

Cancer Mortality Exploration

w203 Teaching Team

Background

In this lab, imagine that your team is hired by a health government agency. They would like to understand factors that predict cancer mortality rates, with the ultimate aim of identifying communities for social interventions, and of understanding which interventions are likely to have the most impact. Your team was hired to perform an exploratory analysis to help the agency address their goals.

Data

You are given a dataset for a selection of US counties, “**cancer.csv**”. The dependent (or target) variable in this data is named “deathRate”.

The labels of some of the variables are listed below; the rest of the variables should be self-explanatory.

avgAnnCount:	"2009-2013 mean incidences per county"
povertyPercent:	"Percent of population below poverty line"
popEst2015:	"Estimated population by county 2015"
PctPrivateCoverage:	"Percentage of the population with private insurance coverage"
PctPublicCoverage:	"Percentage of the population with public insurance coverage"

Objective

Perform an exploratory analysis to understand how county-level characteristics are related to cancer mortality.