

SCHMIDT FUTURES ASSIGNMENT

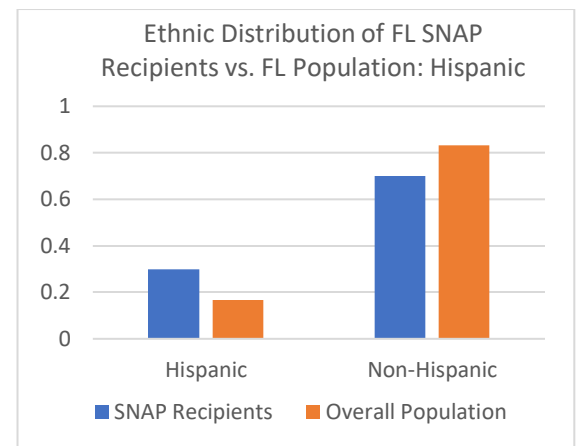
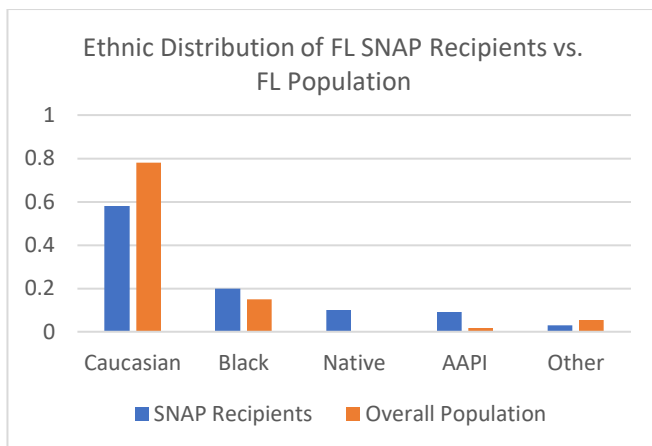
PART 1

INTRODUCTION

The provided data includes 20000 randomly-selected Florida SNAP recipients. It is broken down by ethnicity, gender, county, employment status, primary language, household size, education, and age. Of these, I predict that ethnicity, gender, and language will have the largest effect on claiming SNAP benefits. I will analyze the data for all categories *except* employment status and education, which do not compare to the general population.

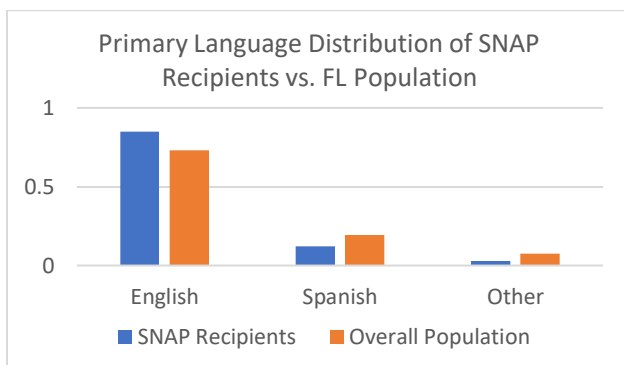
ETHNICITY

In order to analyze this data, I needed to consolidate the given ethnicity labels. I also needed to break out Hispanic as a separate category. I standardized spelling and capitalization, and chose to use the labels “Caucasian”, “Black”, “Native”, “Asian American/Pacific Islander”, and “Other”.



