

## Lab 5: Working with Geospatial Data

For this week's lab, we are going to look at college going outcomes for California school districts.

We will be using several files:

- "DistrictAreas2122.shp": Shapefile of the school district boundaries
- "District College Outcomes.csv": College outcomes by school district
- "college locations.csv" College locations

Please download the Lab 5 R script on the course Canvas page to get started.

### Part I. Cleanup

1. Read in the District Areas Shapefile. Transform the CRS to CA Albers Equal Area.
2. Read in the District College Outcomes csv file. Convert the CDCCode to character. Create a variable that represents the CSU going rate. Create a variable that represents the UC going rate.
3. Left Join the District Areas sf object to the District College outcomes.
4. Read in the college locations csv file. Convert the data frame to a sf object. Transform the CRS to CA Albers Equal Area. Create a sf object for campuses that are part of the CSU system and in LA County. Create a sf object for campuses that are part of the UC system and in LA County.

### Part II. Analysis

1. Create a map of CSU going rates for districts in Los Angeles County with the CSU locations in LA county overlayed. Adjust the color scale so that low values are red, high values are green.
2. Create a map of the UC going rates for districts in Los Angeles County with the UC locations in LA county overlayed. Adjust the color scale so that low values are red, high values are green.
3. Discuss the relationship between college proximity and college going rates in the two maps. Provide an explanation for the patterns observed. Which do you expect to display a stronger spatial relationship and why?
4. Calculate the distances between the School district Centroids and the CSU's in LA county. Compute the average distance by school district. Put the information in a table. Compare this information to the CSU going rate by district. Does the district with the highest CSU going rate also the closest to a CSU campus?
5. Create one additional visualization of your choice that is related to college going rates (map or otherwise). Discuss the findings.
6. ChatGPT acknowledgement: if you used ChatGPT, please explain what you used it for in a few sentences.

### Submission Guidelines

- Submit your writeup that includes your figures as a pdf file (recommended word count: 500 words)
- Submit your R script with the code used to generate your figures

### Grading Rubric

#### Writeup:

- Did you generate the output that was required?
- Does your writeup analyze the data in a thoughtful manner?
- Does your writeup reflect original work that you created, and do you cite external sources when they are used?

#### Script:

- Did your code generate the tables in your writeup?
- Does your code use the tools covered during lecture?
- Is your code organized and well commented?