

1

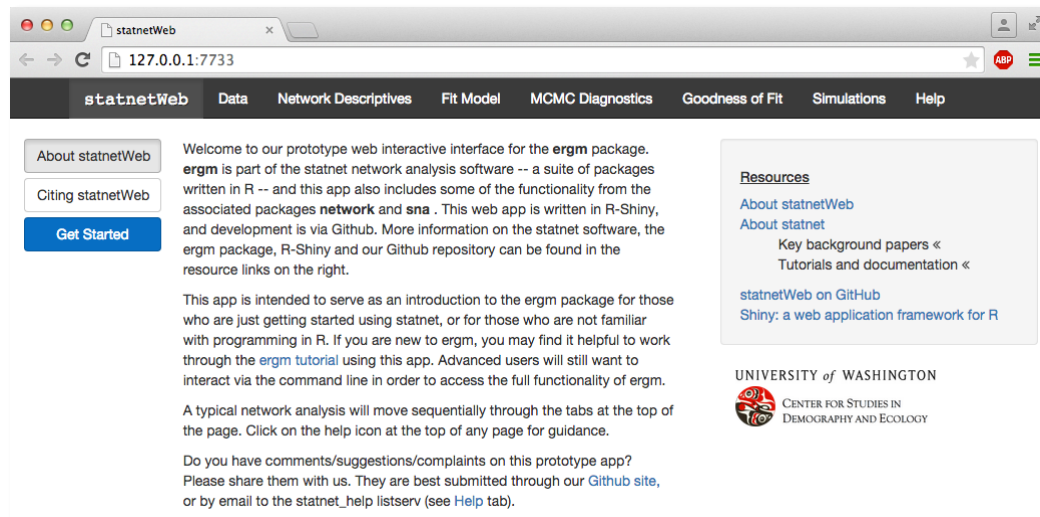
Lab: Descriptive network analyses

Explore with statnetWeb, just a bit

Intro to statnetWeb

statnetWeb is a graphical user interface for network analysis

- An Rshiny App
- Runs in a web browser, like epiweb
- Wiki: <https://github.com/statnet/statnetWeb/wiki>



Using statnetWeb on your computer

- Install the statnetWeb package
 - `install.packages("statnetWeb")`
- Load the package and launch the application
 - `library(statnetWeb)`
 - `run_sw()`

Note: Closing the browser window ends your session, so if you want to save something, do it before you quit

Network data in statnetWeb

- On the “Data” page, you can upload networks from multiple sources

- Internal: built-in networks
- External: R, Excel, Pajek files

The screenshot shows the 'Data' page upload form in statnetWeb. It includes a 'File type' dropdown menu set to 'matrix of relational data (*.csv or *.rds)', a 'Browse...' button, and a status message 'No file selected.'. Below this is a box showing 'name: NA' and 'size: NA'. The form is divided into two sections: 'Matrix Type' and 'Network Attributes'. The 'Matrix Type' section has four radio button options: 'Adjacency matrix' (selected), 'Bipartite adjacency matrix', 'Incidence matrix', and 'Edge list'. The 'Network Attributes' section has four checkbox options: 'directed?' (checked), 'loops?', 'multiple?', and 'bipartite?'.

File type

matrix of relational data (*.csv or *.rds) ▼

Browse... No file selected.

name: NA
size: NA

Matrix Type

- ☒ Adjacency matrix
- ☐ Bipartite adjacency matrix
- ☐ Incidence matrix
- ☐ Edge list

Network Attributes

- ☒ directed?
- ☐ loops?
- ☐ multiple?
- ☐ bipartite?

Examples in statnetWeb

- Load the “faux.mesa.high” network
High school network simulated from Add Health data
- We’ll explore more network concepts using these data

The screenshot shows the 'File type' dropdown menu in the statnetWeb interface. The 'built-in network' option is selected in the main dropdown, and the 'faux.mesa.high' option is highlighted in the expanded list. Below the dropdown, a text box contains the following description:

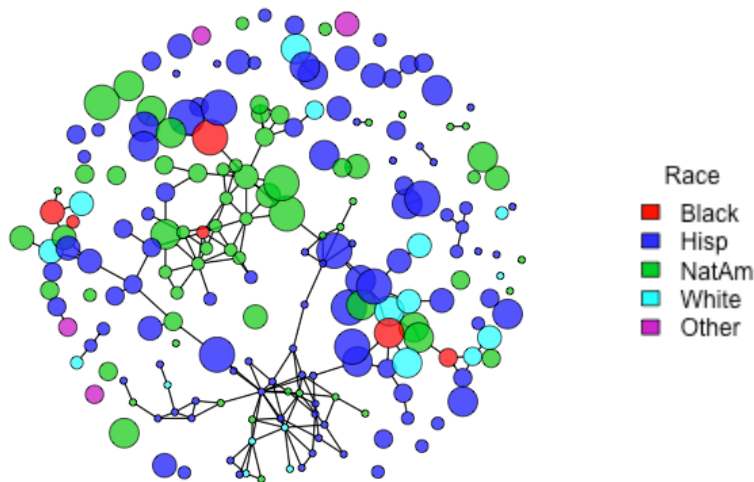
This data set represents a simulation of an in-school friendship network for faux.mesa.high because the school community on which it is based is largely Hispanic and Native American.

faux.mesa.high is a network object with 205 vertices (students) and 1,200 edges (mutual friendships).

The vertex attributes are Grade, Sex, and Race. The Grade attribute has values 7 through 12, indicating each student's grade in school. The Race attribute is based on the answers to two questions: one on Hispanic or Latino ethnicity and one on race.

Attributes

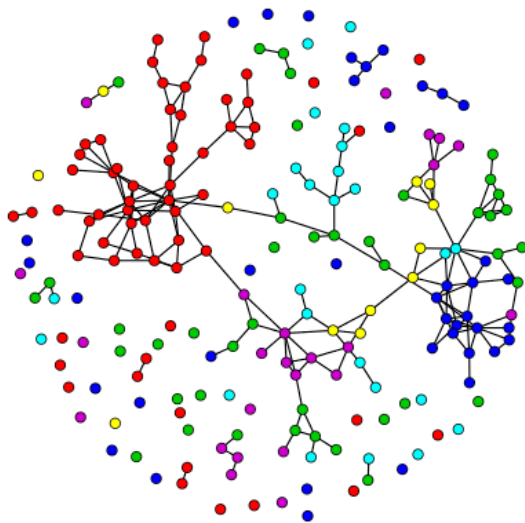
- Individual nodes can have attributes like age, race, sex, etc.
- *Explore:*
 - *Color-code or size nodes on the network plot*
 - *Sort or search attributes in the interactive table*
 - *Look at histograms of attribute counts*
 - *How do these descriptive help you understand the structure of the network?*



Node mixing by attribute

- Collapses the adjacency matrix into categories
- Cell counts = # links between nodes in row and col. categories

Mixing Matrix



Grade
7
8
9
10
11
12

Choose attribute

Grade

Note: Marginal totals can be misleading for undirected mixing matrices.

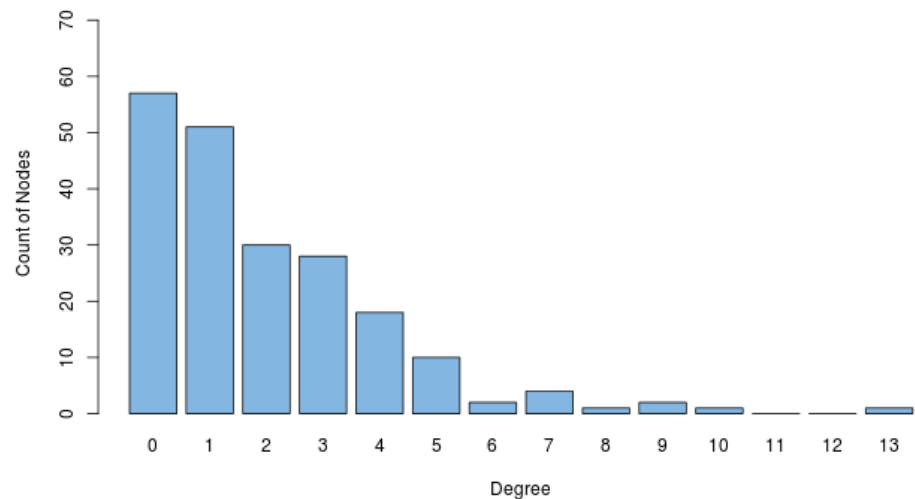
	7	8	9	10	11	12
7	75	0	0	1	1	1
8	0	33	2	4	2	1
9	0	2	23	7	6	4
10	1	4	7	9	1	5
11	1	2	6	1	17	5
12	1	1	4	5	5	6

