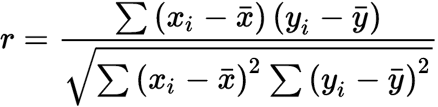
**Appendix**

# **Web Application**

The web application has three tabs used for the visualization of data, each with a different type of visualization or analysis. Below is a brief description of the information contained in each tab:

* **Tab 1: Interactive Map**
  + Contains a map with different layers of data which can be turned on and off. COVID-19 Mortality is broken down into each wave (three in total) and the cumulative deaths Louisiana experienced through April 2021 and is represented by bubbles which vary in size by the relative magnitude of the death rate per 10,000. All other variables are represented by colors for each county. You can turn the bubbles and the outline of Cancer Alley on/off by clicking on their icons in the legend (upper left of map)
* **Tab 2: Correlation Matrix**
  + This diagram is a way of visualizing how the variables in the dataset are correlated with one another. You will be able to examine how the *independent variables* and *dependent variable* correlate with one another, as well as with themselves. In this case, the independent variables are the different demographic and environmental measures, while the dependent variables are the COVID-19 mortality by each wave. The dropdown at the top of the window allows you to change which COVID-19 wave you are choosing as the dependent variable.
  + The color of each box is representative of its correlation coefficient, which is a measure of linear correlation between two variables. It is the covariance of the two variables, divided by the product of their standard deviations.
  + Pearson’s correlation coefficient, *r* , is defined as:



where r = correlation coefficient

= values of the x-variable in a sample

= mean of the values of the x-variable

= values of the y-variable in a sample

= mean of the values of the y-variable

* **Tab 3: Linear Regression**
  + This plot allows you to plot any variable against another, where each point represents a parish in Louisiana. Remember, we generally think about the y-variable as being the *dependent* or *response variable*, and the x-variable as being the *independent* or *explanatory variable*.
  + When you click Linear Regression to be “On,” the results of an Ordinary Least Squares (OLS) regression are displayed based on the variables you selected. This method fits a linear model with coefficients w = (w1, …, wp) to minimize the residual sum of squares between the observed targets in the dataset, and the targets predicted by the linear approximation.
  + In the upper right corner the equation of the line, the R2 value, and the P value for the t statistic are shown:
    - R-squared value : also known as the coefficient of determination and is the proportion of the variance in the dependent variable that is predictable from the independent variable. An R2 of 1 indicates that the regression predictions perfectly fit the data.
    - P > | t | : The "t'' value is computed by dividing the estimated value of the coefficient by its standard error and is a measure of the likelihood that the actual value of the coefficient is not zero (i.e. that the slope of the line is not zero). **If this value is < 0.05 (based on a 95% confidence interval) then it can be considered statistically significant that the coefficients of the model are not zero and there exists a linear relationship between the dependent and independent variables.**