

Executive Summary

This section provides an overview to the project. It should briefly touch on the motivation, data question, data to be used, along with any known assumptions and challenges.

What is one thing that everybody has, yet everyone is different? Life... Life is the single most important thing a person can have. That is why it is so crucial that we sustain and preserve it. How has life been for us since 1980 up into 2014? Here we will take a deep dive using big data into life expectancy and mortality rates using a few different variables. Breaking down every single state by individual county we will determine the male and female life expectancy and mortality rate using ages 0 through 85. My main goal on this journey will be to see how each county/state has progressed or regressed over the 34-year span.

Motivation

Here you will go into more detail about why you have chosen this project.

The word “life” has always felt so strange to me. Is it just a number with age, a long road of A to B, or what you make of it from beginning to end? There is so much that plays a factor in it, so I really wanted to dig deep and look at the life expectancies and mortality rates. What I found was a large data set breaking those variables down county by county, state by state to see how rates have increased or decreased. I am excited to breakdown the male and female mortality risks state by state to better understand how America has progressed from the years 1980 through 2014.

Data Question

Present your question. Feel free to include any research/articles that are relevant or show where others have attempted to answer this question.

1. Has life expectancy increased or decreased in each state from 1980 to 2014?
2. What is the trend in mortality risks for each age group?
3. By county/state does the mortality risk tend to be higher for male or female?
4. Overall, how has mortality risk changed in each age group from 1980 to 2014?

Minimum Viable Product (MVP)

Define your MVP. This should be a description of what your final capstone will look like, including visualizations, how the analysis will be presented, who the intended audience is, etc.

- Map of the United States broken down by state, possibly county showing statistical data on mortality risks or life expectancy.
- Dashboard or other visualization painting a clear picture of potential trends or patterns.

Schedule (through <date of demo day>)

1. Get the Data (**Complete**)
2. Clean & Explore the Data (**July 30th**)
3. Create Presentation of your Analysis (**August 12th**)
 - Should be a presentation, but could include a Jupyter Notebook or dashboard in Excel, Tableau, or PowerBI
4. Internal demos (**August 16th**)
5. Demo Day!! (**August 20th**)

Data Sources

Document the data you use and the source of that data

[United States Life Expectancy and Age-specific Mortality Risk by County 1980-2014 | GHDx \(healthdata.org\)](https://healthdata.org/United-States-Life-Expectancy-and-Age-specific-Mortality-Risk-by-County-1980-2014)

[United States Physical Activity and Obesity Prevalence by County 2001-2011 | GHDx \(healthdata.org\)](https://healthdata.org/United-States-Physical-Activity-and-Obesity-Prevalence-by-County-2001-2011)

Known Issues and Challenges

Explain any anticipated challenges with your project, and your plan for managing them.

- **Issue:** Cleaning multiple large datasets for a possible merge.
Solution: Utilize python to make clear and concise columns easy to digest.
- **Issue:** Making separate columns for male and female values across all datasets.
Solution: Use Python or Excel to pull multiple values from a single column and make multiple columns based on sex with single value.
- **Issue:** Making a visualization that isn't clustered or too busy considering all the different counties.
Solution: Input product into multiple dashboards/visualizations to see which best depicts the final product.