

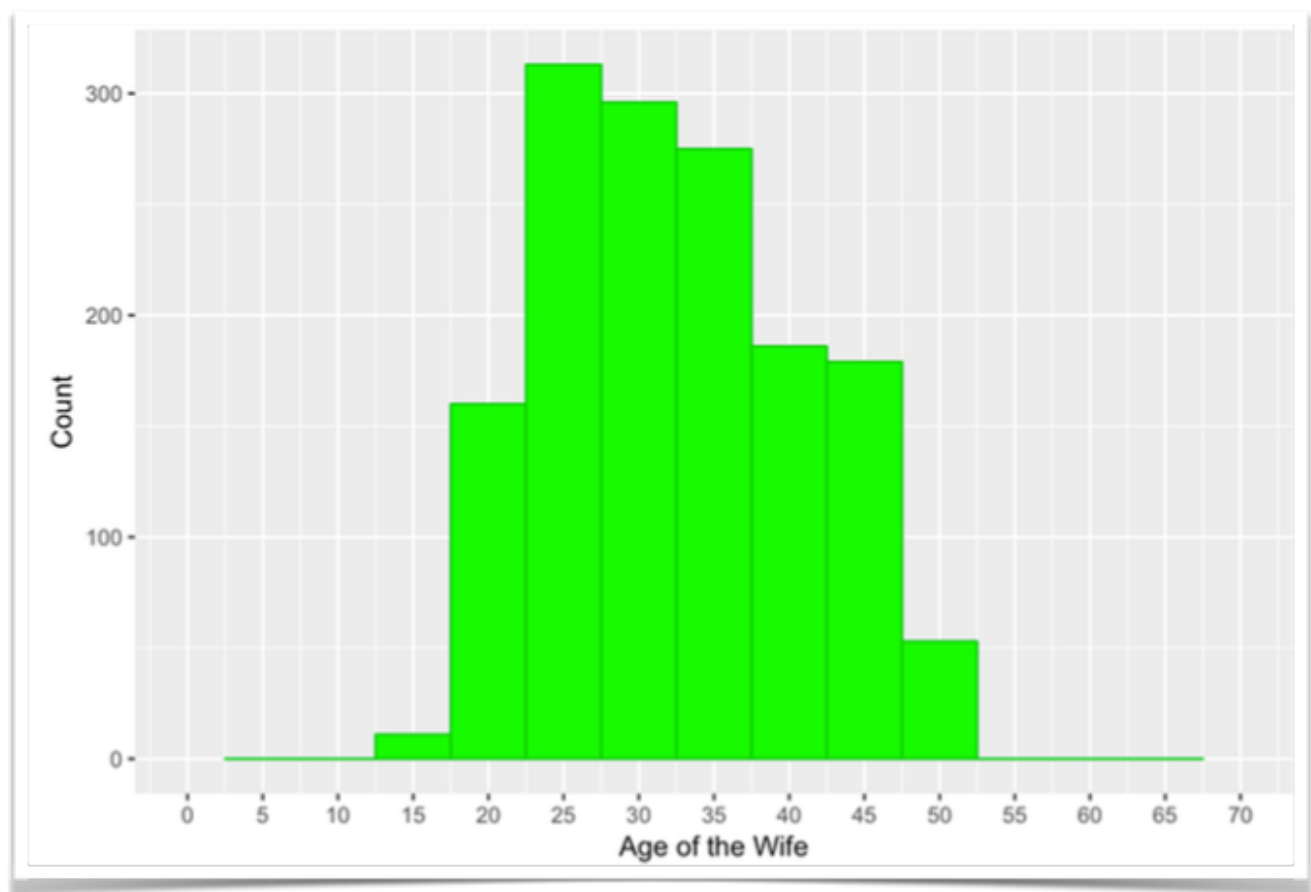
---

# Contraceptive Method Choice

## An Initial Data Analysis On An Indonesian Survey

Laura Cortes - September 21, 2017

---



---

## Introduction

This dataset is a subset of the 1987 National Indonesia Contraceptive Prevalence Survey. The subset was created on June 1997 by Tjen-Sien Lim. The samples are married women who were not pregnant or do not know if they were at the time of the interview.

