Case Study II

```
library(tidyverse)
```

```
## -- Attaching packages --- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.1     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

```
## -- Conflicts ------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
ramenData <- read.csv(file.choose(),header=T)
view(ramenData)
```

Question 1

```
nlevels(as.factor(ramenData$Country))
```

```
## [1] 7
```

There are 7 countries in this data set.

```
ramenData %>%
  group_by(Country) %>%
  summarise(n_brand=nlevels(as.factor(Brand)))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 7 x 2
##   Country     n_brand
##   <chr>         <int>
## 1 China            33
## 2 India             8
## 3 Japan            58
## 4 Singapore        10
## 5 South Korea      35
## 6 Taiwan           47
## 7 USA              49
```

These are the number of brands from each country.

Question 2

```
ramenData %>%
  filter(!is.na(Top.Ten)) %>%
  nrow()
```

```
## [1] 24
```

24 ramen have a “Top Ten” rating.

```
ramenData %>%
  group_by(Brand) %>%
```

```
  filter(!is.na(Top.Ten)) %>%
  summarise()
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 13 x 1
##    Brand
##    <chr>
##  1 A-Sha Dry Noodle
##  2 Koka
##  3 Maruchan
##  4 Myojo
##  5 Nissin
##  6 Nongshim
##  7 Paldo
##  8 Prima
##  9 Prima Taste
## 10 Samyang Foods
## 11 Sapporo Ichiban
## 12 Tseng Noodles
## 13 Wugudaochang
```

These are the brands that have earned a "Top Ten" rating.

Question 3

```
ramenData %>%
  group_by(Brand) %>%
  filter(!is.na(Top.Ten)) %>%
  summarise(count=n()) %>%
  arrange(desc(count))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 13 x 2
##    Brand            count
##    <chr>            <int>
##  1 Prima Taste          5
##  2 Nongshim             4
##  3 Paldo                3
##  4 Myojo                2
##  5 Sapporo Ichiban      2
##  6 A-Sha Dry Noodle     1
##  7 Koka                 1
##  8 Maruchan             1
##  9 Nissin               1
## 10 Prima                1
## 11 Samyang Foods        1
## 12 Tseng Noodles        1
## 13 Wugudaochang         1
```

Prima Taste is the ramen brand that had the most "Top Ten" ratings, having earned it 5 times in total. I personally have never had this brand before, nor have I ever heard of it.

Question 4

```
ramenData %>%
  group_by(Brand) %>%
  mutate(n_Stars=as.double(Stars)) %>%
  summarise(m_rating=mean(n_Stars,na.rm = T)) %>%
  filter(m_rating == 5.00) %>%
  count()
```

```
## Warning: Problem with `mutate()` input `n_Stars`.
## i NAs introduced by coercion
## i Input `n_Stars` is `as.double(Stars)`.
## i The error occurred in group 125: Brand = "Ottogi".
```

```
## Warning in mask$eval_all_mutate(dots[[i]]): NAs introduced by coercion
```

```
## Warning: Problem with `mutate()` input `n_Stars`.
## i NAs introduced by coercion
## i Input `n_Stars` is `as.double(Stars)`.
## i The error occurred in group 149: Brand = "Samyang Foods".
```

```
## Warning in mask$eval_all_mutate(dots[[i]]): NAs introduced by coercion
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1    18
```

A total of 18 brands have an average Star rating of 5.0 stars.

Question 5

```
ramenData %>%
  group_by(Country) %>%
  mutate(n_Stars=as.double(Stars)) %>%
  summarise(m_rating=mean(n_Stars,na.rm = T)) %>%
  arrange(desc(m_rating))
```

```
## Warning: Problem with `mutate()` input `n_Stars`.
## i NAs introduced by coercion
## i Input `n_Stars` is `as.double(Stars)`.
## i The error occurred in group 5: Country = "South Korea".
```

