The Behavioral Risk Factor Surveillance System (BRFSS) is an annual telephone survey of 350,000 people in the United States. We will load a random sample of 20,000 people from the BRFSS survey conducted in 2000.

source("http://www.openintro.org/stat/data/cdc.R")

* Investigate the file.
  1. What data does it contain?
  2. What kind of variables are there?
  3. What type are the parameters contained?
* Describe the object:
  1. How many parameters
  2. How many cases?
  3. How many males/females?
  4. How many older than 25? 35?
  5. Show the distribution of age as a histogram
* Describe one continuous variable:
  1. Calculate the summary statistics
     1. Mean
     2. Median
* Mode

1. Variance
2. Standard deviation
3. Standard error

* IQR

1. Show the data as
   1. a histogram
   2. a boxplot

* repeat i. and ii. for women only
* Describe one categorical variable:
  1. Create a frequency table
  2. Depict the frequency table using a barplot
* Manipulate the dataset
  1. Calculate BMI
  2. Subset the for men and women in two separate datasets
* Create two histograms for both datasets

1. Compare the means between men and women form these datasets.