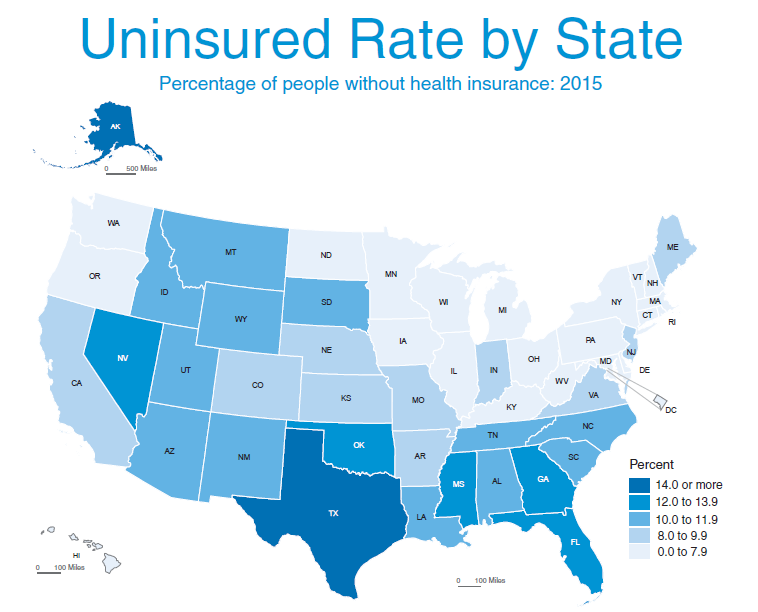
**Introduction**

Health insurance coverage is defined as health insurance that covers basic health care needs excluding disability, dental, vision, accident or prescription medicine plans according to Current Population Survey Annual and Economic Supplement (CPS ASEC). Basic health care services includes in and out-of-area emergency services, inpatient hospital and physician care, outpatient medical services, laboratory and radiology services and preventive services [Va. Code Ann. § 38.2-4300].

An individual can access health care through private insurance provided by an employer or purchase directly through the insurance company. Government also offer health insurance programs which includes federal programs such as Medicare, Medicaid, the Children’s Health Insurance Program (CHIP), Department of Veterans Affairs, individual state health plans and more. Such programs are aim to aid people who cannot access health care due to financial constrain and/or other factors.

Although the federal programs are intended to aid people to have access to basic health care needs, there are still a large number of people who are uninsured in the United States. In 2015, the percentage of people without health insurance coverage for the entire calendar year was 9.1 percent or 29 million from a total population of 321 million. The uninsured rate has improved year-over-year since the enactment of the Affordable Care Act (sometimes known as ACA, PPACA or “Obamacare”) in March 2010. Prior to ACA implementation, there were 49.9 million or 18.4 percent people without access to health care. The ACA is a law, which provides consumers with subsidies (“premium tax credits”) that lower costs for households with incomes between 100 percent and 400 percent of the federal poverty level and expanded the Medicaid program to cover all adults with income below 138 percent of the federal poverty level.

Illustrated in the heat map below is the uninsured rate by state for people without health insurance in 2015. Note that darker shaded colors have higher uninsured rate. Texas had the highest uninsured rate at 17.1 (4,615,000 people in Texas had no health care insurance out of a population size of 27,430,000) versus Massachusetts had the lowest rate at 2.8 (189,000 people in Massachusetts had no health care insurance out of a population size of 1,653,000).



Illustrated in the graph below is number of people who are uninsured by state in 2015 and each state’s population size.

As illustarated n

In 2015, the percentage of people without health insurance coverage for the entire calendar year was 9.1 percent or 29 million

We decided to examine the driving factors that influence the uninsured rate.

**Data Collection**

Prior to conducting our research, we obtained and reviewed datasets from various sources for uninsured rate as well as chosen independent variables “predictors” for all 51 states. We had to find different datasets for poverty rate, unemployment rate, median income, Medicaid, Medicare, age, race/ethnicity, fertility rate and more for all 51 states and merged the datasets together in order to perform our analysis. We used datasets from the timeframe of 2014 to 2015. A final dataset is assembled using these merged datasets.

