Assignment 2: Centrality Measures

Summer 2021
Authors: Team GOAT (Estaban Aramayo, Ethan Haley, Claire Meyer, and Tyler Frankenburg)

The Cocktail DB (https://www.thecocktaildb.com/api.php) is a database of cocktails and ingredients available on the web. In this assignment, we describe how we could use the Cocktail DB's API to generate a network of cocktails and ingredients. We can use some example data to explore how we might be able to predict outcomes from this data using centrality metrics.

Data Source:
The Cocktail DB includes several hundred cocktails, categorized into drink types like 'cocktail', 'shot', or 'non-alcoholic'. Each cocktail also has an ingredient list available.

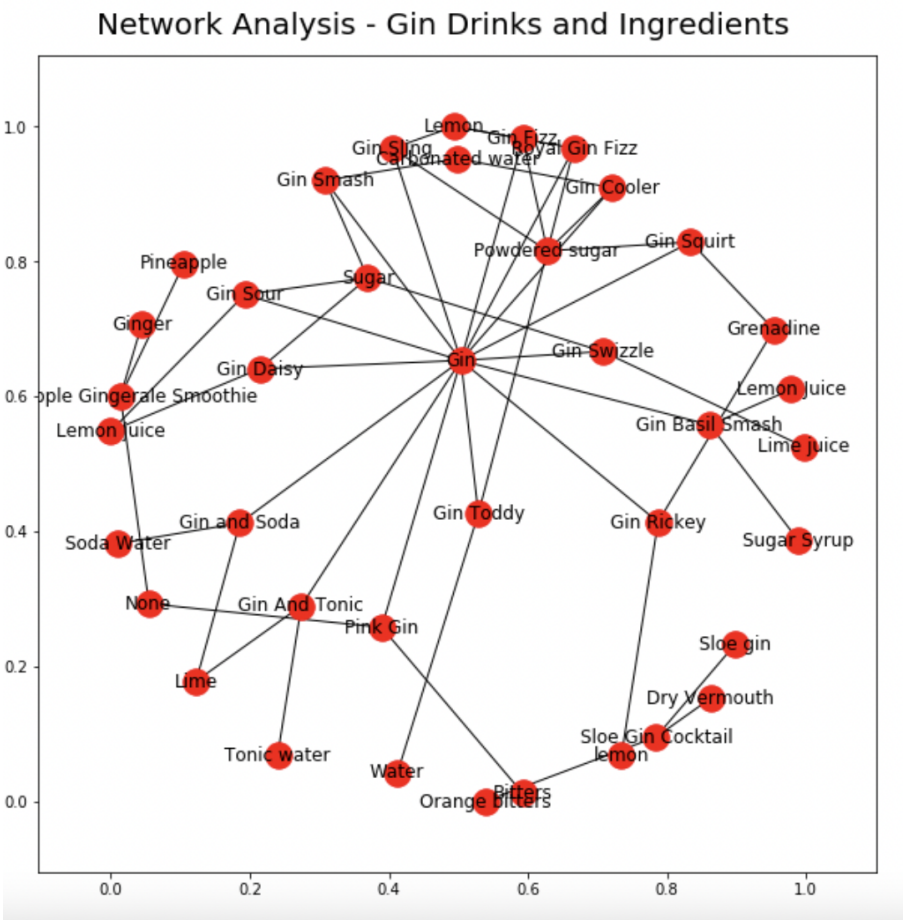
Without digging too deeply into the intricacies of the Cocktail DB API, we can leverage 2 libraries: `requests` to make an API request, and `json` to load the JSON output from the API. We can then iterate through each cocktail output to grab the relevant components. The best way to generate output from this API is through cocktail name search, so we plan to construct a set of functions that search by the first letter of a cocktail, iterating through the alphabet, and populating a dataframe with the cocktail name, category, and ingredients list.

