

Final Project Proposal for Data 602

by Jean Jimenez

FOOD INSECURITY IN THE UNITED STATES

RESEARCH QUESTION

Is there a difference in the food insecurity rate of people of different racial and ethnic groups in the United States? If so, which States are these differences more prominent in?

JUSTIFICATION

Food insecurity is an issue that does not only affect individuals living in the 'Third World'. Food insecurity is a harsh reality in our own backyard. While the government spends billions of dollars to bomb helpless children in the middle east, children and adults in our own back yard are experiencing food insecurity. Programs like SNAP, EBT, and WIC serve as a social safety net to address some of these food insecurity issues. However, the government is increasingly defunding social programs, and the threshold of what is considered a 'food insecure individual' is constantly changing due to different definitions in different states. By highlighting differences in food insecurity rates among different demographic groups, we can divert resources to people and communities that need them the most.

DATA SOURCES

I will be using data from Feeding America. Feeding America is a large nonprofit organization dedicated to fighting hunger across the United States. It gathers data on food insecurity from a variety of sources, including government reports, direct observation, and its extensive network of food banks and community organizations. This data helps them identify who needs help and where the greatest needs are. They make their findings public to raise awareness, drive policy changes, and encourage donations and volunteerism. By sharing information openly, Feeding America aims to promote more collaboration and more effective solutions to hunger nationwide.

The Dataset I will be using has the following Citation. The dataset is NOT publicly available and I had to put in a request for it: Feeding America. "Map the Meal Gap 2023: A Report on County and Congressional District Food Insecurity and County Food Cost in the United States in 2019-2021." 2023.

I will also be using [2020 Census Data](#) as well as [US Census Map Shapefiles](#) to map geospatial data.

LIBRARIES TO POTENTIALLY USE

I will potentially use: -Pandas -Numpy -Matplotlib -seaborn -Plotly -geopandas -scipy

Exploratory Data Analysis and Summary Statistics

In the following, I show summary stats of Food Insecurity Rates for 4 States as well as their distributions. I will compare all states.

There are 1308 total observations in this dataset. States with higher populations (CA, TX, NY) have more observations.

```
In [ ]: import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

```
In [ ]: congressional_data = pd.read_csv('2019_2021_cong_DI_dat.csv')

target_states = ['LA', 'AL', 'NY', 'VA']
congressional_data_filtered = congressional_data[congressional_data['State'].isin(target_states)]

congressional_data_melted = congressional_data_filtered.melt(id_vars=['State', 'year'],
                                                             value_vars=['FI_rate_black', 'FI_rate_hispanic', 'FI_rate_white'],
                                                             var_name='Race_group', value_name='FI_rate_value')
```

```
In [ ]: summary_stats = congressional_data_melted.groupby(['State', 'Race_group'])['FI_rate_value'].agg(['mean', 'std', 'min', 'max'])

print(summary_stats)

plt.figure(figsize=(12, 8))
sns.boxplot(x='State', y='FI_rate_value', hue='Race_group', data=congressional_data_melted)

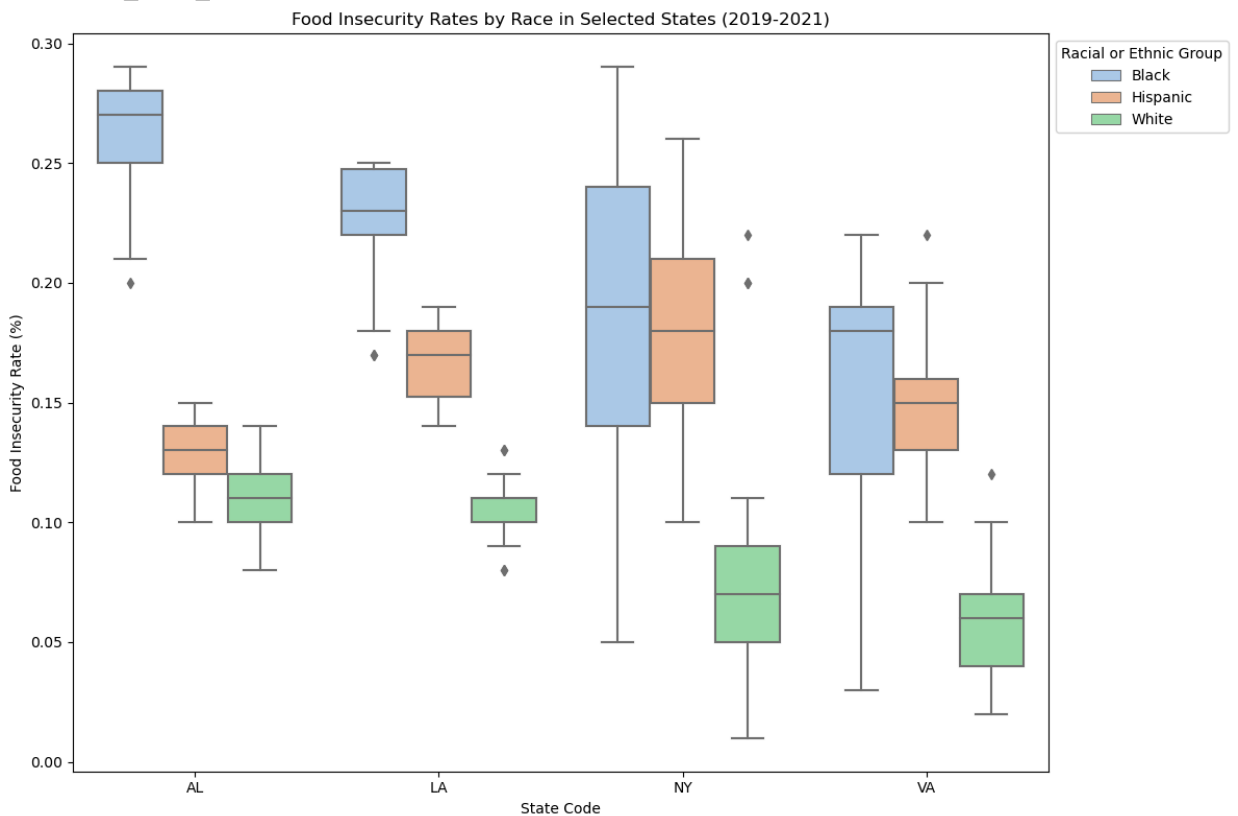
plt.title('Food Insecurity Rates by Race in Selected States (2019-2021)')
plt.ylabel('Food Insecurity Rate (%)')
plt.xlabel('State Code')

