Homework-6

April 26, 2017

1 Analysing the Graph of Stack Overflow

To be completed INDIVIDUALLY and due on May 3 at 3pm.

```
In [ ]: print('Your fitst name + last name')
```

** (70 pts) **

In this assignment, we will look at posts on Stack Overflow during the year of 2015. We have downloaded the data dump for Stack Overflow's posts in 2015. In this assignment, we will analyse the graph of interaction among users. The nodes in this graph will be the users. If $user_i$ has answered a question of $user_j$ (or vice versa), there will be an edge between $node_i$ and $node_j$ in the graph.

```
In [ ]:
```

1.1 Step 1

Analyze the resulting graph in terms of the number of connected components and degree distribution.

```
In [ ]:
```

1.2 Step 2

Working with the largest connected component and using three of the methods we described in the class, find the centrality scores of the nodes in this component. Then use the L_1 distance to perform a comparison of the three methods. Finally, visualize your results.

```
In [ ]:
```

1.3 Step 3

Again working with the largest connected component, use two different methods for identifying communities (partitions of the graph). Visualize your results and compare the output of the two methods using standard clustering comparison methods we discussed in the class.

```
In [ ]:
```