

# NR 592

## Assignment 5

*INSERT YOUR NAME HERE*

*INSERT THE DATE HERE*

*Last Updated: 2019-04-24*

### Instructions

**Show your work.** The response to each question should include all of the commands necessary to reproduce your results, and should be commented to give a naive reader an idea of the intent behind each section of code. Answers to questions should be clearly shown in the code output or spelled out in text or comments.

Before submitting your assignment, ensure that your code for each question works correctly when the lines are run in sequence, starting with an empty environment. You may submit either this R Markdown document with your code entered into the appropriate code blocks, or a regular R script with the question numbers as comments and the answers following. If you choose to Knit the R Markdown document to HTML or PDF, include those files with your submission (but it is not required that it Knit successfully).

### Question 1

Choose one of the following types of data:

1. Ecological community data
  - a. Note: Must have at least ten “sites” (can be locations, years, etc.)
2. Likert survey data
  - a. Note: Must have at least three Likert-type variables and at least one nominal demographic variable
3. Network data
  - a. Note: Must have at least one additional variable (numeric and/or nominal) aside from vertices, edges, and labels. Can be data from a social network, interaction network, bibliometric network, etc.
4. Geospatial data
  - a. Note: Must have at least two additional variables (numeric and/or nominal) aside from coordinates and labels

Find an appropriate data set of that type and load it into R.

If you are having trouble finding data, look through the example data sets given in the Course Project description or data sets attached to published papers in your field.

**Important: Include the data file in your submission and include a link to its source here.**

**# Source**

Delete this and paste a link to the source of your data set here.

**# Load data**

Perform any data cleaning and wrangling necessary to get your data into the format(s) required for the analyses below.

## Question 2

Conduct exploratory analyses of the data, consisting of at least one descriptive statistic and at least one rough exploratory plot. Briefly describe your findings.

Delete this and enter your response here.

## Question 3

Conduct at least one hypothesis test of your data. You do not need to describe the results, but it should be clear from comments and context what test is being performed and why.

## Question 4

Create one linear model (regression, ANOVA, etc.), correlation matrix, distance matrix, or adjacency matrix of your data (whichever is most appropriate). Show the result.

## Question 5

Conduct one of the following sequences of analyses described in Week 5 or Week 6 of the course material, depending on your data type:

1. Ecological community data
  - a. Diversity, Richness, Evenness, and Species accumulation curve
  - b. Diversity, Richness, Evenness, and Multidimensional scaling plot (classical or nonmetric)
  - c. Diversity, Richness, Evenness, and Geospatial plot (if spatial data are included)
  - d. Principal component analysis, Principal component plot
  - e. Cluster analysis, Silhouette plot, and Cluster plot
2. Likert survey data
  - a. Exploratory factor analysis and Confirmatory factor analysis
    - i. Note: Randomly subset your data set into two halves: one for EFA and the other for CFA
  - b. Principal component analysis, Principal component plot
  - c. Cluster analysis, Silhouette plot, and Cluster plot
3. Network data
  - a. Two descriptive statistics (e.g. centrality, diameter, distance, edge density), Network visualization, Community detection
4. Geospatial data
  - a. Geospatial plot
    - i. Note: Must include basemap layer and at least one layer of shapefiles of your choice

- b. Geographic distance matrix, Distance matrix of additional variable for the same points, Mantel test, Hierarchical clustering plots of each distance matrix

**Notes: Analyses should be fully commented, with results shown. If the analyses include a plot, the plot should be as polished and publication-ready as possible. Different analyses are permitted if they are comparable in effort and scope to those given above.**

(pdf / Rmd)