**Homework 8**

**COMP 272 Data Structures II**

In this homework, you will study a social media dataset which is described below.

A small dataset about Facebook pages (November 2017) is the subject of our study. These datasets represent blue verified Facebook page networks of different categories. We are interested in just one category here, the artists. .Nodes represent the pages and edges are mutual likes among them. The csv files contain the edges -- nodes are indexed from 0. For the artist dataset the number of nodes and edges are 50,515 and 819,306 respectively. Each edge is listed as integer pairs x y (this means that the connection is mutual x to y and y to x and hence undirected). You can get more information from https://snap.stanford.edu/data/gemsec-Facebook.html . The CSV file is attached with the homework.

Exercise

Read the data from the file to form an undirected graph using an efficient representation and answer the following questions by developing suitable methods.

1. How many connected components are there?
2. What is the size of the largest connected component?
3. Is there one or more connected components that is a tree? How many are there, if any?

You will use three different algorithms for question 1 above.

1. Using Depth First traversal (see text and other online sources).
2. Reading the edge list on the fly and determining which connected components they go to by using sets and union operation. We will discuss this algorithm in class.
3. Using Breadth first traversal ( see online sources).