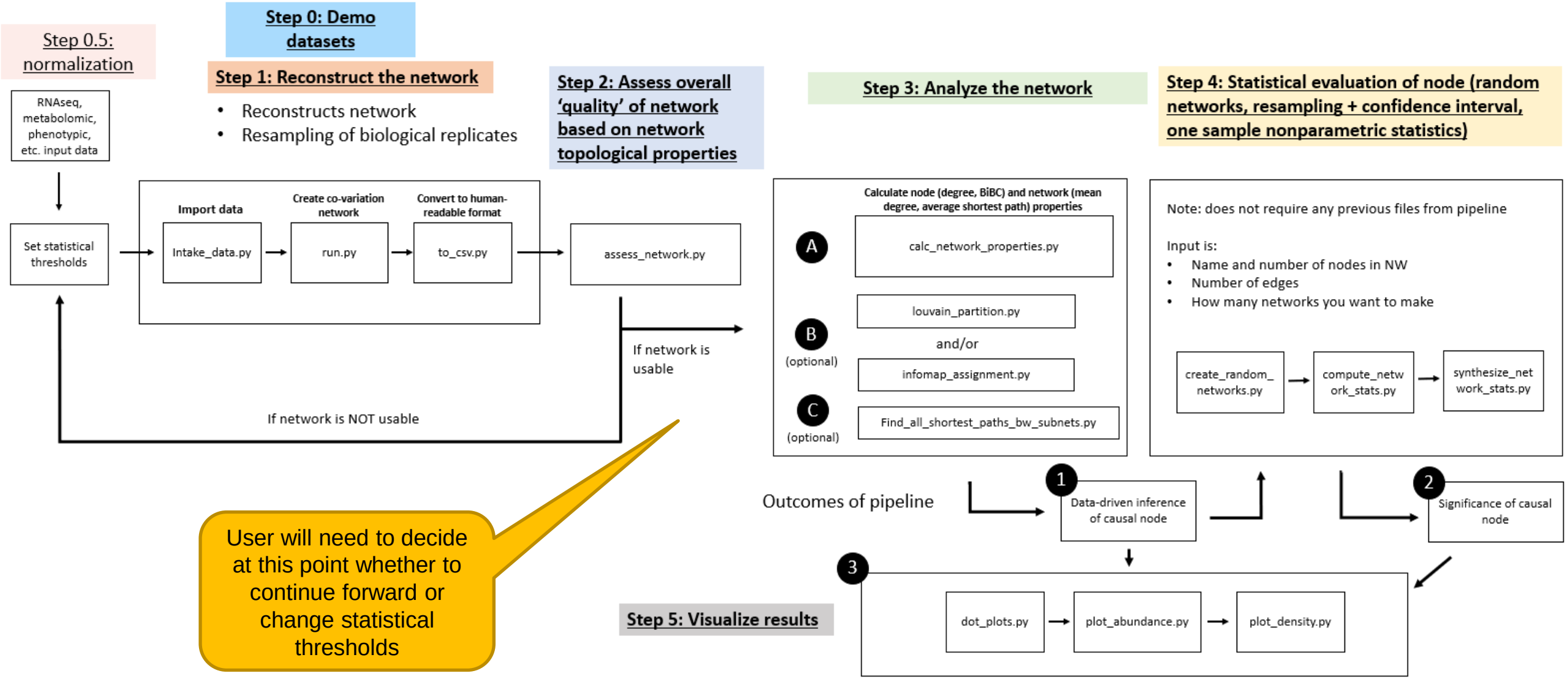
# Overall flow of pipeline



# Reconstruction

| Experiment name | Data matrices (Experiments) | Treatment mapping files |
| --- | --- | --- |
| “Expt1” | “Experiment1.csv” | “Experiment1_map.csv” |
| “Expt2” | “Experiment2.csv” | “Experiment2_map.csv” |

+ Add more

Mapping for parameter types

“type_map.csv”

Continue

There will be a ‘?’ help box pop up for each option to give an example and tell users what the option/text box does

# (Automatically generated from the mapping files)

Comparison thresholds

This table would need to be automatically generated based on the unique data type values in the mapping file

| data type | p-value | Fisher's combined p-value | FDR |
|---|---|---|---|
| gene | [default] | [default] | [default] |
| pheno | [default] | [default] | [default] |
| micro | [default] | [default] | [default] |

Find consistent fold change direction using: (●) Median ( ) Average

Groups to use for fold change calculation:

