Woman gives birth relate to their age and marital status, but not their race

Jiawei Du

Lin Zhu

Siri Huang

Wang Xinyu

2020-10-19

Link: https://github.com/christy723/Woman-gives-birth-relate-to-their-age-and-marital-status-but-not-their-race-.git

Abstract

In this report, we are interested in finding a significant relationship between Quebec women's desire to give birth and their age, marriage, and if they belong to aboriginal people or visible minorities. We use the logistic model to analyze the data, and base on the result we found later, age and marriage are two factors that have an apparent relationship to a woman's desire to give birth. However, there is weak evidence to show woman give birth relate to their race. This weak evidence for the race is essential as it tells us that if aboriginal or visible minority won't influence Quebec women's desire to give birth to a child.

Introduction

Every day, some mothers give birth to their new babies in Quebec, the mothers may have different backgrounds. Some of them are very young, while others are in their middle 40s, some of the mothers belong to aboriginal people while others are not. Most of the mothers are whites, while some mothers belong to a visible minority. Since Canada is a country with multicultural and lots of people have backgrounds, we want to predict the relationship between the mother's desire to give birth to a child and their race, their age, and their marital status. We choose to research this relationship for women in Quebec.

We use Canada General Social Survey to be our data set. Based on our data set, we make a logistic model explore the relationship between Quebec women give birth to a child and their age, marital status and their race. We use R function to help us analyst the ties and finally find only age and marital status have a significant relationship. The report can help socialist better understanding women's desire to pregnant and their background, but there is still a weakness for the survey. If we want to know more detail information about women and if they ever give birth, we should get detail information about where they originally come from.

Data

The data is based on the Canada General Social Survey. The questionnaire is more affordable to gather quantitative information, especially for a large data set, and it is one of the quickest ways to get results. Moreover, the questionnaire is anonymous, allowing respondents to answer questions more truthfully and protect respondents' privacy. However, there are still disadvantages to the use of the questionnaire. We can't force respondents to answer each question correctly, and they may leave empty for some items. Moreover, if we can't explain our problems clearly, they may misunderstand the questions.

The General Social Survey collects results from those people who live in Canada and older than 14 years old. The frame is a list that contains all people to satisfy the two requirements. The sample here is 20602 people who did this survey and send it to us. In this survey, we regard each province as a stratum, and every reply has been recorded to the stratum base on the area where they live. There are several ways to make up the frame, like population Census, billing data, and other data files, Statistics Canada will get a list of all this information from sources like the telephone companies. The survey frame was also made up by the address people register. They use the register address to connect telephone with same address.

People receive the phone call for the interview, and those who do not use the telephone would not be included in this interview. Only half of the interviewers answer the questionnaire correctly. A questionnaire with non-response was not included in the data set. To receive more replies and avoid too much non-response, they will interview more people, and usually, the number of phone calls we made is more than twice the sampling number we thought about before. Moreover, they try to shift the weights of non-response or mis responding.

Model

In this report, we are considering logistic model. Let y_i be the ever_given_birth variable, it represents if the person has given birth to a kid or not.

$$y_i = \begin{cases} 1, & \text{if the } i_{th} \text{ person gives birth to a child} \\ 0, & \text{if the } i_{th} \text{ person not gives birth to a child} \end{cases}$$

Since $y_i \sim bern(p_i)$ where $p_i \sim Pr(y_i = 1)$, we can get our model below:

$$\log \frac{p_i}{1 - p_i} = \beta_0 + \beta_1 x_{age,i} + \beta_2 x_{aboriginal,i} + \beta_3 x_{visminority,i} + \beta_4 x_{evermarried,i}$$

where

$x_{age,i}$ is the age of the ith person

$$x_{aboriginal,i} = \begin{cases} 1, & \text{if the } i_{th} \text{ person is aboriginal} \\ 0, & \text{if the } i_{th} \text{ person is not aboriginal} \end{cases}$$

$$x_{visminority,i} = \begin{cases} 1, & \text{if the } i_{th} \text{ person is visible minority} \\ 0, & \text{if the } i_{th} \text{ person is not visible minority} \end{cases}$$

$$x_{evermarried,i} = \begin{cases} 1, & \text{if the } i_{th} \text{ person married before} \\ 0, & \text{if the } i_{th} \text{ person not married before} \end{cases}$$

We use RStudio to run our model.

