The Ramen Rater

곽영효

2022-08-08

## 출처: Top Ramen Ratings 2022

<https://www.kaggle.com/datasets/ankanhore545/top-ramen-ratings-2022>

## 서론

2002년부터 전세계의 라면 리뷰를 해온 리뷰어 ‘The Ramen Rater’

그가 그동안 라면들에 매긴 별점들에 대한 데이터

## DataFrame 불러오기

df <- read.csv('./TopRamenRatings.csv', header = T,  
 fileEncoding = "euc-kr")  
str(df)

## 'data.frame': 4120 obs. of 6 variables:  
## $ Review..: int 4120 4119 4118 4117 4116 4115 4114 4113 4112 4111 ...  
## $ Brand : chr "MIT" "Sapporo Ichiban" "Binh Tay" "Charming Couple" ...  
## $ Variety : chr "Shallot Sauce Dry Noodle" "Tonkotsu Ramen Japanese Style Noodles" "Mi Kiwi" "Biang Biang Scallion Chicken Sauce" ...  
## $ Style : chr "Pack" "Bowl" "Bowl" "Pack" ...  
## $ Country : chr "Taiwan" "United States" "Vietnam" "Taiwan" ...  
## $ Stars : chr "3.000" "4.500" "3.000" "4.500" ...

library(VIM)

## 필요한 패키지를 로딩중입니다: colorspace

## 필요한 패키지를 로딩중입니다: grid

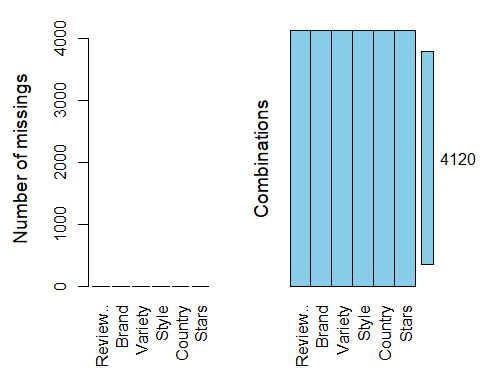
## VIM is ready to use.

## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues

##   
## 다음의 패키지를 부착합니다: 'VIM'

## The following object is masked from 'package:datasets':  
##   
## sleep

aggr(df, prop = F, numbers = T)



summary(df)

## Review.. Brand Variety Style   
## Min. : 1 Length:4120 Length:4120 Length:4120   
## 1st Qu.:1031 Class :character Class :character Class :character   
## Median :2060 Mode :character Mode :character Mode :character   
## Mean :2060   
## 3rd Qu.:3090   
## Max. :4120   
## Country Stars   
## Length:4120 Length:4120   
## Class :character Class :character   
## Mode :character Mode :character   
##   
##   
##

## NA => 0점으로 변환

df$Stars <- as.double(df$Stars)

## Warning: 강제형변환에 의해 생성된 NA 입니다

df[is.na(df$Stars), ]$Stars <- 0

## 범주형으로 변환

df$Brand <- as.factor(df$Brand)  
df$Style <- as.factor(df$Style)  
df$Country <- as.factor(df$Country)

# df[df$Stars >= 4.00 & df$Country == 'South Korea',]$Variety

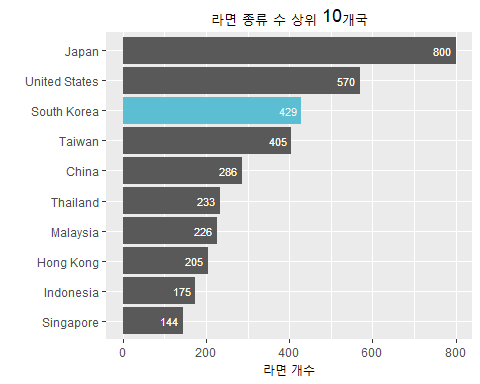
## 항목 수 상위 10개국

country\_10 <- head(table(df$Country)[order(table(df$Country), decreasing = T)], 10)  
  
country\_10

##   
## Japan United States South Korea Taiwan China   
## 800 570 429 405 286   
## Thailand Malaysia Hong Kong Indonesia Singapore   
## 233 226 205 175 144

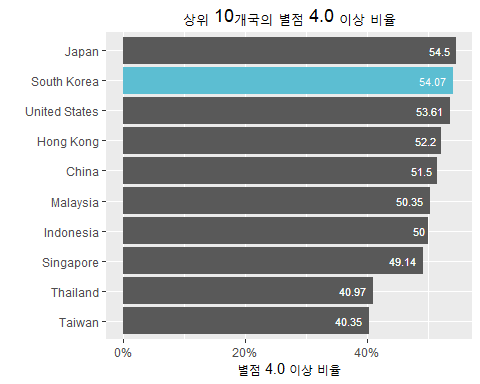
## 상위 10개국의 라면 개수

library(ggplot2)  
country\_10 <- as.data.frame(country\_10)  
  
