

**The City College of  
New York**



---

---

**EMPIRICAL RESEARCH FINAL  
PROJECT:**  
**Socioeconomic Factors Affecting Health  
Insurance In The United States**

---

---

**Hassan Fayyaz | Adeel Arshid | Tanmay Thomas**

**Professor John Schmitz**

**Foundations of Empirical Research (ECO B2100)**

**The City College of New York**

**May 24th, 2022**

---

---

## Abstract:

**Inquiry Question:** Is health insurance unequally distributed among individuals living in the United States?

**Primary Purpose:** To explore what variables have the biggest impact in determining the health insurance status of individuals living in the US? What are the possible disadvantages of not having health insurance? How can we provide health insurance to everyone living in the US by focusing on the key variables of the research? Would this solution actually increase the number of individuals who have health insurance?

**Thesis:** Many would argue that health insurance disparities do not exist: But in reality, health insurance is not distributed equally among people of color and underserved communities because of income inequality, the rising cost of health care, and immigration status.

**Prior belief/ Knowledge:** Individuals who are financially better off, with US-born citizenship status and higher education, are more likely to have health insurance in comparison to the underserved minorities with low income, non-US citizenship status, & lack of access to higher education.

Topic No#	<b><u>Table of Content</u></b>	Page No#
1.	Cover Page	1
2.	Table Of Content	2
3.	Assignment Instruction	3
4.	Introduction	4-6
5.	Literature Review	6-25
6.	Descriptive Statistic	25-42
7.	Feature Engineering	42-52
8.	Conclusion	52-56
9.	Reference	56

## Introduction:

This research aims to find out the socioeconomic factors affecting the health insurance status of individuals and provides a solution to the problem of unequal distribution of health insurance. Having no health insurance is a big problem because individuals with less access to recommended care receive poorer quality of care and experience worse health outcomes than insured individuals (McWilliams, 2009). Moreover, according to research by Kimberly Amadeo, the Former President of World Money Watch, “Every year, around 530,000 people declare medical bankruptcy” (Amadeo, 2021). Many would still argue that health insurance disparities do not exist: But in reality, health insurance is not distributed equally among people of color and underserved communities because of income inequality, the rising cost of health care, and immigration status.

In United States, it's hard to have access to proper healthcare coverage from individual. Not having access healthcare can lead to expensive treatment causing people to go into massive amount of debt. Unequal distribution of health insurance can affect people based of their race, gender, income, immigration status and sexual orientation. During 2019, in United States, almost 30 million individuals didn't have any kind of health coverage. Lack of health coverage can create massive debt on people and their families. Treatment can become expensive for low income families, while other situation can occur where people don't seek medical treatment due to high cost.

This research study analyzes the National Health Interview Survey (NHIS) 2020 data set, which includes information about the health insurance status of everyone living in the United States. Using this data, we focused on finding the most critical variables that determine whether one would have health insurance or not. The key variables used in the data set are age, gender, marital status, race, education, household income. We started our research by exploring and analyzing the data to understand how the data is distributed based on the different variables. This part of the research includes using descriptive statistics in R using SJ-Plots.

After analysis, we did find out some unexpected results. For example, we speculated that

unemployed individuals would be less likely to have health insurance than employed individuals. But the results were the opposite. But to get more accurate outcomes, we performed the feature selection method using Boruta to exclude non-significant variables from the dataset. This process helped us run a logistic regression using only the significant variables selected in the feature engineering process. After running the regression, we found that out of the 29 variables chosen in the feature engineering process, only seven of them were significant in predicting whether one would have health insurance.

## **Literature Review**

In recent years, health insurance has become a necessity of life. Its importance is gaining momentum with the government taking initiatives to promote health insurance and improve health conditions (Sudha & Murugesan, 2021). Health insurance is an individual's agreement with an insurer on their behalf or on behalf of others to cover medical expenses. This cover may include some or the entire medical bill incurred. It helps keep individuals from paying for medical bills they cannot afford out of pocket. It, therefore, acts as a safety net in case of medical emergencies.

Insurance is beneficial because some illnesses and accidents are unexpected, and not being able to afford health care costs can deter one from seeking medical health. Ultimately this may lead to a decrease in their health status and eventually cause death. It has been incredibly beneficial to deal with medical inflation and ensure quality healthcare. Mitrovic & Pesic (2019) argue that insurance is an economic activity significant to the individual and the economy in general. Health insurance may also provide tax benefits and ultimately deliver financial stability. In some countries, health insurance is compulsory for everyone with equitable benefits.

However, there are times when people face discrimination based on race and ethnicity. Society identifies and treats people differently based on social markers such as race. Treadwell (2019) argues that a significant difference in healthcare provision based on race and gender exists. Historically, minority races such as blacks and Hispanics faced unfair treatment compared to whites, a dominant race, especially in the United States. This differentiation based on race has also affected health care provision. Liu et al. (2018) argued that promoting equity in the economic accessibility of health care services to all citizens ensures healthcare justice.

There are various socioeconomic factors associated with health insurance. Some of the benefits of health insurance are that it increases the financial stability of a family or business, allows competitiveness and trade development, and improves the health status of individuals in a country. Improving individuals' health status ensures that the country's economy is stable since people need to be strong enough to work. Health insurance also promotes the country's health sector since medical bills are covered and minimal debts.

However, there are negative socioeconomic factors affecting health insurance. For example, the United States has the highest disparities in health care access based on factors such as employment status, homeownership status, and education level (Griffith et al., 2017). In addition, these factors are often against the minority groups based on social class, gender, race, and economic status. Therefore, this research analyzes the socioeconomic factors affecting health insurance and aims to provide a solution that will help governments and non-for-profit organizations find ways to help our community.

The history of health insurance is significant in understanding the progress made so far. According to Fox and Kongstvedt (2013), health insurance is an invention of the 20th century. In the late 19th century, only a few insurers offered health insurance, but they would only cover accidents that occurred in the workplace. These insurance policies would, later on, evolve to cover non-work-related accidents. The great depression of the 1930s led hospitals to implement other payment forms for medical bills. In the United States, Medicare was introduced in 1965 to cover senior citizens. In the 1980s, there was the development of managed care due to increasing

healthcare costs. Private companies would slowly emerge, offering better policies for people than the government. This participation of private companies led to an increase in the number of applications for health insurance coverage. In 2010, the health care act required most Americans to get basic health insurance coverage.

Health insurance gained momentum and consequently became politicized. Many argue that universal health insurance failed in the first half of the twentieth century because of political influence by insurers, unions, employers, and physicians (Murray, 2007). Stone (1993) argued that the politics of American health insurance is a struggle over the solidarity principle or the logic of actuarial fairness. With actuarial fairness, a sense of responsibility fosters. But it creates differences rather than interdependence and commonalities.

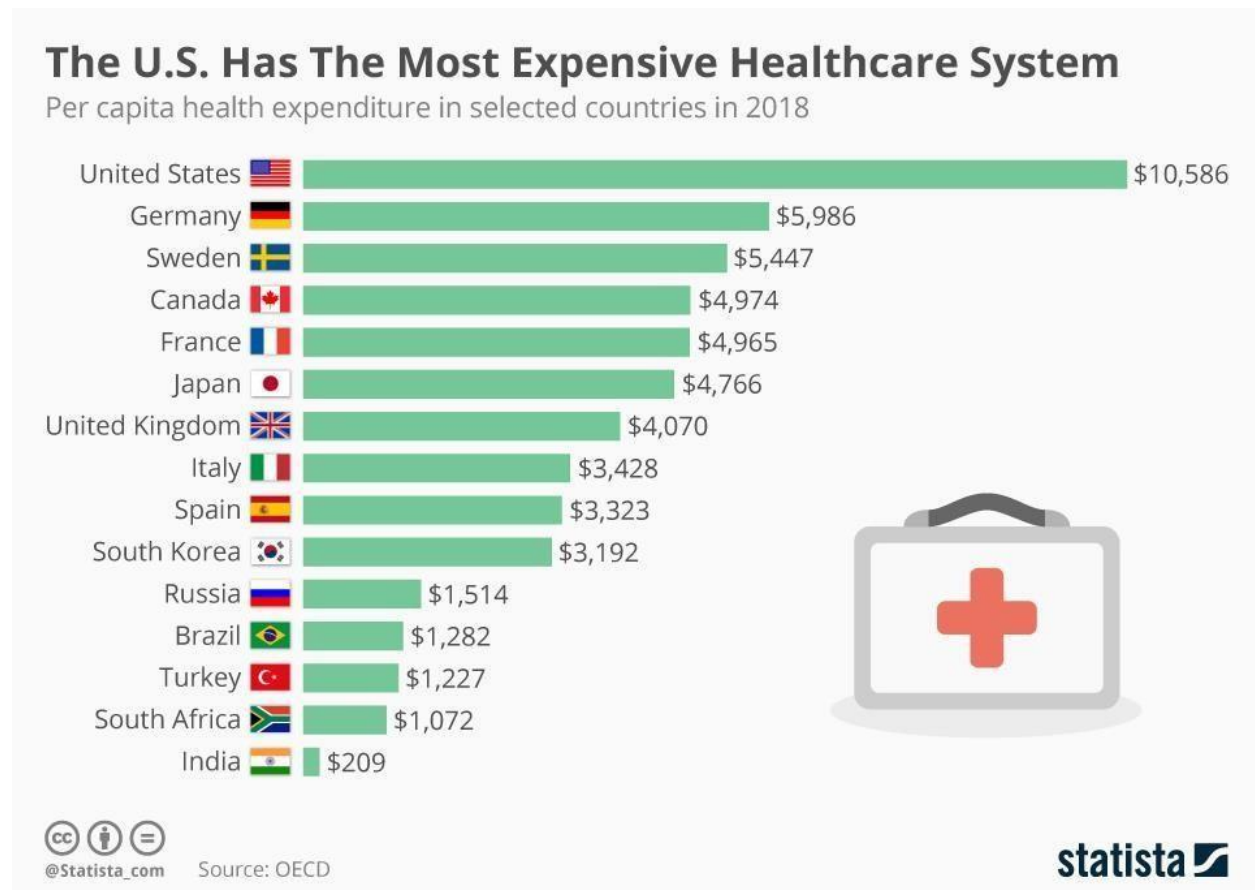
Furthermore, the concept of actuarial fairness creates social divisions based on socioeconomic status. Despite significant changes in the political context in recent years, the logic of actuarial fairness is still deeply embedded in the community and healthcare market today. As a result, the indifference would affect other social aspects such as race, gender, and economic status. Therefore, health insurance politicization favors those of higher socioeconomic status.

The people with higher social status are the ones who can afford health care, yet we try to serve those who need it the least and fail to serve those who need it the most. In 2019, Senator Bernie Sanders criticized the American healthcare system. He mentioned: "the incredible corruption and greed" of the pharmaceutical industry. The graph below shows how much more the United States spends on healthcare than other countries. Still, according to [US News](#), the United States is not even included in the top 10 healthcare systems globally and is way behind the countries that spend less on healthcare. The problem here is that why cannot have the best



healthcare system in the world when we spend the most money compared to other countries. For example, Sweden was ranked #1 by US News, and they spend almost half as much money as we do.

### *Cost & Quality of Healthcare Graph*



Statista. (2019b, August 8). *The U.S. Has the Most Expensive Healthcare System in the World* [Graph]. Statista Infographics. <https://www.statista.com/chart/8658/health-spending-per-capita/>

Now we will dig deeper into the topic and analyze the reasoning behind this problem in the healthcare system. Margot Sanger-Katz briefly explains the condition of the healthcare system in her article: "Why Is U.S. Health Care So Expensive?". This article reveals: why our healthcare is expensive, the areas we spend the most money on, and breaks some myths. Margot

currently writes articles for the New York Times and holds a master's degree in Journalism from Columbia University. This article compares the cost and quality of healthcare in the United States. According to Margot, the U.S. healthcare system is not the best and the worst if we compare it with other countries.

Nonetheless, it is average considering the amount of money we spend. We think our healthcare is lagging because of the relatively low life expectancy. However, we do not realize that it might not be just the fault of the healthcare system itself that has to lead to a shorter life span. Other than the healthcare system, another major cause of the shorter life span is our diet. According to CDC, "The U.S. [obesity prevalence](#) was 42.4% in 2017 – 2018, and it increased by 11.9% since 1999-2000" (*Obesity Is a Common, Serious, and Costly Disease*, 2021).

Therefore, we cannot just blame the healthcare system for not providing optimal results and thus a decreased life expectancy. But still, some flaws also exist in the healthcare system, and there are areas where we spend more than our peers. This article also provides research-based evidence using Dr. Ashish Jha, an Indian-American physician and the Dean of the Brown University School of Public Health. Research reveals that we spend comparatively more money on medical services, which includes the cost of hospitalization, doctors' visits, and prescription drugs.

Along with low-quality healthcare, if Americans do not have health insurance for minorities, it creates an even bigger problem. Firstly, the minorities would not be able to get healthcare, and even if they do, the quality of it will not be high enough. But still, in the end, getting some form of health security is more important than nothing. So instead of focusing on quality, we thought focusing on inclusivity is more important considering the current healthcare

situation of the United States. Therefore, we further analyzed the minorities in the US to find out better ways to serve them and provide them with equal health insurance facilities.

