SENATOR MARKEY'S TEAM GROUP B

TEAM BREAKDOWN

Daniel Strick

Anthony Huang

Carlos Garcia Padilla

John Salloum
Team Lead

Atul Aravind Das

DISCRETIONARY GRANT EQUITY ANALYSIS

- Our team was tasked with analyzing the distribution of federal grants across Massachusetts, with a focus on equity in Gateway Cities and disadvantaged areas.
- We aimed to develop a fully documented, repeatable process that could be applied to future years' data, allowing for future use by the team
- The goal was to help Senator Markey's office identify gaps within the data, visualize where the grants were directed, and understand what demographics were most important regarding where funding should go

DATA STRUCTURE

- Initial dataset contains funding initiatives for Massachusetts, derived from invest.gov
- Each row is a given funding initiative
- Notable features:
  - Agency Name
  - Bureau Name
  - Program Name
  - Category
  - Subcategory
  - Project Name
  - City
  - County
  - Funding Amount Excluding Loans
  - Funding Source
  - Program Type

ISSUES WITH DATA

- Funding projects lack background information/descriptions
- Irregularities in naming schemes, particularly in location names
  - Different cases
  - Inconsistency with higher granularity intra-city labeling
  - Misspellings
- Confusion regarding handling of unique cases:
  - Statewide funding initiatives
  - Inter-city funding initiatives
- Poor formatting, particularly with newlines and spacings

DATA ENRICHMENT

- Decided to drop inter-city and statewide rows from data
- Meaningful analysis required more complex location data
- Assigned sub-county “GEO_ID” value for each location via sub county census tract api call
- Utilized GEO_ID values to query 2020 Decennial data from census api
- Obtained more granular data from census api containing race and income information

DATA ENRICHMENT - A VISUALIZATION

Initiative	City	Funding Amount	…
1	Cityville	$1,000,000	
2	Village Town	$200,000	
3	West County	$5,000,000	

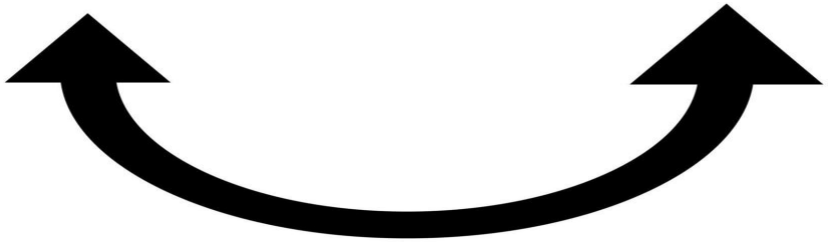
City	GEO_ID
Cityville	0600000US2500334655
Village Town	0600000US2500344385
West County	0600000US2500353050

Sub County Census Tract API

DATA ENRICHMENT - A VISUALIZATION (2)

Initiative	City	Funding Amount	GEO_ID	…
1	Cityville	$1,000,000	0600000US2500334655	
2	Village Town	$200,000	0600000US2500344385	
3	West County	$5,000,000	0600000US2500353050	

GEO_ID	Population	Median Income	…
0600000US2500334655	17062	$106001	
0600000US2500344385	69420	$56906	
0600000US2500353050	177013	$98182	



2020 Decennial Data, Census API

DATA ENRICHMENT - A VISUALIZATION (3)

Initiative	City	Funding Amount	GEO_ID	Population	Median Income
1	Cityville	$1,000,000	0600000US2500334655	17062	$106001
2	Village Town	$200,000	0600000US2500344385	69420	$56906
3	West County	$5,000,000	0600000US2500353050	177013	$98182

VARIABLES CHOSEN

- From https://api.census.gov/data/2020/dec/pl we gained
  - P1_001N - Total Population
  - P1_002N - Population consisting of one race
  - P1_003N - White Population
  - P1_004N - Black or African American Population
  - P1_005N - American Indian and Alaskan Native Population
  - P1_006N - Asian Population
  - P1_007N - Native Hawaiian and Other Pacific Islander Population
  - P1_008N - Population of other race
  - P1_009N - Population of two or more races

- From https://api.census.gov/data/2020/acs/acs5 we gained
  - B19013_001E - Median Household income (Within the last 12 months)
  - B25034_001E - Total Structures Built
  - B25034_009E - Structures Built (1950-59)
  - B25034_010E - Structures Built (1940-49)
  - B25034_011E - Structures Built (1939-pre)