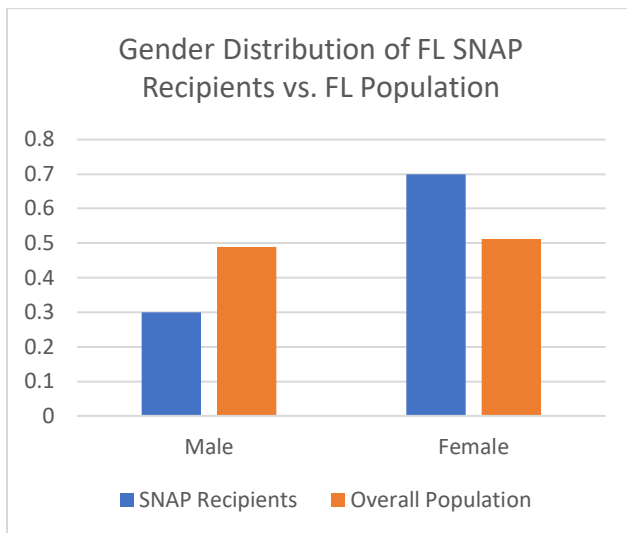
There are a few trends to notice here: Caucasian people represent a larger proportion of the general Floridian population than they do SNAP Recipients. Excluding the “Other” category, all minorities claimed a larger proportion of benefits than their respective proportions of the population, with Native American/Alaska Native-identifying individuals claiming a much larger proportion. Along a similar vein, Hispanics of any race claimed a larger percentage than Non-Hispanics. Given the assumption that the proportion of SNAP-eligible citizens aligns with that of the general population, this data indicates that Caucasians and Non-Hispanics are more frequently leaving their benefits unclaimed. It also indicates that Natives are good about claiming their benefits. Of course, this assumption may be flawed, and a comparison to the ethnicities of the subset of the population who qualifies for benefits would yield a more accurate evaluation. For instance, it is common knowledge that minorities, and natives in particular, are less well-off than the white population in general, which might account for some or all of this result.

PRIMARY LANGUAGE



There were only three language categories present in the SNAP data. Based on the data, non-English speakers, in both the case of Spanish and other primary languages, were underrepresented among SNAP recipients with respect to Florida’s population. Following the hypothesis that SNAP beneficiaries mimic the general population, we conclude that English speakers are more likely to claim their benefits than Spanish or Other speakers. Trying to claim benefits in English and realizing they are even available may be difficult for these speakers, some of whom may not speak English at all, so this result is unsurprising.

GENDER

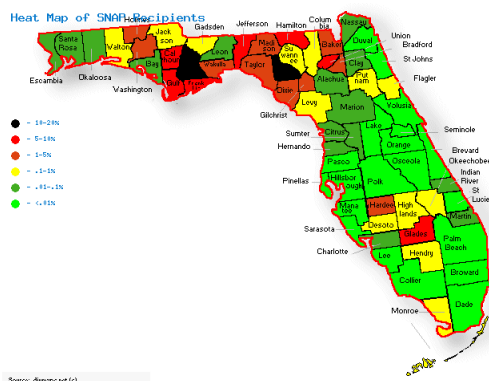
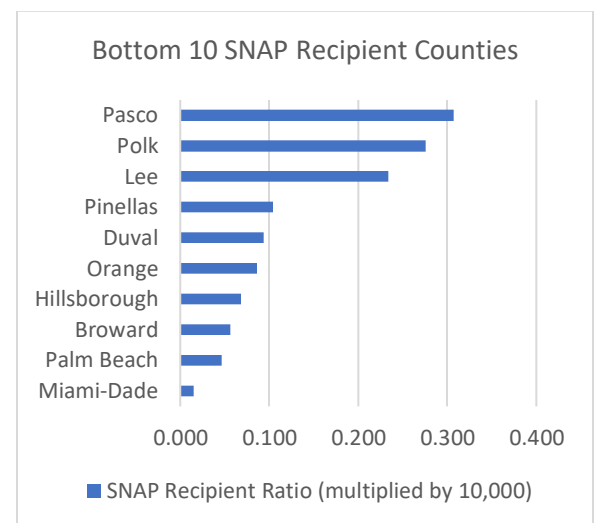
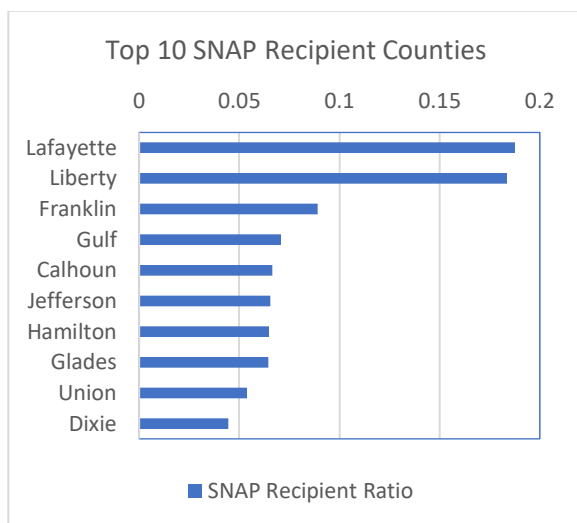


