Deliverable 3

BCNC (Boston Chinatown Neighborhood Center) - Asian Impact & Equity – Team1

Sponsor: Ben Hires

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

Abstract

There is a growing large number of Asians that live in America, 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 mainly focuses 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 is based on IPUMS datasets including ACS and CPS. 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. By analyzing the results we get, we find that there are differences within the Asian subgroups in both salaries and occupation selections. Also, the extent of effects of sex, education, and age are different among Asian and otehr races, which may be caused by cultural differences.

I. Goals of Our Team

- First, collect datasets in the scope of employment; then, retrieve data that is relative to Asians in MA; finally, 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 features of Asians and each Asian subgroup and the possible reasons for those relations.
- By analyzing the data collected, try to find the future trend of employment of Asians.

II. Collected Datasets

IPUMS ACS/CPS data with selected attributes as follows

- ACS (Year Selected from 2015 to 2019)

SAMPLES: 5 (hide) Change

Sample	Density	Note
2015 ACS	1.0%	
2016 ACS	1.0%	
2017 ACS	1.0%	
2018 ACS	1.0%	
2019 ACS	1.0%	

VARIABLES: 24 (hide) Change

Type	Variable	Label	Case Selection
H	YEAR	Census year	--
H	SAMPLE	IPUMS sample identifier	--
H	CBSERIAL	Original Census Bureau household serial number	--
H	STATEFIP	State (FIPS code)	details
H	COUNTYFIP	County (FIPS code)	--
H	CITY	City	--
P	SEX	Sex	--
P	AGE	Age	--
P	RACE (general)	Race [general version]	--
P	RACED (detailed)	Race [detailed version]	details
P	RACASIAN	Race: Asian	details
P	EDUC (general)	Educational attainment [general version]	--
P	EDUCD (detailed)	Educational attainment [detailed version]	--
P	EMPSTAT (general)	Employment status [general version]	--
P	EMPSTATD (detailed)	Employment status [detailed version]	--
P	OCC2010	Occupation, 2010 basis	--
P	IND	Industry	--
P	WKSWORK2	Weeks worked last year, intervalled	--
P	UHRSWORK	Usual hours worked per week	--
P	INCTOT	Total personal income	--
P	INCWAGE	Wage and salary income	--
P	INCBUS00	Business and farm income, 2000	--
P	PWSTATE2	Place of work: state	--
P	PWCOUNTY	Place of work: county	--

- CPS (Year Selected from 2010 to 2021)

SAMPLES: 5 (hide) Change

Sample	Note
IPUMS-CPS, ASEC 2015	
IPUMS-CPS, ASEC 2016	
IPUMS-CPS, ASEC 2017	
IPUMS-CPS, ASEC 2018	
IPUMS-CPS, ASEC 2019	

VARIABLES: 15 (hide) Change

Type	Variable	Label	Case Selection
H	YEAR	Survey year	--
H	STATEFIP	State (FIPS code)	details
H	COUNTY	FIPS county code	--
P	AGE	Age	--
P	SEX	Sex	--
P	RACE	Race	details
P	ASIAN	Asian subgroup	--
P	EMPSTAT	Employment status	--
P	OCC2010	Occupation, 2010 basis	--
P	IND	Industry	--
P	EDUC	Educational attainment recode	--
P	EDUC99	Educational attainment, 1990	--
P	UHRSWORKLY	Usual hours worked per week (last yr)	--
P	INCTOT	Total personal income	--
P	INCWAGE	Wage and salary income	--

We will provide both original datasets without modification in the git repository. As the columns are filled with code to illustrate the value stored, we have changed the code with values that are readable and meaningful. To get what those original codes indicate, please visit https://usa.ipums.org/usa/ for ACS and https://cps.ipums.org/cps/ for CPS, and search for each column name.

III. Preliminary Analysis

All features have been included in the section. The modification on those features will be included in the appendix; generally, we change the code to meaningful English words.

- ACS (American Community Surveys 2015-2019)
  - Total num of rows in the dataset(Asians in MA): 22349
  - Number of people in labor force: 15955
  - Number of people who reported their income: 18787
  - Number of people who reported their wage/salary: 12538

- CPS (Current Population Survey 2010-2021)
  - Total number of rows in the dataset(Asians in MA): 9119
  - Number of people in labor force(who report their wage): 5676
- Dataset resulted from merging ACS & CPS (2015-2019)
  - Total number of rows in the dataset(Asians in MA): 23830
  - Number of people in labor force: 16983
  - Number of people who reported their income: 20000
  - Number of people who reported their wage/salary: 13273

