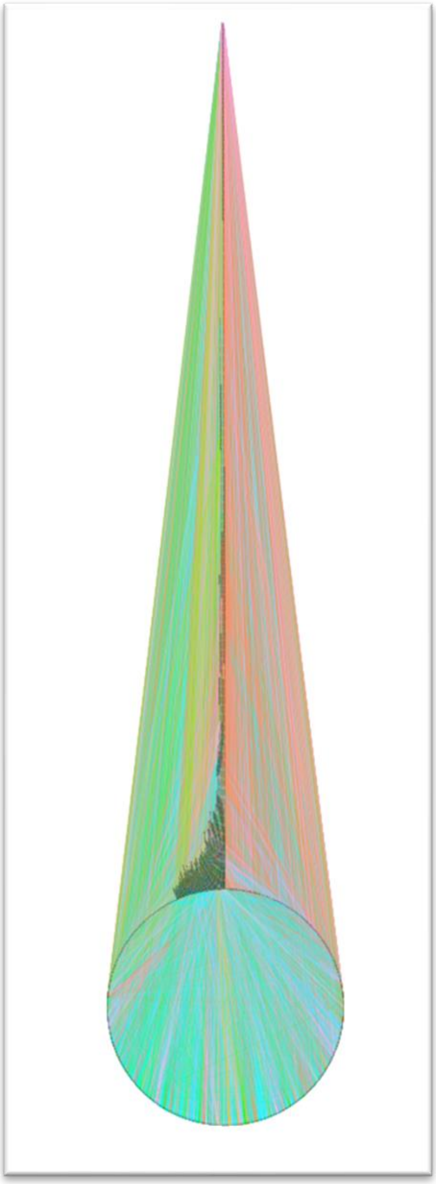


Figure 2

Radial axis layout with six spars and with straight connections

Figure 3:

I am applying a circular layout to the 10 largest communities and a radial axis layout to the other districts. Here the vertical line denotes the bottom (N-10) communities, and the circle contains the top 10 sections. The connections made here are straight lines.



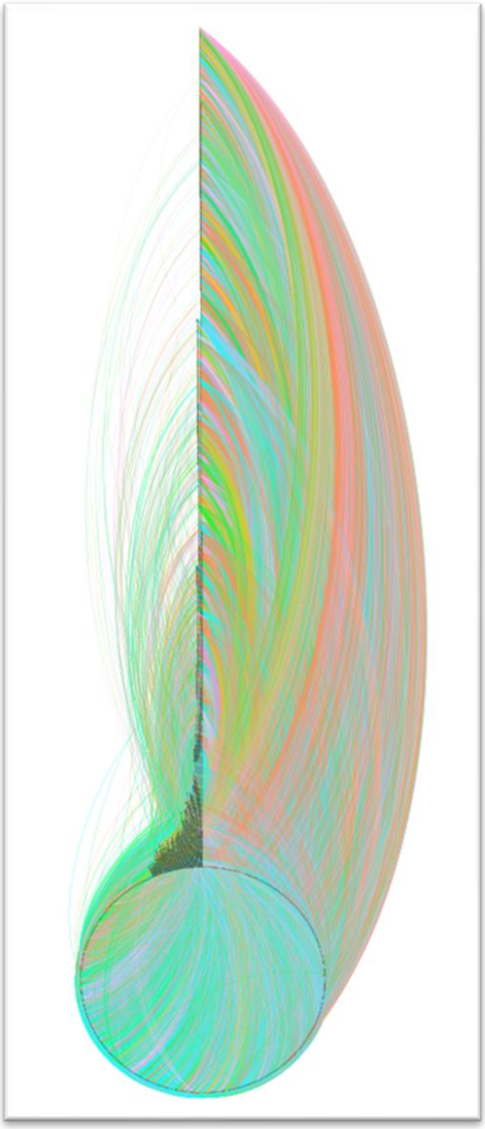


Figure 4:

This art is the same as Figure 3, and only the connections are made randomly but curvedly

The art below (Figure 5) is generated by applying Force Atlas 2 over the nodes. The different colours denote the modularity class. The distance between the two clusters represents the variance in taste/cuisines.

