

Socioeconomic Factors Affecting Health Insurance In The USA

Hassan Fayyaz | Thakur Prasad Ghimire | Adelia Fida
Professor Kevin Foster
Statistics and Introduction to Econometrics (B2000)
The City College of New York
December 19th, 2021

```
library(ggplot2)
library(stats)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##   filter, lag

## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union

library(randomForest)

## randomForest 4.6-14

## Type rfNews() to see new features/changes/bug fixes.

##
## Attaching package: 'randomForest'

## The following object is masked from 'package:dplyr':
##
##   combine

## The following object is masked from 'package:ggplot2':
##
##   margin
```

```
library(corrplot)

## corrplot 0.92 loaded

library(haven)
library(AER)

## Loading required package: car

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
##      recode

## Loading required package: lmtest

## Loading required package: zoo

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric

## Loading required package: sandwich

## Loading required package: survival

library(foreign)
library(nnet)
library(reshape2)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.
3.1 --

## v tibble  3.1.6      v purrr   0.3.4
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.0      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflict
s() --
## x randomForest::combine() masks dplyr::combine()
## x dplyr::filter()         masks stats::filter()
```

```
## x dplyr::lag()           masks stats::lag()
## x randomForest::margin() masks ggplot2::margin()
## x car::recode()          masks dplyr::recode()
## x purrr::some()          masks car::some()

library(stargazer)

##
## Please cite as:

## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

library(ggeffects)
library(gridExtra)

##
## Attaching package: 'gridExtra'

## The following object is masked from 'package:randomForest':
##
##      combine

## The following object is masked from 'package:dplyr':
##
##      combine

library(Boruta)

setwd("C:/Homework EcoB2000/Econometric Final Research")
library(ipumsr)
library(tidyverse)
library(ggplot2)
ddi <- read_ipums_ddi("nhis_00002.xml")
data <- read_ipums_micro(ddi)

## Use of data from IPUMS NHIS is subject to conditions including that users
## should cite the data appropriately. Use command `ipums_conditions()` for more
## details.

data$REGION <- as.factor(data$REGION)
levels(data$REGION) <- c("Northeast", "Midwest", "South", "West")

data$SEX <- as.factor(data$SEX)
levels(data$SEX) <- c("Male", "Female", "Refused", "dont know")
```

```
data$SEXORIEN <- as.factor(data$SEXORIEN)
levels(data$SEXORIEN) <- c("NIU", "Lesbian or gay", "straight", "bisexual", "some
thing else", "dont know", "refused", "NA")

data$MARST <- as.factor(data$MARST)
levels(data$MARST) <- c("NIU", "Married", "Married spouse not there", "Married s
pouse NA", "Widowed", "Divorced", "Separated", "never married", "unknown")

data$RACEA <- as.factor(data$RACEA)
levels(data$RACEA) <- c("white", "Black", "Aleut Alaskan", "American Indian", "As
ian", "Other", "refused", "not ascertained", "unknown")

data$HISPETH <- as.factor(data$HISPETH)
levels(data$HISPETH) <- c("Not Hispanic", "Mexican", "Other Hispanic", "NA")

data$YRSINUS <- as.factor(data$YRSINUS)
levels(data$YRSINUS) <- c("NIU", "Less than 1 year in US", "1-5 years in US", "5
-10 years in US", "10-15 yr in US", "15 or more yr in US", "NA")

data$CITIZEN <- as.factor(data$CITIZEN)
levels(data$CITIZEN) <- c("No not US citizen", "yes US citizen", "refused", "NA"
, "dont know")

data$ARMFEV <- as.factor(data$ARMFEV)
levels(data$ARMFEV) <- c("NIU", "No never active duty", "active only for traini
ng", "yes ever served in armed forces", "refused", "NA", "dont know")

data$EDUC <- as.factor(data$EDUC)
levels(data$EDUC) <- c("NIU", "no school", "less than hs", "12th grade no diplom
a", "HS diploma", "GED", "some college", "assoc deg in tech or occ", "assoc deg ac
ademic", "bachelors", "masters", "professional degree", "doctoral", "refused", "don
t know")

data$EMPSTAT <- as.factor(data$EMPSTAT)
levels(data$EMPSTAT) <- c("NIU", "Employed", "not employed", "dont know")

data$EMPHI <- as.factor(data$EMPHI)
levels(data$EMPHI) <- c("NIU", "no workplace did not offer health insurance", "
yes workplace offer health insurance", "refused", "NA", "dont know")

data$EMPFT <- as.factor(data$EMPFT)
levels(data$EMPFT) <- c("NIU", "parttime", "fulltime", "refused", "NA", "dont know
")
```

```

data$HEALTH <- as.factor(data$HEALTH)
levels(data$HEALTH) <- c("excellent","very good","good","fair","poor","refuse
d","dont know")

is.na(data$HOURSWRK) <- which(data$HOURSWRK > 95) # hours of work each week
is.na(data$HEIGHT) <- which(data$HEIGHT > 94) # height in inches
is.na(data$WEIGHT) <- which(data$WEIGHT > 900) # weight in pounds
is.na(data$BMICALC) <- which(data$BMICALC > 900) # BMI Body Mass Index

data$HINOTCOVE <- as.factor(data$HINOTCOVE)
levels(data$HINOTCOVE) <- c("has health insurance coverage","no health insura
nce coverage","dont know")

data$EMPSTAT<-as.factor(data$EMPSTAT)
levels(data$EMPSTAT)<-c("not employed","employed","dont know","NIU")

```

This analysis i decided to focus on one subgroup that is the black race in the USA. #Data cleaning

```

library(dplyr)

mydata= select(data,-c("YEAR","SERIAL","STRATA","PSU","NHISHID","NHISPID","HH
X"))

```

#Inspect the data and check for missing values

```

sum(is.na(mydata))

## [1] 11249

mydata= na.omit(mydata)#remove all missing values

```

To check the missing values we need to inspect the data.

```

library(sjPlot)

## Learn more about sjPlot with 'browseVignettes("sjPlot")'.

#Health insurance by race

sjPlot::tab_xtab(var.row = mydata$RACEA,
var.col = mydata$HINOTCOVE,
title="Table Health insurance by Race",
show.row.prc =TRUE)

```

Table Health insurance by Race

RACEA

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

white

2236693.8 %

14456.1 %

360.2 %

23847100 %

Black

284890.6 %

2849 %

100.3 %

3142100 %

Aleut Alaskan

18680.9 %

4419.1 %

00 %

230100 %

American Indian

24785.5 %

4214.5 %

00 %

289100 %

Asian

159994.5 %

925.4 %

10.1 %

1692100 %

Other

42593.8 %

286.2 %

00 %

453100 %

refused

1280 %

320 %

00 %

15100 %

not ascertained

118078.7 %

31521 %

40.3 %

1499100 %

unknown

888.9 %

111.1 %

00 %

9100 %

Total

2887192.6 %

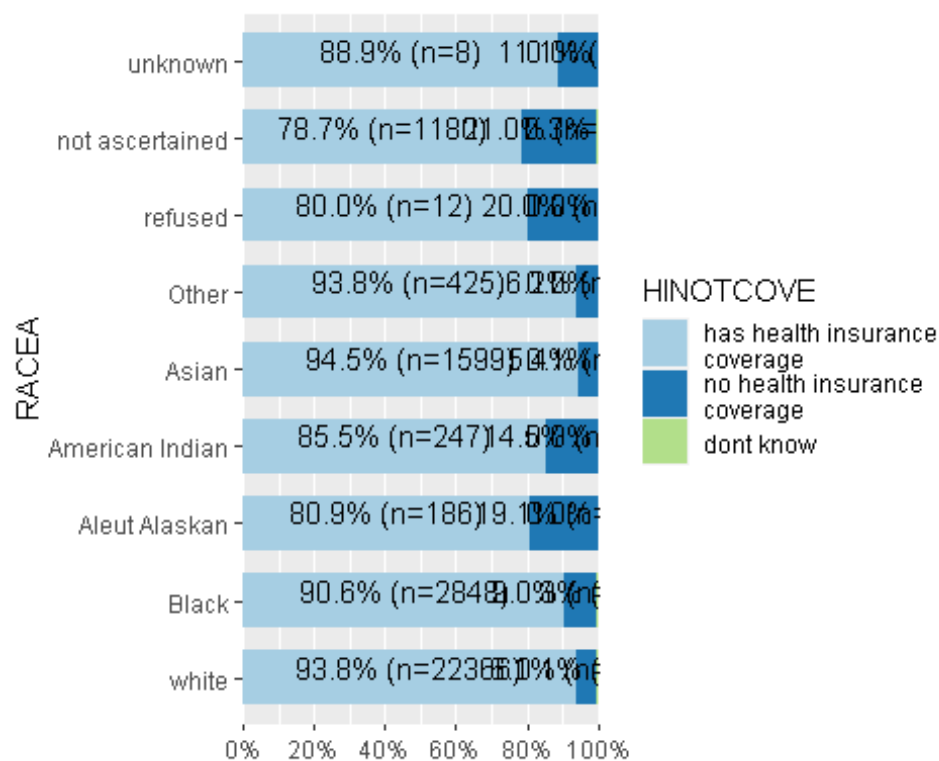
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$RACEA,mydata$HINOTCOVE,margin="row",bar.pos="stack",
coord.flip=TRUE)
```



```
sjPlot::tab_xtab(var.row = mydata$RACEA,
var.col = mydata$HINOTCOVE,
title="Table Health insurance by Race",
show.row.prc =TRUE)
```


Table Health insurance by Race

RACEA

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

white

2236693.8 %

14456.1 %

360.2 %

23847100 %

Black

284890.6 %

2849 %

100.3 %

3142100 %

Aleut Alaskan

18680.9 %

4419.1 %

00 %

230100 %

American Indian

24785.5 %

4214.5 %

00 %

289100 %

Asian

159994.5 %

925.4 %

10.1 %

1692100 %

Other

42593.8 %

286.2 %

00 %

453100 %

refused

1280 %

320 %

00 %

15100 %

not ascertained

118078.7 %

31521 %

40.3 %

1499100 %

unknown

888.9 %

111.1 %

00 %

9100 %

Total

2887192.6 %

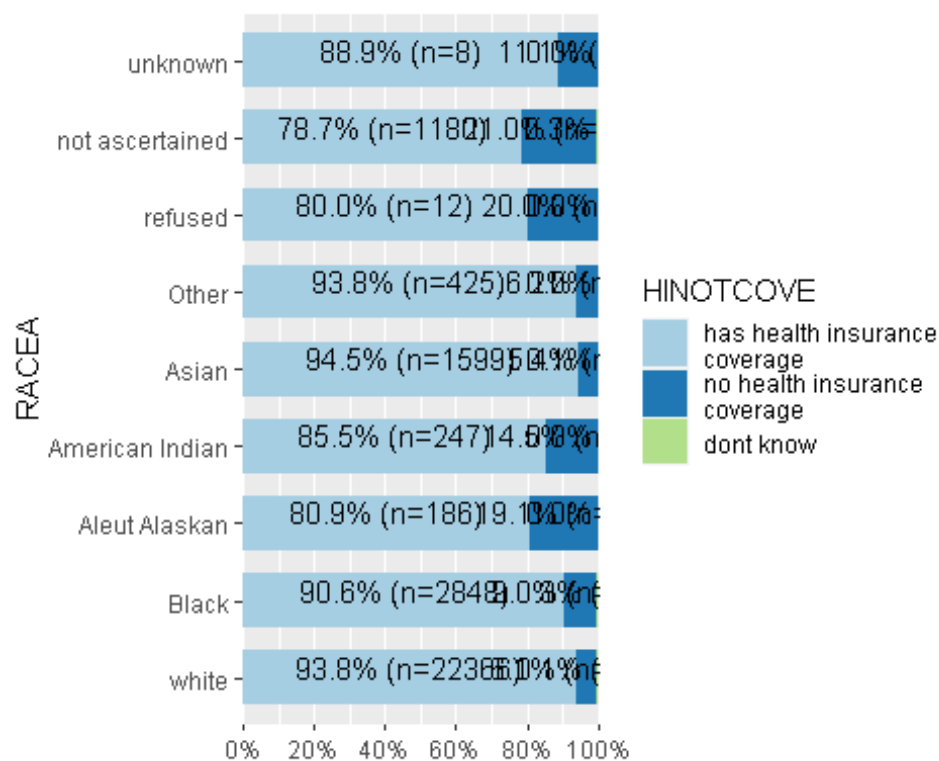
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$RACEA,mydata$HINOTCOVE,margin="row",bar.pos="stack",
coord.flip=TRUE)
```