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       9316
Indian        5627
Other         3265
Vietnamese    2445
Korean        1468
Filipino       800
Japanese       627
Thai           186
Indonesian      96
```

  - CPS dataset is a voluntary survey, so it is biased on the sample selection. ACS dataset is a mandatory survey, but people can leave the information that they do not want to answer blank; the same situation also happens in CPS dataset. Thus, the data filled can also be biased, indicating that the people surveyed are willing to fill this information. 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 modeling may be the trend of specific households rather than the Asians as a whole group.)
  - The data is not large enough to train and predict. There are some sub-Asian groups with less than 100 people in the datasets. Then, the prediction and the general trend of that group are less accurate and representative.

V. Questions Answered

☐ Are there any relations between the income and education of Asians?(See appendix-education distribution) Are the gradients different among different Asian subgroups models?

From appendix-education distribution, we can find that high levels of education leads to high salaries among Asians, which is the same for non-Asian people (U.S. bureau of labor statistics, 2021). 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. 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?

See appendix-*List of Generalized Occupations*, we can find that the most popular career choice for Asians is "IT professionals", followed by "Manager". This trend can be illustrated by various reasons. One reason can be computer engineering is the major with the highest participation rate of Asian students. (statista, 2011) Since there are a lot of students studying CS, it is not surprising to see IT professionals to be popular among 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 among Asians and other races in these aspects?

We can see that Asian people have higher salaries than people of other races. (impactessexcounty.org, 2019). One such reason for this is that most Asians are immigrants rather than U.S. born. They had higher human capital and were employed in high-skill and high-wage occupations (academic.oup.com, 2016).
Also, it is also very interesting to see Asian students, when choosing majors, tend to make quite different choices than students of other races. While trying to find reasons, we found some amusing views. Compared to non-Asian, Asian students are much more likely to major in computer science related majors. This fact may be contributed by the Asian cultures. Asian cultures tend overall to place a high value on computer science related jobs, so a family's wishes for their child are likely an overwhelming influence.(zippia.com, 2017)

- [ ] 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?

  From the Appendix - Correlation between salaries and some features, we can see that careers with the highest earnings are "Physicians and Surgeons", "Financial managers", "Computer and Information Systems Managers", "Software Developers, Applications and Systems Software". These careers are quite consistent with the popular occupations of Asian (see Appendix - List of Generalized Occupations)

- [ ] Are there differences in career choices between different groups of Asians?

  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.

- [ ] 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 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.

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

  Asian people tend to choose computer related majors, which is quite different from people of other races. The field with the highest concentrations of whites is agriculture and natural resources (90 percent). Law and public policy has the highest concentration of African-Americans (14 percent) and Hispanics (10 percent). (newsone.com, 2011)
  We found that Asian people prefer to major in computer related majors, and that computer science related jobs are popular among Asian people (see Appendix - List of Generalized Occupations). Thus, clearly there is indeed a connection between major choices 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?

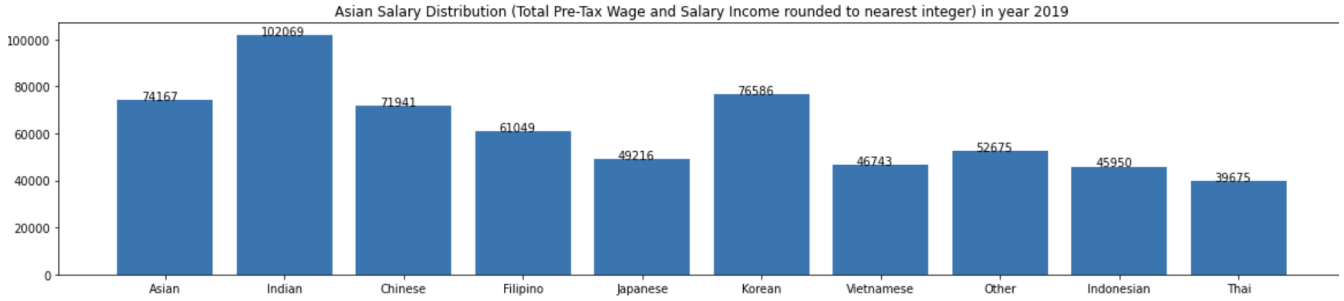
From the figure in the appendix, we can see that the highest salary is within the [36,46] and [46, 56] range, which is quite consistent with our understanding. But, for those careers, like the executive staff, that are for people with more experience in that scope, people who have high salaries tend 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.

VI. Appendix

- **Appendix Catalogue**

- Salary Distribution
  - ACS

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



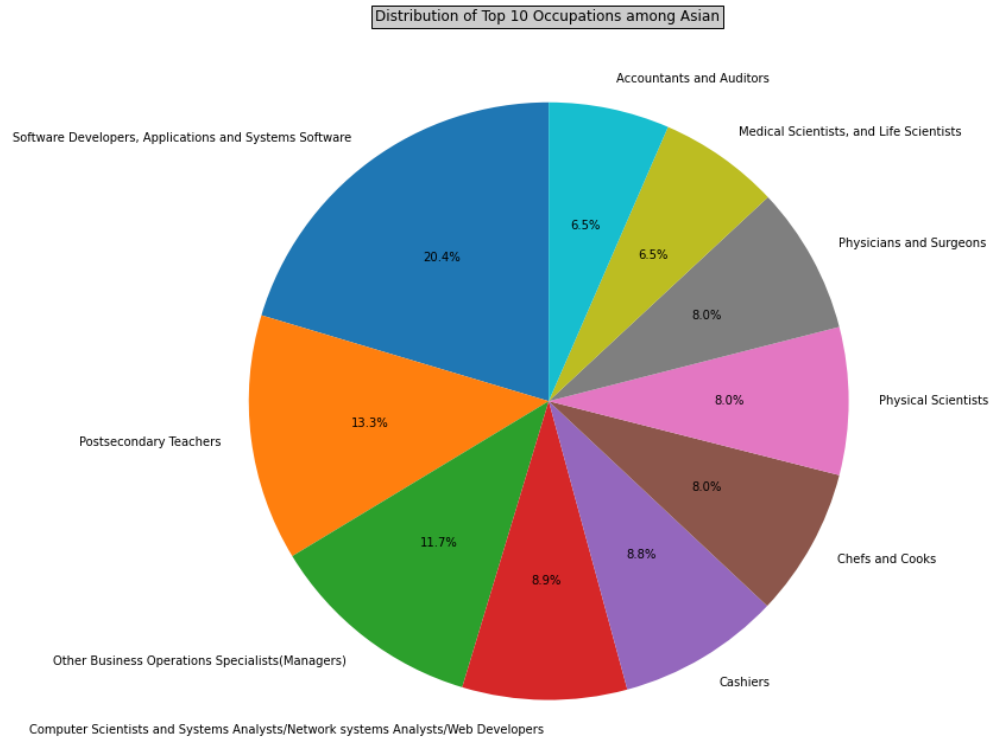
  - CPS

There are totally 5676 labor force in those Asian

- Occupation Distribution of Asian Subgroups (by ACS dataset)

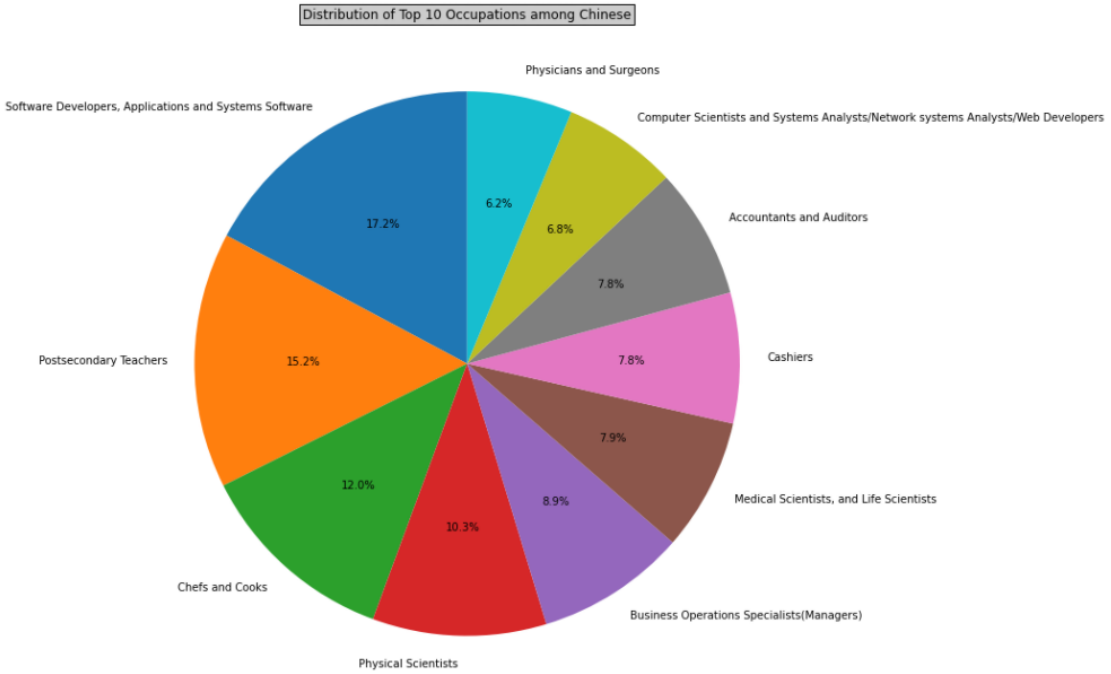
Distribution of Top 10 Occupations among Asian

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



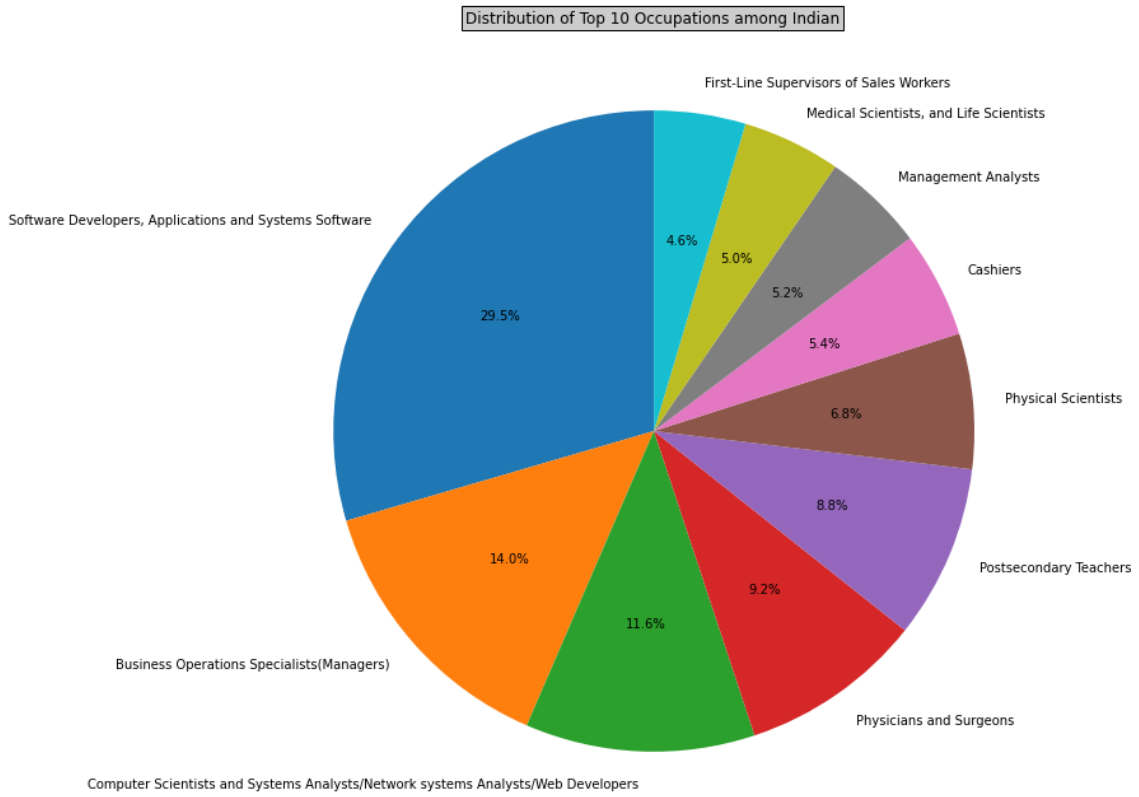
Distribution of Top 10 Occupations among Chinese

