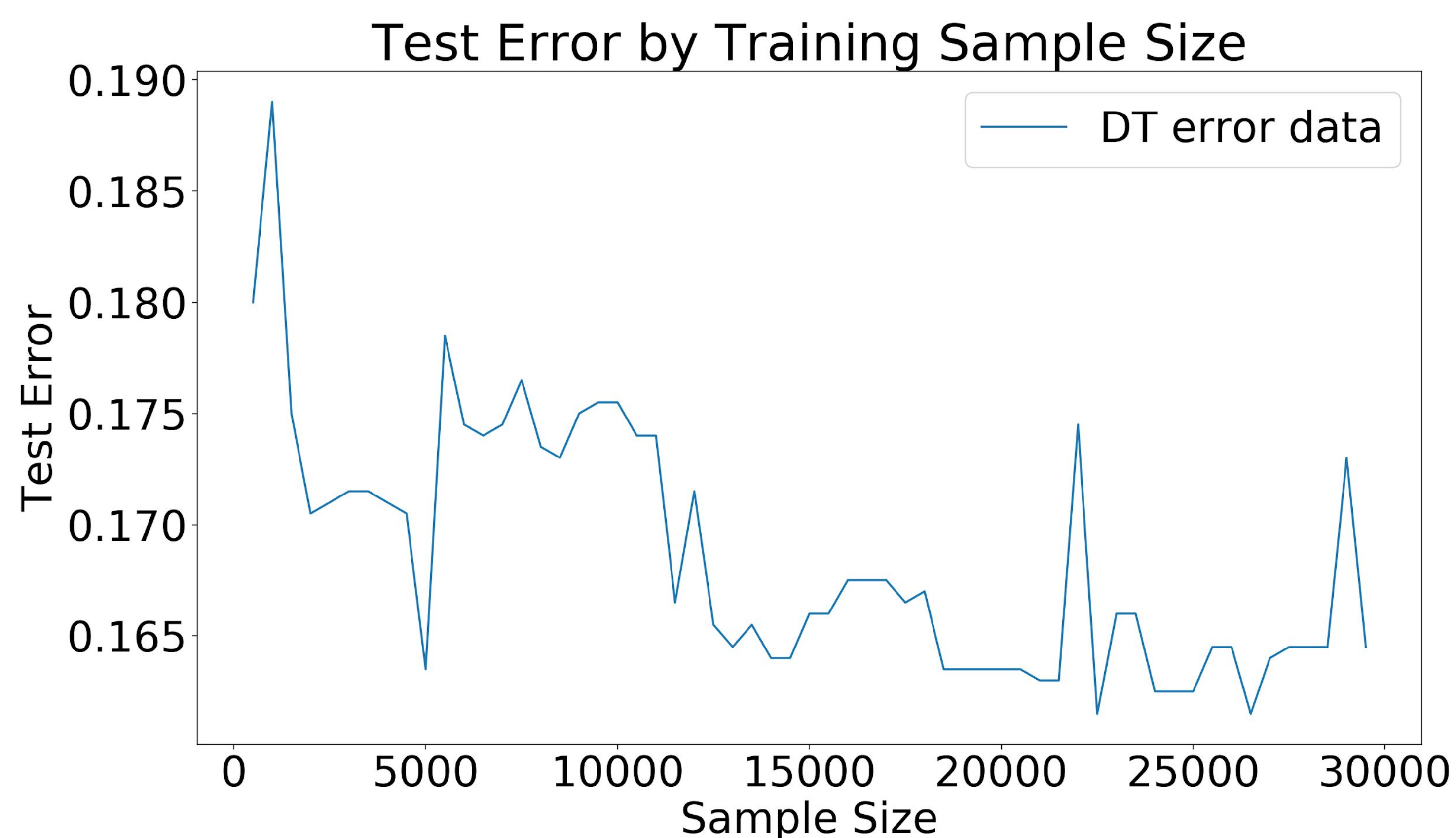


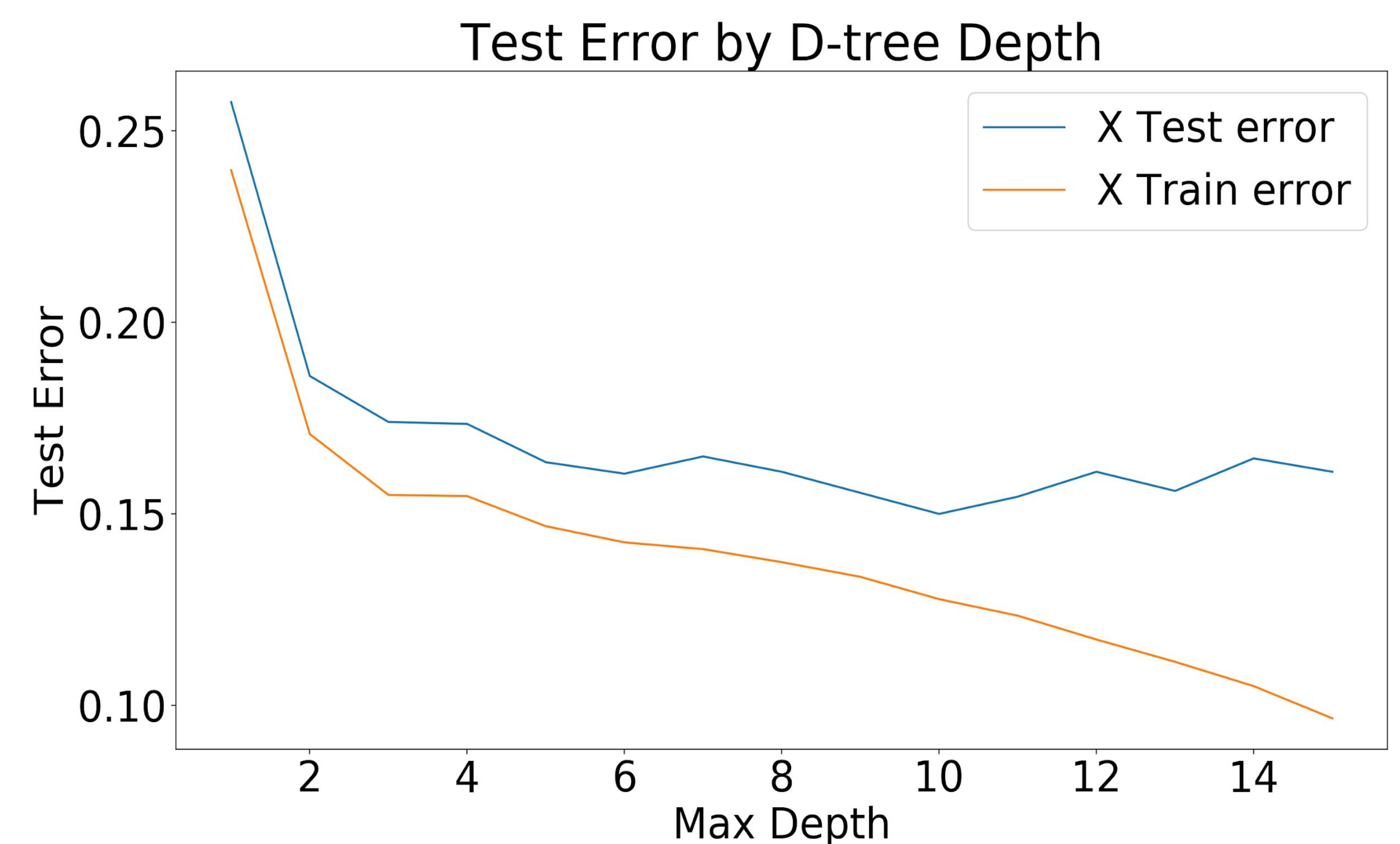
Introduction

Using the UCI adult dataset, a dataset scraped from the census bureau, our group sought to create the most accurate ML model to predict an individual's income based on their attributes. Our goal was to determine if our model exhibited bias.