#health insurance by sex

```
sjPlot::tab_xtab(var.row = mydata$SEX,
var.col=mydata$HINOTCOVE,
```

```
title="Table Health insurance by sex",
show.row.prc =TRUE,drop.empty = TRUE)
```

Table Health insurance by sex

SEX

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

Male

1345691.8 %

11798 %

290.2 %

14664100 %

Female

1541593.4 %

10756.5 %

220.1 %

16512100 %

Refused

00 %

00 %

00 %

0100 %

dont know

00 %

00 %

00 %

0100 %

Total

2887192.6 %

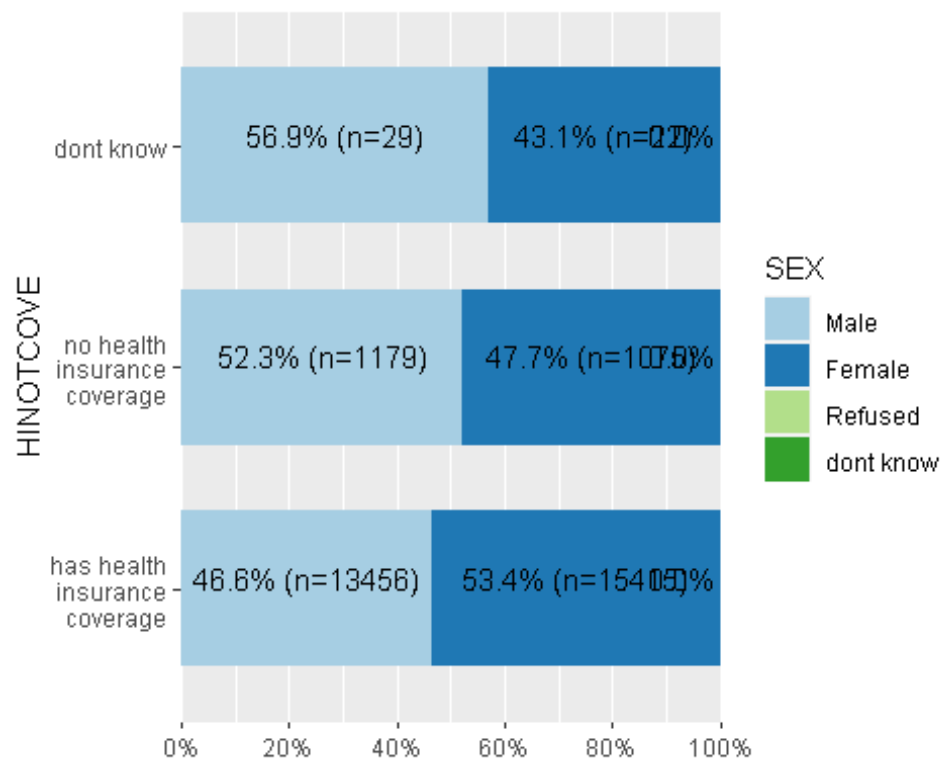
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$HINOTCOVE, mydata$SEX, margin="row", bar.pos="stack",
  ord.flip=TRUE)
```



#Health insurance by marital status

```

sjPlot::tab_xtab(var.row = mydata$EDUC,
                 var.col=mydata$HINOTCOVE,
                 title="Table Health insurance by education status",
                 show.row.prc =TRUE,drop.empty=TRUE)

```

Table Health insurance by education status

EDUC

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

239294.9 %

1224.8 %

70.3 %

2521100 %

no school

3475.6 %

1124.4 %

00 %

45100 %

less than hs

139481.7 %

30517.9 %

70.4 %

1706100 %

12th grade nodiploma

32185.6 %

5314.1 %

10.3 %

375100 %

HS diploma

547289.6 %

62510.2 %

120.2 %

6109100 %

GED

51186.9 %

7713.1 %

00 %

588100 %

some college

410391.4 %

3728.3 %

140.3 %

4489100 %

assoc deg in tech or occ

102191.5 %

948.4 %

10.1 %

1116100 %

assoc deg academic

249193.8 %

1636.1 %

10 %

2655100 %

bachelors

653295.6 %

2954.3 %

50.1 %

6832100 %

masters

338497.5 %

872.5 %

10 %

3472100 %

professional degree

46796.7 %

163.3 %

00 %

483100 %

doctoral

66098.2 %

121.8 %

00 %

672100 %

refused

2592.6 %

13.7 %

13.7 %

27100 %

dont know

6474.4 %

2124.4 %

11.2 %

86100 %

Total

2887192.6 %

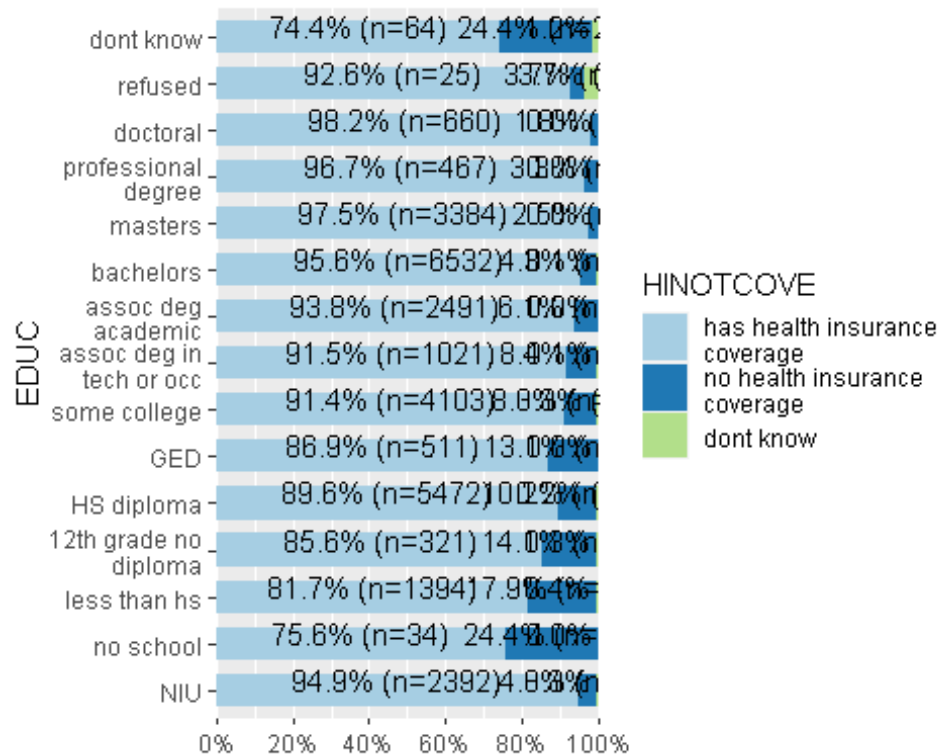
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$EDUC,mydata$HINOTCOVE,  
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



a

```

sjPlot::tab_xtab(var.row = mydata$MARST,
                 var.col=mydata$HINOTCOVE,
                 title = "Table Health insurance by marital status",
                 show.row.prc =TRUE,drop.empty=TRUE)

```

Table Health insurance by marital status

MARST

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

239294.9 %

1224.8 %

70.3 %

2521100 %

Married

1222294.8 %

6675.2 %

100.1 %

12899100 %

Married spouse notthere

57591.7 %

528.3 %

00 %

627100 %

Married spouse NA

266.7 %

133.3 %

00 %

3100 %

Widowed

288897.7 %

622.1 %

50.2 %

2955100 %

Divorced

411091.7 %

3708.3 %

10 %

4481100 %

Separated

35983.3 %

7216.7 %

00 %

431100 %

never married

550986.8 %

82212.9 %

190.3 %

6350100 %

unknown

81489.5 %

869.5 %

91 %

909100 %

Total

2887192.6 %

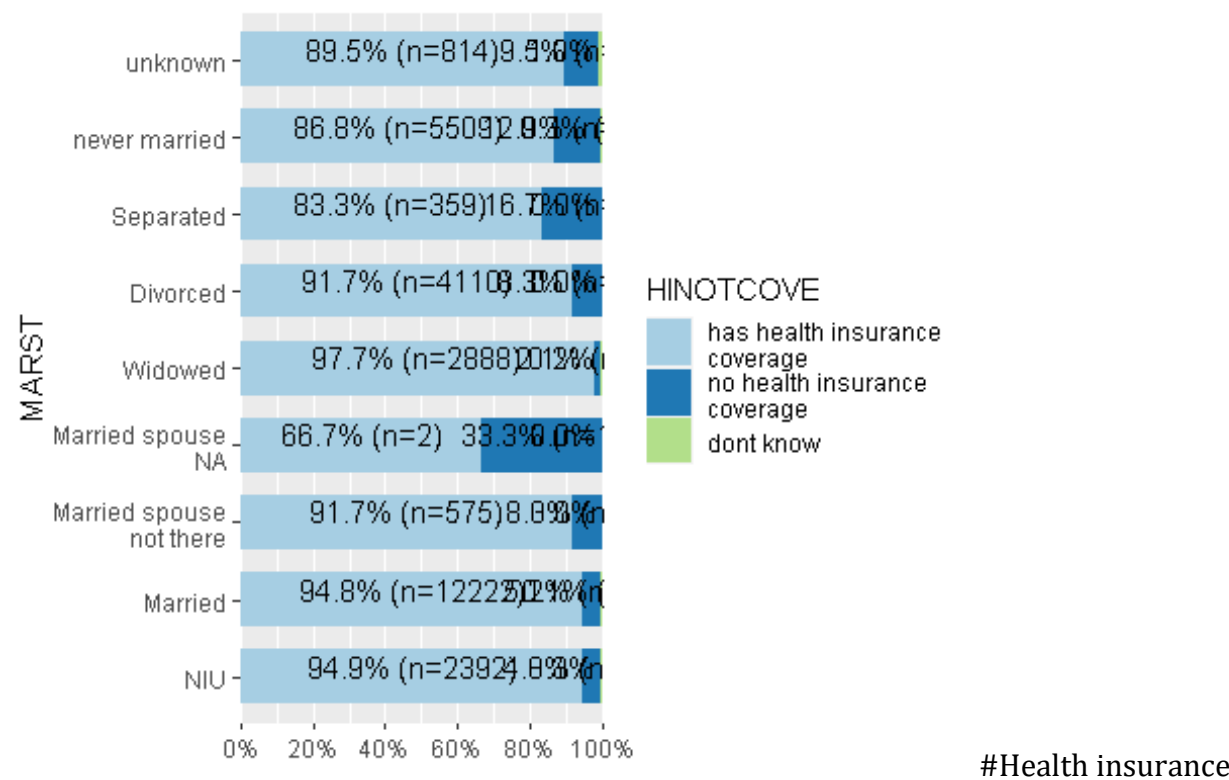
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$MARST,mydata$HINOTCOVE,  
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



status by CITIZEN status

