

## **Deliverable 4**

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 other 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	<a href="#">(hide)</a>	<a href="#">Change</a>
----------	---	------------------------	------------------------

Sample	Density	Note
<a href="#">2015 ACS</a>	1.0%	
<a href="#">2016 ACS</a>	1.0%	
<a href="#">2017 ACS</a>	1.0%	
<a href="#">2018 ACS</a>	1.0%	
<a href="#">2019 ACS</a>	1.0%	

VARIABLES:	24	<a href="#">(hide)</a>	<a href="#">Change</a>
------------	----	------------------------	------------------------

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

- CPS (Year Selected from 2015 to 2019)

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

We will provide both original datasets without modification in the git repository. As the columns are filled with codes to illustrate the value stored, we have changed the codes with values that are readable and meaningful. To get what those original codes indicate, please visit [ACS search](#) and [CPS search](#), and search for each column name.

### III. Preliminary Analysis

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

- 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

Is there any relation between the income and education of Asians? Are the gradients different among different Asian subgroups models?

See appendix - [Mean salary under education](#), we can find that high levels of education leads to high salaries among Asians; the same pattern can also be found for non-Asian people (U.S. Bureau of Labor Statistics, 2021).

Also, from appendix - [Salary Distribution](#) and [Mean salary of Asian subgroups](#), it is quite apparent that Indian Americans earn more than any other Asian ethnicities. This is another proof that high levels of education result in high levels of income. Since education levels among those who identify as Indian are incredibly high: Roughly 76% of those above the age of 25 have graduated from college. (Quartz, 2014)

Besides, we also found that in recent years, the gap of average earnings between Asian people and white people with the same education level has shrunk dramatically. It is reported that, with the same amount of schooling, Asians used to get paid like blacks, but between 1940 and 1970, they started to get paid like whites; and after 1980, they even started to earn more than whites (The Washington Post, 2017).

Though, starting from 1940, many Asian American families did invest, increasingly, in their children's education. It is believed that the shrinkage of salary gap between Asians and white people was not quite relative to Asians' accruing more education, but to some social and historical issues. For example, it is mentioned that after world war II, American racial attitudes towards Asians shifted drastically—people praised Asians for being hardworking, cherishing education and rarely complaining. (The Washington Post, 2017).

What is the salary distribution of Asians? (See appendix-[Salary Distribution](#))

According to the bar graph, the average pre-tax salary of Asians in MA is \$67,553 from 2015 to 2019. If we divide them into subgroups, Indians, Japanese, and Koreans are the top-three subgroups that have the highest average salary, which are \$91,038, \$69,698, and \$66,326 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. (Financial self-sufficiency). 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 (Basu, 2017).

Also, it is 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, which tend to place a high value on computer science related jobs, so a family's wishes for their child are likely an overwhelming influence (Morris, 2017).

Salary gap between males and females among Asians is larger than the ones of other races. This doesn't indicate that Asian women produce less labor force compared to women of other races. Actually, Asian women have the highest average hourly earnings among women of all races, and they even have higher hourly earnings compared to black and Hispanic men. The reason that gender inequality regarding salary is more severe among Asian people is that Asian males tend to earn more and more in recent years—they actually earn the most among males of all races. Both Asian females and Asian males are making progress with regard to salary, but Asian females progress regarding income not as fast as Asian males do. Thus, the salary gap appears. We won't give further discussion of this lag as gender inequality is not in the scope of our project. However, one thing worth mentioning is that white and Asian women have narrowed the wage gap with white men to a much greater degree than black and Hispanic women (Pew Research Center, 2016).

What are the popular career choices within Asians?

See appendix-[List of Generalized Occupations](#), we can find that the most popular career choices for Asians are "IT professionals", followed by "Manager". This trend can be illustrated by various reasons. One of the possible reasons is that computer engineering has been the major with the highest participation rate of Asian American students (Statista, 2011). The case is quite different from 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) (Staff, 2011). Since there are a lot of Asian students studying CS, it is not surprising to see IT professionals to be popular among Asians. The possible motivation for them majoring in computer engineering might be the promising future and high paying jobs of the CS-related industry.

Which career choices in Asians have the highest earnings? Are these the same as the popular occupations of Asians?

From the Appendix - [Mean salary under occupations](#) , 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 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-[Mean salary under sex](#))

According to our dataset, males tend to have higher income than females, and this is not mitigated as the year increases. This finding is also supported by the research done by Amanda Barroso and Anna Brown, which showed that 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 exists a correlation between sex and income; however, the reasons for it are quite complicated.

One key question is that does this indicate more serious gender inequality among Asians? The answer is no, as there are many other factors that can cause such an outcome. One main reason is that Asian mothers are more likely to be stay-at-home mothers compared to white and black mothers (D’Vera Cohn et al., 2014), therefore causing Asian females to have a lower mean salary in total.

Is there any correlation between age and income? What are the possible causes for the differences?

From appendix - [Mean salary under ages](#), we can see that the highest salary is within the [36,46] and [46, 56] range, which is quite consistent with our understanding since people tend to have higher positions when they have been in a company longer and are more experienced. However, for jobs that need more creativity, people who have high salaries tend to be younger. As a result, rather than focusing exclusively on age, the relationship between age and salary should also take into account job

characteristics. From appendix - [Mean salary of ages under occupations](#) , we can see that different occupations have different age peaks when looking at average salary in function of age. For example, we can see that people in "Computer scientists and Systems Analysts/Network Systems Analysts/ Web developers" start to have a higher average salary in their [33, 43]. However, for people in "Postsecondary teachers", they tend to have a much higher average salary after their 53. This comparison is quite reasonable, since for people working in the scope of computer science, they have more production in their middle age, whereas for people as postsecondary teachers, more experience leads to higher salary, hence people after 53 tend to earn more.



## VI. Appendix

- **Appendix Catalogue**

- [General Information of Merged Dataset](#)
- [Average Salary by year](#)
  - [Asians as a whole](#)
  - [Chinese](#)
  - [Filipino](#)
  - [Indian](#)
  - [Indonesian](#)
  - [Japanese](#)
  - [Korean](#)
  - [Thai](#)
  - [Vietnamese](#)
- [Salary Distribution](#)
- [Occupation Distribution of Asian Subgroups](#)
  - [Asian as a whole](#)
  - [Chinese](#)
  - [Filipino](#)
  - [Indian](#)
  - [Indonesian](#)
  - [Japanese](#)
  - [Korean](#)
  - [Thai](#)
  - [Vietnamese](#)
  - [List of Generalized Occupations](#)
- [Correlation between salaries and some features](#)
  - [Mean salary under age](#)
  - [Mean salary under education](#)
  - [Mean salary under sex](#)
  - [Mean salary of Asian subgroups](#)
  - [Mean salary of age under different occupations](#)
  - [Mean salary under occupation](#)
- [Work Cited](#)

## General Information of Merged Dataset

(ACS merged with CPS)

Total num of rows in the dataset: 23830

Number of people in laborforce: 16983

Number of people who reported their income: 20000

Number of people who reported their wage/salary: 13273

Classify by City:

Unknown	17202
Boston	3611
Lowell	1080
Cambridge	1076
Worcester	663
Springfield	198

Name: CITY, dtype: int64

Classify by Sex:

Female	12607
Male	11223

Name: SEX, dtype: int64

Classify by Asian Subgroups:

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

Name: RACED, dtype: int64

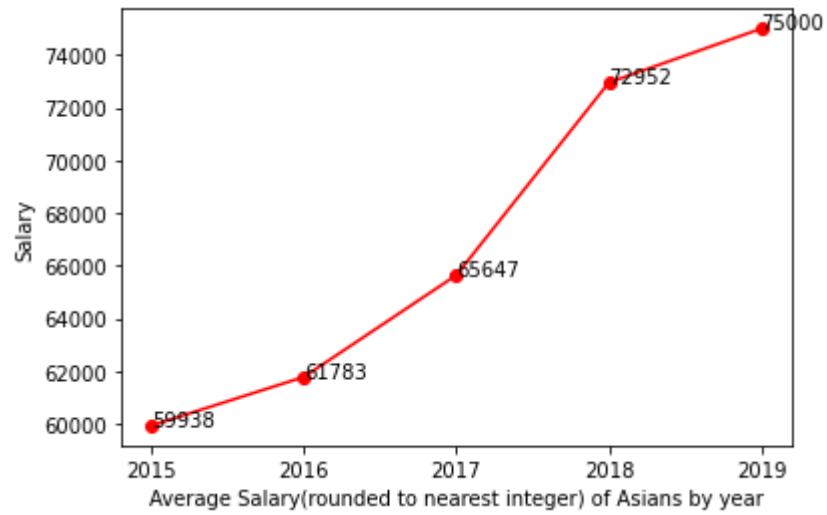
Classify by Employment Status:

Employed	12337
Not in labor force	6847
Unemployed	573
Unknown	4073

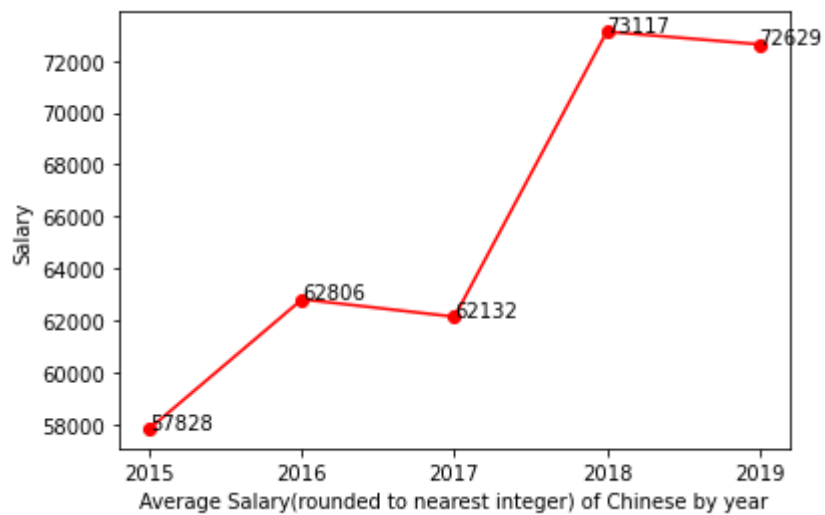
Name: EMPSTAT, dtype: int64

## Average Salary

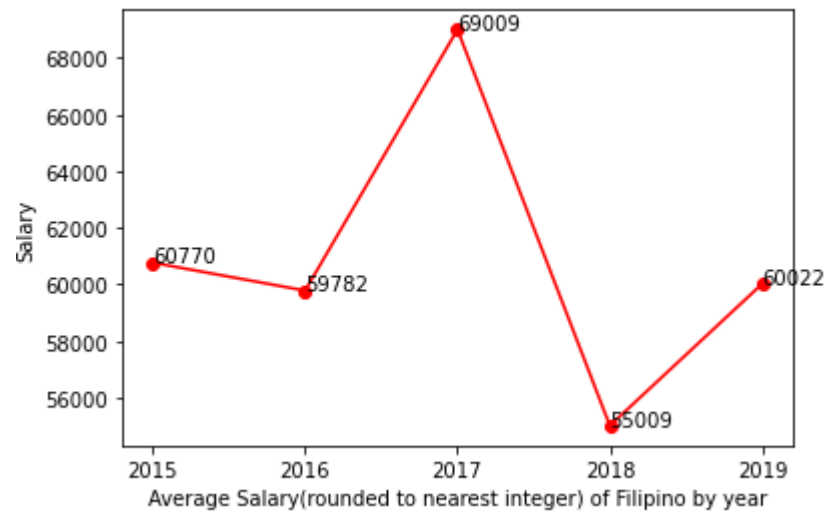
*Average Salary for Asians as a Whole by year*



