**THE MOTHERHOOD WAGE PENALTY**

**Problem Statement:**

Does having children can make women earn less? It is a well-known empirical phenomena that women who do not have children are paid more on average than women who have children as Mothers in the workforce experience additional disadvantage compared to women who are not mothers, including a per-child wage penalty. We will be using the data to investigate the wage penalty size that a mother faces in the USA, we will investigate whether the differences observed represent the casual effect of motherhood and will try to consider the degree to which these differences have changed over time.

Null Hypothesis: there is no significance difference in the mean wages of women with or without children.

Alternative Hypothesis: there is a significance difference in the mean wages of women with or without children.

**Data Source:**

The data is taken to estimate the effects of having children on the young women wages. It is taken from the national longitudinal survey of youth. The data has been downloaded from the PUBL0055 webpage.

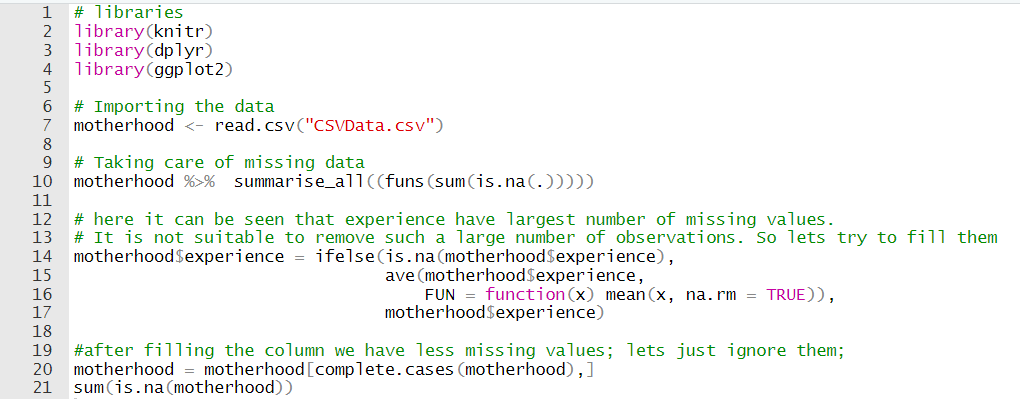
Each observation in the data (N= 2479) represents women (all age between 19 and 30) and the names and description of the variables is included in the data table

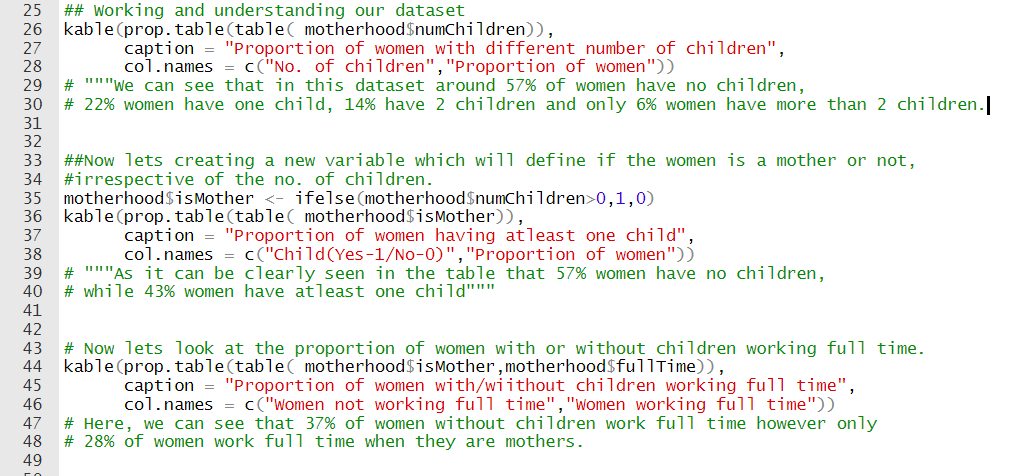
|  |  |
| --- | --- |
| Name of the variable | Description of the variable |
| PUBID | ID of women |
| Wage | Hourly wage, In dollars |
| numChilren | Number of children the women has |
| Age | Age, in years |
| Educ | Level of education   1. Less than HS 2. High school 3. Some college 4. College |
| Experience | Relevant experience in months |
| Tenure | Current job tenure, in years |
| Marstat | Marital status   1. Cohabitting 2. Married 3. No romantic union |
| Fulltime | Employment status   1. Full time = TRUE 2. Part- time = FALSE |
| Y2009 | Year of observation   1. 2009 = 1 2. 2004 = 0 |