```
sjPlot::tab_xtab(var.row = mydata$CITIZEN,  
                 var.col=mydata$HINOTCOVE,  
                 title = "Table Health insurance by citizenship",  
                 show.row.prc =TRUE,drop.empty=TRUE)
```

Table Health insurance by citizenship

CITIZEN

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

No not US citizen

114769.9 %

49029.9 %

40.2 %

1641100 %

yes US citizen

2715094 %

16975.9 %

400.1 %

28887100 %

refused

4481.5 %

916.7 %

11.9 %

54100 %

NA

51989 %

589.9 %

61 %

583100 %

dont know

11100 %

00 %

00 %

11100 %

Total

2887192.6 %

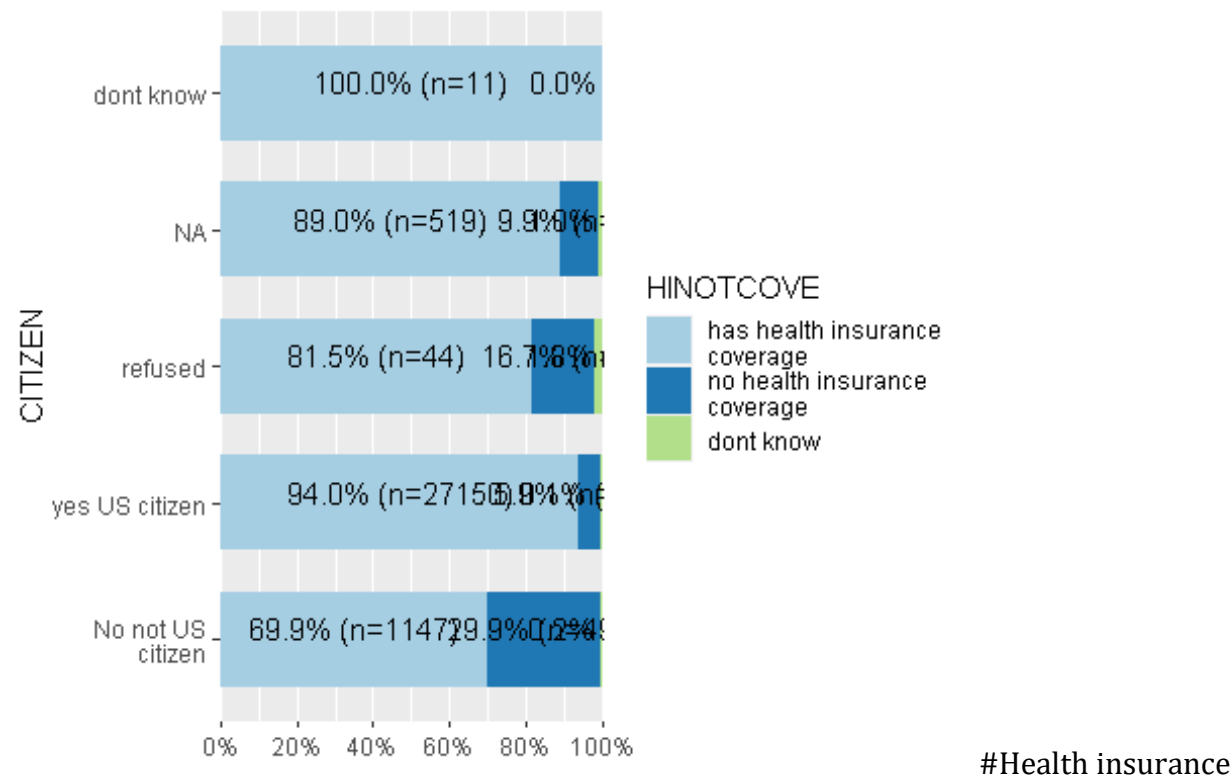
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$CITIZEN,mydata$HINOTCOVE,
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



by Sex orientation

```
sjPlot::tab_xtab(var.row = mydata$SEXORIEN,
  var.col=mydata$HINOTCOVE,
  title = "Table Health insurance by sex orientation",
  show.row =TRUE,drop.empty=TRUE)
```

Table Health insurance by sex orientation

SEXORIEN

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

239294.9 %

1224.8 %

70.3 %

2521100 %

Lesbian or gay

48093.4 %

336.4 %

10.2 %

514100 %

straight

2472792.5 %

19677.4 %

350.1 %

26729100 %

bisexual

32088.4 %

4211.6 %

00 %

362100 %

something else

11089.4 %

1310.6 %

00 %

123100 %

dont know

20292.7 %

167.3 %

00 %

218100 %

refused

15294.4 %

74.3 %

21.2 %

161100 %

NA

48889.1 %

549.9 %

61.1 %

548100 %

Total

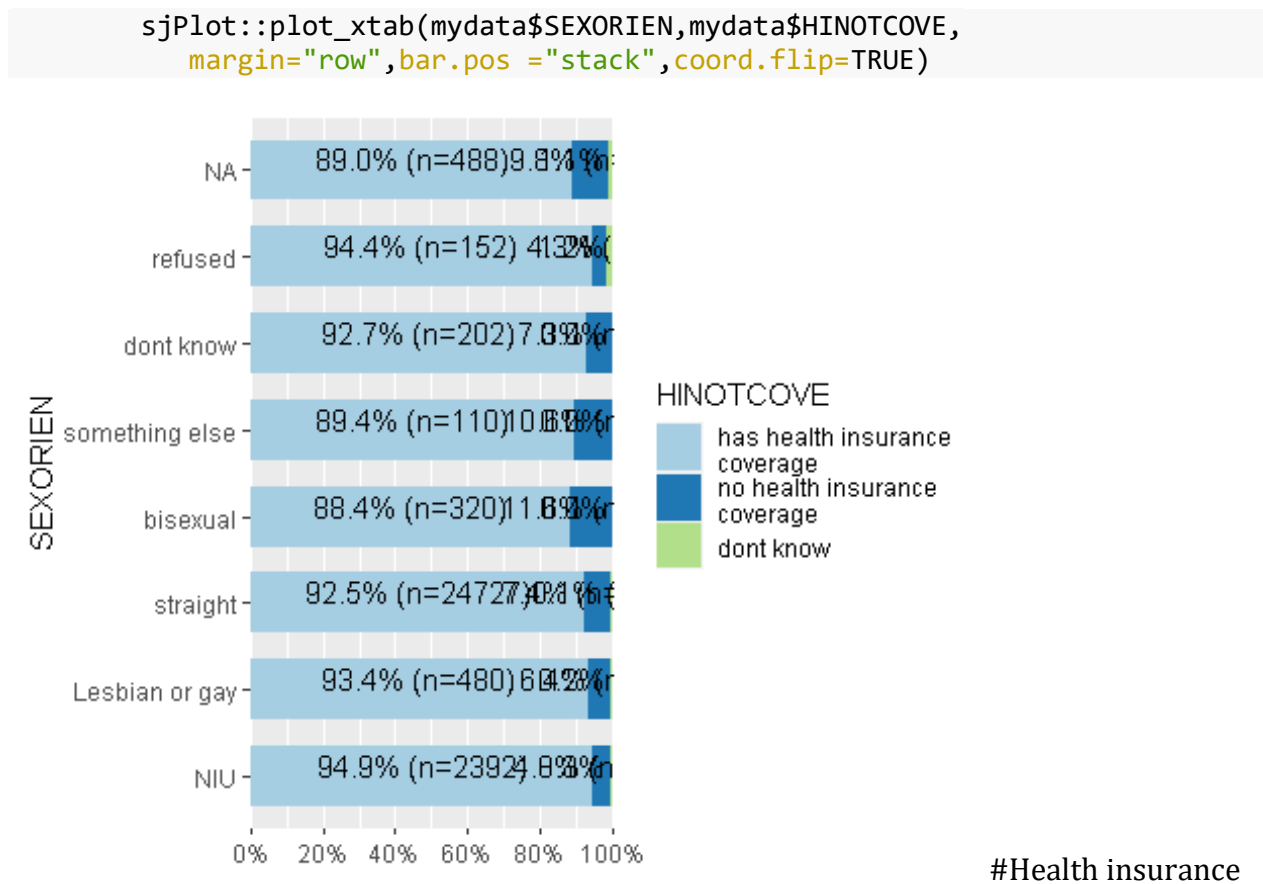
2887192.6 %

22547.2 %

510.2 %

31176100 %

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



by wheather one served in the army

```
sjPlot::tab_xtab(var.row = mydata$ARMFEV,
  var.col=mydata$HINOTCOVE,
  title = "Table Health insurance by army service ",
  show.row.prc =TRUE, drop.empty=TRUE)
```

Table Health insurance by army service

ARMFEV

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

239294.9 %

1224.8 %

70.3 %

2521100 %

No never active duty

2284291.9 %

19798 %

370.1 %

24858100 %

active only for training

48997 %

153 %

00 %

504100 %

yes ever served in armed forces

241097.8 %

532.2 %

10 %

2464100 %

refused

1593.8 %

16.2 %

00 %

16100 %

NA

72188.9 %

8410.4 %

60.7 %

811100 %

dont know

2100 %

00 %

00 %

2100 %

Total

2887192.6 %

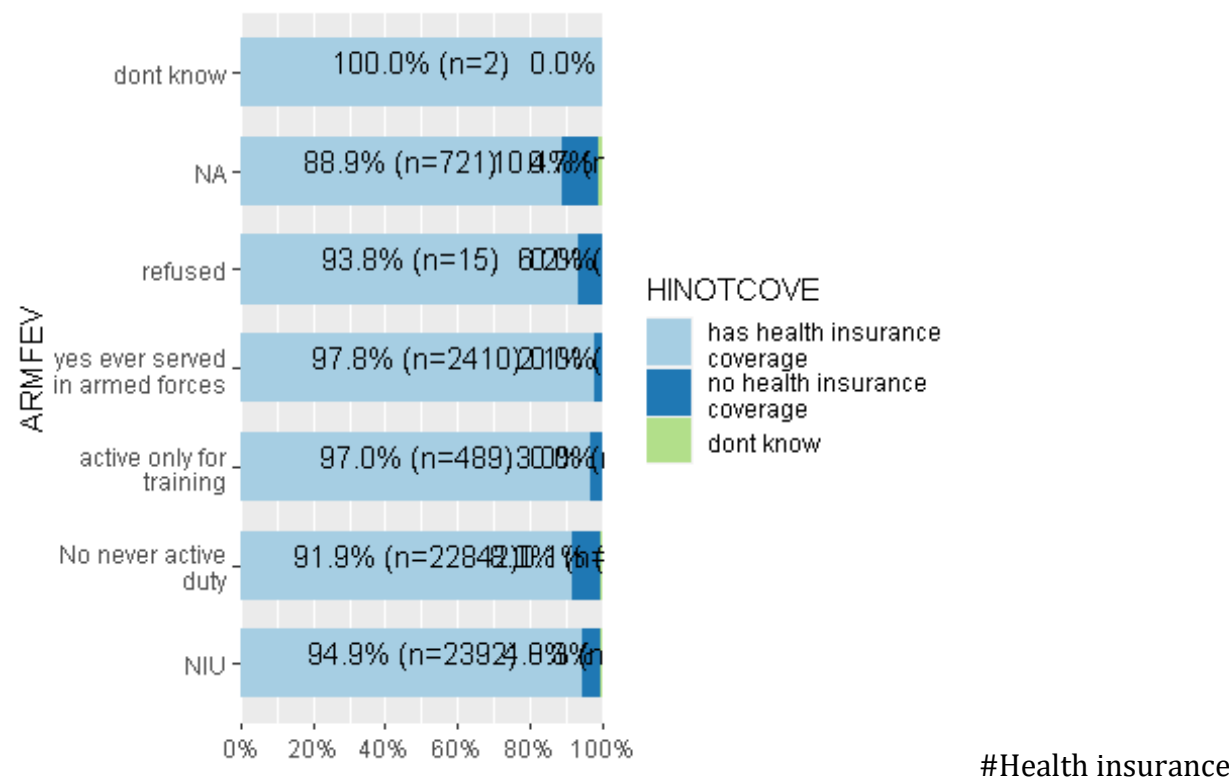
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$ARMFEV,mydata$HINOTCOVE,  
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



by years lived in the united states

