Deliverable 2

Group 1 of BCNC

Group Member: Chenyu Cao, Jiawen Chu, Johanne Antoine, Fangxu Zhou, Zhechao Chen

I. Description of project

There is a growing large number of Asians that live in Boston, but yet the research towards them is limited and insufficient. The motivation of this project is to understand more about the living conditions of Asians in America. This project will mainly focus on the employment of the Asian community in Massachusetts, especially the differences within the Asian community. We think this could largely reflect how (different) Asian communities fit and develop in the local society.

The research will be based on IPUMS datasets including ACS and CPS(voluntary survey). Through this project, we want to understand what kind of job is most appealing to Asians (and how their employment is related to their education). Moreover, the shift of their work and salary in recent years. Last, the possible reasons that may cause the situations above.

II. Goals of our team

- Collect datasets in the scope of employment, retrieve data that is relative to Asians in MA, clear up and combine these data so that they can be more available to the public, for example, the nonprofits, organizers, policymakers, etc.
- By analyzing the data collected, try to find out, if there is any, relations between the employment and Asian people and the possible reasons for those relations.
- By analyzing the data collected, try to find the future trend of employment of Asian people.

III. Preliminary Analysis(ACS will be merged with CPS data in the future)

ACS:

Total num of rows in the dataset(Asians in MA): 22349

Number of people in laborforce: 15955

Number of people who reported their income: 18787

Number of people who reported their wage/salary: 12538

CPS:

Total number of rows in the dataset(Asians in MA): 9119

Number of people in laborforce(who report their wage): 5676

IV. Limitations and potential risks

- **Limitations:**
  - Dataset is small but this is the only dataset that we found to be feasible.
  - Dataset is skewed—most of the samples are Chinese, then Indian, as the graph below shows.

```
Classify by Asian Subgroups:
Chinese       8811
Indian        5210
Other         2993
Vietnamese    2330
Korean        1384
Filipino       747
Japanese       592
Thai           186
Indonesian      96
```

  - Both datasets are voluntary surveys, so they are biased on the samples selected. Also, not all columns should be filled, so the filled data by the samples can also be biased. For example, someone who has a higher salary(income) has a higher probability of filling the data of income.
- **Potential risks of achieving project goal:**
  - The data selected are based on multiple years. There may be people who respond to the survey in multiple years, making the data less representative. (If the number of repetitions is large, the trend we are testing may be the trend of specific households rather than the Asians as a whole group.)
  - The data is not large enough to predict. There are some sub-Asian groups with less than 100 people in the sample. Then, the prediction and the general trend of that group will be less accurate and representative.

V. Datasets plan to collect

- IPUMS ACS/CPS data with selected attributes

VI. Questions plan to answer

- ☐ Are there any relations between the employment and education of Asians?(See appendix-education distribution)

  With the very basic analysis of the data, we can say that there is a positive correlation between levels of education and the salary. We will further analyze more on whether this is unique for Asians and whether the gradient of this model is similar to the model of other races? If it is different, what can be the possible reasons for that.

- ☐ What are the popular career choices and the trend within Asians?

☑ ~~What is the salary distribution of Asians? (Mainly based on ACS data. (see appendix-salary distribution))~~

According to the bar graph, the average pre-tax salary of Asians in MA is $74,167 in 2019. If we divide them into subgroups, Indians, Koreans, and Chinese are the top-three subgroups that have the highest average salary, which are $102,069, $76,586, and $71,941 respectively. Thai, Indonesian, and Vietnamese are the three subgroups that have the lowest average salary. However, one thing to note is that our dataset is gathered from the U.S. Census Bureau's mandatory American Community Survey(ACS) and the U.S. Bureau of Labor Statistics voluntary Current Population Survey(CPS), so there are samples who refused to give the exact number of salaries. As a result, the salary distribution of Asian subgroups may not be highly accurate, and might be skewed.

☐ What is the main difference between Asians and Whites in these aspects?

☐ Which career choices in Asians have the highest earnings? Are these the same as the popular occupations of Asians? If there are any differences, why will these differences be presented?

☑ ~~Are there differences in career choices between different groups of Asians?~~

First of all, for different groups of Asians, the top chosen popular occupations are different. For Chinese, they are more likely to be working as a software developer, postsecondary teacher, and chefs. For Filipinos, they tend to work as nurses and salespersons more. Indians on the other hand, most of them work as software developers or business managers. Chefs or waiters are the top choices for Thais, and personal care service providers or manufacturers for Vietnamese. For the distribution graph of each subgroup, please see appendix-Occupation Distribution of Asian Subgroups.
Secondly, their education levels also differ. (Further analysis needed)

☑ ~~Are there correlations between sex and income among Asians? (See appendix-sex distribution)~~

Males tend to have higher income than females, and this is not mitigated as the year increases. According to research done by Amanda Barroso and Anna Brown, females earned 84% of what males earned in 2020. (2021) In our dataset, females earned around 75% of what males earned in 2019. Thus, there is a correlation between sex and income; however, the reasons for that are quite complicated. One key question is that does that indicate more serious gender inequality among Asians? No, there are many other reasons that cause this outcome. One main reason for that is Asian mothers are more likely to be stay-at-home mothers compared to white and black mothers. (D'Vera Cohn et al., 2014) That causes Asian females to have a lower mean salary in total.

- [ ] Are there any correlations between education and income in Asian subgroups? Are the gradients different among different Asian subgroups models?

  We haven't finished the analysis part of these questions. They seem a little repetitive and intuitively it seems that high education leads to high salary in all kinds of careers. However, we still want to make sure if this intuitive thought is really a ubiquitous fact, or there may be some exceptions for some careers for asian people.

- [ ] What are the most popular majors among asian people? Same as non-asian people? Is there any connection between those majors and popular occupations?

