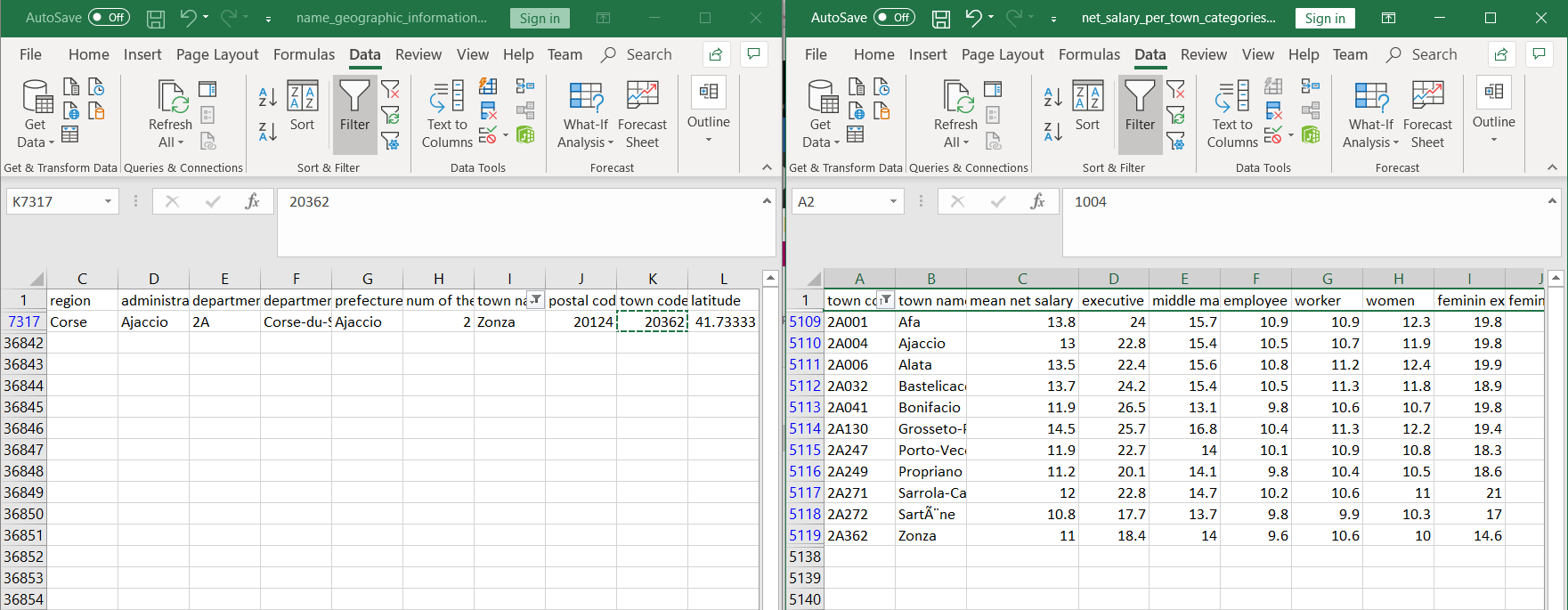
# Data Preparation

I renamed all the excel sheet’s headers into meaningful names instead, so that it will be easier to understand. (eg. In *population.csv*, for cohabitation mode, it was named MOCO instead) I replaced some of the variables into more meaningful values as well. (eg. In *population.csv*, for gender, it was 1 or 2 instead of men and women)

Some of the town code in *net\_salary\_per\_town\_categories.csv* werent in correct format. (eg 20362 is the correct town code – *on the right,* while it is written as 2A362 in the *net\_salary\_per\_town\_categories.csv*)



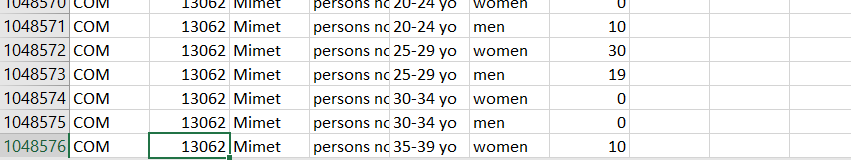
## name\_geographic\_information.csv

This contains the town’s information like town code, town name, latitude, longitude and so on. I will be using this as the main table to refer to town names using town codes. As other excel sheet’s town name might contain some utf-8 errors and this has the most complete information of each town among all the excel sheets.

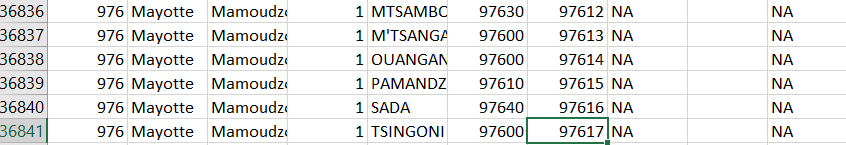
## population.csv

This contains each town's demographics like the town's demographics on age category, gender and cohabitation mode.

However, the table is incomplete as not all town's information is captured, only until town code = **13062**.



While the maximum town code is actually = **97617**.



I will be using this excel sheet to analyze the overall demographics of France, on gender, gender ratio, age category and cohabitation mode.