Here we use age instead of age group because we think there might be a noticeable difference for each period.

The other three variables are categorical because suppose every one responds to the questionnaire correctly; the answer will only be “yes” or “no”. Therefore, we assign the value of “yes” response be one and the value of “no” be 0 to separate them. Moreover, change the amount to be 0 or 1 can help us more clearly analysis the model using RStudio.

Result

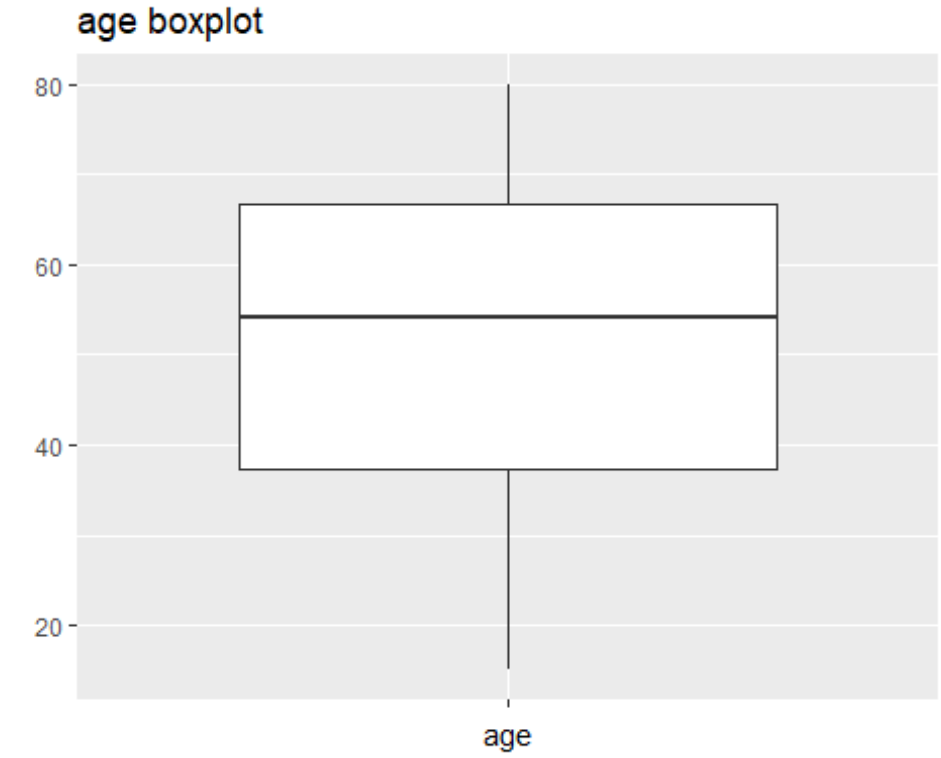

Figure 1. age boxplot

From Figure 1, we find that the median age of women in Quebec is around 54.

