

Fair's Extramarital Affairs Report

Andrea Marcela Huerfano Barbosa

August 18th of 2019

Description

Cross-section data from a survey conducted by Psychology Today in 1969. The main topic is infidelity data, known as Fair's Affairs. The dataset used in this study containing 601 observations on 9 variables. Available in: <http://pages.stern.nyu.edu/~wgreene/Text/tables/tablelist5.htm>

Variables

1. *Affairs*: How often engaged in extramarital sexual intercourse during the past year?
2. *Gender*: Indicating female or male.
3. *Children*: Are there children in the marriage?
4. *Age*: Numeric variable coding age in years:
 - 17.5 = under 20
 - 22 = 20-24
 - 27 = 25-29
 - 32 = 30-34
 - 37 = 35-39
 - 42 = 40-44
 - 47 = 45-49
 - 52 = 50-54
 - 57 = 55 or over.
5. *Religiousness*: 1 = anti, 2 = not at all, 3 = slightly, 4 = somewhat, 5 = very.
6. *Education*:
 - 9 = grade school
 - 12 = high school graduate
 - 14 = some college
 - 16 = college graduate
 - 17 = some graduate work
 - 18 = master degree
 - 20 = Ph.D., M.D., or other advanced degree.
7. *Occupation*: Numeric variable coding occupation according to Hollingshead classification (reverse numbering).
8. *Rating*: Self rating of marriage:
 - 1 = very unhappy.
 - 2 = somewhat unhappy
 - 3 = average
 - 4 = happier than average
 - 5 = very happy.
9. *Years married*: Number of years married
 - 0.125 = 3 months or less
 - 0.417 = 4-6 months
 - 0.75 = 6 months-1 year
 - 1.5 = 1-2 years

- 4 = 3-5 years
- 7 = 6-8 years
- 10 = 9-11 years
- 15 = 12 or more years.

Exploratory Descriptive Analysis

From the plot below, we can realize that the most common number of affairs in the past year is 0, perhaps more interesting is that a few people have had 1,2,3,7 or 12 affairs in the past year. The distribution between men and women is almost the same as we can see in the pie chart because the percentage are so similar.