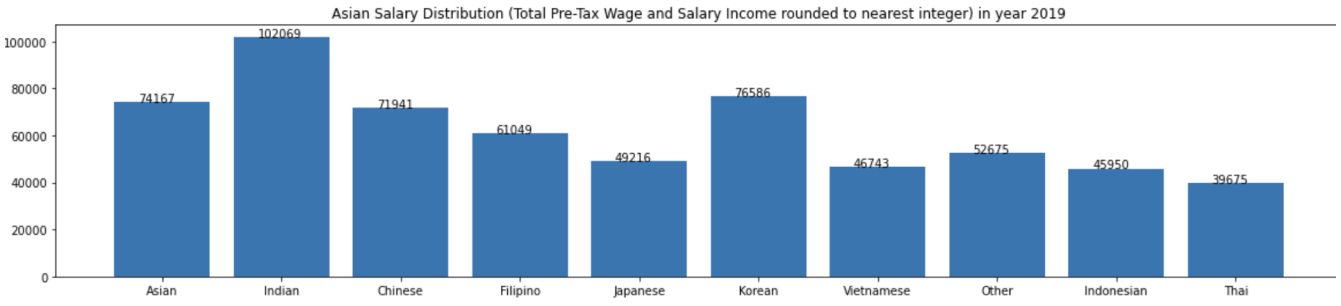
- [ ] Are there any correlations between age and income? Are the correlations the same as the ones of other races? What are the causes of these differences?

  From the figure in the appendix, we can see that the highest salary is within the [36,46] and [46, 56] range, which is quite consistent with our understanding. But, for those careers, like the executive staff, that are for people with more experience in that scope, people who have high salaries tend to be older. Also for those careers that may need more creativity, people who have high salaries tend to be younger. So we still need to dig into some certain careers to compare within them, that comes the following questions. Together with this one, we will further analyze the data to find answers for these questions later.

VII. Appendix

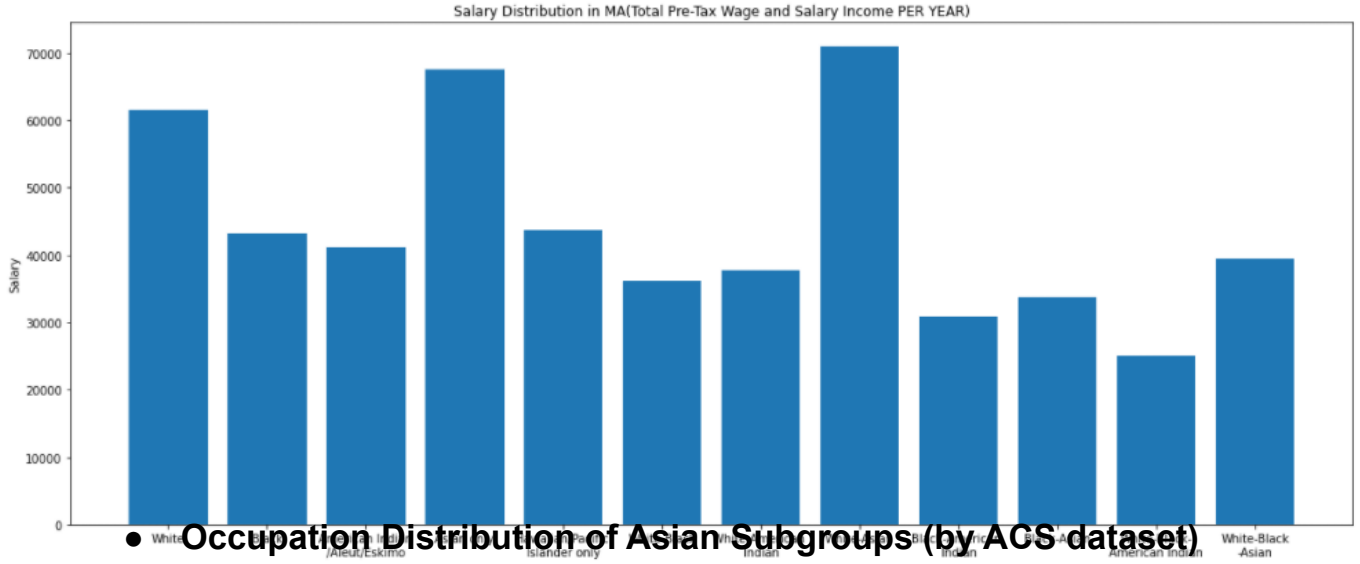
- **Salary Distribution**
  - **ACS**

```
ere are total of 2887 labor force in Asians in 2019
```



