Notes:

Sent row for each person should be in theory corresponding to receive column for that person, fe person 2 sents messages to people - People in the receive sheet should note in column for that person 2 that they were receiving messages from him/her. Since people taking part in the survey didn’t get specifications on how to quantify interactions with others, those aforementioned numbers vary to some extent. We therefore will analyze the average of corresponding rows/columns from „Sent”/”Receive” sheets.

Using networkx built-in functions we have plotted and analyzed a directed graph of interactions in our cohort. Maximum in-degree value has the node number 32 with 54 people sending messages to this person. His out-degree value is 56 meaning that he is sending messages to 56 people. This is also the maximum value in out-degree column.

Also person 22 is very popular – sends and receives messages from 52 people. The balance for person 4: 45/45 and same for person 47 and 68, for person 62: 47/48, person 53: 45/46, person 37: 46/46.

Based only on this characteristic we could conclude that nodes number 32 and 22 are the hubs and are the most popular and influential people in our cohort.

Next we have analyzed the closeness centrality. Closeness centrality of a node is the reciprocal of the sum of the shortest path distances form this node to all other n-1 nodes. Higher values of closeness indicate higher centrality. In our analysis we redefined the ‘distance’ as the reciprocal. (explain…..) Node number 22 shows the highest value in terms of this measure: 4,69. Not far behind are 32 with value 4,44, nodes 45 and 51 both with value 4,52 and node 18 with value 4,58.

Furthermore we have used betweenness centrality to have more insight into the network. Betweenness centrality of a node is the sum ~~of the fraction~~?? of all-pairs shortest paths that pass through this node. We have used also weights of links as an argument in this measure so the edges do not have equal impact but the ones with higher values have more impact than those with lower values. The most important node in this measure turned out to be number 22 with value 2375,5, then node 73 with value 1950, node 51 with value 1350, and node 18 with value 1330.

Then we used eigenvector centrality which computes the centrality for a node based on centrality on its neighbors. The highest eigenvalue has node 22 (0,177). Next one is node 62 (0,167), following node 32 (0,163), and node 47 (0,161).

The clustering coefficient is the highest for nodes 59 (0,0424), 51 (0,0412) and 20 (0,04).

Based on this analysis we conclude that the most popular and influential people are 32 and 22.