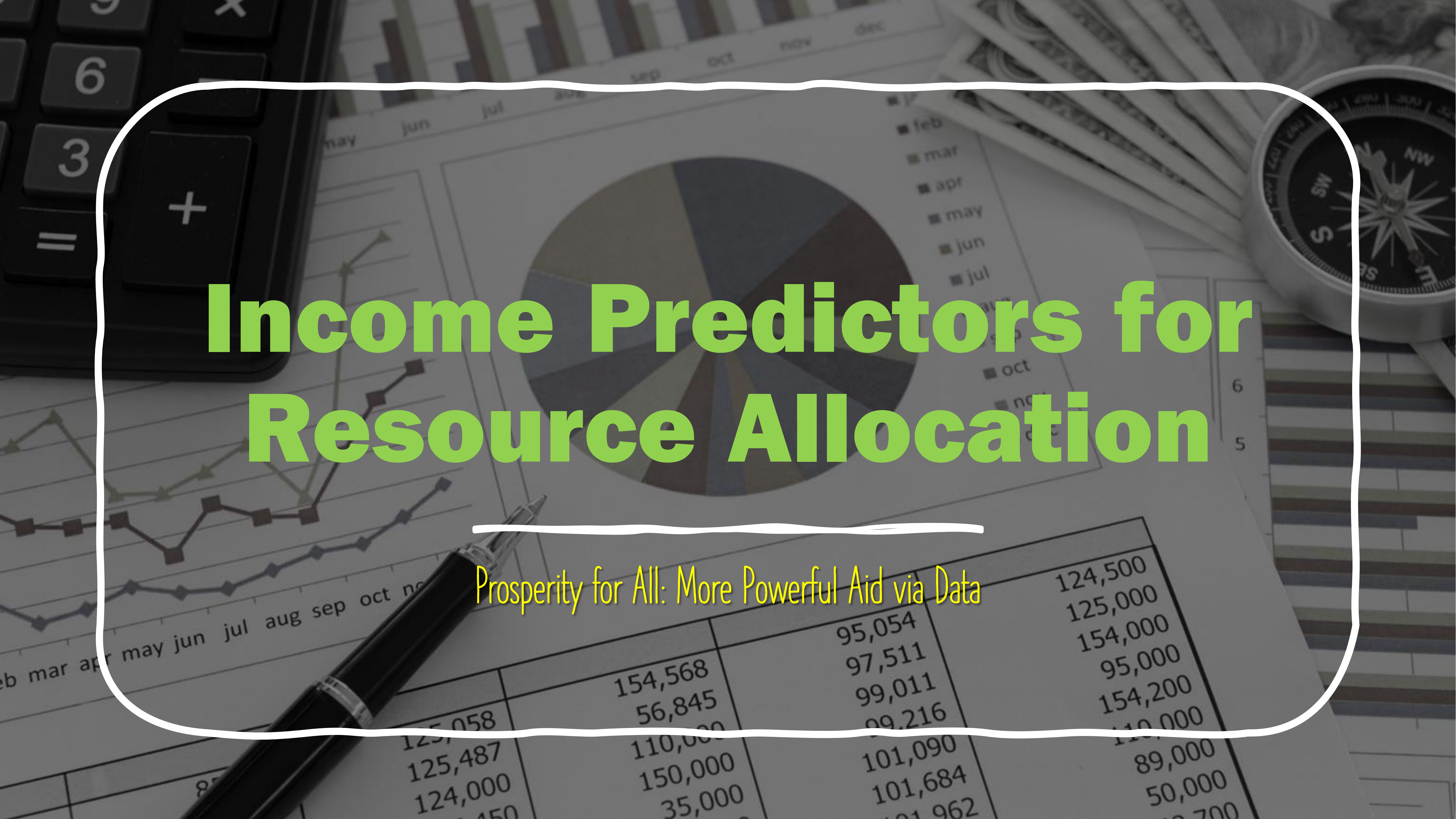


Income Predictors for Resource Allocation

Prosperity for All: More Powerful Aid via Data

The Goals of ‘Prosperity for All’

- <u>Aims for our charity are two-fold.</u>

- Improving socioeconomic mobility for those most in need.
- Developing and lobbying for new programs for the underserved.

Using Funds Now

- The coffers already contain usable funds.
- We need to identify segments of the population in the dataset which can utilize our aid.