```
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')
```



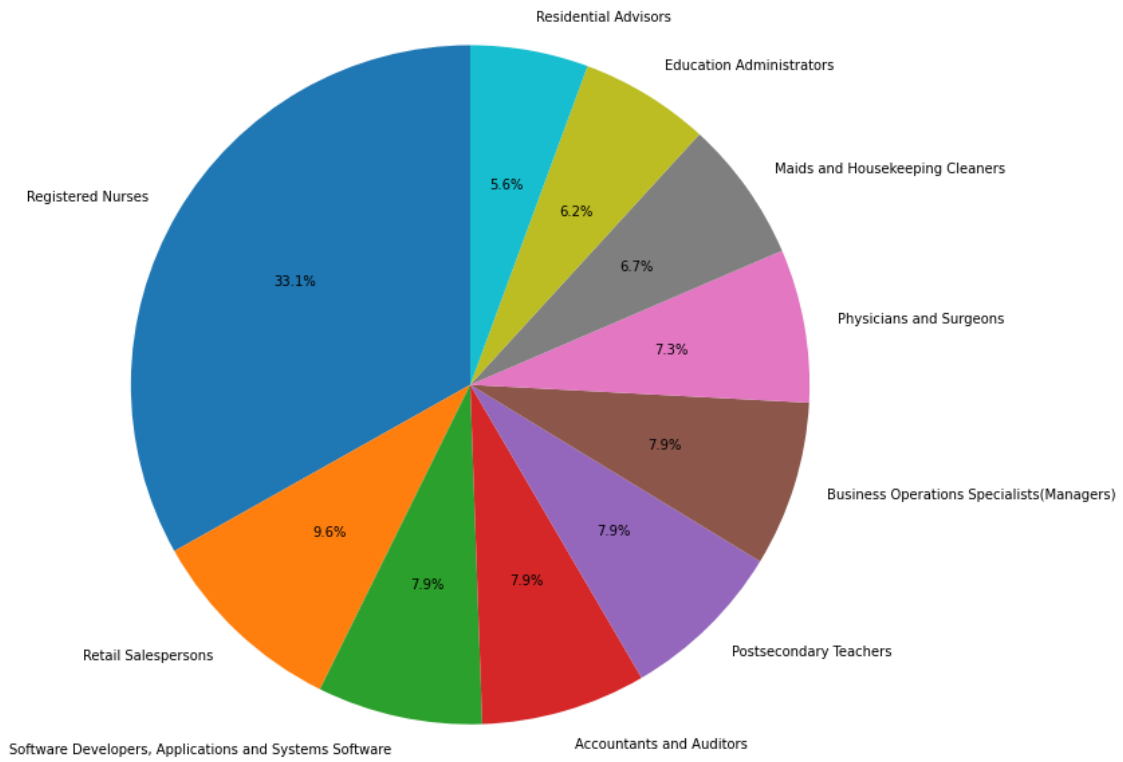
Distribution of Top 10 Occupations among Indian