Degree metrics

- Node level: The number of edges “adjacent” to a node
 - Every node has a degree $\deg(i)$
 - Di-graphs have in- and out- degrees, $\text{ideg}(i)$ and $\text{odeg}(i)$
 - Indegree: the number of arcs that terminate at n_i
 - Outdegree: the number of arcs that originate from n_i
- Network level: The degree distribution
 - Well-known parametric degree distributions: Uniform, Binomial, Poisson, Power-law
 - *The degree distribution in an empirical network may or may not resemble any of these*

Degree distribution

- To view it in statnetWeb:

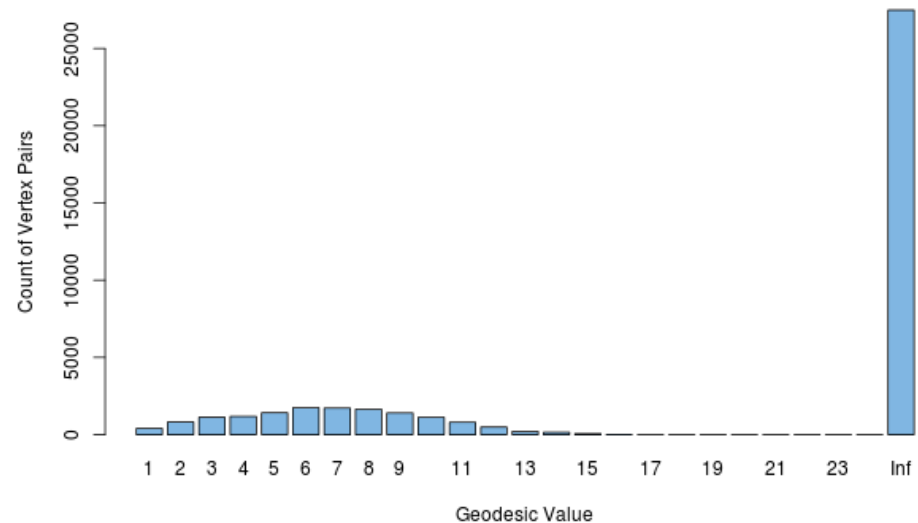


Connectivity measures: Geodesic

- Nodes are **reachable** if there is a path between them.
- A **geodesic** is the shortest path between two nodes
 - Two nodes have an infinite geodesic distance if they are unreachable

Geodesic distribution

- To view it in statnetWeb:

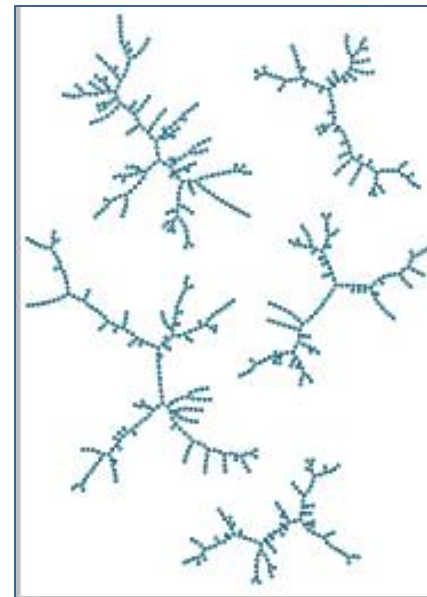


The last bar represents the node pairs with infinite geodesic distance

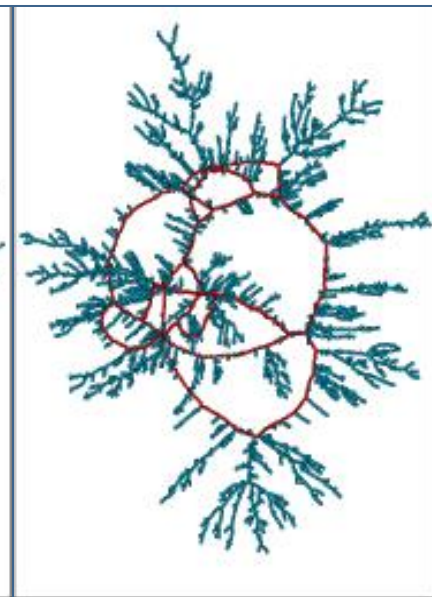
Connectivity measures: Components

- If some node pairs are unreachable, the graph will have multiple “**components**”
 - subgraphs of reachable nodes
- The component size distribution is another basic property of the network

Multiple components



1 giant component



NB: Think about how this connectivity comes to be created...