Racial discrimination is rampant in American society. It permeates every aspect of our community in which it exists. It exists in economic accomplishment, including wages, income, credit extension, prices paid, among others. It also exists in health care provision with different health measures used on different races with no medical explanation (Harris et al., 2006). In addition, marginalized groups live in undesirable residential areas, which affect their access to healthcare. Those environments have health-damaging conditions and consequently limit the socio-economic status. Discrimination can also affect health through access to good healthcare services, creating exposure to stressors such as financial strain and unemployment. These differences are also seen in health insurance practices, as well as seen below:

### **Racial/ Ethnic Differences in Health Insurance**

Health insurance allows expanded access to affordable healthcare to families. Therefore, it is imperative to make health insurance accessible to as many people as possible. However, pre-existing societal inequalities and disparities are present in the market for health insurance therefore racial minorities face significant barriers to access health insurance. In Artiga et al, (2021) a study that investigates differences in health insurance coverage grouped by race found that the Affordable Care Act(ACA) expanded access to healthcare for minorities at a higher rate than the majority racial population, but despite the improvements ushered in with the ACA, racial minorities still significantly lag Caucasians in health insurance coverage

#### **Immigration & health insurance**

In “Healthcare inequality issues among immigrant elders after neoliberal welfare reform: empirical findings from the United States“, the author Younsook Yeo studies the effect of welfare policy reforms on healthcare usage of immigrant elders in the United States. The specific reform that the

paper studies is the welfare reform in 1996 that restricted access of public benefits to new immigrants unless they lived in a state that funded and provided those benefits, accrued 10 years of work history in the US, or became naturalized citizens of the United States. The immigrant elders are compared to native elders to study the effect of the reforms on their healthcare usage.

The paper uses the National Health Information Survey(NHIS) data, a cross sectional survey representative of the US population. The NHIS survey is the primary source for information on health, healthcare access, and health behaviors for the non-institutionalized civilian population that resides in the United States. To represent the pre-reform period, the paper uses the samples from 1993-1996 and uses the samples from 2002-2013 to represent the post-reform period. The data was analyzed using DD analysis which calculated the difference between the control and the experimental groups, the native elders and the immigrant elders to the differences in pre and post-reform period. Additionally, the author used multivariate analyses that compared significant factors associated with healthcare use during both welfare periods.

The author found that while inequalities in healthcare usage predated the welfare reforms of 1996, the inequalities grew significantly as a result of the reforms. The enabling factors and the social structural factors explaining the healthcare usage of elder immigrants in the post-reform period lead the author to suggest that welfare reform may be the motivating cause of these inequalities in healthcare.

### **Immigration & Healthcare under Trump's Administration**

In the United States, people of color such as black, Hispanic and Asians are more likely to live without having access to health care and are more vulnerable. In 2009 Affordable Care Act tackled this issue where people of color were able to have access to health care as well as undocumented immigrants. During the Trump administration, he wanted to go against what the Affordable Care Act was meant to do. According to the article "Immigration and Health Care Under the Trump Administration" by Wendy E. Parmet states, "It is therefore not surprising that the first year of the

Trump administration, which has focused its domestic agenda on restricting immigration and repealing the ACA, has proven especially perilous for immigrants who need health care”. Repealing the ACA can impact many insured Americans who were not able to obtain healthcare before the ACA. In 2015, 42% of undocumented immigrants were living without health insurance relatively to 11% of U.S. citizens didn’t have health insurance. It has always been a challenge for immigrants to get access to proper health insurance. For example, Personal Responsibility and Work Opportunity Reconciliation Act, which was introduced in 1996, prohibited immigrants who were undocumented for getting any sort of programs that were federally funded such as Medicare and Medicaid.

## **Effect Of Covid On Minorities Employment In Early**

### **Days Of The Pandemic**

Couch et al. (2020) studied the impact of COVID-19 on minority unemployment in the early days of the pandemic. The study used data from the Current Population Survey from April-June 2020. The unique challenge of the economic crises brought on by the pandemic is argued to be nuanced in relation to previous crises because of the importance of government mandated closures causing widespread unemployment. The author argues that this is the reason why the COVID-19 crises disproportionately affected Hispanic workers the most, unlike previous crises. The excess distribution of Hispanic workers in the largely informal sector made them the most prone to losing employment. The study also found that unemployment amongst African American workers rose less when compared to previous crises. As the pandemic progressed and job losses were recovered in May and June, the study found that these gains largely went to White workers as unemployment among Hispanic workers remained depressed when the initial recovery began.

## **COVID-19 Health Impact On Different Races**

Gemelas et al., 2021 concluded that structural racism had exacerbated the impact of COVID-19 on

people of color through differential employment loss. The study aimed to investigate whether there had been greater employment loss in employment in people of color from frontline and non-frontline jobs during the pandemic.

The study used the Current Population Survey provided jointly by the Bureau of Labor Statistics and the U.S. Census Bureau. The CPS is the primary source of data on the labor force population in the United States. The study grouped by race and ethnicity, tabulating the change in the percentage employed quarterly and checked status for if the worker was a frontline worker. The author found that the largest declines in employment by race were among Black, Asian American, and Hispanic groups. These declines were even larger when arranged by sector comparing non-frontline groups to the other groups.

## **Longer Term Effect Of Covid On Racial Employment**

Borjas and Cassidy(2020) studied the adverse effect of the COVID-19 pandemic on immigrant employment in the United States. The paper relied on the Current Population Survey Basic Monthly Files to substantiate the disproportionate losses to the immigrants in the pandemic labor market. In previous contracting labor markets, immigrants were likelier than native men to be employed. However, this was not true of COVID-19. The author found that part of the reason for this increase was because immigrants were less likely to work jobs that allowed you to work from home. The worst effects of the pandemic labor market were felt by undocumented immigrant men with their rate of job loss outpaced that of legal immigrants too. The effect of this was likely further exacerbated by the challenges to access public benefits for undocumented people.

Tahmasbi et al. (2021) studies the effect of COVID-19 on Sinophobia. Due to the belief that COVID-19 originated in China. They studied two large datasets consisting of posts from Twitter and 4chan and found that the onset of the pandemic and its origin story had caused an increase in Sinophobic content on the internet. However, the discrimination faced by the community was not restricted to the internet. Due to the belief that Chinese people were primary

spreaders of the diseases, Chinese business districts were boycotted, Chinese businesses shut down in large swaths, and there remained overwhelming fear and resentment towards the community through the course of the pandemic.

## **Impact of Covid-19 on Gender Equality**

The paper studies the change in the continuation of health insurance, access to formal medical care, and medication adherence during periods of unemployment. The author looks into an observed behavior termed “intensive mothering” where women sacrifice their own healthcare for the healthcare needs of their family, discouraging them from spending time on their own healthcare. The paper also argues that men are incentivized to provide income and health insurance due to their predominant status as breadwinners in traditional family structures.

The data used in the paper consists of 100 in-depth interviews conducted with unemployed individuals, both male and female from 2013-2015. The interviews consisted of 84 individuals from unemployment centers in rural and urban Pennsylvania. The sample also included 16 interviews from the pilot study that used snowball sampling. The interviews were coded using a process that borrowed from inductive and deductive theoretical traditions.

The paper found that upon losing employment women were more likely than men to stop seeking healthcare that they had previously maintained for themselves. The loss of employment however caused more men to forego obligations to provide their family with health insurance. There is a contrast drawn that loss of employment causes women to forgo their own healthcare needs, while the same results in men forgoing the responsibility to provide their family with healthcare. Gendered frames in family obligations remain influential in caregiving. A unique toll is placed on women’s health by the intersection of economic inequalities and changing gender norms with respect to healthcare duties in modern society.

The article by **Alon Titan published by the National Bureau of economic research** studies the difference in the economic downturn caused by Covid-19, to other economic

downtowns. Contrasting that while economic downturns usually disproportionately affect the employment of men more, this was not the case in 2020. The economic downturn caused by the Covid-19 pandemic affected industries where women were employed at higher rates, therefore causing larger unemployment losses in the demographic. The closing of childcare facilities and schools increased childcare responsibilities for women, especially working mothers. The authors expect this to continue due to high returns for experience in the labor market.

Despite the pessimistic case for gender equality in the short term. The paper is optimistic about the disruption promoting gender equality in the long run, eroding social norms requiring more fathers to take up primary childcare and providing flexible work schedules for mothers to better manage professional and personal responsibilities.

Also, during the Covid-19 pandemic, many Americans who didn't have access to the healthcare were impacted the most. More than 114 million people were impacted by the covid-19 pandemic causing more than 500,000 deaths. Most of covid-19 related cases impacted minorities and people of color such as Hispanics, Blacks, and Native Americans. Systemic racism also plays a role in creating disparities among people from receiving proper access to healthcare. Throughout the pandemic, Hispanic people had the most amount of covid-19 cases followed by black and then Natives. According to the article by Samantha Artiga "COVID-19 Cases and Deaths by Race/Ethnicity: Current Data and Changes Over Time" it states, "Between Spring 2020 and the early part of Summer 2020, AIAN, Hispanic, and Black people had higher death rates compared to White and Asian people, with a particularly high death rate among AIAN people. Death rates fell over the course of Summer 2020 and disparities narrowed, although death rates for AIAN people remained higher compared to other groups. Deaths peaked in December 2020 across groups, with the highest death rates among AIAN and Hispanic people". Without proper health coverage, they were not able to receive proper treatment and were impacted the most by the pandemic.



## Effect On Immigrants/undocumented farm workers

Having access to healthcare for many immigrants and undocumented people in the United States can be challenging. Many immigrants in the United States were working in the agricultural sector. Working in the agricultural sector can be very physically demanding as well as working conditions can be rough. According to the article “Health care service utilization of documented and undocumented hired farmworkers in the U.S” by Tianyuan Luo and Cesar L. Escalante it states, “several studies have pointed out that the strenuous, rigorous nature of farm work and its greater physical demands, prevailing working conditions on farms, and the workers’ lack of health knowledge and information could have significant adverse effects on the health of hired farmworkers in the U.S.” (Pg. 923). Many small percent of farm workers in the United States were covered by health insurance. For example, in California, 50% of immigrants working on farms were covered with health insurance. One key argument that the author provides is that, if farmers across the United States are not fully covered with health coverage, it can cause more \$374 billion in damages. Compared to documented farmer workers and undocumented, undocumented farmers more uninsured compared with documented. Reasons could occur such as not able to speak English, low income, and their reading proficiency. Policies like Personal Responsibility and Work Opportunity Reconciliation Act of 1996 aimed at limiting the health coverage that is provided to immigrants’ farmworkers.

## Lack of Health Insurance from Employers

Lack of health coverage can occur due to employment and income. U.S. Census Bureau reports estimate of 42 million people lives below poverty level. In the article “*Racial Disparities in Health Status and Access to Healthcare: The continuation of inequality in the United States Due to Structural Racism*” it states, “In 2011, the U.S. Census Bureau reported that 42.7 million people, about 14.3 percent of the U.S. population, were below the poverty level. The poverty rate for African Americans was 25.8 percent compared to 11.6 percent for Caucasians” (pg. 1122). This shows that people are color are at more risk of

being uninsured. Many African Americans also have high unemployment rate compared to Caucasians, which can lead to

disparity among other minority from getting access to proper health insurance. Unequal pay also impact can create racial health disparities. “The government has failed to enforce Title VII of the Civil Rights Act of 1964 and the Equal Act of 1963 by allowing the continuation of employee referral programs that have disparate impact on minorities but not protecting workers who bring racial discrimination cases, and not collecting data related to discrimination in hiring and pay” (Pg.1128). This shows that when employers discriminate against people of color, such as African Americans, it prevents them from getting access to healthcare and also not getting a job that will provide them with health benefits. In 2009, 55% of minorities didn’t had any sort of health coverage, in which 32% were Latinos, 21% were African American, 28% were Native Americans and only 13% were Caucasians. “As a result of their lack employer-sponsored healthcare insurance and their poverty, these minority families are disproportionately unable to afford to pay for healthcare” (Pg. 1131). Not having access to healthcare can be costly and undertreatment can also result in death.

Inequality in health insurance can also happened because of government tax policies. It happens because health insurance that employer provide are not subject to income. In the article “Health Insurance and Income Inequality” by Robert kaestner and Darren lubotsky it states, “Government tax policy for health care affects inequality because employer-provided health insurance is generally not taxed as individual’s income at the federal or state level, or through the payroll tax” (Pg. 59). Most of health insurance that is provided by the employers are mainly beneficial to families with high income.

## **Impact of Affordable Care Act**

People in United States faced many barriers that prevented them from obtaining proper access to health care. One of the policies that was created to help Americans to access health insurance was the Affordable Care Act. This policy helped all Americans access to healthcare that they can easily afford. People in United States struggled to access proper health insurance due to inequality and disparities. Affordable Care Act successfully lowered the uninsured rates of Americans and reduced the coverage gap between people of color. Jesse C. Baumgartner (2020) states in an article from The Commonwealth Fund

that “By 2017-2019, Asian Americans had the lowest uninsured rate of any racial or ethnic group in the U.S. Uninsured rates have fallen among all Asian American subgroups since the passage of the ACA, but not uniformly”. Affordable Care Act helped people in United States to receive health coverage that they were unable to obtain before. Before Affordable Care Act was introduced working-age people of color in United States were more uninsured compared to whites. Factors of disparities among health coverage in united states can happen because of race or their ethnicity. According to “The Impact of the ACA Medicaid Expansions on Health Insurance Coverage through 2015 and Coverage Disparities by Age, Race/Ethnicity, and Gender” by George L. Wehby, it states, “Extensive disparities in coverage have also historically existed by race/ethnicity, especially among young adults with markedly higher pre-ACA uninsured rate overall among Hispanics (31%) and non- Hispanic blacks (21%) compared to whites (12%) (United States Census Bureau 2010)” (Pg. 1250). After Affordable Care Act was implemented, it significantly reduces the amount of uninsured people in the United States.