```
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')
```



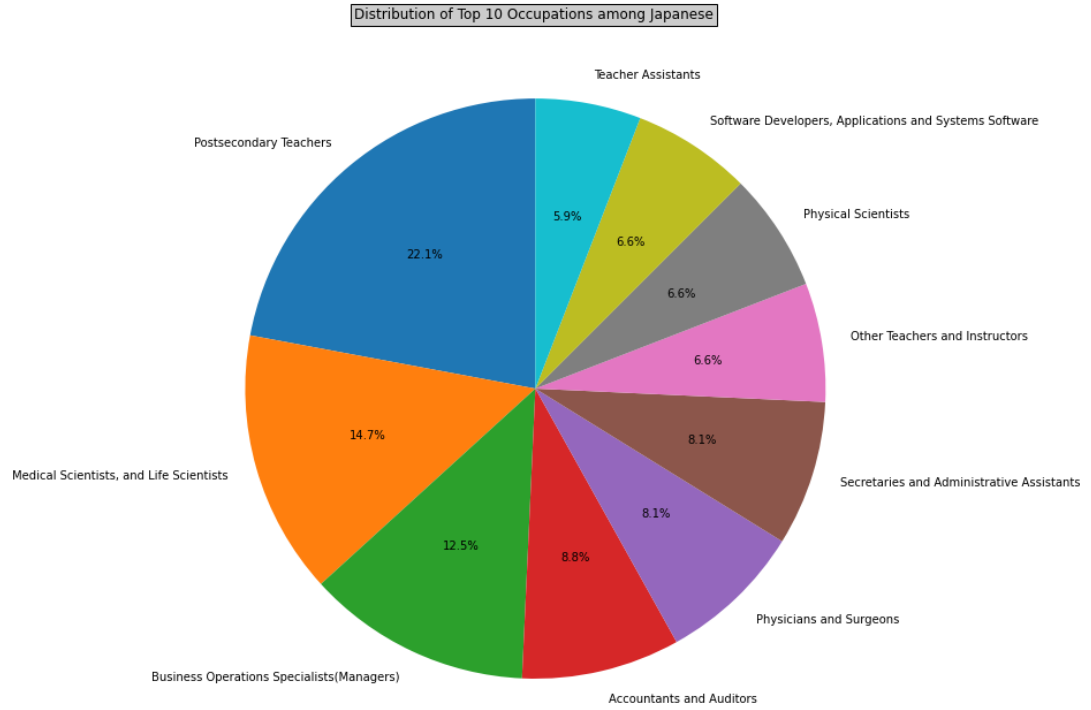
Distribution of Top 10 Occupations among Filipino

