

This lab uses the `fake_news` data from the `bayesrules` library.

## I. Spot Fake News With This One Simple Rule!

1. Get help on the `fake_news` data. When and where are these articles from?
2. We are primarily interested in the `type` of an article: whether it is real or fake news. Whether or not an article has exclamation points in its title (`title_has_excl`) can tell us something about the type.
  - (a) Make a summary of the number of articles of each type and how many of each have exclamation points in their titles.
  - (b) Make a visualization to show the difference in exclamation point usage between real and fake news.
3. Let  $B$  be the event “an article is fake.” Use the data to estimate  $P(B)$  and  $P(B^c)$ . These are your **priors**.
4. Let  $A$  be the event “the article has an exclamation point in its title.” Use the data to estimate  $P(A|B)$  and  $P(A|B^c)$ . This is the **likelihood function**  $L(\cdot|A)$ .
5. Use the **law of total probability**  $P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$  to compute  $P(A)$  and then check that the result matches  $P(A)$  if you estimate it directly from the data.
6. Suppose you observe an article with an exclamation point in the title. Use Bayes’ Rule to compute the **posterior** probability  $P(B|A)$  that the article is fake.

## II. Florida Man Angry About Long Headlines

Let’s look for other variables that may indicate a difference between real and fake news.

1. Make a visualization to compare the `anger` sentiment score between real and fake news.
2. Make a visualization to compare the number of words in the title (`title_words`) between real and fake news.
3. Explore other variables. Do any seem to be strong indicators of fake news?

## III. Are We STILL Dealing With Fake News?!

1. Create a new variable `title_has_caps` which is true when the title contains a word in all capitals and false otherwise. Make a table summarizing this variable and the article `type`.
2. Suppose you observe an article with a word in all caps in its title. Repeat the analysis in Part I to compute the posterior probability that the article is fake.
3. Suppose you observe an article whose title has both a word in all caps and an exclamation point. What is the posterior probability that the article is fake?