Examining the Racial, Environmental, and Economic Influences on COVID-19 Mortality in Louisiana

ENVS3: Environment and Society

Profs. Rich Howarth and J.T. Erbaugh

# Assignment Description

In this assignment, you will investigate the relationship between COVID-19 mortality and social, economic, as well as environmental variables. The goal of the assignment is to combine your background knowledge with data analysis and visualization to identify factors that might have affected COVID-19 mortality in Louisiana. You will begin by reading two short papers on racial burdens of environmental emissions in the United States and hypothesized drivers of disproportionate COVID-19 mortality in Louisiana. You will then use a web application specifically designed for this course to analyze parish-level data. Using the analytical tools and visualizations you generate from the web application, you will develop and test hypotheses related to what might be driving COVID-19 mortality in Louisiana.

# Pre-Assignment Literature

You are expected to read two short papers before beginning this assignment:

1. [C. W. Tessum, D. A. Paolella, S. E. Chambliss, J. S. Apte, J. D. Hill, J. D. Marshall, PM2.5 polluters disproportionately and systemically affect people of color in the United States. Sci. Adv. 7, (2021).](https://advances.sciencemag.org/content/7/18/eabf4491)
2. [Terrell, Kimberly A., and Wesley James. "Racial Disparities in Air Pollution Burden and COVID-19 Deaths in Louisiana, USA, in the Context of Long-Term Changes in Fine Particulate Pollution." Environmental Justice (2020).](https://www.liebertpub.com/doi/10.1089/ENV.2020.0021)

These papers are meant to provide you with key context and background information which will

*be incredibly helpful to you as you begin working with the dataset and answering the questions below.*

# Louisiana Dataset

The dataset you will be working with includes 13 unique variables for all 64 parishes (counties) of Louisiana. Below is a description of each variable and the data sources where they were retrieved from:

 First Wave Covid Deaths per 10k: Deaths per 10,000 people for the first COVID-19 wave, defined as from 1/22/2020 - 7/1/2020.4

 Second Wave Covid Deaths per 10k: Deaths per 10,000 people for the second COVID-19 wave, defined as being from 7/1/2020 - 11/1/2020.4

 Third Wave Covid Deaths per 10k: Deaths per 10,000 people for the third COVID-19 wave, defined as being from 11/1/2020 - 4/28/21.4

 Covid Deaths per 10k thru 4/28/21: Cumulative deaths per 10,000 people for the COVID-19 pandemic from 1/22/2020 - 4/28/21.4

 Population: Population for the parish based upon 2020 Census.1

 Median Household Income: Median household income.2

 Percent 65 and over: Percentage of total county population that is 65 years of age and older.2

 Percent Black: Percentage of population that is non-Hispanic Black or African American.2

Poverty Rate: percentage of population living below the poverty level. The Census Bureau uses a set of money income thresholds that vary by family size and composition to determine who is in poverty. If a family's total income is less than the family's threshold, then that family and every individual in it is considered in poverty. The official poverty thresholds do not vary geographically, but they are updated for inflation using the Consumer Price Index (CPI-U). The official poverty definition uses money income before taxes and does not include capital gains or noncash benefits (such as public housing, Medicaid, and food stamps).10

 Percent Rural: Percentage of population living in a rural area.5

 Respiratory Hazard (RH) Weighted Ave: The sum of hazard quotients for toxins that affect the respiratory system. Hazard quotients are defined as the ratio of the potential exposure to a substance and the level at which no adverse effects are expected (calculated as the exposure divided by the appropriate chronic or acute value). Because different air toxics can cause similar adverse health effects, combining hazard quotients from different toxics is often appropriate. A hazard index (HI) of 1 or lower means air toxics are unlikely to cause adverse noncancer health effects over a lifetime of exposure. The RH data is provided by census tract, so an average for the parish is computed as the weighted average by the population for the entire parish.3

 TRI Total On Site Air Release (pounds): The amount of toxic emissions released on site at all facilities in pounds within a parish for the entire year of 2019.**9**

 Cancer Alley: Whether parish is a part of “Cancer Alley” in Louisiana. Variable is binary (meaning it is either a 0 or 1 value), where 1 = Cancer Alley Parish.7

Sources:

1Retrieved from usafacts.org, data from 2020 US census.

2Retrieved from the 2020 County Health Rankings, data from 2018 census.

3Data from the 2014 National Air Toxics Assessment (EPA).

4Retrieved from usafacts.org, data from the CDC and Louisiana Department of Health.

5Retrieved from the 2020 County Health Rankings, data from 2010 census.

7Based on the definition of Cancer Alley in James et al. 2012 (<https://www.mdpi.com/1660-4601/9/12/4365/htm>).

9Data is from the 2019 Toxics Release Inventory (TRI) and is summed for all facilities within a parish.

10Data from the 2020 Census.

# Short Answer Questions

1. Look at the list of variables and make sure you understand what they represent. Based on the pre-assignment papers you read and your general knowledge, what do you predict to be the three or four most important variables associated with COVID-19 mortality in Louisiana? Please write these down as a set of hypotheses. What reasons do you have for thinking these variables are important?
2. Begin by looking at the Interactive Map. What patterns emerge between race, class, and COVID-19 mortality? Do these patterns change across the different waves (Wave 1, Wave 2, and Wave 3)?
3. Examine which variables correlate with the COVID-19 “hot spots” on the Interactive Map and then use the Correlation Matrix for each COVID-19 wave. Which three to four variables correlate the highest (both positively and negatively) with each wave’s COVID-19 mortality rate? ***\*\*Keep in mind:*** *correlation does not necessarily imply causation!*
4. Test your hypotheses from Question 1 by examining bivariate relationships between your *dependent variables (COVID-19 morality)* and *independent variables (variables that explain the dependent variable).* Use the Linear Regression to see how well the bivariate relationship can be described by a line with a positive or negative slope.
	1. Which variables have a statistically significant linear relationship with each wave of COVID-19?
	2. Of those that were statistically significant, which had the best fit (highest R2 value) for each COVID-19 wave?
5. Based on your analysis, what general conclusions can you make between socioeconomic, social, and environmental variables and COVID-19 mortality in Louisiana?
6. Reflect on the data and your analytical process. What are the benefits and drawback from this approach? What other variables would you want to include in future analysis?