graph1 <- ggplot(country\_10, aes(x = reorder(Var1, Freq),  
 y = Freq)) + geom\_bar(stat = 'identity') +   
 labs(title = "라면 종류 수 상위 10개국") +  
 theme(plot.title = element\_text(hjust = 0.5))  
  
graph1 <- graph1 + coord\_flip() + labs(x='', y='라면 개수') +   
 geom\_text(aes(label=Freq), size=3, hjust=1.25, color='#FFFFFF')  
  
graph1 + geom\_bar(country\_10[country\_10$Var1=='South Korea', ],  
 mapping = aes(x=reorder(Var1, Freq), y=Freq), fill='#5CBED2',   
 stat='identity') +  
 geom\_text(aes(label=Freq), size=3, hjust=1.25, color='#FFFFFF')



## 4점 이상의 별점을 받은 라면의 비율

df\_country\_10 <- df[df$Country %in% country\_10$Var1,]  
df\_country\_10 <- table(df\_country\_10[df\_country\_10$Stars >= 4.00,]$Country)  
df\_country\_10 <- head(df\_country\_10[order(df\_country\_10, decreasing = T)], 10)  
df\_country\_10 <- as.data.frame(df\_country\_10)  
  
for (i in 1:10){  
 df\_country\_10$Freq[i] <- df\_country\_10$Freq[i] / country\_10$Freq[i]}  
  
graph2 <- ggplot(df\_country\_10, aes(x = reorder(Var1, Freq),  
 y = Freq)) + geom\_bar(stat = 'identity') +   
 labs(title = "상위 10개국의 별점 4.0 이상 비율") +  
 theme(plot.title = element\_text(hjust = 0.5))  
  
graph2 <-graph2 + coord\_flip() + labs(x='', y='별점 4.0 이상 비율') +  
 scale\_y\_continuous(labels = scales::percent) +  
 geom\_text(aes(label=round(Freq,4)\*100), size=3,  
 hjust=1.25, color='#FFFFFF')   
  
graph2 + geom\_bar(df\_country\_10[df\_country\_10$Var1=='South Korea', ],  
 mapping = aes(x=reorder(Var1, Freq), y=Freq), fill='#5CBED2',   
 stat='identity') +  
 geom\_text(aes(label=round(Freq,4)\*100), size=3, hjust=1.25,   
 color='#FFFFFF')

