

BIOST 509: In-Class Exercise 2

Instructor: Amy Willis, Biostatistics, UW

Due date: 5:00pm on October 11, 2019 via Canvas

Instructions

All questions for this homework relate to the dataset “drug_reviews.csv”. This dataset contains reviews of drugs posted online. The drug name (`urlDrugName`), the rating given by the reviewer (`rating`), effectiveness of the drug given by the reviewer (`effectiveness`), as well as `sideEffects`, the `condition` it was intended to treat, reviews of the benefits (`benefitsReview`), reviews of the side effects (`sideEffectsReview`), and other comments (`commentsReview`).

The dataset is available from Canvas in **Pages/Week 2 materials** or **Files/datasets**.

Read in the dataset and answer the following questions. Submit your answers to the below questions in a R Script (.R), Word (.doc or .docx) or pdf file to Canvas. Don’t forget to include both your code and answers in your response.

Questions

1. What were the levels of effectiveness observed in the data?
2. How many different drugs are recorded in the dataset?

Answer the following questions *using the **tidyverse** functions introduced in lecture* (`filter`, `group_by`, `summarize`, etc.).

3. What was the mean rating of all drugs in the dataset?
4. What was the maximum rating given to a drug reported as “Ineffective”?
5. What was the mean rating for each level of effectiveness?
6. Optional: Which drug was observed to be “Ineffective” the largest number of times?