

Assignment: Spatial Diversity

Student Name; Z620: Quantitative Biodiversity, Indiana University

03 February, 2017

OVERVIEW

This assignment will emphasize primary concepts and patterns associated with spatial diversity, while using R as a Geographic Information Systems (GIS) environment. Complete the assignment by referring to examples in the handout.

After completing this assignment you will be able to:

1. Begin using R as a geographical information systems (GIS) environment.
2. Identify primary concepts and patterns of spatial diversity.
3. Examine effects of geographic distance on community similarity.
4. Generate simulated spatial data.

Directions:

1. Change “Student Name” on line 3 (above) with your name.
2. Complete as much of the assignment as possible during class; what you do not complete in class will need to be done on your own outside of class.
3. Use the handout as a guide; it contains a more complete description of data sets along with the proper scripting needed to carry out the assignment.
4. Be sure to **answer the questions** in this assignment document. Space for your answer is provided in this document and indicated by the “>” character. If you need a second paragraph be sure to start the first line with “>”.
5. Before you leave the classroom, **push** this file to your GitHub repo.
6. When you are done with the assignment, **Knit** the text and code into an html file.
7. After Knitting, please submit the completed assignment by creating a **pull request** via GitHub. Your pull request should include this file *spatial_assignment.Rmd* and the html output of Knitr (*spatial_assignment.html*).

1) R SETUP

In the R code chunk below, provide the code to:

1. Clear your R environment
2. Print your current working directory,
3. Set your working directory to your “/Week4-Spatial” folder, and

2) LOADING R PACKAGES

In the R code chunk below, do the following:

1. Install and/or load the following packages: **vegan**, **sp**, **gstat**, **raster**, **RgoogleMaps**, **maptools**, **rgdal**, **simba**, **gplots**, **rgeos**

Question 1: What are the packages **simba**, **sp**, and **rgdal** used for?

Answer 1:

3) LOADING DATA

In the R code chunk below, use the example in the handout to do the following:

1. Load the Site-by-Species matrix for the Indiana ponds datasets: `BrownCoData/SiteBySpecies.csv`
2. Load the Environmental data matrix: `BrownCoData/20130801_PondDataMod.csv`
3. Assign the operational taxonomic units (OTUs) to a variable 'otu.names'
4. Remove the first column (i.e., site names) from the OTU matrix.

Question 2a: How many sites and OTUs are in the SiteBySpecies matrix?

Answer 2a:

Question 2b: What is the greatest species richness found among sites?

Answer 2b:

4) GENERATE MAPS

In the R code chunk below, do the following:

1. Using the example in the handout, visualize the spatial distribution of our samples with a basic map in RStudio using the `GetMap` function in the package `RgoogleMaps`. This map will be centered on Brown County, Indiana (39.1 latitude, -86.3 longitude).

Question 3: Briefly describe the geographical layout of our sites.

Answer 3:

In the R code chunk below, do the following:

1. Using the example in the handout, build a map by combining lat-long data from our ponds with land cover data and data on the locations and shapes of surrounding water bodies.

```
# 1. Import TreeCover.tif as a raster file.  
  
# 2. Plot the % tree cover data  
  
# 3. Import water bodies as a shapefile.  
  
# 4. Plot the water bodies around our study area, i.e., Monroe County.  
  
# 5. Convert lat-long data for ponds to georeferenced points.  
  
# 6. Plot the refuge pond locations
```

Question 4a: What are datums and projections?

Answer 4a:

5) UNDERSTANDING SPATIAL AUTOCORRELATION

Question 5: In your own words, explain the concept of spatial autocorrelation.

Answer 5:

6) EXAMINING DISTANCE-DECAY

Question 6: In your own words, explain what a distance decay pattern is and what it reveals.

Answer 6:

In the R code chunk below, do the following:

1. Generate the distance decay relationship for bacterial communities of our refuge ponds and for some of the environmental variables that were measured. Note: You will need to use some of the data transformations within the *semivariogram* section of the handout.

```
# 1) Calculate Bray-Curtis similarity between plots using the `vegdist()` function
# 2) Assign UTM latitude and longitude data to 'lats' and 'lons' variables
# 3) Calculate geographic distance between plots and assign to the variable 'coord.dist'
# 4) Transform environmental data to numeric type, and assign to variable 'x1'
# 5) Using the `vegdist()` function in `simba`, calculate the Euclidean distance between the plots for
# 6) Transform all distance matrices into database format using the `liste()` function in `simba`:
# 7) Create a data frame containing similarity of the environment and similarity of community.
# 8) Attach the columns labels 'env' and 'struc' to the dataframe you just made.
# 9) After setting the plot parameters, plot the distance-decay relationships, with regression lines in
# 10) Use `simba` to calculate the difference in slope or intercept of two regression lines
```

Question 7: What can you conclude about community similarity with regards to environmental distance and geographic distance?

Answer 7:

7) EXAMINING SPECIES SPATIAL ABUNDANCE DISTRIBUTIONS

Question 8: In your own words, explain the species spatial abundance distribution and what it reveals.

Answer 8:

In the R code chunk below, do the following:

1. Define a function that will generate the SSAD for a given OTU.
2. Draw six OTUs at random from the IN ponds dataset and plot their SSADs as kernel density curves. Use **while loops** and **if** statements to accomplish this.

```
# 1. Define an SSAD function
# 2. Set plot parameters
# 3. Declare a counter variable
# 4. Write a while loop to plot the SSADs of six species chosen at random
```

8) UNDERSTANDING SPATIAL SCALE

Many patterns of biodiversity relate to spatial scale.

Question 9: List, describe, and give examples of the two main aspects of spatial scale

Answer 9:

9) CONSTRUCTING THE SPECIES-AREA RELATIONSHIP

Question 10: In your own words, describe the species-area relationship.

Answer 10:

In the R code chunk below, provide the code to:

1. Simulate the spatial distribution of a community with 100 species, letting each species have between 1 and 1,000 individuals.

```
# 1. Declare variables to hold simulated community and species information
```

```
# 2. Populate the simulated landscape
```

While consult the handout for assistance, in the R chunk below, provide the code to:

1. Use a nested design to examine the SAR of our simulated community.
2. Plot the SAR and regression line.

```
# 1. Declare the spatial extent and lists to hold species richness and area data
```

```
# 2. Construct a 'while' loop and 'for' loop combination to quantify the numbers of species for progress
```

```
# 3. Be sure to log10-transform the richness and area data
```

In the R code chunk below, provide the code to:

1. Plot the richness and area data as a scatter plot.
2. Calculate and plot the regression line
3. Add a legend for the z-value (i.e., slope of the SAR)

Question 10a: Describe how richness relates to area in our simulated data by interpreting the slope of the SAR.

Answer 10a:

Question 10b: What does the y-intercept of the SAR represent?

Answer 10b:

SYNTHESIS

Load the dataset you are using for your project. Plot and discuss either the geographic Distance-Decay relationship, the SSADs for at least four species, or any variant of the SAR (e.g., random accumulation of plots or areas, accumulation of contiguous plots or areas, nested design).