Boston Chinatown
Neighborhood Center
波士頓華埠社區中心

Analysis of Poverty in the Asian Community

Team 2
Team Members: Vivek Unnikrishnan, Danny Xu, Boyu (Tom) Yu, Nathan Vollertsen, Dexter Yanzheng Wu

Table of Contents

Description of Project

There is a lack of access and understanding of Asian data around the social determinants and outcomes of health. The vision of the project is to make the related data using the recent Census, American Community Survey data, and other resources more available to nonprofits, organizers, policymakers, elected officials, academics, and average citizens.
Our project will be based on using IPUMS data which uses ACS data from the Census. Our main focus will be on the poverty statistics in the Asian community. We want to answer a few relevant questions regarding the poverty statistics using the data analysis we will be doing.

Goals of our team

- Collect data from data sources pertaining to poverty levels, including children living in poverty
- Separate out Asian demographic data from data, importing into Pandas/Data analysis framework of choice.
- Find the common key to merge datasets if possible and find trends in poverty levels among the Asian community

Completed Deliverables

1. Answered all the questions that were proposed at the start of the project.
2. Collected and preprocessed ACS data with a special focus on the poverty aspect of the data
3. Collected FDIC data for banking sources to make specific correlations regarding having accounts and poverty levels.
4. Create a website for showing comparisons between 2019 (pre covid) and 2020 (covid) data grouped by different parameters for either the Massachusetts region or for the entire United States region. The website can be accessed at: https://census-poverty-explorer.herokuapp.com/

Preliminary Analysis

ACS Data:

Data point	**Value**
Number of rows in the data set	5 million
Number of Asians (2020 Data)	184197
Number of Asians in Poverty (Income to Poverty Threshold less than 100%)	23283
Number of Asians with no Medicare access	4888
Number of Asians with no public insurance	4562

Number of Asians in poverty with Food stamp	4416

FDIC data:

Data point	**Value**
Number of rows in the data set	70203
Number of Asians (2020 Data)	2164
Bank Account type unknown	1750

Dataset view:

YEAR	SAMPLE	SERIAL	CBSERIAL	HHWT	CLUSTER	REGION	STATEICP	COUNTYICP	STRATA	...	EMPSTATD	INCTOT	FTOTINC	INCWAGE	INCSS	INCWELFR	INCSUPP	INCOTHER	POVERTY	ERSCOR50
2020	202001	4	2020010000189	74.0	2020000000041	32	Alabama	1170	120001	...	Not in Labor Force	11100	9999999	0	0	0	11100	0	0	999.9
2020	202001	40	2020010003115	23.0	2020000000401	32	Alabama	0	160001	...	Not in Labor Force	7000	9999999	7000	0	0	0	0	0	25.2
2020	202001	90	2020010006855	53.0	2020000000901	32	Alabama	550	90001	...	Not in Labor Force	1500	9999999	1500	0	0	0	0	0	75.2
2020	202001	148	2020010010714	67.0	2020000001481	32	Alabama	0	250001	...	Armed forces--at work	21000	9999999	21000	0	0	0	0	0	999.9
2020	202001	179	2020010012630	46.0	2020000001791	32	Alabama	0	30201	...	Not in Labor Force	160	9999999	160	0	0	0	0	0	10.0
2020	202001	237	2020010017141	127.0	2020000002371	32	Alabama	0	250001	...	Armed forces--at work	21000	9999999	21000	0	0	0	0	0	999.9
2020	202001	246	2020010018293	48.0	2020000002461	32	Alabama	0	200001	...	Not in Labor Force	0	9999999	0	0	0	0	0	0	999.9
2020	202001	257	2020010019097	63.0	2020000002571	32	Alabama	0	230001	...	Not in Labor Force	11100	9999999	0	0	0	11100	0	0	999.9
2020	202001	331	2020010025214	12.0	2020000003311	32	Alabama	0	200001	...	Armed forces--at work	20000	9999999	20000	0	0	0	0	0	999.9
2020	202001	349	2020010026262	52.0	2020000003491	32	Alabama	0	30101	...	At work	8300	9999999	8300	0	0	0	0	0	27.9

Datasets used

- IPUMS Data ACS/CPS data with selected attributes
- FDIC Data for getting details about banking sources

Questions Answered

What is the poverty level amongst Asian communities and what is the breakdown by age, gender and other filters?

Filter **Graph**

Sex

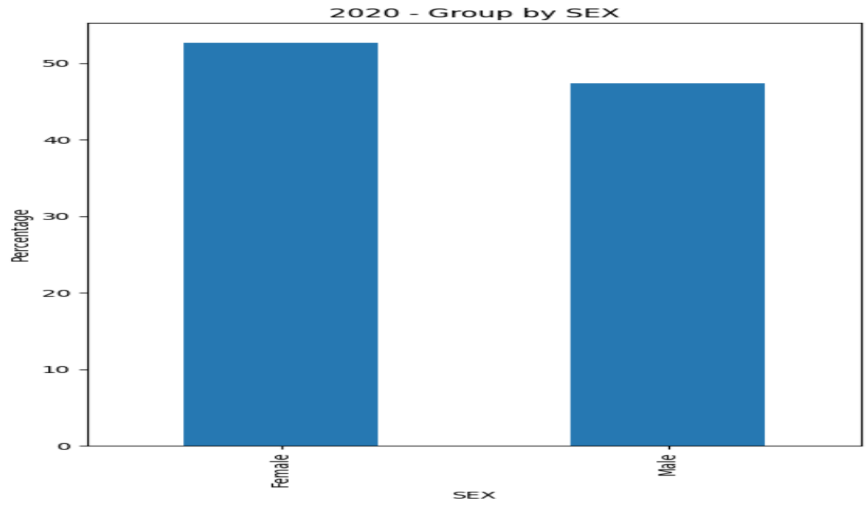

Fig 1: Percentage of Asians per gender in poverty in United States
Data:
Female: 52.64
Male: 47.36