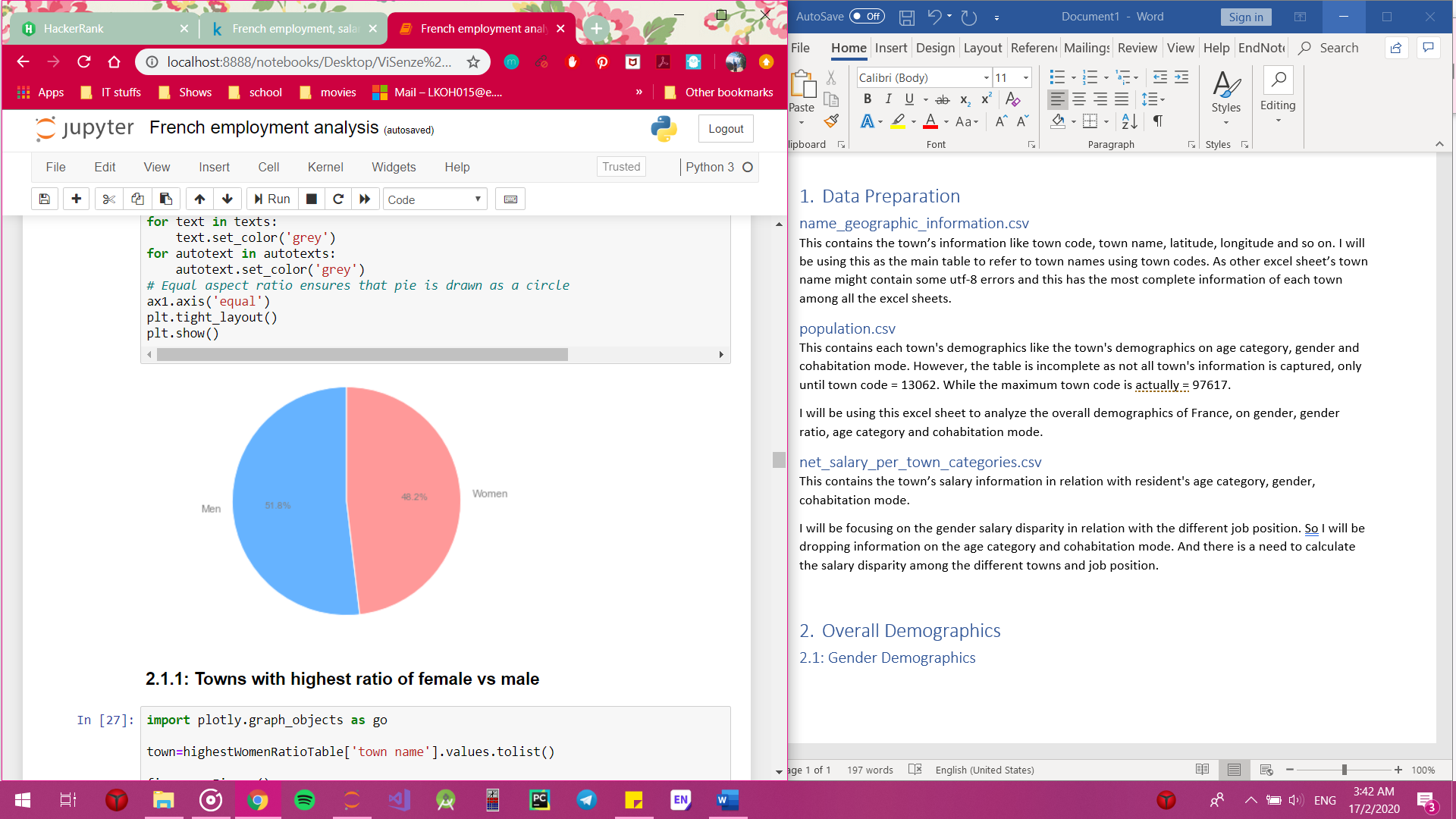
## net\_salary\_per\_town\_categories.csv

This contains the town’s salary information in relation with resident's age category, gender, cohabitation mode.

I will be focusing on the gender salary disparity in relation with the different job position. So I will be dropping information on the age category and cohabitation mode. And there is a need to calculate the salary disparity among the different towns and job position.

