

# Social Categorizations in CRAN Data

*ASA OR Spring Meeting, May 21 2020*

Simon Couch |    @simonpcouch



# To Start, A Question...

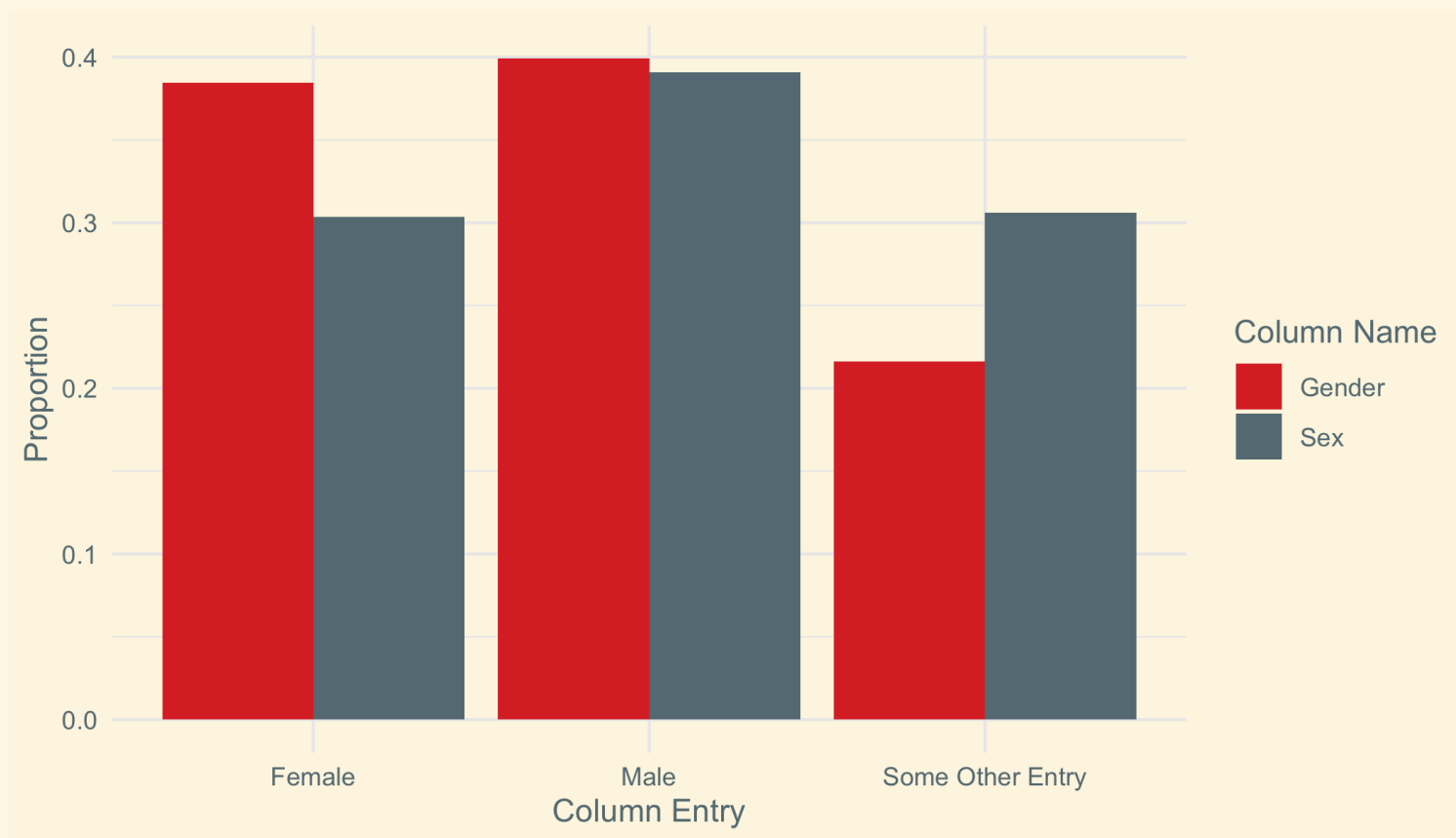
# To Start, A Question...