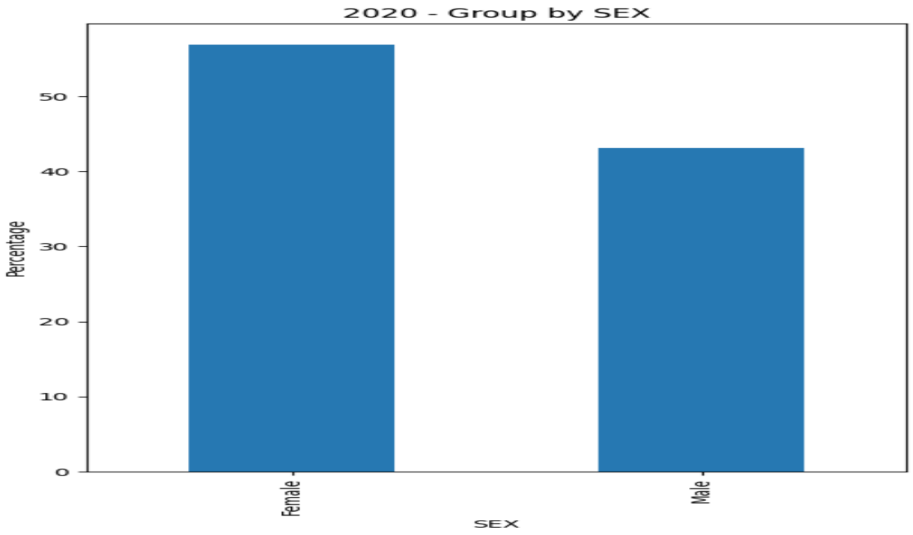

Fig 2: Percentage of Asians per gender in Poverty in Massachusetts
Data:
Female: 56.86
Male: 43.13

Age

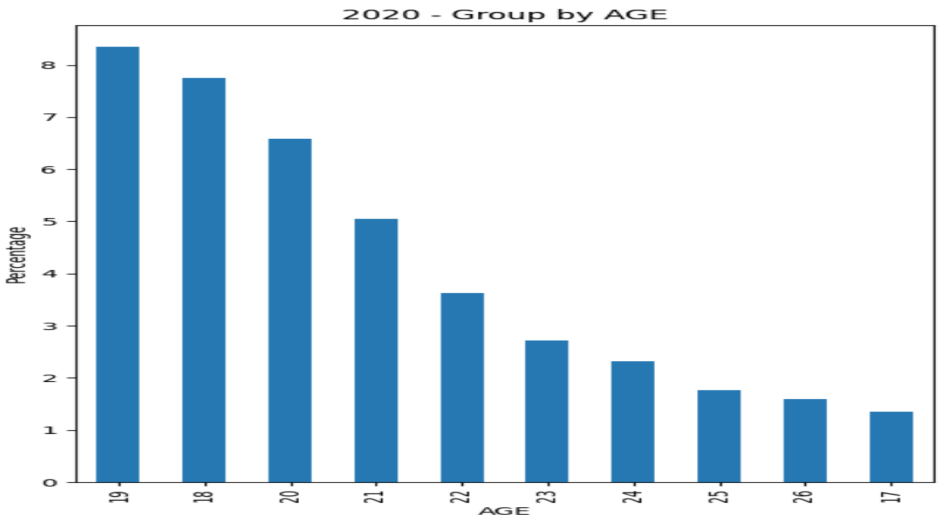

Fig 3: Percentage of Asians per age group in Poverty in United States

Data:

19	8.36
18	7.76
20	6.58

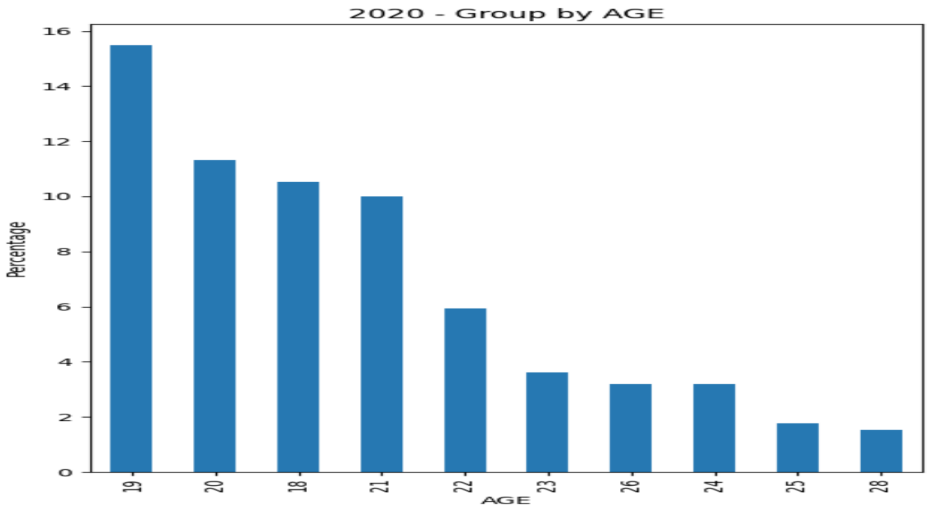

Fig 4: Percentage of Asians per age group in Poverty in Massachusetts

Data:

19	15.48
20	11.31
18	10.54

Communities

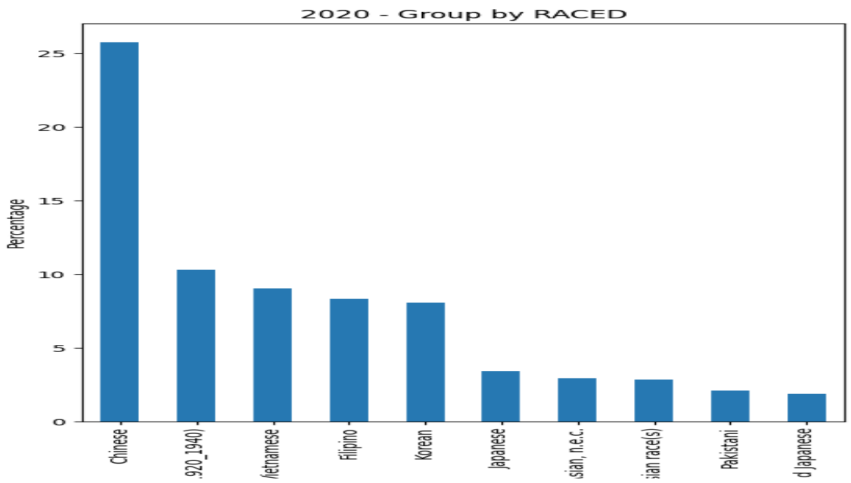

Fig 5: Percentage of Asians per community in Poverty in United States

Data:

Chinese25.73
Asian Indian 10.29
Vietnamese 9.02

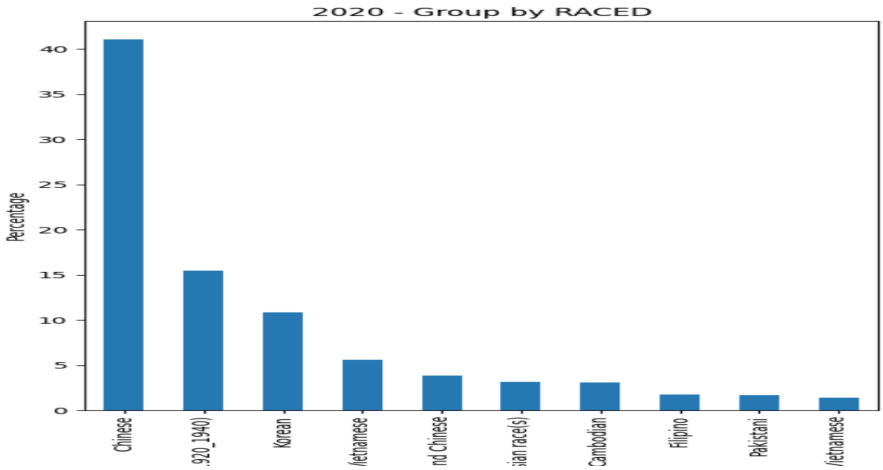

Fig 6: Percentage of Asians per community in Poverty in Massachusetts

Data:

Chinese41.05
Asian Indian 15.48
Korean 10.87

Answer: Here from these three sections we can see the relationship between the gender, age, and poverty level among the many different groups of Asian people. We see that 19 is the age where poverty is highest of all ages, that females are in more poverty than males, and that the Chinese are the largest group of Asians with poverty.

What percentage of them are part of community help programs, food assistance and financial service programs?