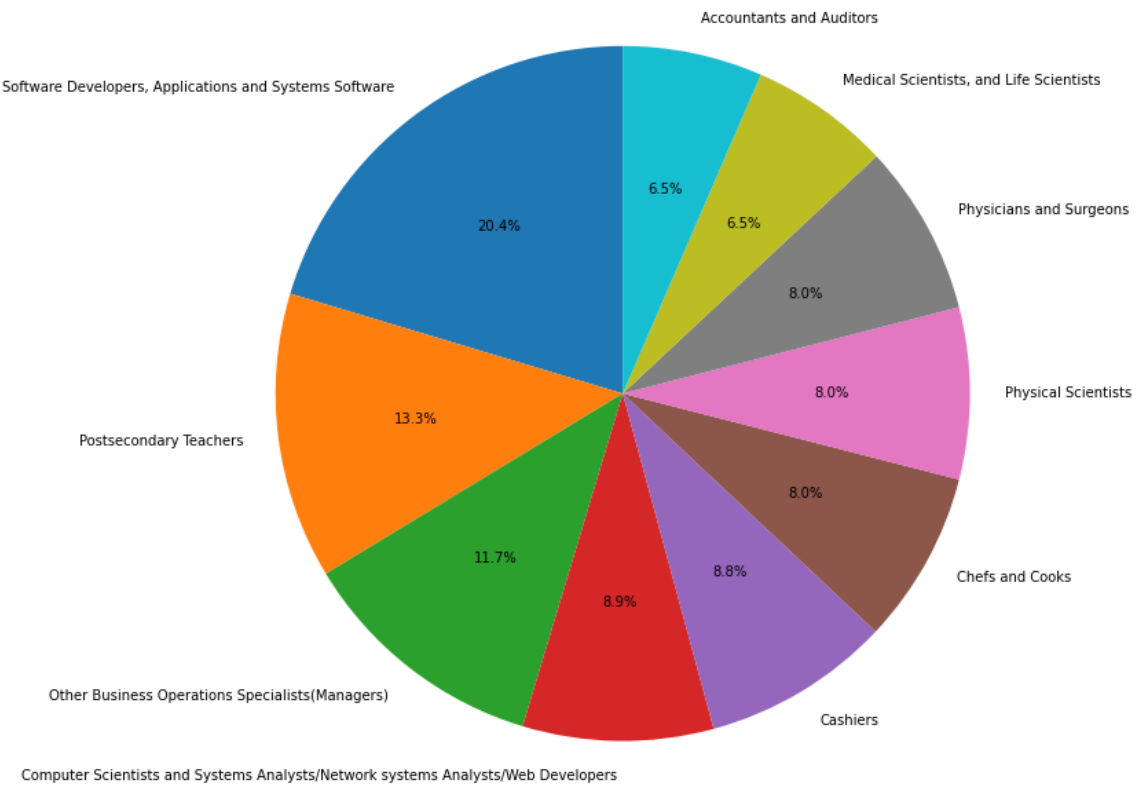
  - **CPS**

There are totally 5676 labor force in those Asian



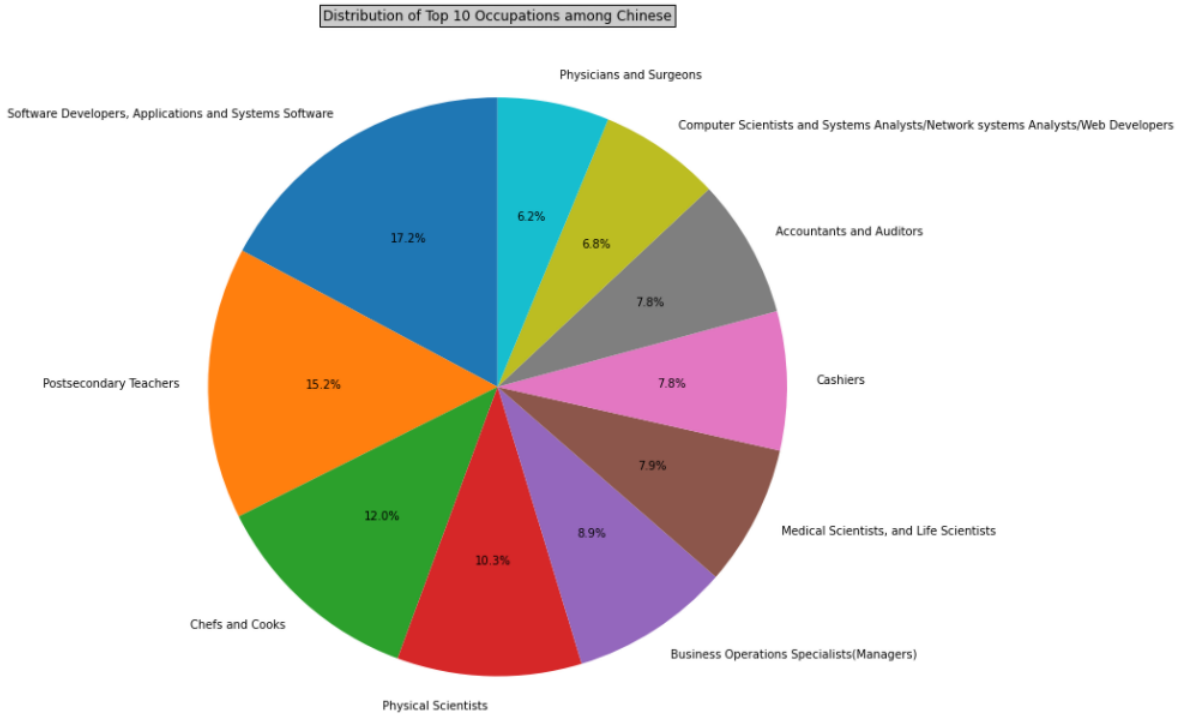
- **Occupation Distribution of Asian Subgroups (by ACS dataset)**

There are 4347 people work as top 10 occupations, about 0.272453776245691 of the total laborforce