```
sjPlot::tab_xtab(var.row = mydata$YRSINUS,
                 var.col=mydata$HINOTCOVE,
                 title = "Table Health insurance by years in the US",
                 show.row.prc =TRUE,drop.empty=TRUE)
```

Table Health insurance by years in the US

YRSINUS
HINOTCOVE
Total
has health insurancecoverage
no health insurancecoverage
dont know
NIU
2518993.8 %

16076 %

450.2 %

26841100 %

Less than 1 year inUS

1164.7 %

635.3 %

00 %

17100 %

1-5 years in US

23174.5 %

7825.2 %

10.3 %

310100 %

5-10 years in US

34084.2 %

6415.8 %

00 %

404100 %

10-15 yr in US

32978.9 %

8821.1 %

00 %

417100 %

15 or more yr in US

269387.2 %

39212.7 %

40.1 %

3089100 %

NA

7879.6 %

1919.4 %

11 %

98100 %

Total

2887192.6 %

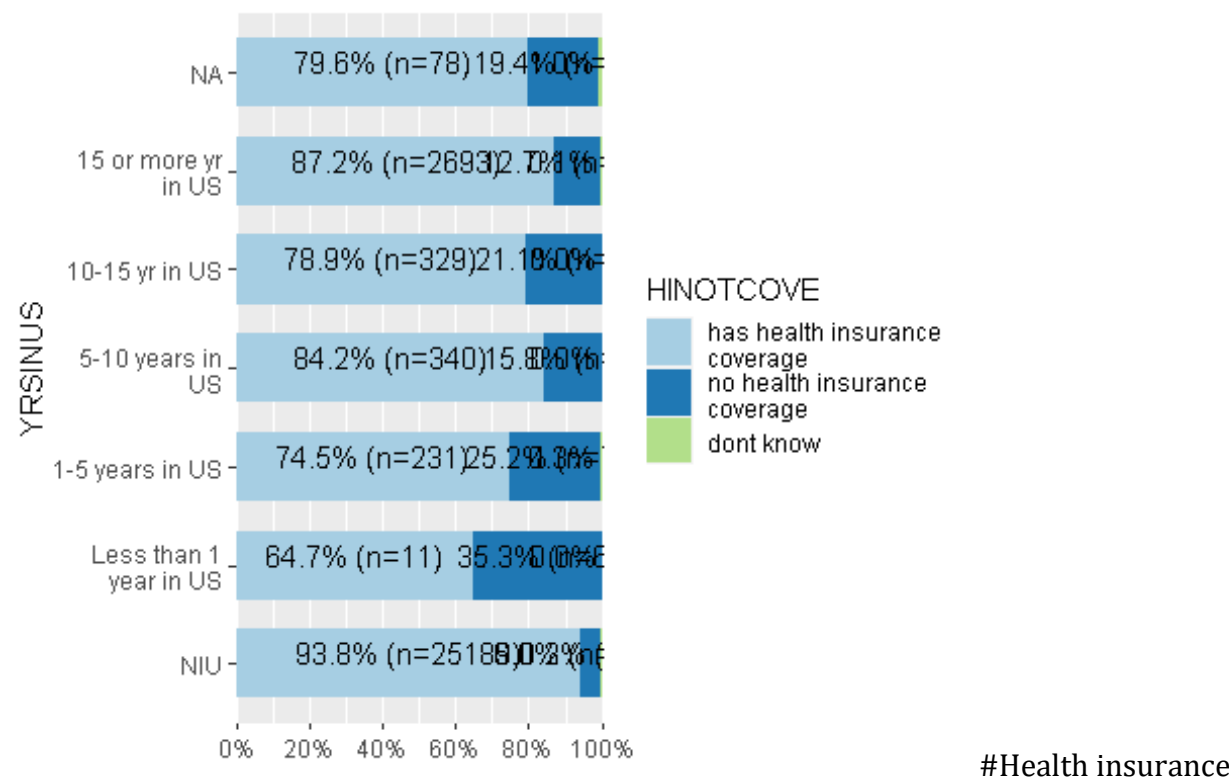
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$YRSINUS,mydata$HINOTCOVE,  
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



by employment status

```
sjPlot::tab_xtab(var.row = mydata$EMPSTAT,  
                 var.col=mydata$HINOTCOVE,  
                 title = "Table Health insurance by employment status",  
                 show.row.prc =TRUE,drop.empty=TRUE)
```

Table Health insurance by employment status

EMPSTAT

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

not employed

239294.9 %

1224.8 %

70.3 %

2521100 %

employed

1443991.1 %

13938.8 %

150.1 %

15847100 %

dont know

1127994.3 %

6565.5 %

230.2 %

11958100 %

NIU

76189.5 %

839.8 %

60.7 %

850100 %

Total

2887192.6 %

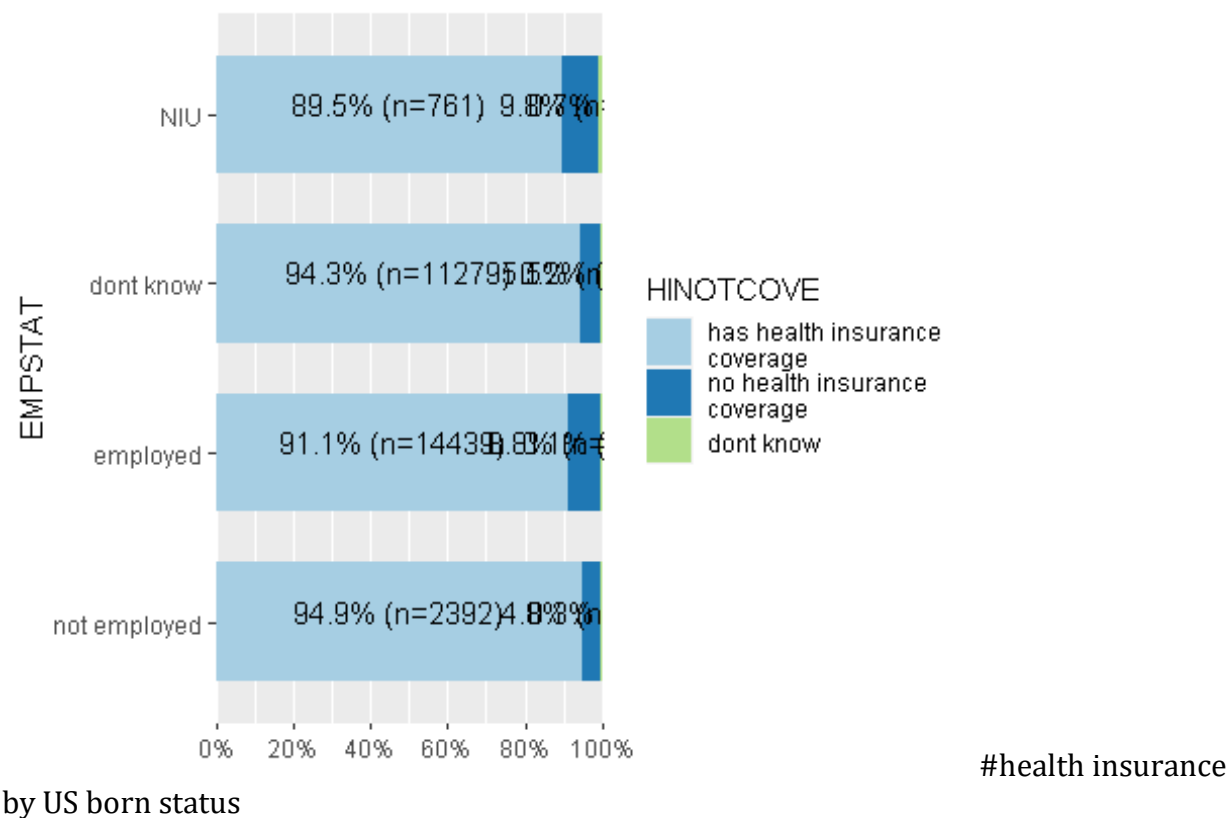
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$EMPSTAT,mydata$HINOTCOVE,  
margin="row",bar.pos="stack",coord.flip=TRUE)
```



by US born status

```
sjPlot::tab_xtab(var.row = mydata$EMPSTAT,
  var.col=mydata$HINOTCOVE,
  title = "Table Health insurance by whether employer offer health insurance",
  show.row.prc =TRUE,drop.empty=TRUE)
```

Table Health insurance by whether employer offer health insurance

EMPSTAT

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

1444994.1 %

8655.6 %

360.2 %

15350100 %

no workplace did not offer health insurance

339177.6 %

97622.3 %

50.1 %

4372100 %

yes workplace offer health insurance

1098396.4 %

4013.5 %

100.1 %

11394100 %

refused

4100 %

00 %

00 %

4100 %

NA

1575 %

525 %

00 %

20100 %

don't know

2980.6 %

719.4 %

00 %

36100 %

Total

2887192.6 %

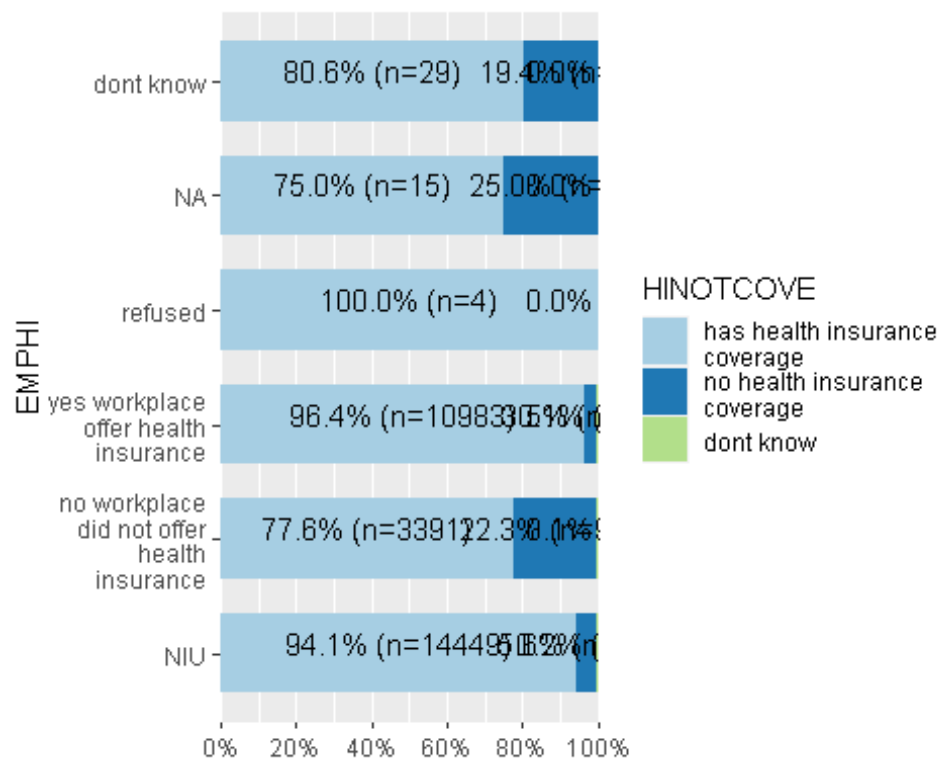
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$EMPHI,mydata$HINOTCOVE,
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



```

sjPlot::tab_xtab(var.row = mydata$HEALTH,
                  var.col=mydata$HINOTCOVE,
                  title = "Table Health insurance by health status",
                  show.row.prc =TRUE)