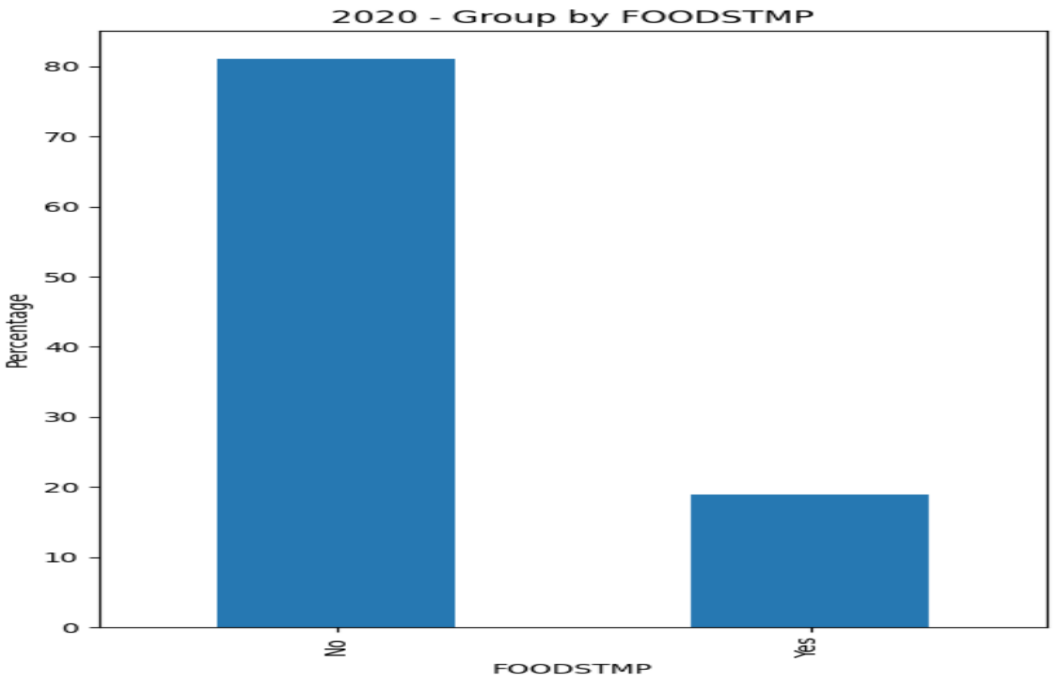

Fig 7: Food Stamp Data for Asians in poverty in the United States

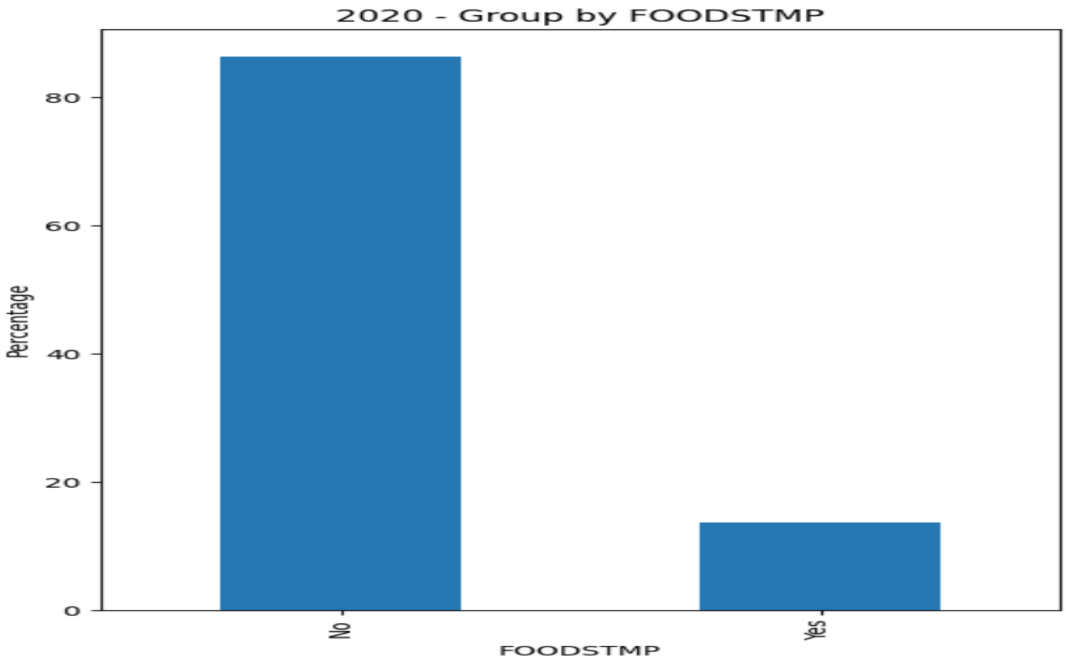

Fig 8: Food Stamp Data for Asians in poverty in Massachusetts

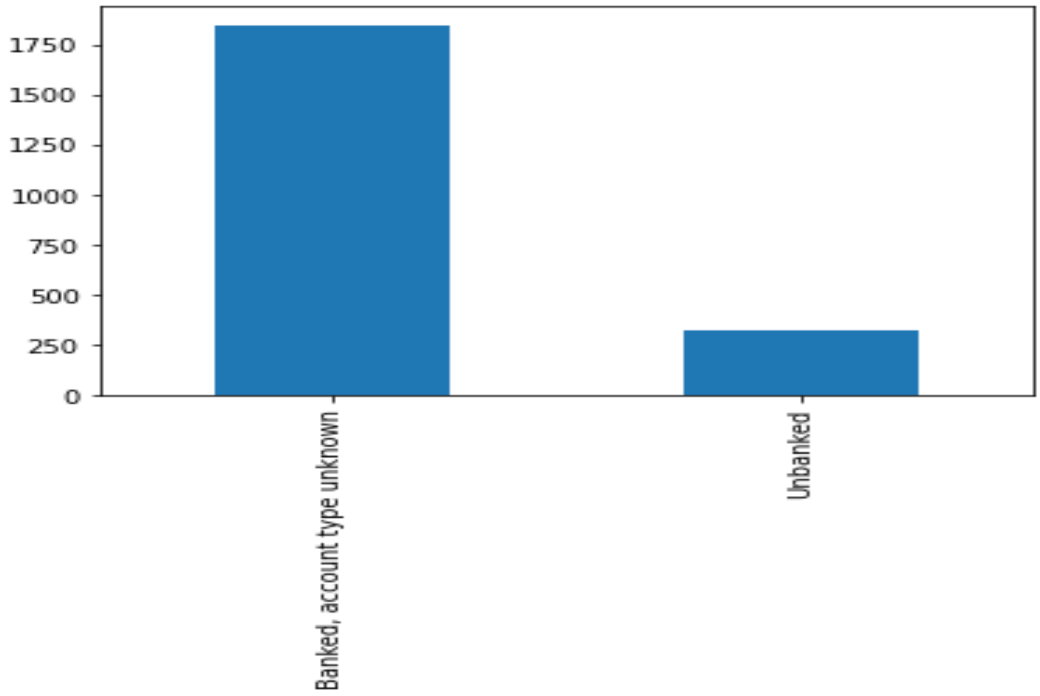

Fig 9: Bank Account type for Asians in Poverty

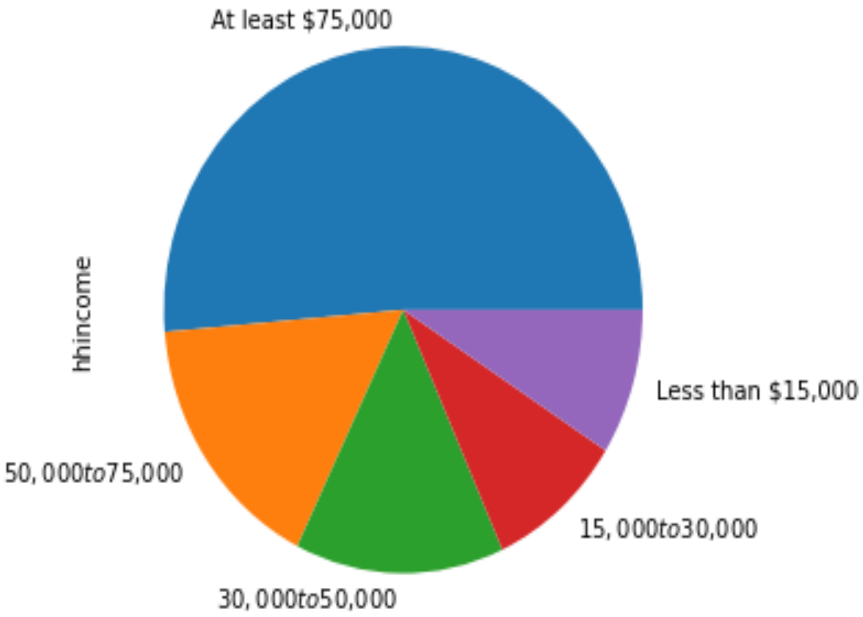

Fig 10: Income for Asians in the United States

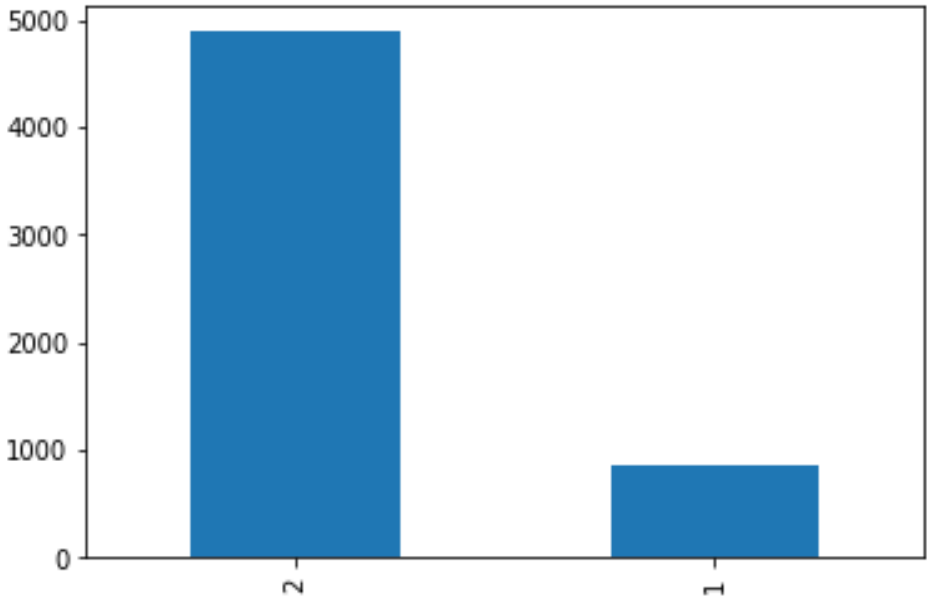

Fig 11: Public health insurance coverage for Asians in the United States ("2" stands for no, "1" stands for yes)

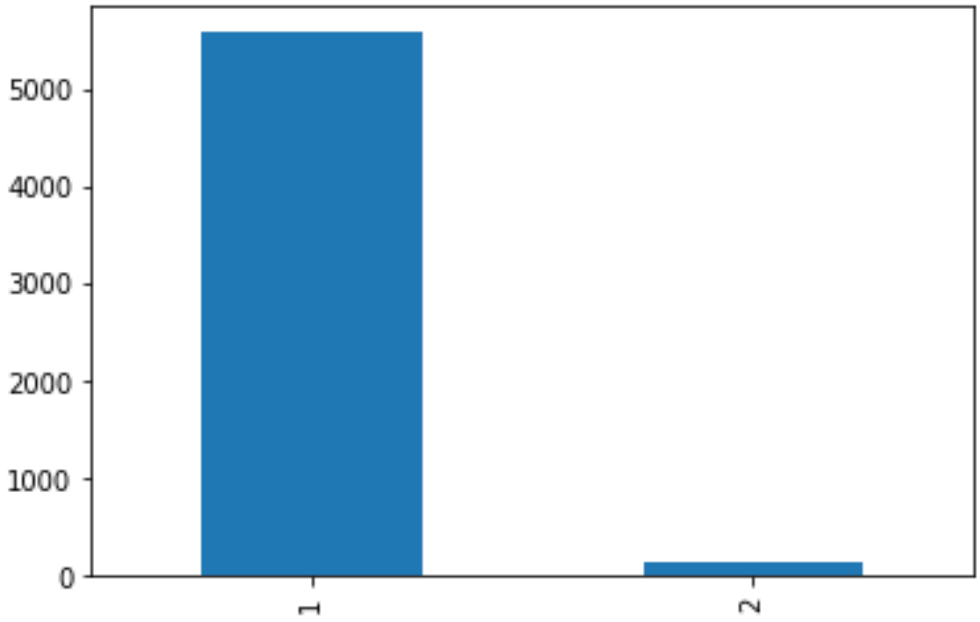

Fig 12: Private health insurance coverage for Asians in the United States ("2" stands for no, "1" stands for yes)

Answer: The above figures help to answer this question. The health insurance data between public and private gives an interesting finding of how public insurances are less for the Asian population. The food stamp data as well as the community help graph help to get similar inferences.

