Our Project is about Food Recipe graph

**Dataset**

https://www.kaggle.com/datasets/saldenisov/recipenlg

The dataset have each recipe ,and it’s ingredients

A screenshot of a computer

Description automatically generated

We focus on title ,NER column (individual ingredient)

**Then we construct Graph**

**Node** : individual ingredient

**Edge :** recipe using the 2 ingredient (undirected ,weighted)

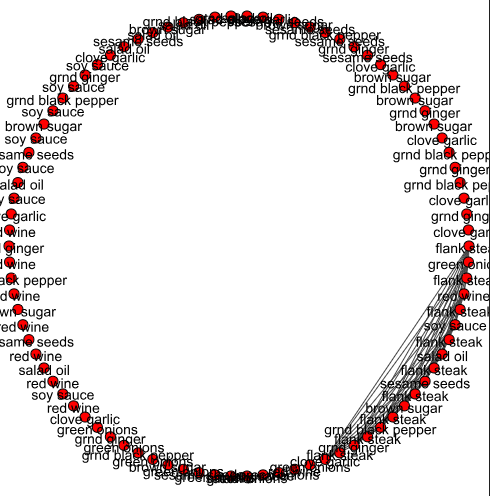
**Sampling**:

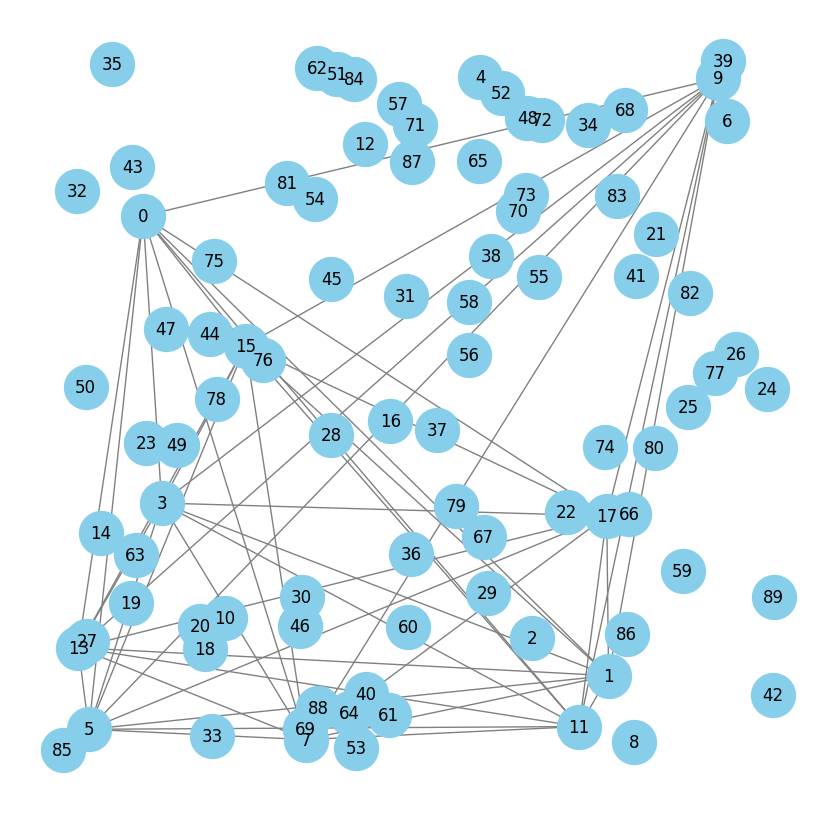
The dataset is very large (large recipe ,large NER) ,so we worked on sample

-working on whole dataset using big data techniques is future work

**Graph**:

We tried to plot graph on many way to understand it:



A number connected to a network

Description automatically generated with medium confidence

We **used Igraph , networkx**

**Then Degree ,avg degree**

A computer screen shot of a code

Description automatically generated

A computer screen shot of a code

Description automatically generated

meanly each individual NER used in one reciepe

Then

Edge list ,adj list

A screenshot of a computer

Description automatically generated

Adj matrix:

A screenshot of a computer

Description automatically generated

**3- get if special graph**

A screen shot of a computer program

Description automatically generated

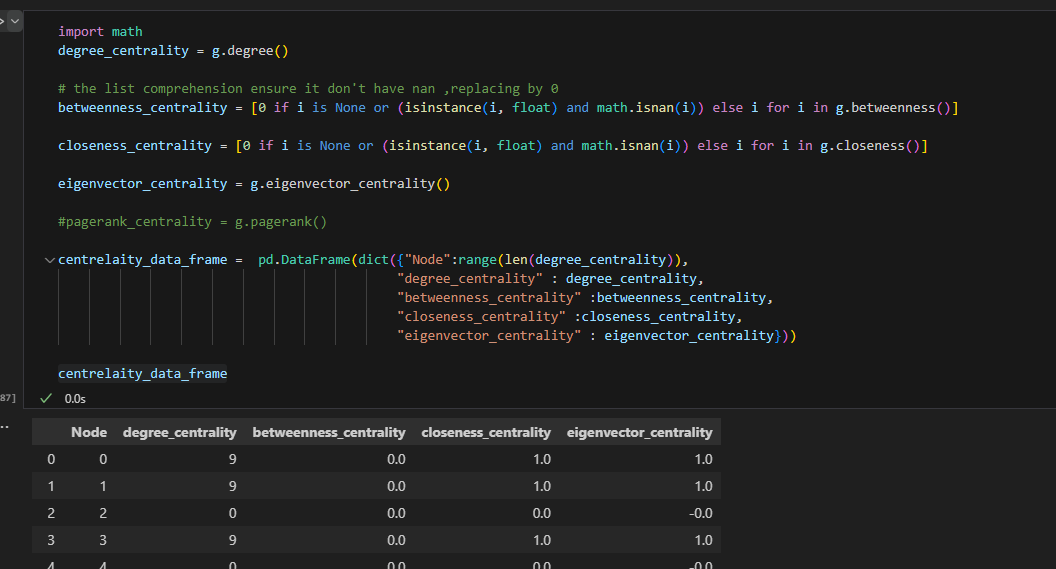
This result no ,we tried to check if near : still no ,but with visualize :

A colorful circle with dots and lines connected to each other

Description automatically generated

-part complete ,other isolated: This mean that some ingredients are complete (mostly used themselve in any reciepe) ,other like isolated ,may on one or two reciepes

**4-Centrality:**



A graph of a number of different types of squares

Description automatically generated with medium confidence

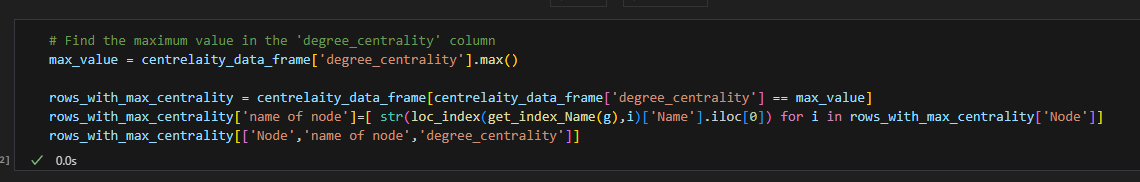
**For our project** : The most important relating of centrality of our dataset is Degree centrality (what is main important ingredients i should have in restaurant as appera in many reciepes can be many so no need eignvalue)  :

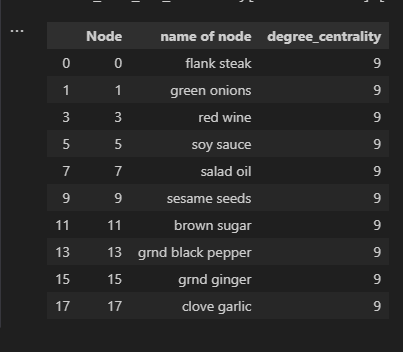
Plotting it: (with size represent deg)

A diagram of a network

Description automatically generated

Then getting most degree:





And those must be found in restaurant