

# Assignment1

*Samuel Kigamba*

*Sep 1, 2019*

## DATA 607: Assignment1 – Loading Data into a Data Frame - Mushroom Data

See raw data under “<https://github.com/igukusamuel/DATA-607-Assignment-1/blob/master/Mushroom%20Data.csv>”

Load data into a data frame named MD directly from the original url.

```
mushroom <- read.table("https://archive.ics.uci.edu/ml/machine-learning-databases/mushroom/agaricus-lep
```

Review the column titles.

```
names(mushroom[c(1:5),c(1:5)])
```

```
## [1] "V1" "V2" "V3" "V4" "V5"
```

Review the data frame structure as well as a summary of the data items.

```
summary(mushroom[c(1:5),c(1:5)])
```

```
##  V1    V2    V3      V4    V5
##  e:3  b:1  f:0  w      :2  f:1
##  p:2  c:0  g:0  g      :1  t:4
##      f:0  s:4  n      :1
##      k:0  y:1  y      :1
##      s:0      b      :0
##      x:4      c      :0
##                      (Other):0
```

Review first 5 rows of the data frame

```
head(mushroom[c(1:5),c(1:5)])
```

```
##   V1 V2 V3 V4 V5
## 1  p  x  s  n  t
## 2  e  x  s  y  t
## 3  e  b  s  w  t
## 4  p  x  y  w  t
## 5  e  x  s  g  f
```

Review last 5 rows in the data frame to confirm that the data has loaded correctly.

```
tail(mushroom, 5)
```

```
##      V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17 V18 V19
## 8120 e  k s  n f  n a c b  y  e  ?  s  s  o  o  p  o  o
## 8121 e  x s  n f  n a c b  y  e  ?  s  s  o  o  p  n  o
## 8122 e  f s  n f  n a c b  n  e  ?  s  s  o  o  p  o  o
## 8123 p  k y  n f  y f  c n  b  t  ?  s  k  w  w  p  w  o
## 8124 e  x s  n f  n a c b  y  e  ?  s  s  o  o  p  o  o
##      V20 V21 V22 V23
## 8120    p    b    c    l
## 8121    p    b    v    l
## 8122    p    b    c    l
## 8123    e    w    v    l
## 8124    p    o    c    l
```

Rename the columns (note the default naming of columns is V1, V2,..., V23)

```
names(mushroom) <- c("classes", "cap-shape", "cap-surface", "cap-color", "bruises?", "odor", "gill-attach")
```

Confirm name change

```
names(mushroom[c(1:5),c(1:5)])
```

```
## [1] "classes"      "cap-shape"    "cap-surface"  "cap-color"    "bruises?"
```

Replace abbreviations used in column named classes

```
levels(mushroom$classes) [levels(mushroom$classes)=="p"] <- "poisonous"
levels(mushroom$classes) [levels(mushroom$classes)=="e"] <- "edible"
```

Confirm abbreviations change in column named classes

```
mushroom[c(1:5),c(1:5)]
```

```
##      classes cap-shape cap-surface cap-color bruises?
## 1 poisonous          x           s          n          t
## 2   edible          x           s          y          t
## 3   edible          b           s          w          t
## 4 poisonous          x           y          w          t
## 5   edible          x           s          g          f
```