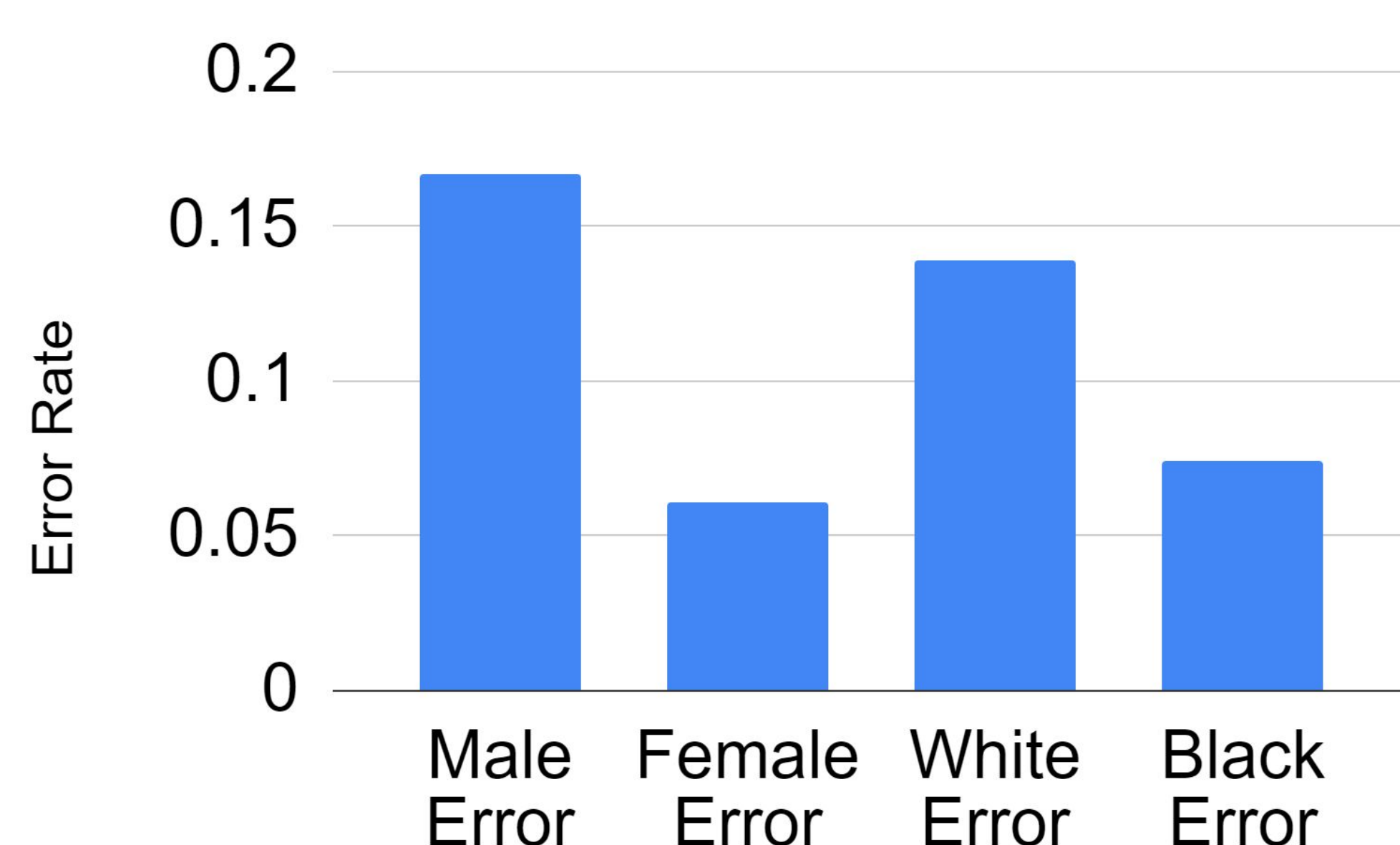
Our Process

We set out to create a ML model that had the lowest error on our test data. We found that the optimal training set size for our model was 22,500 and the optimal D-tree max depth was 10.



