

report_test

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>.

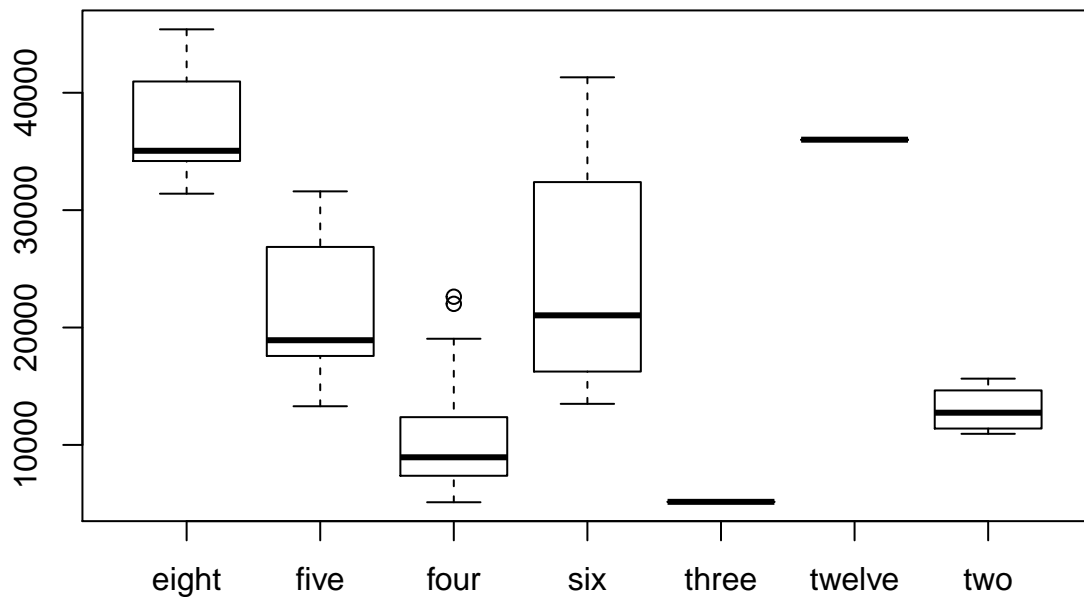
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```
card <- read.csv("CarPrice_Assignment.csv")
coln <- colnames(card)
print(coln)
```

```
## [1] "car_ID"          "symboling"      "CarName"        "fueltype"
## [5] "aspiration"      "doornumber"     "carbody"        "drivewheel"
## [9] "enginelocation"  "wheelbase"      "carlength"      "carwidth"
## [13] "carheight"       "curbweight"     "enginetype"     "cylindernumber"
## [17] "enginesize"      "fuelsystem"     "boreratio"      "stroke"
## [21] "compressionratio" "horsepower"     "peakrpm"        "citympg"
## [25] "highwaympg"      "price"
```

Including Plots

You can also embed plots, for example



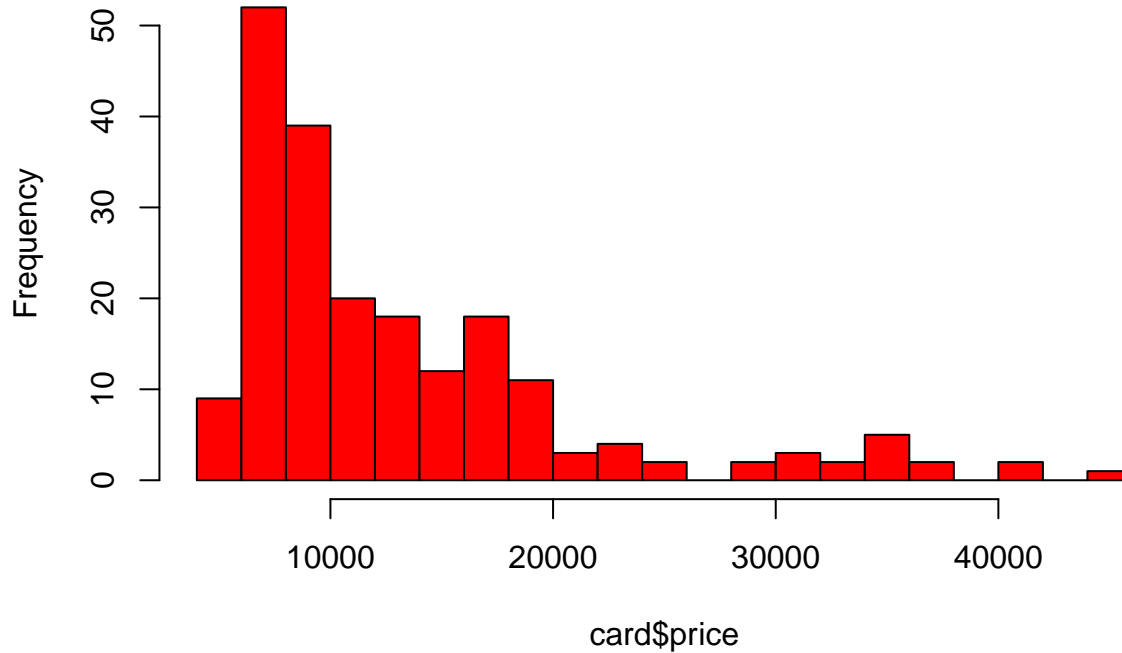
```
summary(card)
```