## Results and Findings

From the study, we found out the health insurance status of individuals based on their race: Asian race had 94.5% likelihood of having health insurance, followed by the white race 93.8%, then the black race with a 90.6%, then American Indian with 85.5%, and finally Aleut Alaskan at 85.5%. Looking at the influence of race on health insurance by observing the chi-square statistic and its p-value, the chi-square test was significant at 5% level; the test shows that an individual's race was highly associated with health insurance. This data set was interesting to explore as we were speculating that the white race would possibly have the highest percentage of individuals with health insurance. Still, it turns out that the Asian race is more likely to have health insurance than the white race.

The next factor influencing health insurance was marital status. Widows had the highest number of individuals with health insurance (97.7%) and then followed by the married people who had 94.8% health insurance, and thirdly 91.7% of people who were divorced had health insurance. Separated and unmarried individuals were the least likely to have health insurance. Therefore, marital status was a significant factor influencing healthcare status. Another factor that influenced health insurance was citizenship; citizens had higher chances of having health insurance (94%) than non-citizens (69.9%).

Next, we looked at the effects of sexual orientation in determining the health insurance status of an individual. Our research found out that lesbians and gays had a higher probability of having health insurance (93.4%) than straight people (92.5%), followed by bisexuals having a rate of 88.7%, which is the lowest percentage among the three sexual orientations. Sexual orientation also proved to be a significant factor influencing health insurance.

Whether someone served in the army was also an essential factor influencing individuals' health insurance coverage. Those who have served had a higher rate (97.8%) of health insurance than those who have never joined the military (91.9%). Another factor that significantly influenced the health insurance status was the number of years someone has stayed in America. The findings were that people who remained in America for 15 or more years had the highest rate of health insurance (87.2%), those who stayed between 5-10 years in the US had (84.2%), the group of people with the least number of people with health insurance was the one who lived in the US for less than one year (64.7%). The more years someone spends in America, the higher their chance to get health insurance.

Next, we looked at the effects of sexual orientation in determining the health insurance status of an individual. Our research found out that lesbians and gays had a higher probability of having health insurance (93.4%) than straight people (92.5%), followed by bisexuals having a rate of 88.7%, which is the lowest percentage among the three sexual orientations. Sexual orientation also proved to be a significant factor influencing health insurance.

Whether someone served in the army was also an essential factor influencing individuals' health insurance coverage. Those who have served had a higher rate (97.8%) of health insurance than those who have never joined the military (91.9%). Another factor that significantly influenced the health insurance status was the number of years someone has stayed in America. The findings were that people who remained in America for 15 or more years had the highest rate of health insurance (87.2%), those who stayed between 5-10 years in the US had (84.2%), the group of people with the least number health insurance was the one who lived in the US for less than one

year (64.7%). The more years one has lived in the United States, the higher are their chances of having health insurance.

Education status was also a significant factor impacting the health insurance status of individuals. The people with higher education had a higher chance of having health insurance. According to education level, people with some college education had a 91.4% rate of having health insurance, which is the highest rate compared to people with no education or high school diploma.

However, when we compared the employment status factor and its relationship with health insurance, we found out that unemployed people had a higher rate of health insurance (94.3%) than employed individuals who had a rate of (91.1%). This output was unexpected. We expected opposite results because it would make more sense for individuals with a stable income to have health insurance and not the other way around. According to researchers from the past Healthcare, inequality is the new income inequality. According to Dan Mangan, a CNBC reporter, "the rich in the United States — despite being healthier on average than the poor — have become the biggest buyers of health care, a dramatic shift in spending patterns across income groups, according to a new Harvard study" (Mangan, 2016).

Moreover, in 2001 research, the U.S. government agency named "Agency for Healthcare Research and Quality" published a research article Health Insurance Coverage and Income Levels for the U.S. noninstitutionalized population under the age of 65. According to this research, "Approximately one-fourth of persons in poor near-poor, low-income families were not covered by health insurance at any time during the year" (Crimmel, 2004). Therefore, this research

illustrates that in 2021 minorities will be more likely to have health insurance than in 2001. This increase in insured individuals has been an enormous achievement in the past 20 years.

But before we conclude our results, we also need to consider other possible variables that could have led employed individuals to be less likely to have insurance than the unemployed. For example, maybe unemployed people have high insecurity of not being able to pay expensive medical bills and thus have started to take advantage of health insurance services. In contrast, the employed individuals are less insecure as they are more likely able to pay medical bills.

Since other factors could have also influenced the health insurance coverage, we decided to explore more variables. The next factor we explored was the health insurance status of a person based on whether a person was born in the U.S. or not. The people born in the U.S. had a higher health insurance rate (94%) than those born outside America (84.9%).

Moving on to the following variable, we concentrated on understanding the health insurance rate based on whether the employer offered health insurance or not. From this analysis, we learned that those whose employers provided health insurance had a higher rate of health insurance coverage (96.4%) than those whose employers did not offer health insurance had a rate of 77.6%.

After analyzing the data and comparing the different variables found in the data set, we moved on to creating a logistic regression model to understand further the relationship between the various factors mentioned and the health insurance status of people living in the United States of America. This model aims to analyze the odds of someone having health insurance, given that person's data, including various variable factors.

## **The variables that stood out to be significant in the model were:**

- Sample person weight.
- Citizenship status.
- Private health insurance.
- Health insurance coverage by children's program.
- Health insurance coverage by other state-sponsored plans.

These variables were the only ones with p-values less than 0.05 (level of statistical significance). The odds ratios of sample weight were 1; hence that's not impactful as it means people with more weight are as likely to be insured as to any other people. But looking at the citizenship status, the odds of one having health insurance was 0.13 times more than one who is not a citizen. People covered by the children's program were 1.42 times more likely to have health insurance. Those covered by state-sponsored programs were 1.21 times more likely to have insurance.

## **Descriptive statistics**

The following figures show the summary statistics of the data. The tables and the plots show the frequency analysis of the variables grouped by the dependent variable. In addition, the tables test the association between each variable and the dependent variable (HINOTCOVE) Health insurance coverage.

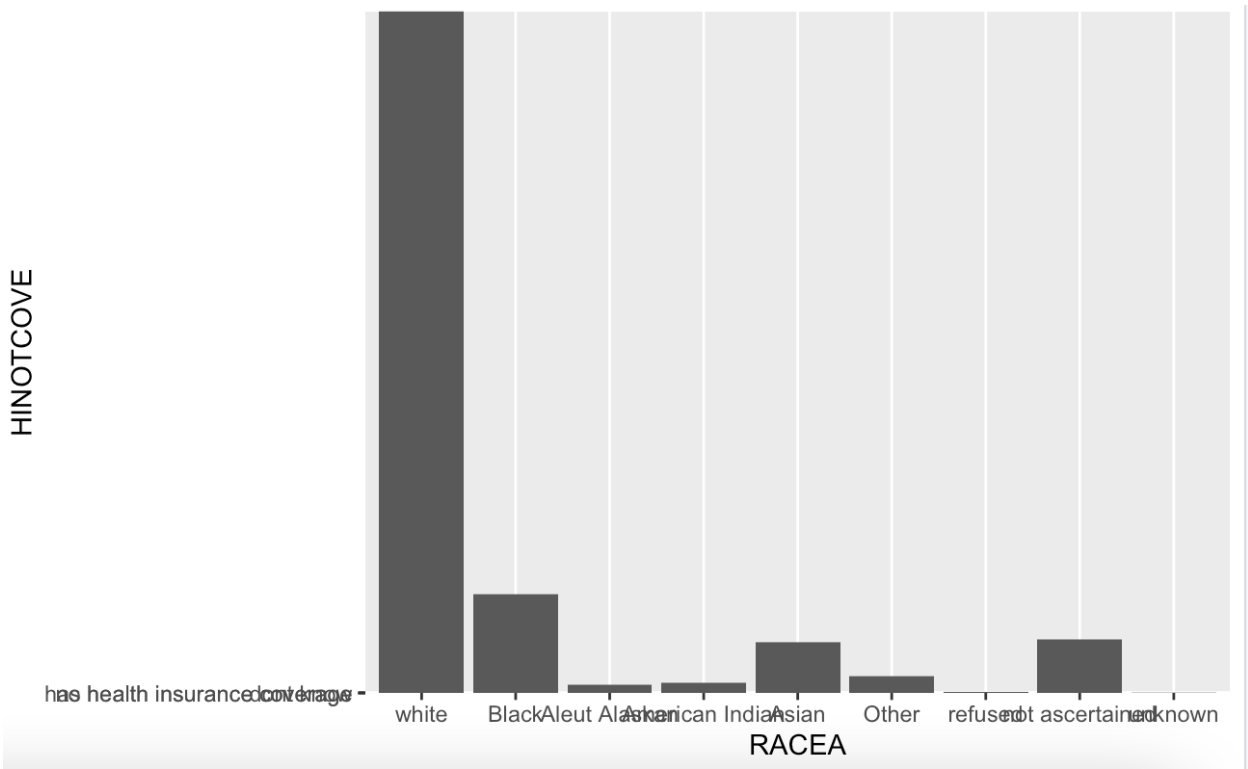


**Table 1: Health insurance by race**

<b>Table Health insurance by RACE</b>				
<i>RACEA</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
white	22366 93.8 %	1445 6.1 %	36 0.2 %	23847 100 %
Black	2848 90.6 %	284 9 %	10 0.3 %	3142 100 %
Aleut Alaskan	186 80.9 %	44 19.1 %	0 0 %	230 100 %
American Indian	247 85.5 %	42 14.5 %	0 0 %	289 100 %
Asian	1599 94.5 %	92 5.4 %	1 0.1 %	1692 100 %
Other	425 93.8 %	28 6.2 %	0 0 %	453 100 %
refused	12 80 %	3 20 %	0 0 %	15 100 %
not ascertained	1180 78.7 %	315 21 %	4 0.3 %	1499 100 %
unknown	8 88.9 %	1 11.1 %	0 0 %	9 100 %
<b>Total</b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$\chi^2=582.200 \cdot df=16 \cdot \text{Cramer's } V=0.097 \cdot \text{Fisher's } p=0.000$

***Figure 1: Health insurance by race***

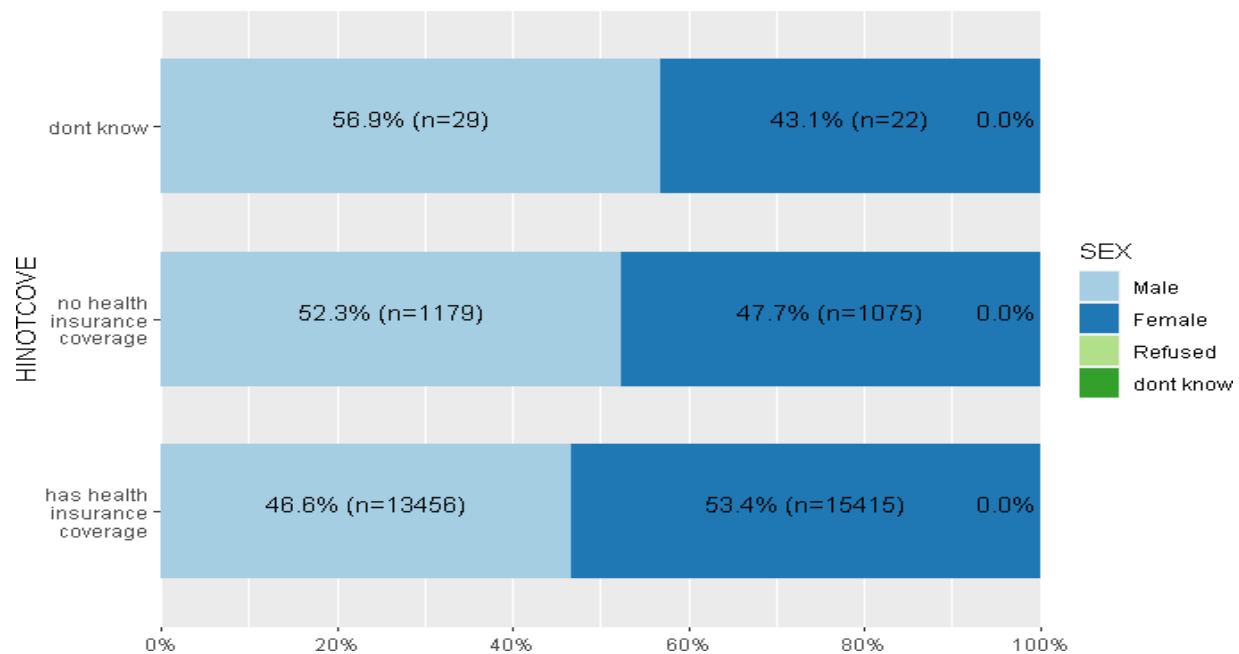


The Asian race had the highest percentage of people with health insurance at 94.5% ; the American Indian had the lowest percentage at 85.5%.

