Exploratory\_Analysis.R

skhim

2023-03-26

#Creating data frame  
library(readr)

## Warning: package 'readr' was built under R version 4.2.3

chocolate <- read\_csv("data/chocolate.csv")

## New names:  
## Rows: 2224 Columns: 19  
## ── Column specification  
## ──────────────────────────────────────────────────────── Delimiter: "," chr  
## (6): company, company\_location, country\_of\_bean\_origin, first\_taste, se... dbl  
## (13): ...1, ref, review\_date, cocoa\_percent, rating, counts\_of\_ingredien...  
## ℹ Use `spec()` to retrieve the full column specification for this data. ℹ  
## Specify the column types or set `show\_col\_types = FALSE` to quiet this message.  
## • `` -> `...1`

View(chocolate)  
  
chocolate <- data.frame(chocolate)  
  
#Finding class of each column  
library(tibble)

## Warning: package 'tibble' was built under R version 4.2.3

column\_names = colnames(chocolate)  
column\_types = sapply(chocolate, class)  
tbbl = as\_tibble(data.frame(column\_names, column\_types))  
  
library(knitr)  
library(xtable)

## Warning: package 'xtable' was built under R version 4.2.3

#xtable::xtable(tbbl)  
knitr::kable(tbbl, caption = "Table of Variables' Type", align = "lc")

Table of Variables’ Type

| column\_names | column\_types |
| --- | --- |
| …1 | numeric |
| ref | numeric |
| company | character |
| company\_location | character |
| review\_date | numeric |
| country\_of\_bean\_origin | character |
| cocoa\_percent | numeric |
| rating | numeric |
| counts\_of\_ingredients | numeric |
| bean | numeric |
| vanilla | numeric |
| vanila | numeric |
| lecithin | numeric |
| salt | numeric |
| sugar | numeric |
| sweetener\_without\_sugar | numeric |
| first\_taste | character |
| second\_taste | character |
| third\_taste | character |

knitr::kable(summary(chocolate), caption = "Table of Summary Statistics")

Table of Summary Statistics

|  | …1 | ref | company | company\_location | review\_date | country\_of\_bean\_origin | cocoa\_percent | rating | counts\_of\_ingredients | bean | vanilla | vanila | lecithin | salt | sugar | sweetener\_without\_sugar | first\_taste | second\_taste | third\_taste |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Min. : 1.0 | Min. : 5 | Length:2224 | Length:2224 | Min. :2006 | Length:2224 | Min. : 42.00 | Min. :1.000 | Min. :1.000 | Min. :1 | Min. :0.0000 | Min. :0.0000 | Min. :0.0000 | Min. :0.00000 | Min. :0.0000 | Min. :0.00000 | Length:2224 | Length:2224 | Length:2224 |
|  | 1st Qu.: 556.8 | 1st Qu.: 776 | Class :character | Class :character | 1st Qu.:2011 | Class :character | 1st Qu.: 70.00 | 1st Qu.:3.000 | 1st Qu.:2.000 | 1st Qu.:1 | 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.0000 | 1st Qu.:0.00000 | 1st Qu.:1.0000 | 1st Qu.:0.00000 | Class :character | Class :character | Class :character |
|  | Median :1112.5 | Median :1381 | Mode :character | Mode :character | Median :2014 | Mode :character | Median : 70.00 | Median :3.250 | Median :3.000 | Median :1 | Median :1.0000 | Median :0.0000 | Median :0.0000 | Median :0.00000 | Median :1.0000 | Median :0.00000 | Mode :character | Mode :character | Mode :character |
|  | Mean :1112.5 | Mean :1337 | NA | NA | Mean :2014 | NA | Mean : 71.49 | Mean :3.199 | Mean :3.076 | Mean :1 | Mean :0.6902 | Mean :0.1565 | Mean :0.2154 | Mean :0.01664 | Mean :0.9631 | Mean :0.03417 | NA | NA | NA |
|  | 3rd Qu.:1668.2 | 3rd Qu.:1928 | NA | NA | 3rd Qu.:2016 | NA | 3rd Qu.: 74.00 | 3rd Qu.:3.500 | 3rd Qu.:4.000 | 3rd Qu.:1 | 3rd Qu.:1.0000 | 3rd Qu.:0.0000 | 3rd Qu.:0.0000 | 3rd Qu.:0.00000 | 3rd Qu.:1.0000 | 3rd Qu.:0.00000 | NA | NA | NA |
|  | Max. :2224.0 | Max. :2490 | NA | NA | Max. :2020 | NA | Max. :100.00 | Max. :4.000 | Max. :6.000 | Max. :1 | Max. :1.0000 | Max. :1.0000 | Max. :1.0000 | Max. :1.00000 | Max. :1.0000 | Max. :1.00000 | NA | NA | NA |