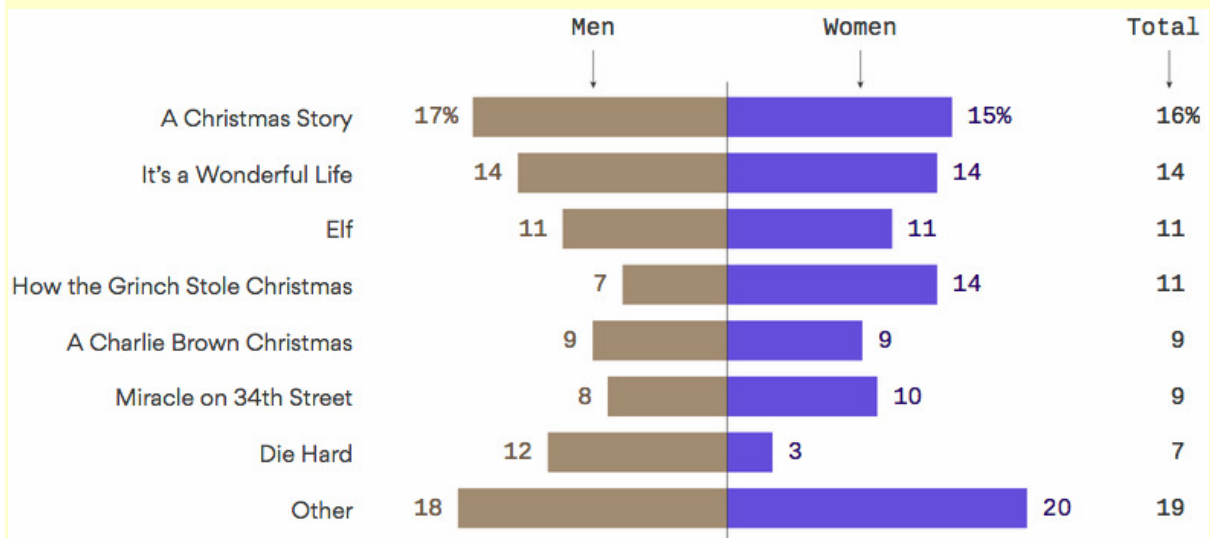# Strange Statistics

(24 December 2017)

Jim Stevenson

Reading Axios on Christmas Eve day, I was struck by what appeared at first to be a strange graph showing preferences for Christmas movies divided between men and women. The thing that struck me as strange was the computation for the total votes: the percentages were the average of the men and women percentages:

(https://www.axios.com/americas-favorite-christmas-movies-2519380482.html, retrieved 12/24/2017)

America's Favorite Christmas Movies

David Nather, 24 December 2017 06:00 AM ET

What, not serious enough for you? Too bad. We're betting that this Axios/SurveyMonkey poll is as serious as most people want to get today. Here's a look at everyone who didn't list "Die Hard" as their favorite Christmas movie.

	Men	Women	Total
A Christmas Story	17%	15%	16%
It's a Wonderful Life	14	14	14
Elf	11	11	11
How the Grinch Stole Christmas	7	14	11
A Charlie Brown Christmas	9	9	9
Miracle on 34th Street	8	10	9
Die Hard	12	3	7
Other	18	20	19

Data: SurveyMonkey,[1] conducted Dec. 19-21, 2017; Chart: Andrew Witherspoon / Axios

Everyone knows you don't combine percentages as if they were regular numbers. They represent fractions of a whole. That is, 25% of 100 people is not the same as 25% of 80 people (25 in the first case and 20 in the second). In general, if M is the number of men and F the number of women, and if the survey said 30% of men voted for an item and 20% of women did the same, then the total percentage would be

$$\frac{.3M + .2F}{M + N} \times 100$$

So if there were 100 men and 50 women, the resulting percentage would be 100 x (30 + 10)/150 = 27%, whereas the average of 30 and 20 (percent) is 25 (percent).

---

[1]   https://www.surveymonkey.com/mp/survey-methodology/

But I wondered if the sampling used the same number of women as men, since then M = F and the resulting percentage would be

$$\frac{.3 + .2}{2} \text{ x } 100 = 25$$

or just the average of the two percentages.

So I checked the link to SurveyMonkey and sure enough, they adjusted the samples to represent an equal number of men and women:

4. **Our survey scientists carefully adjust the data so that it's representative of the population.**

5. **What does that mean? When some groups in our sample don't exactly match the larger population, we use advanced statistics to balance them.**
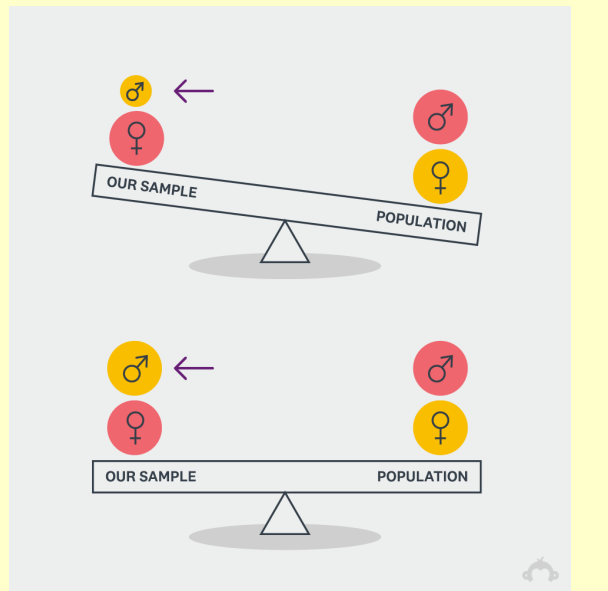
So in this special case, one can average the percentages of two equal samples to get the total percentage.

© 2018 James Stevenson