**Table 2: Health insurance by sex****Table Health insuracne by sex**

<i>SEX</i>	<i>HINOTCOVE</i>			<b><i>Total</i></b>
	has health insurance coverage	no health insurance coverage	dont know	
Male	13456 91.8 %	1179 8 %	29 0.2 %	14664 100 %
Female	15415 93.4 %	1075 6.5 %	22 0.1 %	16512 100 %
Refused	0 0 %	0 0 %	0 0 %	0 100 %
dont know	0 0 %	0 0 %	0 0 %	0 100 %
<b><i>Total</i></b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2 = \text{NaN} \cdot df = 6 \cdot \text{Cramer's } V = \text{NaN} \cdot \text{Fisher's } p = 0.000$$

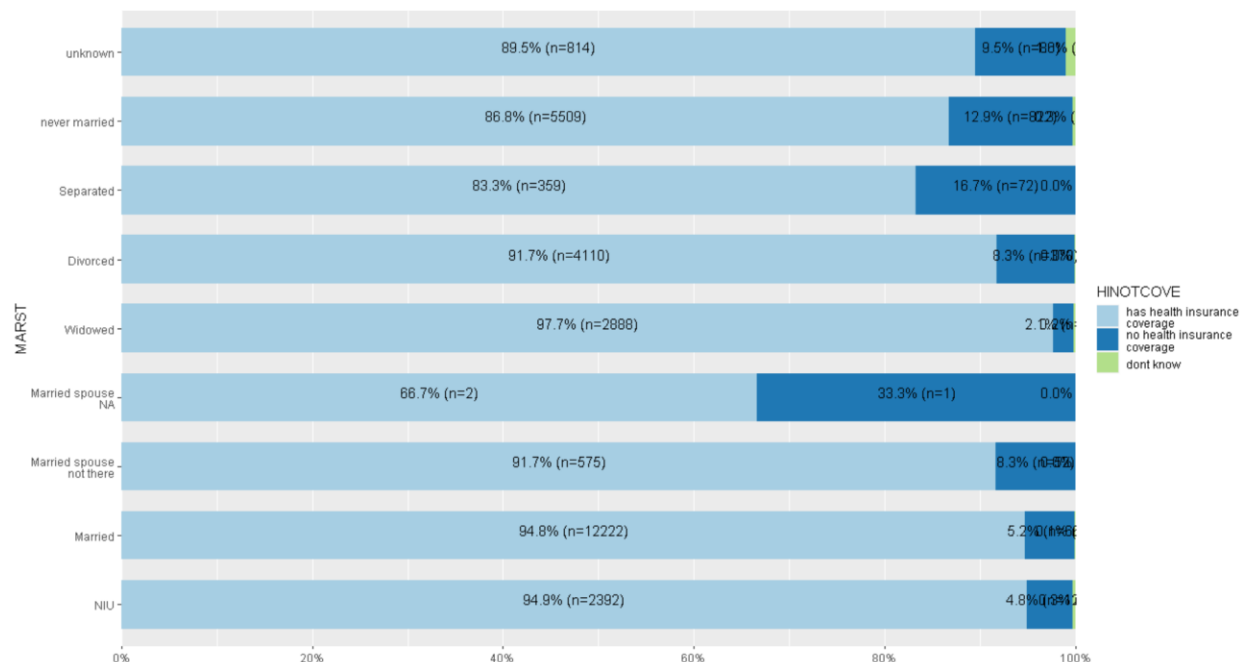
**Figure 2: plot of health insurance by sex**

Females had a higher likelihood of having health insurance with a 53.4% in comparison to males with a 46.6%.

**Table 3: Health insurance by marital status**

<i>MARST</i>	<i>HINOTCOVE</i>			<b><i>Total</i></b>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
Married	12222 94.8 %	667 5.2 %	10 0.1 %	12899 100 %
Married spouse not there	575 91.7 %	52 8.3 %	0 0 %	627 100 %
Married spouse NA	2 66.7 %	1 33.3 %	0 0 %	3 100 %
Widowed	2888 97.7 %	62 2.1 %	5 0.2 %	2955 100 %
Divorced	4110 91.7 %	370 8.3 %	1 0 %	4481 100 %
Separated	359 83.3 %	72 16.7 %	0 0 %	431 100 %
never married	5509 86.8 %	822 12.9 %	19 0.3 %	6350 100 %
unknown	814 89.5 %	86 9.5 %	9 1 %	909 100 %
<b><i>Total</i></b>	<b>28871 92.6 %</b>	<b>2254 7.2 %</b>	<b>51 0.2 %</b>	<b>31176 100 %</b>

$\chi^2=665.620 \cdot df=16 \cdot \text{Cramer's } V=0.103 \cdot \text{Fisher's } p=0.000$

**Figure 3: plot of health insurance by marital status**

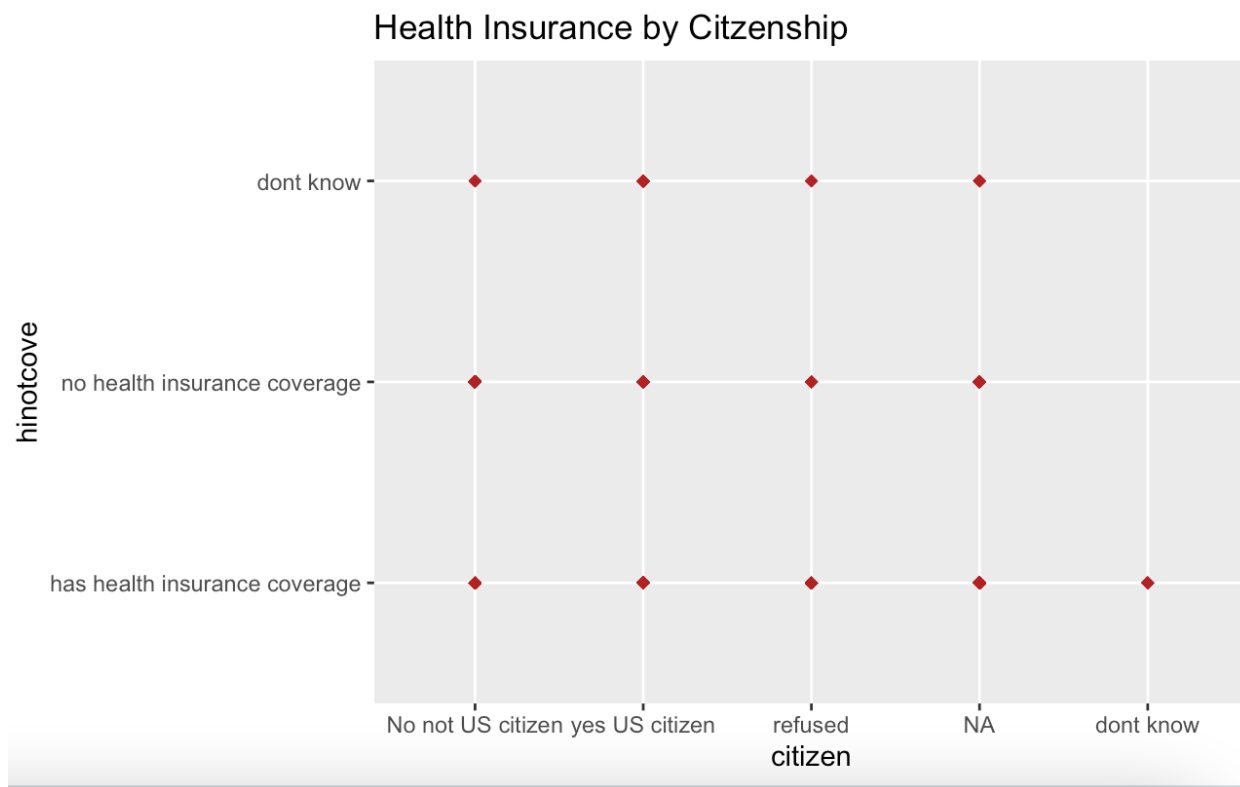
Widows had the highest percentage of health insurance (97.7%), then followed by married people at 94.8%, the lowest percentage was by married spouse NA(66.7%) and the separated (83.3%).

**Table 4: Health insurance by citizenship**

**Table Health insurance by citizenship status**

<i>CITIZEN</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
No not US citizen	1147 69.9 %	490 29.9 %	4 0.2 %	1641 100 %
yes US citizen	27150 94 %	1697 5.9 %	40 0.1 %	28887 100 %
refused	44 81.5 %	9 16.7 %	1 1.9 %	54 100 %
NA	519 89 %	58 9.9 %	6 1 %	583 100 %
dont know	11 100 %	0 0 %	0 0 %	11 100 %
<b><i>Total</i></b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=1385.976 \cdot df=8 \cdot \text{Cramer's } V=0.149 \cdot \text{Fisher's } p=0.000$$

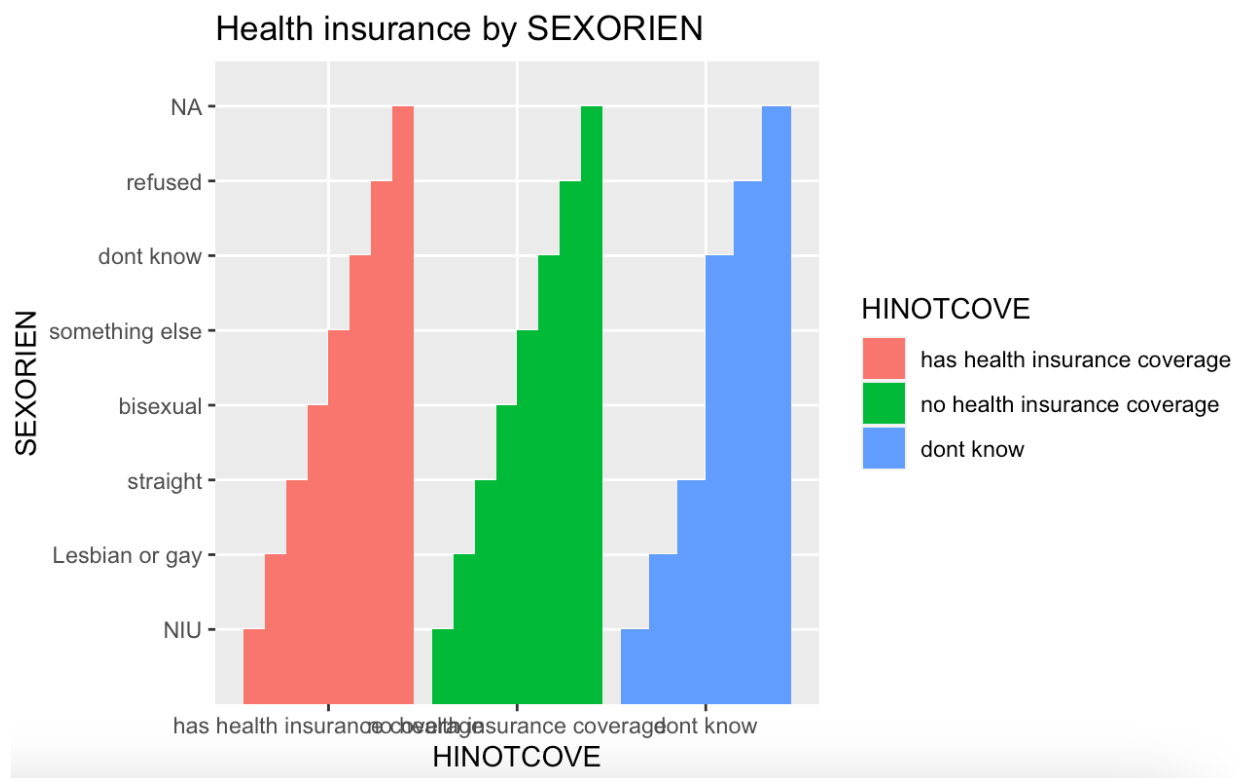
**Figure 4: Health insurance by citizenship**

US citizens had the highest percentage of people with health insurance (94%), non-US citizens had the lowest rate of health insurance (69.9%).

**Table 5: Health insurance by sex orientation****Table Health insurance by sex orientation**

<i>SEXORIEN</i>	<i>HINOTCOVE</i>			<b><i>Total</i></b>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
Lesbian or gay	480 93.4 %	33 6.4 %	1 0.2 %	514 100 %
straight	24727 92.5 %	1967 7.4 %	35 0.1 %	26729 100 %
bisexual	320 88.4 %	42 11.6 %	0 0 %	362 100 %
something else	110 89.4 %	13 10.6 %	0 0 %	123 100 %
dont know	202 92.7 %	16 7.3 %	0 0 %	218 100 %
refused	152 94.4 %	7 4.3 %	2 1.2 %	161 100 %
NA	488 89.1 %	54 9.9 %	6 1.1 %	548 100 %
<b><i>Total</i></b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=88.085 \cdot df=14 \cdot \text{Cramer's } V=0.038 \cdot \text{Fisher's } p=0.000$$

**Figure 5: Health insurance by sex orientation**

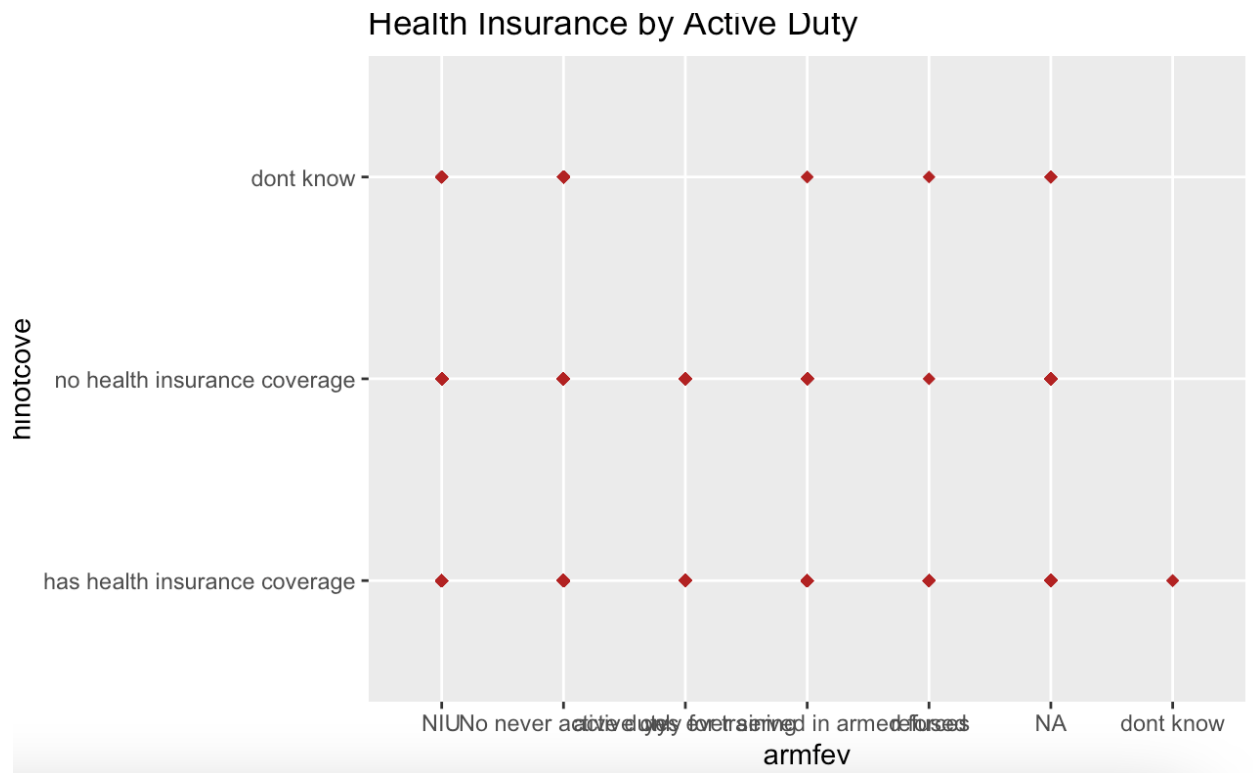
The above plot shows people who refused to inform about their sexual orientation had the highest percentage of insurance coverage (94.4%), and the lowest percent (88.4%) was that of bisexuals.