```

Table Health insurance by health status

HEALTH

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

excellent

768692.1 %

6427.7 %

190.2 %

8347100 %

very good

1004893.3 %

7026.5 %

140.1 %

10764100 %

good

743791.8 %

6568.1 %

110.1 %

8104100 %

fair

284092.8 %

2137 %

60.2 %

3059100 %

poor

84795.4 %

404.5 %

10.1 %

888100 %

refused

7100 %

00 %

00 %

7100 %

dont know

685.7 %

114.3 %

00 %

7100 %

Total

2887192.6 %

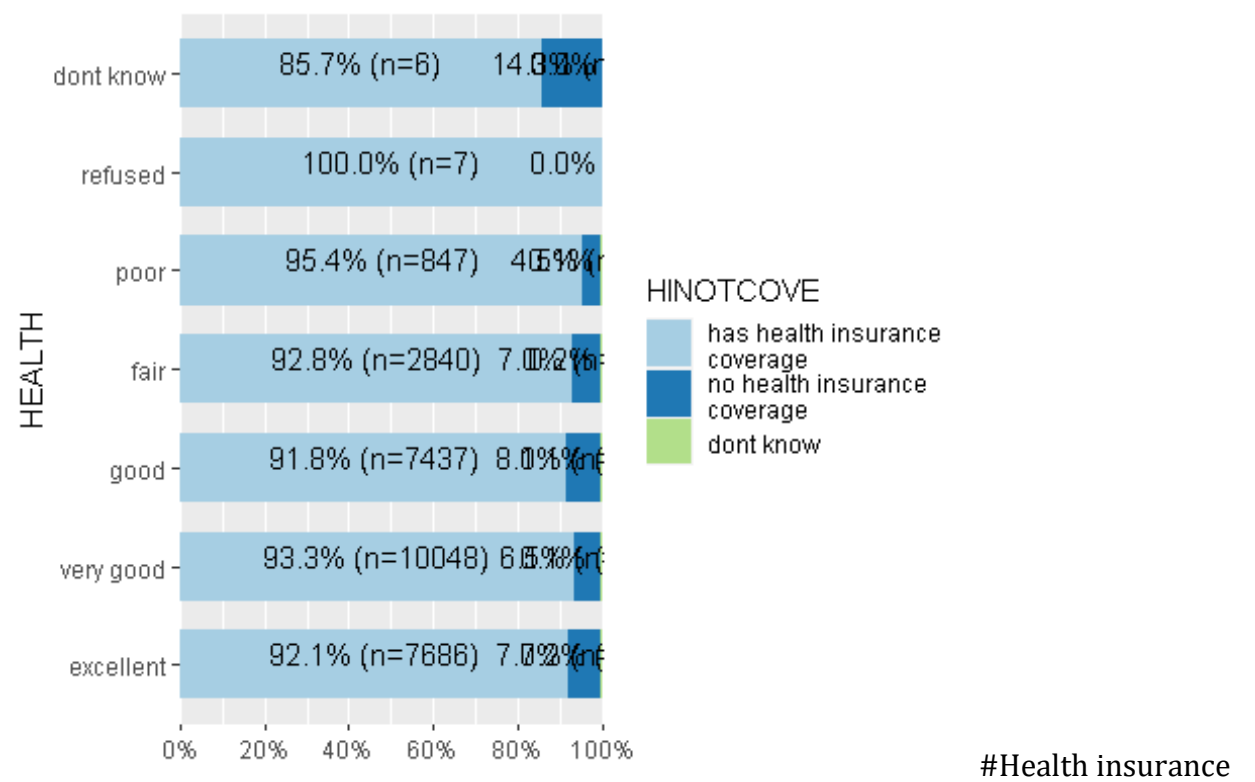
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$HEALTH,mydata$HINOTCOVE,  
margin="row",bar.pos="stack",coord.flip=TRUE)
```



status by medical status

```
sjPlot::tab_xtab(var.row = mydata$HIMCAREE,  
var.col=mydata$HINOTCOVE,  
title = "Table Health insurance by whether one homedicare",  
show.row.prc =TRUE,drop.empty=TRUE)
```

Table Health insurance by whether one homedicare

Covered by Medicare:Recode

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

NIU

239294.9 %

1224.8 %

70.3 %

2521100 %

No

1679588.7 %

213211.3 %

10 %

18928100 %

Yes, information

9660100 %

00 %

00 %

9660100 %

Yes, but no information

18100 %

00 %

00 %

18100 %

Unknown-refused

313 %

00 %

2087 %

23100 %

Unknown-don't know

311.5 %

00 %

2388.5 %

26100 %

Total

2887192.6 %

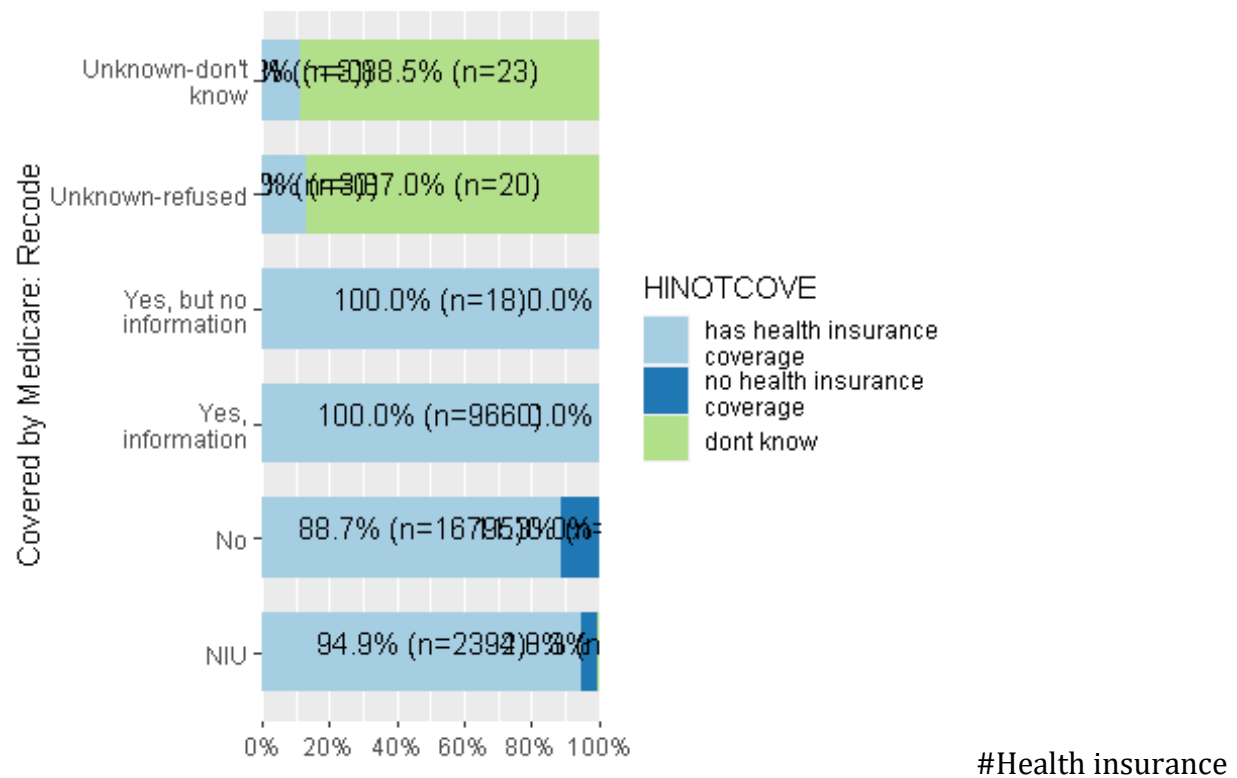
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$HIMCAREE,mydata$HINOTCOVE,  
  margin="row",bar.pos="stack",coord.flip=TRUE)
```



status by medicaid status

```
sjPlot::tab_xtab(var.row = mydata$HIMCAIDE,
  var.col=mydata$HINOTCOVE,
  title = "Table Health insurance by whether one has madicaid",
  show.row =TRUE,drop.empty=TRUE)
```

Table Health insurance by whether one has madicaid

Covered by Medicaid:Recode

HINOTCOVE

Total

has health insurancecoverage

no health insurancecoverage

dont know

No

2552891.8 %

22548.1 %

170.1 %

27799100 %

Yes, information

3300100 %

00 %

00 %

3300100 %

Yes, but no information

23100 %

00 %

00 %

23100 %

Unknown-refused

1140.7 %

00 %

1659.3 %

27100 %

Unknown-don't know

933.3 %

00 %

1866.7 %

27100 %

Total

2887192.6 %

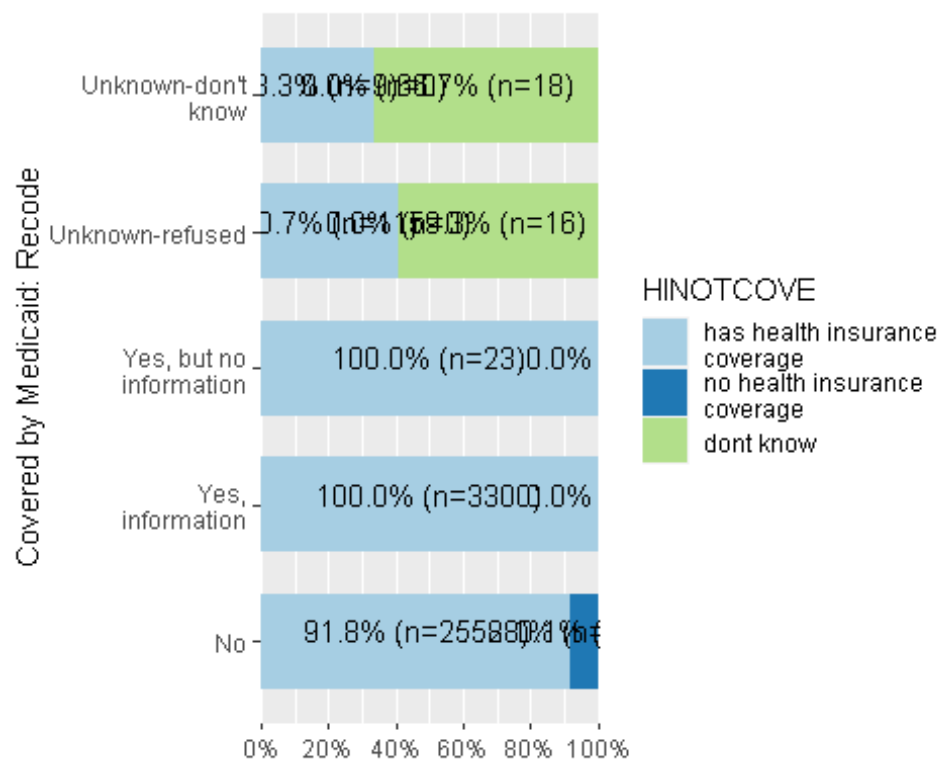
22547.2 %

510.2 %

31176100 %

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

```
sjPlot::plot_xtab(mydata$HIMCAIDE, mydata$HINOTCOVE,
  margin="row", bar.pos="stack", coord.flip=TRUE)
```



#Logistic Regression

Logistic Regression

The dependent variable health insurance “HINOTCOVE” has 3 levels, thus will be recoded

to include 2 levels in order to fit a logistic regression model

```
library(tidyverse)
library(haven)
library(stargazer)
library(ggeffects)
library(gridExtra)
```

The dependent variable health insurance “HINOTCOVE” has 3 levels, thus will be recoded

to include 2 levels in order to fit a logistic regression model

```
mydata$H_insurance = ifelse(mydata$HINOTCOVE == "has health insurance coverage", "has insurance coverage", "no insurance") #drop the - HINOTCOVE
colnames(mydata)
```

```
## [1] "REGION"      "PERNUM"      "SAMPWEIGHT"  "LONGWEIGHT"  "PARTWEIGHT"
## [6] "ASTATFLG"    "CSTATFLG"    "AGE"         "SEX"          "SEXORIEN"
## [11] "MARST"       "RACEA"       "HISPETH"     "YRSINUS"     "USBORN"
## [16] "CITIZEN"     "ARMFEV"     "EDUC"        "EMPSTAT"     "HOURSWRK"
## [21] "PAIDSICK"    "EMPHI"      "EMPFT"       "FAMTOTINC"   "HEALTH"
## [26] "HEIGHT"     "WEIGHT"     "BMICALC"     "HINOTCOVE"   "HIPPRIVATEE"
## [31] "HICHIPE"    "HIMILITE"    "HISTATEE"    "HIMCAIDE"    "HIMCAREE"
## [36] "HINOTCOV"    "HIPPRIVATE"  "SMOKEV"      "SMOKFREQNOW" "CVDDIAG"
## [41] "CVDTEST"     "H_insurance"
```