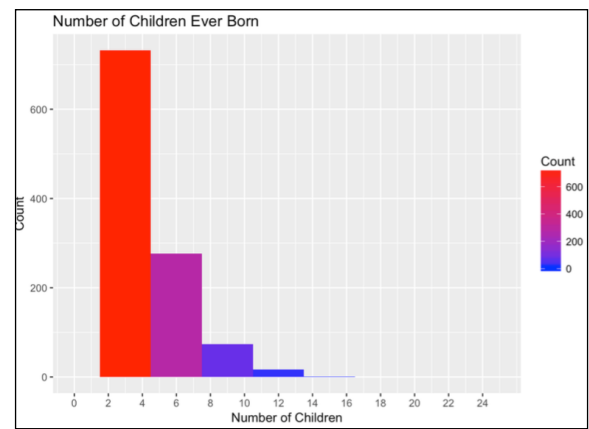
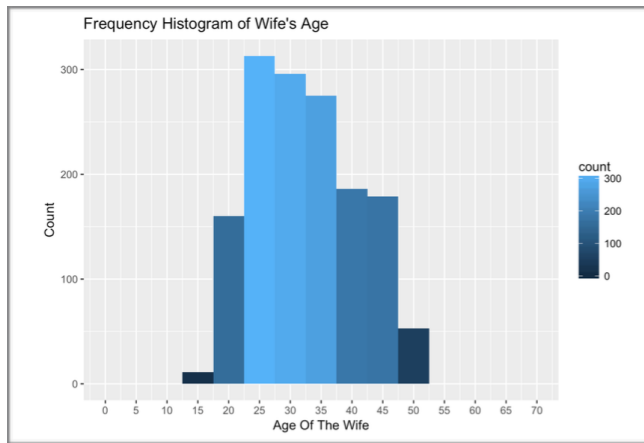
| CMC DATASET                  |                 |                                     |
|------------------------------|-----------------|-------------------------------------|
| Attribute Overview           |                 |                                     |
| Information                  | Type            | Values                              |
| Wife's age                   | numerical       |                                     |
| Wife's education             | categorical     | 1=low, 2, 3, 4=high                 |
| Husband's education          | categorical     | 1=low, 2, 3, 4=high                 |
| Number of children ever born | numerical       |                                     |
| Wife's religion              | binary          | 0=Non-Islam, 1=Islam                |
| Wife's now working           | binary          | 0=Yes, 1=No                         |
| Husband's occupation         | categorical     | 1, 2, 3, 4                          |
| Standard-of-living index     | categorical     | 1=low, 2, 3, 4=high                 |
| Media exposure               | binary          | 0=Good, 1=Not good                  |
| Contraceptive method used    | class attribute | 1=No-use, 2=Long-term, 3=Short-term |
|                              |                 |                                     |

## The project

The goal is to predict the contraceptive method of choice of a woman based on her demographic and socio-economic characteristics. The method of choice could be no-use, long-term methods, or short-term methods.

### 1. Numbers

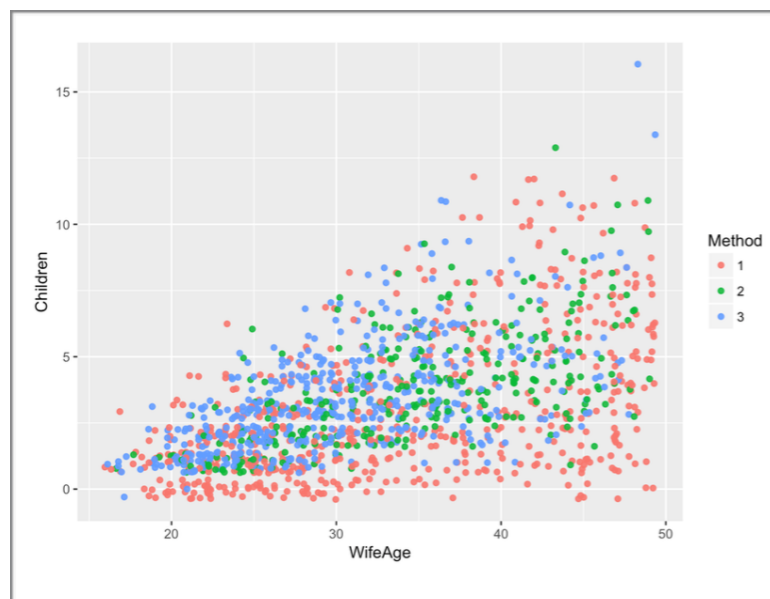
The histograms about wife's age and about their number of children ever born show that a good number of women are 25 to 35 years old, and that most of them have gave birth



to 2 to 4 children. By applying the summary function, we also get to know that most of the women practice Islam, do not work, and that more than half of them use no contraceptive method.