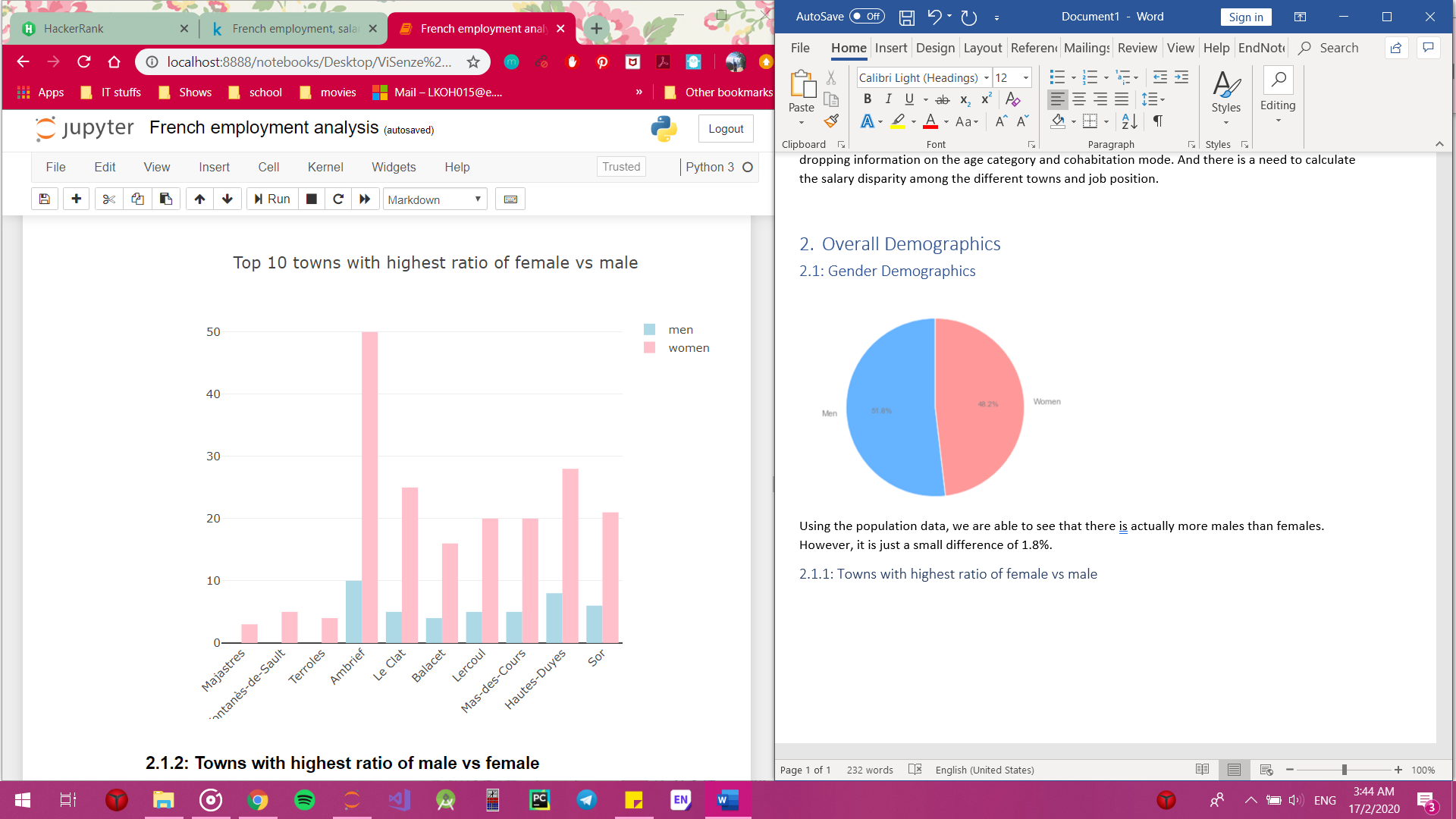
# Overall Demographics

## 2.1: Gender Demographics



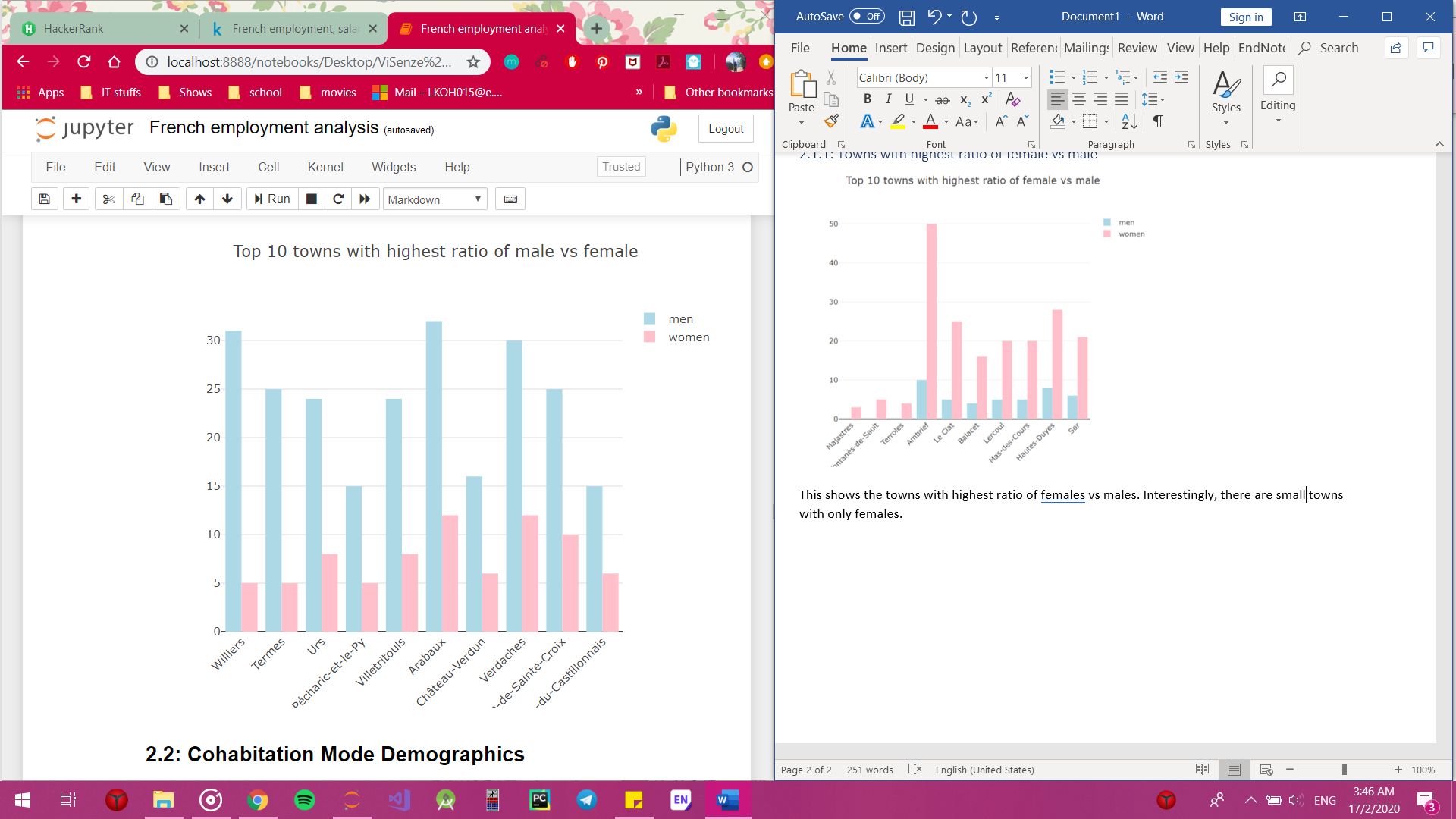
Using the population data, we are able to see that there is actually more males than females. However, it is just a small difference of 1.8%.

