Brainstorming

Datasets Considered

Shopping Behaviour Dataset:

The dataset captures customer shopping behaviour across different demographics and purchase attributes. It includes information about items purchased, spending, customer background, and shopping preferences. The data helps analyse consumer trends, purchasing patterns, and seasonal behaviour. It is suitable for visualization, segmentation, and predictive modelling.

Chocolate Sales Data:

This dataset contains detailed records of chocolate sales, including product details, sales quantities, revenue, and customer segments. It is designed for sales forecasting, trend analysis, and business intelligence, helping businesses optimize pricing strategies, inventory management, and customer targeting.

Worldwide Travel Cities (Chosen Dataset):

This dataset contains curated travel information for 560 cities across the globe, offering a rich blend of structured data and subjective experience-based ratings. It is designed to support applications in travel recommendation systems, climate analysis, tourism research, and trip planning.

Possible Visualisations

Create a tree-map plot, would be a great way to visualise which countries/continents are more populated.

We could build a brief, interactive app using streamlit. The app would take a few user inputs (like budget or chosen continent) and get a recommended travel city. It's very relevant to the dataset and would be a great way make use of the data.

Areas to Explore/Analyse

We can explore the different budgets in the budget column to see if they influence duration, length of trips.

Analyse the data to see if there are any patterns or correlations between the budget and the climate of the different cities.

Possible research questions:

- **Question 1 –** Are luxury travels more common with destinations with higher temperatures?
  Hypothesis 1 Higher destination temperature is higher for luxury packages than mid-range and budget
- **Question 2 –** How does personal budget level influence duration of trip?
  Hypothesis 2 Budget packages have shorter length of stay than mid-range or luxury packages

Feature Engineering Ideas

Possibly make a column about places near the equator, linking each city. Or we could use the already existing temperatures column instead to categorise places by their typical climate. The latter is easier to read, understand, because how we make the first idea tangible, what would be the values with the column.

Potential Target Audience

Our target audience can literally be anyone wants to travel, but is not sure about the particulars of the trip, such the country they want to visit, temperature they might want to