

Assignment1_Mushroom

Priya Shaji

January 31, 2019

Read the file in table format and create a data frame from it

```
Mushroom_Family <- read.csv(url("https://archive.ics.uci.edu/ml/machine-learning-databases/mushroom/agaricus-lepiota.data"), header = FALSE)
```

Mushroom_Family is the dataframe created. Dimensions of the dataframe:

```
dim(Mushroom_Family)
```

```
## [1] 8124 23
```

The first six rows of the data frame are:

```
head.matrix(x = Mushroom_Family, 6L)
```

```
##   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
```

Mushroom_Sibling is the subset of the Mushroom_Family dataframe from the selected columns:

```
Mushroom_Sibling <- Mushroom_Family[,c(1,4,6,10,11,23)]
```

Renaming the column names of the subset dataframe:

```
names(Mushroom_Sibling) = c("Type", "Cap_Color", "Odor", "Gill_Color",
"Stem_Shape", "Habitat")
```

Column names of Mushroom_Sibling dataframe are renamed:

```
names(Mushroom_Sibling)
```

```
## [1] "Type"      "Cap_Color" "Odor"      "Gill_Color" "Stem_Shape"
## [6] "Habitat"
```

Replace the abbreviation used in column Type:

1. Convert Columns to characters

```
Mushroom_Sibling$Type <- as.character(Mushroom_Sibling$Type)
```

2. Replace abbreviation with meaningful names

```
Mushroom_Sibling$Type[Mushroom_Sibling$Type == "e"] <- "edible"
Mushroom_Sibling$Type[Mushroom_Sibling$Type == "p"] <- "poisonous"
```

3. If needed to maintain columns as factors

```
Mushroom_Sibling$Type <- as.factor(Mushroom_Sibling$Type)
```

Display first six rows of the dataframe to verify the change

```
head.matrix(Mushroom_Sibling, 6L)
```

```
##      Type Cap_Color Odor Gill_Color Stem_Shape Habitat
## 1 poisonous      n    p         k         e         u
## 2  edible      y    a         k         e         g
## 3  edible      w    l         n         e         m
## 4 poisonous      w    p         n         e         u
## 5  edible      g    n         k         t         g
## 6  edible      y    a         n         e         g
```

Replace abbreviations of column Cap_Color

```
Mushroom_Sibling$Cap_Color <- as.character(Mushroom_Sibling$Cap_Color)
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "n"] <- "brown"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "b"] <- "buff"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "c"] <- "cinnamon"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "g"] <- "gray"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "r"] <- "green"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "p"] <- "pink"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "u"] <- "purple"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "e"] <- "red"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "w"] <- "white"
Mushroom_Sibling$Cap_Color[Mushroom_Sibling$Cap_Color == "y"] <- "yellow"
Mushroom_Sibling$Cap_Color <- as.factor(Mushroom_Sibling$Cap_Color)
```

Replace Abbreviations of column Odor

```
Mushroom_Sibling$Odor <- as.character(Mushroom_Sibling$Odor)
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "a"] <- "almond"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "l"] <- "anise"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "c"] <- "creosote"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "y"] <- "fishy"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "f"] <- "foul"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "m"] <- "musty"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "n"] <- "none"
```

```
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "p"] <- "pungent"
Mushroom_Sibling$Odor[Mushroom_Sibling$Odor == "s"] <- "spicy"
Mushroom_Sibling$Odor <- as.factor(Mushroom_Sibling$Odor)
```

Replace abbreviation of column Gill_Color

```
head.matrix(Mushroom_Sibling, 3L)
```

```
##      Type Cap_Color   Odor Gill_Color Stem_Shape Habitat
## 1 poisonous  brown pungent         k         e         u
## 2  edible   yellow almond         k         e         g
## 3  edible   white  anise         n         e         m
```

```
Mushroom_Sibling$Gill_Color <- as.character(Mushroom_Sibling$Gill_Color)
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "k"] <- "black"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "n"] <- "brown"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "b"] <- "buff"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "h"] <-
"chocolate"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "g"] <- "gray"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "r"] <- "green"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "o"] <- "orange"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "p"] <- "pink"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "u"] <- "purple"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "e"] <- "red"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "w"] <- "white"
Mushroom_Sibling$Gill_Color[Mushroom_Sibling$Gill_Color == "y"] <- "yellow"
Mushroom_Sibling$Gill_Color <- as.factor(Mushroom_Sibling$Gill_Color)
```

Replace abbreviation of column Stem_Shape

```
Mushroom_Sibling$Stem_Shape <- as.character(Mushroom_Sibling$Stem_Shape)
Mushroom_Sibling$Stem_Shape[Mushroom_Sibling$Stem_Shape == "e"] <-
"enlarging"
Mushroom_Sibling$Stem_Shape[Mushroom_Sibling$Stem_Shape == "t"] <- "tapering"
Mushroom_Sibling$Stem_Shape <- as.factor(Mushroom_Sibling$Stem_Shape)
```

Replace abbreviations of column Habitat

```
Mushroom_Sibling$Habitat <- as.character(Mushroom_Sibling$Habitat)
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "g"] <- "grasses"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "l"] <- "leaves"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "m"] <- "meadows"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "p"] <- "paths"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "u"] <- "urban"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "w"] <- "waste"
Mushroom_Sibling$Habitat[Mushroom_Sibling$Habitat == "d"] <- "woods"
Mushroom_Sibling$Habitat <- as.factor(Mushroom_Sibling$Habitat)
```

Display the table to verify:

```
head.matrix(x = Mushroom_Sibling, 6L)
```

```
##      Type Cap_Color   Odor Gill_Color Stem_Shape Habitat
## 1 poisonous   brown pungent    black enlarging   urban
## 2  edible    yellow almond    black enlarging grasses
## 3  edible    white  anise    brown enlarging meadows
## 4 poisonous   white pungent    brown enlarging   urban
## 5  edible    gray   none    black tapering grasses
## 6  edible    yellow almond    brown enlarging grasses
```

By exploring and analyzing the dataset and the data description of Mushrooms ,the following conclusions can be drawn:

- 1) Edible and poisonous mushrooms can be analyzed based on two strong factors: Odor and Habitat With Odor being the strongest factor.
- 2) Therefore Let's calculate the frequencies of both the factors affecting the edibility of the mushrooms:

```
table = table(Mushroom_Sibling$Type,
Mushroom_Sibling$Odor,Mushroom_Sibling$Habitat)
result_data = as.data.frame(table)
names(result_data)[1] = 'Type'
names(result_data)[2] = 'Odor'
names(result_data)[3] = 'Habitat'
result_data
```

```
##      Type   Odor Habitat Freq
## 1  edible  almond grasses  176
## 2 poisonous  almond grasses    0
## 3  edible  anise grasses  176
## 4 poisonous  anise grasses    0
## 5  edible creosote grasses    0
## 6 poisonous creosote grasses    0
## 7  edible  fishy grasses    0
## 8 poisonous  fishy grasses    0
## 9  edible   foul grasses    0
## 10 poisonous  foul grasses  576
## 11  edible  musty grasses    0
## 12 poisonous  musty grasses    0
## 13  edible   none grasses 1056
## 14 poisonous  none grasses   36
## 15  edible  pungent grasses    0
## 16 poisonous  pungent grasses  128
## 17  edible  spicy grasses    0
## 18 poisonous  spicy grasses    0
## 19  edible  almond leaves    0
## 20 poisonous  almond leaves    0
## 21  edible  anise leaves    0
## 22 poisonous  anise leaves    0
## 23  edible creosote leaves    0
## 24 poisonous creosote leaves    0
```

## 25	edible	fishy	leaves	0
## 26	poisonous	fishy	leaves	192
## 27	edible	foul	leaves	0
## 28	poisonous	foul	leaves	192
## 29	edible	musty	leaves	0
## 30	poisonous	musty	leaves	0
## 31	edible	none	leaves	240
## 32	poisonous	none	leaves	16
## 33	edible	pungent	leaves	0
## 34	poisonous	pungent	leaves	0
## 35	edible	spicy	leaves	0
## 36	poisonous	spicy	leaves	192
## 37	edible	almond	meadows	128
## 38	poisonous	almond	meadows	0
## 39	edible	anise	meadows	128
## 40	poisonous	anise	meadows	0
## 41	edible	creosote	meadows	0
## 42	poisonous	creosote	meadows	0
## 43	edible	fishy	meadows	0
## 44	poisonous	fishy	meadows	0
## 45	edible	foul	meadows	0
## 46	poisonous	foul	meadows	0
## 47	edible	musty	meadows	0
## 48	poisonous	musty	meadows	0
## 49	edible	none	meadows	0
## 50	poisonous	none	meadows	36
## 51	edible	pungent	meadows	0
## 52	poisonous	pungent	meadows	0
## 53	edible	spicy	meadows	0
## 54	poisonous	spicy	meadows	0
## 55	edible	almond	paths	48
## 56	poisonous	almond	paths	0
## 57	edible	anise	paths	48
## 58	poisonous	anise	paths	0
## 59	edible	creosote	paths	0
## 60	poisonous	creosote	paths	0
## 61	edible	fishy	paths	0
## 62	poisonous	fishy	paths	192
## 63	edible	foul	paths	0
## 64	poisonous	foul	paths	624
## 65	edible	musty	paths	0
## 66	poisonous	musty	paths	0
## 67	edible	none	paths	40
## 68	poisonous	none	paths	0
## 69	edible	pungent	paths	0
## 70	poisonous	pungent	paths	0
## 71	edible	spicy	paths	0
## 72	poisonous	spicy	paths	192
## 73	edible	almond	urban	0
## 74	poisonous	almond	urban	0

## 75	edible	anise	urban	0
## 76	poisonous	anise	urban	0
## 77	edible	creosote	urban	0
## 78	poisonous	creosote	urban	0
## 79	edible	fishy	urban	0
## 80	poisonous	fishy	urban	0
## 81	edible	foul	urban	0
## 82	poisonous	foul	urban	144
## 83	edible	musty	urban	0
## 84	poisonous	musty	urban	0
## 85	edible	none	urban	96
## 86	poisonous	none	urban	0
## 87	edible	pungent	urban	0
## 88	poisonous	pungent	urban	128
## 89	edible	spicy	urban	0
## 90	poisonous	spicy	urban	0
## 91	edible	almond	waste	0
## 92	poisonous	almond	waste	0
## 93	edible	anise	waste	0
## 94	poisonous	anise	waste	0
## 95	edible	creosote	waste	0
## 96	poisonous	creosote	waste	0
## 97	edible	fishy	waste	0
## 98	poisonous	fishy	waste	0
## 99	edible	foul	waste	0
## 100	poisonous	foul	waste	0
## 101	edible	musty	waste	0
## 102	poisonous	musty	waste	0
## 103	edible	none	waste	192
## 104	poisonous	none	waste	0
## 105	edible	pungent	waste	0
## 106	poisonous	pungent	waste	0
## 107	edible	spicy	waste	0
## 108	poisonous	spicy	waste	0
## 109	edible	almond	woods	48
## 110	poisonous	almond	woods	0
## 111	edible	anise	woods	48
## 112	poisonous	anise	woods	0
## 113	edible	creosote	woods	0
## 114	poisonous	creosote	woods	192
## 115	edible	fishy	woods	0
## 116	poisonous	fishy	woods	192
## 117	edible	foul	woods	0
## 118	poisonous	foul	woods	624
## 119	edible	musty	woods	0
## 120	poisonous	musty	woods	36
## 121	edible	none	woods	1784
## 122	poisonous	none	woods	32
## 123	edible	pungent	woods	0
## 124	poisonous	pungent	woods	0

## 125	edible	spicy	woods	0
## 126	poisonous	spicy	woods	192