Gender provides a straightforward and interesting comparison point. Females represent a much larger proportion of SNAP Recipients than their respective share of the population. This indicates that women are claiming SNAP benefits at a much higher rate than men, who may be leaving their benefits unclaimed.

A few potential explanations for this result: like with ethnicity, women make less than men overall, so may be more likely to qualify to begin with. Women are more likely to be supporting children on their own, as indicated in Florida's single mother vs. single father statistics, which would lead them to claim SNAP benefits. Finally, societal pressure on men related to masculinity might lead them to be ashamed to claim.

COUNTY

For this analysis, I looked at the proportion of how frequently a county was present in the SNAP data to the population of that county.



From the data, it is clear that Lafayette and Liberty counties are most overrepresented among SNAP Recipients, while Miami-Dade is least. These are also the least and most populous counties in Florida respectively. Looking at the map, it appears that Northern Florida is more reliant on SNAP than Southern Florida, and counties are roughly clustered by how many claim benefits. Assuming roughly equal distribution among the general population and SNAP qualifiers, this would indicate that those in less-populated counties are more comfortable with and/or able to claim SNAP benefits. Once again, however, it is possible that people are simply better-off in densely populated counties.

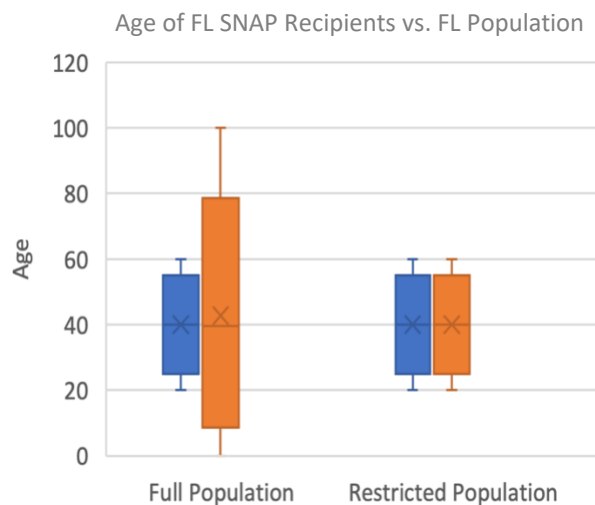
HOUSEHOLD SIZE

The data on household size from the SNAP csv includes sizes of 1-4 members. 4052 values are NaN – over a fifth of the dataset. It is unclear whether this is because the household size was larger than 4 members or just entered in a non-numerical fashion. As such, I conducted two analyses, one ignoring these values and one accounting for them. The data for Florida households is broken down by household types not sizes, so I worked with average size and single-individual households.

	Ignoring NaN Values	Including NaN Values
Average SNAP Recipient Household Size:	1.75	≥ 1.75
Average FL Population Household Size:	2.46	
Percent SNAP Population Household Size of One:	63%	50%
Percent FL Population Household Size of One:	26.6%	

The data suggests that SNAP Recipients are overwhelmingly more likely to belong to a household of just one. This indicates that people may feel less comfortable claiming benefits or be less able to if they have others in their household – partners, children, etc. This could be because those who have partners are already supported and do not need benefits, because people with others in their household would be ashamed to claim SNAP, or because these people, for instance parents, do not have the time to spare to claim benefits.