```
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')
```



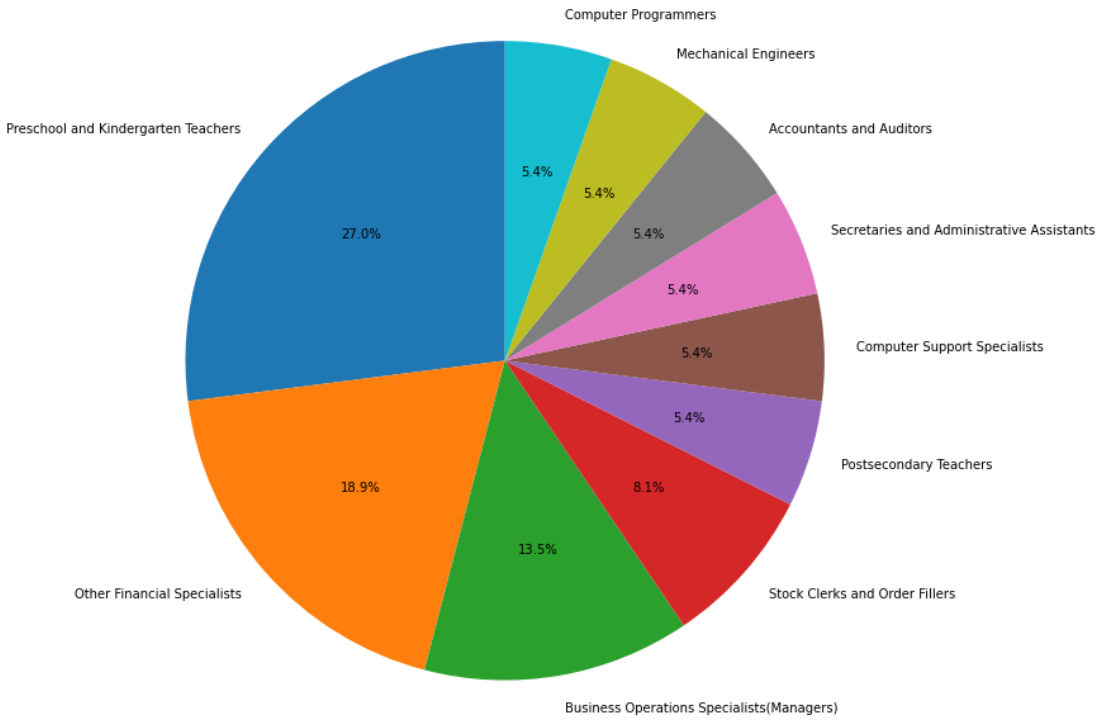
Distribution of Top 10 Occupations among Japanese

```
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')
```



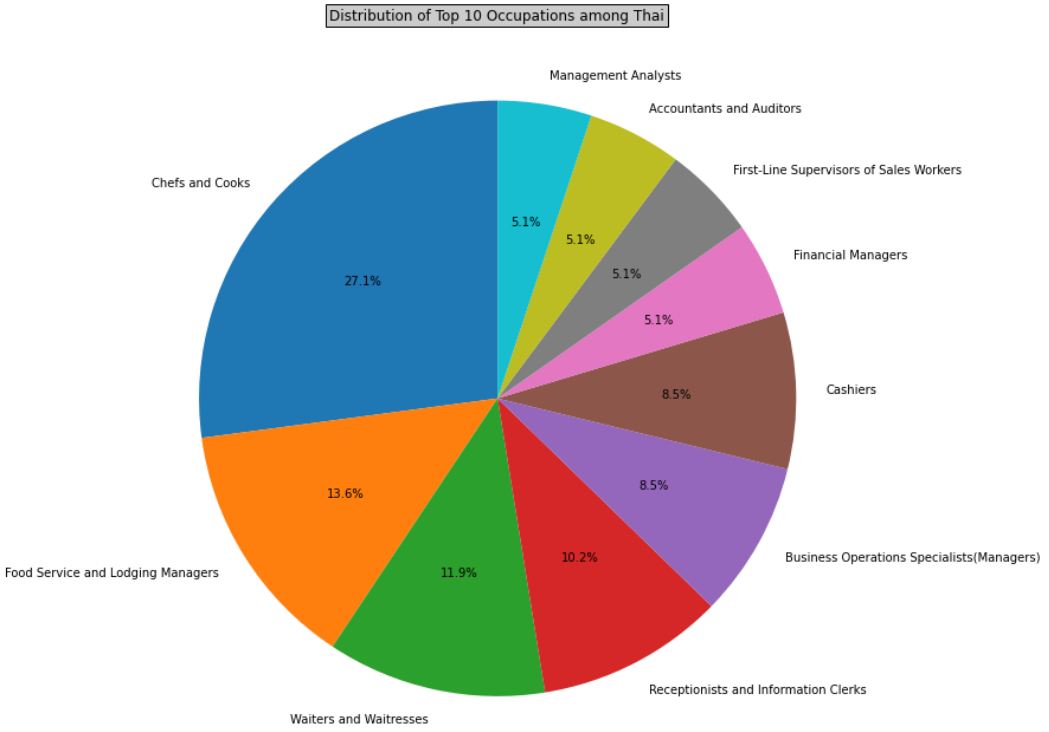
Distribution of Top 10 Occupations among Indonesian