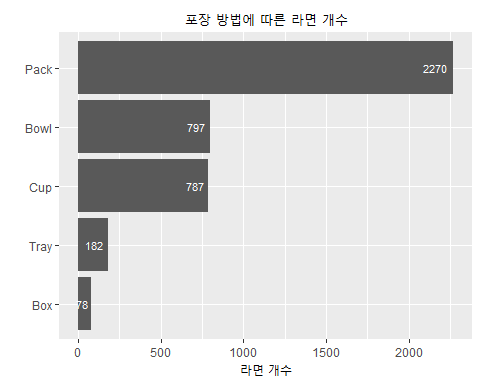
 ##

df[(df$Country == 'South Korea') & (df$Stars == 5),]$Variety

## [1] "Cheese Rapokki"   
## [2] "Cheese Ramyun"   
## [3] "Mac & Cheese Spaghetti"   
## [4] "Shin Light Air Dried Noodle Soup"   
## [5] "Mr Kimchi Stirfried Kimchi Ramen"   
## [6] "Bul Jjamppong"   
## [7] "Buldak Light"   
## [8] "Big Gomtang Instant Noodle Soup Mix"   
## [9] "King Lid Ramen Noodle Soup"   
## [10] "Big 3 Instant Ramen Bowl"   
## [11] "Mr. Kimchi Stirfried Kimchi Ramen"   
## [12] "Jjajang Buldak Ramen Topokki"   
## [13] "Carbo HOT Chicken Flavor Topokki"   
## [14] "Jjajang Ttoekbokki With Noodles"   
## [15] "Pho Rice Noodle HOT"   
## [16] "Spicy Sujebi"   
## [17] "Carbo Buldak Topokki"   
## [18] "Spicy Tteokbokki"   
## [19] "Hi-Myon Spicy Katsuo Udon"   
## [20] "Tsuyu Soy Sauce Udon"   
## [21] "Cheese Ramen"   
## [22] "Buldak Topokki Hot Chicken Flavor Ramen"   
## [23] "Spicy Gamjatang"   
## [24] "Buldak Ramen Topokki"   
## [25] "Bowl Noodle Soup Hot & Spicy Flavor"   
## [26] "Samyang Ramen Spicy Flavor"   
## [27] "Ramen Snack Chicken"   
## [28] "Samyang Ramen Hot"   
## [29] "Korean Gomtang"   
## [30] "Knight Of Ribs"   
## [31] "Budae Jjigae Ramen"   
## [32] "Neoguri Stir-fry Noodles Spicy Seafood"   
## [33] "Paegaejang Ramen"   
## [34] "Shin Ramyun Black"   
## [35] "Samyang Ramen Classic Edition"   
## [36] "Bul Jjamppong"   
## [37] "Jin Jjambbong Spicy Seafood Ramyun"   
## [38] "Buldak Bokkeummyun Snack"   
## [39] "Stew Buldak Bokkeumtangmyun"   
## [40] "Clam Flavor Soup Kal-guksu"   
## [41] "Bul Jjajangmyeon"   
## [42] "Gold Jjamppong Fried Noodle"   
## [43] "Budae Jjigae"   
## [44] "Non-Fried Ramyun Noodle (Crab Flavor)"   
## [45] "Chal Bibim Myun"   
## [46] "Rice Noodle Seafood Flavour"   
## [47] "Champong Noodle Soup Spicy Seafood Flavor"   
## [48] "Cheese Curry Ramyun"   
## [49] "Zha Wang ((Jjawang) Noodles With Chajang Sauce"   
## [50] "King Bowl Super Spicy Pan Stirfried Noodle"   
## [51] "Raobokki Noodle (Export Version)"   
## [52] "Hurricane Rice Cake (Topokki)"   
## [53] "Jinjja Jinjja (New)"   
## [54] "Jjajangmen Chajang Noodle King Bowl"   
## [55] "Jjamppong Seafood Noodle King Bowl"   
## [56] "Cheese Ramyun (for US market)"   
## [57] "Korean Traditional Beef Gomtangmen"   
## [58] "Soon Veggie Noodle Soup"   
## [59] "Red Nagasaki Jjampong"   
## [60] "Maesaengyitangmyun Baked Noodle"   
## [61] "Cheese Noodle"   
## [62] "Doong Ji Authentic Korean Cold Noodles With Chili Sauce"  
## [63] "Nagasaki Crab Jjampong"   
## [64] "Premuim Noodle House Crab Jjampong"   
## [65] "Namja Ramen (United States version)"   
## [66] "Rice Topokki"   
## [67] "Namja"   
## [68] "Bibim Men Cucumber"   
## [69] "Kokomen Spicy Chicken"   
## [70] "Ramen E Ramen White Broth"   
## [71] "Shin Ramyun Black Onion"   
## [72] "Jinjja Jinjja"   
## [73] "Kko Kko Myun"   
## [74] "DJ DOC Jjolmyeon"   
## [75] "DJ DOC Instant rice Ramen Topokki"

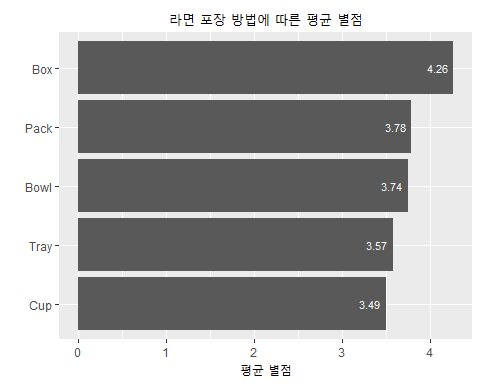
## 라면 포장 방법에 따른 종류 개수

ramen\_style\_df <- as.data.frame(table(df$Style))  
ramen\_style\_df <- ramen\_style\_df[-c(1, 2, 5, 8),]  
  
graph3 <- ggplot(ramen\_style\_df, aes(x = reorder(Var1, Freq),  
 y = Freq)) + geom\_bar(stat = 'identity') +   
 labs(title = "포장 방법에 따른 라면 개수") +  
 theme(plot.title = element\_text(hjust = 0.5))  
graph3 <- graph3 + coord\_flip() + labs(x='', y='라면 개수') +   
 geom\_text(aes(label=Freq), size=3, hjust=1.25, color='#FFFFFF')  
  
graph3



## 라면 포장 방법에 따른 평균 별점

ramen\_style <- c('Pack', 'Bowl', 'Cup', 'Tray', 'Box')  
ramen\_style\_mean <- c('')  
  
for (i in 1:5){  
 ramen\_style\_mean[i] <- round(mean(df[df$Style == ramen\_style[i],]$Stars),2)  
}  
  
ramen\_style\_mean\_df <- data.frame(ramen\_style, ramen\_style\_mean)  
ramen\_style\_mean\_df[2] <- as.double(ramen\_style\_mean)  
  
graph4 <- ggplot(ramen\_style\_mean\_df, aes(x = reorder(ramen\_style,   
 ramen\_style\_mean),  
 y = ramen\_style\_mean)) + geom\_bar(stat = 'identity') +   
 labs(title = "라면 포장 방법에 따른 평균 별점") +  
 theme(plot.title = element\_text(hjust = 0.5))  
graph4 <- graph4 + coord\_flip() + labs(x='', y='평균 별점') +   
 geom\_text(aes(label=ramen\_style\_mean), size=3, hjust=1.25, color='#FFFFFF')  
graph4

length(df[(df$Country == 'South Korea'),]$Variety)

## [1] 429