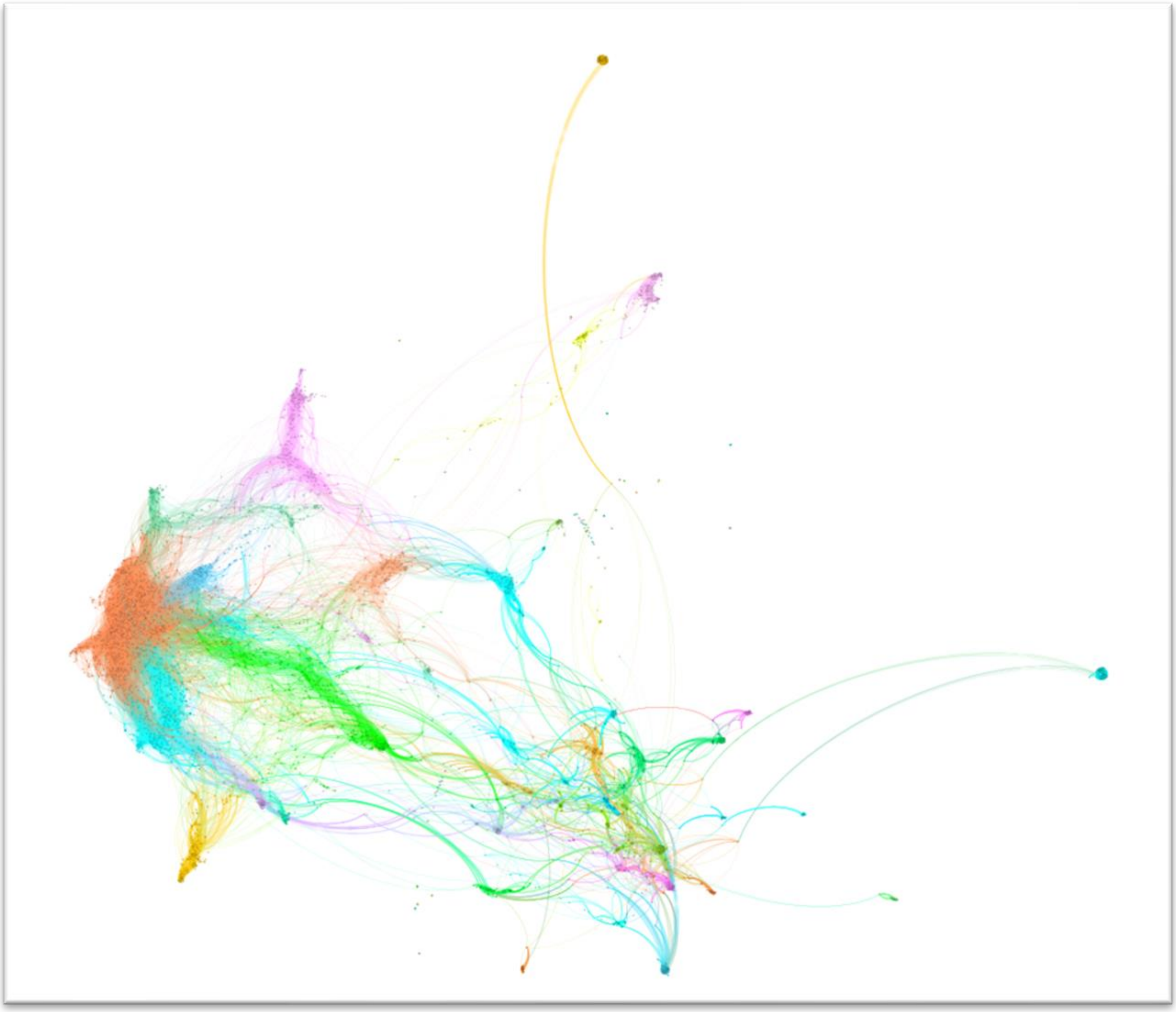


Figure 5: Art made by applying Force Atlas 2 over the communities

- **Summary and Supposition**

  The paper has shown gastronomic guides' highly influential roles in taste-making and restaurant selection criteria. Among the various directions, It constructs a hierarchy of chefs and restaurants and, within the scale, draws firm symbolic boundaries between multiple categories of value, i.e., between one, two and three stars.

- **Tools Used:**

Programming Language	:	Python3.8
Libraries	:	Pandas
Main Libraries Used	:	Pandas(1.4.2), Numpy(1.23.2), scikit-learn (1.1.1), network(2.8.6), NDLib(5.1.0), NLTK (3.6.6), PyTorch(1.12.0), Matplotlib(3.5.2)
Applications/Machines	:	Gephi application version 0.9.7 on Microsoft Windows 11, Google Colab Pro Instance with NVIDIA Tesla V100 GPU and 25GB RAM