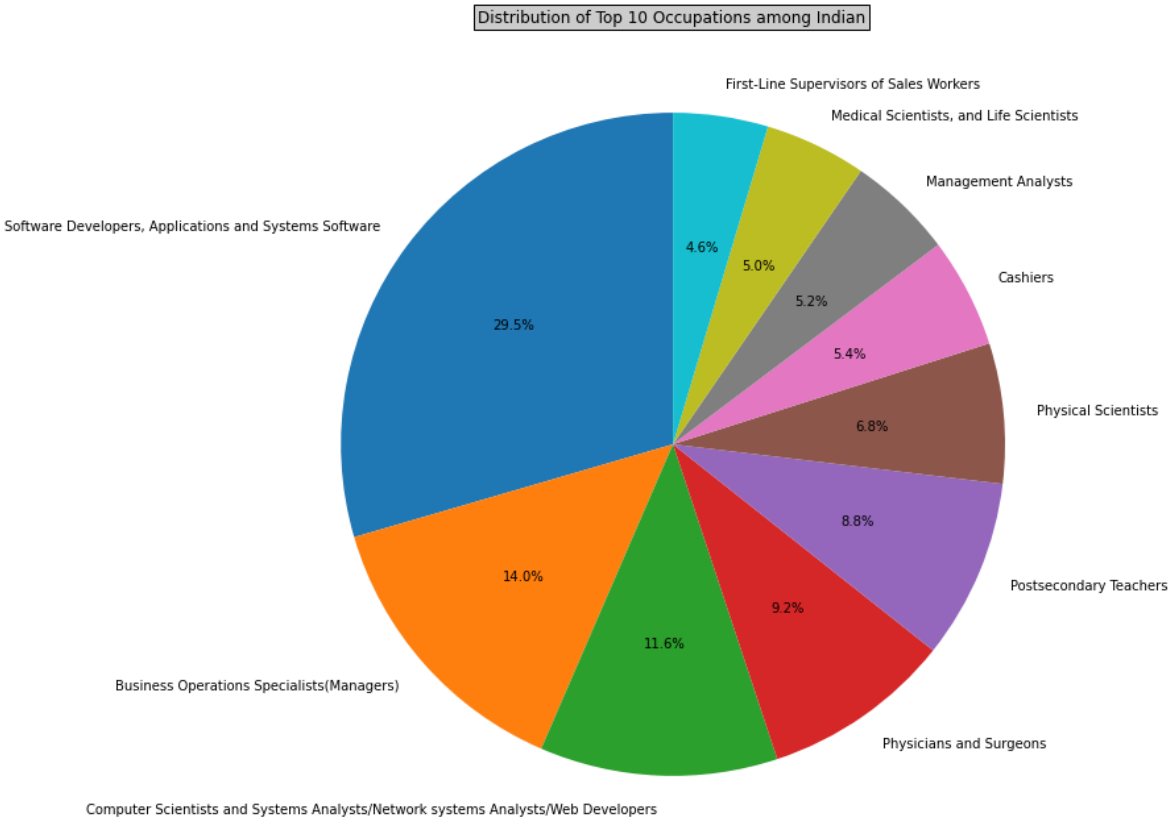
There are 1843 people work as top 10 occupations, about 0.32840342124019956 of the total Chinese Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Chinese')



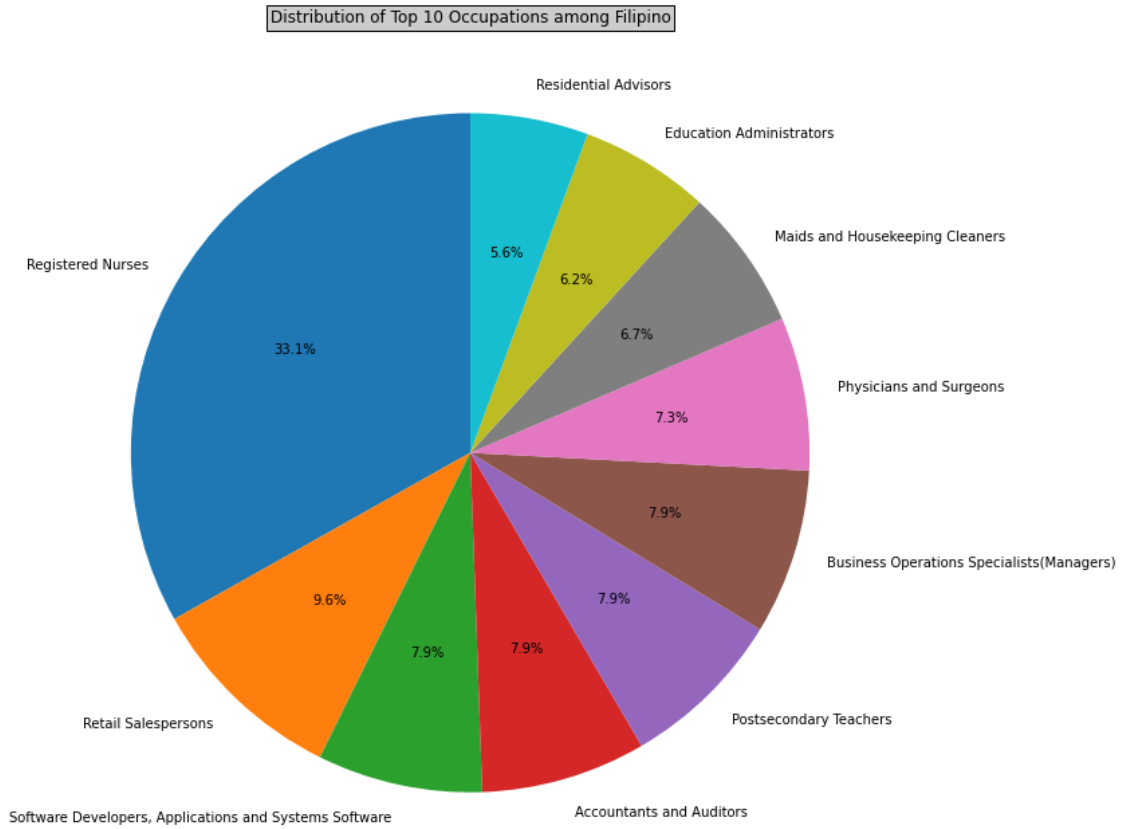
There are 1494 people work as top 10 occupations, about 0.44610331442221557 of the total Indian Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Indian')