leg = plt.legend(title='Racial or Ethnic Group', loc='upper left', bbox_to_anchor=(1, 0.5))
new_labels = ['Black', 'Hispanic', 'White']
for t, l in zip(leg.texts, new_labels):
    t.set_text(l)

plt.tight_layout()
plt.show()
```

		count	mean	std	min	25%	50%	75%	\
State	Race_group								
AL	FI_rate_black	21.0	0.262857	0.025523	0.20	0.2500	0.27	0.2800	
	FI_rate_hispanic	21.0	0.128571	0.015260	0.10	0.1200	0.13	0.1400	
	FI_rate_white	21.0	0.109524	0.017742	0.08	0.1000	0.11	0.1200	
LA	FI_rate_black	18.0	0.223889	0.026377	0.17	0.2200	0.23	0.2475	
	FI_rate_hispanic	18.0	0.167778	0.018329	0.14	0.1525	0.17	0.1800	
	FI_rate_white	18.0	0.105556	0.014642	0.08	0.1000	0.11	0.1100	
NY	FI_rate_black	81.0	0.183827	0.064818	0.05	0.1400	0.19	0.2400	
	FI_rate_hispanic	81.0	0.182346	0.038996	0.10	0.1500	0.18	0.2100	
	FI_rate_white	81.0	0.070741	0.037175	0.01	0.0500	0.07	0.0900	
VA	FI_rate_black	33.0	0.158182	0.050895	0.03	0.1200	0.18	0.1900	
	FI_rate_hispanic	33.0	0.146061	0.028825	0.10	0.1300	0.15	0.1600	
	FI_rate_white	33.0	0.058788	0.024464	0.02	0.0400	0.06	0.0700	

		max
State	Race_group	
AL	FI_rate_black	0.29
	FI_rate_hispanic	0.15
	FI_rate_white	0.14
LA	FI_rate_black	0.25
	FI_rate_hispanic	0.19
	FI_rate_white	0.13
NY	FI_rate_black	0.29
	FI_rate_hispanic	0.26
	FI_rate_white	0.22
VA	FI_rate_black	0.22
	FI_rate_hispanic	0.22
	FI_rate_white	0.12



```
In [ ]: state_counts_all = congressional_data['State'].value_counts()
print("Count of rows per state for all states:")
print(state_counts_all)
```

```
total_count_all = congressional_data.shape[0]
print("Total count of observations for all states:", total_count_all)
```

Count of rows per state for all states:

State

CA	159
TX	108
FL	81
NY	81
IL	54
PA	54
OH	48
GA	42
MI	42
NC	39
NJ	36
VA	33
WA	30
TN	27
IN	27
AZ	27
MA	27
MN	24
MD	24
MO	24
WI	24
SC	21
AL	21
CO	21
LA	18
KY	18
OR	15
CT	15
OK	15
AR	12
NV	12
KS	12
IA	12
MS	12
UT	12
WV	9
NE	9
NM	9
HI	6
ME	6
RI	6
ID	6
NH	6
SD	3
DC	3
DE	3
VT	3
ND	3
MT	3
AK	3
WY	3

Name: count, dtype: int64

Total count of observations for all states: 1308