Hantz_Angrand_HW1

Hantz Angrand

September 1, 2018

Load essential library

```
library(plyr)
```

Load the data

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

head(mushroom)
```

```
##   V1 V2 V3 V4 V5 V6 V7 V8 V9 V10 V11 V12 V13 V14 V15 V16 V17 V18 V19 V20
## 1  p  x  s  n  t  p  f  c  n   k   e   e   s   s   w   w   p   w   o   p
## 2  e  x  s  y  t  a  f  c  b   k   e   c   s   s   w   w   p   w   o   p
## 3  e  b  s  w  t  l  f  c  b   n   e   c   s   s   w   w   p   w   o   p
## 4  p  x  y  w  t  p  f  c  n   n   e   e   s   s   w   w   p   w   o   p
## 5  e  x  s  g  f  n  f  w  b   k   t   e   s   s   w   w   p   w   o   e
## 6  e  x  y  y  t  a  f  c  b   n   e   c   s   s   w   w   p   w   o   p
##   V21 V22 V23
## 1   k   s   u
## 2   n   n   g
## 3   n   n   m
## 4   k   s   u
## 5   n   a   g
## 6   k   n   g
```

Rename column Headers

```
colnames(mushroom)<-c("class", "cap-shape", "cap-surface", "cap-color", "bruises", "odor", "gill-attachm

mush<-data.frame(mushroom)

head(mush)
```

```
##   class cap.shape cap.surface cap.color bruises odor gill.attachment
## 1     p         x           s         n       t    p               f
## 2     e         x           s         y       t    a               f
## 3     e         b           s         w       t    l               f
## 4     p         x           y         w       t    p               f
## 5     e         x           s         g       f    n               f
## 6     e         x           y         y       t    a               f
##   gill.spacing gill.size gill.color stalk.shape stalk.root
## 1            c         n          k           e          e
```

```
## 2            c         b          k           e          c
## 3            c         b          n           e          c
## 4            c         n          n           e          e
## 5            w         b          k           t          e
## 6            c         b          n           e          c
##   stalk.surface.above.ring stalk.surface.below.ring stalk.color.above.ring
## 1                        s                        s                      w
## 2                        s                        s                      w
## 3                        s                        s                      w
## 4                        s                        s                      w
## 5                        s                        s                      w
## 6                        s                        s                      w
##   stalk.color.below.ring veil.type veil.color ring.number ring.type
## 1                      w         p          w           o         p
## 2                      w         p          w           o         p
## 3                      w         p          w           o         p
## 4                      w         p          w           o         p
## 5                      w         p          w           o         e
## 6                      w         p          w           o         p
##   spore.print.color population habitats
## 1                 k          s        u
## 2                 n          n        g
## 3                 n          n        m
## 4                 k          s        u
## 5                 n          a        g
## 6                 k          n        g
```

Select a subset of the data

```
mush<-subset(mush, select=c("class","cap.shape", "cap.color","odor", "habitats"))
head(mush)
```

```
##   class cap.shape cap.color odor habitats
## 1     p         x         n    p        u
## 2     e         x         y    a        g
## 3     e         b         w    l        m
## 4     p         x         w    p        u
## 5     e         x         g    n        g
## 6     e         x         y    a        g
```

Modifying the data to make it more user friendly by replacing abbreviations

```
levels(mush$class)<-c("edible", "poisonous")
levels(mush$class)
```

```
## [1] "edible"    "poisonous"
```

```
levels(mush$cap.shape)<-c("bell","conical","flat","knobbed","sunken","convex")
levels(mush$cap.shape)
```

```
## [1] "bell"    "conical" "flat"    "knobbed" "sunken"  "convex"
```

```
levels(mush$cap.color)<-c("buff","cinnamon","red","gray","brown","pink","green","purple","white","yellow")
levels(mush$cap.color)
```

```
##  [1] "buff"     "cinnamon" "red"      "gray"     "brown"    "pink"
##  [7] "green"    "purple"   "white"    "yellow"
```

```
levels(mush$odor)<-c("almond","creosote","foul","anise","musty","none","pungent","spicy","fishy")
levels(mush$odor)
```

```
## [1] "almond"   "creosote" "foul"     "anise"    "musty"    "none"
## [7] "pungent"  "spicy"    "fishy"
```

```
levels(mush$habitats)<-c("woods","grasses","leaves","meadows","paths","urban","waste")
levels(mush$habitats)
```

```
## [1] "woods"   "grasses" "leaves"  "meadows" "paths"   "urban"   "waste"
```

```
head(mush)
```

```
##       class cap.shape cap.color    odor habitats
## 1 poisonous    convex     brown pungent    urban
## 2    edible    convex    yellow  almond  grasses
## 3    edible      bell     white   anise  meadows
## 4 poisonous    convex     white pungent    urban
## 5    edible    convex      gray    none  grasses
## 6    edible    convex    yellow  almond  grasses
```

Group the data by class

```
edible<-subset(mush,class="edible")
dim(edible)
```

```
## [1] 8124    5
```

```
summary(edible)
```

```
##        class        cap.shape      cap.color         odor
##  edible   :4208   bell   : 452   brown  :2284   none   :3528
##  poisonous:3916   conical:   4   gray   :1840   foul   :2160
##                   flat   :3152   red    :1500   spicy  : 576
##                   knobbed: 828   yellow :1072   fishy  : 576
##                   sunken :  32   white  :1040   almond : 400
##                   convex :3656   buff   : 168   anise  : 400
##                                  (Other): 220   (Other): 484
##     habitats
##  woods  :3148
##  grasses:2148
##  leaves : 832
##  meadows: 292
##  paths  :1144
##  urban  : 368
##  waste  : 192
```