### 2.1.1: Towns with highest ratio of female vs male



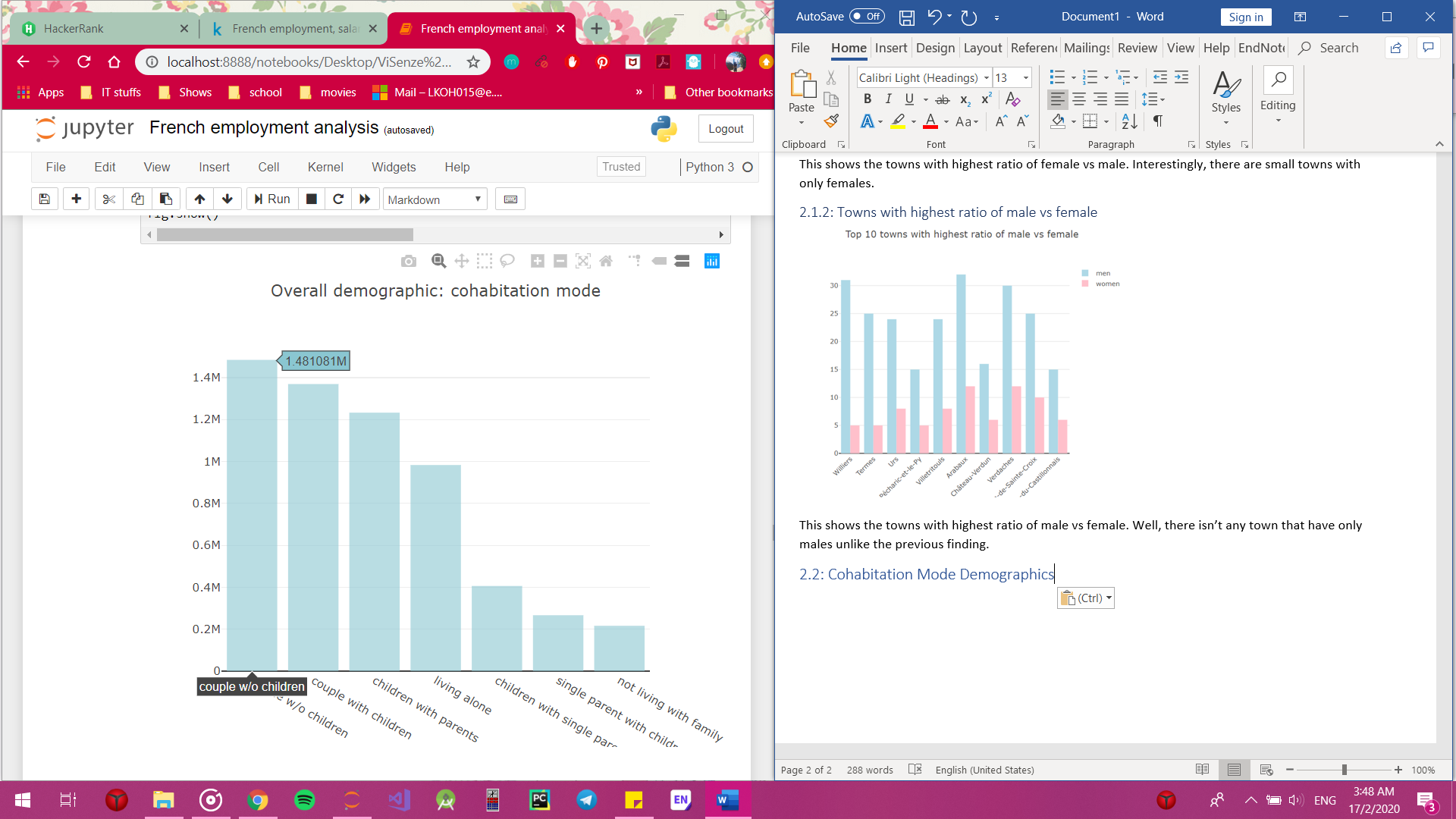
This shows the towns with highest ratio of female vs male. Interestingly, there are small towns with only females.

### 2.1.2: Towns with highest ratio of male vs female



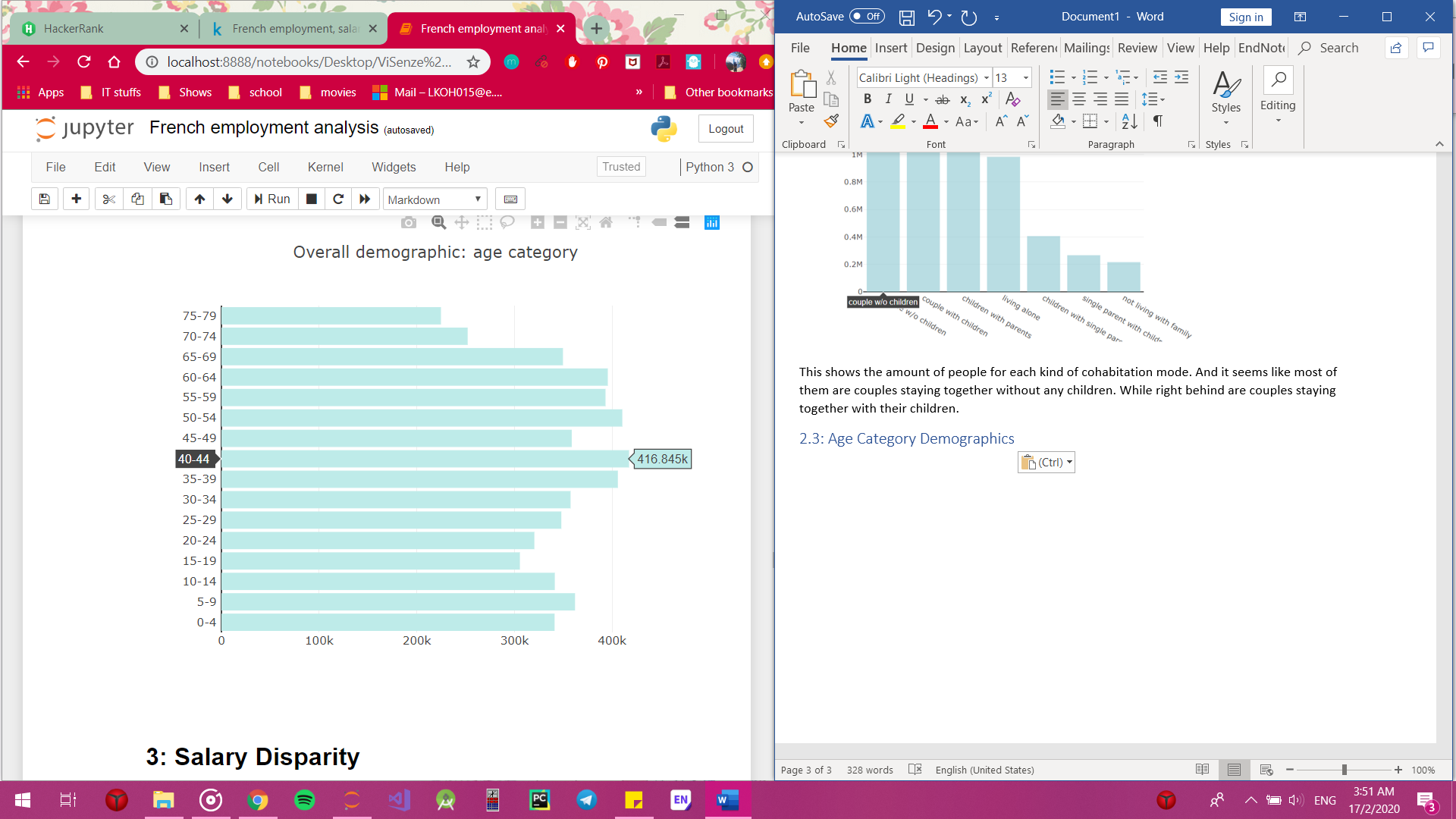
This shows the towns with highest ratio of male vs female. Well, there isn’t any town that have only males unlike the previous finding.

