
Data Handling Summary

Jaylon Kiper

Problem Statement/Background Information

With the data I have gathered for The Butterfly Project, I want to present my findings to the Louisville Community and show how the black maternal mortality rate looks compared to the white maternal mortality rate.

Hypothesis

Why is the black maternal mortality rate so much higher than the white maternal mortality rate?

- African American mothers are dying at three to four times the rate of non-Hispanic white mothers, and infants born to African American mothers are dying at twice the rate as infants born to non-Hispanic white mothers.

Cleaning/Organizing the Data

- Having to clean the dataset on ratios was not difficult at all. Most of the variables were organized when I began to manipulate the data.
 - As I worked in both Excel and Tableau, I had formed two separate graphs to show my data.
 - My Tableau sheet consisted of plotting the years by the ratio as an area graph to show the spikes in how much black maternal mortality increases compared to white maternal mortality.
 - My Excel graph groups all years by five year time periods with the average ratio increasing from the year 1915 to 2016.
-

Sharing My Message

I want to show the Louisville community that this is an important matter when it comes to an African American mother's and child's life. It ultimately comes down to two factors. The first being how African American mothers receive less financial health advantages compared to those who are white. The second is that this type of data can't be manipulated easily by the public. It was very odd that this one topic of interest is non-accessible to everyone to work with.

Challenges with the Data

There were various challenges not really involved in working with this dataset, but in finding one's big and accessible to display better data.

- One issue within finding datasets on maternal health is scarcity.
- Not a lot of data is accessible to the public eye, making all forms of manipulating and working in these health datasets is nigh impossible without lots of information.
- One reason could be confidentiality. Having possible access to a patient's personal information can be dangerous for the health industry.
- All accessible data is recorded until the year 2016 for most of our datasets, making more present data non-researchable.
- The current form this data is represented is majorly accessible only in PDF format, not Excel/CSV format.

Graphs:



