

Data_607_Week_1_Assignment

Jimmy Ng

1/29/2019

[illegible]

```

        `cap-shape` == "f" ~ "flat",
        `cap-shape` == "k" ~ "knobbed",
        `cap-shape` == "s" ~ "sunken",
        TRUE ~ "others"),
    `cap-surface` = dplyr::case_when(`cap-surface` == "f" ~ "fibrous",
        `cap-surface` == "g" ~ "grooves",
        `cap-surface` == "y" ~ "scaly",
        `cap-surface` == "s" ~ "smooth",
        TRUE ~ "others"),
    `cap-color` = dplyr::case_when(`cap-color` == "n" ~ "brown",
        `cap-color` == "b" ~ "buff",
        `cap-color` == "c" ~ "cinnamon",
        `cap-color` == "g" ~ "gray",
        `cap-color` == "r" ~ "green",
        `cap-color` == "p" ~ "pink",
        `cap-color` == "u" ~ "purple",
        `cap-color` == "e" ~ "red",
        `cap-color` == "w" ~ "white",
        `cap-color` == "y" ~ "yellow",
        TRUE ~ "others"))

# check summary, missing value & unique values by column
purrr::map(list(summary,
    function(x) {colSums(is.na(x))},
    function(x) {sapply(x, unique)}),
    function(x){x(df)})

## [[1]]
##   classes      cap-shape      cap-surface
## Length:8124    Length:8124    Length:8124
## Class :character Class :character Class :character
## Mode :character Mode :character Mode :character
##   cap-color
## Length:8124
## Class :character
## Mode :character
##

```

```

## [[2]]
##      classes   cap-shape cap-surface   cap-color
##           0         0         0         0
##
## [[3]]
## [[3]]$classes
## [1] "poisonous" "edible"
##
## [[3]]$`cap-shape`
## [1] "convex" "bell" "sunken" "flat" "knobbed" "conical"
##
## [[3]]$`cap-surface`
## [1] "smooth" "scaly" "fibrous" "grooves"
##
## [[3]]$`cap-color`
## [1] "brown" "yellow" "white" "gray" "red" "pink"
## [7] "buff" "purple" "cinnamon" "green"

# print & plot distribution by column
par(mfrow = c(2, 2))
sapply(df, function(x) round(prop.table(ftable(x)), 2)) %>%
  print %>%
lapply(., barplot) %>%
  invisible

## $classes
## x edible poisonous
##
##      0.52      0.48
##
## $`cap-shape`
## x bell conical convex flat knobbed sunken
##
##      0.06      0.00      0.45 0.39      0.10      0.00
##
## $`cap-surface`

```

```
## x fibrous grooves scaly smooth
##
##      0.29      0.00  0.40   0.31
##
## $`cap-color`
## x brown buff cinnamon gray green pink purple  red white yellow
##
##      0.28 0.02      0.01 0.23   0.00 0.02   0.00 0.18   0.13   0.13
```