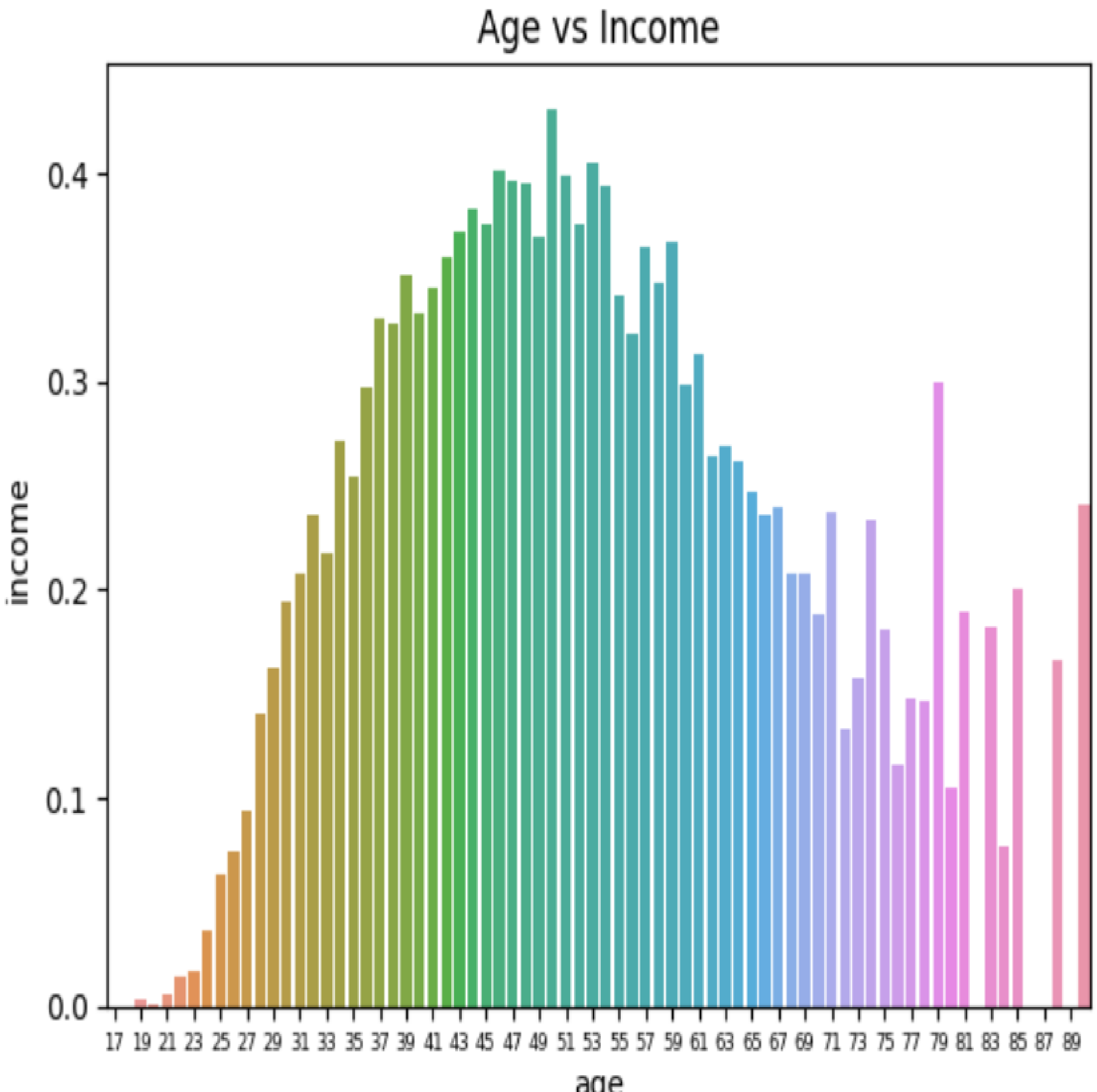
Fixed Demographic Segmentation

This dataset, as well as previous studies, have identified certain demographics which appear to be underserved.

Age, race, and gender are all methods of segmentation which can lead to more efficient distribution of resources to the most underserved.

