Driving Business Decisions Through Sales Data Visualization

IDENTIFYING PRODUCT, TEMPORAL, AND GEOGRAPHIC REVENUE PATTERNS

Project Overview

Recap of the Data, Goals, and Tasks

The dataset used in this project is a transnational data set which contains transactions occurring between 1/12/2010 and 9/12/2011 for a UK-based and registered online retail company. The company mainly sells unique all-occasion gifts. Many customers of the company are wholesalers.

The dataset includes the following attributes: invoice number, stock code, product description, quantity, invoice date, unit price, customer ID, and country. These variables provide a transactional view of product sales across regions and over time. Here is the **link to the dataset on Kaggle:** ⊕ E-Commerce Data

I elected to exclude the UK from the dataset and to look only at the top selling 5 countries after that in an effort to reduce clutter and focus attention.

The primary goal of this project is to explore the dataset to generate new business insights that can directly inform decisions around product strategy, sales forecasting, and regional performance. This is driven by three core questions:

1. **What are the top-selling products by total revenue (quantity × unit price)?**
   This question aims to surface the products that contribute most significantly to overall revenue, supporting inventory management and marketing prioritization.
2. **How do sales trend over time?**
   Understanding seasonality, spikes, and slumps enables better forecasting and performance analysis.
3. **Which countries bring in the most revenue?**
   Identifying regional performance helps guide logistics and market targeting.

To address these questions, the project focuses on the following tasks derived from asking **how** the users are going to interact with the data and **what** does a task seek to learn about the data:

- **Aggregating and ranking total revenue per product**, allowing users to explore individual product performance using interactive visualizations (e.g., tooltips or highlighting).
- **Grouping total sales by time periods** (day, week, month) and visualizing trends using dynamic line charts with zoom/pan features.

- **Comparing total sales across countries**, with filters and interactions that enable exploration of country-specific sales information.

These tasks are guided by high-level analytic goals: recognizing consistent top-performing products, identifying temporal patterns in sales activity, and evaluating geographic distribution of revenue. Each of these insights informs specific design decisions for the visualizations, such as chart types, interactivity, and user navigation features.

Visualization Implementation

The visualization was implemented using Altair (a Python library) and Jupyter Notebooks. A screenshot of the final visualization is provided below.
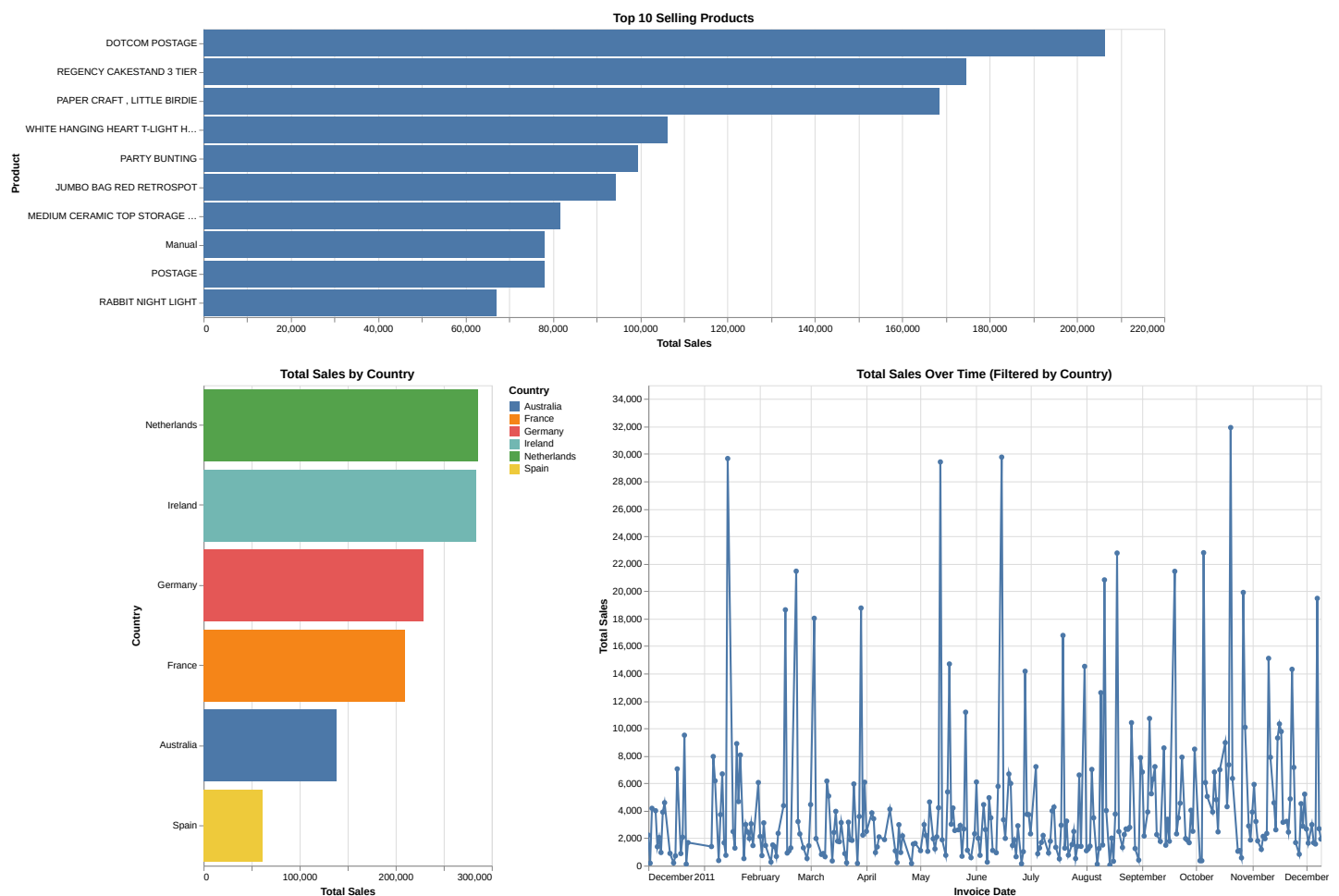
Alternatively, the interactive visualization can be accessed at the following URL:
https://lbeeler1.github.io/altair-sales-insights-dashboard/Visualizations/dashboard.html

