Centrality measures in Primary School network data.

Graph Dataset is taken from the website: <http://www.sociopatterns.org/datasets/primary-school-temporal-network-data/>

* In the 2 and 3 column, those ids are the nodes of the people and each line represents a undirected edge.
* The edge weights are the time duration that the two individuals call it I and j were in contact in multiple of 20 seconds.
* The 4th and 5th columns are node attributes of the classnames of node I and j respectively.

Steps:

1. Extract the tar.gz file from the website mentioned above and import libraries in python such as

networkx, and matplotlib.pyplot

1. Load the csv file using either pandas read\_csv or another way like csv library using csv.reader
2. Create the graph using nx.Graph() functions and using add\_nodes\_from
3. Retrive the Density and Diameter of the graph to get some statistics.
4. For each group of classnames (1-5 depending on what class they start with), sort the nodes by degree count and extract say the first 10 nodes
   1. This can then tell us which group or school class had the most interactions with students and teachers.
5. For these top nodes, we can compute the time it takes for the particular student to be in contact with the other students and teachers
6. For each of these top nodes, use networkx’s closeness\_centrality to find the shortest paths to all other nodes, we can then see which classname group has more centrality.

Doing these computations for computing highest degree and shortest path can show which groups/classes socialized longer and see which ones could contract diseases.