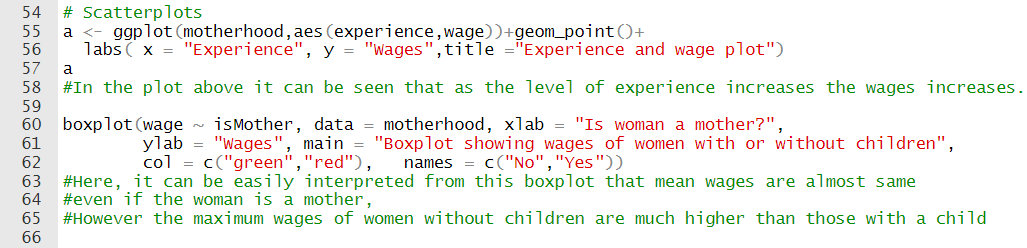
**Data Preparation for visualisation:**

Our first step after importing the data was to look for missing values in our data. We came across a fact that experience column had largest number of missing values. It was not suitable to remove such a large number of observations. So we tried to fill them using mean experience. However after imputing tye missing values for experience we still had some missing values which we thought would be best to remove instead of imputing as the are almost negligible proportion. After cleaning our data we went for preliminary analysis; that includes looking into our data and checking what the variables are and what would be the best modelling technique;

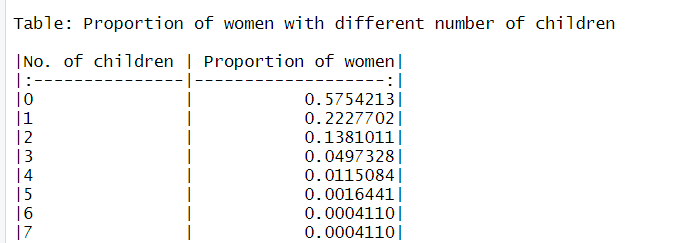
Basically the steps could be counted as EDA or exploratory data analysis.



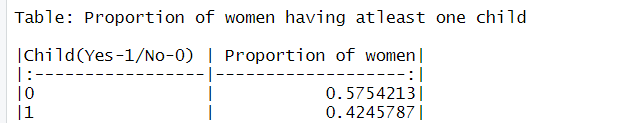


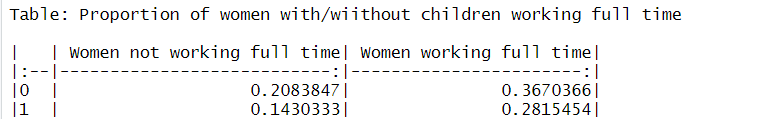


**Graphs and tables:**

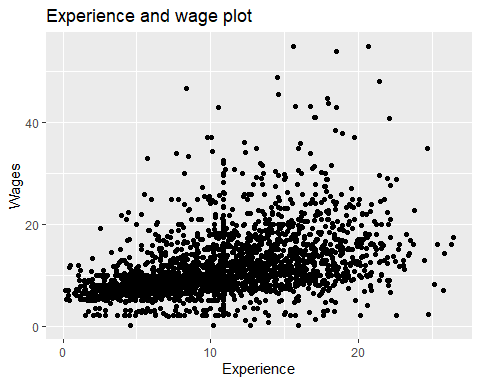


We can see that in this dataset around 57% of the women have exactly no child, 22% women in our data have only one child while 14% have 2 children and only 6% women have more than 2 children.

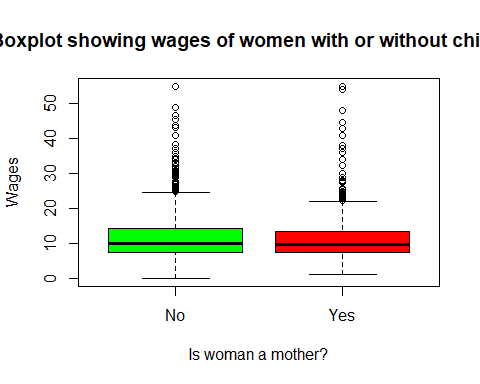
As it can be clearly seen in the table that 57% women have exactly no children, while 43% of the women in our data have at least one child.



Here, we can see that 37% of the women who do not have any children work full time whereas only 28% are seen working full time when they are mothers.