Design Rationale

My design focuses on answering three core business questions: identifying top-selling products, understanding sales trends over time, and analyzing revenue by country. To address these goals effectively, I selected three visualizations tailored to the type of data and the insight each question seeks to provide:

1. **Bar Chart of Total Revenue by Product**
   I used a horizontal bar chart to display products ranked by total revenue (calculated as quantity × unit price). This format allows users to quickly compare performance across products. Interactive tooltips enhance usability by revealing specific revenue values and product details, making it easier to identify top performers.
2. **Line Chart of Sales Over Time**
   To visualize trends, I used a scatter plot with a line chart with filters for different time intervals (daily, weekly, monthly). This helps users detect patterns such as seasonality, spikes, or downturns. Line charts are ideal for showing continuous changes over time, and zoom/pan features make exploration more flexible.
3. **Bar Chart of Revenue by Country**
   A horizontal bar chart shows total revenue aggregated by country. I chose a uniform color for all bars to emphasize comparison rather than categorical differences. This visualization highlights regional performance and market distribution, and users can hover over each country to get specific figures as well as explore sales over time for subsets of countries through interactivity between this chart and the sales over time chart.

All visualizations follow consistent design principles: clear axis labels, sorted categories for easy scanning, and minimalist strategic color schemes to maintain focus. These choices support intuitive navigation, accurate interpretation, and alignment with the tasks identified in the planning phase.

Evaluation Methodology

Evaluation Type: Insight-Based (Qualitative)

This evaluation will assess whether users can gain useful insights from the visualizations. The focus is on evaluating how effectively the visualizations communicate key patterns in the dataset, including sales performance, temporal trends, and geographic comparisons.

Goals

Assess whether users can:

1. Identify top-selling products.
2. Understand revenue trends over time, including drill-down capabilities.
3. Compare total revenue by country.

Participants

The evaluation will include 2 participants who are familiar with basic data analysis and interpretation.

Procedure

1. **Introduction**
   Each participant will be introduced to the dataset and visualizations. They will be told that the visualizations represent one year of retail transaction data and will be given specific tasks to complete.
2. **Presentation**
   Participants will be shown the interactive visualizations created using Altair in Python.
3. **Tasks**
   Participants will be asked to complete the following tasks using the visualizations:
   ○ Task 1: Identify which product generated the most revenue.
   ○ Task 2: Describe any trends in revenue over time.
   ○ Task 3: Determine which countries had the highest and lowest total revenue.
4. **Think-Aloud Protocol**
   Participants will be encouraged to think aloud as they interact with the visualizations—explaining what they are seeing, interpreting, or finding confusing.
5. **Post-Task Interview**
   After completing the tasks, participants will be asked follow-up questions:
   ○ What was easy or difficult about interpreting the visualizations?
   ○ Which chart was most helpful and why?
   ○ Is there anything you would improve?

Insight Capture Sheet

Task	Insight Observed	Accurate?	Time Taken	Notes on Behavior
Top Product	"Dotcom Postage"	Yes	5 sec	Located it quickly using the bar chart.
Revenue Trend	"The highest sales for the top 5 countries excluding the UK were on Oct 20, 2011)	Yes	43 sec	Paused at the drill-down interaction. The data itself is very up and down.
Country Revenue	"The Netherlands had the highest overall"	Yes	30 sec	Interpreted bar heights easily.

Insights

- Both participants were able to identify the top product quickly using the bar chart, suggesting that it communicated rankings effectively.
- Both participants hesitated when interpreting the time series visualization, indicating a need for more specific instructions.
- Participants successfully identified countries with high and low revenues.

Results

The results of the evaluation indicated that participants were able to accurately answer questions about the data using the visualization, the visualization was perceived as easy to use, and task completion times were within an acceptable range.

Synthesis of Findings + Conclusion

This project successfully highlighted key patterns in sales data by focusing on product performance, temporal trends, and geographic revenue distribution. The visualizations provided clear, actionable insights: top-selling products were easily identifiable through the ranked bar chart, the scatter + line chart revealed sales fluctuations and seasonal trends, and the country-based bar chart offered a straightforward comparison of regional performance.

What worked particularly well was the clarity and focus of the visualizations. By limiting the scope to three primary questions, each chart could be designed with intention and simplicity, helping avoid clutter and confusion. The use of interactive elements such as tooltips and time filters further enhanced the user experience by enabling deeper exploration without overwhelming the main visuals.

In future iterations, I would refine the design by experimenting with more dynamic dashboard layouts or adding interactivity to all charts in a more integrated way. I would also consider incorporating additional metrics like profit margins, business revenue goals with drill down, or customer segments to broaden the analytical value.