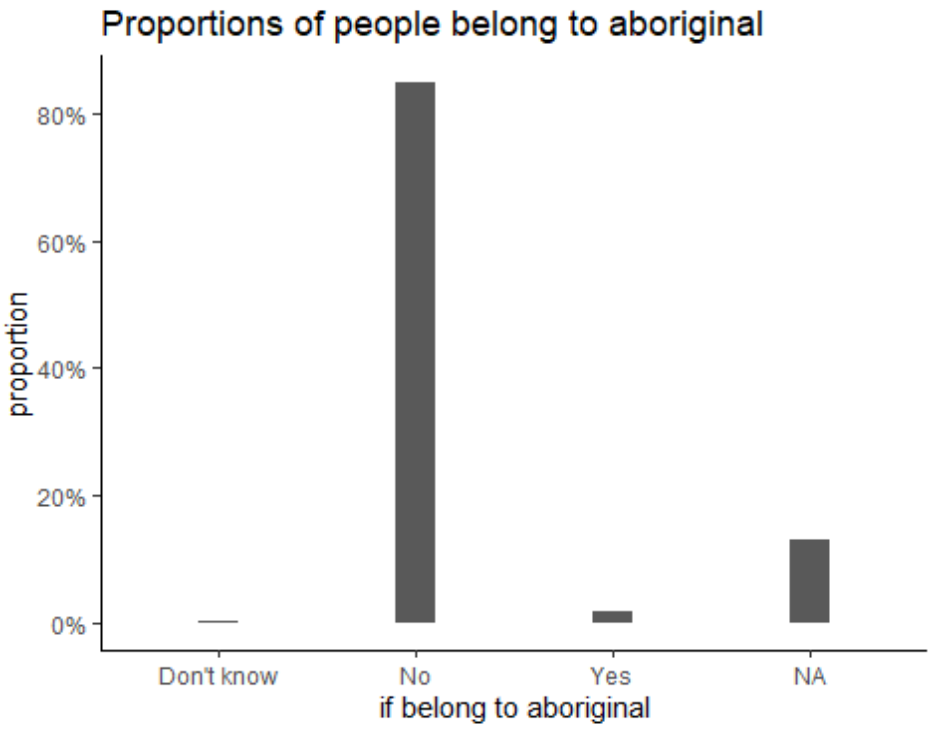

Figure 2. Proportions of people belong to aboriginal

Figure 2, we find that only around 4% women in Quebec belong to aboriginal.

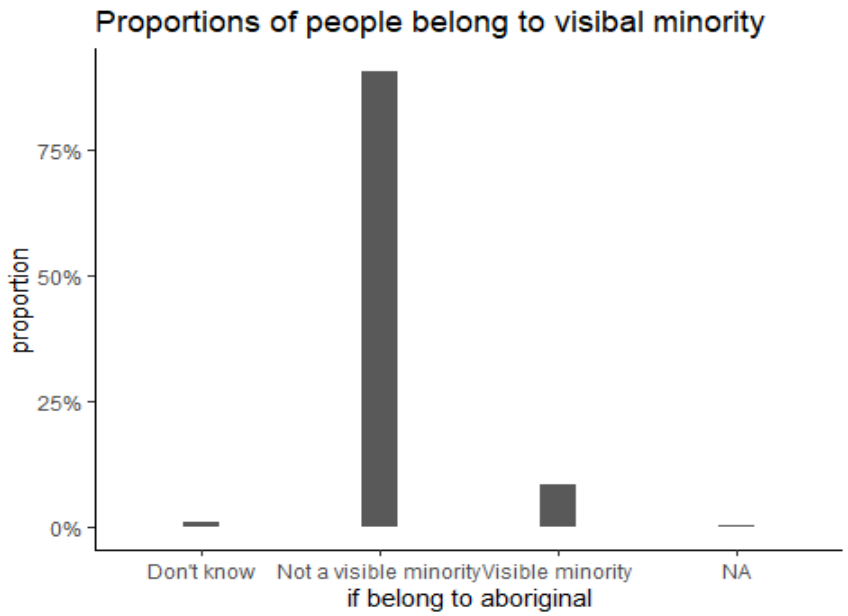

Figure 3. Proportions of people belong to visible minority

In this Figure 3, we find that only around 10% women in Quebec belong to visible minority.