#Lets inspect our variable and change it to a factor

```
table(mydata$H_insurance)
```

```
##
## has insurance coverage      no insurance
##                28871                2305
```

```
class(mydata$H_insurance)
```

```
## [1] "character"  
mydata$H_insurance = as.factor(mydata$H_insurance)
```

Drop irrelevant variables

```
mydata = mydata[, -c(1:9)]  
  
dim(mydata) #this is a very big dataset with 31176 rows and 45 columns  
## [1] 31176    33
```

NOTE: We will use a small sample for the modelling to utilize the small PC RAM

Sample 10% of the data

```
sample_size = floor(0.01*nrow(mydata))  
set.seed(777)  
picked = sample(seq_len(nrow(mydata)), size = sample_size)  
mydata_sample = mydata[picked,]  
dim(mydata_sample)  
## [1] 311    33
```

Feature selection using BORUTA PACKAGE

```
library(Boruta)  
set.seed(123)  
boruta.train <- Boruta(H_insurance~. , data = mydata_sample, doTrace = 2)  
  
## 1. run of importance source...  
## 2. run of importance source...  
## 3. run of importance source...  
## 4. run of importance source...  
## 5. run of importance source...  
## 6. run of importance source...  
## 7. run of importance source...
```

```
## 8. run of importance source...
## 9. run of importance source...
## 10. run of importance source...
## 11. run of importance source...
## 12. run of importance source...
## After 12 iterations, +0.88 secs:
## confirmed 7 attributes: CITIZEN, HIMCAIDE, HIMCAREE, HINOTCOV, HINOTCOVE
and 2 more;
## rejected 6 attributes: CVDDIAG, CVDTEST, HICHIPE, HISTATEE, SEXORIEN and
1 more;
## still have 19 attributes left.
## 13. run of importance source...
## 14. run of importance source...
## 15. run of importance source...
## 16. run of importance source...
## After 16 iterations, +1.1 secs:
## confirmed 1 attribute: HOURSWRK;
## rejected 4 attributes: HEIGHT, HIMILITE, HISPETH, SMOKFREQNOW;
## still have 14 attributes left.
## 17. run of importance source...
## 18. run of importance source...
## 19. run of importance source...
## 20. run of importance source...
## After 20 iterations, +1.5 secs:
## rejected 2 attributes: ARMFEV, HEALTH;
## still have 12 attributes left.
```

```
## 21. run of importance source...
## 22. run of importance source...
## 23. run of importance source...
## After 23 iterations, +1.7 secs:
## rejected 1 attribute: YRSINUS;
## still have 11 attributes left.
## 24. run of importance source...
## 25. run of importance source...
## 26. run of importance source...
## After 26 iterations, +1.8 secs:
## rejected 2 attributes: RACEA, SMOKEV;
## still have 9 attributes left.
## 27. run of importance source...
## 28. run of importance source...
## 29. run of importance source...
## After 29 iterations, +2 secs:
## rejected 2 attributes: FAMTOTINC, PAIDSICK;
## still have 7 attributes left.
## 30. run of importance source...
## 31. run of importance source...
## 32. run of importance source...
## After 32 iterations, +2.1 secs:
## confirmed 1 attribute: EMPHI;
## still have 6 attributes left.
## 33. run of importance source...
## 34. run of importance source...
```



```
## 35. run of importance source...
## 36. run of importance source...
## 37. run of importance source...
## 38. run of importance source...
## 39. run of importance source...
## 40. run of importance source...
## 41. run of importance source...
## After 41 iterations, +2.5 secs:
## confirmed 1 attribute: EDUC;
## rejected 1 attribute: MARST;
## still have 4 attributes left.
## 42. run of importance source...
## 43. run of importance source...
## 44. run of importance source...
## 45. run of importance source...
## 46. run of importance source...
## 47. run of importance source...
## 48. run of importance source...
## 49. run of importance source...
## 50. run of importance source...
## 51. run of importance source...
## 52. run of importance source...
## 53. run of importance source...
## 54. run of importance source...
## 55. run of importance source...
## 56. run of importance source...
```

```
## 57. run of importance source...
## 58. run of importance source...
## 59. run of importance source...
## 60. run of importance source...
## 61. run of importance source...
## 62. run of importance source...
## 63. run of importance source...
## 64. run of importance source...
## 65. run of importance source...
## 66. run of importance source...
## 67. run of importance source...
## 68. run of importance source...
## 69. run of importance source...
## 70. run of importance source...
## 71. run of importance source...
## 72. run of importance source...
## 73. run of importance source...
## 74. run of importance source...
## 75. run of importance source...
## 76. run of importance source...
## 77. run of importance source...
## 78. run of importance source...
## 79. run of importance source...
## 80. run of importance source...
## 81. run of importance source...
## 82. run of importance source...
```

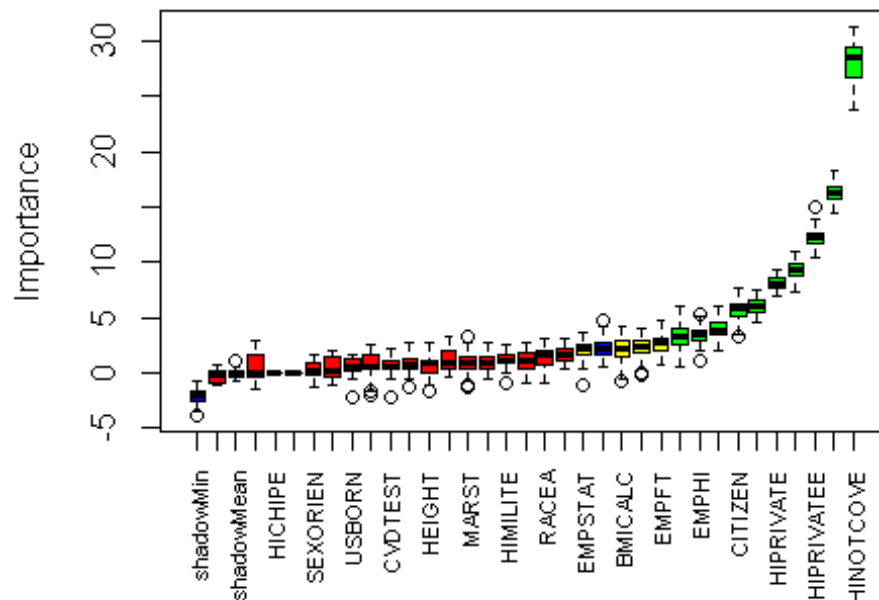
```
## 83. run of importance source...
## 84. run of importance source...
## 85. run of importance source...
## 86. run of importance source...
## 87. run of importance source...
## 88. run of importance source...
## 89. run of importance source...
## 90. run of importance source...
## 91. run of importance source...
## 92. run of importance source...
## 93. run of importance source...
## 94. run of importance source...
## 95. run of importance source...
## 96. run of importance source...
## 97. run of importance source...
## 98. run of importance source...
## 99. run of importance source...

print(boruta.train)

## Boruta performed 99 iterations in 4.967409 secs.
## 10 attributes confirmed important: CITIZEN, EDUC, EMPHI, HIMCAIDE,
## HIMCAREE and 5 more;
## 18 attributes confirmed unimportant: ARMFEV, CVDDIAG, CVDTEST,
## FAMTOTINC, HEALTH and 13 more;
## 4 tentative attributes left: BMICALC, EMPFT, EMPSTAT, WEIGHT;

plot(boruta.train, xlab = "", xaxt = "n")
lz<-lapply(1:ncol(boruta.train$ImpHistory),function(i)
  boruta.train$ImpHistory[is.finite(boruta.train$ImpHistory[,i]),i])
names(lz) <- colnames(boruta.train$ImpHistory)
Labels <- sort(sapply(lz,median))
```

```
axis(side = 1, las=2, labels = names(Labls),
     at = 1:ncol(boruta.train$ImpHistory), cex.axis = 0.7)
```



```
final.boruta <- TentativeRoughFix(boruta.train)
print(final.boruta)

## Boruta performed 99 iterations in 4.967409 secs.
## Tentatives roughfixed over the last 99 iterations.
## 13 attributes confirmed important: BMICALC, CITIZEN, EDUC, EMPHI,
## EMPSTAT and 8 more;
## 19 attributes confirmed unimportant: ARMFEV, CVDDIAG, CVDTEST, EMPFT,
## FAMTOTINC and 14 more;
```

```
boruta.df <- attStats(final.boruta)
head(boruta.df)
```

```
##          meanImp  medianImp  minImp  maxImp  normHits decision
## SEXORIEN 0.2741338 0.1380992 -1.339645 1.683586 0.00000000 Rejected
## MARST    0.9675209 0.9379554 -1.360392 3.345252 0.09090909 Rejected
## RACEA    1.4581311 1.6416994 -1.002576 3.045947 0.04040404 Rejected
## HISPETH  0.5173526 -0.0742569 -1.538081 2.938490 0.01010101 Rejected
## YRSINUS  0.7580682 0.6467762 -1.342705 2.698256 0.03030303 Rejected
## USBORN   0.4734069 0.5912763 -2.188271 1.718444 0.00000000 Rejected
```

```

features=getSelectedAttributes(final.boruta, withTentative = F)
features

## [1] "CITIZEN"      "EDUC"         "EMPSTAT"      "HOURLSWRK"    "EMPHI"
## [6] "WEIGHT"      "BMICALC"      "HINOTCOVE"    "HIPPRIVATEE"  "HIMCAIDE"
## [11] "HIMCAREE"     "HINOTCOV"     "HIPPRIVATE"

final.features=c("bhr", "basedp", "pkhr", "X.mphr.b.", "sbp", "dp", "maxhr", "mbp",
                 "dpmaxdo", "age", "gddpeakdp")
my.features=mydata_sample[,features]

```

Re-introduce the y class variable

```
my.features$H_insurance = mydata_sample$H_insurance
```

Now lets fit a logistic regression model

```

mymodel =glm(H_insurance ~ .,data=my.features,family = "binomial")

## Warning: glm.fit: algorithm did not converge

summary(mymodel)

##
## Call:
## glm(formula = H_insurance ~ ., family = "binomial", data = my.features)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.409e-06 -2.409e-06 -2.409e-06 -2.409e-06  2.409e-06
##
## Coefficients: (2 not defined because of singularities)
##                                Estimate Std. Error z    Pr(>|z|)
alue
## (Intercept)                    -2.657e+01  3.033e+05
0
## CITIZENyes US citizen          -3.576e-08  1.174e+05
0
## CITIZENNA                      -3.715e-08  2.347e+05
0
## EDUCno school                  -4.519e-06  4.262e+05
0
## EDUCless than hs               6.409e-09  2.322e+05
0
## EDUC12th grade no diploma     -3.527e-09  3.250e+05

```

```

0
## EDUCHS diploma -3.183e-09 2.084e+05
0
## EDUCGED -2.300e-08 2.639e+05
0
## EDUCsome college -3.663e-09 2.041e+05
0
## EDUCassoc deg in tech or occ -4.199e-09 2.462e+05
0
## EDUCassoc deg academic -3.236e-09 2.170e+05
0
## EDUCbachelors -3.965e-09 2.075e+05
0
## EDUCmasters -7.843e-09 2.142e+05
0
## EDUCprofessional degree -2.999e-09 2.439e+05
0
## EDUCdoctoral -3.456e-09 2.614e+05
0
## EMPSTATemployed 4.763e-09 1.934e+05
0
## EMPSTATdont know 4.286e-11 1.662e+05
0
## EMPSTATNIU NA NA
NA
## HOURSWRK -8.448e-11 2.272e+03
0
## EMPHIIno workplace did not offer health insurance 1.614e-10 7.360e+04
0
## EMPHIyes workplace offer health insurance NA NA
NA
## WEIGHT 2.239e-12 9.630e+02
0
## BMICALC 7.601e-11 7.059e+03
0
## HINOTCOVno health insurance coverage 5.313e+01 1.354e+05
0
## HIPRIVATEE -6.141e-10 8.250e+04
0
## HIMCAIDE -1.836e-09 8.529e+04
0
## HIMCAREE 9.427e-10 6.399e+04
0
## HINOTCOV 1.051e-07 1.445e+05
0