```
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')
```



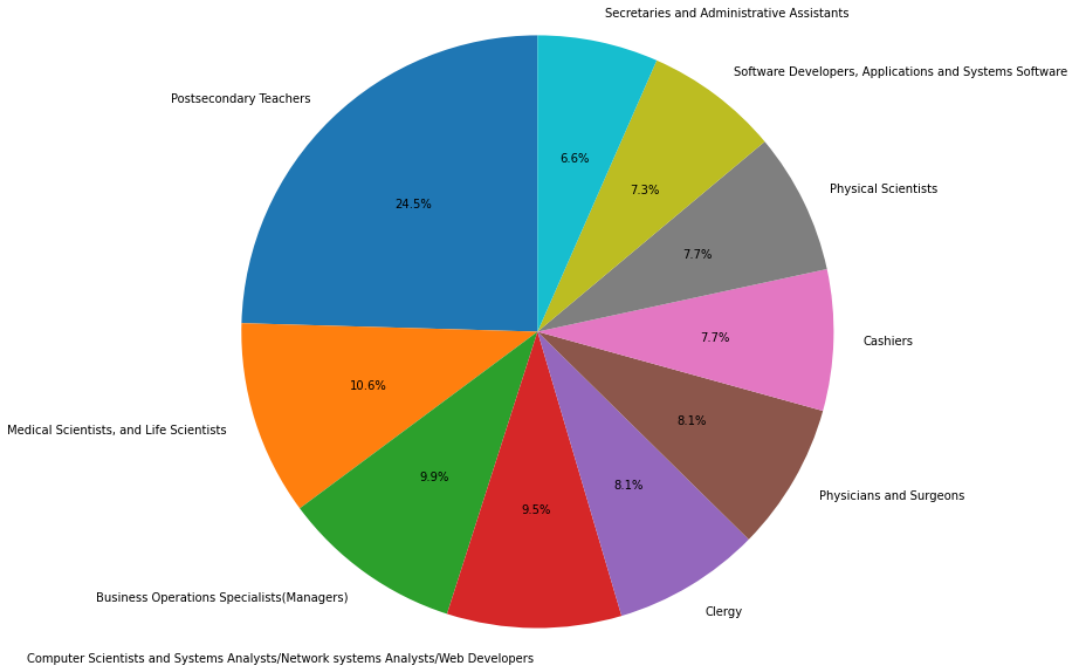
Distribution of Top 10 Occupations among Thai

```
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')
```



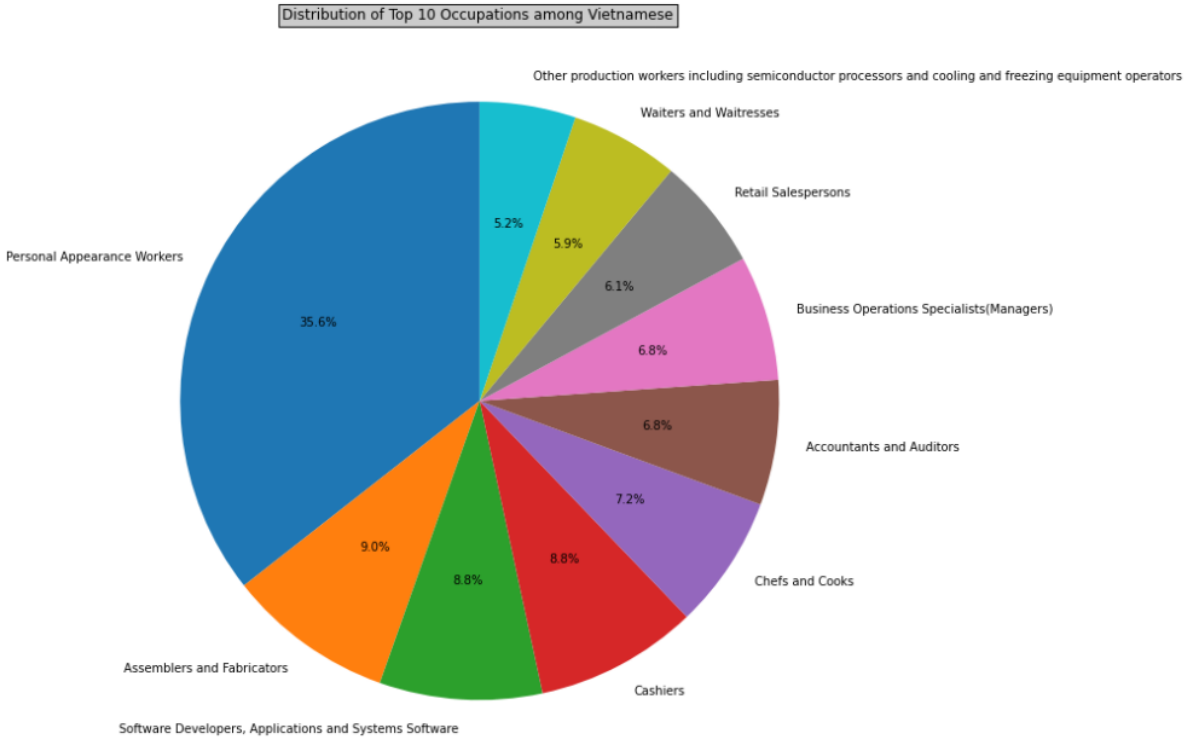
Distribution of Top 10 Occupations among Korean

```
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')
```



Distribution of Top 10 Occupations among Vietnamese