**Table 6: Health insurance by army service record****Table Health insurance by army service**

<i>ARMFEV</i>	<i>HINOTCOVE</i>			<b><i>Total</i></b>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
No never active duty	22842 91.9 %	1979 8 %	37 0.1 %	24858 100 %
active only for training	489 97 %	15 3 %	0 0 %	504 100 %
yes ever served in armed forces	2410 97.8 %	53 2.2 %	1 0 %	2464 100 %
refused	15 93.8 %	1 6.2 %	0 0 %	16 100 %
NA	721 88.9 %	84 10.4 %	6 0.7 %	811 100 %
dont know	2 100 %	0 0 %	0 0 %	2 100 %
<b><i>Total</i></b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=184.177 \cdot df=12 \cdot \text{Cramer's } V=0.054 \cdot \text{Fisher's } p=0.000$$

**Figure 6: Plot of Health insurance by army service record**

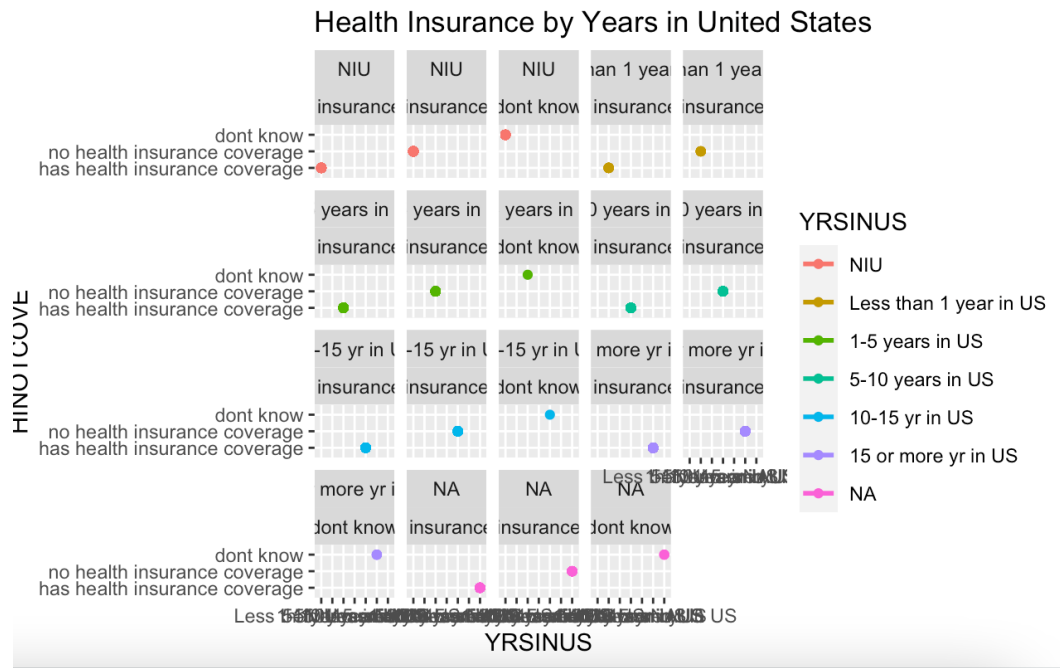
Using the above plot, we observed that individuals who did not respond to the active military service question had the highest percent(100%) of health insurance coverage, and NA had the lowest percent (88.9%) of health insurance coverage. On the other hand, the second-highest percentage (97.8%) of individuals who received health insurance participated in the armed forces. Therefore, we can conclude that participating in the armed forces is beneficial and can increase the chances of individuals having health insurance. Table 5 also illustrates the calculated chi-square value (184.177) and degree of freedom equal 16.

**Table 7: Health insurance by years in the US****Table Health insurance by years in the US**

<i>YRSINUS</i>	<i>HINOTCOVE</i>			<b><i>Total</i></b>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	25189 93.8 %	1607 6 %	45 0.2 %	26841 100 %
Less than 1 year in US	11 64.7 %	6 35.3 %	0 0 %	17 100 %
1-5 years in US	231 74.5 %	78 25.2 %	1 0.3 %	310 100 %
5-10 years in US	340 84.2 %	64 15.8 %	0 0 %	404 100 %
10-15 yr in US	329 78.9 %	88 21.1 %	0 0 %	417 100 %
15 or more yr in US	2693 87.2 %	392 12.7 %	4 0.1 %	3089 100 %
NA	78 79.6 %	19 19.4 %	1 1 %	98 100 %
<b><i>Total</i></b>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$\chi^2=560.111 \cdot df=12 \cdot \text{Cramer's } V=0.095 \cdot \text{Fisher's } p=0.000$

## Health Insurance by Years in United States

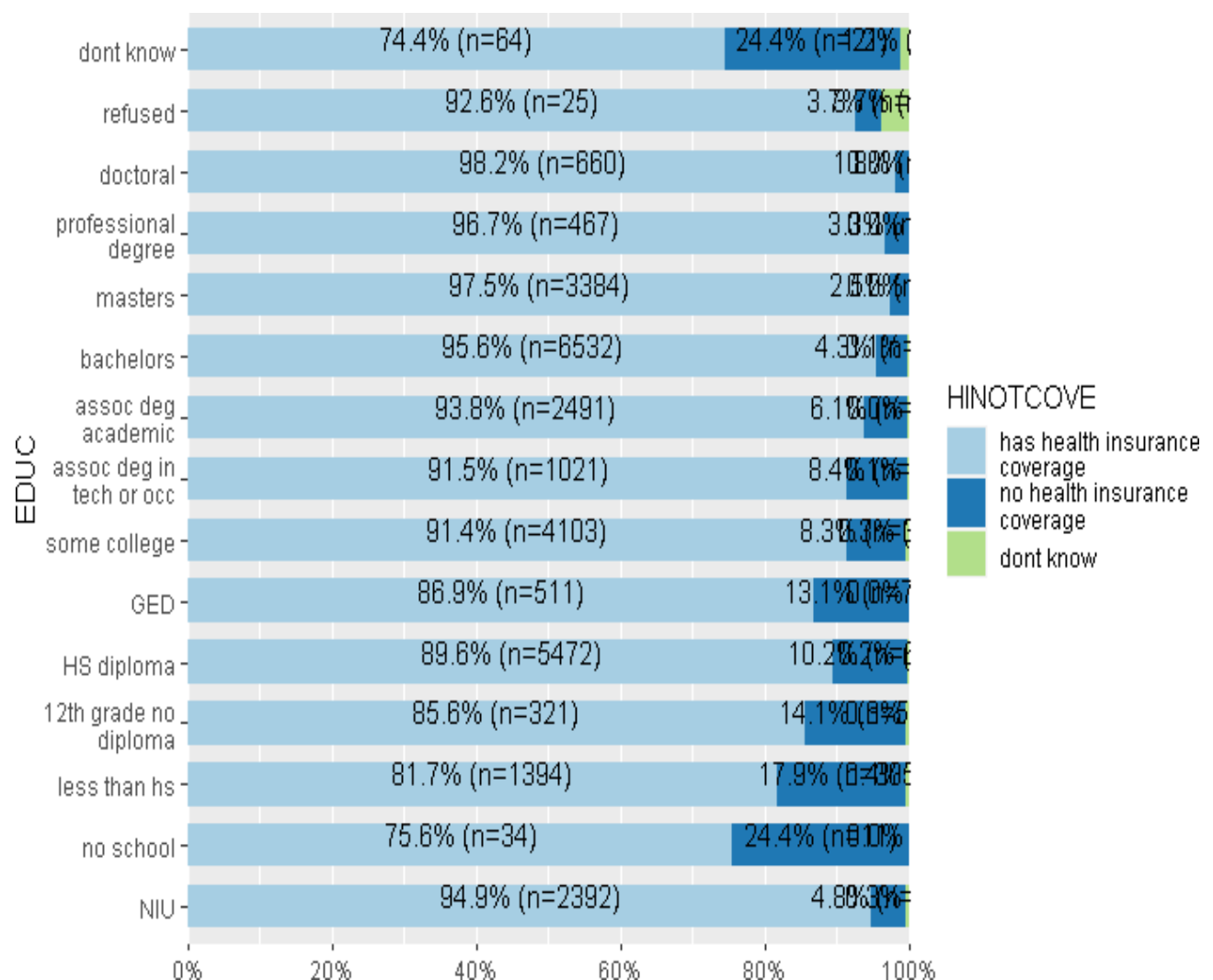


Plot 7 shows NIU had the highest percent (93.8%) of health insurance, and those who lived in the US for less than a year had 64.7% had health insurance coverage.

**Table 8: Health insurance by education status**

<i>EDUC</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
no school	34 75.6 %	11 24.4 %	0 0 %	45 100 %
less than hs	1394 81.7 %	305 17.9 %	7 0.4 %	1706 100 %
12th grade no diploma	321 85.6 %	53 14.1 %	1 0.3 %	375 100 %
HS diploma	5472 89.6 %	625 10.2 %	12 0.2 %	6109 100 %
GED	511 86.9 %	77 13.1 %	0 0 %	588 100 %
some college	4103 91.4 %	372 8.3 %	14 0.3 %	4489 100 %
assoc deg in tech or occ	1021 91.5 %	94 8.4 %	1 0.1 %	1116 100 %
assoc deg academic	2491 93.8 %	163 6.1 %	1 0 %	2655 100 %
bachelors	6532 95.6 %	295 4.3 %	5 0.1 %	6832 100 %
masters	3384 97.5 %	87 2.5 %	1 0 %	3472 100 %
professional degree	467 96.7 %	16 3.3 %	0 0 %	483 100 %
doctoral	660 98.2 %	12 1.8 %	0 0 %	672 100 %
refused	25 92.6 %	1 3.7 %	1 3.7 %	27 100 %
dont know	64 74.4 %	21 24.4 %	1 1.2 %	86 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

 $\chi^2=820.846 \cdot df=28 \cdot \text{Cramer's } V=0.115 \cdot \text{Fisher's } p=0.000$

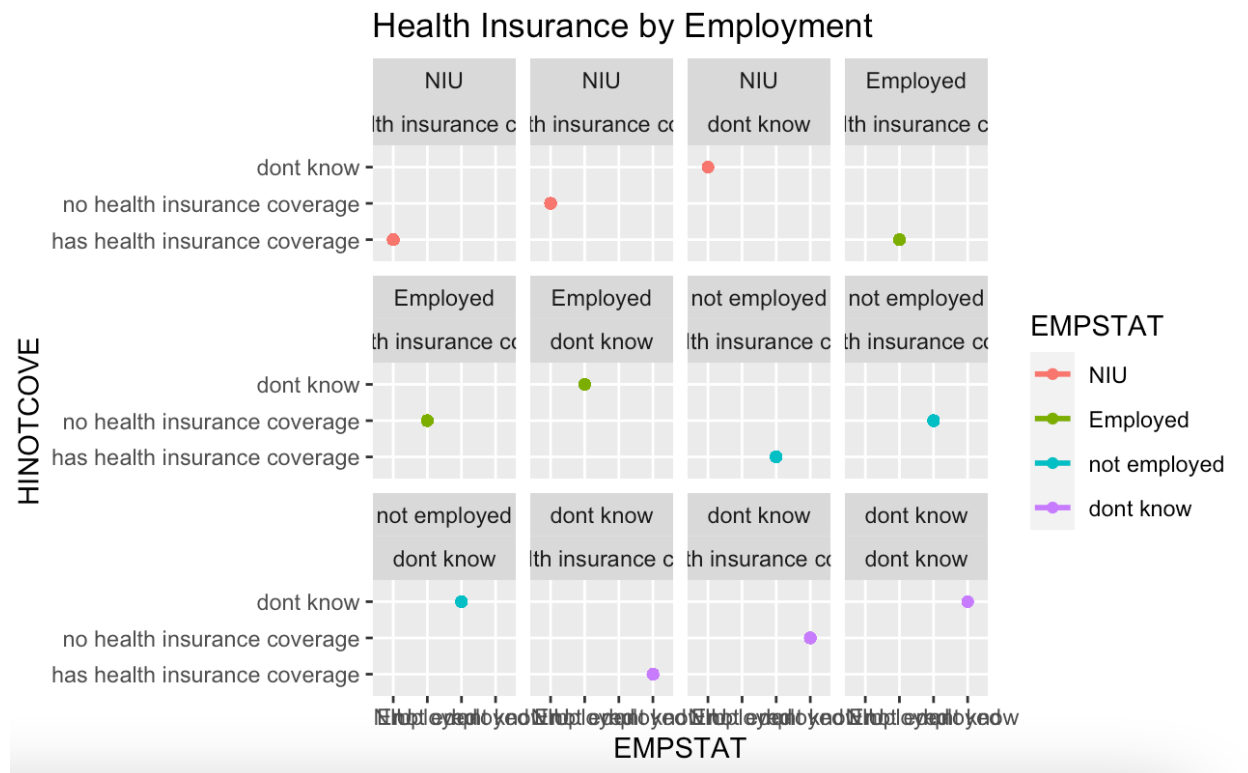
***Figure 8 Plot of Health insurance by education status***

Based on education status, individuals with a doctoral degree had the highest percentage of health insurance coverage (98.2%), and the population with the lowest rate (75.6%) of individuals with health insurance coverage are the ones who have less than a school degree. Therefore, we can conclude that being more educated is directly proportional to health insurance coverage. Thus, having more education increases the probability of individuals having health insurance and vice versa.