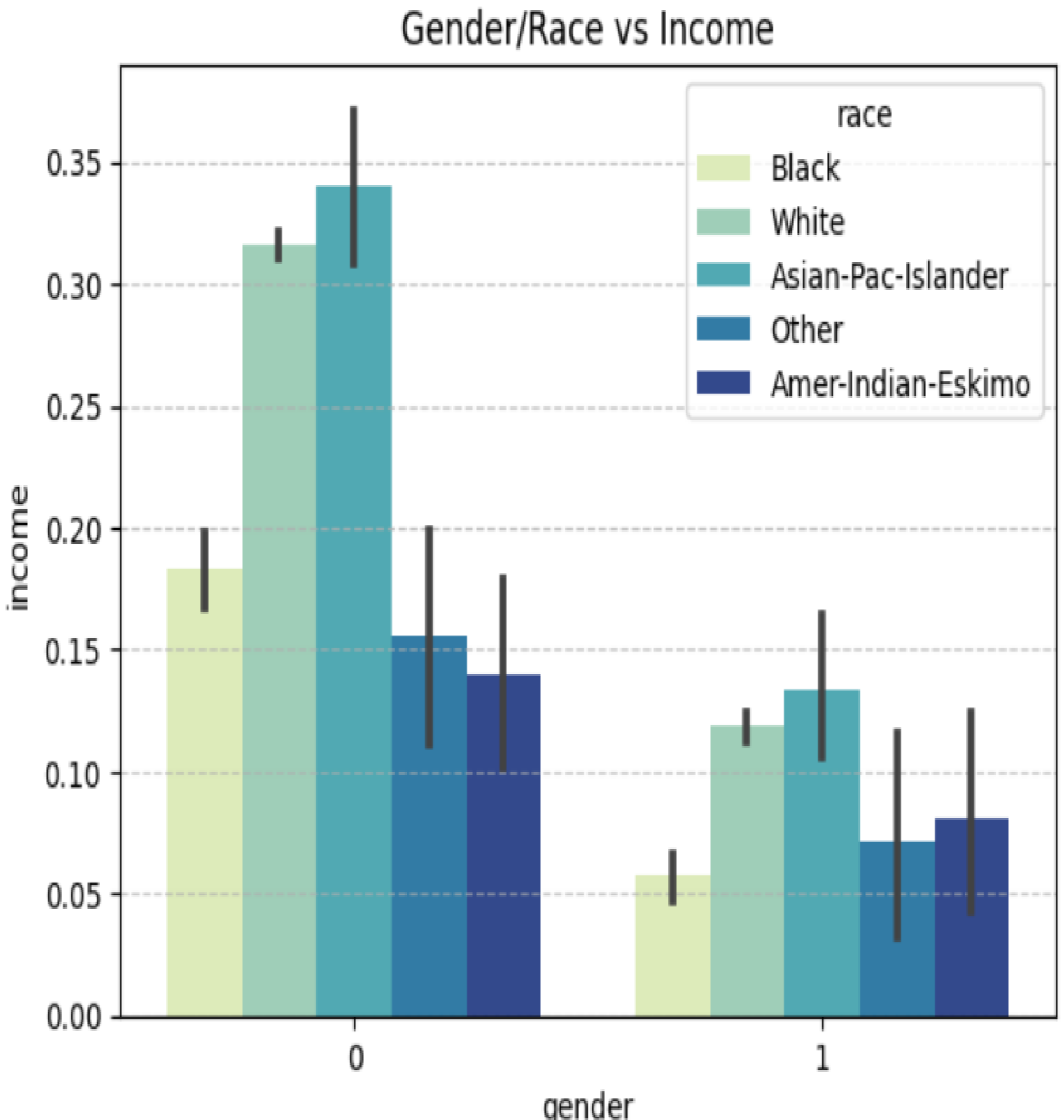
What the Demographics Illustrate

- Age: This data doesn't necessarily account for net wealth but it does clearly demonstrate the fact that young and old portions of the population (those outside of peak earning age groups) can both potentially qualify as deserving targets for our aid.