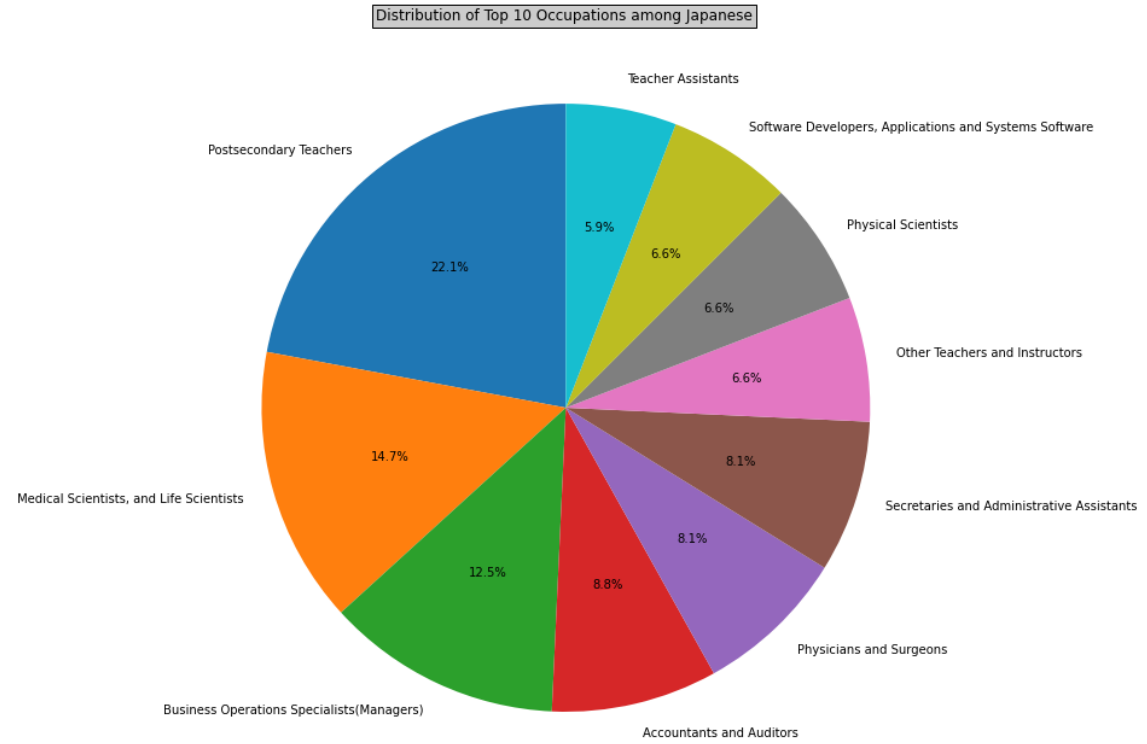
There are 178 people work as top 10 occupations, about 0.326605504587156 of the total Filipino Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Filipino')



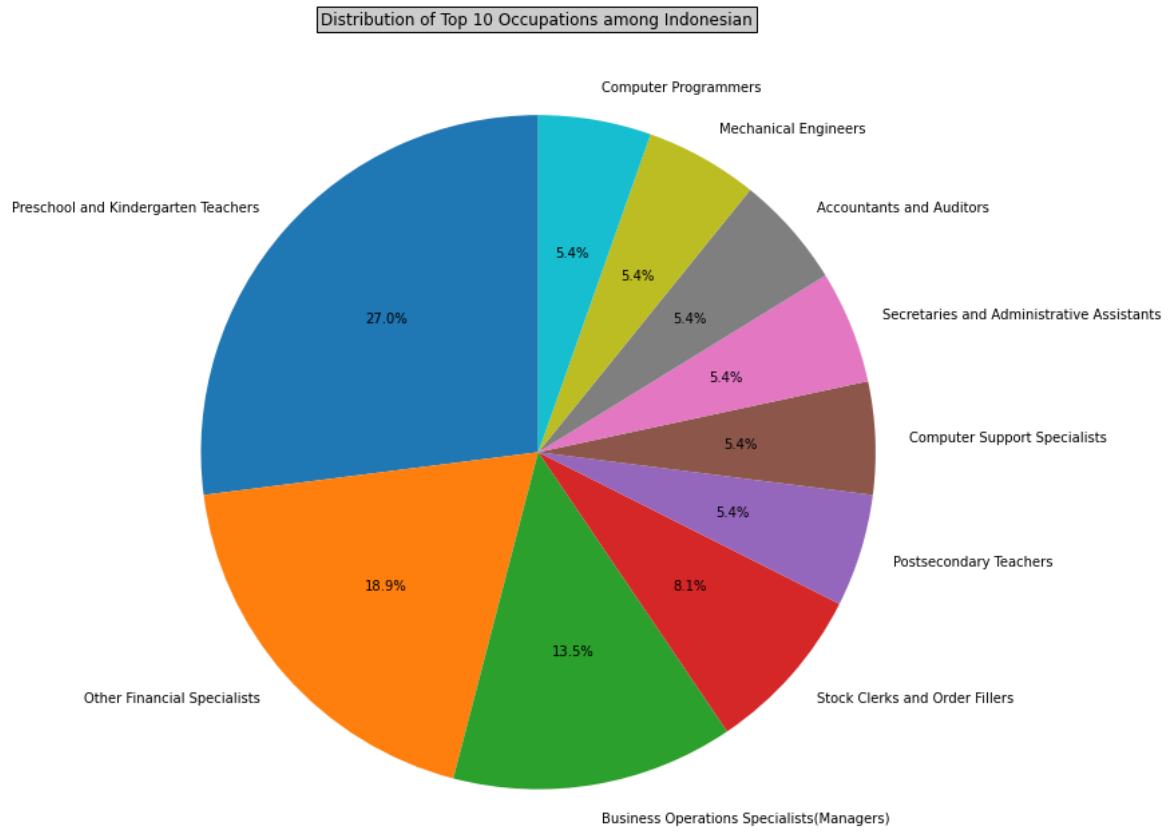
There are 136 people work as top 10 occupations, about 0.3726027397260274 of the total Japanese Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Japanese')