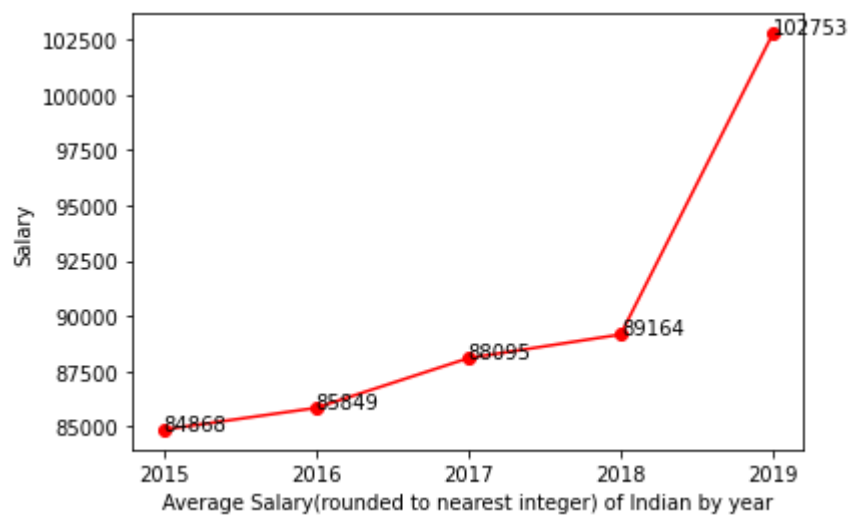
*Average Salary for Chinese by year*



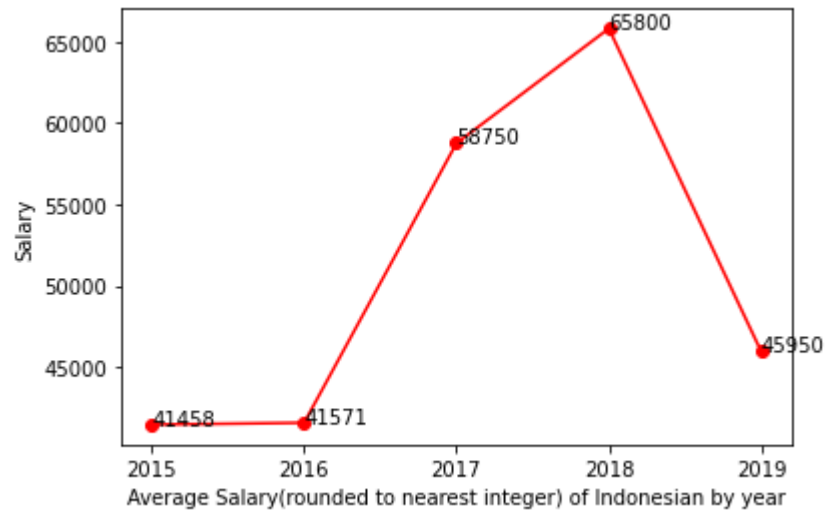
*Average Salary for Filipino by year*



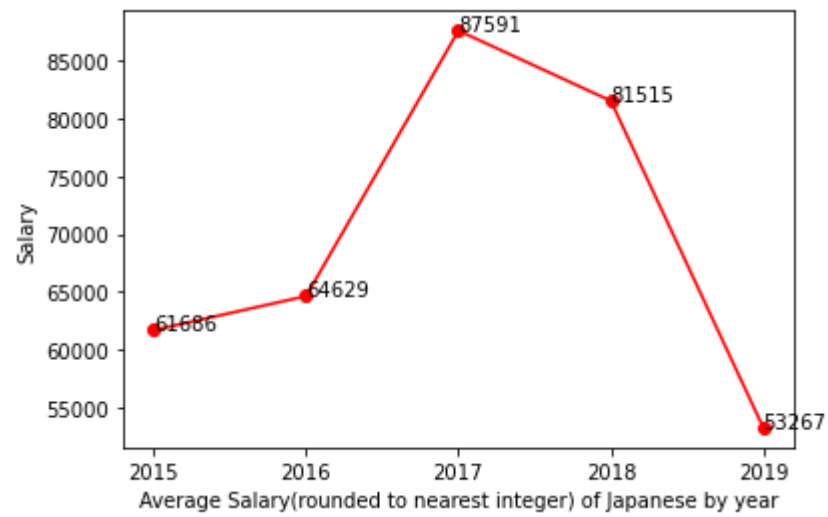
*Average Salary for Indian by year*



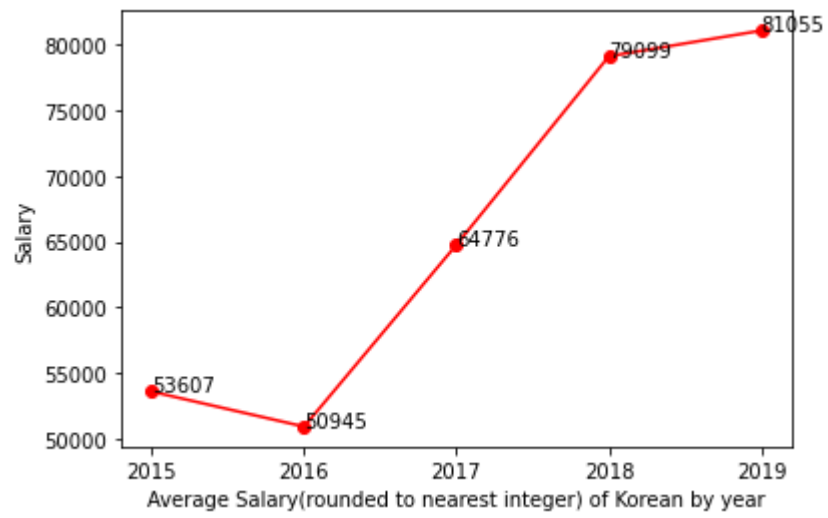
*Average Salary for Indonesian by year*



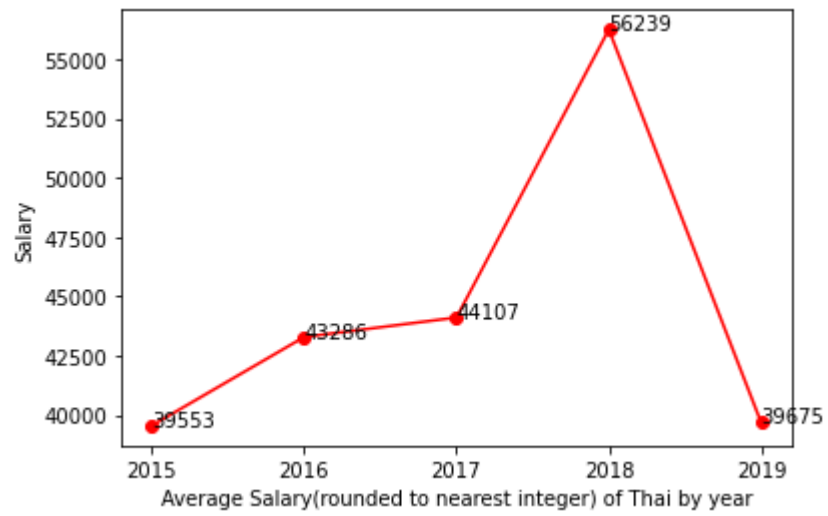
*Average Salary for Japanese by year*



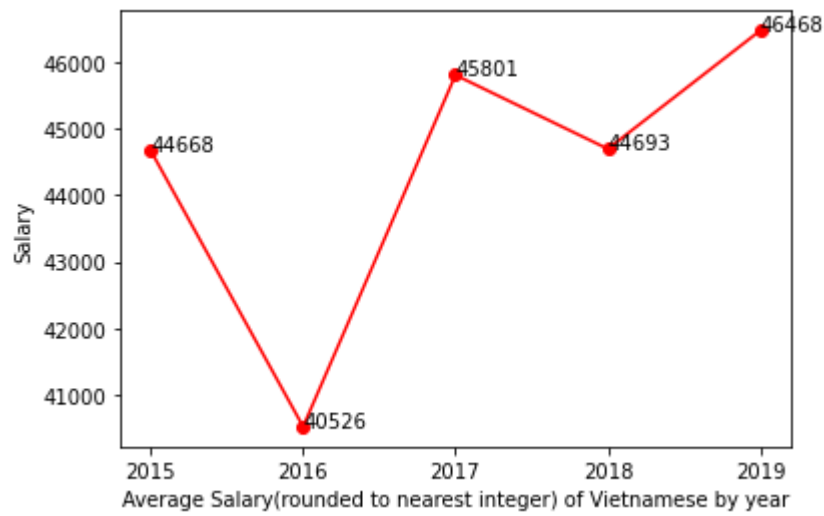
*Average Salary for Korean by year*



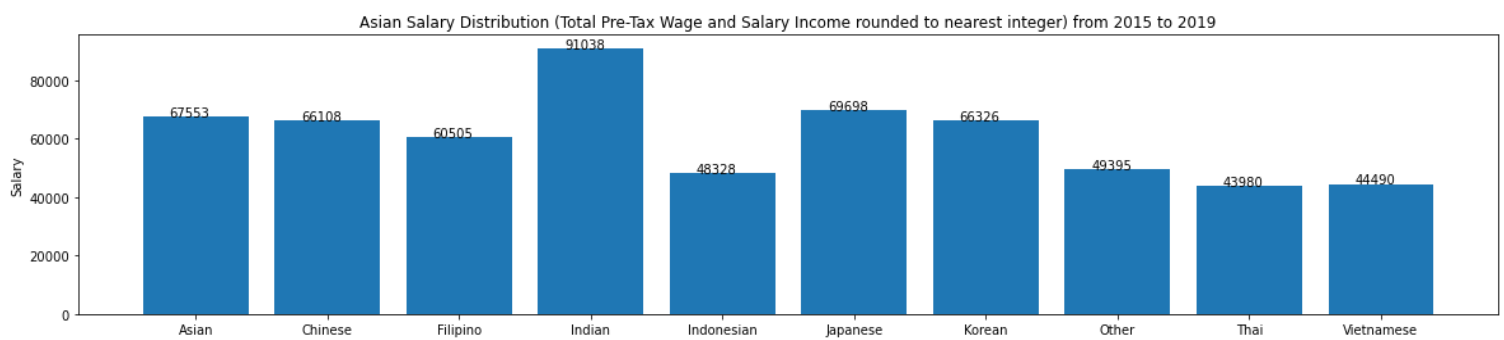
*Average Salary for Thai by year*



*Average Salary for Vietnamese by year*



## Salary Distribution

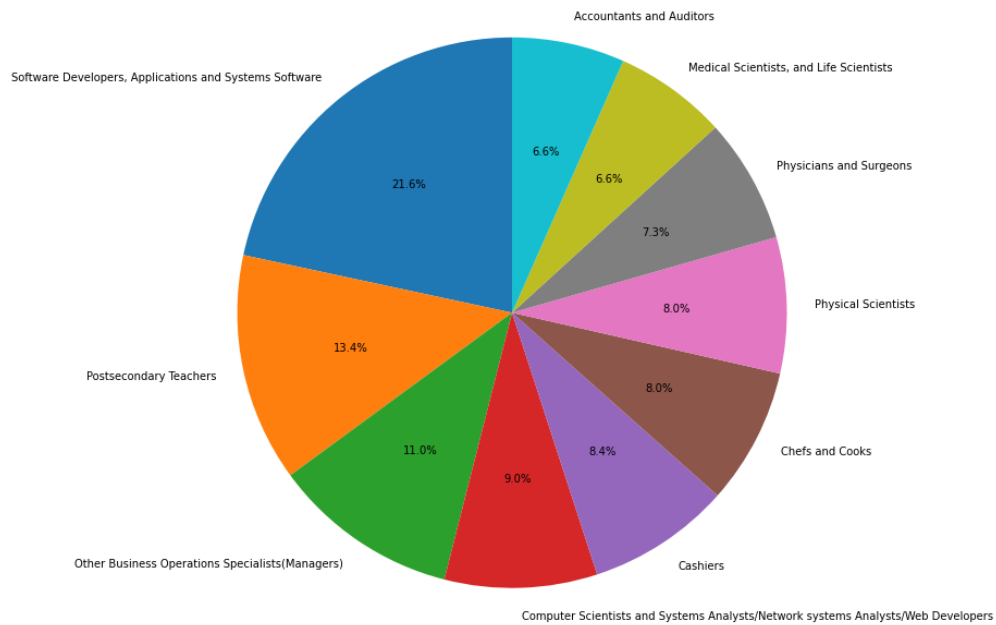


# Occupation Distribution

## *Distribution of Top 10 Occupations among Asian*

There are 4210 people work as top 10 occupations, about 0.3171852633165072 of the total laborforce

Distribution of Top 10 Occupations among Asian

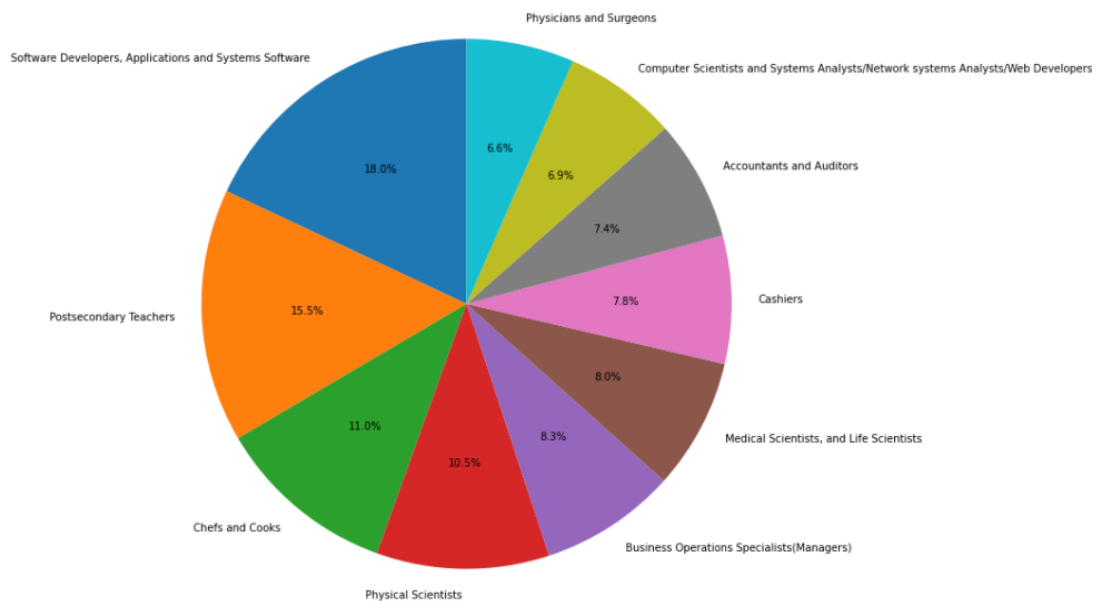


## *Distribution of Top 10 Occupations among Chinese*

There are 1766 people work as top 10 occupations, about 0.3428460493108134 of the total Chinese Laborforce

