

Disha_Project

Setup

```
library(knitr)

## Warning: package 'knitr' was built under R version 3.5.3

knitr::opts_chunk$set(echo = FALSE)
#knitr::opts_knit$set(root.dir= 'Users\AlstonQ\Documents\DR\ ')
knitr::opts_chunk$set(error = FALSE)
```

Load packages

Load data

Part 1: Data

In this case, the BRFSS Survey takes place every year from all 50 states and the District of Columbia, Puerto Rico, and Guam and collect survey data over a limited point- in-time (usually one to three months) from American Samoa, Federated States of Micronesia, and Palau. This makes the data generalized. The survey is conducted through both landline telephone and cellular telephone-based. In conducting the BRFSS landline telephone survey, interviewers collect data from a randomly selected adult in a household. In conducting the cellular telephone version of the BRFSS questionnaire, interviewers collect data from an adult who participates by using a cellular telephone and resides in a private residence or college housing. This makes the data collection random.

Part 2: Research questions

Research question 1: Does the number of responses for general health of household depend on the last time they had a checkup?

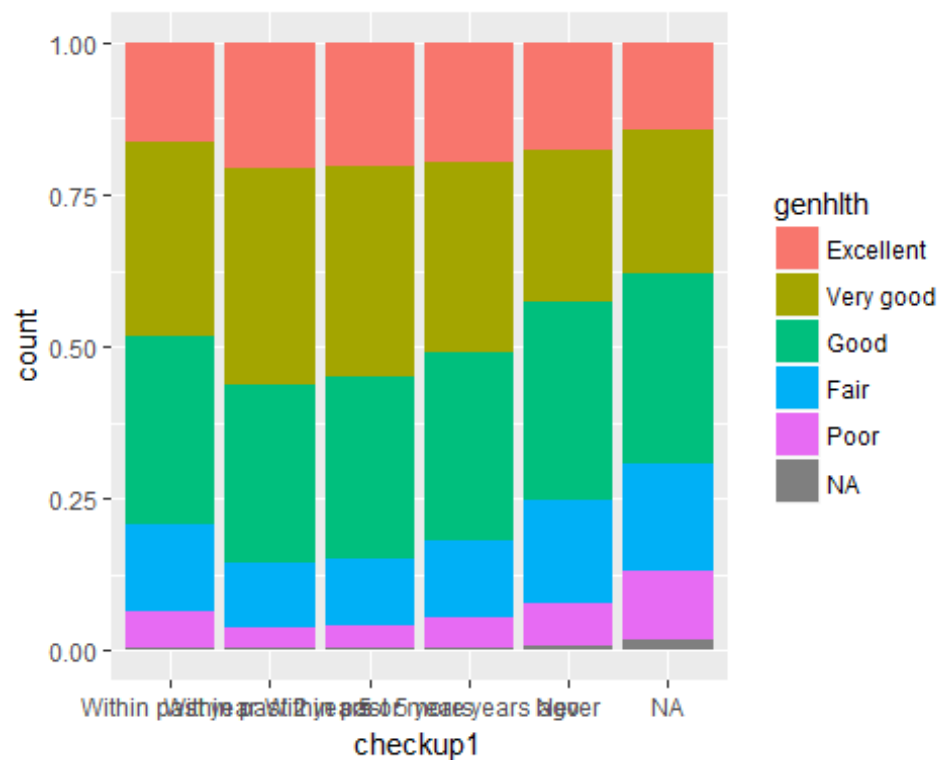
Research question 2: People with children at home and how often they smoke and drink.

Research question 3: A person being detected with cholestrol as compared to his exercises and food?

Part 3: Exploratory data analysis

Research question 1: Does the number of responses for general health of household depend on the last time they had a checkup?

The answer to question seems to be a no. As the amount of people who answer either of the options seem to be fairly equal even though the last time they had a check up was 2 or 5 years ago or even never. This tells us how unreliable the data is until further investigation.



Research question 2: People with children at home and how often they smoke and drink.

The table shows the number of people who smoke(3 categories) and drink(0 or more than 100 drinks) with number of children in household. This data can help child services to identify children who might be affected by the adults and can be a victim of domestic violence.

##	smokday2	Every day	Some days	Not at all
## children	alccount			
## 0	0	18394	6456	51435
##	100	19352	7923	58744
## 1	0	3211	1090	4614
##	100	3995	1780	6580
## 2	0	2334	817	3204
##	100	3081	1423	5931
## 3	0	1117	356	1420
##	100	1345	630	2178
## 4	0	443	152	555
##	100	445	208	596
## 5	0	147	53	175
##	100	154	65	165
## 6	0	52	21	81
##	100	58	28	57

## 7	0	17	8	18
##	100	11	6	16
## 8	0	7	5	12
##	100	4	0	8
## 9	0	2	2	3
##	100	2	2	3
## 10	0	0	2	2
##	100	6	2	0
## 11	0	0	1	0
##	100	0	0	0
## 12	0	0	0	0
##	100	2	0	1
## 13	0	0	0	1
##	100	1	0	1
## 15	0	1	0	0
##	100	0	0	0
## 17	0	0	0	1
##	100	0	0	0

Research question 3: A person being detected with cholestrol as compared to his exercises and food?

Here the people who have green vegetables and exercise also tend to have high cholestrol. The food habit and exercise routine doesnt have an affect any result of cholestrol as the data is equally divided into both categories..

##		exerany2	Yes	No
## toldhi2	greens			
## Yes	bad		7023	6865
##	good		111250	45505
## No	bad		8202	6500
##	good		157943	45057