Fruit And The City: Measuring The Accuracy And Bias Of Public Science

Sarah Neeves, B.S. Geography and Environmental Systems, GIScience Certificate
Mentor: Dillon Mahmoudi, Assistant Professor, Geography and Environmental Systems

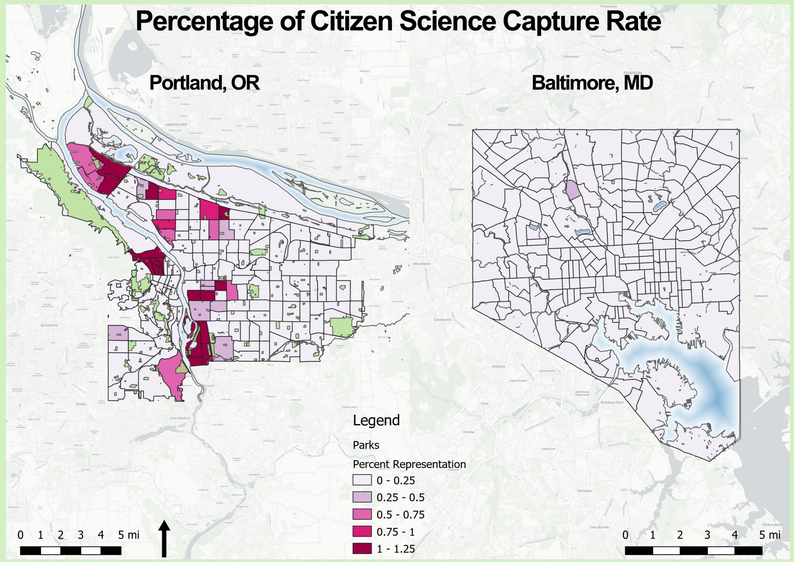
Background

Participatory science can be a valued method for researchers as it allows for mass data collection. Without participatory science, data may be overly expensive or impossible to obtain. Participatory science data collection can produce biased or inaccurate data if participants are not educated in the subject matter, there is not a wide variety of participants, or if there is not sufficient spatial coverage of data. Using geographic information system applications, this study aims to compare complete public tree inventories to Falling Fruit inventory, a participatory science database for fruit producing trees in urban cities. This presentation examines for accuracy as well as analyze the spatial patterns of the data to detect spatial biases along lines of income and race/ethnicity. Early findings of this exploratory study has produced metrics to calculate accuracy and bias of the Falling Fruit database across Baltimore, Maryland and Portland, Oregon.

Methods

Tree inventories for the 200 largest cities/metropolitan areas within the United States were researched and collected if data was available. Only about 15-20 city tree inventories were collected. The entire Falling Fruit database was collected as well. After tree data was collected, census data was also collected. Data collected from the census included: city boundaries, county census tracts, water features, and census tract data for median household income, and race/ethnicity in 2020. Edible fruit tree species were then filtered from city tree inventories to create a fruit tree inventory from city data. Lastly, the percentage of capture rate of Falling Fruit was calculated by dividing the Falling Fruit data by the city fruit tree inventory.

Results



Percent Capture Rate of Falling Fruit Data	Portland Average Median Household Income	Baltimore Average Median Household Income
0% - 25%	$75,539	$55,221
25% - 50%	$99,701	$62,188
50% - 75%	$91,896	No Data
75% - 100%	$86,734	No Data
Over 100%	$80,024	No Data

Percent Capture Rate of Falling Fruit Data	Percent Portland BIPOC	Percent Baltimore BIPOC
0% - 25%	30.0%	72.7%
25% - 50%	20.5%	72.5%
50% - 75%	28.7%	No Data
75% - 100%	28.3%	No Data
Over 100%	21.8%	No Data

Data and references available upon request: sneeves1@umbc.edu

Conclusion

After looking at the data, a trend was found that as Falling Fruit capture rate increases per tract percentage-wise, the percentage of Non-Hispanic White populations also seems to increase. There also seems to be a great disparity in capture rate between Baltimore and Portland. While this trend does appear in our findings, these early findings only show a comparison between two cities.

Discussion

Because this study only compares two cities, there could also be other circumstantial factors that contribute to the difference in capture rate. One possible factor that could contribute to this disparity is the difference in climates. Portland's climate may allow for more fruit trees to grow which might contribute to the fact that Portland's tree count is much higher than Baltimore's. Another factor that can contribute to the differences in capture rate is that the socio-spatial structure is different for both cities. Maryland has one of the highest median household incomes in the United States. Oregon has a lower median household income relative to Maryland. However, Portland has a higher median household income compared to the rest of Oregon, while Baltimore has a lower median household income compared to the rest of Maryland.

Next Steps

This study is an ongoing study, so the next steps are to look at data from other cities that fall within the largest 200 cities in the United States. So far, tree inventories from about 15 other cities have been collected. One thing that will be changed in this study going further is normalizing the data by looking at tree density per census tract instead of the city-wide count.