## 2.2: Cohabitation Mode Demographics



This shows the amount of people for each kind of cohabitation mode. And it seems like most of them are couples staying together without any children. While right behind are couples staying together with their children.

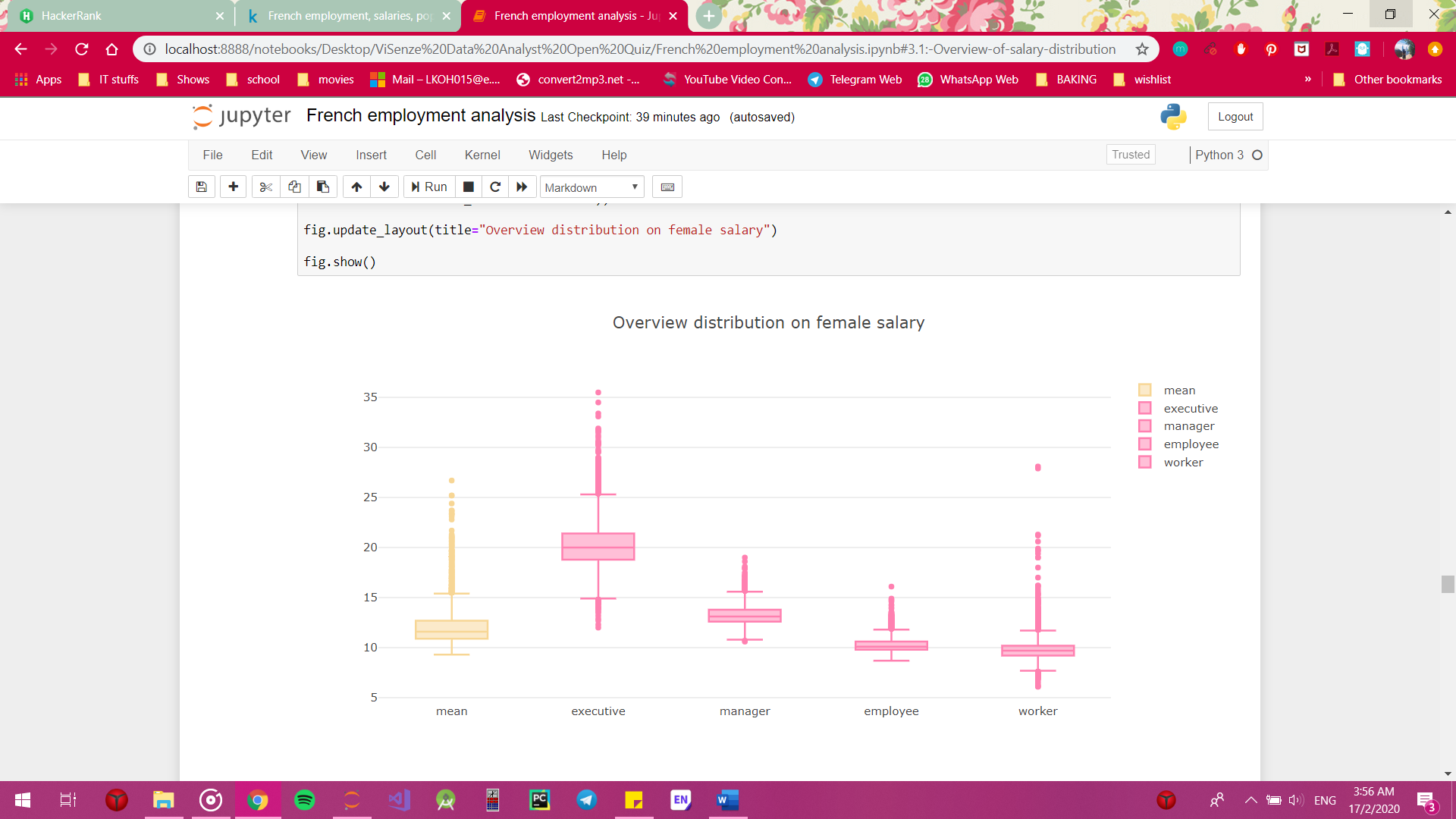
2.3: Age Category Demographics

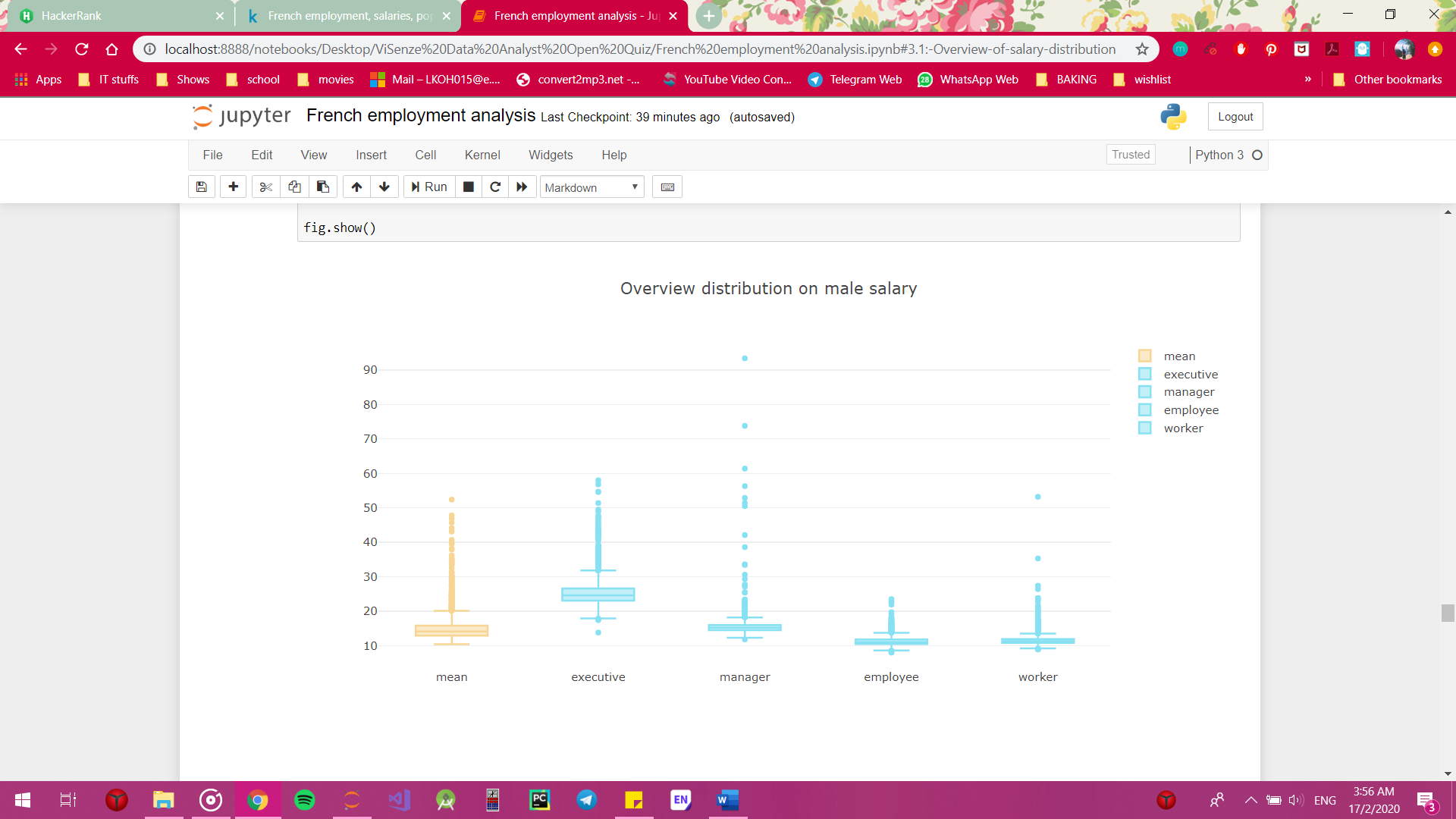


This shows the amount of people in each age category. From this, we are able to see that France do not have an aging population, where 18-64 years old have a small proportion compared to 65 years old onwards. We can clearly see that the number of working adults is much higher than the elderlies.

