Practical1

Kent Hervey D. Gener BSIT2-B

2024-03-05

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.

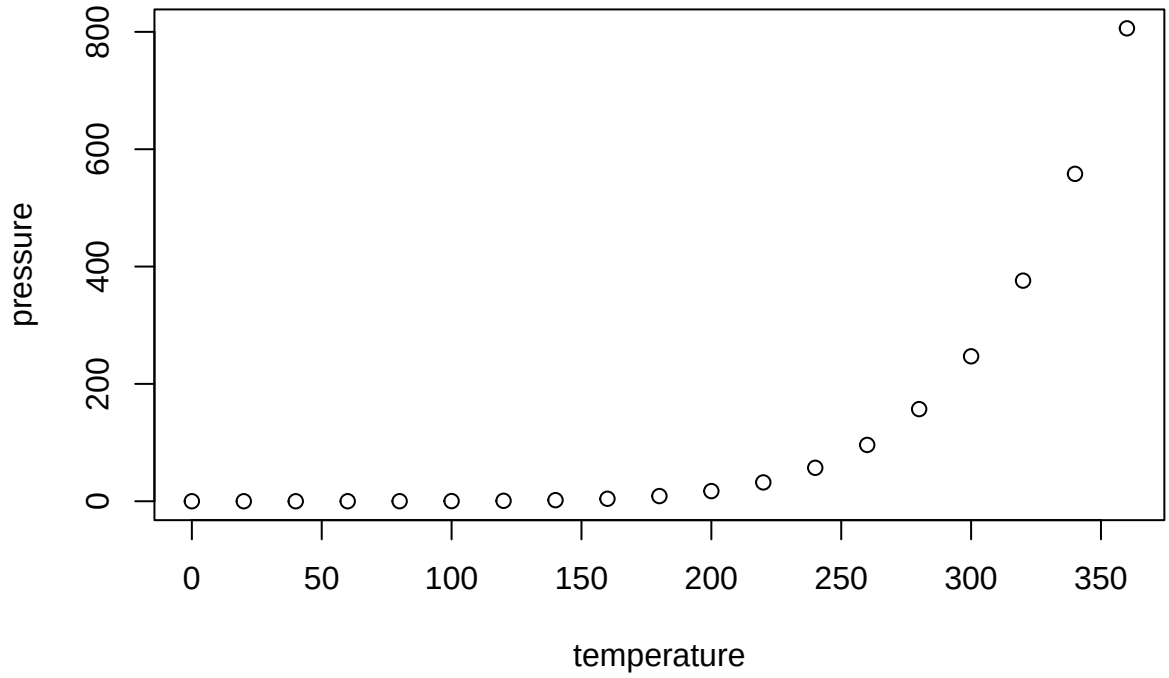
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
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. # A

```
warpbreaks
```

```
##    breaks wool tension
## 1      26    A       L
## 2      30    A       L
## 3      54    A       L
## 4      25    A       L
## 5      70    A       L
## 6      52    A       L
## 7      51    A       L
## 8      26    A       L
## 9      67    A       L
## 10     18    A       M
## 11     21    A       M
## 12     29    A       M
## 13     17    A       M
## 14     12    A       M
## 15     18    A       M
## 16     35    A       M
## 17     30    A       M
## 18     36    A       M
## 19     36    A       H
## 20     21    A       H
```

```
## 21     24    A       H
## 22     18    A       H
## 23     10    A       H
## 24     43    A       H
## 25     28    A       H
## 26     15    A       H
## 27     26    A       H
## 28     27    B       L
## 29     14    B       L
## 30     29    B       L
## 31     19    B       L
## 32     29    B       L
## 33     31    B       L
## 34     41    B       L
## 35     20    B       L
## 36     44    B       L
## 37     42    B       M
## 38     26    B       M
## 39     19    B       M
## 40     16    B       M
## 41     39    B       M
## 42     28    B       M
## 43     21    B       M
## 44     39    B       M
## 45     29    B       M
## 46     20    B       H
## 47     21    B       H
## 48     24    B       H
## 49     17    B       H
## 50     13    B       H
## 51     15    B       H
## 52     15    B       H
## 53     16    B       H
## 54     28    B       H
```

1.

```
numeric_cols <- sapply(warpbreaks, is.numeric)
integer_cols <- sapply(warpbreaks, is.integer)
print(numeric_cols)
```

```
##  breaks    wool tension
##    TRUE   FALSE   FALSE
```

```
print(integer_cols)
```

```
##  breaks    wool tension
##   FALSE   FALSE   FALSE
```

2.

```
warpbreaks[, numeric_cols] <- lapply(warpbreaks[, numeric_cols], as.integer)
```

```
## Warning in `[<-.data.frame`(`*tmp*`, , numeric_cols, value = list(26L, 30L, :
## provided 54 variables to replace 1 variables
```

```
print(warpbreaks)
```

```
##    breaks wool tension
## 1      26    A       L
## 2      26    A       L
## 3      26    A       L
## 4      26    A       L
## 5      26    A       L
## 6      26    A       L
## 7      26    A       L
## 8      26    A       L
## 9      26    A       L
## 10     26    A       M
## 11     26    A       M
## 12     26    A       M
## 13     26    A       M
## 14     26    A       M
## 15     26    A       M
## 16     26    A       M
## 17     26    A       M
## 18     26    A       M
## 19     26    A       H
## 20     26    A       H
## 21     26    A       H
## 22     26    A       H
## 23     26    A       H
## 24     26    A       H
## 25     26    A       H
## 26     26    A       H
## 27     26    A       H
## 28     26    B       L
## 29     26    B       L
## 30     26    B       L
## 31     26    B       L
## 32     26    B       L
## 33     26    B       L
## 34     26    B       L
## 35     26    B       L
## 36     26    B       L
## 37     26    B       M
## 38     26    B       M
## 39     26    B       M
## 40     26    B       M
## 41     26    B       M
```

```
## 42     26    B       M
## 43     26    B       M
## 44     26    B       M
## 45     26    B       M
## 46     26    B       H
## 47     26    B       H
## 48     26    B       H
## 49     26    B       H
## 50     26    B       H
## 51     26    B       H
## 52     26    B       H
## 53     26    B       H
## 54     26    B       H
```

3.

