

# Case Study I: Ramen

*Abbie M Popa*

For our first case study, we will analyze recent reviews of ramen! The data file we will investigate is located on the github course site. The assignment is worth 65 points total, you will lose 2.5 points for each hour it is late. Results should be turned in as a knit pdf, one file per group. Please include all group members names at the top of your document.

**Load** the data in the following way:

- 1) Download the .RData file to your computer
- 2) Load the data to your current R session. Change “dir” to be the location of the data on YOUR computer, then, run the following code

```
load("dir/ramen.RData")
```

The data contains reviews from Ramen Rater. It includes variables for Brand, Variety, Serving Style, Country of Origin, Number of Stars, and if it has ever won a place in the Top Ten Ramens of the Year.

**Directions:** Get into groups of 2 or 3 and answer the following questions, please show all code:

- 1) (3 pts) How many different brands are reviewed in the data set? Do you recognize any of them?
- 2) (3 pts) What years do we have “top ten” data from? (Hint: This will require turning the Top Ten column into a string column and using some of our string functions)
- 3) (3 pts) Which ramens are from the United States (hint: beware of abbreviations!)?
- 4) (3 pts) Has any brand of ramen won the #1 slot in the Top Ten Ramens list more than once? Which brand?
- 5) (3 pts) Which brand has the highest average star rating? How often has this brand won a slot in the top 10?
- 6) (5 pts) Do certain countries tend to package their ramen in a similar style?
- 7) (5 pts) What is the average, maximum, and minimum rating for ramens from the US? What about for ramens from Singapore? Based on your findings, which country do you expect to land in the Top Ten more often? Check your prediction.
- 8) (5 pts) Which country makes the most ramen? Which country makes the best ramen? How did you define “best”? Why? Was the country that makes the best ramen also the country that makes the most ramen?
- 9) (5 pts) How many ramens are considered spicy as part of their variety (hint, don’t forget that “spicy” and “Spicy” are both spicy)? Are spicy ramens typically rated higher than non-spicy ramens?
- 10) (5 pts) Based on this data set, what effect does saltiness have on ramen ratings?
- 11) (5 pts) Are there certain styles, countries, brands, or flavors of ramen that tend to use more or less salt? Feel free to be creative in how you parse “variety” into flavors.
- 12) [Open Ended] (10 pts) You have been hired by a new ramen start up, Roamin’ Ramen. They want to better understand the ramen market and ask you to break down ramen into 5 collections of “similar” ramens. How do you go about it? Discuss what you find.
- 13) [Open Ended] (10 pts) Roamin’ Ramen wants to make the next Top Ten Ramen winner. What recommendations do you have for them based on this data set? Be sure to use the columns provided, and generate your own ideas by examining the Variety column. Support your assertions with statistics (like mean, min, max or others you prefer) and at least one plot.