# Salary Disparity

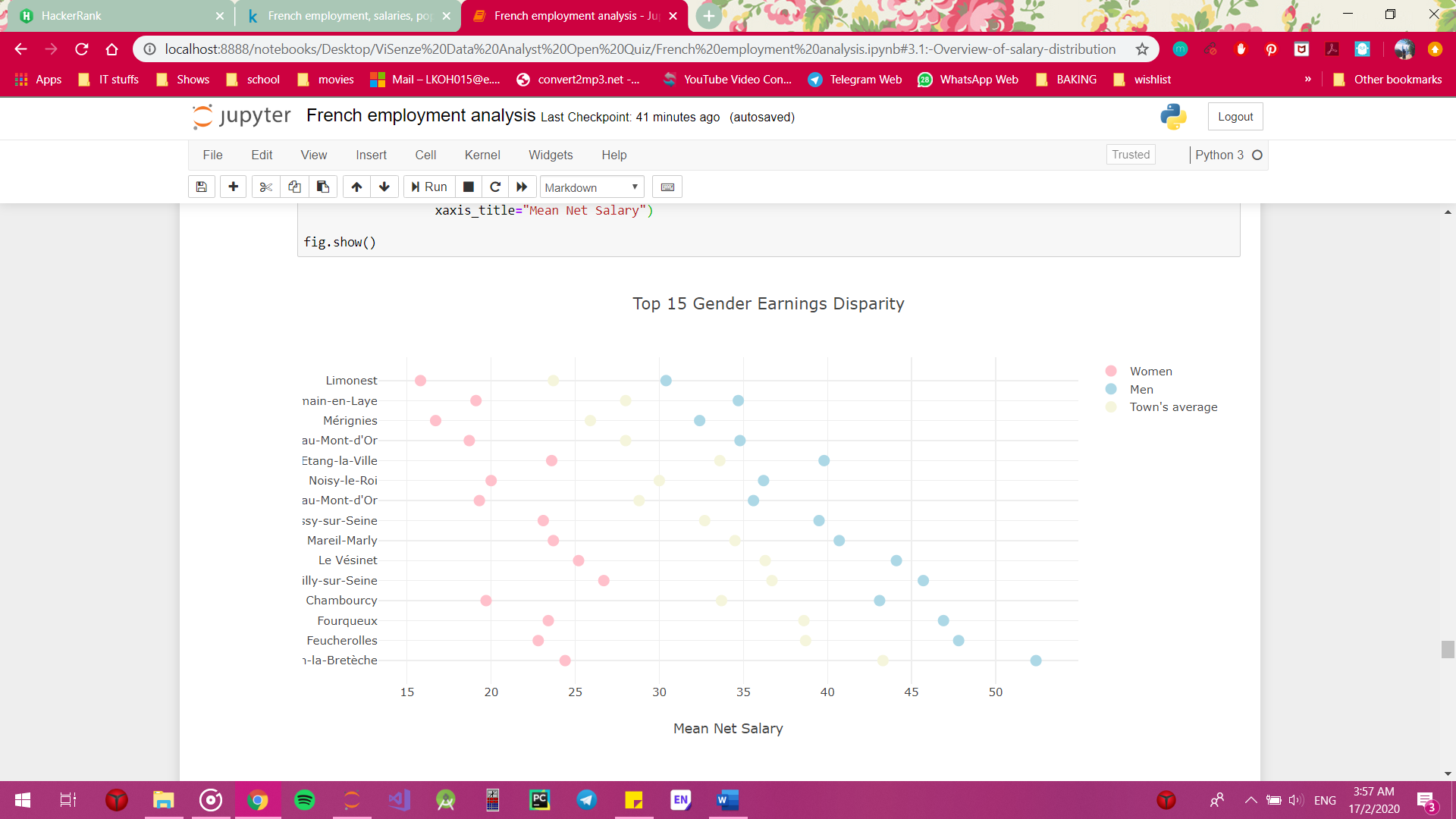
## 3.1: Overview of salary distribution





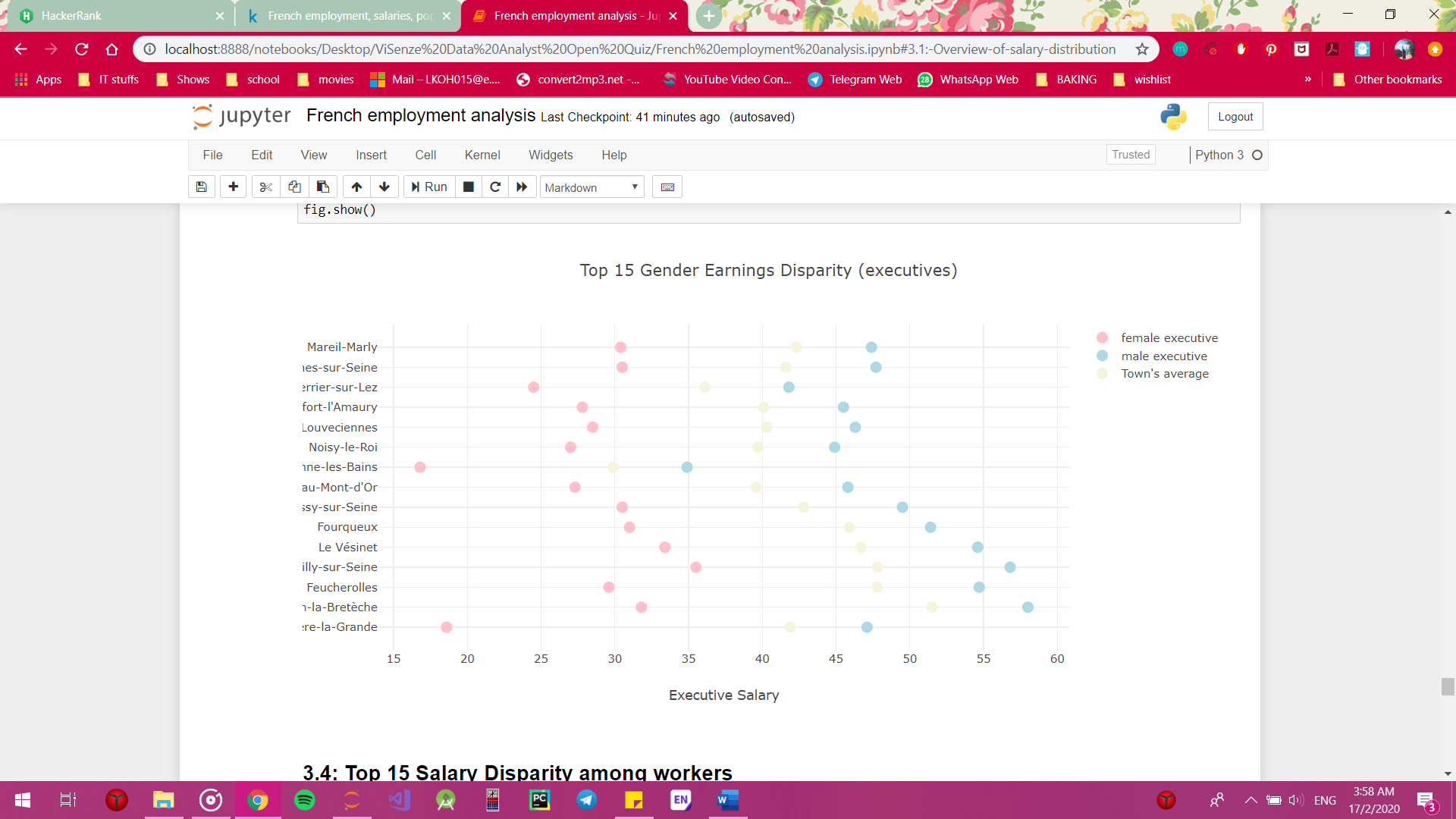
It can be seen from the box plot that the salary difference between female executive and female manager is huge. The range of female salary is quite small compared to the male salary. The maximum of female salary is 35.5 while the maximum of male salary is 93.4. All of