MEASURING LEAD POISONING

- Used pre-1960 housing as an indicator of lead poisoning

- Housing Data:
  - B25034_001E - Total Structures Built
  - B25034_009E - Structures Built (1950-59)
  - B25034_010E - Structures Built (1940-49)
  - B25034_011E - Structures Built (1939-pre)

- This was calculated as a percentage:
  - Percent Pre-1960s Housing = (B25034_009E + B25034_010E + B25034_011E) / B25034_001E
  - Percent Pre-1960s Housing = (pre40s + 40s + 50s) / All Housing

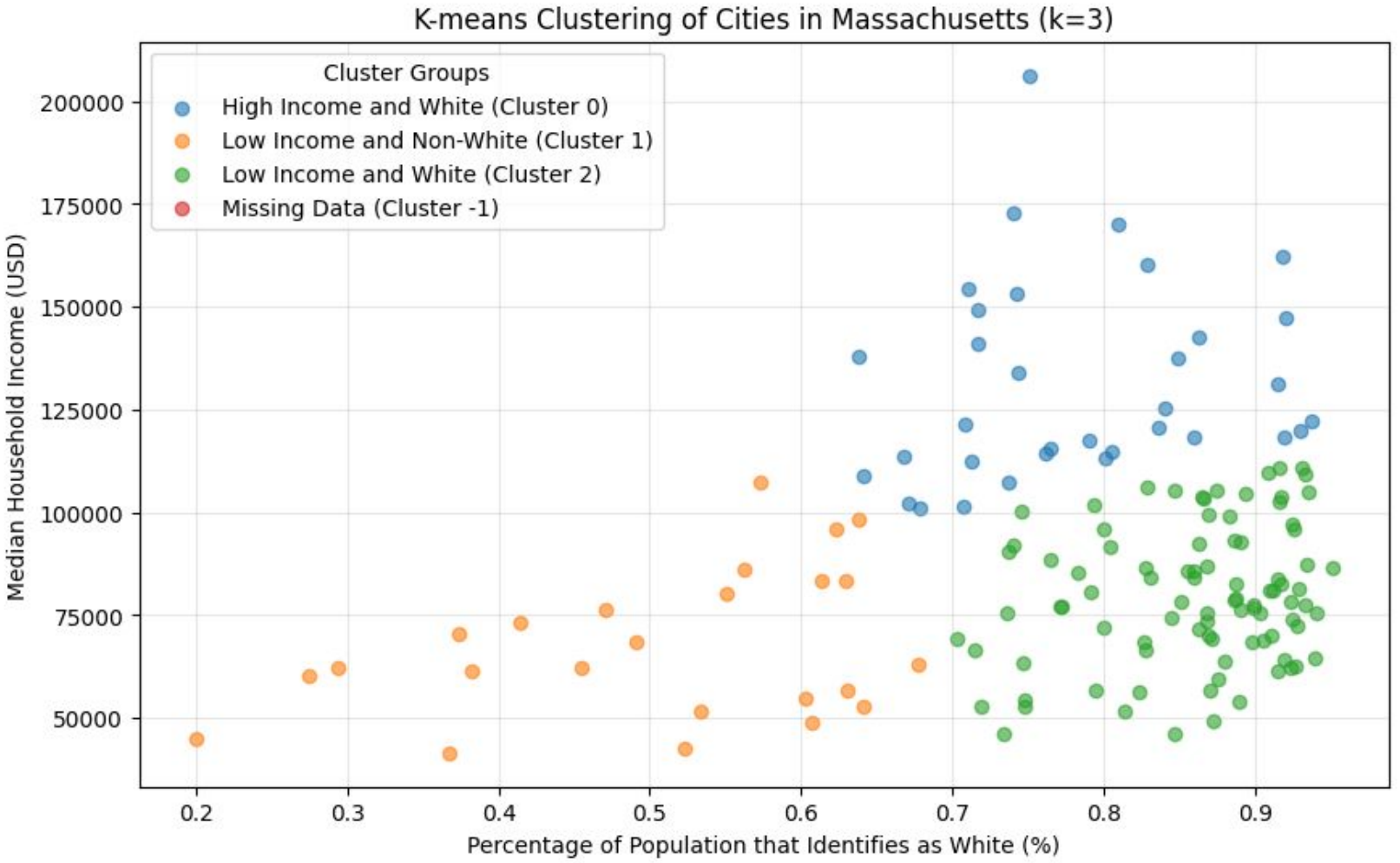
FEATURES CREATED

- Wanted to make some rudimentary groupings based on demographics

- “Majority White” racial groupings:
  - White Population / Total Population
  - >79% - “Majority White”
  - ≤79% - “Not Majority White”