AGE



The ages of the individuals in the SNAP dataset range from 20 years of age to 60 years of age. It makes sense that no children would be claiming benefits on their own. I found the min, max, median, and 1st and 3rd quartile data from the SNAP dataset and estimated the same values from the Florida demographics site, which gave age ranges. I took the mean of the proper range as the value in each case. I did this analysis first on the general population, and then restricted the age range to the same one as the SNAP dataset, namely 20-60 years of age. Interestingly, there appears to be almost no difference beyond the range of the population. The median age of SNAP recipients is 40 and the mean age of SNAP Recipients is 39.86. These values are more or less comparable to the median age of Florida's population: 38.7. The last consideration is that there does not appear to be an age limit on SNAP benefits, so it is odd that everyone in the data is 60 or younger. Elderly people are underrepresented in the data.

SUMMARY

Based on the assumption that the general Florida population's demographics track with the demographics of those who qualify for SNAP, the data suggests that the following groups are underrepresented: Caucasians, Non-Hispanics, and people who identify as an "Other" ethnicity (beyond Caucasian, Black, Native, or AAPI); people whose primary language is not English; men; people living in densely populated counties; people who share their households with others; and people over the age of 60. I am sure that with more time, interesting observations would arise from intersections of these demographics or from examining education/employment data.

PART 2

"I received a raise last year so based on what I saw a few years ago I am no longer eligible."

This individual falsely believes they are ineligible. They may have outdated information or a faulty memory. A change in income may change eligibility, so people who receive raises need to know whether or not they still qualify. Understanding the **latest eligibility requirements** is a pain point.

"My father is a libertarian so does not believe in government handouts."

This individual is feeling pressured by an outside party not to accept the benefits. Specifically, they are concerned with a political perspective. Facing **social stigma** or **family pressure** is a pain point. Confronting **conflicting political beliefs** is a pain point.

"The last I heard people on social security are not eligible for food stamps because we exceed the maximum pay."

This individual falsely believes they are ineligible. They are going based on hearsay rather than facts. They are confused because they fall into a certain subset already receiving government aid. Dispelling **incorrect rumors** is a pain point. Understanding **eligibility requirements for specific subgroups** is a pain point.

“My grandmother does not have much familiarity with technology and is not able to apply in person right now due to COVID-19.”

This person’s grandmother is unsure of how to apply when she cannot follow the same procedure she always has. She is not well-versed in technology. Learning **alternate paths** to applying is a pain point. **Using technology** is a pain point.

“I receive decent pay as a nurse but it does not stretch far given I have seven kids.”

This individual incorrectly assumes that because their base salary is not traditionally low, they do not qualify. They do not realize that household size is taken into account when looking at eligibility. Understanding that people who fall out of the **stereotypical SNAP beneficiary** could qualify is a pain point.

“My son doesn’t have time to apply for me.”

This person believes the process is too complicated to navigate without help. Complexity, or **perceived complexity** of the application process is a pain point.

“What is SNAP? I don’t think I’ve heard of it before. Maybe a different name?”

This person does not know the acronym for the food stamps program. Using a **nickname** that does not imply the program’s purpose is a pain point.

“It’s been 4 months since I applied, and I tried re-submitting my documents 3 times. They keep saying that my documents are incorrect, but why?”

This person has tried unsuccessfully to submit an application several times and does not understand why their documents are rejected. **Understanding the application** is a pain point. Receiving **feedback** on errors is a pain point. The **time taken to respond** is a pain point.

