Assignment 3

Rosetta Wang

Omeed Tavasoli

Kamal Bains

**Dataset:** fallingfruit.org

**Research Interests:** Our purpose is to explore the ecological diversity of urban areas. Some interesting metrics we will be able to obtain from our work will be what the average diversity is across the entire network as well as understand how important individual plants are to the ecological diversity of the urban area.

**What does each node represent?**

Each node is a single plant at a single location

**Number of nodes:** approximately 30k

There are approximately 30,000 marked plants of the database collected at fallingfruit.org in the Bay area.

**What does each link represent?**

First we must consider **what the weight of each link represents**. Since we are interested in measuring diversity, the weight will be higher with closer nodes and higher with nodes of plants that are more taxonomically different. In other words, the weight is negatively correlated with physical distance and positively correlated with taxonomic difference.

**Number of links:**

We want to avoid having a link between every node because that is not interesting to visualize.

Nodes that are very far apart will not be linked and nodes that have the same species and subspecies will not be linked.

The weight between each node will be calculated and if it does not reach a certain threshold, it will not be linked. Let’s call this the **linking threshold**. Because of the way the weight will be calculated, the two plants that are closer in location and more different in taxonomy will have links and vice versa.

Since developing the number of links is the backbone of our project, we have not yet done this analysis.

**Centrality and Average Degree:** a higher centrality will represent how much a node contributes to the diversity in the area. For example, imagine if we have a garden of orange trees. If we were to add an apricot tree, we would see a graph form where the apricot tree is connected to each orange tree but each orange tree is only connected to the single apricot tree. This graph will have a low average degree (representing the low ecological diversity), but the apricot tree will have a high centrality, underscoring its importance in the ecological diversity of the area. With these metrics, we hope to capture the ecological diversity of the bay area as well as explore what nodes (ie which plants) provide the most diversity.

**Part II:**

We found the paper “Demographic and landscape-level urban foraging trends in the USA derived from web and mobile app usage” which was published in the Journal of Urban Ecology. The paper actually uses the Falling Fruit dataset to examine how ecological diversity and foraging availability differs across demographics (race, income, zip code).

<https://academic.oup.com/jue/article/3/1/jux006/4056551?fbclid=IwAR3LEj82oy1to0rlgNKBS32dzKdH_EMIi2oVBn-T3GiSwsAnu-1OL2pZVVs>

I think the methods we are creating to measure diversity are interesting. It seems to me that in these times, diversity is being increasingly realized to be important. Whether that is genetic diversity, racial, or otherwise. This networked method of measuring diversity can be applied not only to plants but to other areas such as in a human population.

https://www.w3schools.com/python/ref\_string\_split.asp