Error Discrepancies

When we calculated the error by features, we found interesting discrepancies. Surprisingly, our model illustrated that there were significantly lower error rates for black adults and women than there were for white adults and men.



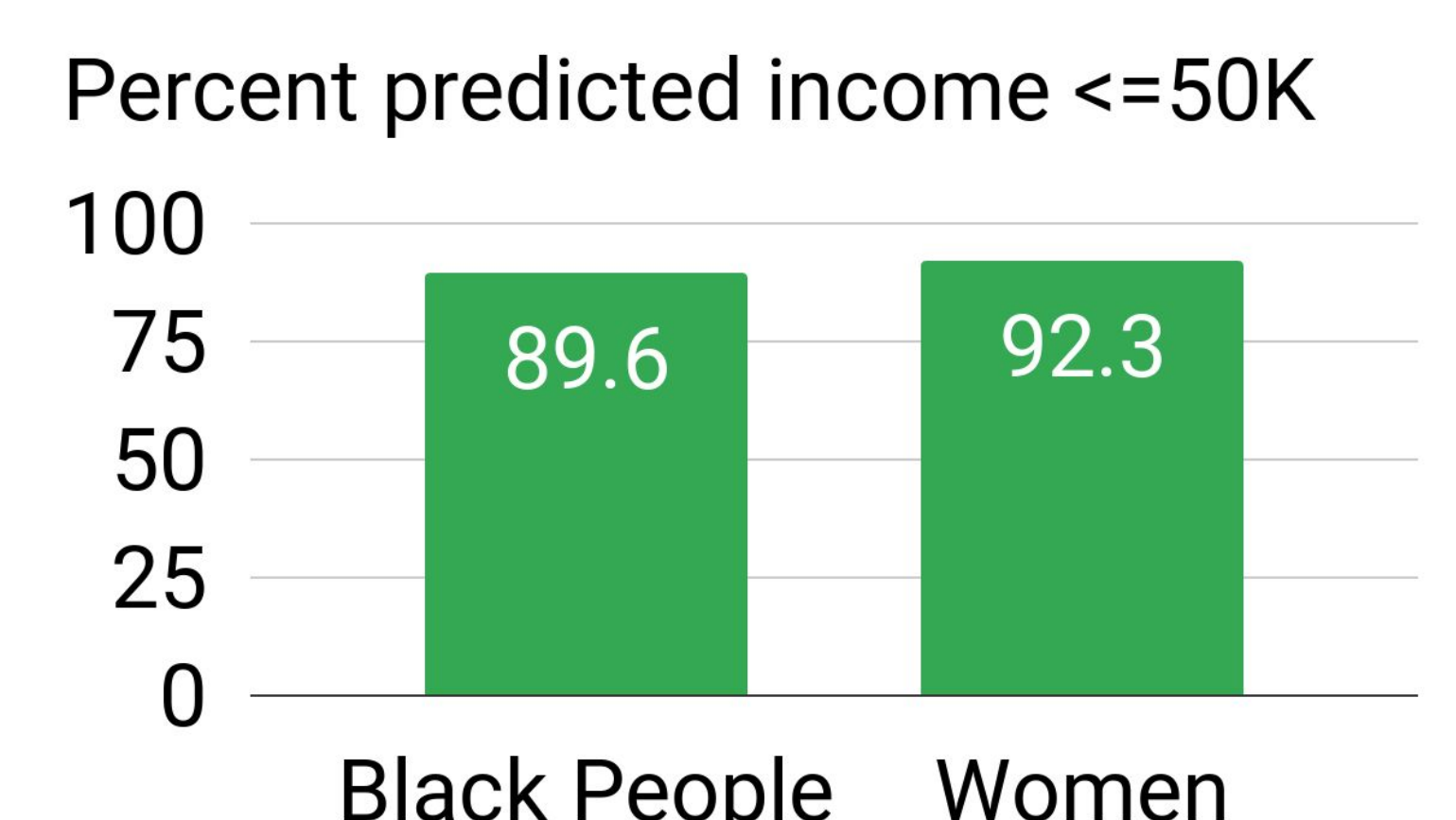
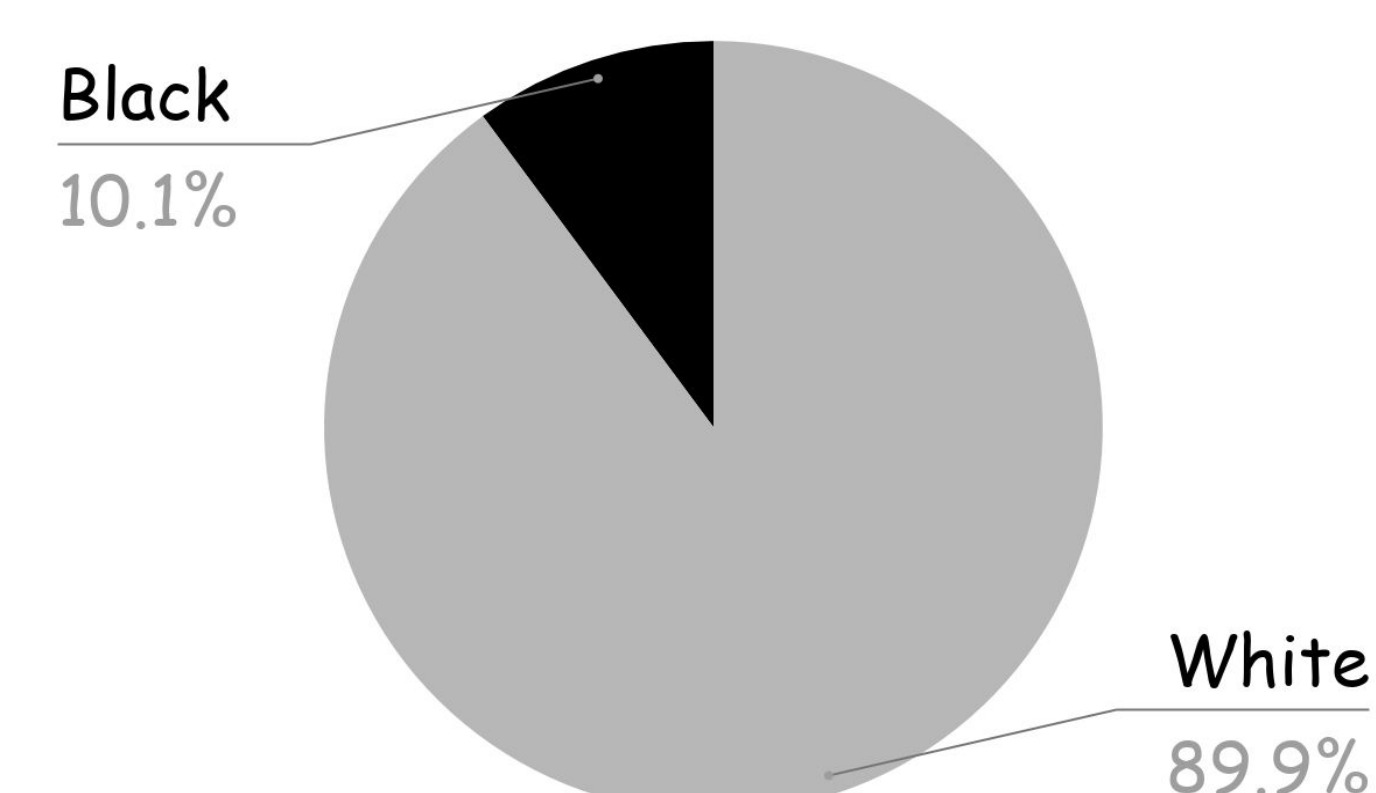