```
##      car_ID      symboling      CarName      fueltype      aspiration
## Min.   : 1      Min.   : -2.0000      peugeot 504 : 6      diesel: 20      std :168
## 1st Qu.: 52      1st Qu.: 0.0000      toyota corolla: 6      gas :185      turbo: 37
## Median :103      Median : 1.0000      toyota corona : 6
## Mean   :103      Mean   : 0.8341      subaru dl : 4
## 3rd Qu.:154      3rd Qu.: 2.0000      honda civic : 3
## Max.   :205      Max.   : 3.0000      mazda 626 : 3
##                                     (Other) :177
## doornumber      carbody      drivewheel      enginelocation      wheelbase
## four:115      convertible: 6      4wd: 9      front:202      Min. : 86.60
## two : 90      hardtop : 8      fwd:120      rear : 3      1st Qu.: 94.50
##                                     hatchback :70      rwd: 76      Median : 97.00
##                                     sedan :96      Mean : 98.76
##                                     wagon :25      3rd Qu.:102.40
##                                     Max. :120.90
##
##      carlength      carwidth      carheight      curbweight      enginetype
## Min. :141.1      Min. : 60.30      Min. : 47.80      Min. :1488      dohc : 12
## 1st Qu.:166.3      1st Qu.:64.10      1st Qu.:52.00      1st Qu.:2145      dohcv: 1
## Median :173.2      Median :65.50      Median :54.10      Median :2414      l : 12
## Mean :174.0      Mean :65.91      Mean :53.72      Mean :2556      ohc :148
## 3rd Qu.:183.1      3rd Qu.:66.90      3rd Qu.:55.50      3rd Qu.:2935      ohcf : 15
## Max. :208.1      Max. :72.30      Max. :59.80      Max. :4066      ohcv : 13
```

```
##
##      cylindernumber  enginesize      fuelsystem  boreratio      rotor: 4
##      stroke
## eight : 5      Min.   : 61.0    mpfi      :94    Min.   :2.54    Min.   :2.070
## five  : 11     1st Qu.: 97.0    2bbl      :66    1st Qu.:3.15    1st Qu.:3.110
## four  :159     Median :120.0    idi       :20    Median :3.31    Median :3.290
## six   : 24     Mean    :126.9    1bbl      :11    Mean    :3.33    Mean    :3.255
## three : 1      3rd Qu.:141.0    spdi      : 9    3rd Qu.:3.58    3rd Qu.:3.410
## twelve: 1      Max.     :326.0    4bbl      : 3    Max.     :3.94    Max.     :4.170
## two   : 4
##      compressionratio  horsepower      peakrpm      citympg
## Min.   : 7.00      Min.   : 48.0    Min.   :4150    Min.   :13.00
## 1st Qu.: 8.60      1st Qu.: 70.0    1st Qu.:4800    1st Qu.:19.00
## Median : 9.00      Median : 95.0    Median :5200    Median :24.00
## Mean    :10.14     Mean    :104.1    Mean    :5125    Mean    :25.22
## 3rd Qu.: 9.40      3rd Qu.:116.0    3rd Qu.:5500    3rd Qu.:30.00
## Max.     :23.00     Max.     :288.0    Max.     :6600    Max.     :49.00
##
##      highwaympg      price
## Min.   :16.00      Min.   : 5118
## 1st Qu.:25.00      1st Qu.: 7788
## Median :30.00      Median :10295
## Mean    :30.75     Mean    :13277
## 3rd Qu.:34.00      3rd Qu.:16503
## Max.     :54.00     Max.     :45400
##
```

