**Executive Summary**

My passion in helping others has led me to pursue a data analytics career within the healthcare industry. My primary goal of my capstone project is to continue to bring awareness to the healthcare disparities among breast cancer mortality rates among women of color. There is sufficient data to support that black and brown women die at higher rates from breast cancer compared to white women.

There are several factors that play into why women of color are dying at higher rates, one being having access to adequate healthcare facilities and treatment. According to the American Cancer Society, women aged 40 to 44 years of age should start having mammography screenings at least once a year. However, women aged 45 to 49 years of aged are recommend getting a mammograph screen annually. For many women, mammograms are the best way to find breast cancer early, when it is easier to treat and before it is big enough to feel or cause symptoms. At this current time, the American Cancer Society states that a mammogram is the best way to find breast cancer for most women of screening age.

For my capstone, I want to look at the mortality rates of women in Davidson County, provide the breakdown of deaths by race and age and to look at the number of FDA mammograph facilities available by zip code within Davidson County.

**Data Question**

I would like to see which zip codes have the most mammograph facilities and insurance and compare the death rates to see if there is a correlation to having access to these facilities. Would like to see if there is a correlation to having access to a mammogram facility for screening and the death rates.

**Minimum Viable Product (MVP)**

Tools: Python/Jupyter Notebooks for gathering and analysis of data and Tableau for data visualization

I would like to provide a map showing the death rates among women of color by zip codes or congressional districts and illustrate if there is a health disparity in having access to a mammogram facility. The focus is to bring awareness of the health disparity that exist within our communities and to encourage our healthcare officials and or local officials to invest in providing more healthcare resources to combat the breast cancer mortality rates among women of color.

**Data Sources**

*Document the data you use and the source of that data*

**Main data: make sure to find death rates by Zip code instead of county level.**

State Cancer Profiles: NIH – National Cancer Institute

<https://statecancerprofiles.cancer.gov/deathrates/index.php?stateFIPS=47&areatype=county&cancer=055&race=00&sex=2&age=001&type=death#results>

CDC – US Cancer Statistics -State and County

<https://gis.cdc.gov/Cancer/USCS/#/StateCounty/>

Congressional Districts – US Cancer Statistics – Tennessee (Davidson is District 5)

<https://gis.cdc.gov/Cancer/USCS/#/CongressionalDistricts/>

Mammography facilities – where to find them by zip code

<https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfMQSA/mqsa.cfm>

Old Congressional District Map – to find the outline in case if I can’t map the deaths/death rates by zip code.

<https://www.capitol.tn.gov/districtmaps/Congress12.html>

Davidson County Congressional District is 5. The old CG district 5, before recent rezoning, consist of all Davidson & Dickson counties and most of Cheatham County.

My Congressional District – can find some census data here.

<https://www.census.gov/mycd/?st=47&cd=05>

Census Reporter – still a working progress in building my table.

<https://censusreporter.org/>

**Schedule (through )**

1. Get the Data (11/11/2022)
2. Clean & Explore the Data (12/3/2022)
3. Create Presentation of your Analysis (12/15/2022)

* Should be a presentation, but could include a Jupyter Notebook or dashboard in Excel, Tableau, or PowerBI

1. Internal demos (12/20/2022)
2. Demo Day!! (1/5/2022)