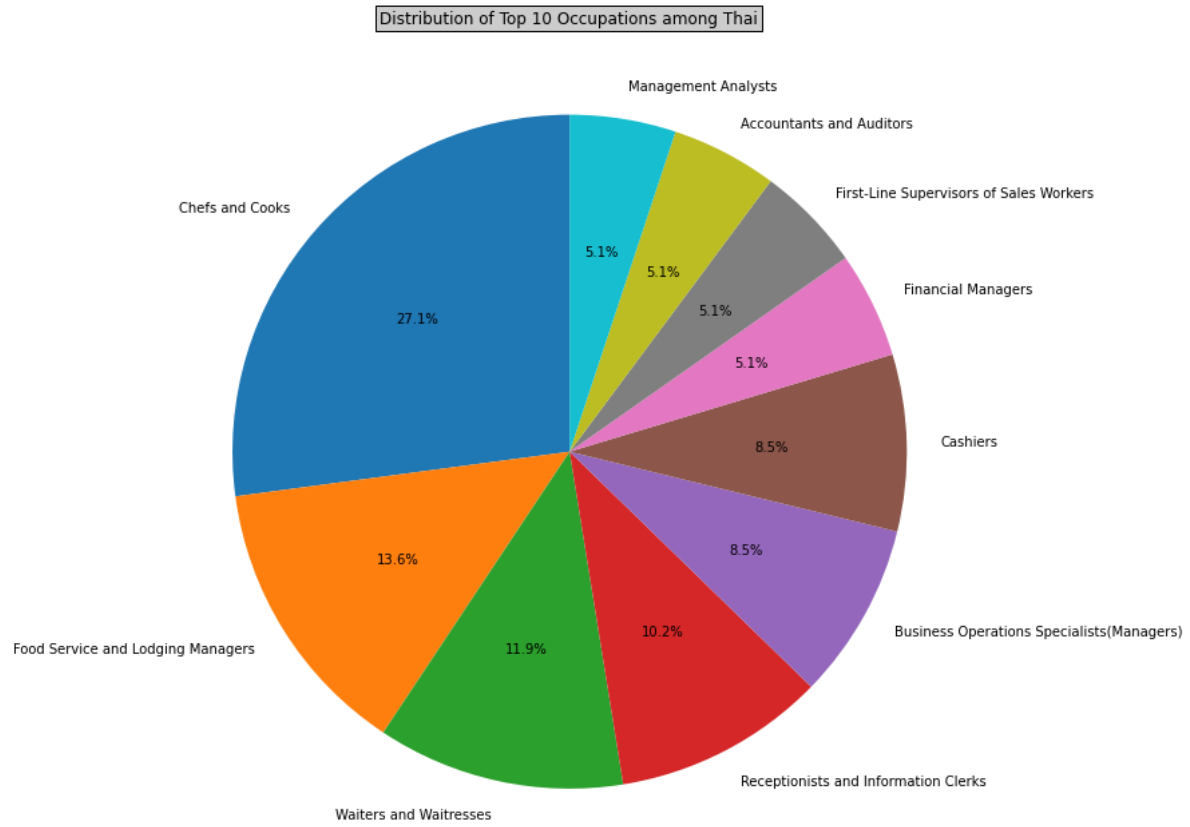
There are 37 people work as top 10 occupations, about 0.5873015873015873 of the total Indonesian Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Indonesian')



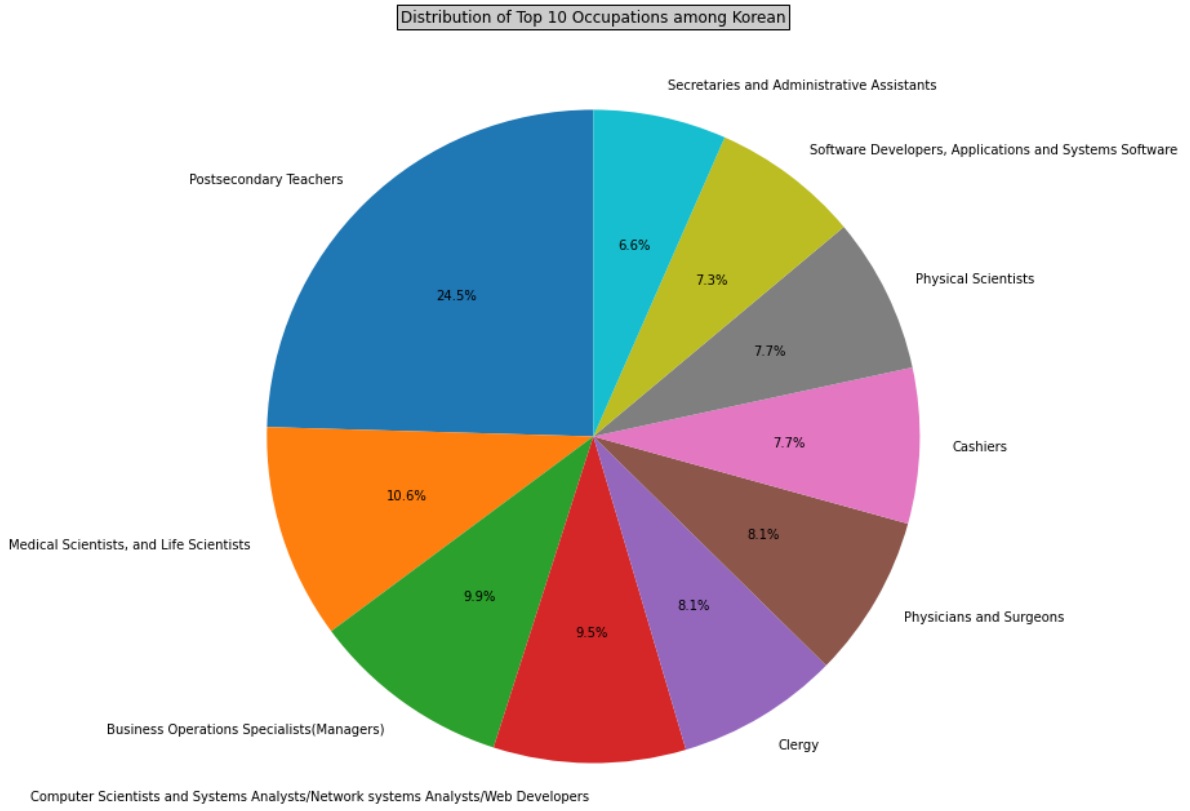
There are 59 people work as top 10 occupations, about 0.4306569343065693 of the total Thai Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Thai')