Example output:

name	id	category	photoURL	ingredients
A1	17222	Cocktail	https://www.thecocktaildb.com/images/media/dri...	[Gin, Grand Marnier, Lemon Juice, Grenadine]
ABC	13501	Shot	https://www.thecocktaildb.com/images/media/dri...	[Amaretto, Baileys irish cream, Cognac]
Ace	17225	Cocktail	https://www.thecocktaildb.com/images/media/dri...	[Gin, Grenadine, Heavy cream, Milk, Egg White]
Adam	17837	Ordinary Drink	https://www.thecocktaildb.com/images/media/dri...	[Dark rum, Lemon juice, Grenadine]
AT&T	13938	Ordinary Drink	https://www.thecocktaildb.com/images/media/dri...	[Absolut Vodka, Gin, Tonic water]

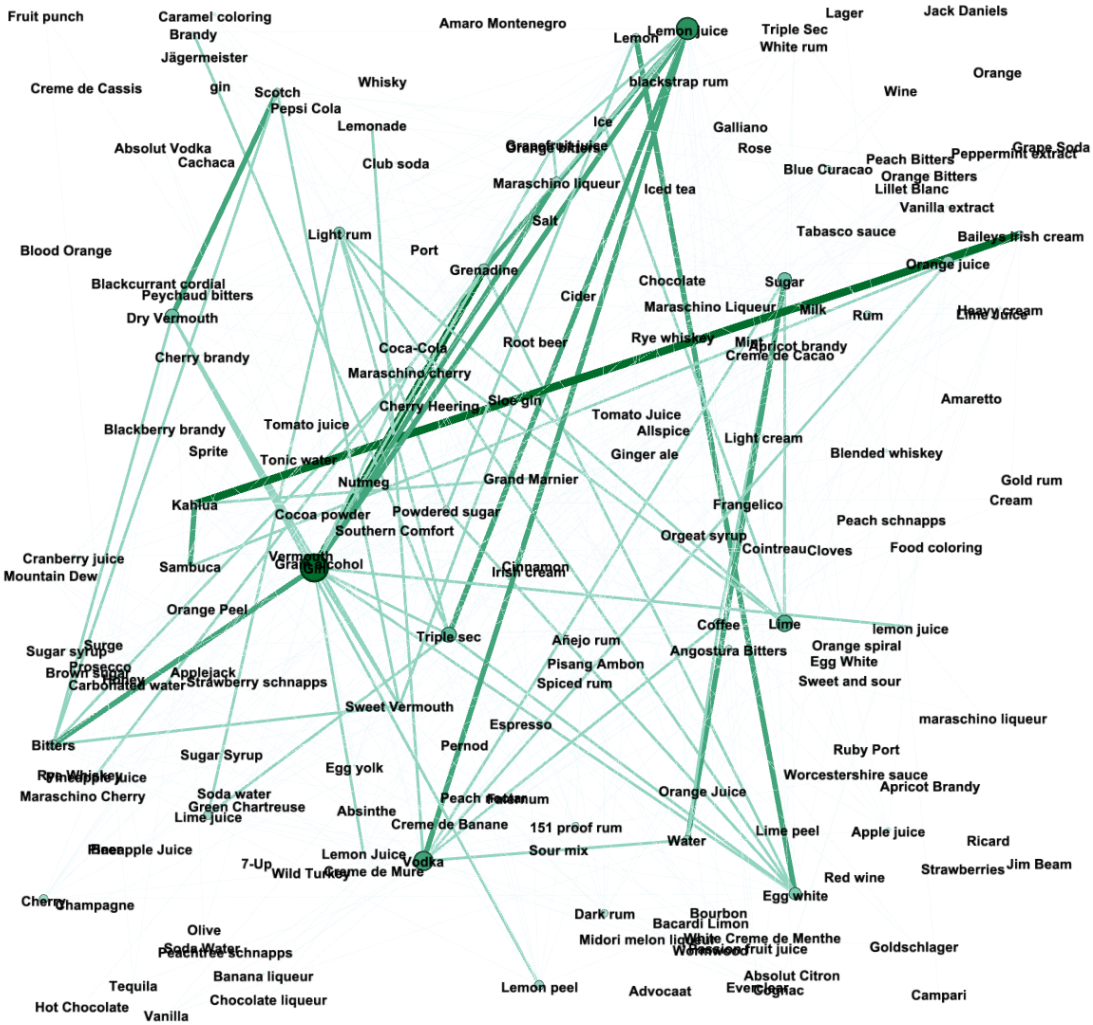
From there, we can either create a graph with multiple node types (ingredients and cocktails).

Example visualization of 'Gin' drinks:



We could also go further and create a bipartite graph of cocktails and their ingredients. From there, we can understand how, for example, ingredients co-occur in cocktails, or which cocktail ingredients to keep on hand to maximize the ability to improvise.

Example visualization of ingredients:



Hypothesis:

We would like to explore these relationships, and compare degree centrality against these cocktail categories. We would expect ‘Shot’ category cocktails to have lower degree centrality, with shorter and more streamlined ingredients lists.