Are there any problems with the living conditions of Asians reflected by these datasets?

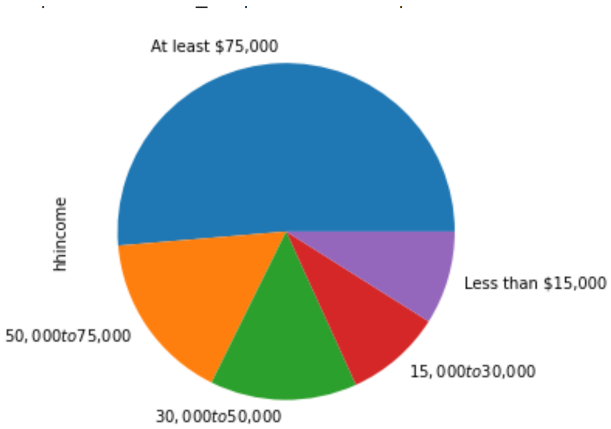

Fig 12: Income for Asians in the United States

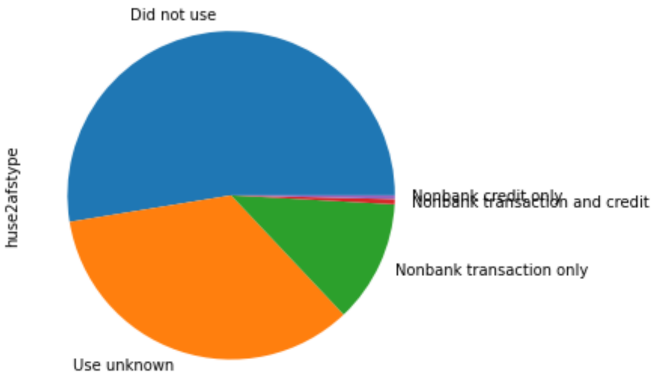

Fig 13: Types of banking transactions done by Asians in the United States

Answer: From the first pie chart we can observe that at least a quarter of the Asian population are falling under the 2022 federal poverty level. The second pie chart shows that there is almost more than half of the Asian population that do not use any bank/unbank service whatsoever. Gathering all the information provided by these plots, we can conclude that roughly half of the Asian citizens in America do not use any bank/unbank service and roughly one-fourth of the Asian individuals are living in poverty.

Are there differences between different groups of Asians?

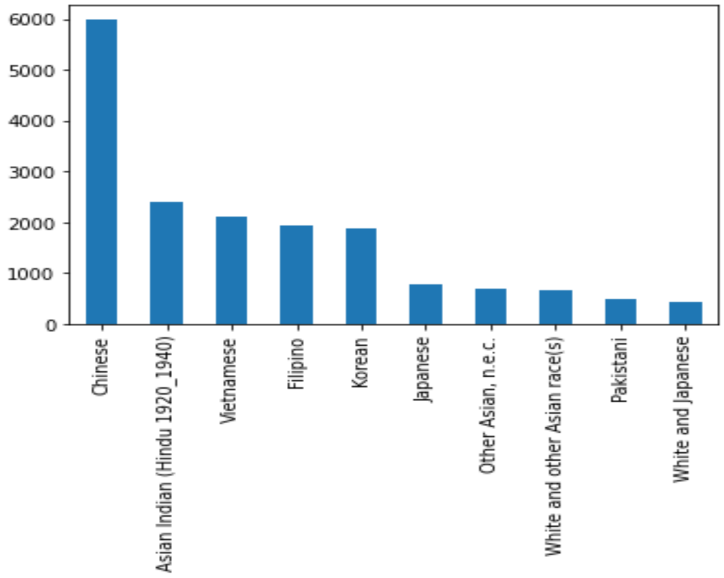

Fig 14: Count of Asians in poverty in the United States (Count vs Community)

Answer: Yes, there are differences between the group of Asians. From the graphs above we can see that overall there are more Chinese than any other Asian group. This is true for other data sources we found as well. The data is skewed due to this, but the differences are clear

How does COVID-19 impact income and different Asian groups from pre-COVID to the present time?