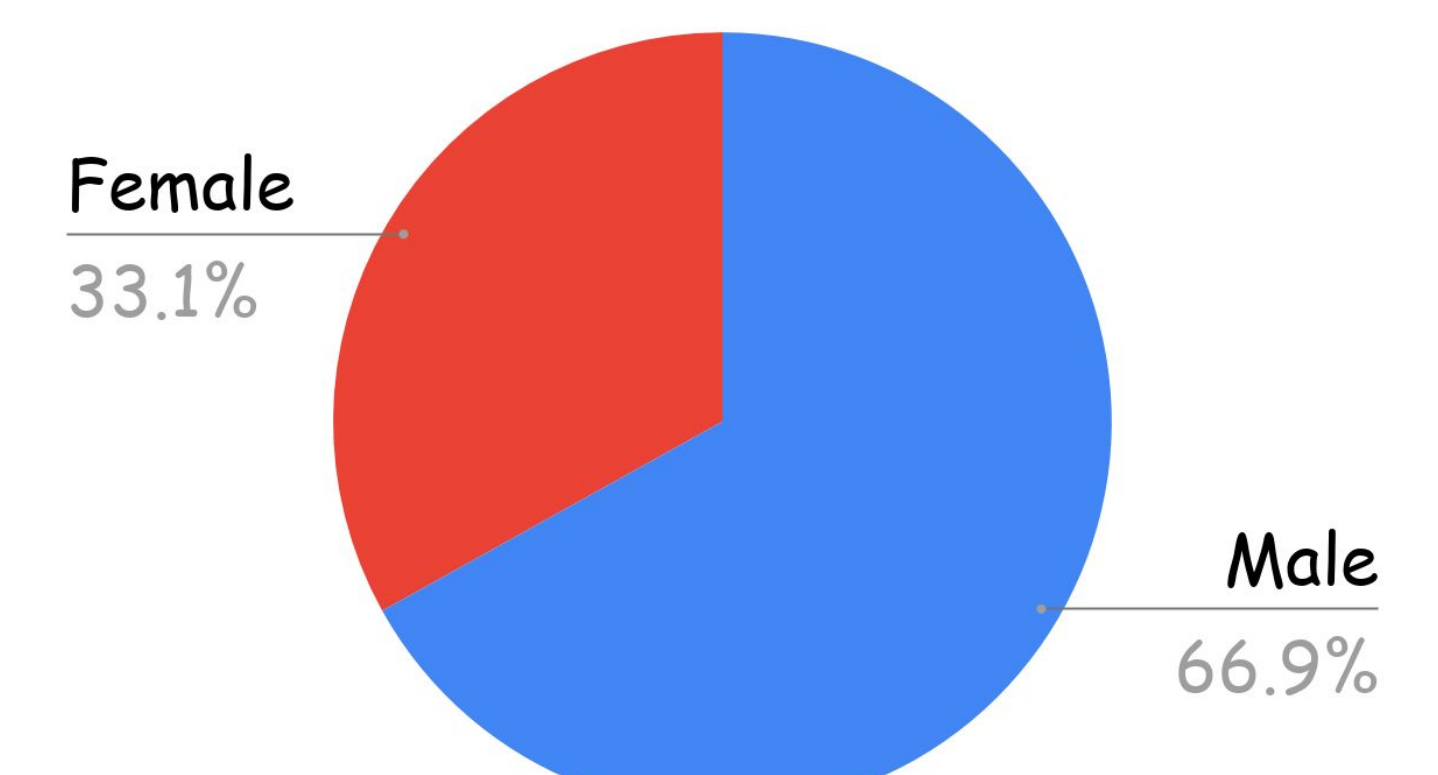
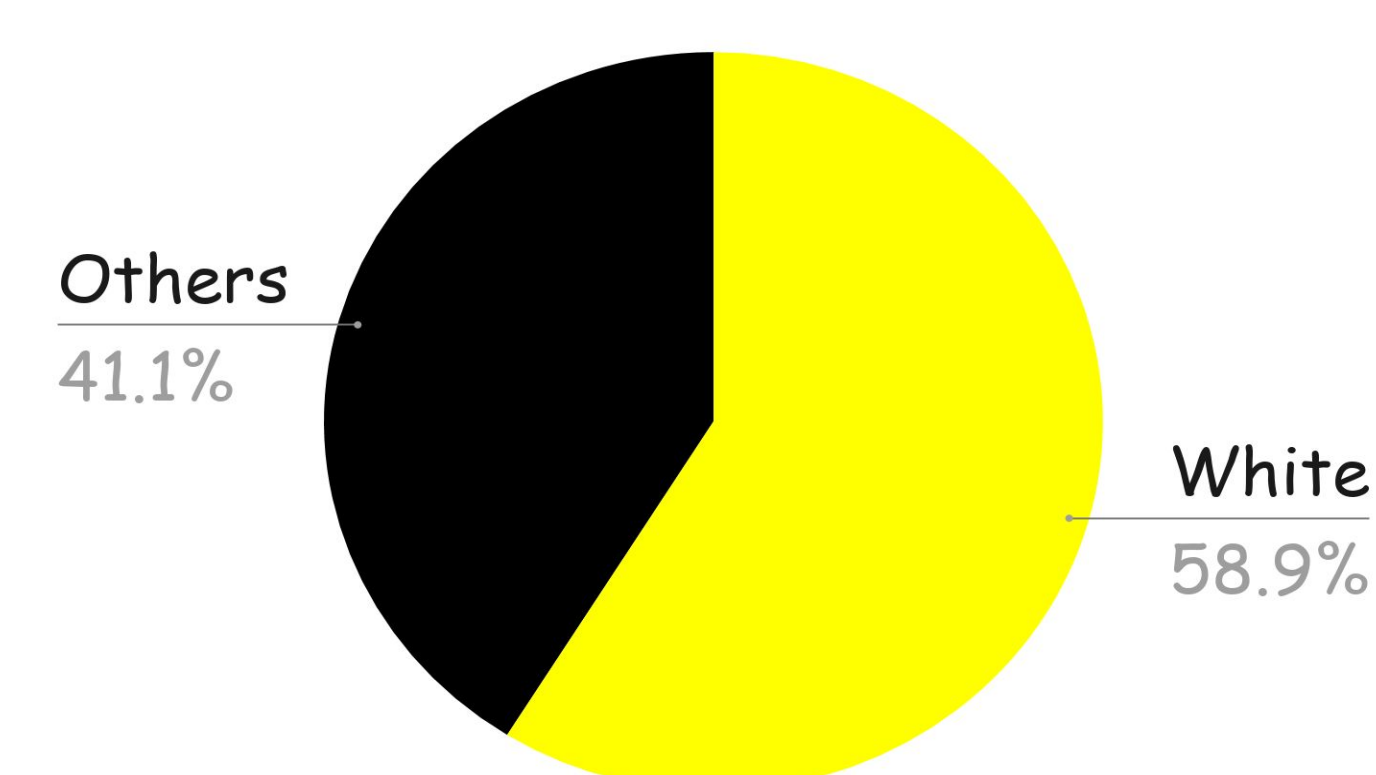
Hypotheses