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

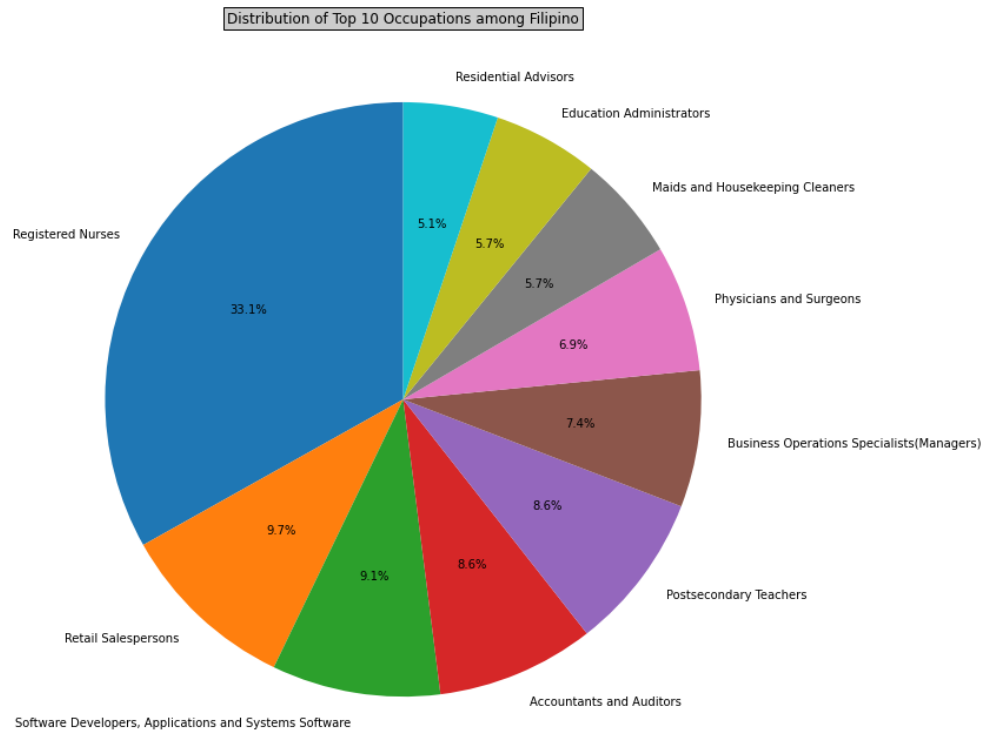
Distribution of Top 10 Occupations among Chinese





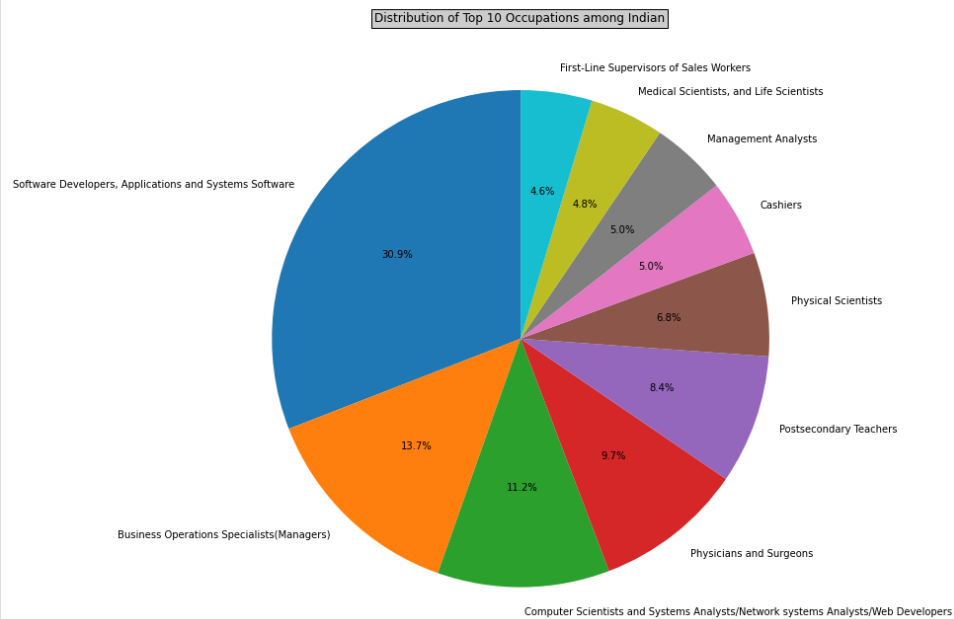
## Distribution of Top 10 Occupations among Filipino

There are 175 people work as top 10 occupations, about 0.3295668549905838 of the total Filipino Laborforce  
 Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Filipino')



## Distribution of Top 10 Occupations among Indian

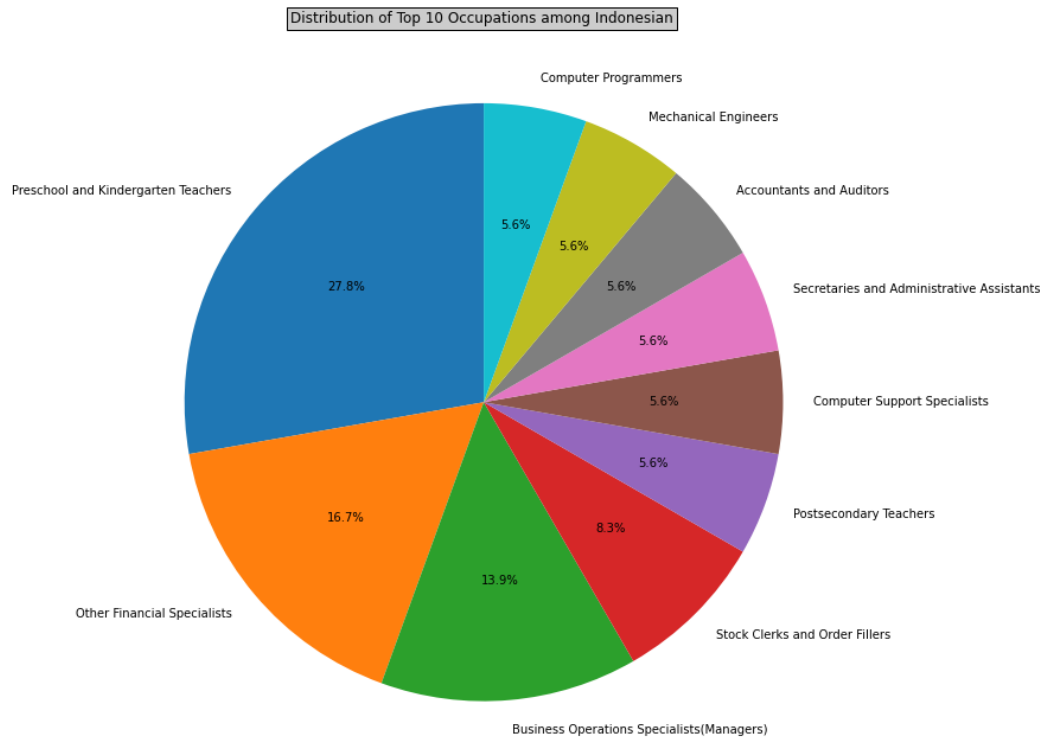
There are 1511 people work as top 10 occupations, about 0.46563944530046225 of the total Indian Laborforce  
 Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Indian')



## Distribution of Top 10 Occupations among Indonesian

There are 36 people work as top 10 occupations, about 0.6 of the total Indonesian Laborforce