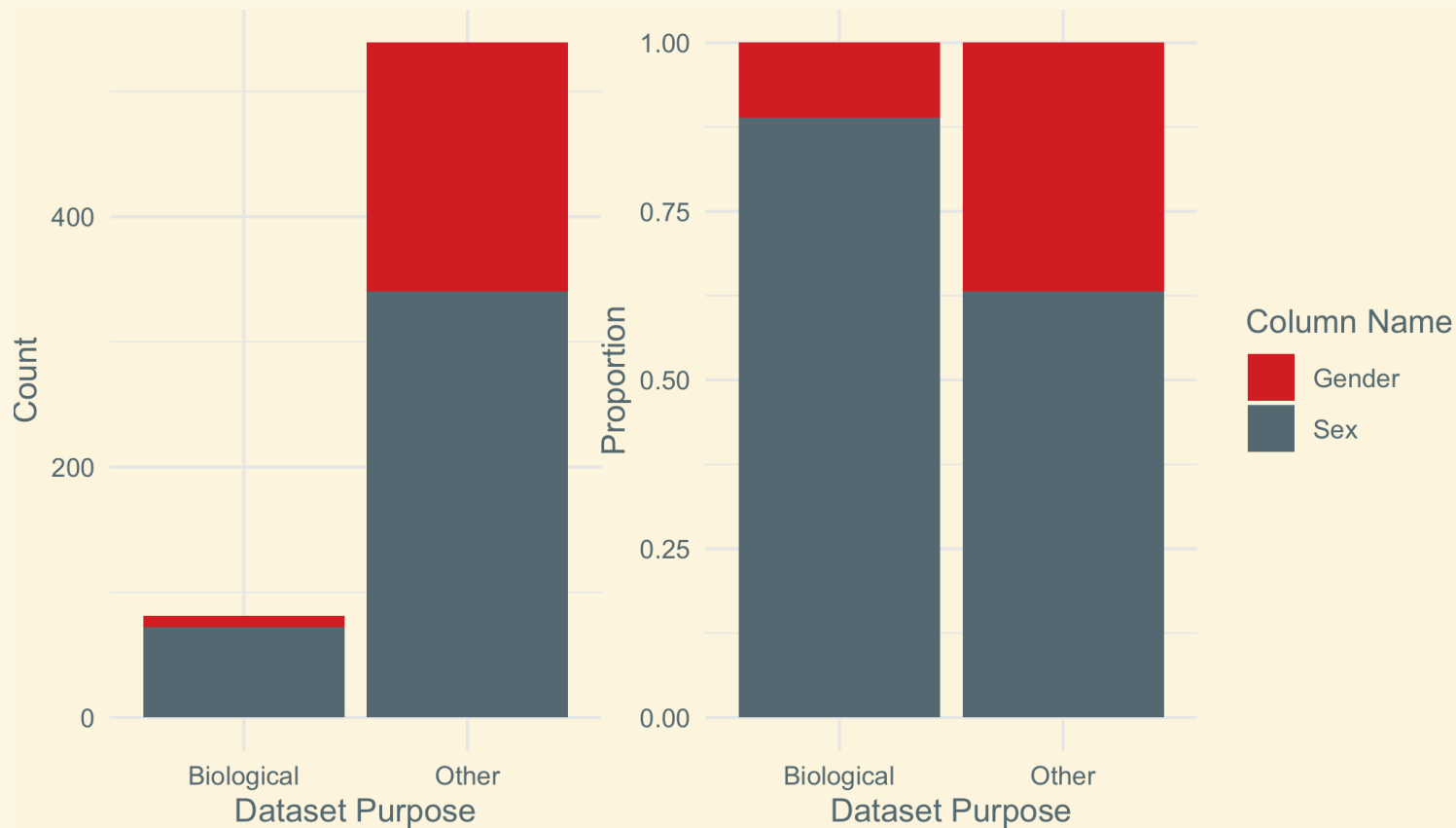
*A colleague comes to you with a dataset containing a column called "Gender"—what are the 2 most common entries in that column?*

# To Start, A Question...

*A colleague comes to you with a dataset containing a column called "Gender"—what are the 2 most common entries in that column?*

*Suppose, instead, the column was called "Sex"—does that change your answer?*





*"[W]hen handling the complex issue of group identity, we should look to purposes and effects, consequences and functions... what must be addressed is who is defining, how is the definition constructed, and why is the definition being propounded."*

- Cheryl Harris, *Whiteness as Property* (1993, pp. 1763-4)

# Thank You!

Special thanks to Kjersten Whittington, Yaejoon Kwon, Kristin Bott, and Kelly McConville for their valuable insight throughout this project.

Source code, bibliography, full paper, and slides available on GitHub at:  
*[github.com/simonpcouch/social\\_divisions\\_in\\_data](https://github.com/simonpcouch/social_divisions_in_data)*

   @simonpcouch