## 2. Initial Findings

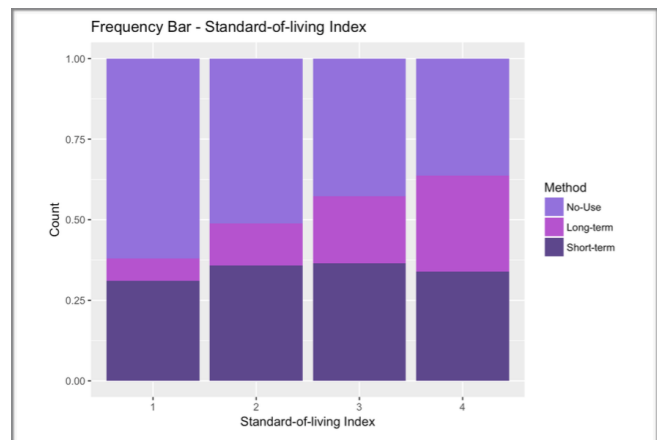
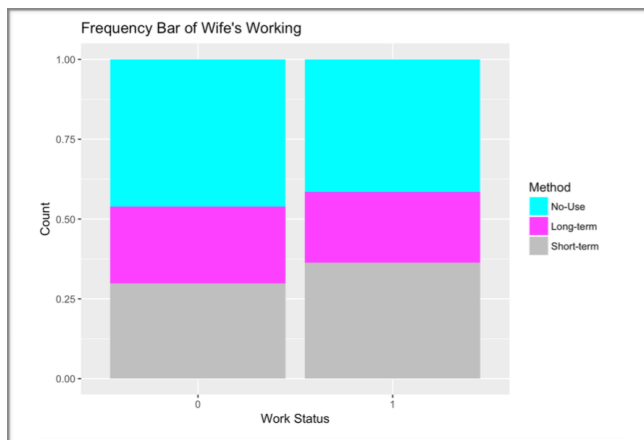
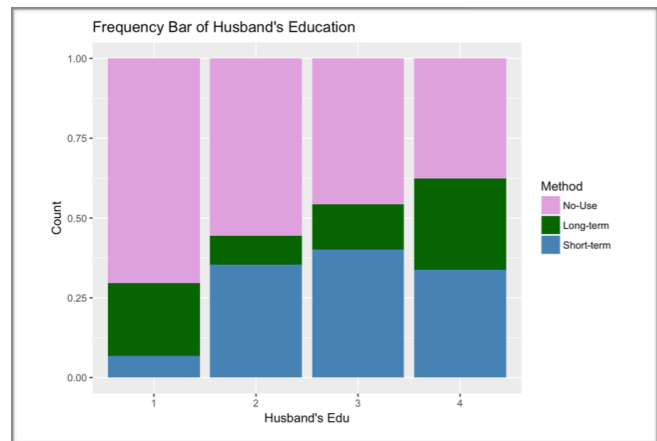
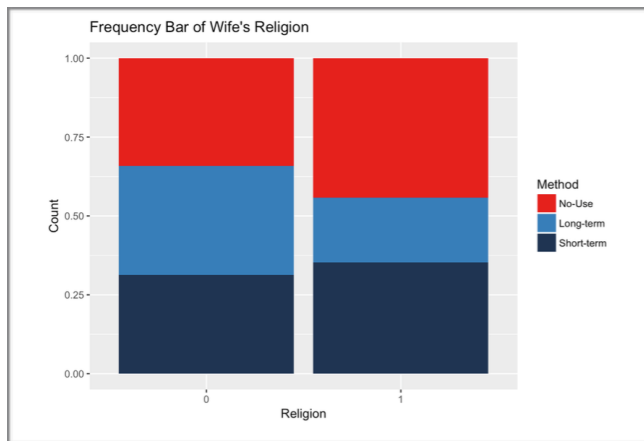
It is difficult to identify a trend as 8 out of 10 of our attributes are not numerical. The scatter plots that we get, for example age on number of children, having the contraceptive method as a color, show a weak distribution. This correlation we will need to investigate further.



further.

### 3. Some Hypotheses (Alternative)

- If the woman practice Islam, she will not use a contraceptive method
- If the husband & wife are educated, she will use a contraceptive method
- If the wife works, she will use a contraceptive method
- The higher the couple's standard-of-living index, the more they will use a method



The frequency bars above, except the third one, seemed to reject the null hypothesis, making us consider the first, second, and fourth alternative hypotheses we made. Thus, next we used Pearson's Chi-squared test on each of the hypotheses, and the results showed the same: only the null hypothesis related to the third one was not rejected. The fact that the wife works does not affect the method of her choice.

---

### 3. Next Steps

We will try to use create mosaic plots for these hypotheses to see how they look like.