

Data Analysis Portfolio Project: Candy Sales Analysis

Project Brief

This portfolio project demonstrates end-to-end data analysis skills by exploring a fictional candy company's sales data. The analysis will include data cleaning, exploration, visualisation, business insights, and geospatial analysis. The aim is to extract meaningful insights that can inform strategic decisions across sales, logistics, and pricing.

Project Objective

Analyse candy sales data to answer key business questions and demonstrate core data analysis competencies. The project will provide a comprehensive look at sales performance, logistics efficiency, regional trends, pricing strategies, and seasonal patterns.

Data Sources

The data for this project is sourced from the Maven Analytics Data Playground. You can find the dataset here: [Maven Analytics - US Candy Distributor](#)

The following datasets are used in the project:

candy_sales.csv	Sales transaction records
candy_targets.csv	Sales targets by product/division
candy_factories.csv	Factory location and details
candy_products.csv	Product information including pricing
uszip.csv	US ZIP code geospatial data (coordinates, region)

Analysis & Business Questions

1. Sales Performance Analysis

- Question: Which divisions and products are generating the most revenue?
- Skills: Data aggregation, grouping, visualisations (bar charts, line graphs)

2. Factory to Customer Shipping Analysis

- Question: What is the average shipping distance from factories to customers?
- Skills: Data merging, geospatial analysis, distance calculations (using coordinates)

3. Regional Sales Trends

- Question: Which geographic regions are outperforming or underperforming in candy sales?
- Skills: Geographic analysis, heatmaps, regional comparisons

4. Seasonality & Sales Trends

- Question: Are there seasonal trends in candy sales throughout the year?
- Skills: Time series analysis, seasonal decomposition, trend visualisation