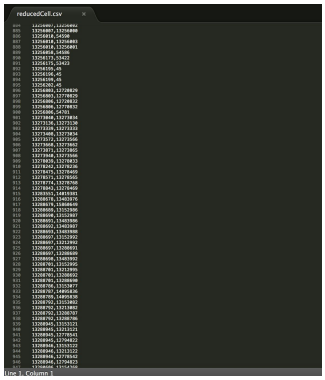
Cluster Viewer Documentation

Project GitHub:

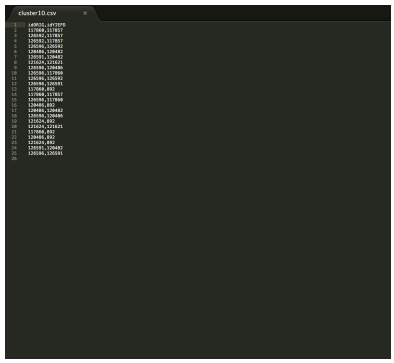
https://github.com/jinnkafka/ResearchProject

Project Description:

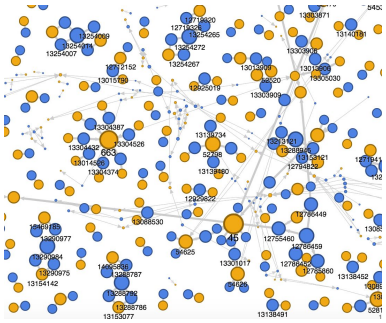
Given a CSV files contains thousands lines of relational numbers, the Cluster Visualizer (CV) parses all the numbers and generate multiple CSV files that representing a smaller subset of the original CSV numbers. More importantly, these subsets are all clusters from the original numbers.



Original Large Database



Individual Cluster Database



Original Graph Visualization

Individual Cluster Visualization

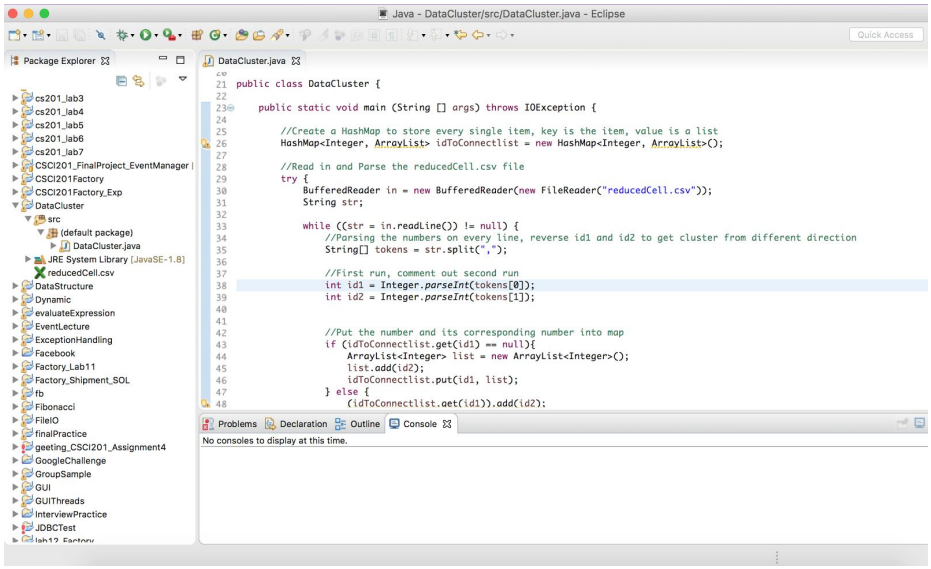
Step-by-Step Instruction:

1. Download DataCluster.java from the project's Github repository. At line 29:

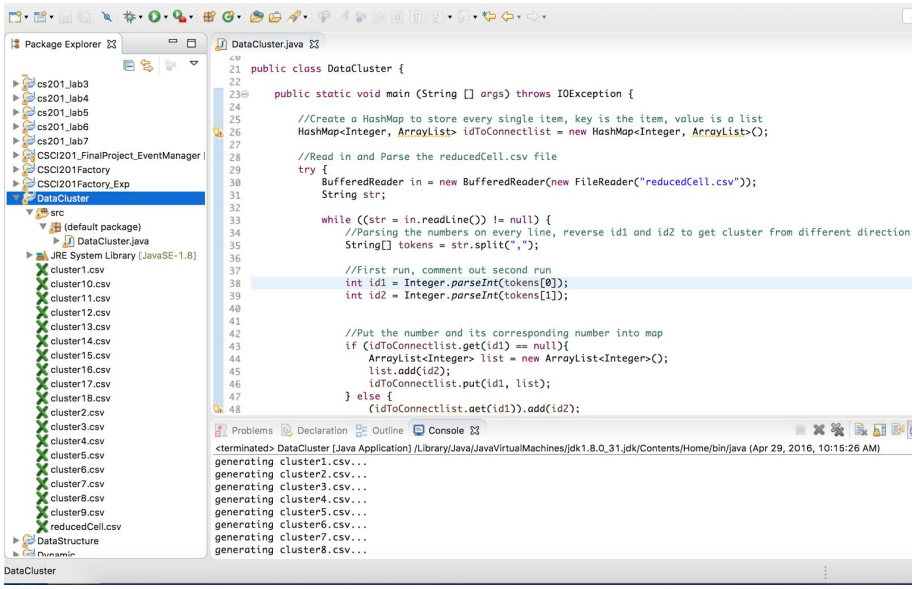
```
BufferedReader in = new BufferedReader(new FileReader("reducedCell.csv"));
```

Replace “reducedCell.csv” with the name of your file, and make sure your CSV file is in the same folder as DataCluster.java.

The best way to run DataCluster.java is inside an Integrated Development Environment (IDE) such as Eclipse or IntelliJ. If running on a terminal, the user will need to compile the project first before running.



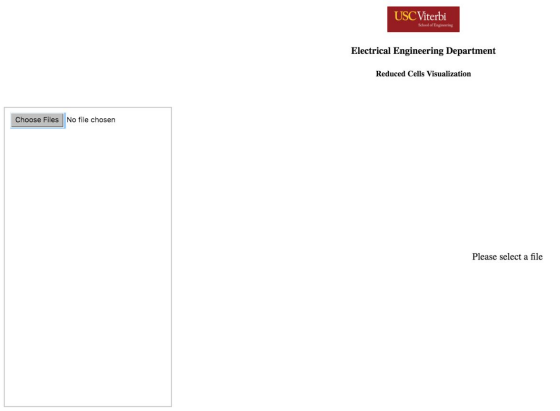
Before running the program



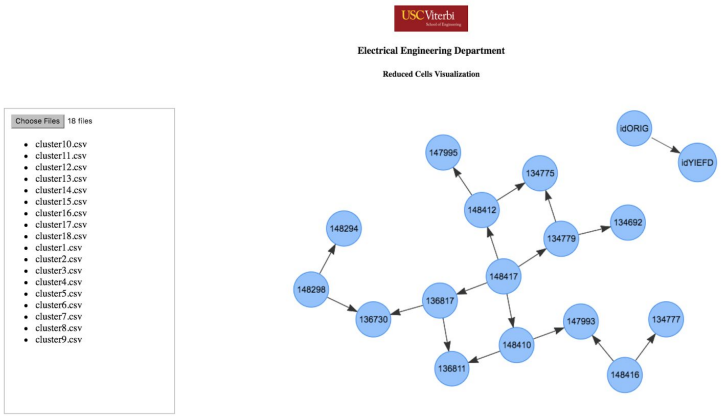
After running the program

If the generated files are not showing in the project folder after running the program, simply refresh the project folder.

2. Go to http://www-scf.usc.edu/~chenjin/ClusterViewer/reduced_cells_graph.html



And then click “Choose Files” button on top left corner. Choose all the CSV files you want to view. Click on the individual file name to see the graph visualization.



Since the project doesn’t have a database, every time user :

- Refresh the webpage
- Upload new files

will erase all the existing files. The user will need to upload the files again to view the graphs.

Questions and Suggestions:

Please email chenjin@usc.edu, or leave an issue ticket on the project Repository, if you have any questions and suggestions. Thank you for using Cluster Viewer !