The below group names in the dropdowns will be automatically generated from the groups in the mapping files

Fold change = Group 1 / Group 2

Group 1 [Select group] Group 2 [Select group]

Continue

Correlation thresholds

| data type 1 | data type 2 | p-value | Fisher's combined p-value | FDR |
|---|---|---|---|---|
| gene | gene | [default] | [default] | [default] |
| pheno | pheno | [default] | [default] | [default] |
| micro | micro | [default] | [default] | [default] |
| gene | pheno | [default] | [default] | [default] |
| gene | micro | [default] | [default] | [default] |
| micro | pheno | [default] | [default] | [default] |

This table would need to be automatically generated based on all pairs of the unique data type values in the mapping file

◉ Find consistent edges across ALL groups (default)

○ Find consistency across a proportion (between 0.5 and 1) "0.66"

Show an example value here between 0.5 and 1

I want to find consistent correlations across the same group in all

I want to find consistent correlations across only selected groups

**If option 1 (default):**

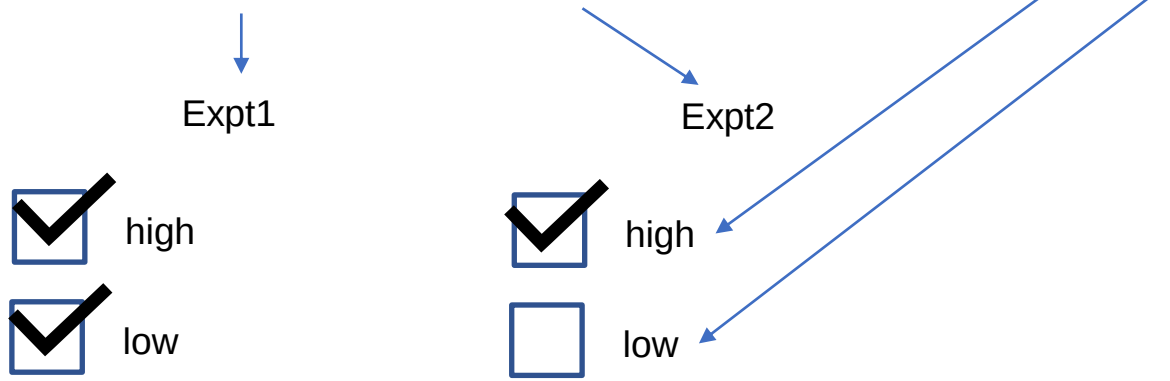
Choose the group to find consistent correlations in: ◉ high ○ low

**If option 2:**

Choose which groups to find consistent correlations between

These will be automatically populated from the mapping files in the first step

Expt1 and Expt2 are automatically populated from the experiment names in the first step




Create network

# Network assessment

Note: The input for this code is just one file, but the user will need to be able to view the output to decide if they want to use different thresholds for their network or move forward with analyzing their network

# Network analysis

**Cluster identification in network**

☑ Find clusters using the Louvain algorithm

☑ Find clusters using Infomap

Note: If the user wants to use outputs of modularity codes, they will need to first download the outputs and conduct cluster analysis separately, then feed the mapping file (output) to the BiBC script

**Bipartite Betweenness Centrality (BiBC) calculations**

What groups do you want to use for your BiBC analysis?

Group 1 “0” Group 2 “1”

Mapping file “mapping_file.csv”

☑ Normalize your Bipartite Betweenness Centrality results by the number of nodes in the subnetworks (checked by default)

**Random networks**

☑ Create random networks and analyze using the same groups and mapping file from BiBC calculation step (checked by default)

**Please note: depending on the size of the network, making/analyzing random networks can take an extensive amount of time.**

Analyze

# Visualization

**Dot plots**

Property to plot on X-axis of dot plots: BiBC [default]

Property to plot on Y-axis of dot plots: Node_degrees [default]

These drop-down menus would be automatically populated based on the names in the node_properties.txt file

Number of all top nodes to plot: "10"

Number of top nodes to plot per data type: "5"

Plot the top ____ percent of nodes: %

**Abundance plots**

Color the bar plots based on: Experiment group [default]

Group the bar plots based on: Experiment number [default]

For simplicity, the code right now only supports these two options

The mapping files and data files will be the same ones used in the reconstruction step, not sure if we want user input here for them

List any additional nodes you would like to plot on the abundance and density plots

"gene1"
"gene5"

This is a text box where the user can type any names of nodes they want to plot

Plot results