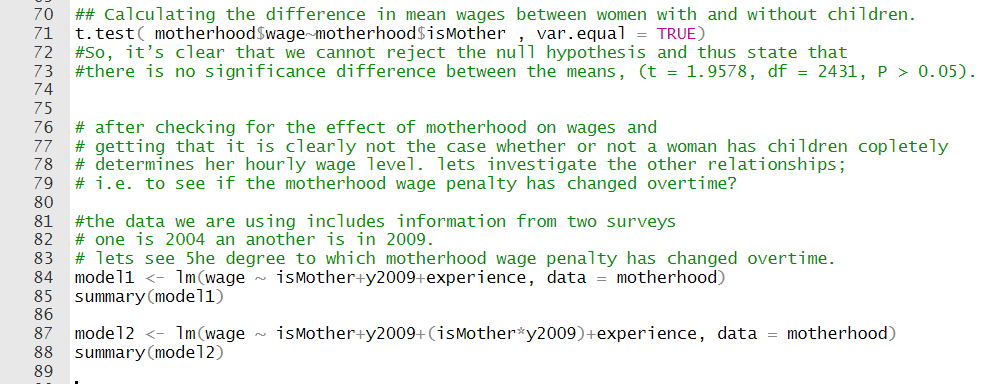
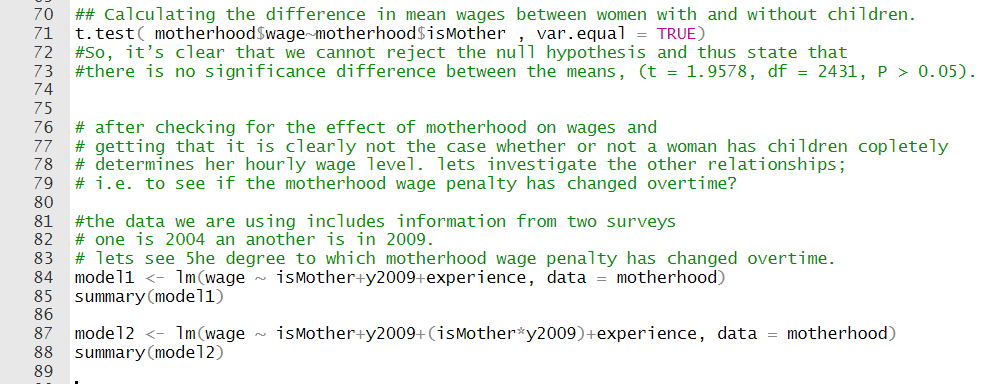
In the plot above it can be seen that as the level of experience increases the wages increases.

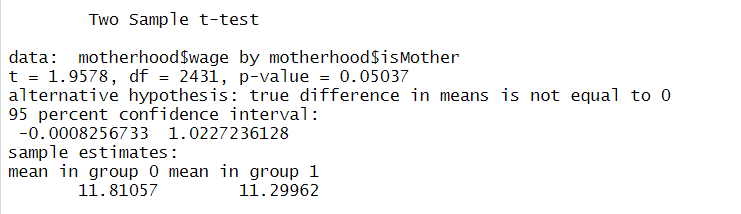


Here, it can be easily interpreted from this boxplot that mean wages are almost same even if the woman is a mother, However the maximum wages of women without children are much higher than those with a child.

