Correlation in Cause of Death: Demographic and Temporal Relationships

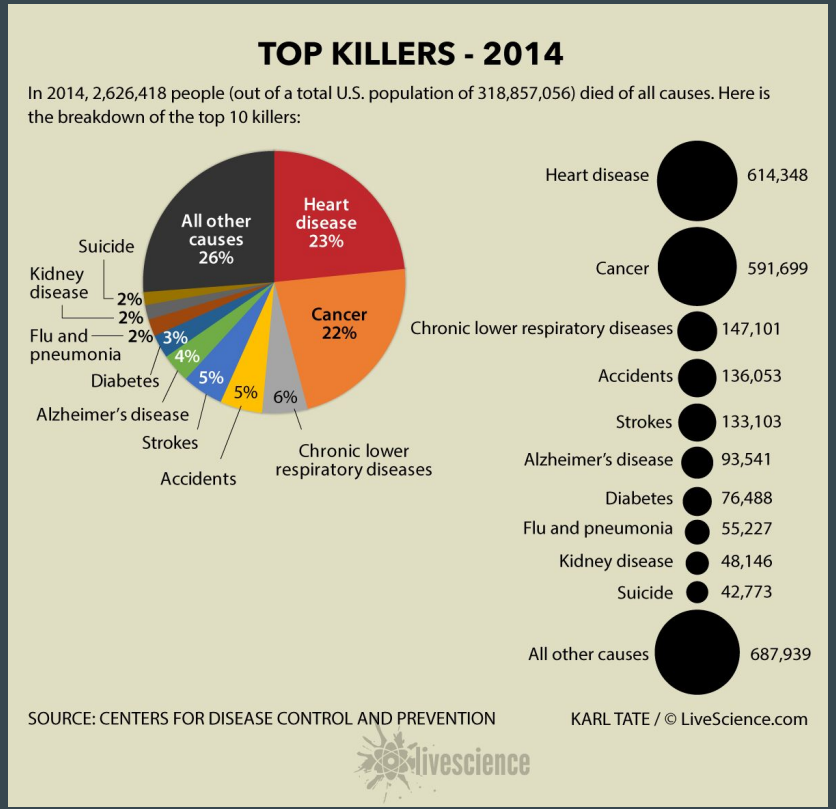
Derek Gorthy, Bryce Melvin, Michael Condon

The goal

Many studies on causes of death focus on specific diseases and specific age ranges. Our project will focus on the correlation between the manner of death, the place of death, and personal demographics; we also intend to look at temporal factors such as month and day.

Prior work

The prior work is CDC doing work on the dataset since it is CDC's dataset. There are also statistical analysis of causes of deaths in America, by sites like LiveScience.



Datasets

Death in the United States (2014), U.S. Census Bureau (from kaggle.com):

2.4 Million objects,

30 object attributes

Possible second dataset for data on income, sex, race to find trends between cause of death, time of death etc and income

Tools

Python

Github

R (For easy statistics)

Evaluation

We can use our results to supply doctors with information about the likelihoods of deaths from a certain disease, as a fact-check for political statements about controversial issues (e.g. deaths by shootings, violence in low income neighborhoods), and as a tool for education about healthy lifestyles