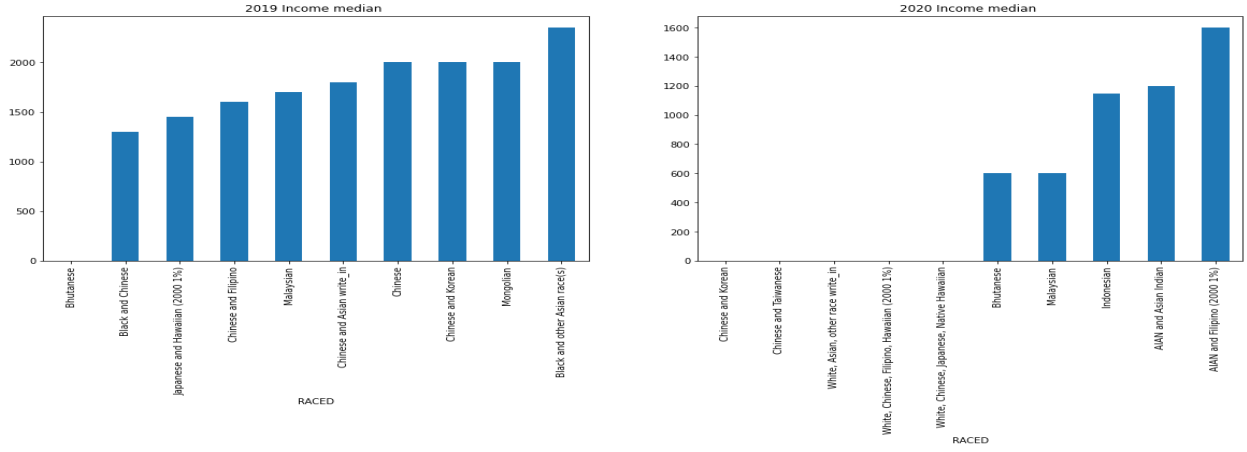

Fig 15: Pre covid and covid data for the median income for different Asian communities

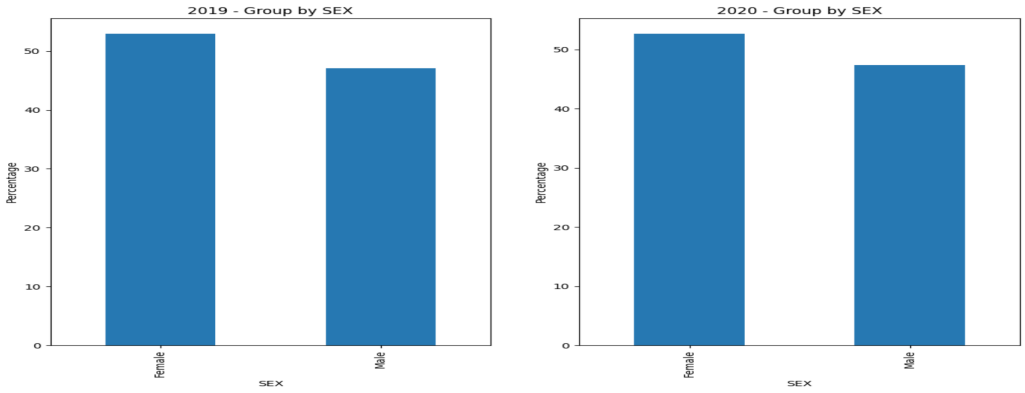

Fig 16: Pre covid and covid data for the percentage of Asians grouped by Sex

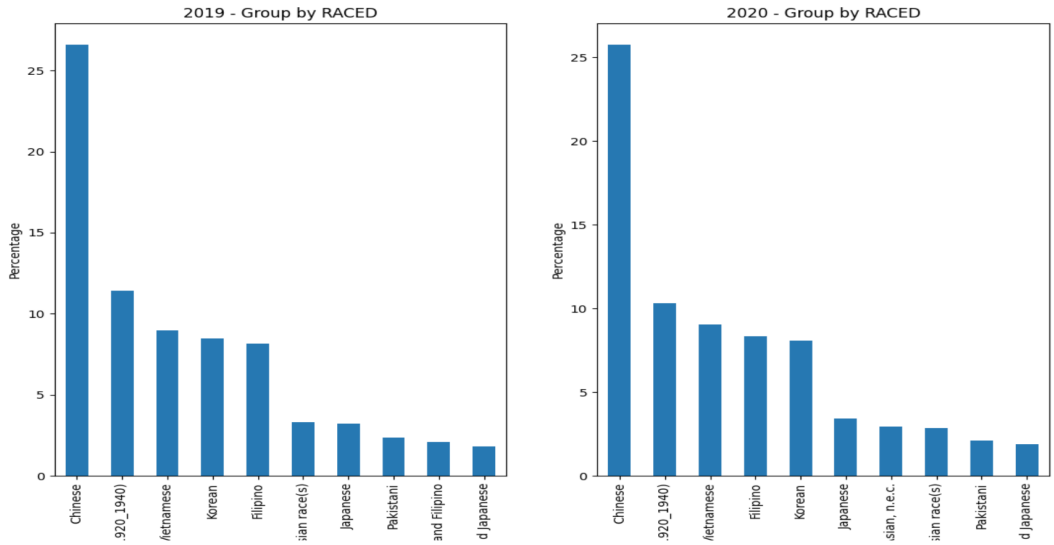

Fig 17: Pre covid and covid data for the percentage of Asians grouped by community

Answer: We compared the percentage of Asians group by different factors pre Covid and during Covid. We compared median total personal income and family income between 2019 and 2020 for Asians. We found that the personal income of people in poverty decreased by ~7% (from 42050 to 42000) while the personal income decreased by ~0.1% (negligible). This indicates that COVID 19 has decreased the personal income from the previous year but for people not in poverty the effect cannot be seen and for people in poverty we can see that COVID has further decreased their income level. We also see that compared to 2019, in 2020 the median income has decreased for many groups. For example, for Asian Indians, we see there is a ~16% decrease while for Chinese there is a 10% decrease in income.

Is there a relationship between Asians in a different neighbourhood and their poverty level, if so what are the different relationships?