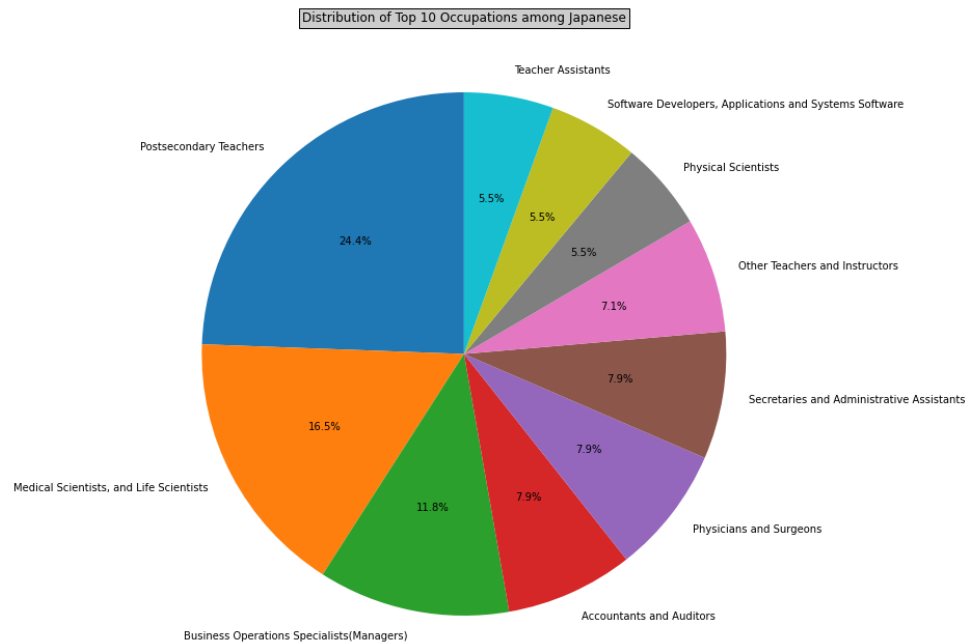
Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Indonesian')



## Distribution of Top 10 Occupations among Japanese

There are 127 people work as top 10 occupations, about 0.4057507987220447 of the total Japanese Laborforce

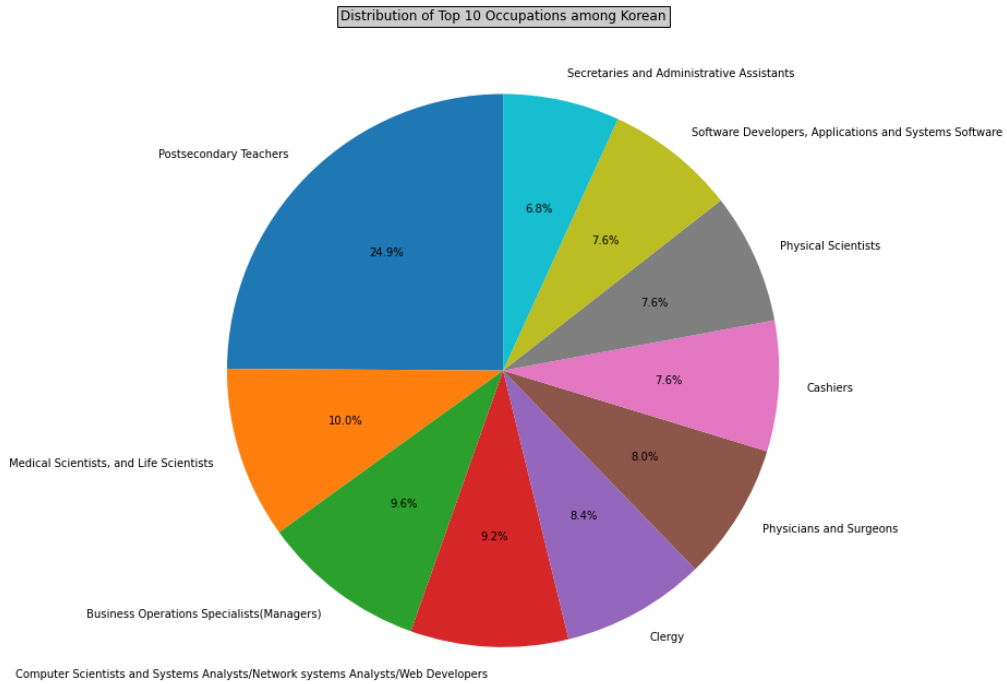
Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Japanese')



### *Distribution of Top 10 Occupations among Korean*

There are 249 people work as top 10 occupations, about 0.28987194412107103 of the total Korean Laborforce

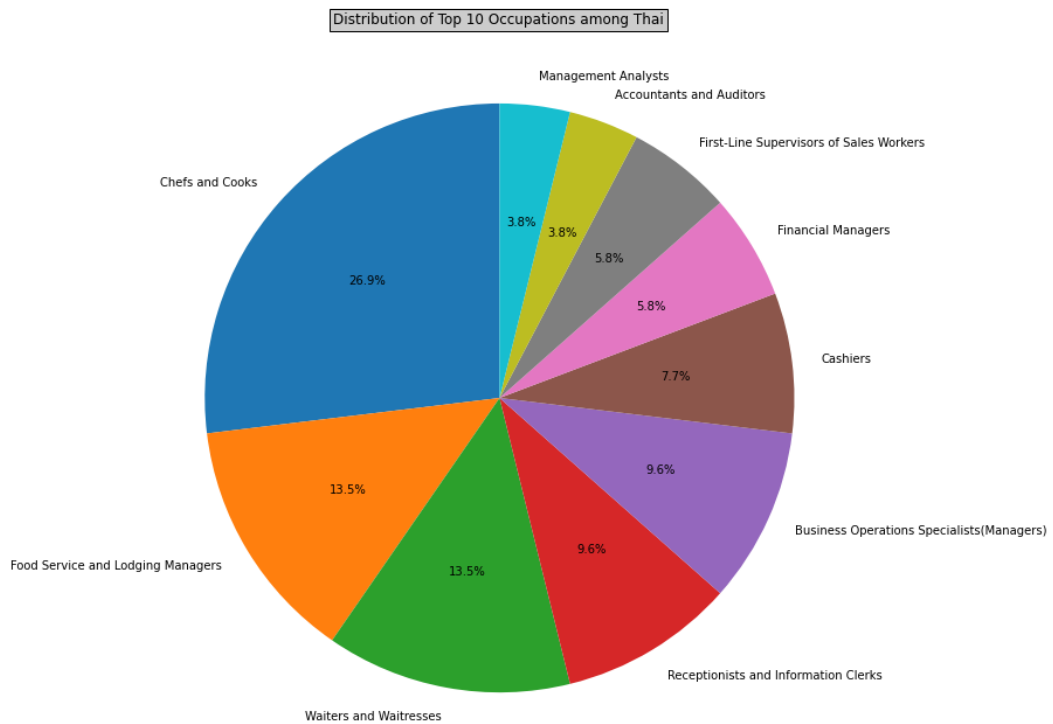
Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Korean')



### *Distribution of Top 10 Occupations among Thai*

There are 52 people work as top 10 occupations, about 0.45614035087719296 of the total Thai Laborforce

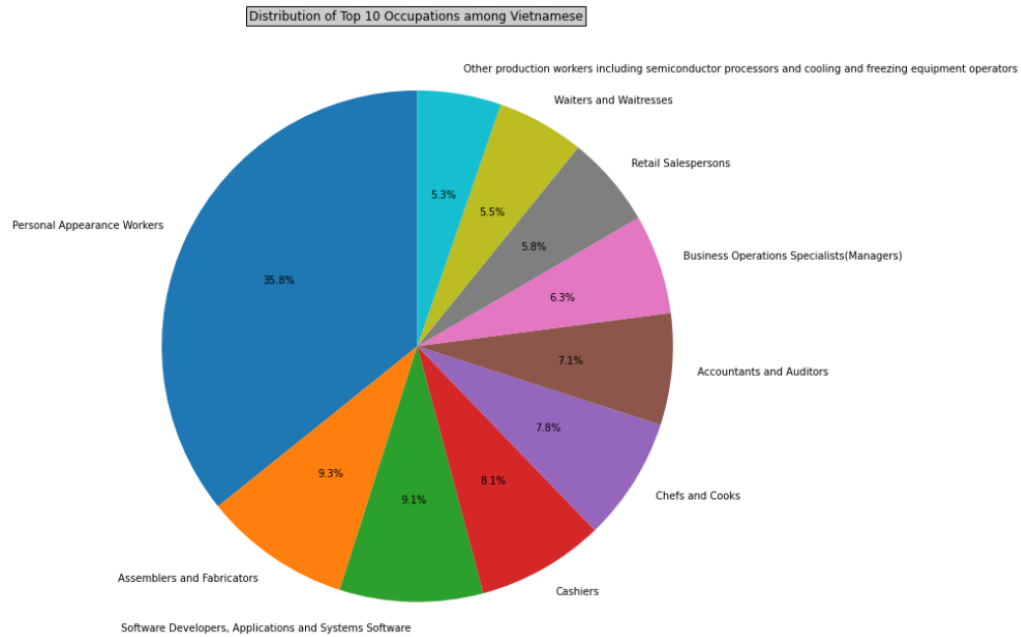
Text(0.5, 1.08, 'Distribution of Top 10 Occupations among Thai')