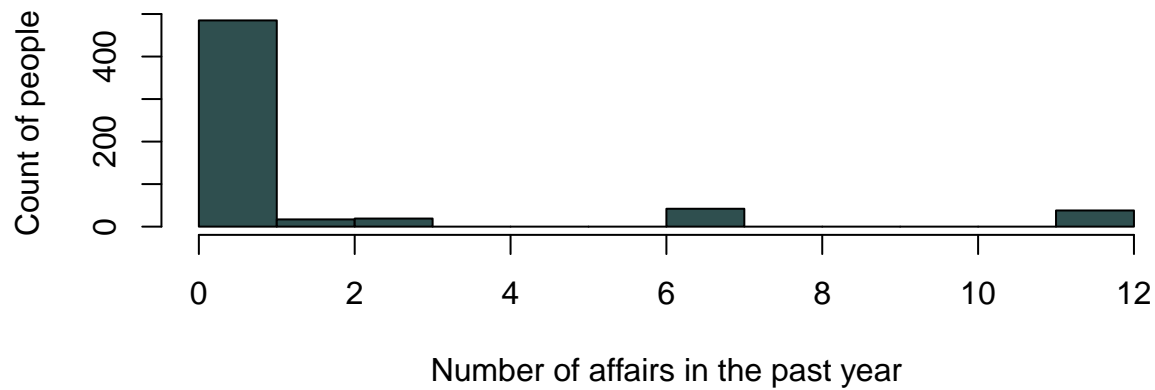


Figure 1: Distribution of affairs

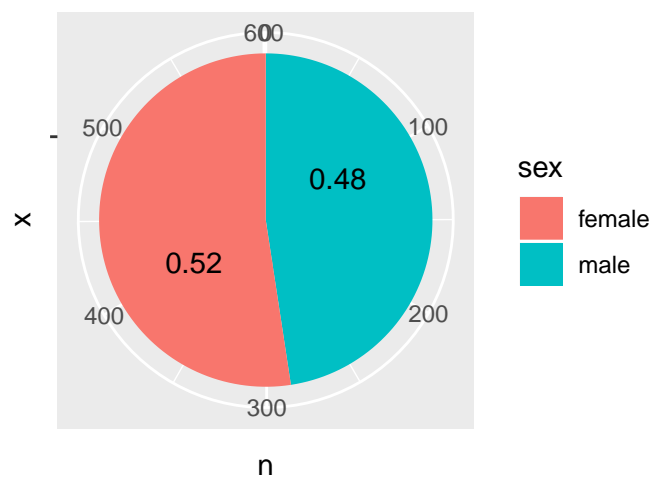


Figure 2: Population distribution by gender

In Figure 3 we can see a similar the behavior for the number of affairs by gender because the majority people in each level didn't have affairs in the past year, futhemore a few amount of women and men have had between 1 and 12 affairs.

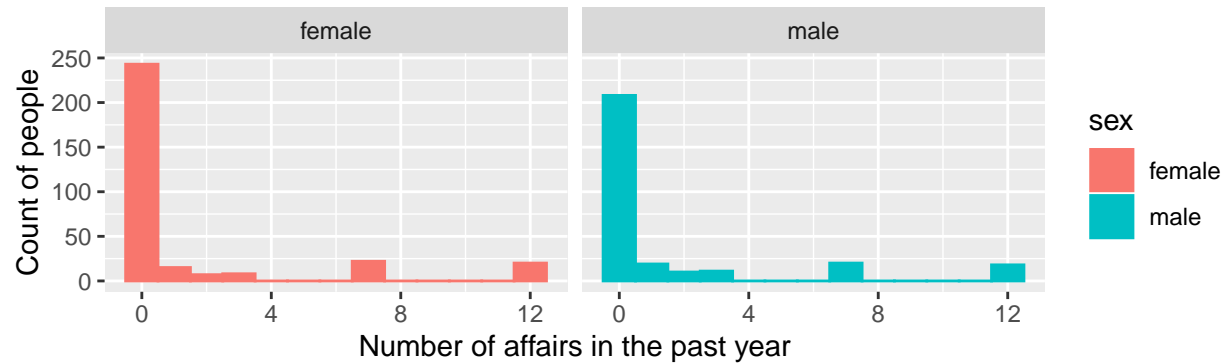


Figure 3: Plot of the number of affairs in the past year by gender

From the Figure 4, we can realize that women with low educational levels (9-15 asociated with grades school and some college) have more affairs than men in the same educational level. Perhaps, for educational level higher than 16(college graduate), men tend to have more affairs than women in the same educational level.