```
hist(card$price, col = 2, breaks = 20)
```

Histogram of card\$price



```
d1=read.table("student-mat.csv",sep=";",header=TRUE)
d2=read.table("student-por.csv",sep=";",header=TRUE)

d3=merge(d1,d2,by=c("school","sex","age","address","famsize","Pstatus","Medu","Fedu","Mjob","Fjob","reason"),all=TRUE)
print(nrow(d3)) # 382 students
```

```
## [1] 382
```

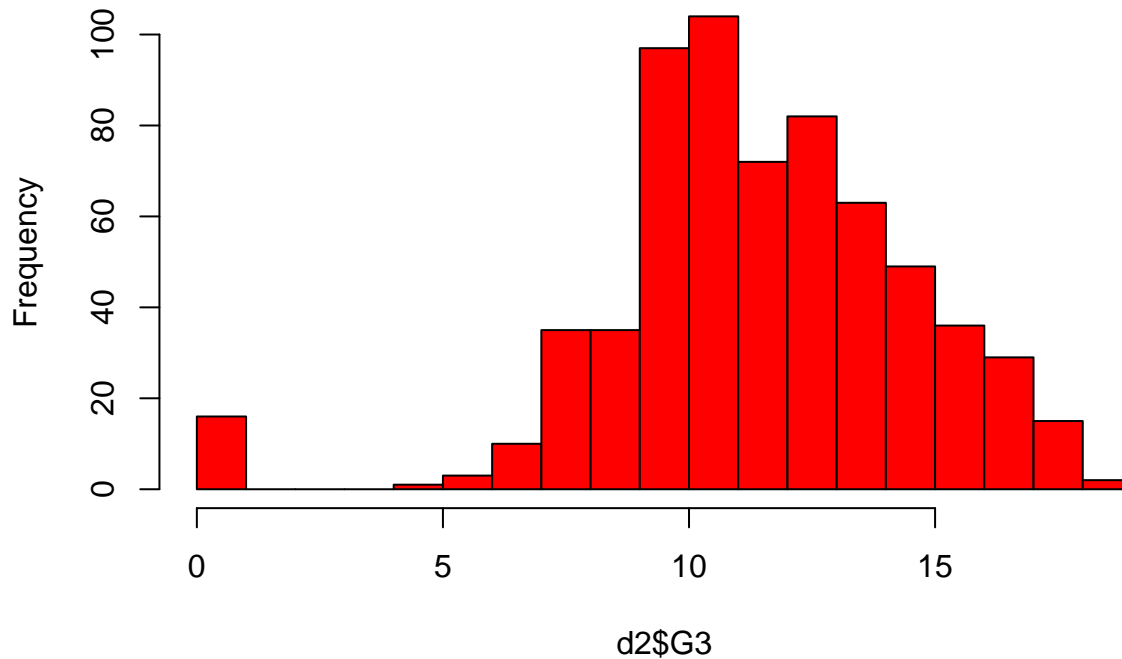
```
summary(d2)
```

```
## school sex age address famsize Pstatus Medu
## GP:423 F:383 Min. :15.00 R:197 GT3:457 A: 80 Min. :0.000
## MS:226 M:266 1st Qu.:16.00 U:452 LE3:192 T:569 1st Qu.:2.000
## Median :17.00 Median :2.000
## Mean :16.74 Mean :2.515
## 3rd Qu.:18.00 3rd Qu.:4.000
## Max. :22.00 Max. :4.000
## Fedu Mjob Fjob reason guardian
## Min. :0.000 at_home :135 at_home : 42 course :285 father:153
## 1st Qu.:1.000 health : 48 health : 23 home :149 mother:455
## Median :2.000 other :258 other :367 other : 72 other : 41
## Mean :2.307 services:136 services:181 reputation:143
## 3rd Qu.:3.000 teacher : 72 teacher : 36
## Max. :4.000
## traveltime studytime failures schoolsup famsup paid
```