## Distribution of Top 10 Occupations among Vietnamese

There are 397 people work as top 10 occupations, about 0.3077519379844961 of the total Vietnamese Laborforce

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

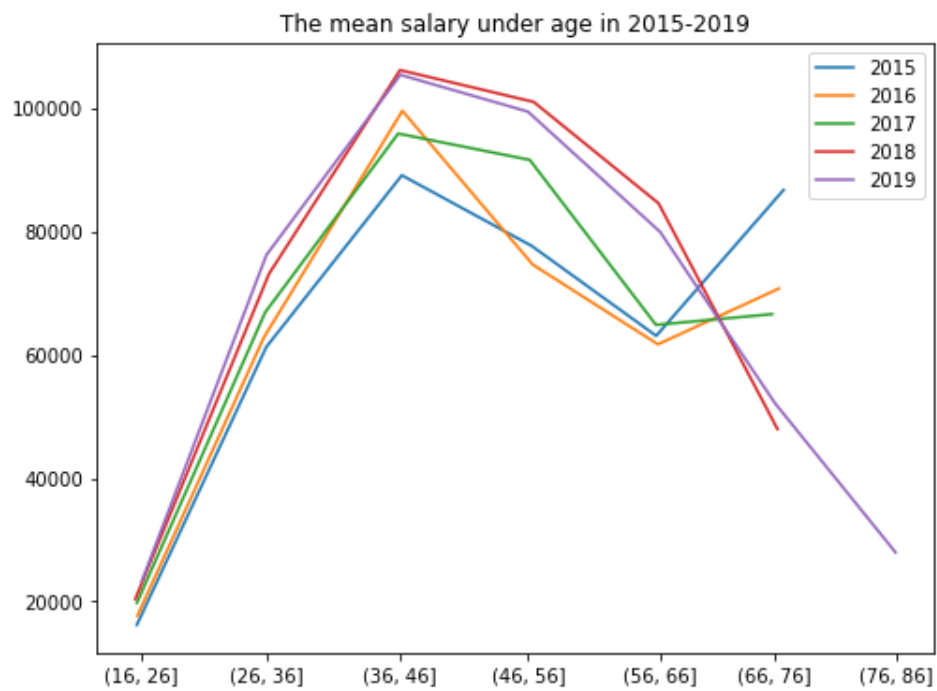


## List of Generalized Occupations

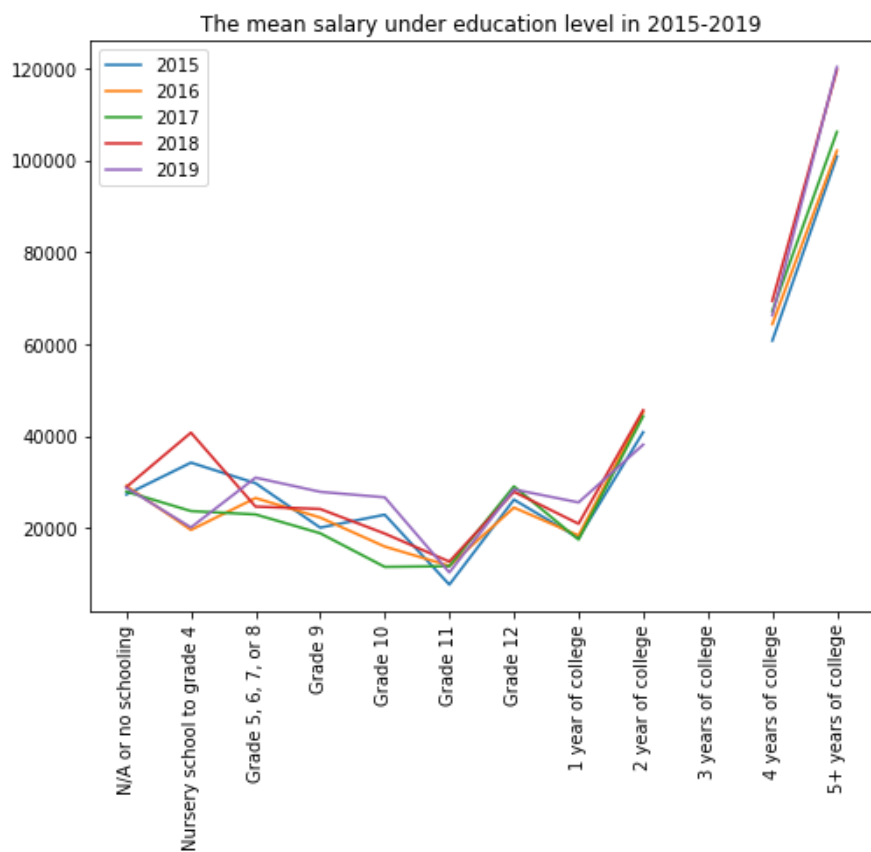
IT Professionals	2301
Managers	1409
Healthcare Professionals	1207
Office and Administrative Support	1174
Education Professionals	1126
Sales	944
Researchers & Scientists	907
Business Operators	876
Food Service Providers	807
Other Service Providers	763
Production Workers	761
Social Service Providers & Military	339
Art & Media	275
Transportation	247
Consturction	103
Agriculture	4

## Correlation between salaries and some features

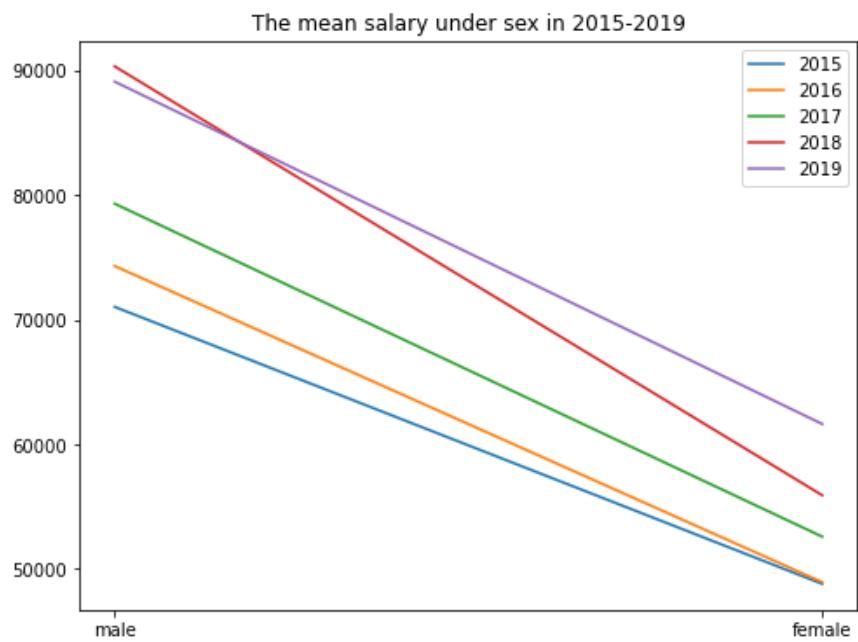
*Mean salary under age in 2015 - 2019*



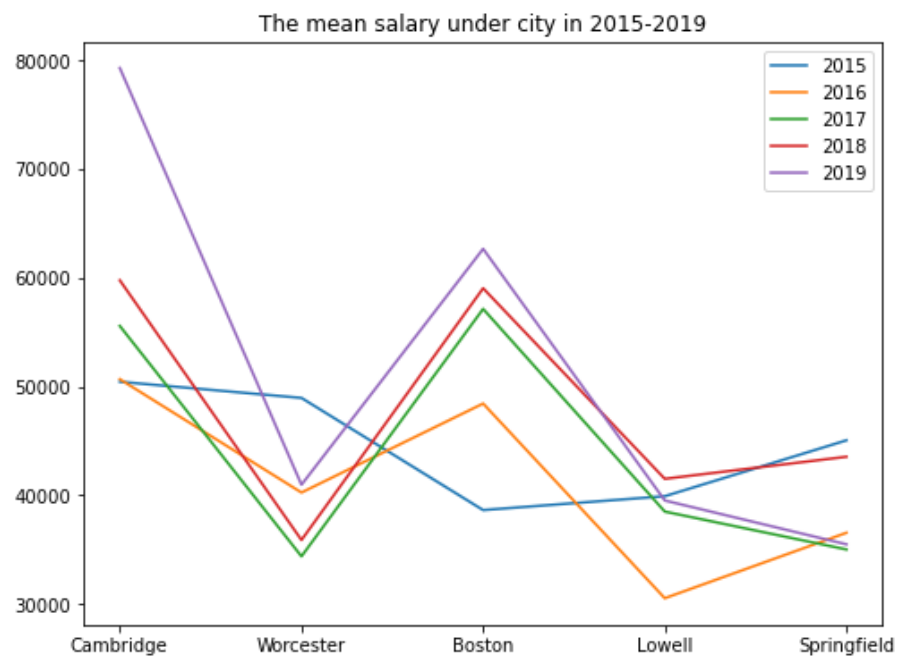
*Mean salary under education level in 2015 - 2019*