Answer: From the dataset that we collected and analyzed we are unable to determine if there is a relationship between the different neighbourhoods because we don't have the data of who lives in what neighbourhood. Since we don't have the residence of where these people live in we can't answer this and it is limited.

Limitations and Potential Risks

From Deliverable 1

- While the Census data we looked at provided important information, some of the data are already dated too long to the present time (2022), and with COVID-19 impacts, data on income can change drastically because of the current situation.
- The dataset is skewed with the majority of the data contributed by Chinese and Indians. This would make it difficult to give broad ideas(Data processing such as box cox transformation to the race population column might be helpful to create a normalized distribution, which could be beneficial for predictions and classifications).
- Data regarding access to financial systems is currently not available and we are still searching for the same(Might need to consider the financial aspect from other public systems due to the limitations).
- The data set is encoded so for readability purposes clean up needed to be done.
- The current dataset looks small but the data is cleaned and ready to use(For the purpose of expanding the columns of features, we could work some presumptions such as linearity of the dataset and adding features like the production of two original columns in order to gain more a more generalized sense of the relationship among the features, similar to the deep learning dataset manipulations).
- In order to build a non-biased model for the purpose of better understanding the response value, the Poverty values, adjusted weights for the *Race* class might be introduced and work on calculating the best-fit weights would be challenging.

From Deliverable 2

- FDIC data is smaller compared to census data with a lot of data missing or unavailable for Asian households. This causes analysis to be difficult
- Community program data is limited and currently, only government-run community program data could be found which might not provide a good picture of how it affects Asian households.
- FDIC data needs to be cleaned to understand due to the heavy encoding and its raw nature of it.

From Deliverable 3

- 2020 Data has discrepancies mainly due to the data collection issues by the census due to the COVID 19 pandemic

- Due to the above limitation, we cannot fully guarantee the accuracy of the analysis even though our results are close to Census estimates provided by the Government.

Website

We created a website that would help play around and derive conclusions from the Census data for Asians in poverty by selecting 2 filters:

1. Use data of the entire United States or only Massachusetts.
2. The column to group the data by and provide a normalized graph

This website helps to understand the relationship of data in Massachusetts or the United States and how it is affected across different parameters like Sex, age or food stamp data. This is based on the discussion with the client about having an easy interface to understand and derive results from the data

URL: https://census-poverty-explorer.herokuapp.com/

Work distribution

Here are the cards we worked on during the course of the project

Epics

Type	Key	Summary	Assignee	Reporter	P	Status
	BT2-23	Final deliverable	Unassigned	Vivek Unnikrishn...	=	DONE
	BT2-17	Deliverable 3	Unassigned	Vivek Unnikrishn...	=	DONE
	BT2-13	Complete Mid Term Presentation	Unassigned	Vivek Unnikrishn...	=	DONE
	BT2-7	Completing Deliverable 2	Unassigned	Vivek Unnikrishn...	=	DONE
	BT2-1	Deliverable 0 and Deliverable 1	Unassigned	Vivek Unnikrishn...	=	DONE

Vivek Unnikrishnan

Type	Key	Summary	Assignee	Reporter	P	Status
	BT2-24	Complete simple flask application	Vivek Unnikrishn...	Vivek Unnikrishn...	=	DONE
	BT2-18	Analyse Income Data	Vivek Unnikrishn...	Vivek Unnikrishn...	=	DONE
	BT2-11	Complete coding on the Jupyter Notebook	Vivek Unnikrishn...	Vivek Unnikrishn...	=	DONE
	BT2-6	Deliverable-1 Document	Vivek Unnikrishn...	Vivek Unnikrishn...	=	DONE
	BT2-3	Complete Deliverable 0 Document	Vivek Unnikrishn...	Vivek Unnikrishn...	=	DONE

Nathan

Type	Key	Summary	Assignee	Reporter	P	Status
	BT2-26	Answer remaining questions	Nathan Vollertsen	Vivek Unnikrishn...	=	DONE
	BT2-20	Preprocess Income Data	Nathan Vollertsen	Vivek Unnikrishn...	=	DONE
	BT2-8	Checking into Financial Data	Nathan Vollertsen	Vivek Unnikrishn...	=	DONE
	BT2-4	Track issues with the dataset	Nathan Vollertsen	Vivek Unnikrishn...	=	DONE

Dexter

Type	Key	Summary	Assignee	Reporter	P	Status
	BT2-22	Preprocess Education Data	Dexter Yanzheng...	Vivek Unnikrishn...	=	DONE
	BT2-15	Slide 2	Dexter Yanzheng...	Vivek Unnikrishn...	=	DONE
	BT2-9	Checking into Food Data	Dexter Yanzheng...	Vivek Unnikrishn...	=	DONE
	BT2-2	Read about Census Dataset	Dexter Yanzheng...	Vivek Unnikrishn...	=	DONE

Danny

Type	Key	Summary	Assignee	Reporter	P	Status
☑	BT2-25	Complete document	D dannyxu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-16	Slide 3	D dannyxu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-10	Checking into Help programs	D dannyxu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-5	Create Jupyter Notebook	D dannyxu	VU Vivek Unnikrishn...	=	DONE

Boyu (Tom) Yu

Type	Key	Summary	Assignee	Reporter	P	Status
☑	BT2-21	Collect Education data	BY Bo Yu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-19	Collect income data	BY Bo Yu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-14	Slide 1	BY Bo Yu	VU Vivek Unnikrishn...	=	DONE
☑	BT2-12	Create Document and add details	BY Bo Yu	VU Vivek Unnikrishn...	=	DONE