```
## Min. :1.000 Min. :1.000 Min. :0.0000 no :581 no :251 no :610
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.0000 yes: 68 yes:398 yes: 39
## Median :1.000 Median :2.000 Median :0.0000
## Mean :1.569 Mean :1.931 Mean :0.2219
## 3rd Qu.:2.000 3rd Qu.:2.000 3rd Qu.:0.0000
## Max. :4.000 Max. :4.000 Max. :3.0000
## activities nursery higher internet romantic famrel
## no :334 no :128 no : 69 no :151 no :410 Min. :1.000
## yes:315 yes:521 yes:580 yes:498 yes:239 1st Qu.:4.000
## Median :4.000
## Mean :3.931
## 3rd Qu.:5.000
## Max. :5.000
## freetime goout Dalc Walc health
## Min. :1.00 Min. :1.000 Min. :1.000 Min. :1.00 Min. :1.000
## 1st Qu.:3.00 1st Qu.:2.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:2.000
## Median :3.00 Median :3.000 Median :1.000 Median :2.00 Median :4.000
## Mean :3.18 Mean :3.185 Mean :1.502 Mean :2.28 Mean :3.536
## 3rd Qu.:4.00 3rd Qu.:4.000 3rd Qu.:2.000 3rd Qu.:3.00 3rd Qu.:5.000
## Max. :5.00 Max. :5.000 Max. :5.000 Max. :5.00 Max. :5.000
## absences G1 G2 G3
## Min. : 0.000 Min. : 0.0 Min. : 0.00 Min. : 0.00
## 1st Qu.: 0.000 1st Qu.:10.0 1st Qu.:10.00 1st Qu.:10.00
## Median : 2.000 Median :11.0 Median :11.00 Median :12.00
## Mean : 3.659 Mean :11.4 Mean :11.57 Mean :11.91
## 3rd Qu.: 6.000 3rd Qu.:13.0 3rd Qu.:13.00 3rd Qu.:14.00
## Max. :32.000 Max. :19.0 Max. :19.00 Max. :19.00
```

```
hist(d2$G3, col =2, breaks = 20)
```

Histogram of d2\$G3

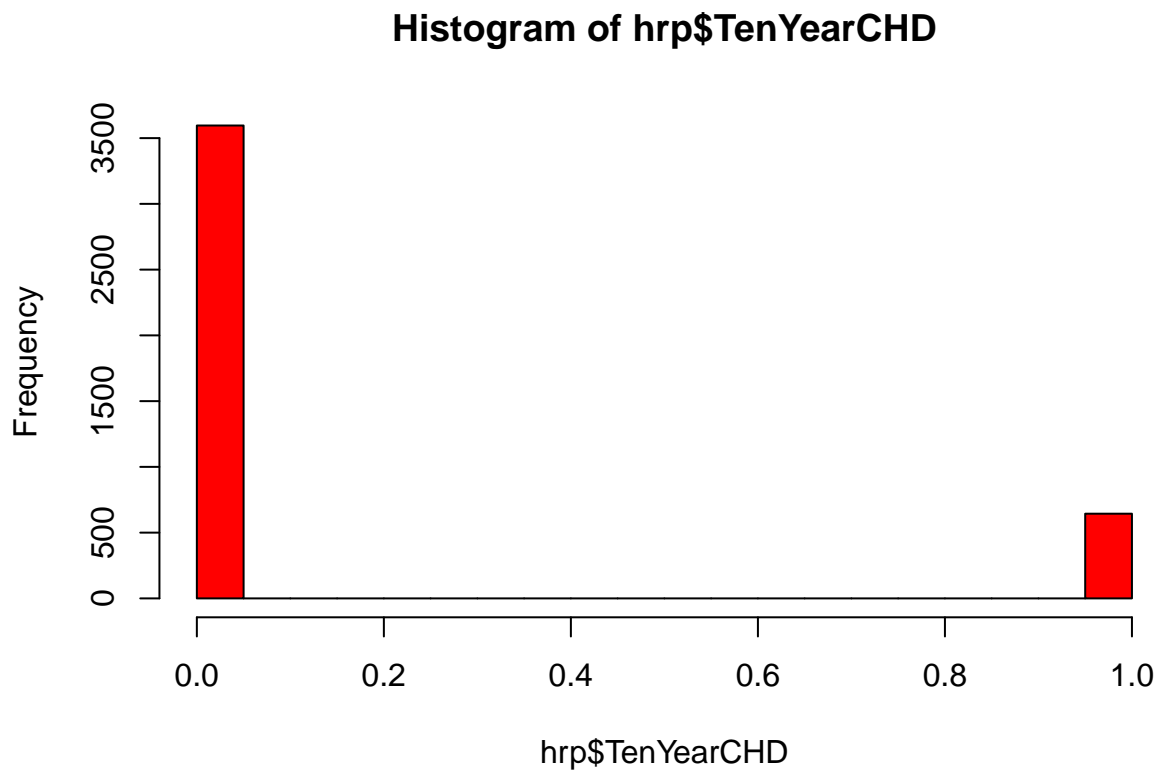


