Network Data Sets

Networks Data Sets Build into R Packages

Data Set	Description	Package	Additional Information	How to Access*

Sampson's Monks	This includes a small, directed social network of likes among monks preparing for ordination at a monastery.	ergm	https://cran.r-project.org/web/packages/ergm/ergm.pdf	Use Import Script:
				`>source("./Import Scripts/Sampson's Monks.R")`
				
				Access Orignial Format from Package:
				`>install.packages("ergm")`
				`>data("sampson", package = "ergm")`

Florentine Families	This is an adjacency matrix for a small, directed network of the marriages held by the Florentine family and others in Rome.	network	https://cran.r-project.org/web/packages/network/network.pdf	Use Import Script:
				`>source("./Import Scripts/Florentine Families.R")`
				
				Access Orignial Format from Package:
				`>install.packages("network")`
				`>data("flo", package = "network")`

Enron Emails	This is a large, directed network of the e-mail relationships among Enron employees preceding investigations.	igraphdata	https://cran.r-project.org/web/packages/igraphdata/igraphdata.pdf	Use Import Script:
				`>source("./Import Scripts/Enron Emails.R")`
				
				Access Orignial Format from Package:
				`>install.packages("igraphdata")`
				`>data("sampson", package = "igraphdata")`

French Hospital RFID	This is a medium-sized, directed network of contacts between patients and healthcare workers in a French hospital across 4 days.			Use Import Script:
				`>source("./Import Scripts/French Hospital.R")`
				
				Access Orignial Format from Package:
				`>install.packages("igraphdata")`
				`>data("rfid", package = "igraphdata")`

U.S. Airports	This is a large, directed network of passenger flights between airports within the US in December, 2010.			To Use Import Script:
				`>source("./Import Scripts/US Airports.R")`
				
				Access Orignial Format from Package:
				`>install.packages("igraphdata")`
				`>data("USairports", package = "igraphdata")`

US Governmental Network Data Sets

Data Set	Description	Additional Information	How to Access*
Supreme Court Opinions	This is a data set for a medium-sized, bimodal network of relationships among Supreme Court justices based on judicial opinion citations.	For more information, contact Professor Rolfe.	See professor Rolfe for access to data.
Congressional Bill Co-Sponsorship	This is a data set for a large, directed network showing cosponsor-ships among legislators for the 93rd through 114th congresses.	http://fowler.ucsd.edu/cosponsorship.htm	To Use Import Script: `>source("./Import Scripts/Congressional Cosponsorship.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Cosponsorships*

International Relations Network Data Sets

Data Set	Description	Additional Information	How to Access*
IMF CPIS	This is an edgelist for a large, directed network of total cross-border short- and long-term equity holdings between banking economies of sovereign countries.	http://cpis.imf.org/	Use Import Script: `>source("./Import Scripts/IMF CPIS.R")` Location of Raw Data: *Network Analysis Working Directory/Data/IMF CPIS*
Big, Allied, and Dangerous (Terror Networks)	This website has data on prominent terrorist organizations and their social networks among other organizations.	http://www.start.umd.edu/baad/database	Organization attribute data is in the class working directory. However, to get the network connections data, you must to contact the BAAD prodject at U. Albany: *https://www.start.umd.edu/baad-frequently-asked-questions#q1*

International Relations Network Data Sets (Continued)

Data Set		Description	Additional Information	How to Access*
Correlates of War	Militarized Interstate Disputes	This is an edgelist for a large, directed network of militarized interstate disputes from 1870 to 2009; incident-level data also available online.	http://www.correlatesofwar.org/data-sets/MIDs	Use Import Script: `>source("./Import Scripts/COW MID.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Correlates of War/MID*
	Alliances	This is an edgelist for a large, undirected network of formal alliances between states from the 1800's to present.	http://www.correlatesofwar.org/data-sets/formal-alliances	Use Import Script: `>source("./Import Scripts/COW Alliances.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Correlates of War/Alliances*
	Trade	This is an edgelist for a large, directed network of trade between states from 1870 to 2009; nation-level trade data also available online.	http://www.correlatesofwar.org/data-sets/bilateral-trade	Use Import Script: `>source("./Import Scripts/COW Trade.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Correlates of War/Trade*
MJ Peterson's Data on the IMF and World Bank		This is a non-relational data set on the educational backgrounds of World Bank and IMF employees, which can be used for a medium-sized, bimodal network.	Contact MJ Perterson for information.	Contact MJ Peterson for access to data.