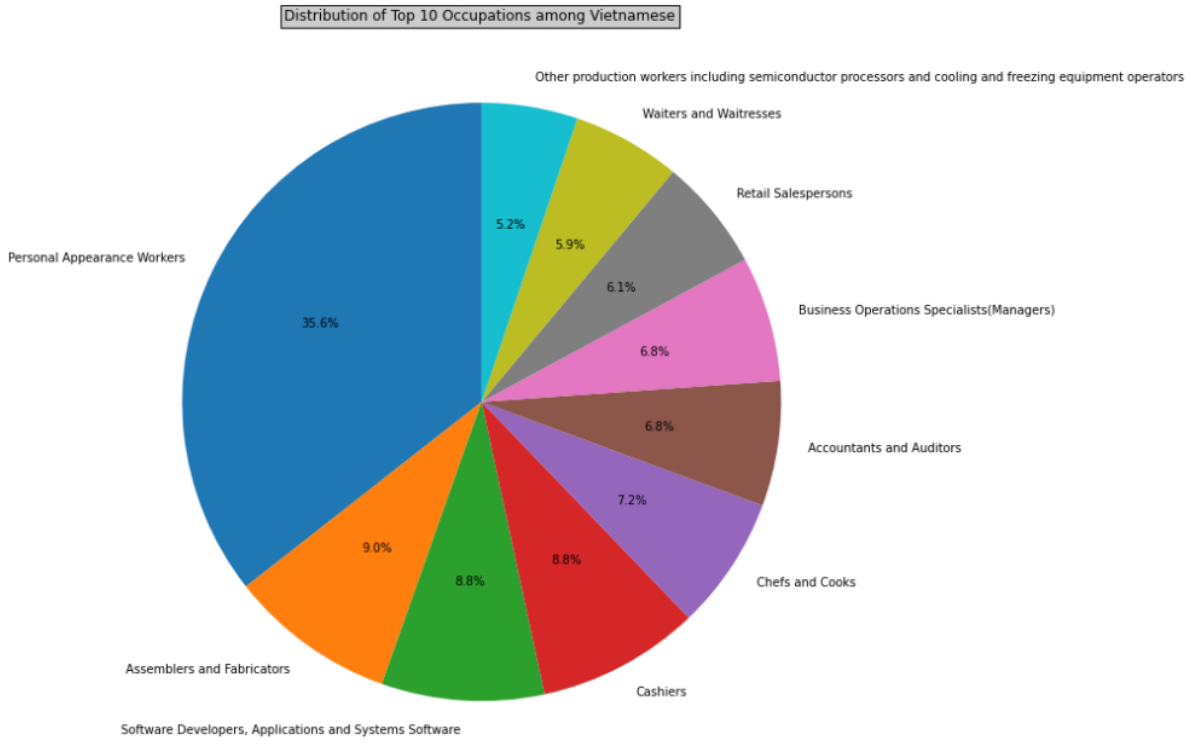
There are 273 people work as top 10 occupations, about 0.2941810344827586 of the total Korean Laborforce

Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Korean')



There are 444 people work as top 10 occupations, about 0.30833333333333335 of the total Vietnamese Laborforce

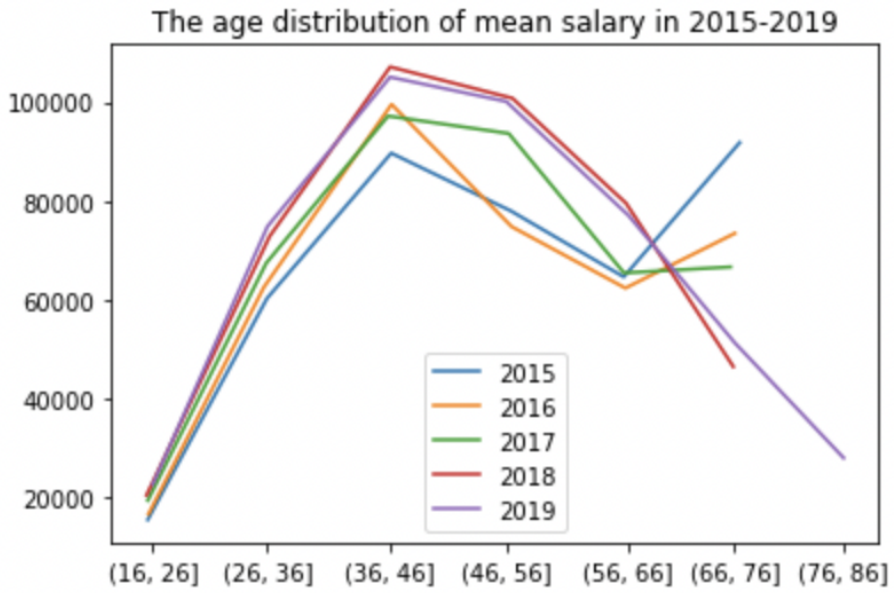
Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Vietnamese')