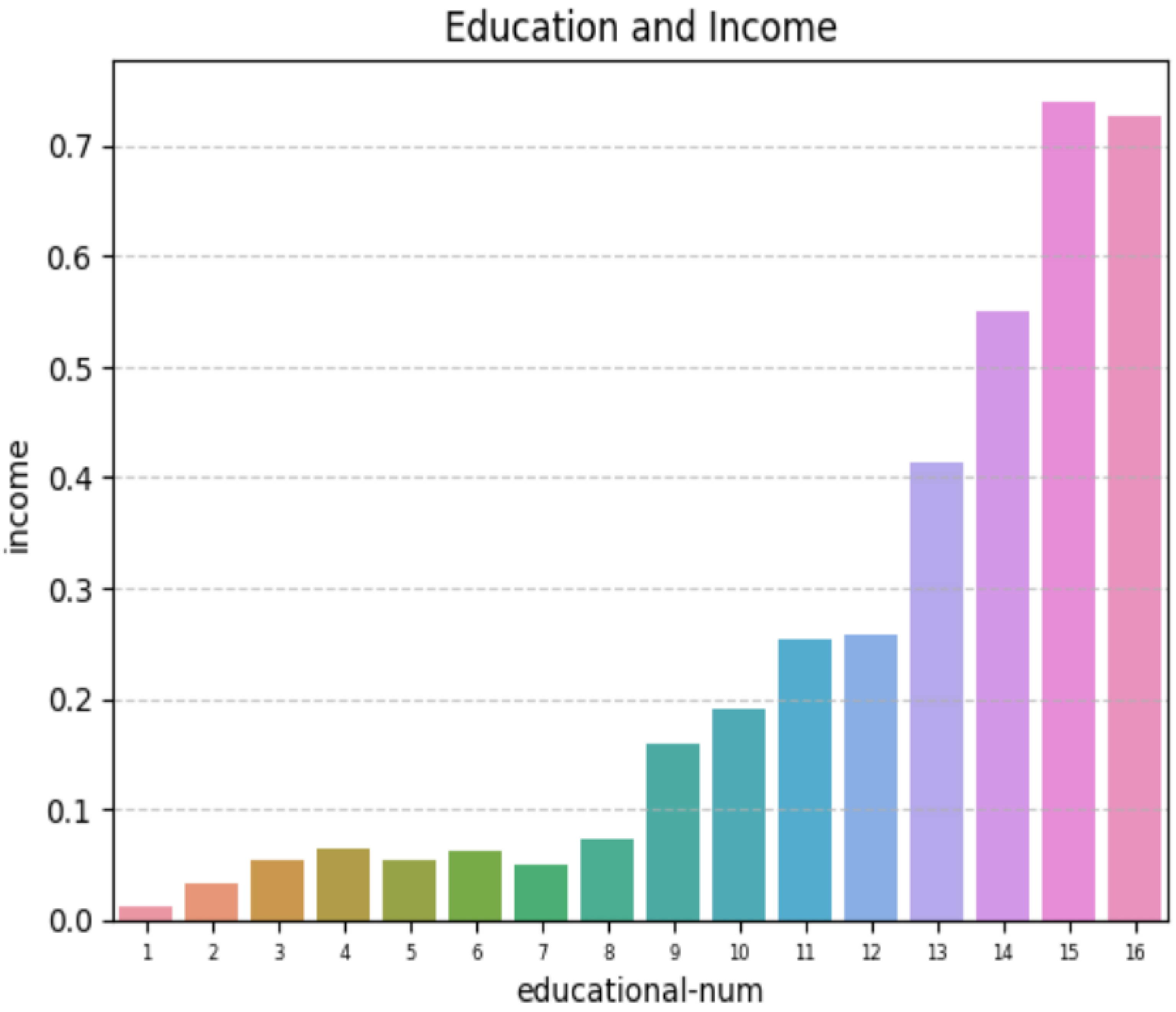
Further Demographic Inspection

- Gender and race can both be useful segmentation markers for identifying targets for our aid.
- Causation cannot be determined with this type of data but these correlations provide an excellent starting point for target research.
- As can be seen here, with males represented by ‘0’ and females by ‘1’, there are likely underserved segments which can be aided immediately.



Education!

- Educational attainment levels are both highly predictive of socioeconomic mobility and responsive to effective aid programs.
- There have been countless studies on the relationship between education and income and our dataset replicates these findings perfectly.