Network Data Sets from Previous Classes

Data Set		Description	More Information	How to Access*
Swiss Climate Policy	Alliance-Opposition	These are adjacency matrixes for a small weighted directed network of stakeholders in the Swiss climate policy debate, indicating which nodes saw which other nodes as allies or opposition. The data includes several node attributes.	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27427	Use Import Script: `>source("./Import Scripts/Swiss Climate Alliance-Opposition.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Swiss Climate Policy*
	Collaboration	These are adjacency matrixes for a small unweighted undirected network of stakeholders in the Swiss climate policy debate, indicating which nodes collaborated with which other nodes. The data includes several node attributes.	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27427	Use Import Script: `>source("./Import Scripts/Swiss Climate Alliance-Opposition.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Swiss Climate Policy*
	Influence	These are adjacency matrixes for a small unweighted directed network of stakeholders in the Swiss climate policy debate, indicating which nodes saw which other nodes as influential. The data includes several node attributes.	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27427	Use Import Script: `>source("./Import Scripts/Swiss Climate Collaboration.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Swiss Climate Policy*
	Preference Distance	These are adjacency matrixes for a small weighted undirected network of stakeholders in the Swiss climate policy debate, indicating how different node's policy preferences are. The data includes several node attributes.	https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/27427	Use Import Script: `>source("./Import Scripts/Swiss Climate Preference Distance.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Swiss Climate Policy*

Network Data Sets from Previous Classes (Continued)

Data Set		Description	How to Access*
Game of Thrones	Regions	This is an edgelist for a small, undirected network of regions based on shared borders from within the Game of Thrones story line.	Use Import Script: `>source("./Import Scripts/Game of Thrones Regions.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*
	Locations	This is an edgelist for a medium sized, undirected network of locations weighted by distance based on the Game of Thrones story line. It contains several node and edge attributes and is more detailed than the Regions network.	Use Import Script: `>source("./Import Scripts/Game of Thrones Locations.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*
	Interactions	This is an edgelist for a medium-sized, undirected, weighted network of interactions between characters in the Game of Thrones Novels.	Use Import Script: `>source("./Import Scripts/Game of Thrones Interactions.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*
	Kills	This is an edgelist for a medium-sized, directed, unweighted network between characters in the Game of Thrones Novels, indicating who has killed who.	Use Import Script: `>source("./Import Scripts/Game of Thrones Kills.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*
	Marriages	This is an edgelist for a medium-sized, undirected, unweighted network of marriages between characters in the Game of Thrones Novels.	Use Import Script: `>source("./Import Scripts/Game of Thrones Marriage.R")` Location of Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*
	Like/Dislike	This is an adjacency matrix for a medium-sized, directed network of likes and dislikes between characters within the Game of Thrones story line.	Use Import Script: `>source("./Import Scripts/GoT LikeDislike")` To Get Raw Data: *Network Analysis Working Directory/Data/Game of Thrones*

Network Data Sets from Previous Classes (Continued)

Data Set		Description	How to Access*
Game of Thrones	TV Series	This is an edgelist for a medium sized undirected, unweighted network of interactions between characters in the universe of the Game of Thrones TV Series. It seems to be based on data from seasons one and two.	Use Import Script: `>source("./Import Scripts/Game of Thrones TV.R")` Location of Raw Data: *http://www.zengxiaohan.com/media/code/data/a-song-of-ice-and-fire-desc.json*
House of Cards	Social Network	This includes an edgelist for a small, undirected network of observed relationships among characters in the House of Cards story line.	Use Import Script: `>source("./Import Scripts/House of Cards.R")` Location of Raw Data: *Network Analysis Working Directory/Data/House of Cards*
	Counter-factual Social Network	This includes an edgelist for a small, undirected network of relationships among characters in the House of Cards story line, imagining that Claire and Frank had never married.	Use Import Script: `>source("./Import Scripts/House of Cards Counterfactual.R")` Location of Raw Data: *Network Analysis Working Directory/Data/House of Cards*
Celebrity Sex Networks		This is a large, undirected network of sexual relationships between prominent celebrities, as reported by a website on celebrity relationships.	Use Import Script: `>source("./Import Scripts/Celebrity Sex.R")` To Get Raw Data: *Network Analysis Working Directory/Data/Celebrity Sex*
Senate Tweets		This is a dataset for a large undirected network of senators based on the similarities of their tweets.	Contact Noah Kouchekinia noah.a.k@me.com for access to data and information.

Some Important Notes:

*For the R language in the "How to Access" column uses relative paths. For them to work, your working directory must be set to the working directory provided for the class using the `setwd()` function.

*For most of the data sets, you have the option to access the raw data or use an import script.

- Accessing the raw data means it will be read into R in the format it was originally provided in. It is then up to you to arrange the data into network objects.
- The import scripts will read in the data, create network objects, and provide a short blurb explaining what has been done. The import scripts are purpose built for this class. The network objects they create are compatible with the weekly syntax files used for course assignments.

This list is only a small subset of the network data available. For more options, look at:

- Socio-Patterns
- Stanford Network Analysis Project
- Network Repository