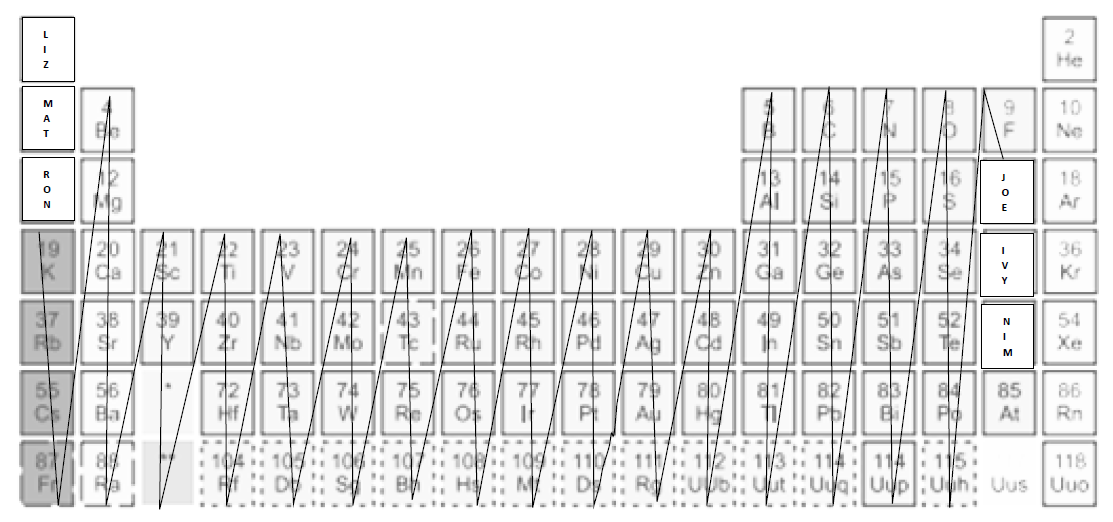
# Assignment 1

**Problem 1: List all the benefits and limitations of this visualization that contains the student names of a Chemistry class sorted high to low by their Assignment 1 scores:**



***Benefits.*** One of the benefits The density of the network was determined using the Statistics tab of Gephi, Under Network Overview -> Graph Density. A screenshot is provided to the right, along with some other statistical tests run on the graph. The desired metric for the purposes of answering this problem is Graph Density.

***Limitations.*** The density of the network was determined to be 0.112 – which means that the network is sparse. Out of all the possible connections that could be made in the network, only 11.2% of the nodes are connected.

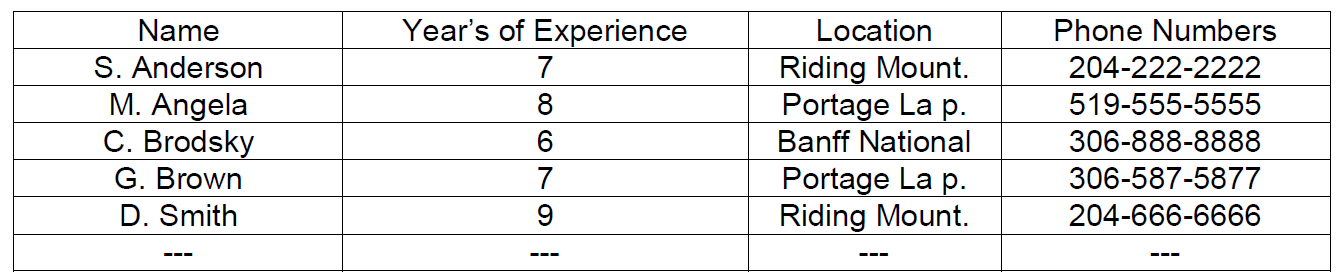
***Problem 1 Bonus: Propose a beautiful visualization for this scenario.***

***Method.*** The density of the network was determined using the Statistics tab of Gephi, Under Network Overview -> Graph Density. A screenshot is provided to the right, along with some other statistical tests run on the graph. The desired metric for the purposes of answering this problem is Graph Density.

*Attempt #1: Unique visualization based on communities.*

*Attempt #2: Visualization close to the original.*

**Problem 2: How effective is the following table in achieving the goal of displaying a list of 1000 experienced members with their year of experience? Can you propose a better way to reorder the table?**



***Effectivity.*** For easy identification, the network was configured such that the size of the node in the visualization represents the degree of the node itself. Moreover, a gradient colouring scheme was applied to make the most prominent node stand out further. Finally, the layout used for visualization was Force Atlas. A screenshot of the corresponding network with the highest-degree node isolated is shown below.

***Proposition for Improvement.*** Based on the dataset visualization which was also double-checked against the database using the Data Laboratory tab, the node with the highest degree is Lentil Soup with a degree of 114. This implies that users who left a comment on the Lentil Soup recipe are more likely to leave a comment on other recipes as well.

***Problem 2 Bonus: Propose a beautiful visualization for this scenario.***

***Method.*** The density of the network was determined using the Statistics tab of Gephi, Under Network Overview -> Graph Density. A screenshot is provided to the right, along with some other statistical tests run on the graph. The desired metric for the purposes of answering this problem is Graph Density.

*Attempt #1: Multi-dimensional visualization based on location and years of experience.*

*Attempt #2: Improvement of visualization by geographical location.*

**Problem 3: Given a set of tuples (team, score), draw a blue square for each team such that the side length of the *i*th square corresponds to the score of the *i*th team.**

***Method.*** The network was filtered by size to identify the node with the highest degree of betweenness. Finally, a colour gradient was applied to further emphasize this result, and the Force Atlas layout was utilized. The corresponding data visualization is shown below.

***Challenges.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.

***Status.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.

**Problem 4:** Enhance the visualization that you created for Q3 by drawing a black square on top of each team’s square. The side length of the black square equals the average score of all the teams.

***Method.*** The network was filtered by size to identify the node with the highest degree of betweenness. Finally, a colour gradient was applied to further emphasize this result, and the Force Atlas layout was utilized. The corresponding data visualization is shown below.

***Challenges.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.

***Status.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.

**Problem 5:** Create a scatterplot of the countries from the given dataset. Show the label of a country only if the country has a non-zero radius. Only the first 5 letters of the country name should be displayed.

***Method.*** The network was filtered by size to identify the node with the highest degree of betweenness. Finally, a colour gradient was applied to further emphasize this result, and the Force Atlas layout was utilized. The corresponding data visualization is shown below.

***Challenges.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.

***Status.*** While the degree signifies the “most famous node,” The implication follows that upon inspection of the list of recipes a user has commented on, it is most likely that the Red Lentil Curry recipe would be on that list.