From analyzing our errors, we hypothesized that either:

1. Our data set has very few entries for women and black adults.
2. Our model predicts black adults and women almost always make a certain income range either $>50k$ or $\leq 50k$

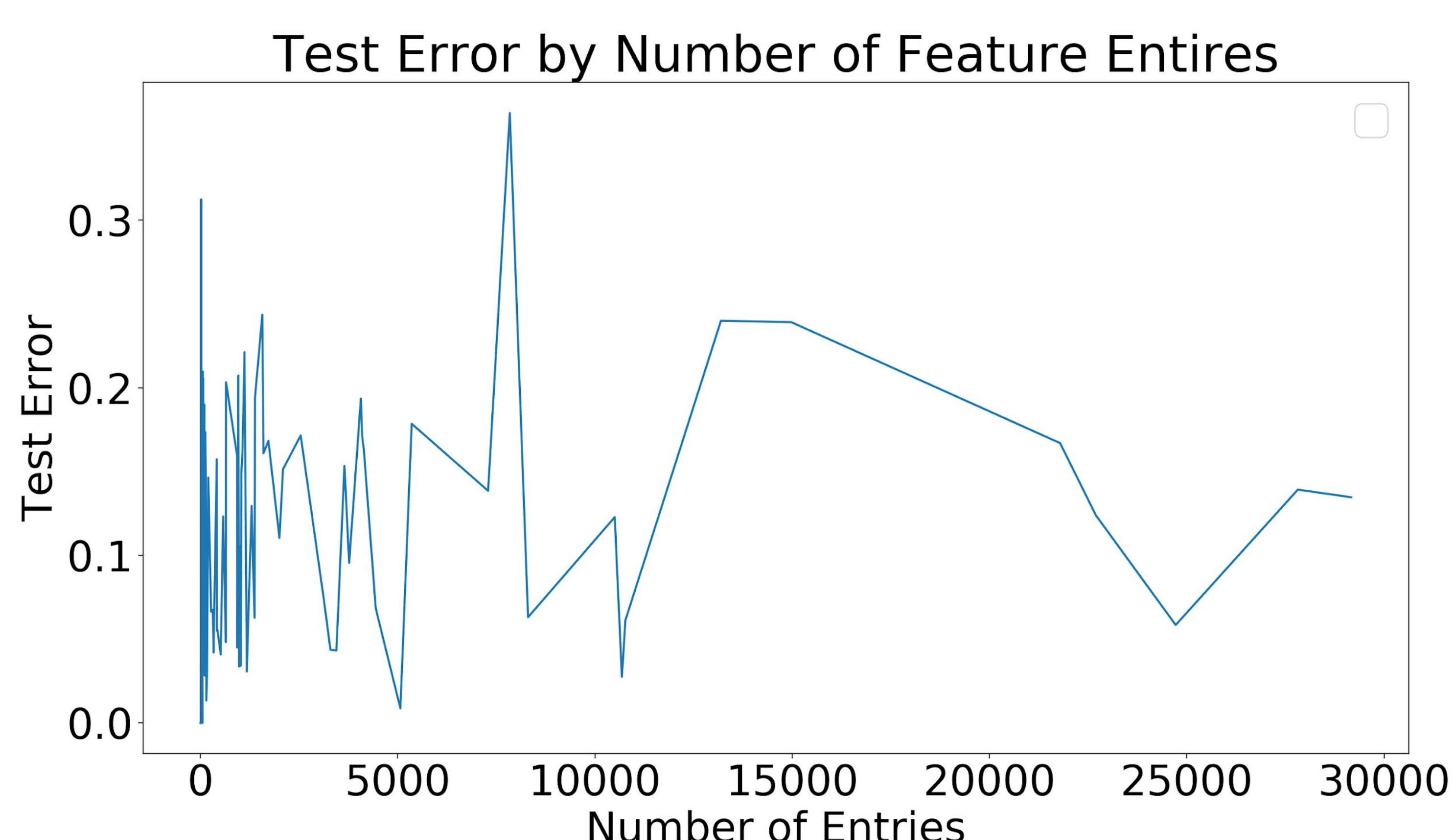
Data Imbalance

To better understand our model's bias, we looked into the breakdown of the census data by feature. The inequality in number of entries in the census data is astounding. Additionally, our model almost always predicts that women and black adults make $\leq 50k$.



Hypothesis Confirmation

To test our hypothesis, we graphed the test error by number of feature entries. The graph shows that for features in our dataset with a low number of entries, the errors vary greatly, due to our model not having enough data to determine patterns.



Conclusion

As you can see, the census data was overwhelmingly white men. When we trained a model on this data, the model unsurprisingly exhibited bias. The census is supposed to be an accurate representation of the population, and it is clearly not. ML models are only as good as their data, and even 'unbiased' data always has bias.