**Table 9: Health insurance by employment status****Table Health insurance by employment status**

<i>HINOTCOVE</i>				
<i>EMPSTAT</i>	has health insurance coverage	no health insurance coverage	dont know	<i>Total</i>
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
Employed	14439 91.1 %	1393 8.8 %	15 0.1 %	15847 100 %
not employed	11279 94.3 %	656 5.5 %	23 0.2 %	11958 100 %
dont know	761 89.5 %	83 9.8 %	6 0.7 %	850 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=163.537 \cdot df=6 \cdot \text{Cramer's } V=0.051 \cdot \text{Fisher's } p=0.000$$

***Figure 9: Health insurance by employment status***

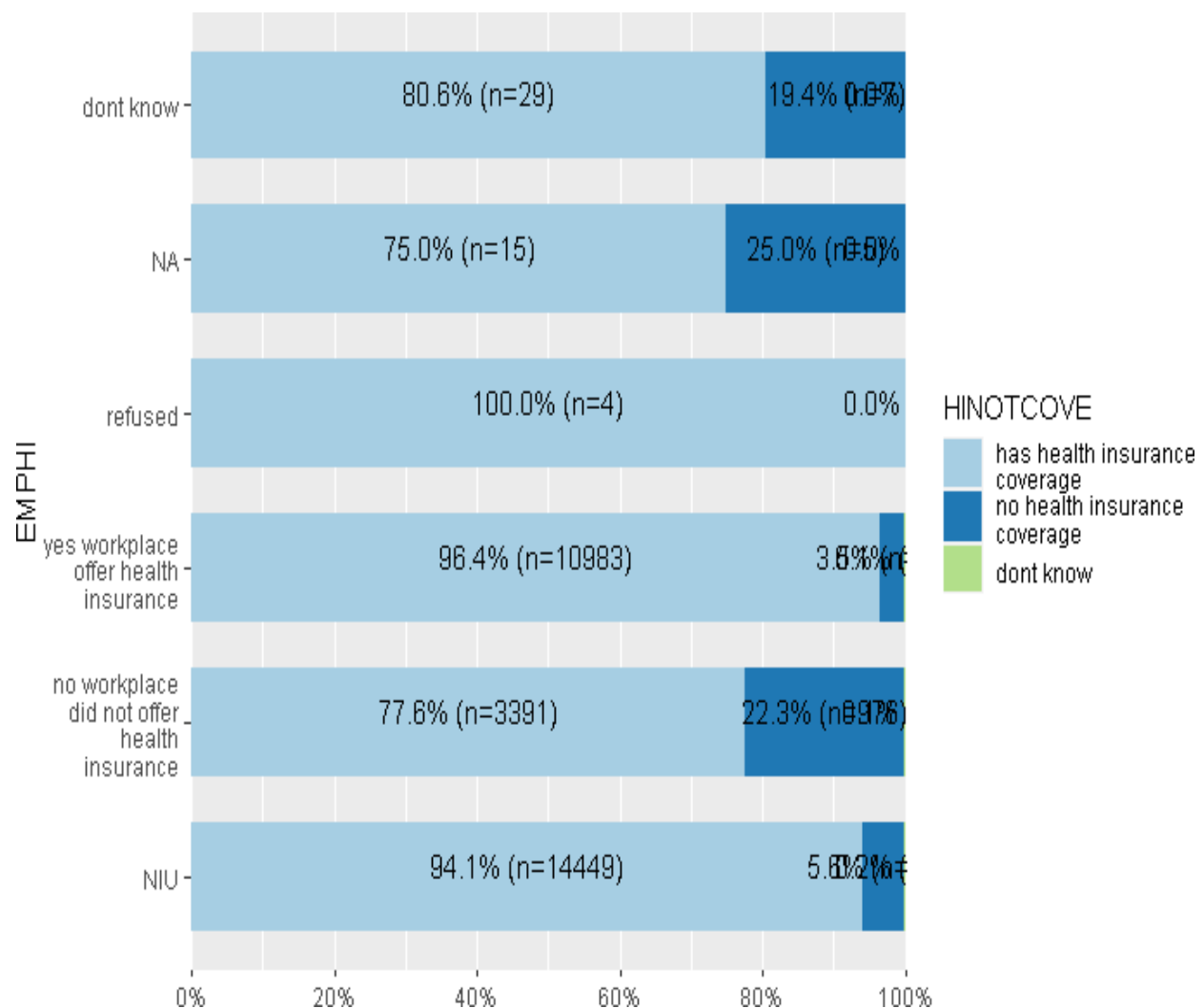
Based on the employment status NIU had the highest percentage (94.9%) of health insurance coverage, and those who did not respond employment status had the lowest percentage (89.5%) of health insurance coverage.



**Table 10: Health insurance by whether employers offer health insurance****Table Health insurance by whether employer offer health insurance**

<i>EMPHI</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	14449 94.1 %	865 5.6 %	36 0.2 %	15350 100 %
no workplace did not offer health insurance	3391 77.6 %	976 22.3 %	5 0.1 %	4372 100 %
yes workplace offer health insurance	10983 96.4 %	401 3.5 %	10 0.1 %	11394 100 %
refused	4 100 %	0 0 %	0 0 %	4 100 %
NA	15 75 %	5 25 %	0 0 %	20 100 %
dont know	29 80.6 %	7 19.4 %	0 0 %	36 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=1804.196 \cdot df=10 \cdot \text{Cramer's } V=0.170 \cdot \text{Fisher's } p=0.000$$

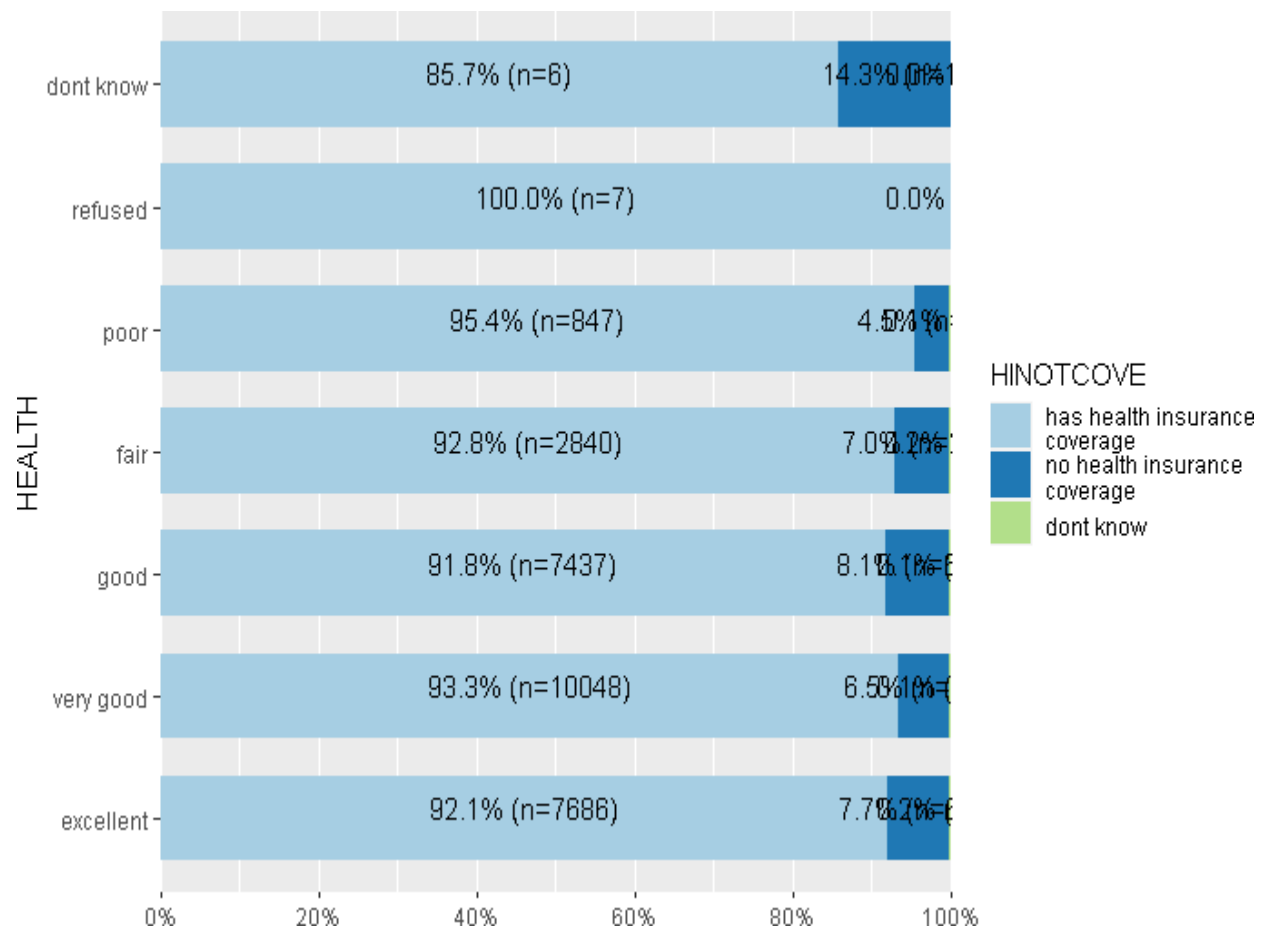
**Figure 10: Plot of Health insurance by whether employers offer health insurance**

From figure 10, we determine that those who refused to respond whether health insurance coverage was offered by the employers or not had 100% health insurance coverage, and NA had the lowest percent (75%) of health insurance coverage.

**Table 11: Health insurance by health status****Table Health insurance by health status**

<i>HEALTH</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
excellent	7686 92.1 %	642 7.7 %	19 0.2 %	8347 100 %
very good	10048 93.3 %	702 6.5 %	14 0.1 %	10764 100 %
good	7437 91.8 %	656 8.1 %	11 0.1 %	8104 100 %
fair	2840 92.8 %	213 7 %	6 0.2 %	3059 100 %
poor	847 95.4 %	40 4.5 %	1 0.1 %	888 100 %
refused	7 100 %	0 0 %	0 0 %	7 100 %
dont know	6 85.7 %	1 14.3 %	0 0 %	7 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=34.636 \cdot df=12 \cdot \text{Cramer's } V=0.024 \cdot \text{Fisher's } p=0.000$$

