# Assigment

Okky Mabruri

13/09/2020

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

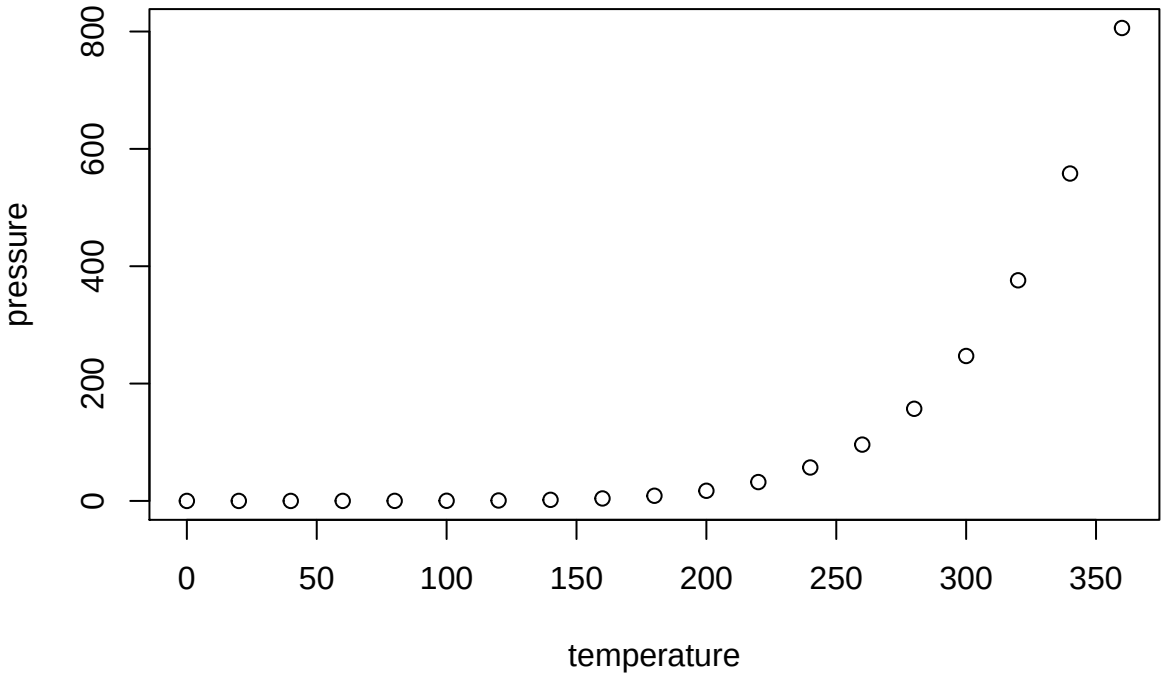
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
summary(cars)
```

```
##      speed           dist       
##  Min.   : 4.0   Min.   :  2.00  
##  1st Qu.:12.0   1st Qu.: 26.00  
##  Median :15.0   Median : 36.00  
##  Mean   :15.4   Mean   : 42.98  
##  3rd Qu.:19.0   3rd Qu.: 56.00  
##  Max.   :25.0   Max.   :120.00
```

## 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.

## set working directory

setwd('ProgAssignment3-data/')

## read csv data

hospitaldata <- read.csv('hospital-data.csv') outcome <- read.csv("outcome-of-care-measures.csv", colClasses = "character")

## check outcome data

head(outcome) nrow(outcome) names(outcome)

## make hist plot from col "Hospital.30.Day.Death..Mortality..Rates.from.Heart..

outcome[, 11] <- as.numeric(outcome[, 11]) hist(outcome[, 11])

## Finding the best hospital in a state

## Handling ties. If there is a tie for the best hospital for a given outcome, then the hospital names should be sorted in alphabetical order and the first hospital in that set should be chosen (i.e. if hospitals "b", "c",and "f" are tied for best, then hospital "b" should be returned).The function should use the following template.

best <- function(state, outcome) { ## Read outcome data outcome <- read.csv("outcome-of-care-measures.csv", colClasses = "character")

```
## Check that state and outcome are valid
if (!(state %in% data$State)) {
    result <- "invalid state"
}


## Return hospital name in that state with lowest 30-day death

## rate
```

}

outcome[, 11] <- sapply(outcome[, 11], as.numeric) outcomeclean <- na.omit(outcome[11])

?na.omit

new <- outcome[outcome[-11] != is.na(outcome[11])]

new <- subset(outcome, colnames(outcome[11]) == 'Not Available')

summary(outcome) sapply(outcome, class) data <- subset(outcome, State == “TX” & colnames(outcome[11]) == numeric(3), select = c(colnames(outcome[2]), colnames(outcome[11]))) data mean(data, na.rm = TRUE) subset(mindata, ) subset(outcome, State == ‘TX’, select=c(City, County.Name))