IT Professionals	2291
Managers	1496
Office and Administrative Support	1285
Healthcare Professionals	1237
Education Professionals	1196
Sales	1069
Researchers & Scientists	930
Business Operators	923
Other Service Providers	922
Food Service Providers	894
Production Workers	819
Social Service Providers & Military	369
Art & Media	334
Transportation	291
Consturction	126
Agriculture	11

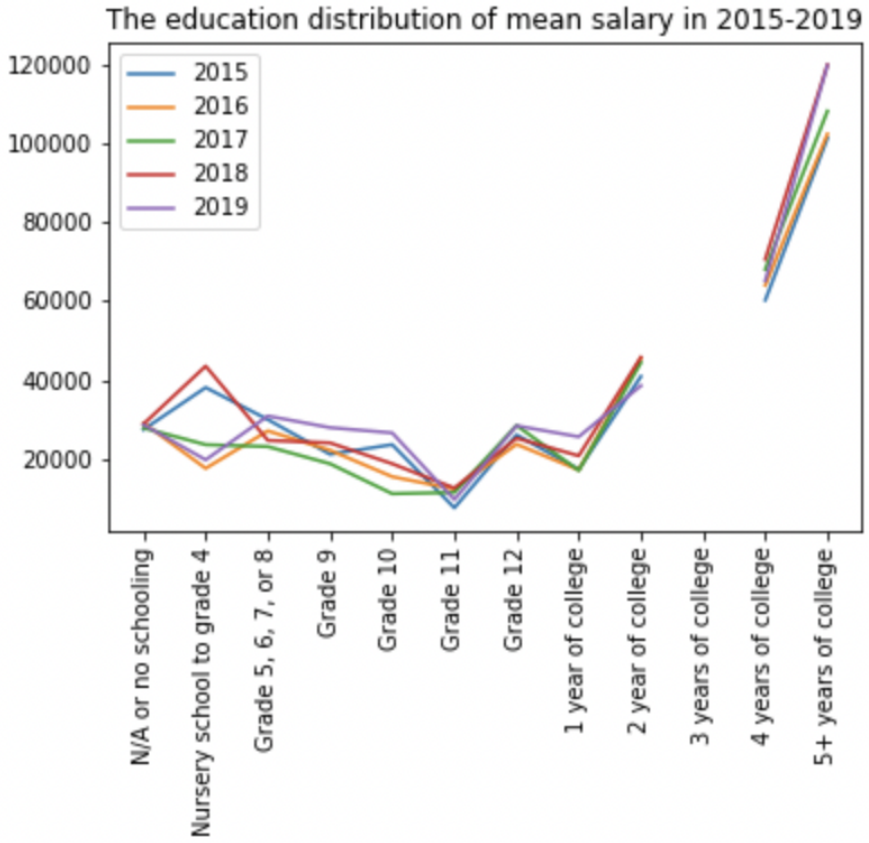
Name: OCC2010, dtype: int64

Age distribution



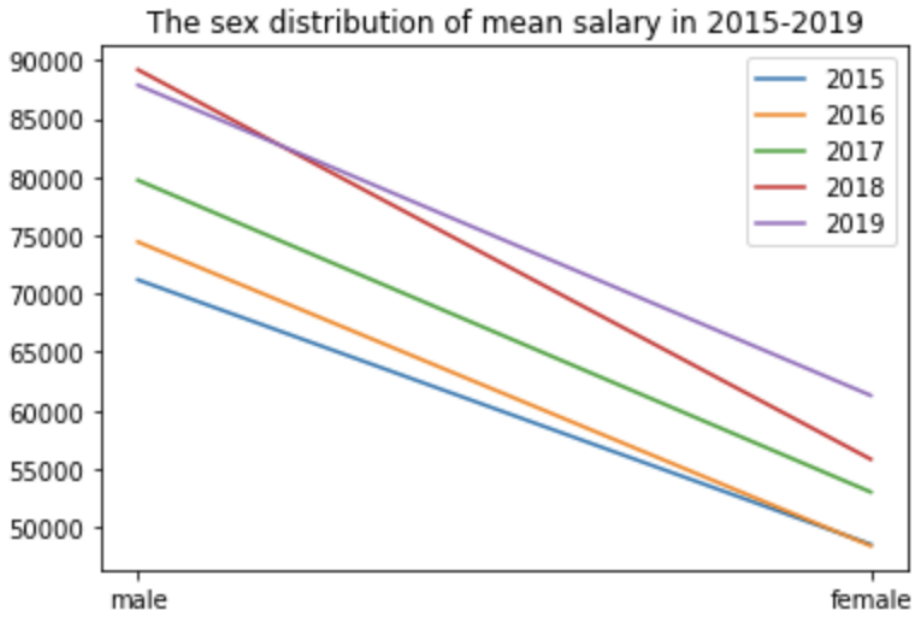
The highest salary is within the [36,46] and [46, 56] range.

Education distribution



We can clearly see that as the education level grows, the mean salary increases a lot.

Sex distribution



Work Cited:

Amanda Barroso and Anna Brown, “Gender pay gap in U.S. held steady in 2020”, Pew Research Center, 25 May 2021, https://www.pewresearch.org/fact-tank/2021/05/25/gender-pay-gap-facts/

D’Vera Cohn, Gretchen Livingston, and Wendy Wang, “Chapter 2: Stay-at-Home Mothers by Demographic Group”, Pew Research Center, 8 April 2014, https://www.pewresearch.org/social-trends/2014/04/08/chapter-2-stay-at-home-mothers-by-demographic-group/