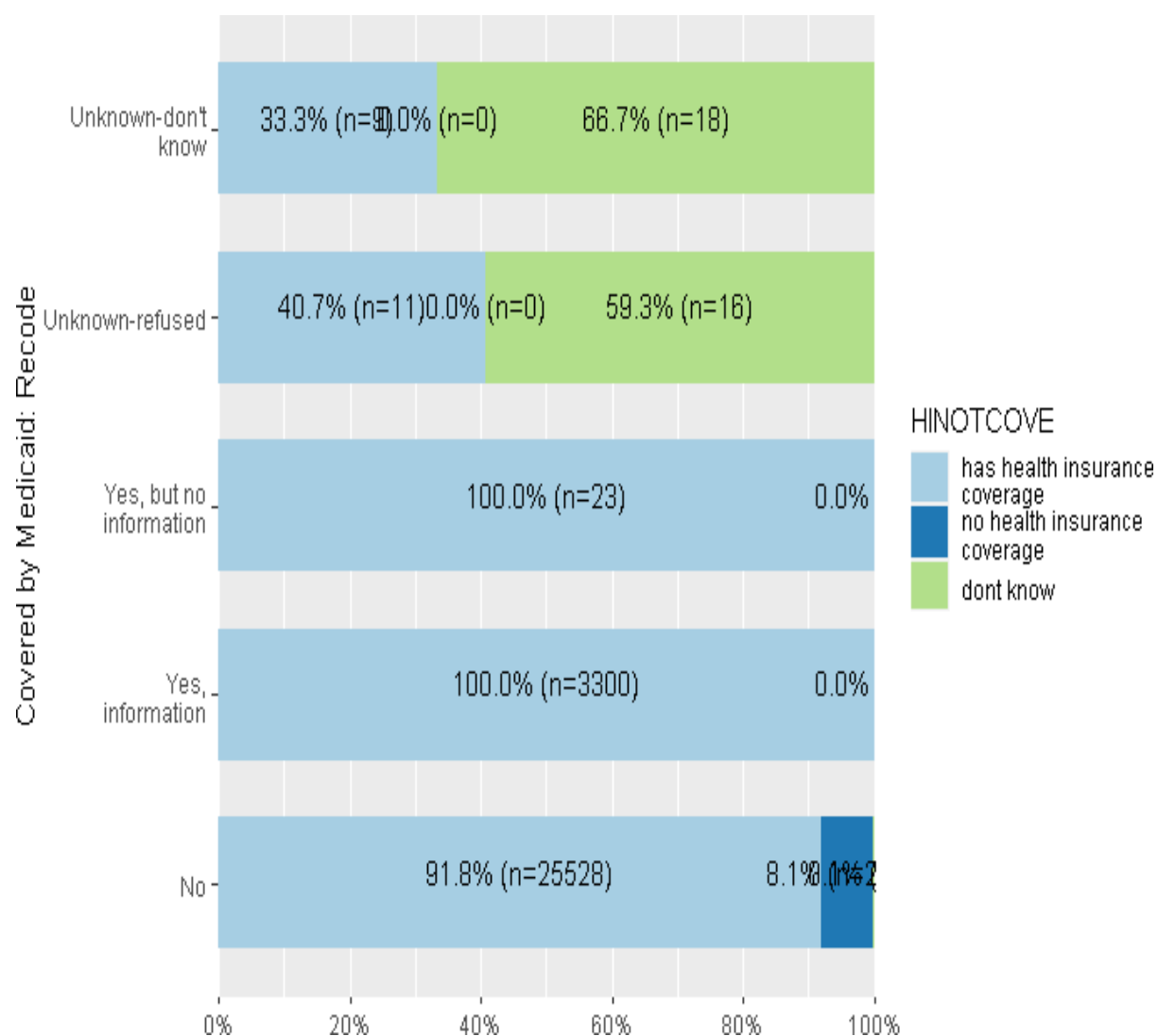
***Figure 11: Health insurance by health status***

Based on health insurance coverage status from figure number 11, we determine that participants who refused to respond to their insurance status had the highest percentage (100%) of health insurance coverage. In contrast, those with good health status have the lowest 91.8% health insurance coverage rates. Therefore, this analysis illustrates that the healthier individuals are less likely to care about their health insurance status than unhealthy individuals as they are less likely to visit the doctor and have a lower chance of getting sick.

**Table 12: Health insurance by Medicaid status****Table Health insurance by whether one has medicaid**

<i>Covered by Medicaid: Recode</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
No	25528 91.8 %	2254 8.1 %	17 0.1 %	27799 100 %
Yes, information	3300 100 %	0 0 %	0 0 %	3300 100 %
Yes, but no information	23 100 %	0 0 %	0 0 %	23 100 %
Unknown-refused	11 40.7 %	0 0 %	16 59.3 %	27 100 %
Unknown-don't know	9 33.3 %	0 0 %	18 66.7 %	27 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$$\chi^2=13400.181 \cdot df=8 \cdot \text{Cramer's } V=0.464 \cdot \text{Fisher's } p=0.000$$

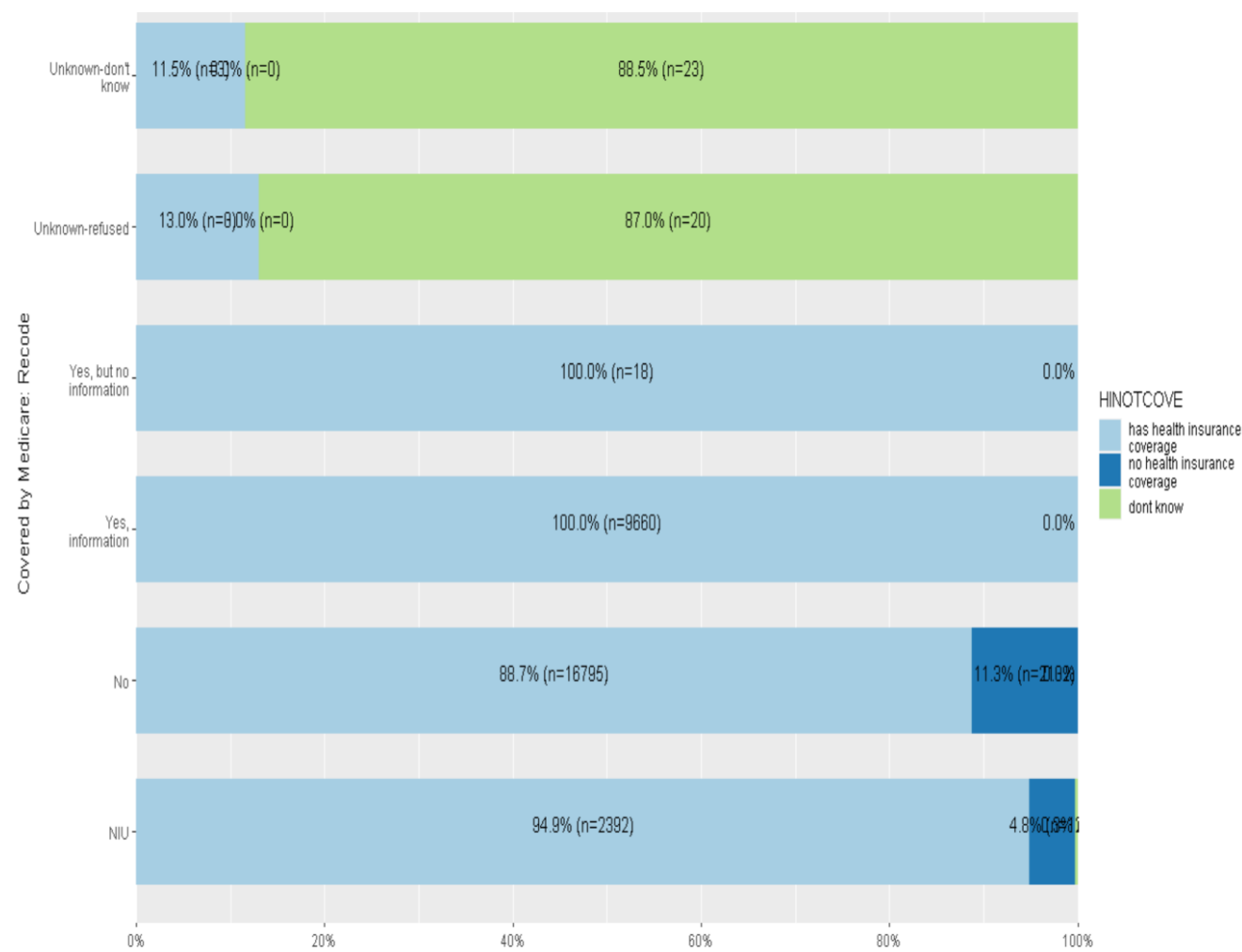
***Figure 12: Health insurance by Medicaid status***

Based on Medicaid status from figure 12 shows that 100 percent of health insurance coverage had Medicaid, and the lowest 33.3 percent who are unknown.

**Table 13: Health insurance by Medicare status****Table Health insurance by whether one has medicare**

<i>Covered by Medicare: Recode</i>	<i>HINOTCOVE</i>			<i>Total</i>
	has health insurance coverage	no health insurance coverage	dont know	
NIU	2392 94.9 %	122 4.8 %	7 0.3 %	2521 100 %
No	16795 88.7 %	2132 11.3 %	1 0 %	18928 100 %
Yes, information	9660 100 %	0 0 %	0 0 %	9660 100 %
Yes, but no information	18 100 %	0 0 %	0 0 %	18 100 %
Unknown-refused	3 13 %	0 0 %	20 87 %	23 100 %
Unknown-don't know	3 11.5 %	0 0 %	23 88.5 %	26 100 %
<i>Total</i>	28871 92.6 %	2254 7.2 %	51 0.2 %	31176 100 %

$\chi^2=24302.186 \cdot df=10 \cdot \text{Cramer's } V=0.624 \cdot \text{Fisher's } p=0.000$

***Figure 13: Plot of Health insurance by Medicare status***

From figure 13, those who are taking Medicare have the highest percentage (100%) of health insurance coverage, and those who do not know or are unknown have the lowest, 11.5 percent had health insurance coverage.



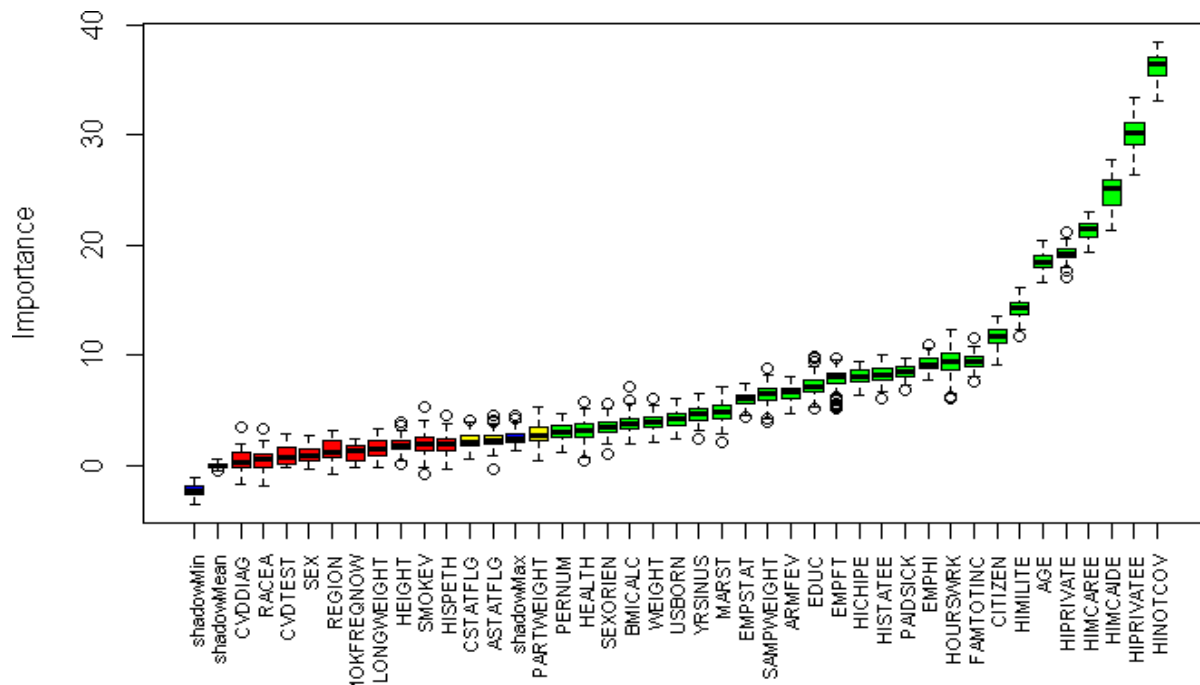
## Feature Engineering

The data was taken through feature selection to exclude non-significant variables from the dataset. The feature selection process reduces the number of features to a minimum and a more manageable number to be able to fit a model that performs well in predicting whether a person has health insurance. In R the Boruta package was used to perform feature selection.

Boruta performed 99 iterations in 2.10329 mins. Twenty-seven attributes confirmed importance: AGE, ARMFEEV, BMICALC, CITIZEN, EDUC, and 22 more; 10 attributes confirmed unimportant: CVDDIAG, CVDTEST, HEIGHT, HISPETH, LONGWEIGHT and five more; 3 tentative attributes left: ASTATFLG, CSTATFLG, PARTWEIGHT.

The following regression plot shows the variables plotted based on importance. The Blue boxplots represent the minimum, average, and maximum Z scores of a shadow attribute. The red, yellow, and green boxplots reflect the Z scores of rejected, tentative, and verified qualities, respectively. It is now time to make a judgment on tentative characteristics. The tentative characteristics will be identified as confirmed or rejected by comparing their median Z scores to the median Z score of the best shadow attribute.

**Regression Graph: Health Insurance Based On The Key Variables**



**The best selected features selected were as follows:**

"PERNUM", "SAMPWEIGHT", "PARTWEIGHT", "ASTATFLG", "CSTATFLG", "AGE", "SEXORIEN", "MARST", "YRSINUS", "USBORN", "CITIZEN", "ARMFEV", "EDUC", "EMPSTAT", "HOURSWRK", "PAIDSICK", "EMPHI", "EMPFT", "FAMTOTINC", "HEALTH", "WEIGHT", "BMICALC", "HIPRIVATEE", "HICHIPE", "HIMILITE", "HISTATEE", "HIMCAIDE", "HIMCAREE", "HINOTCOV", "HIPRIVATE".