the data shown have a lot of outliers, which is understandable as the data is huge and it includes thousands of different values.

## 3.2: Top 15 Salary Disparity (mean net salary)



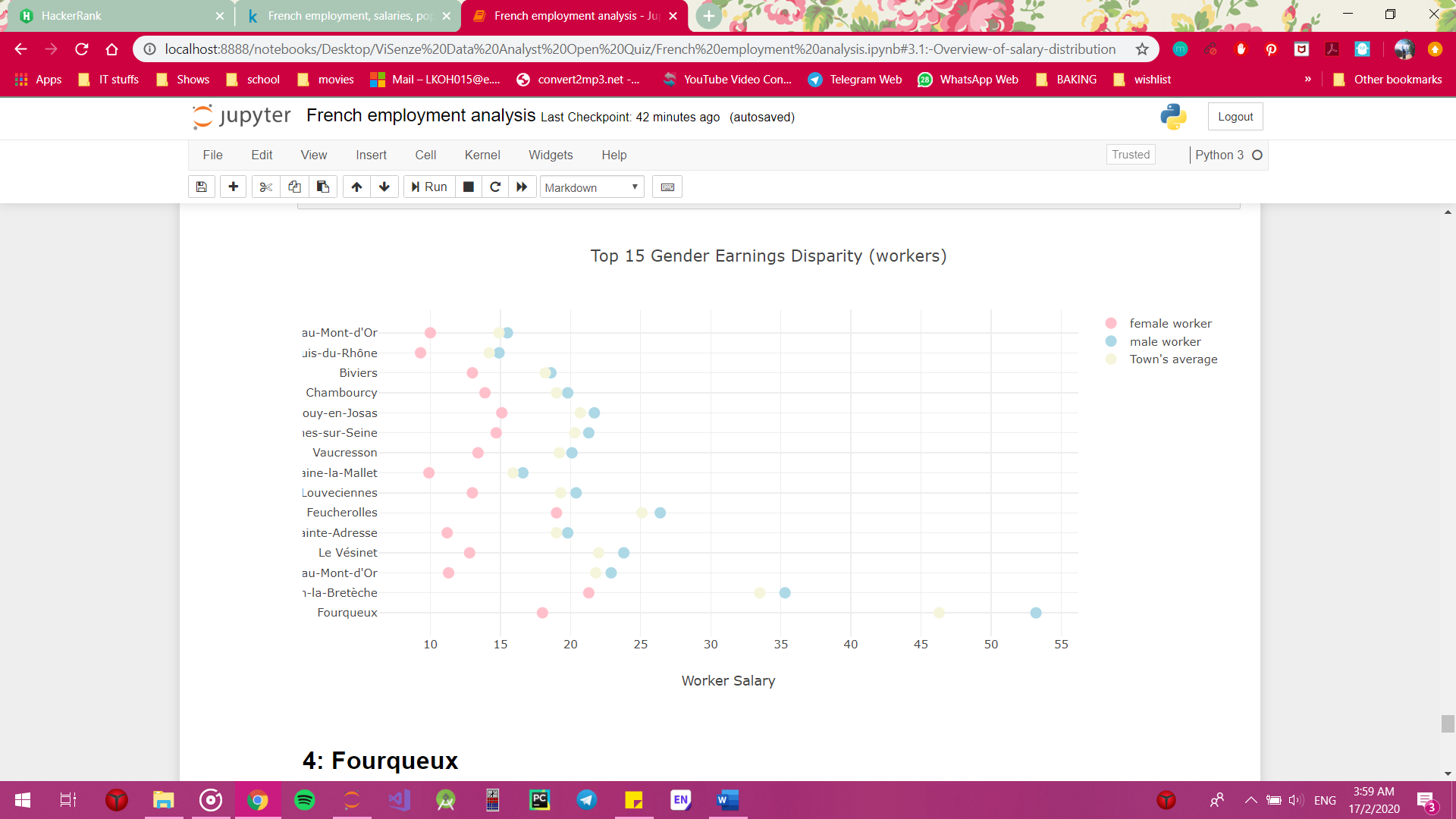
This graph shows the top 15 salary disparity between male and female executives. It can be seen that even though executive is the highest position, the salary disparity still stays the same.

## 3.2: Top 15 Salary Disparity among executives



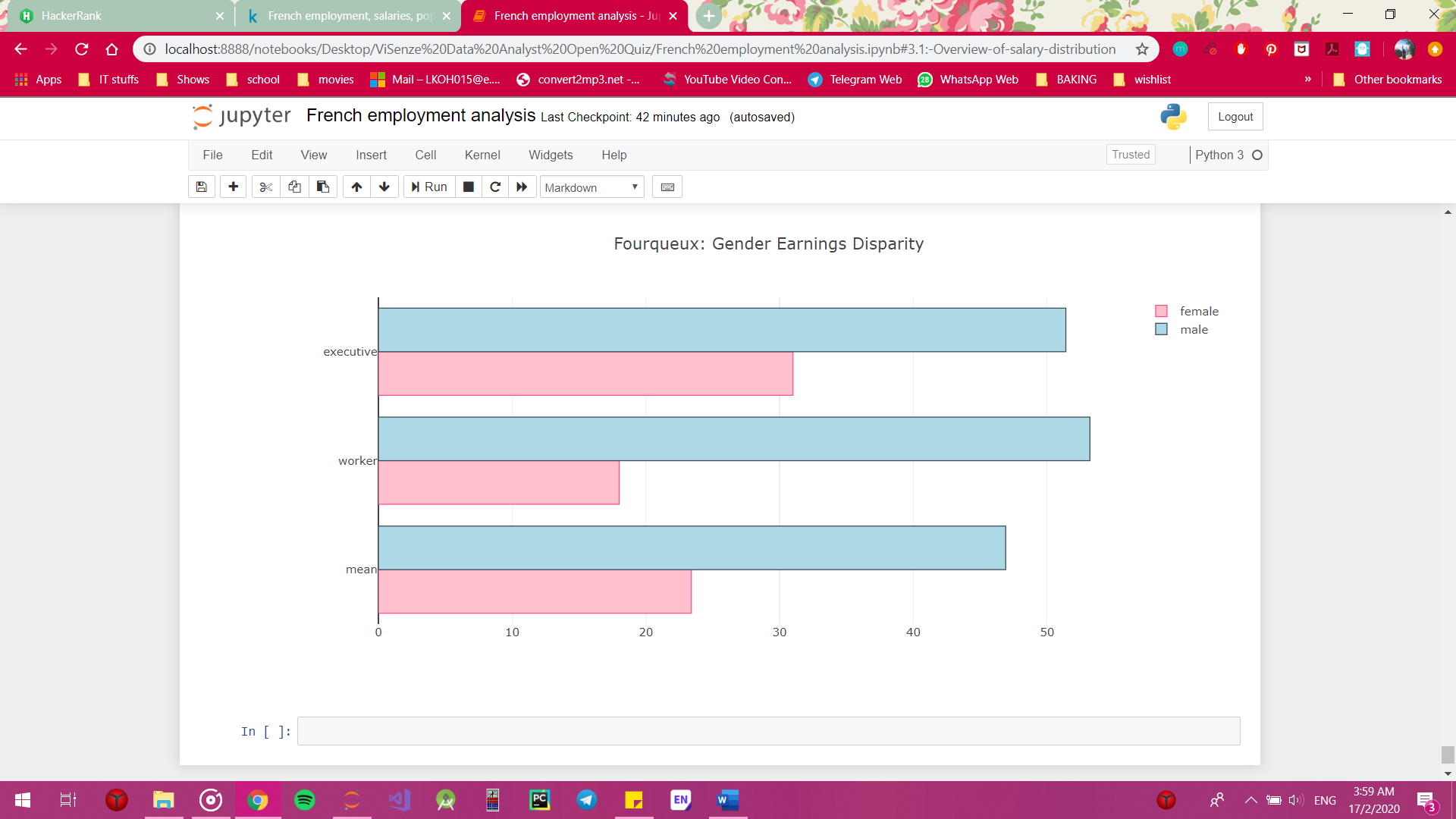
This graph shows the top 15 salary disparity between male and female executives. It can be seen that even though executive is the highest position, the salary disparity still stays the same.

## 3.4: Top 15 Salary Disparity among workers



This graph shows the top 15 salary disparity between male and female workers. It can be seen that **Fourqueux**'s pay disparity among the male and female workers is huge compared to the others. It can be seen that Fourqueux male worker's salary is even higher than Fourqueux male executive (from the previous graph-3.2 Top 15 Salary Disparity among executives) which is a little odd.

# 4: Fourqueux



The result shown in the previous graphs are a little odd. Hence, we place the data together so we can have a clearer picture of the data. It would have been interesting to find out more about Fourqueux, like the age, gender and cohabitation demographics. However, due to the missing data in population.csv, Fourqueux demographic is not available to analyse.