- Median income groupings:
  - <$65,000
  - $65,000-$90,000
  - $90,000-$115,000
  - >$115,000

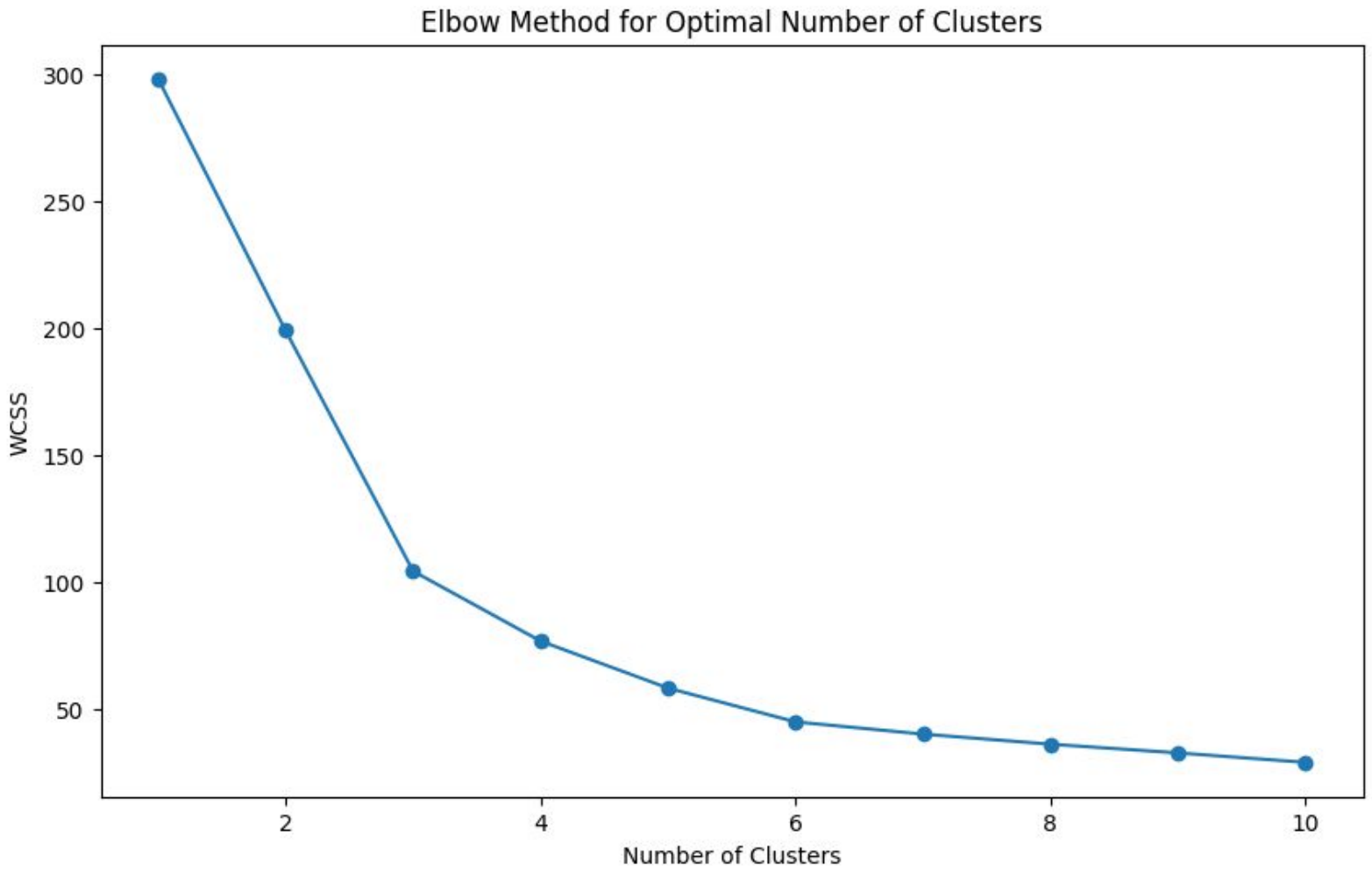
CLUSTERING ON INCOME AND RACE



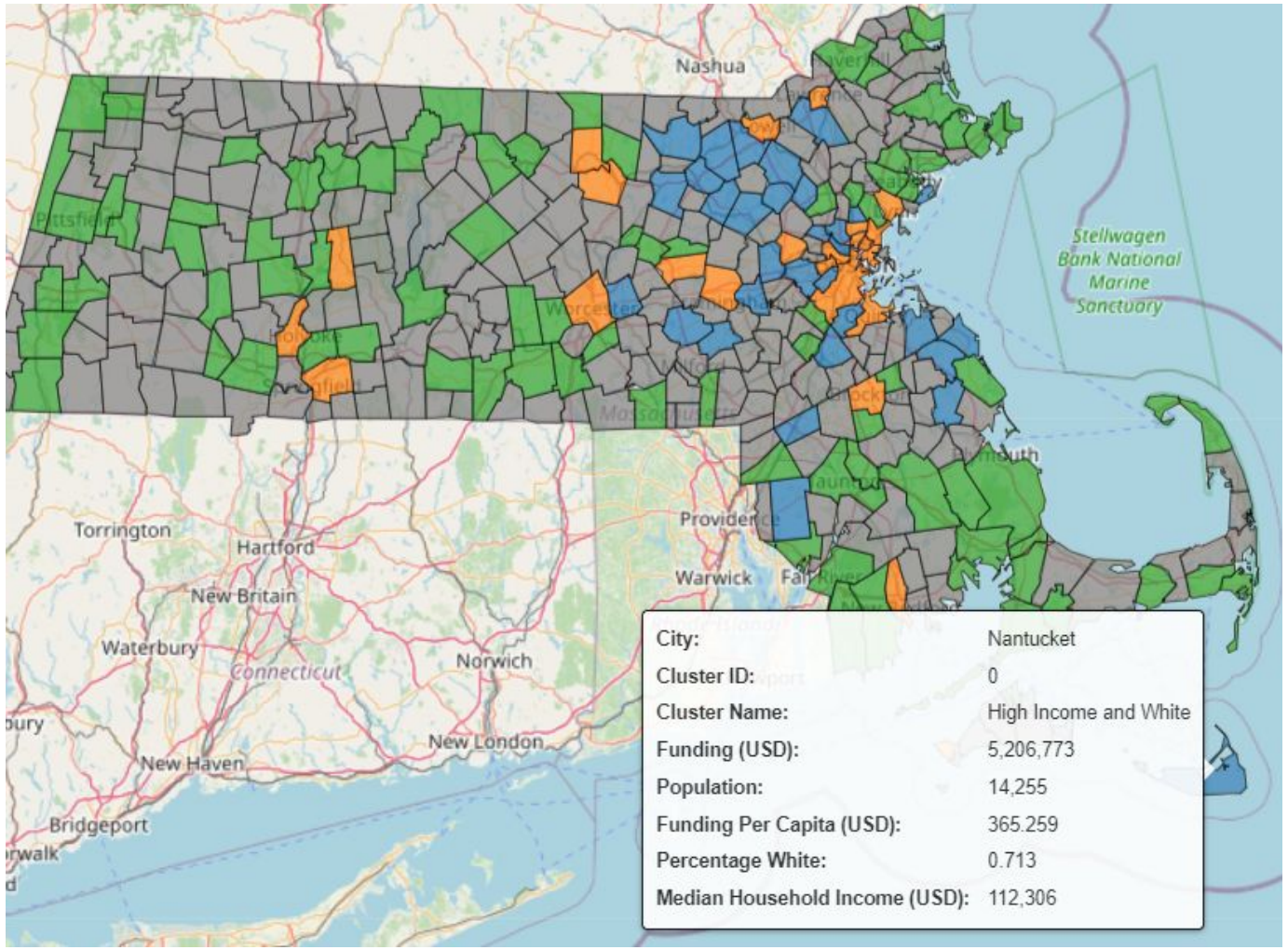
Clusters:

- Orange = Low income, Bipoc
- Green = Low income, White
- Blue = High Income, White

CLUSTERING ON INCOME AND RACE



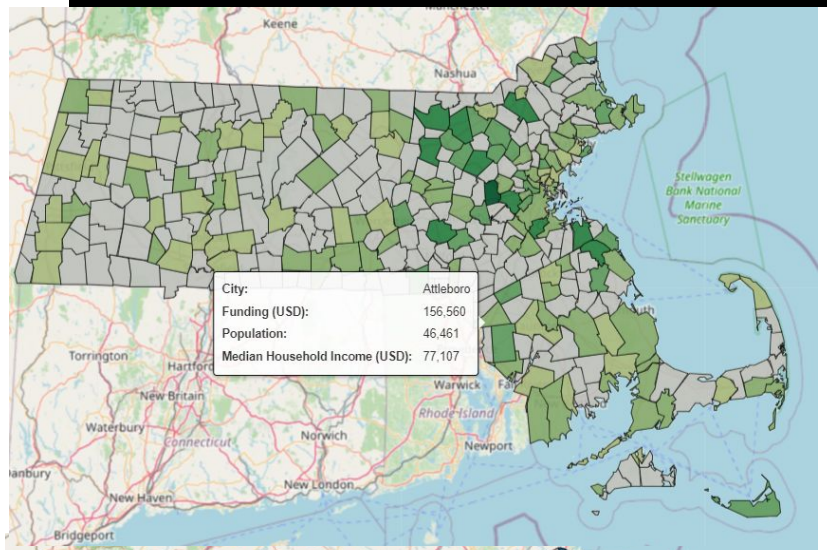
Visualizing Clustering with Geography



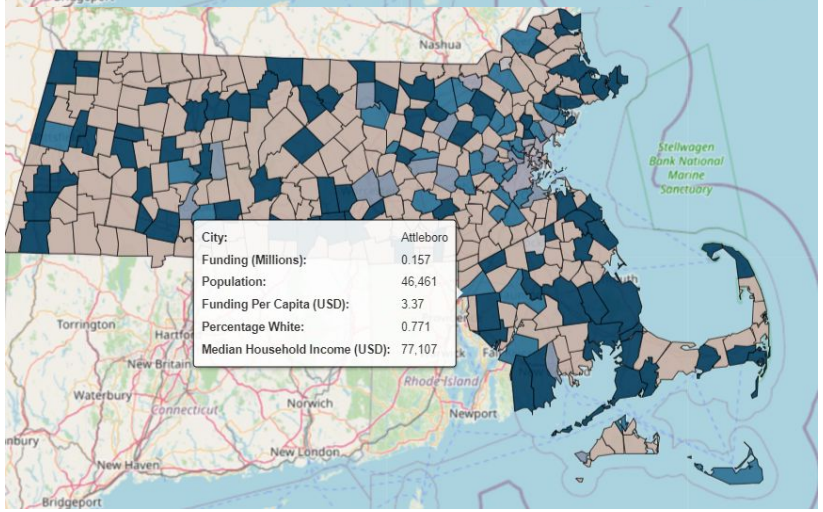
Clusters:

- Orange = Low income, Bipoc
- Green = Low income, White
- Blue = High Income, White

