

New Yorker cartoons exercise:

This data, `new_yorker_cartoons.xlsx`, was published publicly by researcher Matt Michel. He used it for an academic research project studying how cartoonists portrayed different characters – the racial and gender identity and other things. It only includes cartoons with human characters. Notice that each row in the dataset represents one character – so there might be multiple rows for each cartoon. This is a good example of how you can build your own dataset and be able to say something definitive from it.

Here is their methodology:

Methods

Our subscription to The New Yorker started in 2014, so we analyzed the gender and race of the cartoon characters in all 47 issues of 2014 (sometimes The New Yorker publishes one issue that covers two weeks). Issues that were not available from the magazine rack in my bathroom were obtained through The New Yorker's online archive.

For each cartoon we recorded the race ("white" or "non-white"), gender ("male" or "female"), occupation (as best we could) and if the character was talking. The name of the cartoonist was also recorded. Only cartoons with humans were included. We focus exclusively on gender and racial statistics here and will discuss the other recorded data in a future article.

While gender was easy to discern for most characters (excluding babies), race was not as simple. We settled on a simple rule of recording every character that had some sort of shading (that was not due to a shadow) as a non-white character. If in doubt of a character's race (which was rare), we recorded them as non-white (so our estimates are likely overestimates).

Assignment:

Using Pivot Tables (plus some additional calculations), answer the following questions:

- 1) What percentage of the characters were non-white?
- 2) What percentage of the characters were women?
- 3) Which cartoonist had the highest percentage of non-white characters?
- 4) What percentage of the characters were both non-white and female?
- 5) Write a paragraph that summarizes your key findings. (you can use the results from questions above and/or do your own additional analysis)