### Model fitting

The aim of the study was to identify factors that impact the likelihood or the odds of someone having health insurance, given data about them is used as the predictors. The best model to fit on

data with a categorical dependent variable is the logistic regression model. The model was fitted on the data with features that were selected by the Boruta package in R. The model had only seven significant predictors out of the 29 selected features. The model has fitted again with the seven predictors, and their odds ratios were as follows:

<b>H insurance</b>			
<i>Predictors</i>	<i>Odds Ratios</i>	<i>CI</i>	<i>p</i>
(Intercept)	0.00	0.00 – 0.01	<b>&lt;0.001</b>
Sample Person Weight	1.00	1.00 – 1.00	<b>0.001</b>
CITIZEN: yes US citizen	0.11	0.06 – 0.20	<b>&lt;0.001</b>
CITIZEN: refused	0.00	NA – 29925883778150993920.00	0.980
CITIZEN: NA	0.45	0.13 – 1.31	0.169
Covered by other state-sponsored health plan:Recode	1.33	1.25 – 1.42	<b>&lt;0.001</b>
Covered by Medicaid:Recode	0.23	0.11 – 0.45	<b>&lt;0.001</b>
Covered by Medicare:Recode	0.52	0.35 – 0.76	<b>0.001</b>
Has no health insurance(excluding single service plans)	290.14	142.51 – 660.87	<b>&lt;0.001</b>
Has any private health insurance	0.17	0.10 – 0.27	<b>&lt;0.001</b>
Observations	3117		
R <sup>2</sup> Tjur	0.691		

**From the figure above, the odds ratios can be interpreted as follows:**

- The sample weight does not change the odds of someone having health insurance.
- The odds of someone having health insurance were 0.11 more for US citizens than non-US citizens.
- People covered by the other state-sponsored health coverage were 1.33 times more likely to have health insurance than people who did not have the coverage.
- People covered by the Medicaid health coverage were 0.23 times more likely to have health insurance than people who did not have the coverage.
- People covered by the Medicare health coverage were 0.52 times more likely to have health insurance than people who did not have the coverage.
- People with any private health insurance were 0.17 times more likely to have health insurance than people who did not have the coverage.

## **Conclusion:**

In conclusion, health insurance is an essential aspect of life. Accidents and illnesses are inevitable and consequently, so are medical expenses. Health insurance helps prepare for the often-unexpected medical expenses. Employers usually provide medical coverage to employees and sometimes also to their immediate families. In the long run, medical insurance help keep the economy afloat and promotes health.

However, some differences exist in health insurance based on social attributes. From the feature engineering process, we learned the seven essential variables that predict the health insurance status of individuals are citizenship status, Medicaid status, the employer provides insurance, family income, military status, age, and private health insurance status.

Based on the existing data, people of color are less likely to be insured than their white counterparts. This segregation is mainly due to political influence, such as enacted policies. These policies disadvantage people of color and especially immigrants. Other factors associated with these discrepancies affect the health sector, such as housing and employment for the minority. In most cases, the minority have jobs that do not offer medical coverage. This problem further expanded during the Trump administration that significantly disadvantaged people of color.

In recent times, other policies have helped narrow the gap in health insurance coverage. These policies include the Affordable Care Act, allowing the poor to afford health insurance. In addition, insurance policies such as Medicaid have been developed under the ACA to benefit the poor communities, mainly compromising Blacks and Hispanics. Finally, the ARPA strengthened the ACA, instituted under the Biden administration. ARPA allowed most people under the federal poverty line to access health insurance and healthcare. This action further narrowed the gap in health insurance coverage but did not eliminate it.

Solving the disparities in health insurance in the United States can lead to an overall improved country's economic welfare. It will also result in a healthier nation and consequently a healthier workforce. In addition, children will be able to get vaccinated on time. As a result, we can catch cancer-like illnesses early on, and pregnant mothers can get proper prenatal care to benefit their unborn children. This improvement would ultimately reduce the overall mortality rate.

In the long run, it will solve one of the country's social issues, which is racism and segregation. Diseases affect everyone equally, and therefore, everyone should have equal access to healthcare. Therefore, individuals who live in the U.S must know the importance of getting

health coverage. And the role of the government should be to enact more policies and take actions to bridge the gap and provide health insurance to everyone regardless of their background and other features.

## References

- Kronick R. (2009). Health insurance coverage and mortality revisited. *Health services research, 44*(4), 1211–1231. <https://doi.org/10.1111/j.1475-6773.2009.00973.x>
- Leah A. Burke, MD and Andrew M. Ryan, PhD. The Complex Relationship between Cost and Quality in US Health Care, February 2014, *Journal American Medical Association*  
*Journal of Ethics* Volume 16, Number 2: 124-130.
- Brown, E. Richard and Ojeda, Victoria D. and Wyn, Roberta and Levan, Rebecka (2000)  
 Racial and Ethnic Disparities in Access to Health Insurance and Health Care. Other.  
 UNSPECIFIED.
- Artiga, S., Hill, L., Orgera, K., & Damico, A. (2021). Health coverage by race and Ethnicity, 2010–2019. *Kaiser Family Foundation*.
- McWilliams. (2009, June). *NCBI - WWW Error Blocked Diagnostic*. National Center for Biotechnology Information. Retrieved December 16, 2021, from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2881446/>
- How Health Care Inequality Increases Costs for Everyone*. (2021, April 1). The Balance. <https://www.thebalance.com/health-care-inequality-facts-types-effect-solution-4174842#toc-health-care-inequality>
- Mangan, D. (2016, July 6). *Wealthy spending more on health care than poor and middle class, reversing trend*. CNBC. Retrieved December 16, 2021, from <https://www.cnbc.com/2016/07/06/wealthy-spending-more-on-health-care-than-poor-and-middle-class-reversing-trend.html>

- Crimmel, B. C. (2004, May). *Medical Expenditure Panel Survey (MEPS) & Agency for Healthcare Research and Quality (AHRQ)*. Health Insurance Coverage and Income Levels for the U.S. Noninstitutionalized Population under Age 65, 2001. Retrieved December 16, 2021, from [https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj9i9Hf0en0AhXrmOAKHTjCi8QFnoECBgQAQ&url=https%3A%2F%2Fmeps.ahrq.gov%2Fdata\\_files%2Fpublications%2Fst40%2Fstat40.pdf&usg=AOvVaw25RrPFSt1Smq6Y9o3msd8z](https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwj9i9Hf0en0AhXrmOAKHTjCi8QFnoECBgQAQ&url=https%3A%2F%2Fmeps.ahrq.gov%2Fdata_files%2Fpublications%2Fst40%2Fstat40.pdf&usg=AOvVaw25RrPFSt1Smq6Y9o3msd8z)
- Artiga, S., Hill, L., Orgera, K., & Damico, A. (2021). Health coverage by race and Ethnicity, 2010–2019. *Kaiser Family Foundation*.
- Boulware, L. E. (2020). Race disparities in the COVID-19 pandemic—Solutions lie in policy, not biology. *JAMA Network Open*, 3(8), e2018696-e2018696.
- Cutler, D. (2020, April). How will COVID-19 affect the health care economy? In *JAMA Health Forum* (Vol. 1, No. 4, pp. e200419-e200419). American Medical Association.
- Fox, P. D., & Kongstvedt, P. R. (2013). A history of managed health care and health insurance in the United States. *Essentials of managed health care. Burlington, Sixth Edition: Jones & Bartlett Learning*.
- Ganesan, D., Harshini, R., Nagulu, S. P., & Surya, V. S. An Empirical Evaluation On Proclivity Of Customers Towards Health Insurance During Pandemic.
- Gemelas, J., Davison, J., & Ing, S. (2021). Inequities in employment by race, ethnicity, and sector during COVID-19. *Journal of racial and ethnic health disparities*, 1-6.
- Griffith, K., Evans, L., & Bor, J. (2017). The Affordable Care Act reduced socioeconomic disparities in health care access. *Health Affairs*, 36(8), 1503-1510.



- Harris, R., Tobias, M., Jeffreys, M., Waldegrave, K., Karlsen, S., & Nazroo, J. (2006). Racism and health: The relationship between experience of racial discrimination and health in New Zealand. *Social science & medicine*, 63(6), 1428-1441.
- Johnson-Staub, C., & Weerasinghe, I. (2021). Supporting Mental Health Policies and Practices through the American Rescue Plan.
- Liu, J., Xu, J., Zhang, T., & Ma, Y. (2018). Ethical reflections on the equity of the current basic health insurance system reform in China: a case study in hunan province. *Cambridge Quarterly of Healthcare Ethics*, 27(3), 447-458.
- Mitrović, A., & Pešić, M. (2019). IMPORTANCE OF HEALTH INSURANCE FOR EMPLOYEES IN HOTEL MANAGEMENT AND TOURISM IN THE REPUBLIC OF SERBIA. *CHALLENGES OF TOURISM AND BUSINESS LOGISTICS IN THE 21 ST CENTURY*, 43.
- Murray, J. E. (2007). *Origins of American health insurance: a history of industrial sickness funds*. Yale University Press.
- O'Mahen, P. N., & Petersen, L. A. (2021, June). Subsidizing Affordable Care Act Marketplace Insurance for Adults Below the Federal Poverty Level. In *JAMA Health Forum* (Vol. 2, No. 6, pp. e210871-e210871). American Medical Association.
- Parmet, W. E. (2018). Immigration and health care under the Trump administration. *Health Affairs*.
- Shanske, D. (2021). Getting Beyond Ad Hoc Fiscal Federalism: A proposal for a Default Federal Liquidity Facility for the States.
- Shrivastava, S. R. B., & Shrivastava, P. S. (2021). Addressing health sector discrimination: need of the hour. *Menoufia Medical Journal*, 34(1), 404.

Stone, D. A. (1993). The struggle for the soul of health insurance. *Journal of Health Politics, Policy and Law*, 18(2), 287-317.

SUDHA, M. S., & MURUGESAN, S. (2021). A STUDY ON AWARENESS LEVEL AND THE INFLUENCE OF EDUCATIONAL QUALIFICATION TOWARDS THE HEALTH INSURANCE POLICIES (WITH SPECIAL REFERENCE TO CHENNAI CITY). *PalArch's Journal of Archaeology of Egypt/Egyptology*, 18(1), 4683-4689.

Sun, C., & Zhai, Z. (2020). The efficacy of social distance and ventilation effectiveness in preventing COVID-19 transmission. *Sustainable cities and society*, 62, 102390.

Treadwell, H. M. (2019). Wages and women in health care: the race and gender gap.

Cooper, Z. (2016, February 12). *Why Is Healthcare So Expensive?* Yale Insights.<https://insights.som.yale.edu/insights/why-is-healthcare-so-expensive>

Sanger-Katz, M. (2018, March 13). *Why Is U.S. Health Care So Expensive? Some of the Reasons You've Heard Turn Out to Be Myths*. The New York Times.  
<https://www.nytimes.com/2018/03/13/upshot/united-states-health-care-resembles-rest-of-world.html>

Statista. (2019, August 8). *The U.S. Has the Most Expensive Healthcare System in the World*. Statista Infographics. <https://www.statista.com/chart/8658/health-spending-per-capita/>

*Obesity is a Common, Serious, and Costly Disease*. (2021, September 30). Centers for DiseaseControl and Prevention. <https://www.cdc.gov/obesity/data/adult.html>

Luo, & Escalante, C. L. (2018). Health care service utilization of documented and

- undocumented hired farmworkers in the U.S. *The European Journal of Health Economics*, 19(7), 923–934. <https://doi.org/10.1007/s10198-017-0939-x>
- Alon, T., et al. (2020). The impact of COVID-19 on gender equality, National Bureau of economic research.
- Tahmasbi, F., Schild, L., Ling, C., Blackburn, J., Stringhini, G., Zhang, Y., & Zannettou, S. (2021, April 1). "go eat a bat, Chang!": On the emergence of Sinophobic Behavior on web communities in the face of covid-19: *Proceedings of the web conference 2021*. ACM Conferences. Retrieved May 24, 2022, from <https://dl.acm.org/doi/abs/10.1145/3442381.3450024>
- Borjas, G. J., & Cassidy, H. (2020, May 25). *The adverse effect of the COVID-19 Labor Market Shock on immigrant employment*. NBER. Retrieved May 24, 2022, from <https://www.nber.org/papers/w27243>
- Couch, K. A., Fairlie, R. W., & Xu, H. (2020, September 14). *Early evidence of the impacts of covid-19 on minority unemployment*. Journal of Public Economics. Retrieved May 24, 2022, from [https://www.sciencedirect.com/science/article/pii/S0047272720301511?casa\\_token=IQ-TrVVcZVcAAAAA%3AIU8QbFPQfGka7VpPJGHcd\\_X3u57fzkiSmEe9tYd\\_-LvwwqIGLmo92f5ekipfDhqCJAPLmj5OV5Mas](https://www.sciencedirect.com/science/article/pii/S0047272720301511?casa_token=IQ-TrVVcZVcAAAAA%3AIU8QbFPQfGka7VpPJGHcd_X3u57fzkiSmEe9tYd_-LvwwqIGLmo92f5ekipfDhqCJAPLmj5OV5Mas)
- Wehby, & Lyu, W. (2018). The Impact of the ACA Medicaid Expansions on Health Insurance Coverage through 2015 and Coverage Disparities by Age, Race/Ethnicity, and Gender. *Health Services Research*, 53(2), 1248–1271. <https://doi.org/10.1111/1475-6773.12711>
- Courtemanche, Marton, J., Ukert, B., Yelowitz, A., Zapata, D., & Fazlul, I. (2019). The three-year impact of the Affordable Care Act on disparities in insurance coverage. *Health Services Research*, 54(1), 307–316. <https://doi.org/10.1111/1475-6773.13077>

Kaestner, & Lubotsky, D. (2016). Health Insurance and Income Inequality. *The Journal of Economic Perspectives*, 30(2), 53–77. <https://doi.org/10.1257/jep.30.2.53>

Jennifer Tolbert, K. O. F. @\_K. O. on T. (2020, November 12). *Key facts about the uninsured population*. KFF. Retrieved May 24, 2022, from <https://www.kff.org/uninsured/issue-brief/key-facts-about-the-uninsured-population/>

Latoya Hill Follow @hill\_latoya on Twitter and Samantha Artiga Follow @SArtiga2 on Twitter  
Published: Feb 22, 2022. (2022, February 22). *Covid-19 cases and deaths by Race/ethnicity: Current data and changes over time*. KFF. Retrieved May 24, 2022, from <https://www.kff.org/coronavirus-covid-19/issue-brief/covid-19-cases-and-deaths-by-race-ethnicity-current-data-and-changes-over-time/>