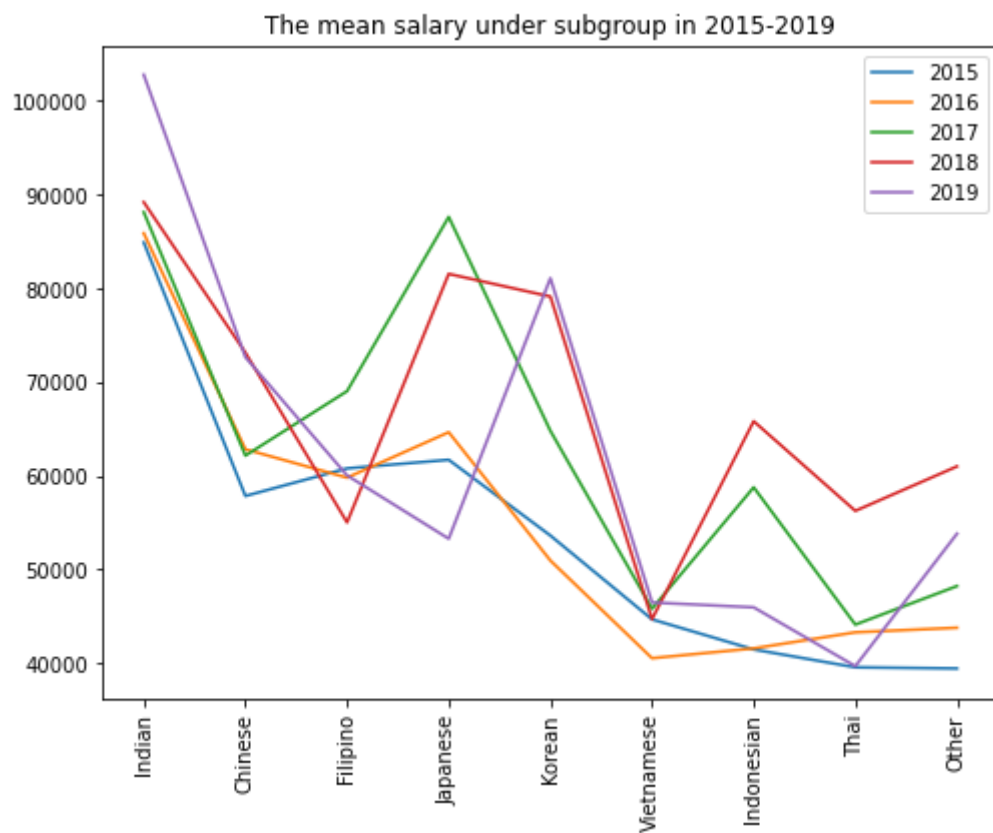
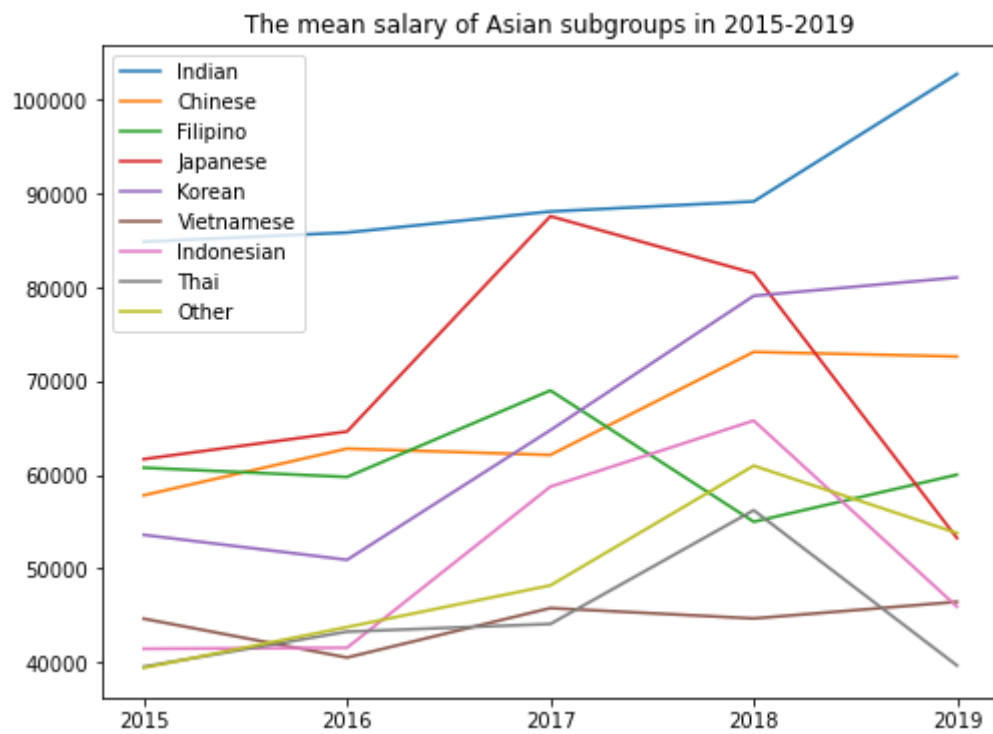
*Mean salary under sex in 2015 - 2019*



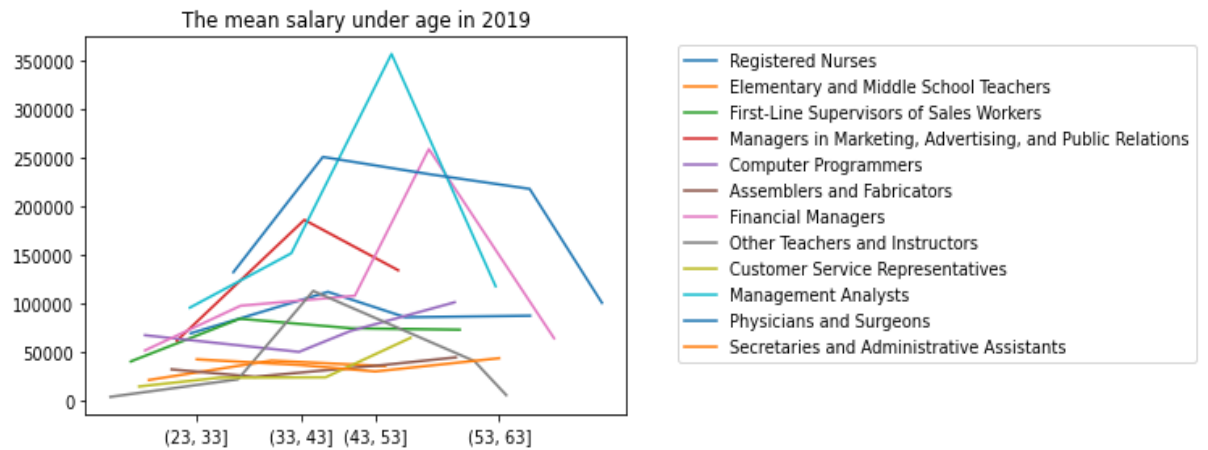
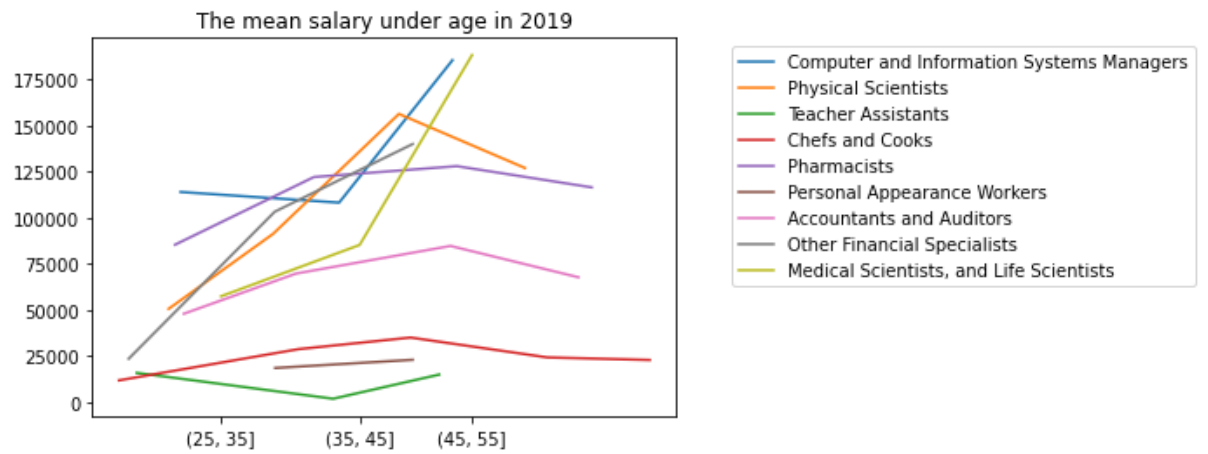
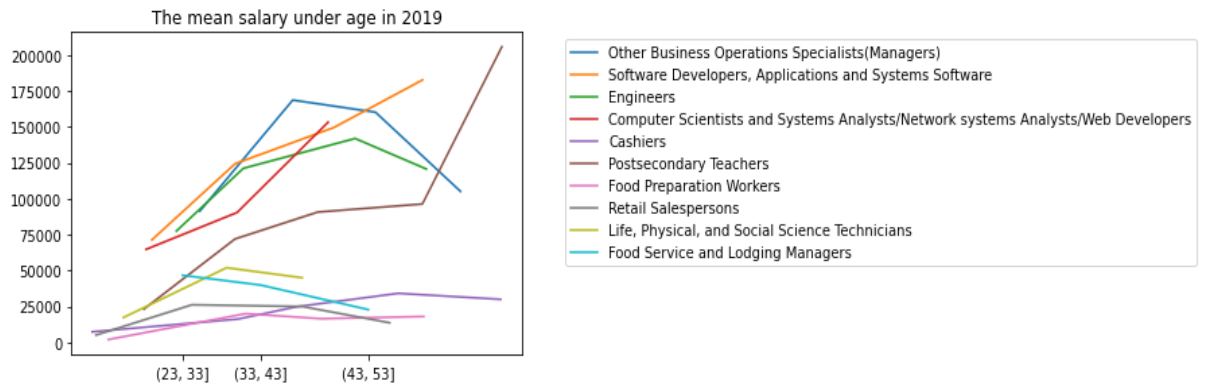
*Mean salary under city in 2015 - 2019*



