One Smart Cookie

The goals:

- Analyzing sales, trends, and additional factors to predict future outcomes for cookie sales.
- Studying the ways that COVID has affected the amount of sales and the methods in which people purchase their cookies.
- Looking into whether or not low-income and high-income areas have noticeable differences in sales and cookie preference.
- Seeing whether or not the income of an area will predict the level of cookie sales.

Machine Learning Inquiries:

Can median income of an area predict cookie sales?

Simple linear regression using median income (as is) as a predictor for overall individual sales. This could also be expanded and applied by year.

Can cookie type sales predict income of an area?

Bucket income into 'low' and 'high' in order to perform Logistic Regression by using the cookie types as predictors.

Mapping the Data

For the D3 Leaflet:

We will be displaying Mecklenburg median housing income over the various zip codes. In these areas, we'll use a bubble map to show the top sales, with the size of the bubbles representing the sales volume. Utilizing the pop-ups, we will be displaying the name of each county with their zip codes, their average income, and their most popular cookie.

Charts for Additional Data

Was there a noticeable difference in sales volume from 2019-2022?

To compare overall sales per year, we will use a line graph and overlay each year for comparison, making it an interactive chart.

-X-Axis will be the year, Y-Axis will be the Sales Amount.

<u>Was there a large difference in sales when comparing areas with high income vs low income?</u>

For this, we will use a bar chart to measure the highest sales amount for each area. We will also compare the highest vs lowest sales areas in another bar chart to see what the biggest gap in sales is between the two.

-X-Axis will be the Area, Y-Axis will be Sales Amount. One color for each year.

<u>Was there a noticeable change in shipped/girl delivered/in hand orders from 2019-2022?</u>

(Are we able to predict the trend for next year?)

Here we will use a stacked bar chart to show each year and their number of orders as per method of delivery.

-X-Axis Years, Y-Axis Number of orders, Bars/colors – Order Method

<u>What was the highest selling cookie per year?</u>

As a side comparison, we will make a pie chart for the specific cookies and how well they each sold every year.

-Three Pie Charts, one per year, with cookie types represented by their color theme. Will include the ability to hover and see the exact percentage for each wedge.

<u>Was there a difference in cookie popularity per area each year?</u>

We will have a chart for top cookie in each area as well, using a bubble chart for each year.

 -X-Axis will be cookie type, Y-Axis will be Amount Sold, and the Bubbles will represent areas.