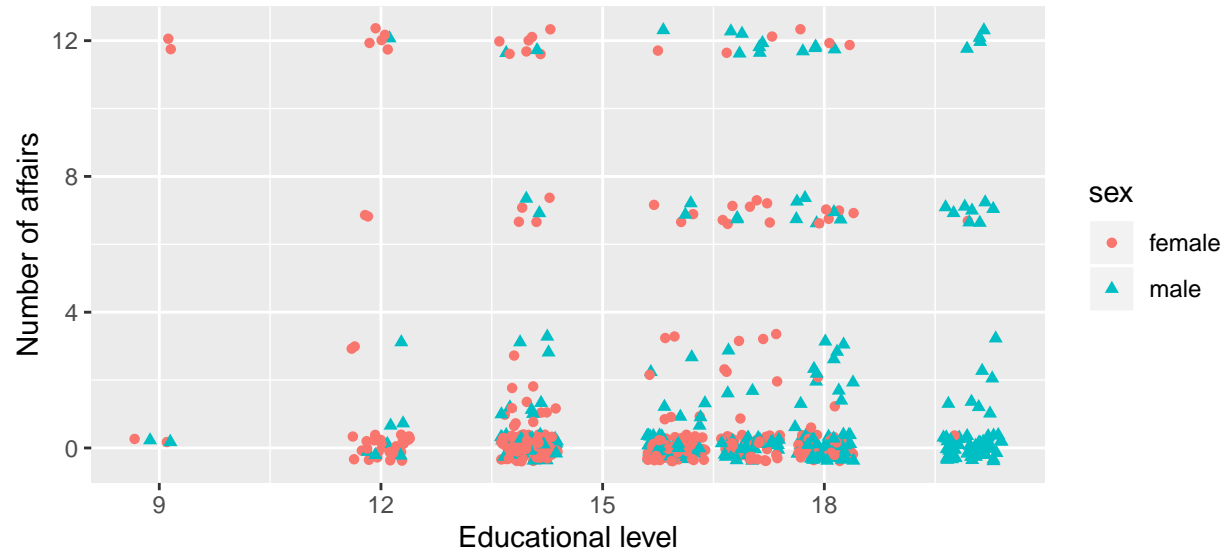


Figure 4: Number of affairs for each combination of educational level and gender

From Table 1, we can see quiet clearly that people who have children tend to have more affairs that people without children. The table shows the proportion in these levels.

Table 1: Have children

Child	%
no	0.28
yes	0.72

From the Figure 5, when the rating of marriage increases people tends to have less number of affairs, perhaps more interesting is that there are a great difference when people rate their marriage with 2 (somewhat unhappy) or 3(average) because the number of affairs tend to decrease strongly between these levels another interesting fact is when people rate their marriage with 4(happier than average) or 5(very happy) for the 75% of the marriages in these levels there are not affairs in the past years.