```
poisonous<-subset(mush, class="poisonous")
dim(poisonous)
```

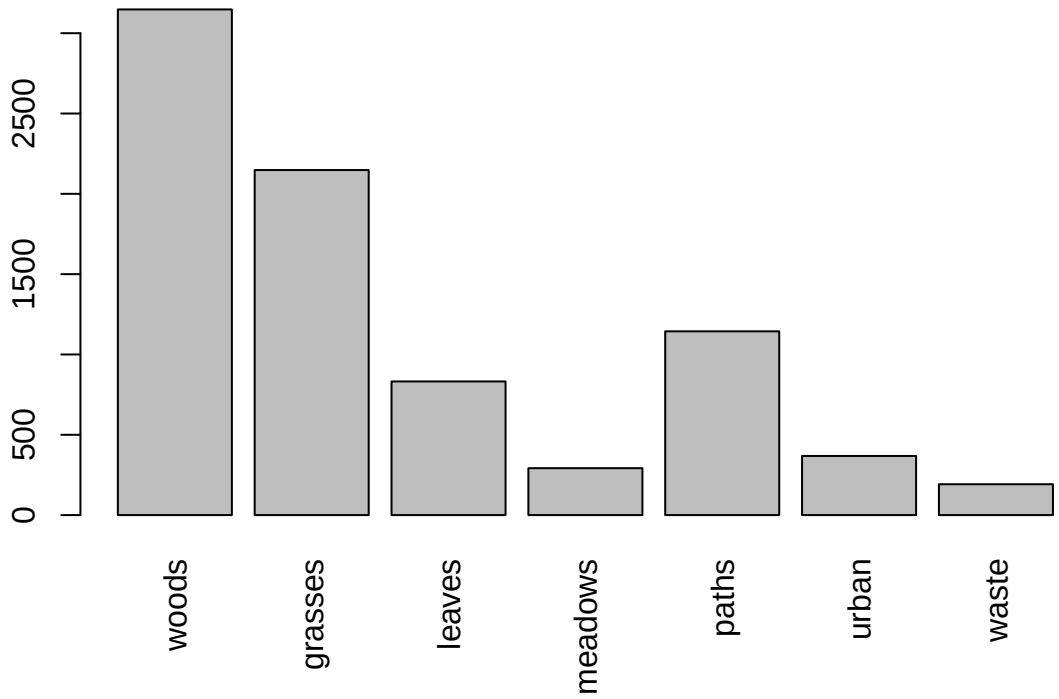
```
## [1] 8124    5
```

```
summary(poisonous)
```

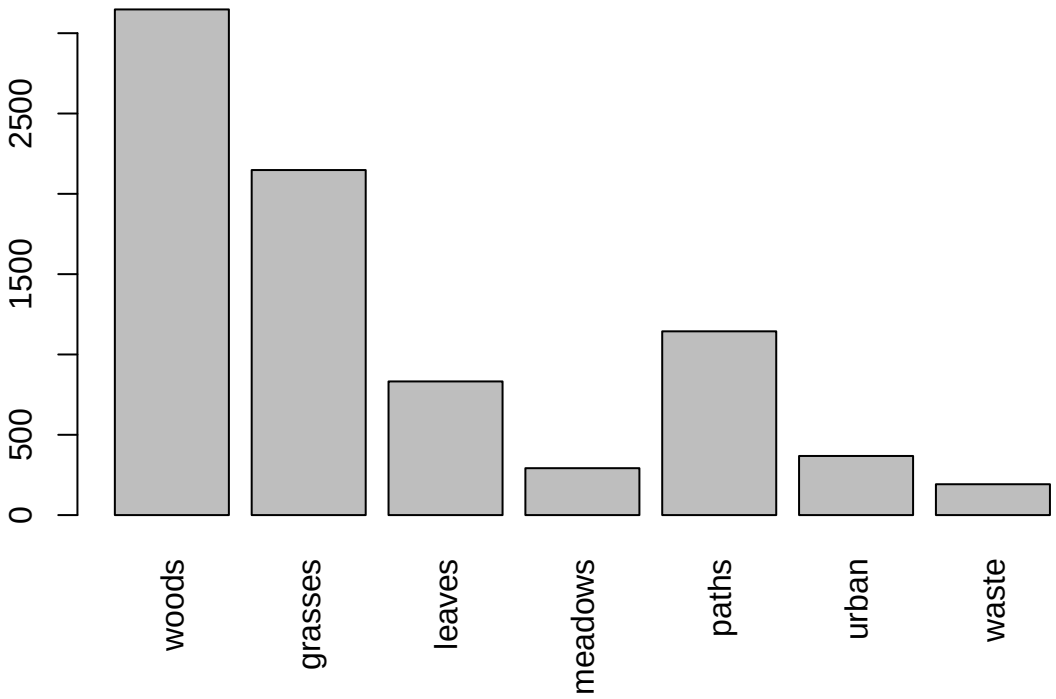
```
##        class        cap.shape      cap.color         odor
##  edible   :4208   bell   : 452   brown  :2284   none   :3528
##  poisonous:3916   conical:   4   gray   :1840   foul   :2160
##                   flat   :3152   red    :1500   spicy  : 576
##                   knobbed: 828   yellow :1072   fishy  : 576
##                   sunken :  32   white  :1040   almond : 400
##                   convex :3656   buff   : 168   anise  : 400
##                                  (Other): 220   (Other): 484
##     habitats
##  woods  :3148
##  grasses:2148
##  leaves : 832
##  meadows: 292
##  paths  :1144
##  urban  : 368
##  waste  : 192
```

Plotting some data

```
plot(edible$habitats, las=3)
```




```
plot(poisonous$habitats, las=3)
```


0
500
1500
2500
woods
grasses
leaves
meadows
paths
urban
waste