Kaggle Data Visualization Exercise

Exercise Overview

Objective: Create an interactive D3.js visualization using a simple Kaggle dataset

Time: Approximately 20 minutes

Difficulty: Beginner-Intermediate

Dataset: "Chocolate Sales Data" (alternative options provided)

Submit the prompt (text format) and code with the compression: dryjins@gmail.com

Step-by-Step Implementation with Structured Prompts

Step 1: Data Acquisition & Processing

Use this prompt to generate data preprocessing code:

```
"Help me preprocess this Chocolate Sales Dataset from Kaggle. I need to:
1. Load the CSV data using d3.csv()
2. Parse dates and convert sales values to numbers
3. Group sales by product category and month
4. Calculate the total sales for each category"
```

Step 2: Visual Mapping Design

Use this prompt to generate visualization design code:

```
"Design a D3.js visualization for chocolate sales data that:
1. Creates a stacked bar chart showing monthly sales
2. Uses different colors for each chocolate category
3. Includes a legend for product categories
4. Implements tooltip showing sales details on hover"
```

Step 3: Rendering & Interaction

Use this prompt to complete the visualization:

```
"Complete my D3.js visualization by adding:
1. Smooth transitions when filtering data
2. Responsive design to fit different screen sizes
3. A simple filter control to show/hide specific chocolate categories
4. Proper axes formatting with month names and sales values"
```

Evaluation Criteria

1. **Data Processing**: Correct loading and transformation
2. **Visual Representation**: Clear communication of patterns
3. **Interactivity**: Effective user interaction implementation
4. **Code Quality**: Well-structured and commented
5. **Optimization**: Efficient rendering

Alternative Datasets

If preferred, you may use these alternatives:

- "Mobiles Dataset (2025)"
- FIFA 18 Player Statistics
- Starbucks Locations Dataset

Extensions (If Time Permits)

- Link multiple visualizations using crossfilter.js
- Integrate additional chart types
- Add time-series analysis components