```
## Warning in mask$eval_all_mutate(dots[[i]]): NAs introduced by coercion
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
## # A tibble: 7 x 2
##   Country     m_rating
##   <chr>          <dbl>
## 1 Singapore       4.13
## 2 Japan           3.98
## 3 South Korea     3.79
## 4 Taiwan          3.67
## 5 USA             3.46
## 6 China           3.42
## 7 India           3.40
```

Singapore is the country that makes the best ramen. Our group defined the "best" as having the highest overall average Star rating because that means a majority of people who have had ramen from this country have rated it highly, which also means it has the best taste. And Singapore had the highest average Star rating at 4.13 stars. Thus, it must taste pretty good.

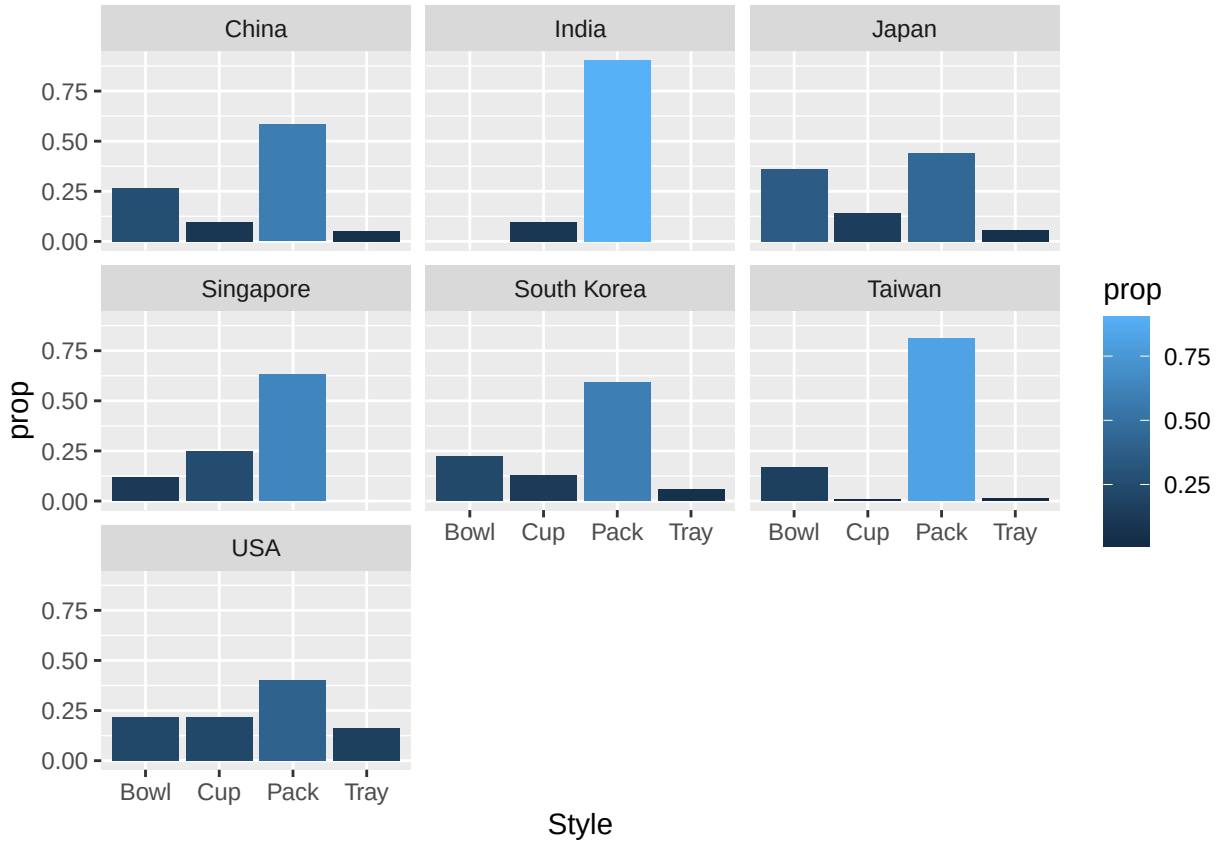
Question 6

Part A

```
data_style <- ramenData %>%
  filter(Style=="Bowl" | Style=="Cup" | Style=="Pack" | Style=="Tray")
```

Part B

```
data_style <- ramenData %>%
  filter(Style=="Bowl" | Style=="Cup" | Style=="Pack" | Style=="Tray") %>%
  ggplot()+
  geom_bar(aes(x=Style,fill=stat(prop),y=stat(prop),group=1))+
  facet_wrap(~Country)
data_style
```



Part C

Based on the plot above, the most popular Style of ramen is Pack. That seems to make sense as ramen that comes in a Pack is relatively the cheapest out of the other Styles.