We obtained our datasets from the following sources:

* Kaiser Family Foundation estimates based on the Census Bureau's March 2016 Current Population Survey (CPS: Annual Social and Economic Supplement) – <http://kff.org/>
* U.S. Department of Health & Human Serves – <https://aspe.hhs.gov/evaluations>
* United States Census Bureau – https://www.census.gov/data/datasets/2011/demo/health-insurance/cps-revision.html
* Healthcare.gov – <https://www.healthcare.gov/glossary/affordable-care-act>

**Independent Variables “Predictors”**

Through examining our datasets and research, we found that there are many factors contribute to healthcare uninsured rate, which is the dependent variable in our regression analysis. However, we only selected independent variables “predictors” that we determined are the primarily drivers that can assist in explaining uninsured rate.

The independent variables and associated definitions as defined by the United States Census Bureau contributing to uninsured rate are as follows:

|  |  |  |
| --- | --- | --- |
| Item # | Data | Definition/Description |
| 1 | Poverty Rate | The poverty rate is the percentage of the number of people who fall below the poverty line and the total population |
| 2 | Unemployment Rate | The percentage of the labor force that is employed. |
| 3 | Median Income | The median household income in the U.S. is $56,500 according to the most recent government data. Approximately half of all Americans will fall below this household income amount, and the other half will fall above. |
| 4 | Medicaid Rate | Medicaid is a health care program that assists low-income families or individuals in paying for long-term medical and custodial care costs. Percentage of people under the Medicaid program |
| 5 | Medicare Rate | Medicare is the federal health insurance program for people who are 65 or older. Percentage of people under the Medicare program |
| 6 | Race | Percentage of demographic uninsured: White, Black, Hispanic, Asian, etc. |
| 7 | Age | Percentage of age groups uninsured: age groups such as Children 0-18, Adults 19-25, Adults 26-34, Adults 35-54, Adults 55-64 and 65+ |
| 8 | Adult with and without children | Percentage of Adult with and without children uninsured |
| 9 | Percent of graduates with debt (4-year or above) | Percentage of graduates with debt |
| 10 | Fertility Rate | The ratio of live births in an area to the population of that area |

**Regression Analysis**

We performed a regression analysis on the uninsured rate (dependent variable) and associated independent variables “predictors.” We wanted to analyze how well the predictors could explain the uninsured rate for each state. As such, we utilized the ten independent variables to develop a multiple regression model.

Prior to running the regression model, we performed preliminary assessments on each independent variable cause and effect on the dependent variable (uninsured rate). These independent variables have a positive and negative effect on the uninsured rate to increase or decrease.

|  |  |  |  |
| --- | --- | --- | --- |
| Item # | Independent Variable “Predictor” | Cause and Effect (+/-) | Reasoning |
| 1 | Poverty Rate | Positive | Poverty rate indicates that people do not have excessive income to purchase health care insurance resulting increase uninsured rate |
| 2 | Unemployment Rate | Positive | Unemployment rate indicates that people do not have health coverage through employment or private insurance thus increase uninsured rate |
| 3 | Median Income | Positive | Depending on the median income of an household, it will indicate if an individual can afford health coverage |
| 4 | Medicaid Rate | Negative | Individuals under Medicaid program (governmental health coverage program for the elderly 65+ will lower the uninsured rate |
| 5 | Medicare Rate | Negative | Individuals under Medicare program (governmental health coverage program for individuals who earn at and/or below poverty line) will lower the uninsured rate |
| 6 | Race | Positive | Depending on the race/ethnicity: Blacks, Asians, Whites, Hispanics can affect the uninsured rate. |
| 7 | Age | Negative | Adults will more likely to receive health coverage through employment or private, thus reduce the uninsured rate |
| 8 | Adult with and without children | Negative | Adult with children will most likely to have health coverage compare to adult without children |
| 9 | Percent of graduates with debt (4-year or above) | Positive | Graduates with debt will likely not have health coverage, thus driving up the uninsured rate |
| 10 | Fertility Rate | Positive | Fertility rate will drive down uninsured rate because parents need coverage for the newborn |

After our preliminary assessments for the independent variables, we ran a correlation matrix to get overall view of multicollinearity among uninsured rate and independent variables. Next, we used the Variance Inflation Factor (VIF) approach to identify and eliminate coefficients with high (VIF > 5). VIF measures the impact of collinearity among the variables or measures how much the variance of the estimated regression coefficients are inflated as compared to when the independent variables are not linearly related. We eliminated coefficients with VIF one by one and re-ran the regression model each time.

We also eliminated coefficient with p-value exceeding alpha > 0.05 from the regression until there were none left.