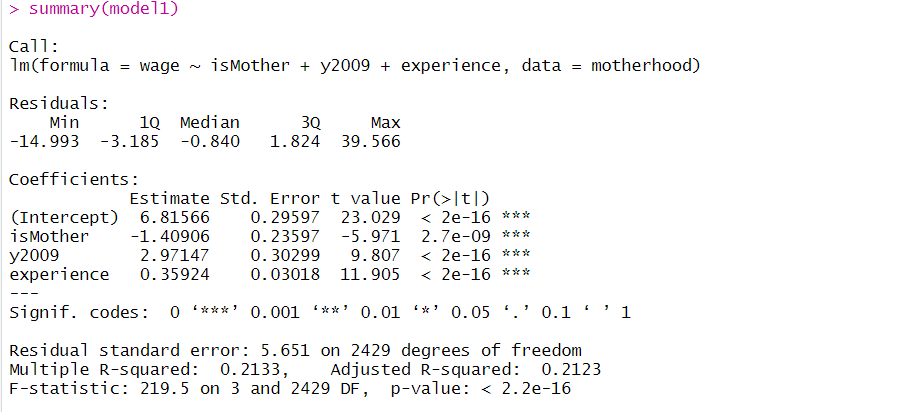
**Analysis:**

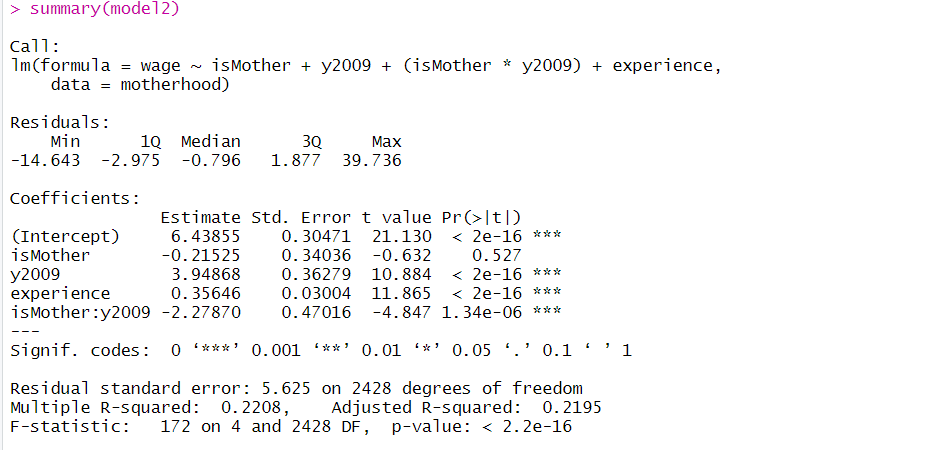
After our preliminary analysis we came to the t-test to see if our null hypothesis is to be accepted or rejected. And after checking for the effect of motherhood on wages and getting that if it is clearly a case or not the case whether a woman who have children completely determines her wage level after that we will investigate the other relationships i.e. to see if the motherhood wage penalty has changed overtime?





If the p-value is inferior or equal to the significance level 0.05, we can reject the null hypothesis and accept the alternative hypothesis. In other words, we can conclude that the mean values of group A and B are significantly different. So, here we came to the conclusion that it is clearly a case where a women with children earn less than a women who does not have children. Thus we reject our null hypothesis.



 Models have negative effect of motherhood on wages, i.e. if the woman becomes a mother than her wages decreases, However model 2 have more negative effect than model 1 but the variable is not significant for model2 but is significant for model1.

Coefficients Shows the regression beta coefficients and their statistical significance. Predictor variables, that are significantly associated to the outcome variable, are marked by stars.

Based on model 2 the model is:

wage = 6.43855 + (-0.21525)\*(isMother) + (3.94868)\*(y2009)+ (0.35646)\*(experience) + (-2.27870)\*(isMother:y2009),

Here are a few real time questions which we may want to look at:

1. Interviewed in 2004, A woman without children, with 10 months of relevant job experience.

wage1 = 6.43855 **+** (**-**0.21525)**\***0 **+** (3.94868)**\***0 **+** 0.35646**\***10 **+** (**-**2.27870)**\***(0)  
wage1

## [1] 10.00315

Wage of woman interviewed in 2004 without children, with 10 months of relevant job experience is 10.00315

1. Interviewed in 2004, A woman with children, with 10 months of relevant job experience

wage2 = 6.43855 **+** (**-**0.21525)**\***1 **+** (3.94868)**\***0 **+** 0.35646**\***10 **+** (**-**2.27870)**\***(1**\***0)  
wage2

## [1] 9.7879

Wage of a woman interviewed in 2004 with children and 10 months of relevant job experience is 9.7879

1. Interviewed in 2009, A woman without children, with 10 months of relevant job experience

wage3 = 6.43855 **+** (**-**0.21525)**\***0 **+** (3.94868)**\***1 **+** 0.35646**\***10 **+** (**-**2.27870)**\***(1**\***0)  
wage3

## [1] 13.95183

Wage of a woman interviewed in 2009 without children and with a 10 months of relevant job experience is 13.95183.

1. Interviewed in 2009, A woman with children, with 10 months of relevant job experience

wage4 = 6.43855 **+** (**-**0.21525)**\***1 **+** (3.94868)**\***1 **+** 0.35646**\***10 **+** (**-**2.27870)**\***(1**\***1)  
wage4

## [1] 11.45788

Wages of a woman interviewed in 2009 with children and a relevant 10 months job experience is 11.45788.

**Consideration and evaluation of results:**

Mothers have to suffer a penalty with respect to women who are not mothers. Women without children have more wages than women with children. Also it has been noticed that the size of motherhood penalty have changed overtime, wages have increased overtime irrespective of motherhood.

**Conclusion**

We came with the below conclusions in our analysis.

* We reject our null hypothesis and conclude that there is a significant difference in mean wage of mothers and non- mothers.
* The wages have increased overtime for all the women but the increase in wages of non-mothers is more than increase in wages of mothers.