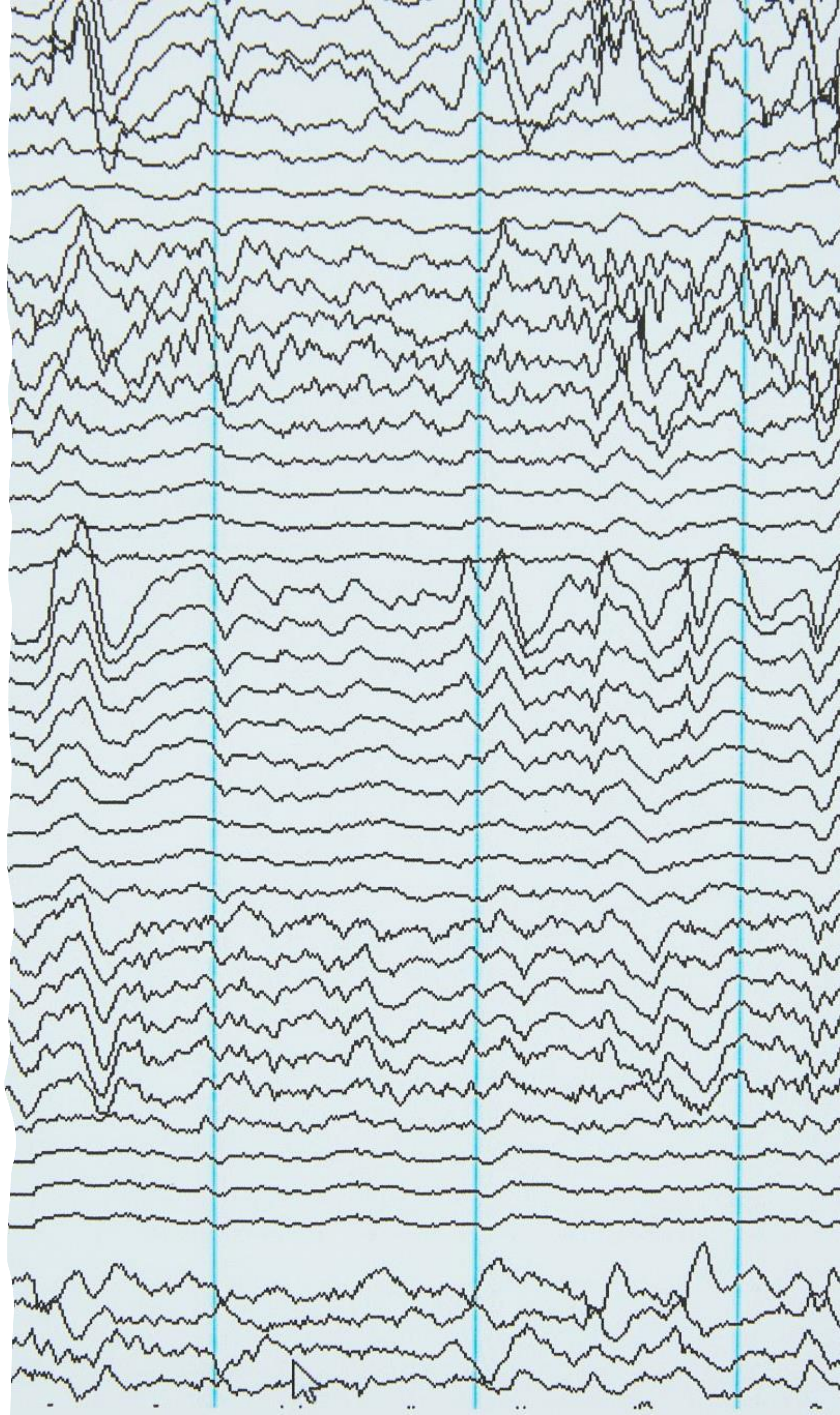


Lobbying for the Future

- ‘Prosperity for All’ believes in helping those in need now and identifying ways to improve socioeconomic mobility in the future.
- We will be using predictive analysis in our consultations with aid organizations and the government in order to effectively build aid and resource programs while at the same time making sure the right people and organizations get our help.