# *Mean salary of Asian subgroups in 2015 - 2019*

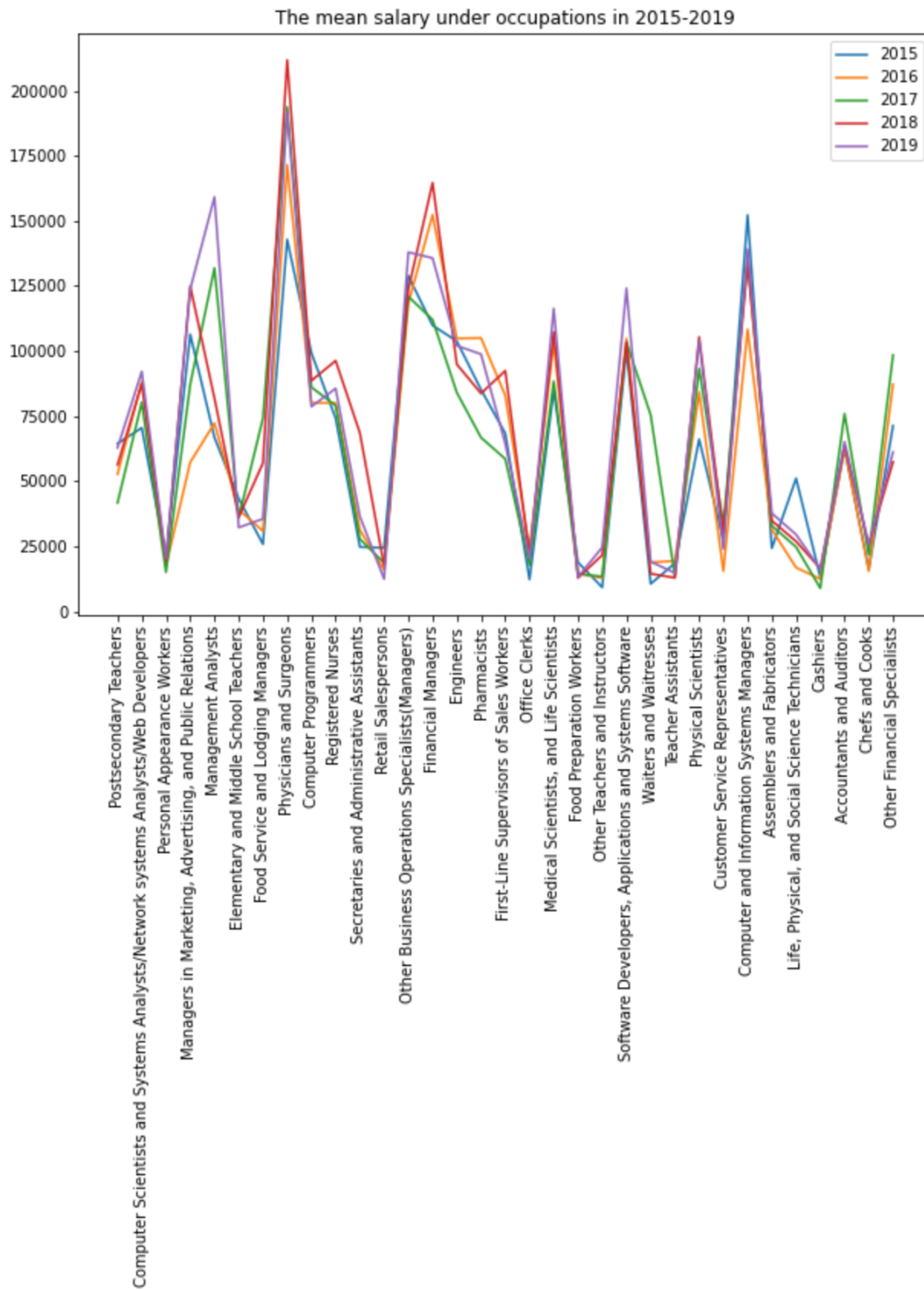


## Mean salary of ages under different occupations in 2019





## *Mean salary under occupations*



## Work Cited

Barroso, A., & Brown, A. (2021, May 25). Gender pay gap in U.S. held steady in 2020. Retrieved November 30, 2021, from

<https://www.pewresearch.org/fact-tank/2021/05/25/gender-pay-gap-facts/>

Basu, S. (2017, March). Wage assimilation of immigrants: A comparison of ‘new’ and ‘old’ Asian source countries. Retrieved November 30, 2021, from

<https://academic.oup.com/migration/article/5/1/1/2566844>

Cohn, D., Livingston, G., & Wang, W. (2014, April 8). Stay-at-home moms by Demographic Group. Retrieved November 30, 2021, from

<https://www.pewresearch.org/social-trends/2014/04/08/chapter-2-stay-at-home-mothers-by-demographic-group/>

Financial self-sufficiency: Median Household Income, by Race/Ethnicity, 2015-19. (n.d.). Retrieved November 30, 2021, from

<https://impactsexcounty.org/financial-self-sufficiency/median-household-income-by-race-ethnicity>

Hinrichs, P. (2015, March 31). Racial and ethnic differences in college major choice. Retrieved November 30, 2021, from

<https://www.clevelandfed.org/newsroom-and-events/publications/economic-trends/2015-economic-trends/et-20150331-racial-and-ethnic-differences-in-college-major-choice.aspx>

IPUMS. (n.d.). University of Minnesota. Retrieved November 30, 2021, from

<https://www.ipums.org/>

Morris, R. (2017, October 15). A detailed look at college major distributions by Race. Retrieved November 30, 2021, from

<https://www.zippia.com/advice/college-major-distributions-by-race/>

Staff, N. (2011, May 24). Study: College majors are highly segregated by race, gender. Retrieved November 30, 2021, from

<https://newsone.com/1252915/college-majors-highly-segregated-by-race-and-gender/>

Statista. (2011, May 24). U.S. college majors with the highest participation of Asian students 2009. Retrieved November 30, 2021, from

<https://www.statista.com/statistics/225726/college-majors-with-the-highest-participation-of-asian-students-in-the-united-states/>

U.S. Bureau of Labor Statistics. (2021, April 21). Earnings and unemployment rates by educational attainment, 2020. Retrieved November 30, 2021, from <https://www.bls.gov/emp/chart-unemployment-earnings-education.htm>

Pew Research Center. (2016, July 1). Racial, gender wage gaps persist in U.S. despite some progress. Retrieved November 30, 2021, from <https://www.pewresearch.org/fact-tank/2016/07/01/racial-gender-wage-gaps-persist-in-u-s-despite-some-progress/>

The Washington Post. (2017, April). The real secret to Asian American success was not education. Retrieved November 30, 2021, from <https://www.washingtonpost.com/news/wonk/wp/2016/11/19/the-real-secret-to-asian-american-success-was-not-education/>

QUARTZ. (2014, October 23). Indians in the US make the most because they studied the most. Retrieved November 30, 2021, from <https://qz.com/285610/indians-in-the-us-make-the-most-because-they-studied-the-most/>