```
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')
```



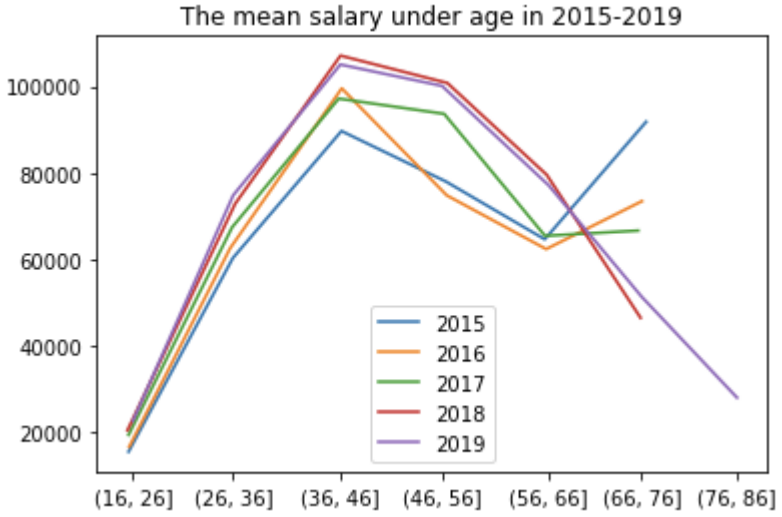
List of Generalized Occupations

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

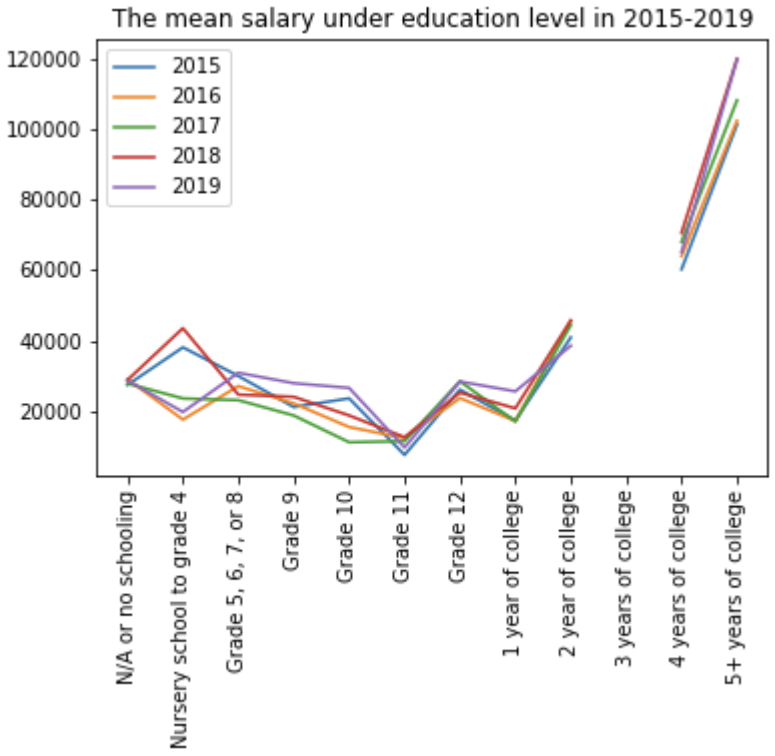
- Correlation between salaries and some features

Mean salary under age in 2015 - 2019



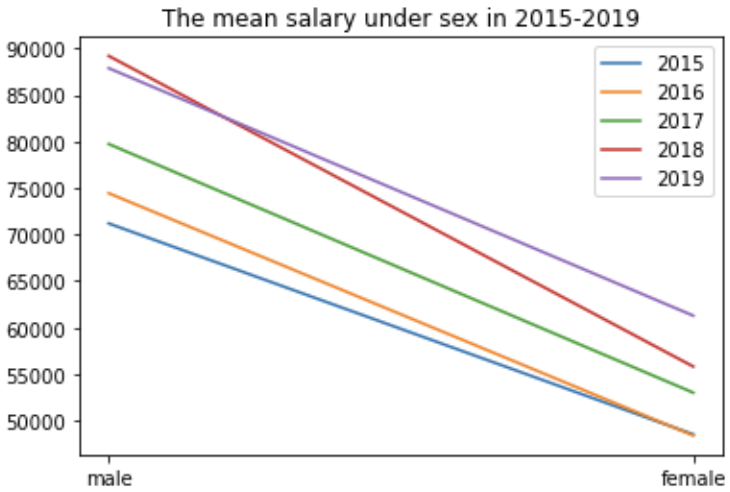
The highest salary is within the [36,46] and [46, 56] range.

Mean salary under education level in 2015 - 2019

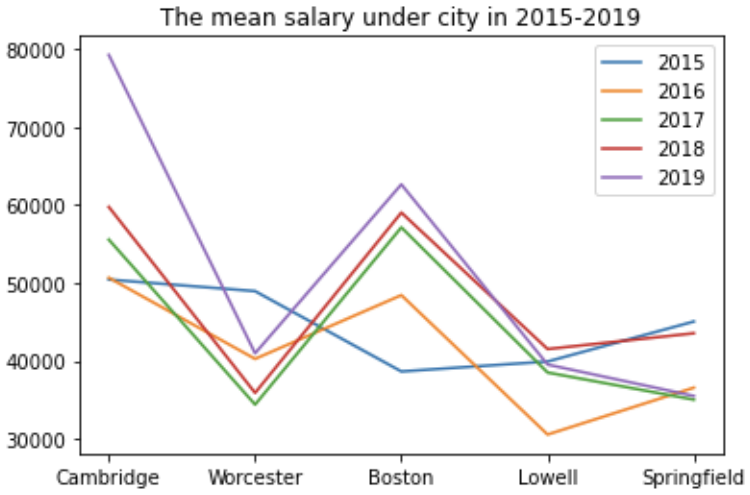


We can clearly see that as the education level grows, the mean salary increases a lot.

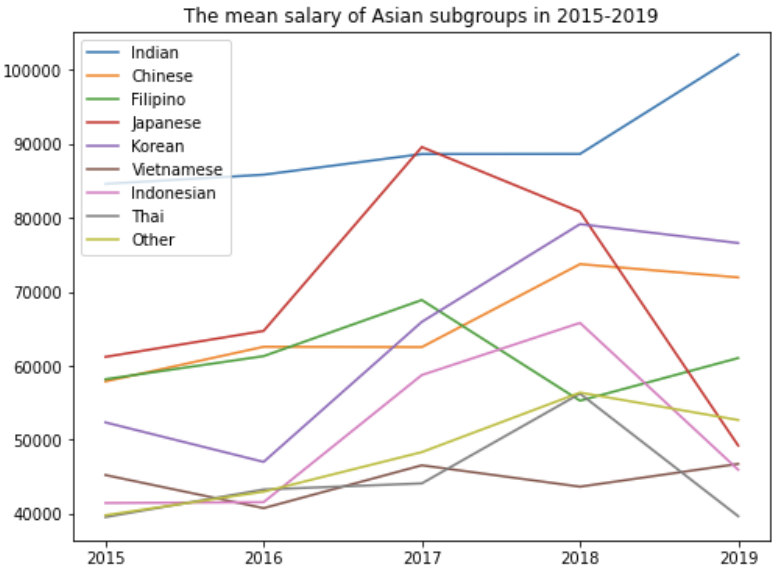
Mean salary under sex in 2015 - 2019



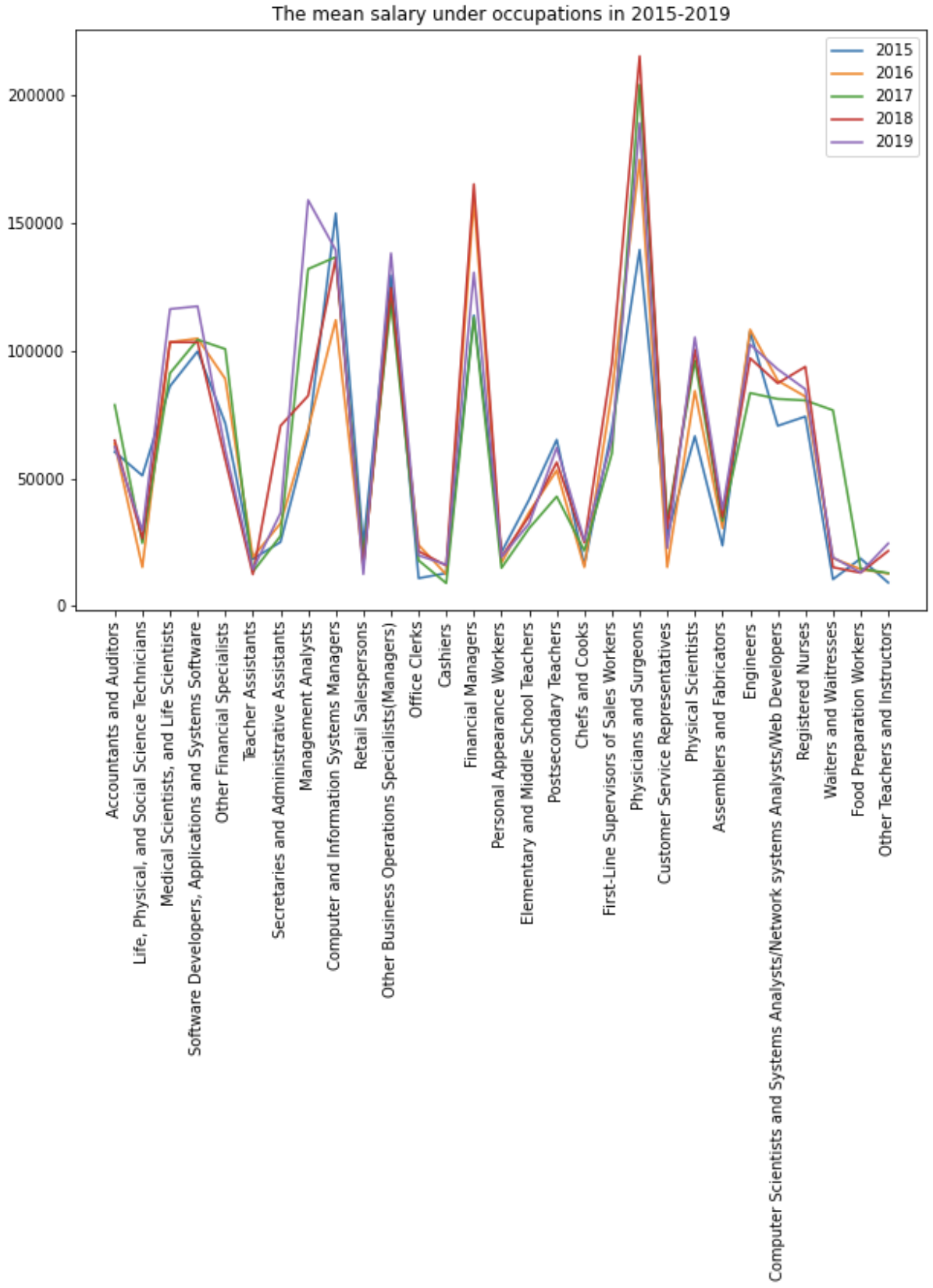
Mean salary under city in 2015 - 2019



Mean salary of Asian subgroups in 2015 - 2019



Mean salary under occupations in 2015 - 2019



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/

IPUMS-CPS, University of Minnesota, www.ipums.org.

Impact Essex Country, "Median Household Income, by Race/Ethnicity, 2015-19", impactessexcountry.org, Accessed by 17 Nov 2021, https://impactessexcounty.org/financial-self-sufficiency/median-household-income-by-race-ethnicity

Statista, "College majors with the highest participation rate of Asian students in the U.S. in 2009", statista.com, May 24, 2011, https://www.statista.com/statistics/225726/college-majors-with-the-highest-participation-of-asian-students-in-the-united-states/

Wage assimilation of immigrants: A comparison of 'new' and 'old' Asian source countrie, October 20 2016, https://academic.oup.com/migration/article/5/1/1/2566844

Racial and Ethnic Differences in College Major Choice https://www.clevelandfed.org/newsroom-and-events/publications/economic-trends/2015-economic-trends/et-20150331-racial-and-ethnic-differences-in-college-major-choice.aspx

Study: College Majors Are Highly Segregated By Race, Gender https://newsone.com/1252915/college-majors-highly-segregated-by-race-and-gender/

U.S. Bureau of labor statistics: Education pays, April 21, 2021, https://www.bls.gov/emp/chart-unemployment-earnings-education.htm

A detailed look at college major distributions by race, October 15, 2017, https://www.zippia.com/advice/college-major-distributions-by-race/