```

```
## HIPRIVATE -1.851e-10 8.037e+04
0
## Pr(>|z|)
## (Intercept) 1
## CITIZENyes US citizen 1
## CITIZENNA 1
## EDUCno school 1
## EDUCless than hs 1
## EDUC12th grade no diploma 1
## EDUCHS diploma 1
## EDUCGED 1
## EDUCsome college 1
## EDUCassoc deg in tech or occ 1
## EDUCassoc deg academic 1
## EDUCbachelors 1
## EDUCmasters 1
## EDUCprofessional degree 1
## EDUCdoctoral 1
## EMPSTATemployed 1
## EMPSTATdont know 1
## EMPSTATNIU NA
## HOURSWRK 1
## EMPHIno workplace did not offer health insurance 1
## EMPHIyes workplace offer health insurance NA
## WEIGHT 1
## BMICALC 1
## HINOTCOVno health insurance coverage 1
## HIPRIVATEE 1
## HIMCAIDE 1
## HIMCAREE 1
## HINOTCOV 1
## HIPRIVATE 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 2.0613e+02 on 310 degrees of freedom
## Residual deviance: 1.8043e-09 on 284 degrees of freedom
## AIC: 54
##
## Number of Fisher Scoring iterations: 25
```

Now lets fit a logistic regression model

```
mymodel = glm(H_insurance ~.
, data=my.features, family = "binomial")
```

```
## Warning: glm.fit: algorithm did not converge
summary(mymodel)

##
## Call:
## glm(formula = H_insurance ~ ., family = "binomial", data = my.features)
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.409e-06 -2.409e-06 -2.409e-06 -2.409e-06  2.409e-06
##
## Coefficients: (2 not defined because of singularities)
##                                     Estimate Std. Error z v
alue
## (Intercept)                      -2.657e+01  3.033e+05
0
## CITIZENyes US citizen              -3.576e-08  1.174e+05
0
## CITIZENNA                        -3.715e-08  2.347e+05
0
## EDUCno school                     -4.519e-06  4.262e+05
0
## EDUCless than hs                  6.409e-09  2.322e+05
0
## EDUC12th grade no diploma         -3.527e-09  3.250e+05
0
## EDUCHS diploma                   -3.183e-09  2.084e+05
0
## EDUCGED                          -2.300e-08  2.639e+05
0
## EDUCsome college                 -3.663e-09  2.041e+05
0
## EDUCassoc deg in tech or occ      -4.199e-09  2.462e+05
0
## EDUCassoc deg academic            -3.236e-09  2.170e+05
0
## EDUCbachelors                    -3.965e-09  2.075e+05
0
## EDUCmasters                      -7.843e-09  2.142e+05
0
## EDUCprofessional degree          -2.999e-09  2.439e+05
0
## EDUCdoctoral                    -3.456e-09  2.614e+05
0
## EMPSTATemployed                   4.763e-09  1.934e+05
```



```

0
## EMPSTATdont know          4.286e-11  1.662e+05
0
## EMPSTATNIU                NA          NA
NA
## HOURSWRK                  -8.448e-11  2.272e+03
0
## EMPHIIno workplace did not offer health insurance 1.614e-10  7.360e+04
0
## EMPHIyes workplace offer health insurance        NA          NA
NA
## WEIGHT                    2.239e-12  9.630e+02
0
## BMICALC                   7.601e-11  7.059e+03
0
## HINOTCOVEno health insurance coverage          5.313e+01  1.354e+05
0
## HIPRIVATEEE               -6.141e-10  8.250e+04
0
## HIMCAIDE                  -1.836e-09  8.529e+04
0
## HIMCAREE                   9.427e-10  6.399e+04
0
## HINOTCOV                   1.051e-07  1.445e+05
0
## HIPRIVATE                 -1.851e-10  8.037e+04
0
##
## Pr(>|z|)
## (Intercept)                1
## CITIZENyes US citizen       1
## CITIZENNA                   1
## EDUCno school               1
## EDUCless than hs            1
## EDUC12th grade no diploma   1
## EDUCHS diploma              1
## EDUCGED                     1
## EDUCsome college            1
## EDUCassoc deg in tech or occ 1
## EDUCassoc deg academic      1
## EDUCbachelors               1
## EDUCmasters                 1
## EDUCprofessional degree     1
## EDUCdoctoral                1
## EMPSTATemployed             1
## EMPSTATdont know           1

```

```
## EMPSTATNIU NA
## HOURSWRK 1
## EMPHIIno workplace did not offer health insurance 1
## EMPHIyes workplace offer health insurance NA
## WEIGHT 1
## BMICALC 1
## HINOTCOVEno health insurance coverage 1
## HIPRIVATEE 1
## HIMCAIDE 1
## HIMCAREE 1
## HINOTCOV 1
## HIPRIVATE 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
## Null deviance: 2.0613e+02 on 310 degrees of freedom
## Residual deviance: 1.8043e-09 on 284 degrees of freedom
## AIC: 54
##
## Number of Fisher Scoring iterations: 25
```

Odds ratios

```
exp(mymodel$coefficients[c(3,29,69,70,72:76)])
```

```
## CITIZENNA HIPRIVATE <NA> <NA> <NA> <NA> <NA>
<NA>
## 1 1 NA NA NA NA NA
NA
## <NA>
## NA
```

Print the model

```
tab_model(mymodel)
```

```
## Warning: Model matrix is rank deficient. Parameters EMPSTATNIU, EMPHIyes w
orkplace offer
## health insurance were not estimable.

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred

## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: algorithm did not converge
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
```

[illegible]

[illegible]

[illegible]

[illegible]

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning in regularize.values(x, y, ties, missing(ties), na.rm = na.rm):
## collapsing to unique 'x' values
```

H_insurance

Predictors

Odds Ratios

CI

p

(Intercept)

0.00

0.00 – Inf

1.000

CITIZEN: yes US citizen

1.00

0.00 – Inf

1.000

CITIZEN: NA

1.00

0.00 – Inf

1.000

EDUC: no school

1.00

0.00 – Inf

1.000

EDUC: less than hs

1.00

0.00 – Inf

1.000

EDUC: 12th grade nodiploma

1.00

0.00 – Inf

1.000

EDUC: HS diploma

1.00

0.00 – Inf

1.000

EDUC: GED

1.00

0.00 – Inf

1.000

EDUC: some college

1.00

0.00 – Inf

1.000

EDUC: assoc deg in techor occ

1.00

0.00 – Inf

1.000

EDUC: assoc deg academic

1.00

0.00 – Inf

1.000

EDUC: bachelors

1.00

0.00 – Inf

1.000

EDUC: masters

1.00

0.00 – Inf

1.000

EDUC: professional degree

1.00

0.00 – Inf

1.000

EDUC: doctoral

1.00

0.00 – Inf

1.000

EMPSTAT: employed

1.00

0.00 – Inf

1.000

EMPSTAT: dont know

1.00

0.00 – Inf

1.000

Total hours worked lastweek or usually

1.00

0.00 – Inf

1.000

EMPHI: no workplace didnt offer healthinsurance

1.00

0.00 – Inf

1.000

Weight in pounds withoutshoes

1.00

0.00 – Inf

1.000

Body Mass Index,calculated from publiclyreleased height andweight variables

1.00

0.00 – Inf

1.000

Has no health insurance(excluding single serviceplans): HINOTCOVE: nohealth insurance coverage

118848596401786454026000.00

0.00 – Inf

1.000

Covered by private healthinsurance: Recode

1.00

0.00 – Inf

1.000

Covered by Medicaid:Recode

1.00

0.00 – Inf

1.000

Covered by Medicare:Recode

1.00

0.00 – Inf

1.000

Has no health insurance(excluding single serviceplans)

1.00

0.00 – Inf

1.000

Has any private healthinsurance

1.00

0.00 – Inf

1.000

Observations

311

R2 Tjur

1.000

Odds ratios

```
exp(mymodel$coefficients[c(3,29,69,70,72:76)])
```

## CITIZENNA HIPRIVATE	<NA>	<NA>	<NA>	<NA>	<NA>
<NA>					
##	1	1	NA	NA	NA

Print the model

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning: glm.fit: fitted probabilities numerically 0 or 1 occurred
## Warning in regularize.values(x, y, ties, missing(ties), na.rm = na.rm):
## collapsing to unique 'x' values
```

H_insurance

Predictors

Odds Ratios

CI

p

(Intercept)

0.00

0.00 – Inf

1.000

CITIZEN: yes US citizen

1.00

0.00 – Inf

1.000

CITIZEN: NA

1.00

0.00 – Inf

1.000

EDUC: no school

1.00

0.00 – Inf

1.000

EDUC: less than hs

1.00

0.00 – Inf

1.000

EDUC: 12th grade nodiploma

1.00

0.00 – Inf

1.000

EDUC: HS diploma

1.00

0.00 – Inf

1.000

EDUC: GED

1.00

0.00 – Inf

1.000

EDUC: some college

1.00

0.00 – Inf

1.000

EDUC: assoc deg in techor occ

1.00

0.00 – Inf

1.000

EDUC: assoc deg academic

1.00

0.00 – Inf

1.000

EDUC: bachelors

1.00

0.00 – Inf

1.000

EDUC: masters

1.00

0.00 – Inf

1.000

EDUC: professional degree

1.00

0.00 – Inf

1.000

EDUC: doctoral

1.00

0.00 – Inf

1.000

EMPSTAT: employed

1.00

0.00 – Inf

1.000

EMPSTAT: dont know

1.00

0.00 – Inf

1.000

Total hours worked lastweek or usually

1.00

0.00 – Inf

1.000

EMPHI: no workplace didnt offer healthinsurance

1.00

0.00 – Inf

1.000

Weight in pounds withoutshoes

1.00

0.00 – Inf

1.000

Body Mass Index,calculated from publiclyreleased height andweight variables

1.00

0.00 – Inf

1.000

Has no health insurance(excluding single serviceplans): HINOTCOVE: nohealth insurance coverage

118848596401786454026000.00

0.00 – Inf

1.000

Covered by private healthinsurance: Recode

1.00

0.00 – Inf

1.000

Covered by Medicaid:Recode

1.00

0.00 – Inf

1.000

Covered by Medicare:Recode

1.00

0.00 – Inf

1.000

Has no health insurance(excluding single serviceplans)

1.00

0.00 – Inf

1.000

Has any private healthinsurance

1.00

0.00 – Inf

1.000

```
## (Intercept) data$RACEABlack
## no health insurance coverage -45.74414 1.329015
## dont know -19.94692 1.761993
## data$RACEAAleut Alaskan data$RACEAAmerican In
dian
## no health insurance coverage 6.6631456 5.409
2575
## dont know -0.1460416 -0.109
```

```

1718
##                                data$RACEAsian data$RACEAOther data$RACEAref
used
## no health insurance coverage      -6.6641692   -1.453370e+00    2.064651
e+00
## dont know                        0.1058754   -1.961201e+06    -1.789056
e+06
##                                data$RACEAnot ascertained data$RACEAunknown
## no health insurance coverage          8.5302148    8.732779e-01
## dont know                          0.9850883    -2.477825e+05
##                                data$YRSINUSLess than 1 year in US
## no health insurance coverage          4.288396e+00
## dont know                          -1.902456e+06
##                                data$YRSINUS1-5 years in US
## no health insurance coverage          16.1206115
## dont know                          0.4821026
##                                data$YRSINUS5-10 years in US
## no health insurance coverage          10.09047
## dont know                          -38357.62413
##                                data$YRSINUS10-15 yr in US
## no health insurance coverage          12.6532509
## dont know                          0.2970035
##                                data$YRSINUS15 or more yr in US data$YRSINUSN
A
## no health insurance coverage          12.4268147    3.64637
0
## dont know                          -0.7559886    2.38428
9
##                                data$SEXFemale data$SEXRefused data$SEXdont k
now
## no health insurance coverage      -2.813519   -1.124155e+08    21109
351
## dont know                        -1.021459    3.828451e+00    -1244
992
##                                data$MARSTMarried
## no health insurance coverage      -1.2204850
## dont know                        -0.1593693
##                                data$MARSTMarried spouse not there
## no health insurance coverage          1.364185
## dont know                          -32.273536
##                                data$MARSTMarried spouse NA data$MARSTWidowed
## no health insurance coverage          1.972824    -8.6248847
## dont know                        -19.543449    -0.1252851
##                                data$MARSTDivorced data$MARSTSeparated
## no health insurance coverage          3.6775019    4.82337533