HEAT MAPS BY RACE INCOME AND FUNDING

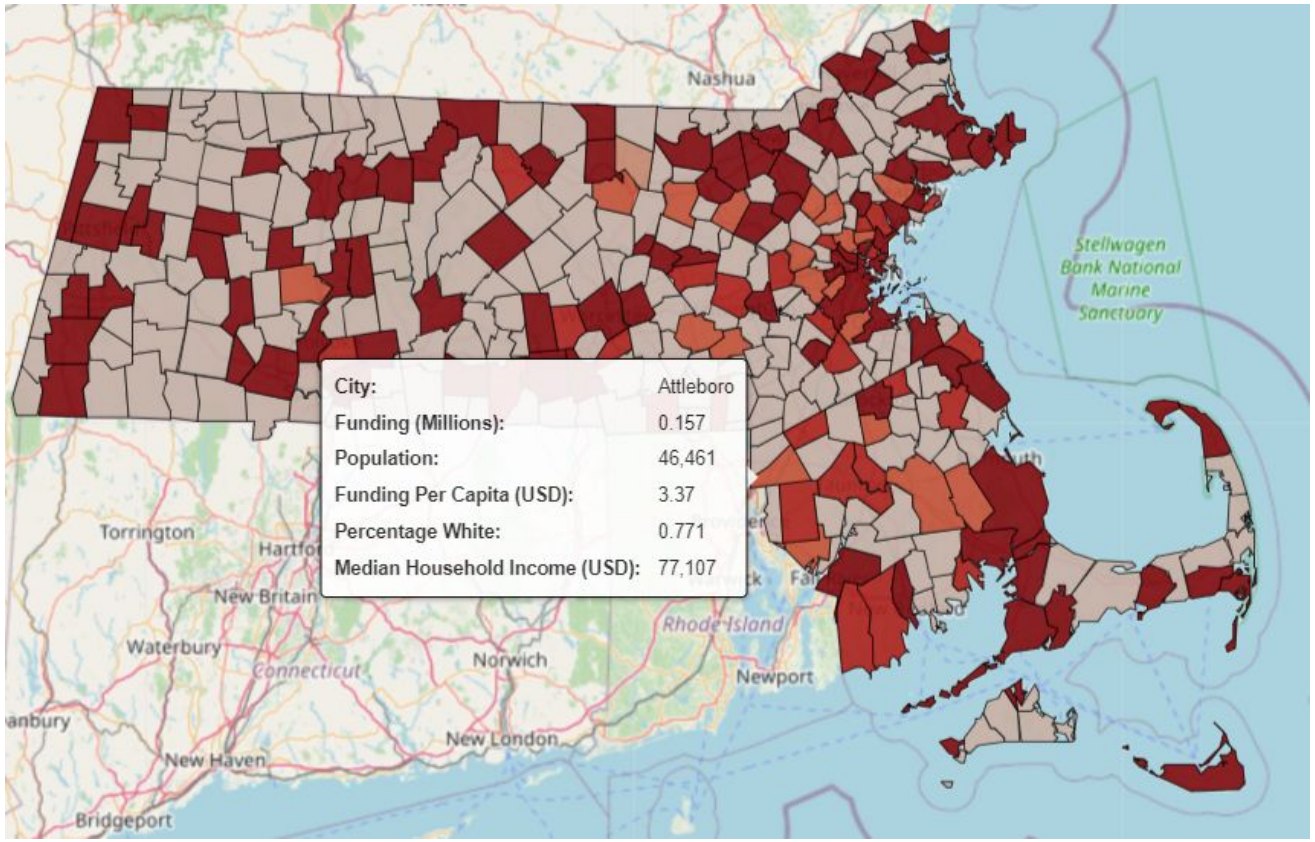


INCOME

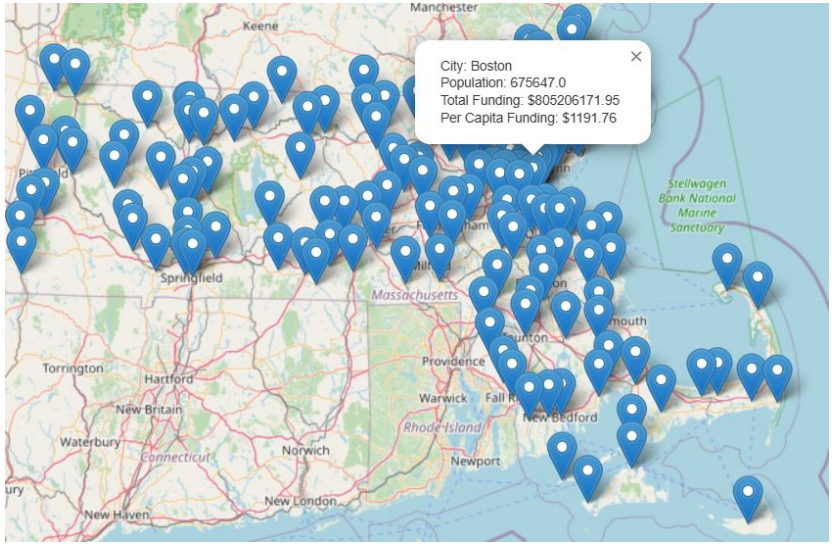


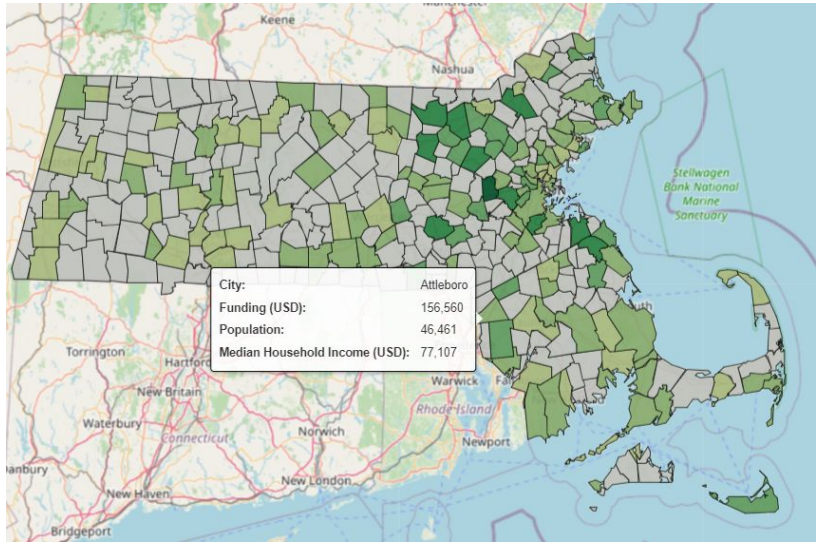
PERCENT WHITE

FUNDING PER CAPITA



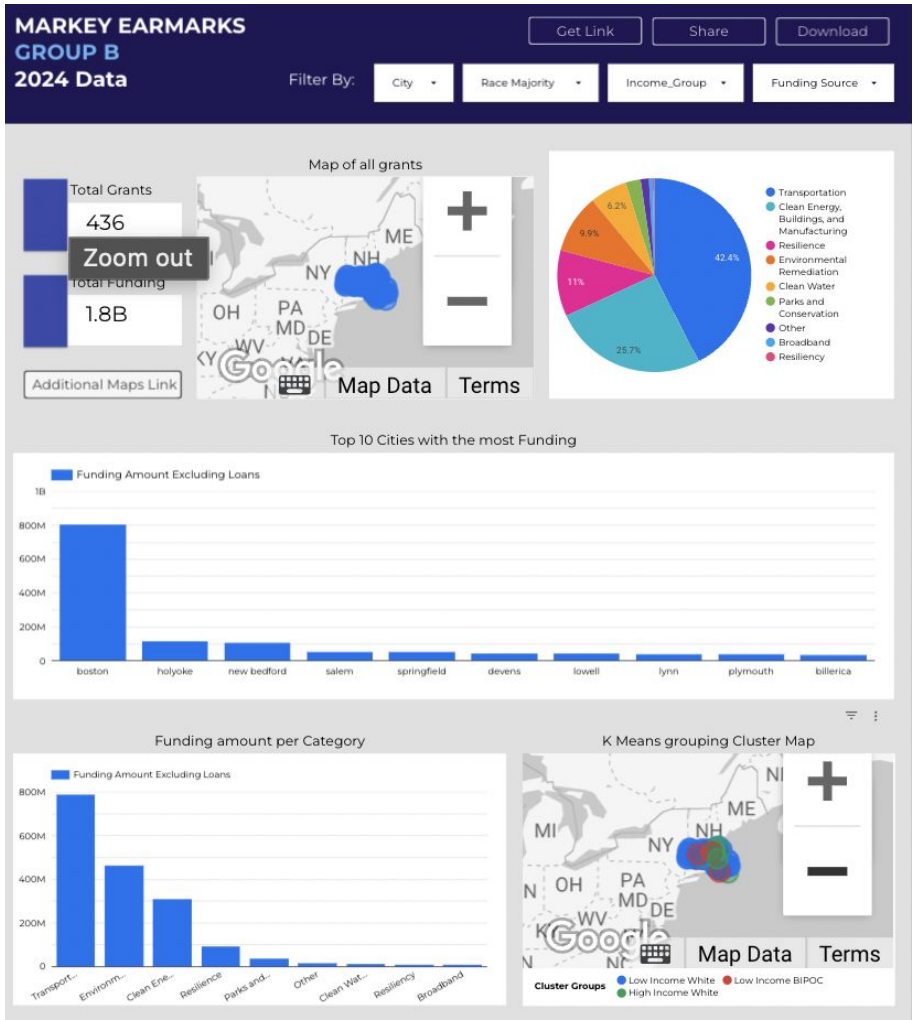
MAPS: EXPLAINING THE DETAILS

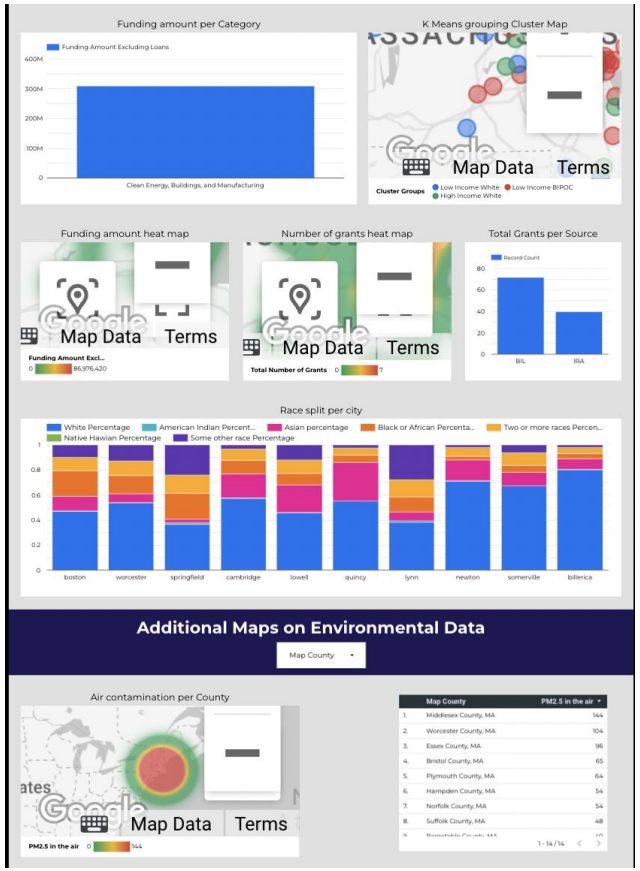




Massachusetts Maps

LOOKER STUDIO DASHBOARD





Interactive Link Here

KEY TAKEAWAYS

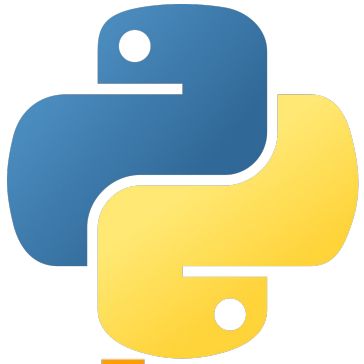
- The cities that the grants are given to can be grouped by the following; High income White Majority, Low income White Majority, and Low income Bipoc Majority
- We also see trends that suggest that the higher the income or the white population, the grants given are generally lower, suggesting focus is on low income, Bipoc majority areas
