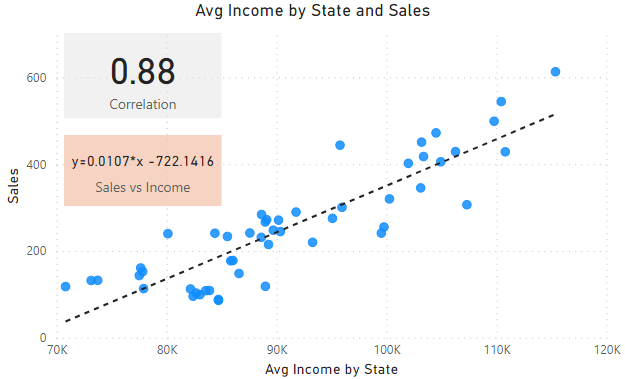
Project Report

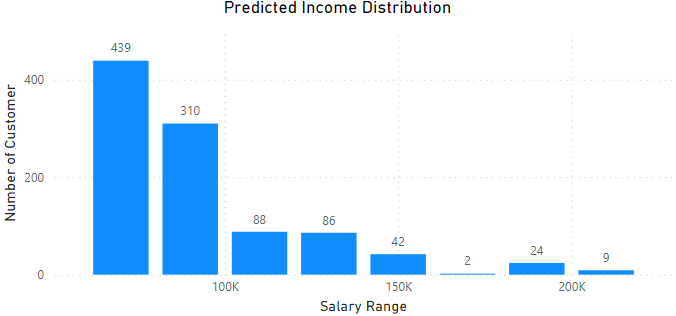
1. **Average sales and Average Income by State**

It assumes that the average sales are the average purchasing values of provided customers in past six months in each state. Creating a scatter plot between average sales and income, we can see a highly positive correlation between income and sales, the square of correlation is at 0.78. It means that state with high average income often spend more on company’s products. So the sales team should put higher focus on rich states such as District of Columbia, Maryland, Washington and California…A Linear regression indicates that if income increase by 1000 dollars then customers could have an extra spend at about 11 dollars.



1. **Recommended Products**

Based on regression model we developed in the previous part and purchasing history of customers, we can infer their incomes. Below is their predicted income distribution. Majority of customers have incomes that are below $100k. The customer with the highest predicted income is John Little, his predicted income is at about $558144



Customer incomes will impact on their choices of products. High value products tend to be selected by rich customers while lower income residents could just afford for cheap or moderate products. Base on this, we can recommend “Shirt” for customers with income below $85k, “Sweater” for people with income between $85k and $160k, “Leather Bag” for customers with income above $160k. We can see that Shirt is the most recommended product then follow by Sweater.



1. **Product Analysis**

The analysis on products shows that products with high ratings often have low return rate. It is quite right in practice, when customers satisfy about the products they rate higher star and then they also less likely return products. Some products with high ratings are Chronograph Watch, Leather Sneakers, Long Dress, Polo Shirt…