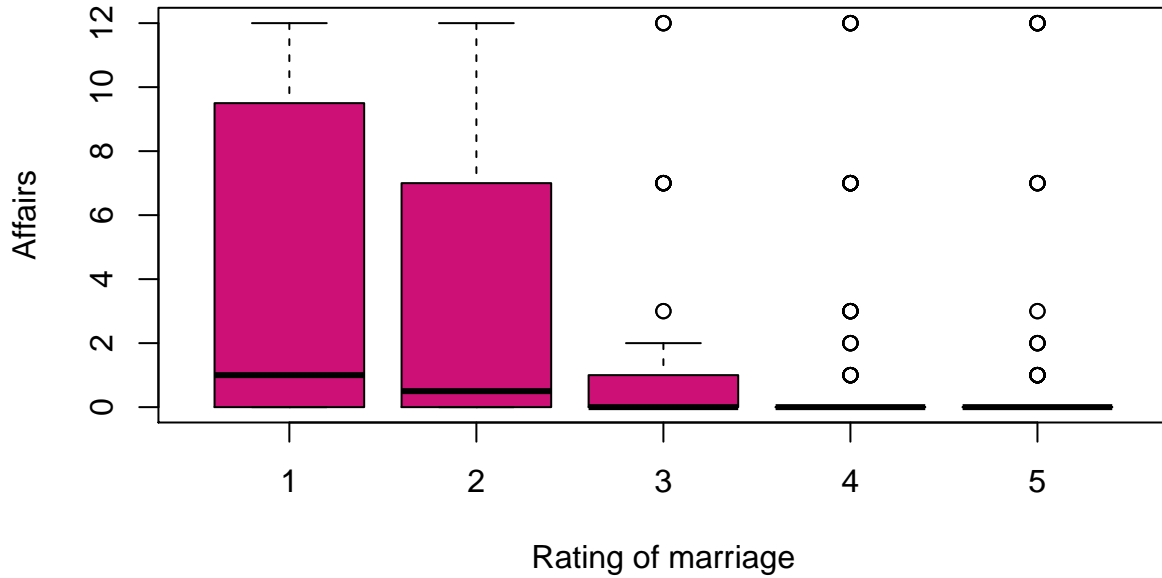
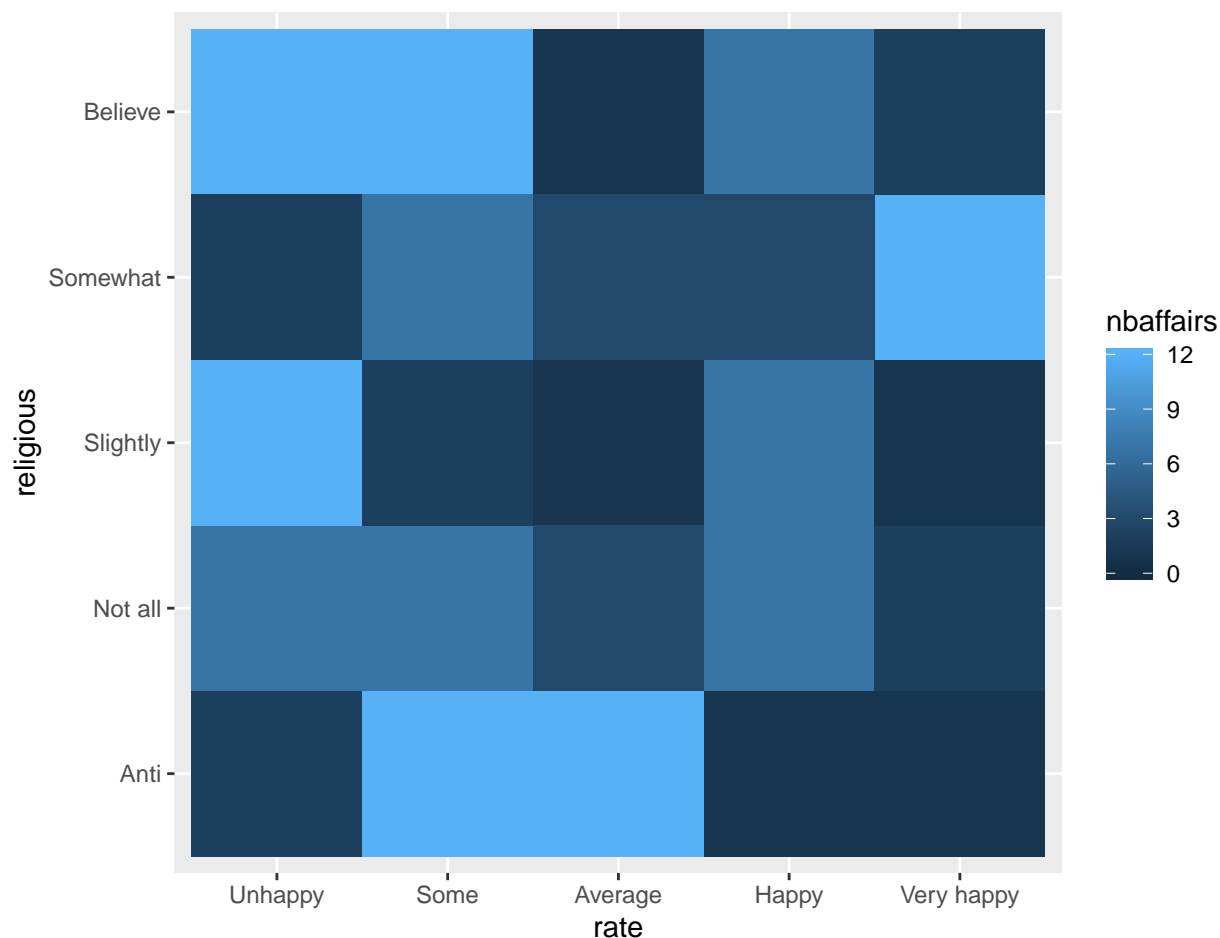


Figure 5: Boxplot for the number of affairs by the rating of marriage

The maximum number of affairs by any combination of rate and religious level appears in the next heatmap.



References

- Yves Croissant (2016). Ecdat: Data Sets for Econometrics. R package version 0.3-1. <https://CRAN.R-project.org/package=Ecdat>
- H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.
- G. Grothendieck (2017). sqldf: Manipulate R Data Frames Using SQL. R package version 0.4-11. <https://CRAN.R-project.org/package=sqldf>
- Sarkar, Deepayan (2008) Lattice: Multivariate Data Visualization with R. Springer, New York. ISBN 978-0-387-75968-5
- Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2019). dplyr: A Grammar of Data Manipulation. R package version 0.8.3. <https://CRAN.R-project.org/package=dplyr>
- Diethelm Wuertz, Tobias Setz and Yohan Chalabi (2017). fBasics: Rmetrics - Markets and Basic Statistics. R package version 3042.89. <https://CRAN.R-project.org/package=fBasics>
- Hao Zhu (2019). kableExtra: Construct Complex Table with 'kable' and Pipe Syntax. R package version 1.1.0. <https://CRAN.R-project.org/package=kableExtra>