**A Step-by-Step Guide to Creating a Gephi Dataset**

Brian Tsz Ho Wong

1. Extract data from the dataset and then sort it in a spreadsheet.

A screenshot of a computer

Description automatically generated

1. Copy all the projects and actors into a new spreadsheet.

A screenshot of a computer

Description automatically generated

1. Go to ‘Data’, use ‘Advanced Filter’ to filter for duplicate items.

A screenshot of a computer

Description automatically generated

1. Copy the filtered list of projects and actors into a new spreadsheet. Assign an ID to each project and actor. Rename the spreadsheet to ‘Nodes’.

A screenshot of a computer

Description automatically generated

1. Go back to the first spreadsheet. Add a column next to the ‘Projects’ and ‘Actors’ columns, then copy the ‘Node’ to this spreadsheet.

A screenshot of a computer

Description automatically generated

1. Use the ‘Xlookup’ function in Excel to match the project and data to their IDs.

A screenshot of a computer

Description automatically generated

1. Open a new spreadsheet. Copy the ‘Actors’’ ID column to the ‘Source’ column and the ‘Projects’’ ID column to the ‘Target’ column. As in this case, these are the organisations that were involved in and invested in the projects. If we treat the ‘Projects’ as ‘Source’ and the ‘Actors’ as ‘Target’, it will imply that these organisations have evolved from the projects, which is contrary to what we want to visualise. Rename the spreadsheet to ‘Edges’.

A screenshot of a computer

Description automatically generated

1. Finally, go back to the ‘Nodes’ spreadsheet, and add the location where the projects and organisations are based, as well as the latitude and longitude coordinates of that location.

A screenshot of a computer

Description automatically generated