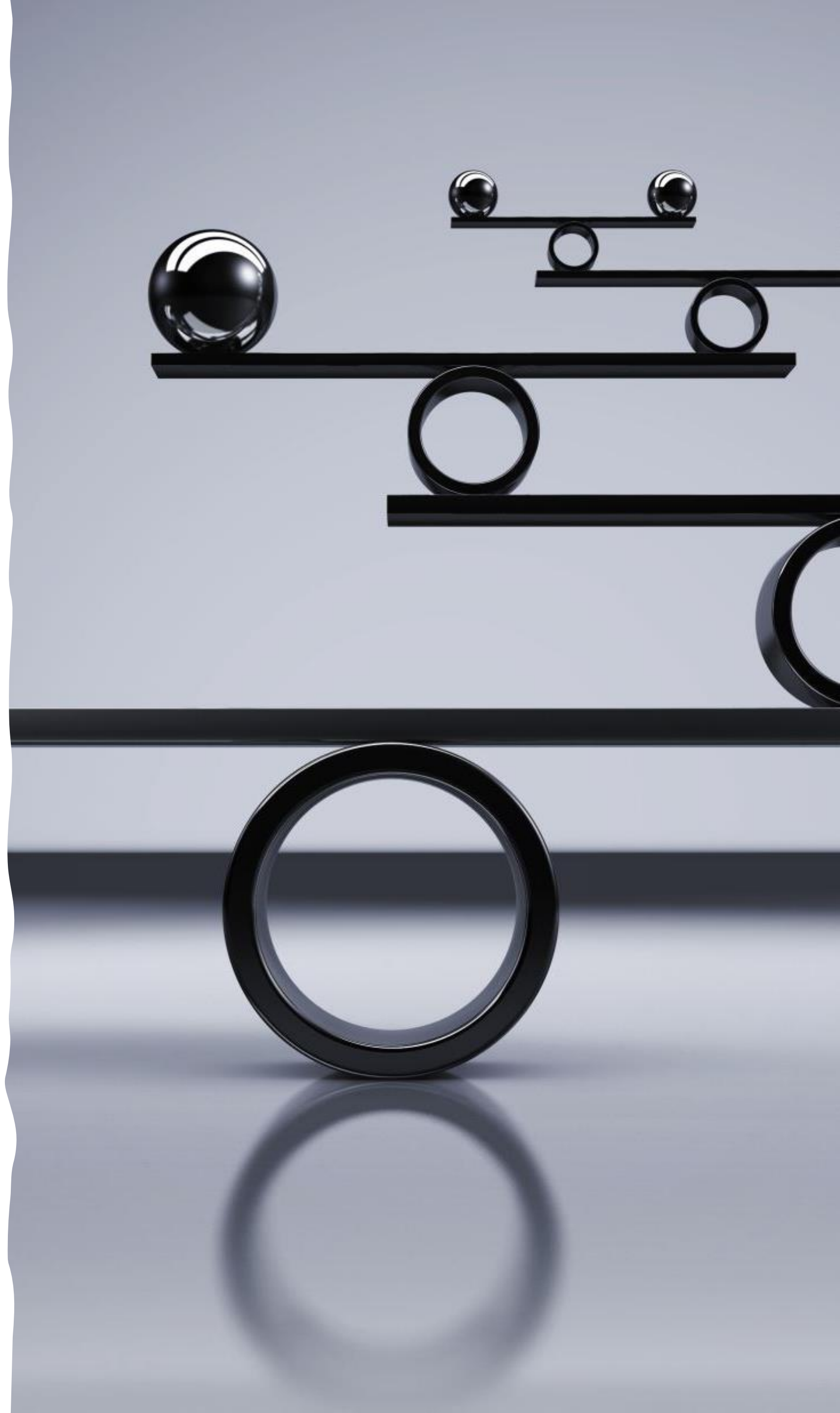
Predictive Analysis

- We tried and tested a number of predictive models on the dataset but got the best performance out of a model called KNN (K-Nearest Neighbors).
- The model attained an f-1 score of .77, meaning it performed fairly well in terms of correctly predicting which income category respondants would land in and avoiding false positive and false negative predictions.
- Of the 12,198 predictions there were 1,252 false negatives, 661 false positives, 1678 true positives, and 8,607 true negatives.



Limitations and Cautions

- Any false negatives or false positives could push resources away from people who actually need it and towards those who don’t.
- The false outcomes in this model are non-negligible and should be accounted for when using this information to determine outcomes.
- Furthermore, this is an unbalanced dataset which does seem to introduce some bias towards accurately predicting false negatives.



Conclusions

- It is vitally important to the people that need our resources, that we allocate them intelligently and compassionately. We can only do this with accurate data and good predictive models for our activism.
- The examination of this data and the predictive models utilized are not perfect but do provide a great starting point for our decision makers.
- The false predictions need to be accounted for but understanding their frequency and likelihood should allow the leaders of our organization to make informed decisions.