```



```

## dont know                -0.1478439        -0.08888516
##                          data$MARSTnever married data$MARSTunknown
## no health insurance coverage      9.8333579398        3.513988
## dont know                0.0005941285        0.113115
##                          data$EDUCno school data$EDUCless than hs
## no health insurance coverage      4.4487949        10.6817782
## dont know                -0.5425468        0.1639214
##                          data$EDUC12th grade no diploma data$EDUCHS di
ploma
## no health insurance coverage      4.56321905        7.354
38194
## dont know                0.06153474        0.085
08369
##                          data$EDUCGED data$EDUCsome college
## no health insurance coverage      6.346087        3.6832674
## dont know                -40.567672        0.1177433
##                          data$EDUCassoc deg in tech or occ
## no health insurance coverage      3.83659977
## dont know                0.04193598
##                          data$EDUCassoc deg academic data$EDUCbachelor
s
## no health insurance coverage      0.73009801        -3.3711346
2
## dont know                -0.05749649        -0.0123675
4
##                          data$EDUCmasters data$EDUCprofessional degree
## no health insurance coverage      -7.1777932        -2.642032
## dont know                -0.0351558        -2.550881
##                          data$EDUCdoctoral data$EDUCrefused
## no health insurance coverage      -4.718972        0.2565127
## dont know                -108.151662        0.3794169
##                          data$EDUCdont know
## no health insurance coverage      4.5593132
## dont know                0.1716821

p <- (1 - pnorm(abs(z), 0, 1)) * 2
pp <- (1 - pnorm(abs(z), 0, 1)) * 2
p

##                          (Intercept) data$RACEABlack
## no health insurance coverage      0        0.18384291
## dont know                0        0.07807054
##                          data$RACEAAleut Alaskan data$RACEAAmerican In
dian
## no health insurance coverage      2.680278e-11        6.328659
e-08

```

```

## dont know                                8.838886e-01          9.130662
e-01
##                                data$RACEAsian data$RACEAOther data$RACEAref
used
## no health insurance coverage    2.661671e-11          0.1461209          0.0389
5599
## dont know                        9.156812e-01          0.0000000          0.0000
0000
##                                data$RACEAnot ascertained data$RACEAunknown
## no health insurance coverage    0.0000000          0.3825116
## dont know                      0.3245807          0.0000000
##                                data$YRSINUSLess than 1 year in US
## no health insurance coverage    1.799681e-05
## dont know                      0.000000e+00
##                                data$YRSINUS1-5 years in US
## no health insurance coverage    0.0000000
## dont know                      0.6297331
##                                data$YRSINUS5-10 years in US
## no health insurance coverage    0
## dont know                      0
##                                data$YRSINUS10-15 yr in US
## no health insurance coverage    0.0000000
## dont know                      0.7664639
##                                data$YRSINUS15 or more yr in US data$YRSINUSN
A
## no health insurance coverage    0.000000          0.000265971
3
## dont know                      0.449656          0.017112147
5
##                                data$SEXFemale data$SEXRefused data$SEXdont k
now
## no health insurance coverage    0.004900256          0.0000000000
0
## dont know                      0.307036880          0.0001289526
0
##                                data$MARSTMarried
## no health insurance coverage    0.2222811
## dont know                      0.8733779
##                                data$MARSTMarried spouse not there
## no health insurance coverage    0.1725095
## dont know                      0.0000000
##                                data$MARSTMarried spouse NA data$MARSTWidowed
## no health insurance coverage    0.0485156          0.0000000
## dont know                      0.0000000          0.9002978
##                                data$MARSTDivorced data$MARSTSeparated

```

```

## no health insurance coverage      0.0002355294      1.411490e-06
## dont know                        0.8824659839      9.291732e-01
##                                data$MARSTnever married data$MARSTunknown
## no health insurance coverage      0.000000      0.000441433
## dont know                        0.999526      0.909939410
##                                data$EDUCno school data$EDUCless than hs
## no health insurance coverage      8.635344e-06      0.0000000
## dont know                        5.874419e-01      0.8697931
##                                data$EDUC12th grade no diploma data$EDUCHS di
ploma
## no health insurance coverage      5.037521e-06      1.91846
5e-13
## dont know                        9.509334e-01      9.32194
9e-01
##                                data$EDUCGED data$EDUCsome college
## no health insurance coverage 2.208609e-10      0.0002302634
## dont know                        0.000000e+00      0.9062710881
##                                data$EDUCassoc deg in tech or occ
## no health insurance coverage      0.0001247495
## dont know                        0.9665497361
##                                data$EDUCassoc deg academic data$EDUCbachelor
s
## no health insurance coverage      0.4653303      0.000748592
7
## dont know                        0.9541497      0.990132382
2
##                                data$EDUCmasters data$EDUCprofessional degree
## no health insurance coverage      7.085443e-13      0.008241032
## dont know                        9.719555e-01      0.010745091
##                                data$EDUCdoctoral data$EDUCrefused
## no health insurance coverage      2.37039e-06      0.7975550
## dont know                        0.00000e+00      0.7043783
##                                data$EDUCdont know
## no health insurance coverage      5.132119e-06
## dont know                        8.636875e-01

exp(coef(test))

##                                (Intercept) data$RACEABlack
## no health insurance coverage 0.042447274      1.089400
## dont know                        0.002770128      1.652603
##                                data$RACEAAleut Alaskan data$RACEAAmerican In
dian
## no health insurance coverage      2.8710855668      2.353683
2968
## dont know                        0.0005218728      0.000226

```

```

6321
##                                data$RACEAAsian data$RACEAOther data$RACEAref
used
## no health insurance coverage      0.4771786      7.740905e-01      3.363631
e+00
## dont know                        1.0604728      1.634167e-06      7.862508
e-06
##                                data$RACEAnot ascertained data$RACEAunknown
## no health insurance coverage      1.843599      1.9411332690
## dont know                        1.520761      0.0004283714
##                                data$YRSINUSLess than 1 year in US
## no health insurance coverage      7.490778e+00
## dont know                        3.621683e-05
##                                data$YRSINUS1-5 years in US
## no health insurance coverage      8.090084
## dont know                        1.643392
##                                data$YRSINUS5-10 years in US
## no health insurance coverage      4.077625e+00
## dont know                        6.660412e-05
##                                data$YRSINUS10-15 yr in US
## no health insurance coverage      5.025370
## dont know                        1.355805
##                                data$YRSINUS15 or more yr in US data$YRSINUSN
A
## no health insurance coverage      2.2694318      2.54270
8
## dont know                        0.6597712      6.28381
4
##                                data$SEXFemale data$SEXRefused data$SEXdont k
now
## no health insurance coverage      0.8879198      1.922209e-06      2.980063e
+05
## dont know                        0.7914009      8.226805e+01      4.879618e
-01
##                                data$MARSTMarried
## no health insurance coverage      0.8991283
## dont know                        0.2342556
##                                data$MARSTMarried spouse not there
## no health insurance coverage      1.233032e+00
## dont know                        1.009964e-06
##                                data$MARSTMarried spouse NA data$MARSTWidowed
## no health insurance coverage      9.345396578      0.2855325
## dont know                        0.000242612      0.3192789
##                                data$MARSTDivorced data$MARSTSeparated
## no health insurance coverage      1.412167      2.012352e+00

```

```

## dont know          0.259964      2.830851e-05
##                    data$MARSTnever married data$MARSTunknown
## no health insurance coverage      2.331157      1.575393
## dont know          1.005424      2.801098
##                    data$EDUCno school data$EDUCless than hs
## no health insurance coverage      3.8066804869      2.732029
## dont know          0.0007055731      4.449634
##                    data$EDUC12th grade no diploma data$EDUCHS di
ploma
## no health insurance coverage      2.047458      1.8
20118
## dont know          1.756039      2.1
69815
##                    data$EDUCGED data$EDUCsome college
## no health insurance coverage 2.336628e+00      1.381742
## dont know          2.892928e-06      2.921221
##                    data$EDUCassoc deg in tech or occ
## no health insurance coverage      1.601143
## dont know          1.466250
##                    data$EDUCassoc deg academic data$EDUCbachelor
s
## no health insurance coverage      1.0784457      0.736564
6
## dont know          0.5917268      0.893468
3
##                    data$EDUCmasters data$EDUCprofessional degree
## no health insurance coverage      0.4122106      5.094354e-01
## dont know          0.7257462      3.536364e-05
##                    data$EDUCdoctoral data$EDUCrefused
## no health insurance coverage      2.555087e-01      1.162746
## dont know          1.231730e-06      31.855801
##                    data$EDUCdont know
## no health insurance coverage      3.10956
## dont know          4.81290

head(pp <- fitted(test))

##   has health insurance coverage no health insurance coverage   dont know
## 1          0.9690415          0.03072019 2.382906e-04
## 2          0.9840552          0.01548139 4.634406e-04
## 3          0.9830292          0.01697074 1.785294e-08
## 4          0.6961160          0.28986741 1.401658e-02
## 5          0.9530880          0.04168275 5.229219e-03
## 6          0.9721090          0.02732734 5.636169e-04

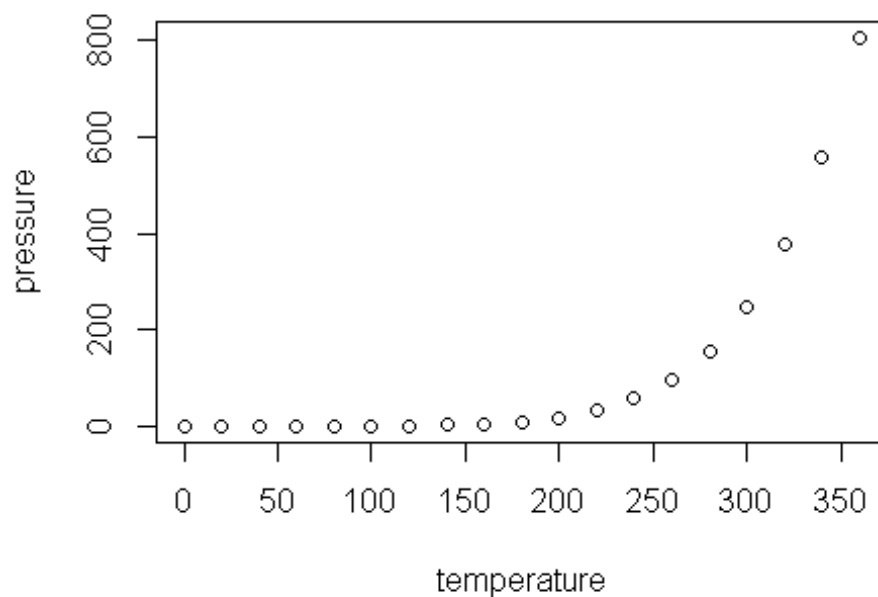
```

```
dses <- data.frame(insurance = c("has health insurance coverage", "Employed"),
                      write = mean(mydata$write))

## Warning: Unknown or uninitialised column: `write`.
## Warning in mean.default(mydata$write): argument is not numeric or logical:
## returning NA
```

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.