ID2230 HW Three: Social Distance Calculator
Deadline: Nov 24, 2024

In this assignment, the goal is to take an input graph and the labels of two vertices and find the distance between them.

For the purpose of this assignment, we will use only the attached files soc-karate.csv and fb-pages-sport.csv as the input files for testing. The first file is small and may be useful to generate your own test cases; a visual representation of this graph is also provided. For the second file, a mapping of IDs to actual sportspersons/organizations is given in the fb-pages-sportNodes.csv file, although it isn't necessary for the program.

Input Format: The input consists of a file in csv format (comma separated values) with each line being in the form i,j which indicates that there's an edge between vertices i and j.

The lists are assumed to be sorted with respect to the first column so that they can be treated similar to an Adjacency list representation with minor modifications in implementation. Note that each edge is present twice: an edge between vertices i and j is present both as i,j and also as j,i.

Output: Your program should accept the labels of two vertices and output a shortest path between them, as well as the distance.

Submit a single file (one program/project file or one zip file with multiple program files) and name your submission file with your roll number.

Example from soc-karate.csv: For input node labels 15, 16, a shortest path is 15,32,2,0,5,16 of length 5.

Examples from fb-sports-pages.csv: For input node labels 12633, 12071, a shortest path is 12633,1331,12071 of length 2. For input node labels 2025, 13500, a shortest path is 2025,13571,8415,9210,13500 of length 5.

Sources: The soc-karate (Zachary's karate club network) is attributed to W. W. Zachary, "An information flow model for conflict and fission in small groups, Journal of Anthropological Research 33, 452-473 (1977)." and the fb-sports-pages file was obtained from https://networkrepository.com/networks.php (The Network Data Repository with Interactive Graph Analytics and Visualization by Ryan A. Rossi and Nesreen K. Ahmed), which has a collection of real-world graphs in multiple domains.