- Environmental remediation was a key point, being the second most present type of gran in the database and almost only being given to low income, Bipoc majority communities, which is where it should be going

KEY TAKEAWAYS

- Towns with lower household incomes receive slightly more funding per capita. However, this relationship is not strong, so it's only a general trend rather than a definitive rule.
- Towns with larger racial minority populations may receive a bit more funding per capita, however, the relationship is very slight.
- In summary, these findings suggest that poorer towns and those with larger racial minority populations may receive more funding, but these are not strong trends.
- This indicates that other factors, beyond income and racial demographics, likely play a role in how funding is distributed across towns.

TOOLS USED

- Pandas - csv reading/writing
- pyplot - making plots
- Google Colab - collaborative code writing
- Looker Studio - dashboard for our data
- Python - programming language used
- Folium - interactive map creation
- requests - https requests



CONCLUSION

- In conclusion, we can see that the grants are mainly going to low income, highly diverse areas, which is where they generally should be going
- Environmental remediation grants were the most popular amongst these areas, showing that there was enough work done in areas with run down property.
- Likewise transportation was high overall, talking up roughly 42% of all funding. In big cities public transportation is more readily available, so this could be removed in future analysis to focus more on other grants.
- The dashboard we created displayed a complete analysis with the possibility of additional analysis with future year’s grants

Thank you!