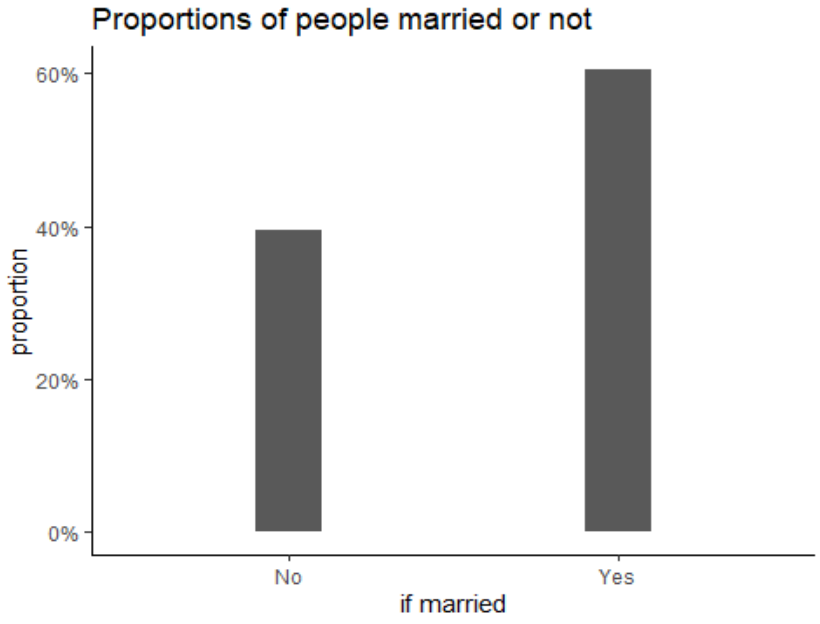

Figure 4. Proportions of people married or not

Figure 4 tells us that 60% women in Quebec married.

logistic model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.108	0.036	-2.981	0.003
age	0.007	0.001	8.861	0.000
aboriginal	0.001	0.068	0.019	0.985
vis_minority	-0.041	0.083	-0.493	0.622
ever_married	0.503	0.028	18.258	0.000

Table 1. logistic model summary

Table 1 is a summary of the logistic model. The p-value for age and ever married is smaller than 0.05, which means there is significant evidence to show there is the meaningful statistical relationship between Quebec woman ever give birth and their age and if they have married. However, we cannot find the relationship between people ever give birth, and their race since the p-value for visible minority and aboriginal is more extensive than 0.05. Table 2 below shows the odds and confidence interval of the model. Based on over odds below, we find out that they get married, and older age can be two factors to make Quebec women have more desire to give birth to the child.

confidence interval

	OR	2.5 %	97.5 %
(Intercept)	0.898	0.837	0.964
age	1.007	1.005	1.008
aboriginal	1.001	0.877	1.143
vis_minority	0.960	0.816	1.130
ever_married	1.654	1.567	1.746

Table 2. Confidence Interval summary

Discussion

As we know from table 3, only age and marital status (ever married or not) have significant evidence to influence the Quebec women ever gives birth. The odds of ever give birth increases 0.07 for each year older, and the odds of ever give birth increases 0.654 for people married before. By both two tables, we can find there are weak evidence to show weather aboriginal or visible minority have influence on Quebec woman ever give birth.

Weakness

In this report, we are considering the relationship between race and ever had children among women. However, through the data set, we can only know if our target population is aboriginal or not and visible minorities. Suppose we want to explore more about the relationship. In that case, we should know more detailed information about people's race, like where they are initially from Africa, Europe, Asia, etc. Moreover, the age range we are researching is too extensive, and we should narrow our age range. And therefore, there might be significant evidence to show the influence of race on ever had children among women.

Next Step

Except for the factors we already consider in our model, there might be other factors that influence women ever birth like there education level, their family income, and if they are immigrant. We need to build another logistic model to explore the relationship between them. And as we said in the weakness, we feel like there is not enough information. To show the detailed race of people, maybe we can make or other data set to help us get a more in-depth relationship.

Reference

Government of Canada, S. (2017, February 27). The General Social Survey: An Overview. Retrieved October 20, 2020, from https://www150.statcan.gc.ca/n1/pub/89f0115x/89f0115x2013001-eng.htm

Gss_cleaning. R by Rohan Alexander and Sam Caetano