```
hrp <- read.csv("framingham.csv")
summary(hrp)
```

```
##      male      age      education      currentSmoker
## Min.   :0.0000   Min.   :32.00   Min.   :1.000   Min.   :0.0000
## 1st Qu.:0.0000   1st Qu.:42.00   1st Qu.:1.000   1st Qu.:0.0000
## Median :0.0000   Median :49.00   Median :2.000   Median :0.0000
## Mean   :0.4292   Mean   :49.58   Mean   :1.979   Mean   :0.4941
## 3rd Qu.:1.0000   3rd Qu.:56.00   3rd Qu.:3.000   3rd Qu.:1.0000
## Max.   :1.0000   Max.   :70.00   Max.   :4.000   Max.   :1.0000
##
##      NA's :105
##      cigsPerDay      BPMeds      prevalentStroke      prevalentHyp
## Min.   : 0.000   Min.   :0.000000   Min.   :0.000000   Min.   :0.0000
## 1st Qu.: 0.000   1st Qu.:0.000000   1st Qu.:0.000000   1st Qu.:0.0000
## Median : 0.000   Median :0.000000   Median :0.000000   Median :0.0000
## Mean   : 9.006   Mean   :0.02962   Mean   :0.005896   Mean   :0.3106
## 3rd Qu.:20.000   3rd Qu.:0.000000   3rd Qu.:0.000000   3rd Qu.:1.0000
## Max.   :70.000   Max.   :1.00000   Max.   :1.000000   Max.   :1.0000
## NA's   :29      NA's   :53
##      diabetes      totChol      sysBP      diaBP
## Min.   :0.000000   Min.   :107.0   Min.   : 83.5   Min.   : 48.0
## 1st Qu.:0.000000   1st Qu.:206.0   1st Qu.:117.0   1st Qu.: 75.0
## Median :0.000000   Median :234.0   Median :128.0   Median : 82.0
## Mean   :0.02571   Mean   :236.7   Mean   :132.4   Mean   : 82.9
## 3rd Qu.:0.000000   3rd Qu.:263.0   3rd Qu.:144.0   3rd Qu.: 90.0
```

```
## Max. :1.00000 Max. :696.0 Max. :295.0 Max. :142.5
## NA's :50
## BMI heartRate glucose TenYearCHD
## Min. :15.54 Min. : 44.00 Min. : 40.00 Min. :0.0000
## 1st Qu.:23.07 1st Qu.: 68.00 1st Qu.: 71.00 1st Qu.:0.0000
## Median :25.40 Median : 75.00 Median : 78.00 Median :0.0000
## Mean :25.80 Mean : 75.88 Mean : 81.96 Mean :0.1519
## 3rd Qu.:28.04 3rd Qu.: 83.00 3rd Qu.: 87.00 3rd Qu.:0.0000
## Max. :56.80 Max. :143.00 Max. :394.00 Max. :1.0000
## NA's :19 NA's :1 NA's :388
```

```
hist(hrp$TenYearCHD, col =2, breaks=15)
```



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