ISTE 782: Visual Analytics

Assignment 3

Due date: November 1, 2024, 11:59 PM

<u>**Part 1:**</u>

Question 1. <u>Network Analysis with Superheroes</u> (35 points):

The Marvel Universe is characterized by the multitude of heroes and also the connections between them. Take a look at the <u>Marvel Universe Social Network data</u> Download the **"nodes.csv"** and **"edges.csv"** files.

Open Gephi. Open a new project. Go to Data Laboratory and load the nodes and edges files in the same workspace. Import the graph as an <u>**undirected**</u> graph. Do the following:

1) In the "Layout" section, run the "Fruchterman-Reingold" visualization algorithm

2) Go to "Filters". Go to "Topology -> Degree Range." Drag the "Degree Range" parameter into the Queries section. Check if there are any heroes that have 2 or fewer connections. Set the minimum value for the degree range to be 2.

3) Go to "Statistics" and run "Network Diameter".

4) Go to "Appearance" and select "Nodes". Go to "Size" option and choose "Ranking". Rank the nodes according to their "Degree".

5) Go to the "Color" option. Choose a "Blue-Yellow-Red" color scheme. Color the nodes that have the lowest degree centrality with the blue color and the ones that have the highest degree centrality with the red color.

6) Go to "File" and export the graph that you created as **"gephi_degree_centrality.pdf"**.

7) Go back to your project and run a modularity analysis with the default options. Report the number of communities that you found.

8) Go to "Appearance" tab and choose "Partition". Choose "Modularity Class" from the dropdown menu. Color the communities as you wish.

9) Check that now you are able to see different communities in your graph with distinguishable colors.

10) Go to "File" and export the graph that you created as **"gephi_communities.pdf"**.

11) Go to "File" again and save the project file that you created as **"gephi_analysis.gephi"**.

Question 2. Summary (15 points):

Write a 200-250 word summary of what you found in the steps above. Make sure that you emphasize what you found in each step of the analysis. Save the summary file as **"q1_summary.pdf"**

Part 2:

Question 3: Oil Pipelines (50 points)

Oil pipelines can be visualized using networks, as well. Open the Wikipedia page on the "List of oil pipelines". Go to Europe section and take a look at the map as well as different pipelines in Europe. You will be creating a network of pipelines for this question.

Step 1 (10 points):

Explore each link under the Europe section and—by writing on a piece of paper—note the countries that they pass through. Two countries are connected if a pipeline passes through them.

After you find a list of all countries in the oil pipeline network, manually create nodes and edges files. Save the file that you created as **"oil_pipeline_edges.csv"** and "**oil_pipeline_nodes.csv**"

Example: Create a **nodes.csv** file with at least two columns: id and label

id	label
1	Baku
2	Tbilisi
3	Ceyhan

In the **edges.csv** file, include the following columns: id, source, target, label. source and target should reference the id from nodes.csv.

id	source	target	Label
1	1	3	Baku–Tbilisi–Ceyhan pipeline

Step 2 (25 points):

Import the data (the network you created) in Gephi

Use the "Fruchterman-Reingold" layout and decorate your network.

Save the plot that you created as **"oil_pipeline_network_graph.png"**.

Step 3 (15 points):

Write a 50-100 word short summary of your findings. Comment on the network. Save the summary file as **"q3_summary.pdf"**

General rules:

Submit **gephi_degree_centrality.pdf, gephi_communities.pdf, gephi_analysis.gephi, q1_summary.pdf, oil_pipeline_edges.csv, oil_pipeline_nodes.csv, oil_pipeline_network_graph.png and q3_summary.pdf.**