PART 3

The most important data I would like to see to be confident in an analysis of this sort is the demographic distribution of Floridians who qualify for SNAP benefits. I believe this population would differ from that of the general population in important ways, so relying on the general population as a proxy is dangerous. Other general information I would like is to see the entire customer journey involved in receiving SNAP benefits – a walkthrough of all processes involved in applying for and receiving food stamps in Florida. A potential leverage point is the complexity of the application process. To this end, useful data would include the number of people who began an application but did not complete it or did not correctly complete it and the number of people who completed an application on behalf of someone else (as well as the demographics of these populations if possible). Another area worth exploring is the clarity of the eligibility requirements. For this, useful data to have would be the number of ineligible individuals who applied and the proportion of claimants who fell outside of typical requirements (e.g. earned more than the minimum for an individual but had multiple dependents, qualified for social security, etc.). A final leverage point could involve the way in which people identified SNAP. For this, it would help to know the ways in which people accessed the SNAP application and the information sites. For instance, if they came from a web search, what did they search for? If they applied in person, how did they know where to go?

PART 4

I hypothesize that Florida senior citizens are chronically underrepresented among those receiving SNAP benefits due to the complexity of the online application process. In the dataset of 20,000 SNAP recipients, the oldest person was 60 years of age. Given the assumption that this dataset is a representative sample of SNAP recipients, a significant demographic – namely, those over the age of 60 – is missing. Based on eligibility requirements I found, nothing precludes elderly Florida citizens from claiming SNAP benefits. The logical conclusion is that there are systemic issues preventing this population from getting their benefits. Older individuals are notoriously less technologically savvy. Especially during Covid-19, online applications are the only safe way for senior citizens, who are high-risk, to apply. As pointed out by one user, however, whose grandmother had trouble applying, using the online system even more necessary in these times can be difficult for this group. Other user feedback indicates that the application process is confusing, with one stating they need help from their child. Another has their application repeatedly rejected, which could be due to a technical error or a different error of understanding, but is certainly a downside to virtual applications, as there is no human present to explain the problem.

There are a few key parts of my hypothesis, each of which should be tested separately. One is that the online application system is complex. Several metrics could be used to validate this somewhat nebulous claim. The number of steps required, number of clicks required, number of documents to fill, number of web pages needed to access, and number of different types of software or hardware required could give objective clues as to complexity. The number or proportion of online applicants who make a technical error while applying or who begin an application without completing it provide more subjective user-experience clues. A second part of my hypothesis is that people over the age of 60 tend to struggle with using technology. Simple comparative studies could be done to test this idea (although there is already undoubtedly information supporting the claim), wherein individuals of different ages could be given a technological task and compared on completion and time taken to complete the task. Task difficulty could be varied as well. Metrics useful on this front could include the percentage of adults over 60 who own or who regularly use a computer, a tablet, or a smartphone (as compared to other age groups). A final part of my hypothesis is that elderly individuals have particular difficulty with online applications. This could be tested by examining the proportion of individuals who applied in person versus online broken down by age. Under my hypothesis, older recipients would be more likely to apply in person. A simple survey asking participants for their age and whether they would prefer to apply to something in person or online could provide useful data. Perhaps the most useful step to take to both verify this hypothesis and begin down the path of making improvements would be to observe elderly users attempt to submit a virtual SNAP application. If a high proportion fail, take a very long time, or are significantly confused by the process, it is likely that this is indeed a problem point for this population. (To be more precise, one would compare to other age groups, but in an observational study, many pain points specific to the older population would probably be evident.) This has the additional advantage of providing insight into which parts of the technology and the SNAP application problem most consistently provide problems for users over 60.

PART 5

Part 1: I took about 4 hours, 30 min in total to load the data, analyze it, and make visualizations (including upwards of an hour working on visualizations I eventually scrapped). I could easily spend several more hours breaking down the data in other ways, and most likely would want to in a real-world setting.

Part 2: I took about 20 minutes to go through each user quote and identify pain points.

Part 3: I spent 30 minutes brainstorming and laying out alternate data sources I would like.

Part 4: I took about 40 minutes to generate a hypothesis and analyze metrics to test it.

Total: 6 hours