B.

1.

```
linesile <- readLines("D:/CS102/exampleFile.txt")
```

```
## Warning in readLines("D:/CS102/exampleFile.txt"): incomplete final line found
## on 'D:/CS102/exampleFile.txt'
```

```
print(linesile)
```

```
## [1] "// Survey data. Created : 21 May 2013"
## [2] "// Field 1: Gender"
## [3] "// Field 2: Age (in years)"
## [4] "// Field 3: Weight (in kg)"
## [5] "M;28;81.3"
## [6] "male;45;"
## [7] "Female;17;57,2"
## [8] "fem.;64;62.8"
```

2.

```
linesile <- "D:/CS102/exampleFile.txt"
lines <- readLines(linesile)
```

```
## Warning in readLines(linesile): incomplete final line found on
## 'D:/CS102/exampleFile.txt'
```

```r
comments <- lines[grepl("^#", lines)]
data <- lines[!grepl("^#", lines)]

print(comments)
```

```
## character(0)
```

```r
print(data)
```

```
## [1] "// Survey data. Created : 21 May 2013"
## [2] "// Field 1: Gender"
## [3] "// Field 2: Age (in years)"
## [4] "// Field 3: Weight (in kg)"
## [5] "M;28;81.3"
## [6] "male;45;"
## [7] "Female;17;57,2"
## [8] "fem.;64;62.8"
```

3.

```r
date <- regmatches(lines[grepl("^#", lines)][1], regexpr("\\d{4}-\\d{2}-\\d{2}", lines[grepl("^#", lines
print(date)
```

```
## character(0)
```

4.

A.

```r
data_lines <- lines[!grepl("^#", lines)]

data_matrix <- sapply(data_lines, function(line) unlist(strsplit(line, ";")))

print(data_matrix)
```

```
## $`// Survey data. Created : 21 May 2013`
## [1] "// Survey data. Created : 21 May 2013"
##
## $`// Field 1: Gender`
## [1] "// Field 1: Gender"
##
## $`// Field 2: Age (in years)`
## [1] "// Field 2: Age (in years)"
##
```

```
## $`// Field 3: Weight (in kg)`
## [1] "// Field 3: Weight (in kg)"
##
## $`M;28;81.3`
## [1] "M"    "28"   "81.3"
##
## $`male;45;`
## [1] "male" "45"
##
## $`Female;17;57,2`
## [1] "Female" "17"     "57,2"
##
## $`fem.;64;62.8`
## [1] "fem." "64"   "62.8"
```

B.

```
splitdata <- strsplit(lines[!grepl("^#", lines)], ";")

maxfields <- max(lengths(splitdata))

maxnum <- t(sapply(splitdata, function(x) c(x, rep(NA, maxfields - length(x)))))

print(maxnum)
```

```
##      [,1]                                    [,2] [,3]
## [1,] "// Survey data. Created : 21 May 2013" NA   NA
## [2,] "// Field 1: Gender"                    NA   NA
## [3,] "// Field 2: Age (in years)"            NA   NA
## [4,] "// Field 3: Weight (in kg)"            NA   NA
## [5,] "M"                                     "28" "81.3"
## [6,] "male"                                  "45" NA
## [7,] "Female"                                "17" "57,2"
## [8,] "fem."                                  "64" "62.8"
```

C.

```
data <- strsplit(lines[!grepl("^#", lines)], ";")

max_fields <- max(lengths(data))

matrix_data <- matrix(unlist(lapply(data, function(x) c(x, rep(NA, max_fields - length(x))))), ncol = ma

print(matrix_data)
```

```
##      [,1]                                    [,2] [,3]
## [1,] "// Survey data. Created : 21 May 2013" NA   NA
## [2,] "// Field 1: Gender"                    NA   NA
```

```
## [3,] "// Field 2: Age (in years)"            NA   NA
## [4,] "// Field 3: Weight (in kg)"            NA   NA
## [5,] "M"                                     "28" "81.3"
## [6,] "male"                                  "45" NA
## [7,] "Female"                                "17" "57,2"
## [8,] "fem."                                  "64" "62.8"
```

D.

```
comment_lines <- lines[grepl("^#", lines)][2:4]

field_names <- unlist(strsplit(comment_lines, ";"))

colnames(matrix_data) <- field_names

print(matrix_data)
```

```
##      <NA>                                    <NA> <NA>
## [1,] "// Survey data. Created : 21 May 2013" NA   NA
## [2,] "// Field 1: Gender"                    NA   NA
## [3,] "// Field 2: Age (in years)"            NA   NA
## [4,] "// Field 3: Weight (in kg)"            NA   NA
## [5,] "M"                                     "28" "81.3"
## [6,] "male"                                  "45" NA
## [7,] "Female"                                "17" "57,2"
## [8,] "fem."                                  "64" "62.8"
```