Introduction

Modern retail transactions are widely dependent on the acceptance of credit cards. Banks issuing the credit cards earn revenue both from any annual fees on the cards paid by the customer and a portion of each transaction value paid by the retailer. Therefore, a bank will earn more revenue if customers accrue a high value in transactions (either more purchases, or more expensive purchases, or both) and are able to pay off any balance on their cards on time. It is important for the bank to find such potential customers and offer their credit services to them. One way that banks do this is by directly sending mail to addresses with a description of their services and how they can be availed.

In this direct mail advertising, businesses send physical mail to every address in a particular region. However, the conversion rate of a person receiving mail into a paying customer is very low. Therefore, a business will prefer getting the best possible customers for the funds being spent in the effort. In order to attract better customers, it is beneficial to target those with more disposable income to spend. As we cannot accurately predict the exact budgets of each and every household in a region, we can use the median household income as a proxy for this measure. It is important to use the median rather than the mean household income as it is more representative of the entire population given that it is not influenced by the outliers that may exist in the region.

However, the problem with this approach is that it automatically excludes any region for which data is not readily available. This can mean new communities and regions that have not been surveyed for a long period of time will likely be not advertised to. In addition, it is quite expensive to survey entire regions for data about their incomes. This results in banks missing out on a considerable amount of revenue and large portions of the population remaining underserved in terms of credit services. To overcome this hurdle, we can link higher median incomes to the types of businesses that open and operate in a given region. This allows us to make approximations about the expected median income of a given region without having to rely on expensive surveys.