**Data Source (URL web address with hyperlink):**

https://www.animenewsnetwork.com/encyclopedia/ratings-anime.php?top50=best\_bayesian&n=250

**Context of Data and Variables:**

Data: Anime Top 250 Best Rated (Bayesian estimate)

**Output from str() function applied to the data object (apply a monospaced font like “Courier New” to the output):**

$ rank : int [1:250] 1 2 3 4 5 6 7 8 9 10 ...

$ title : chr [1:250] "Fullmetal Alchemist: Brotherhood (TV)" "Steins;Gate (TV)" "Clannad After Story (TV)" "your name. (movie)" ...

$ rating : num [1:250] 9.09 9.05 9.05 9.03 8.97 8.93 8.92 8.89 8.89 8.86 ...

$ nb.vote: num [1:250] 5848 4685 5186 1141 6656 ...

**Research Questions to be explore:**

**1.Is the piece that has high population also get high rank?**

**2.What are the pieces getting more than 10,000 votes?**

**Statistical Analysis Plan**

**Population:**

**all pieces in the data list**

**Primary Objective:**

**finding the trends of pieces, i.e. whether piece that has high population also get high rank**

**Secondary Objectives:**

**finding the mode the pieces(What are the pieces getting more than 10,000 votes)**

**Data Collection methods:**

**rvest package scarp data from www.animenewsnetwork.com**

**Variables under consideration:**

**rank, name, rating, nb of voting**

**Missing data procedures: If any data is missing, it will be excluded.**

**Numerical and graphical summaries to be presented: use basic R function and tidyverse function, such as histogram(), summarise()**

**Models to be fitted: linear model**