

Deliverable 1

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

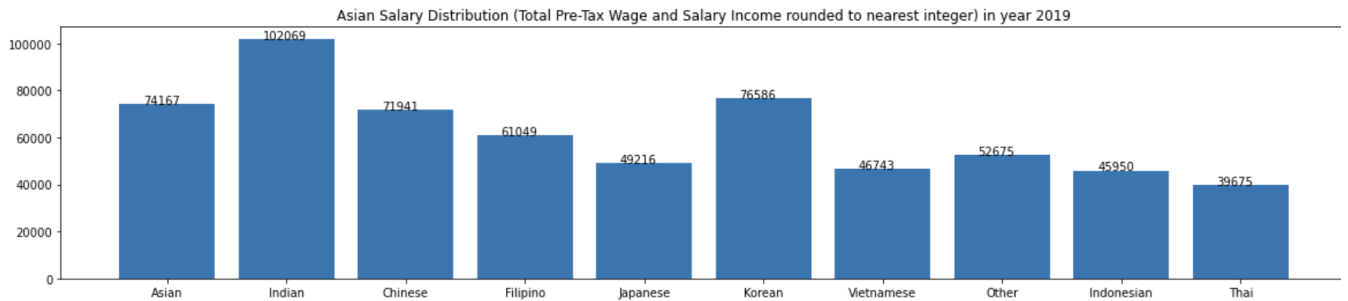
- ☐ Are there any relations between the employment and education of Asians?
- ☐ What are the popular career choices and the trend within Asians?
- ☒ ~~What is the salary distribution of Asians? (Answered)~~
Mainly based on ACS data. (see appendix-salary distribution)
- ☐ What is the main difference between Asians and Whites in these aspects?
- ☐ Which career choices in Asians have the highest earnings?
- ☐ Are there differences between different groups of Asians?
- ☐ Are there correlations between sex and income?

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

