Sales and Customer Analysis for a UK Online Retailer

By Axel Aranda Núñez

Aspiring Data Analyst & Student of Industrial Organization Engineering

Index

1. Business Context & Objective

- Analyze customer and sales trends over one year.

- Identify top products and valuable customers.

- Support better decisions in marketing and inventory.

2. Prepare – Understanding the data

- Public dataset from Kaggle (Online Retail Dataset)

- 500k+ transactions | 2010–2011 | UK-based e-commerce.

- Fields: Invoice, Product, Quantity, Price, Customer, Country.

- Issues: missing values, negative quantities, vague descriptions.

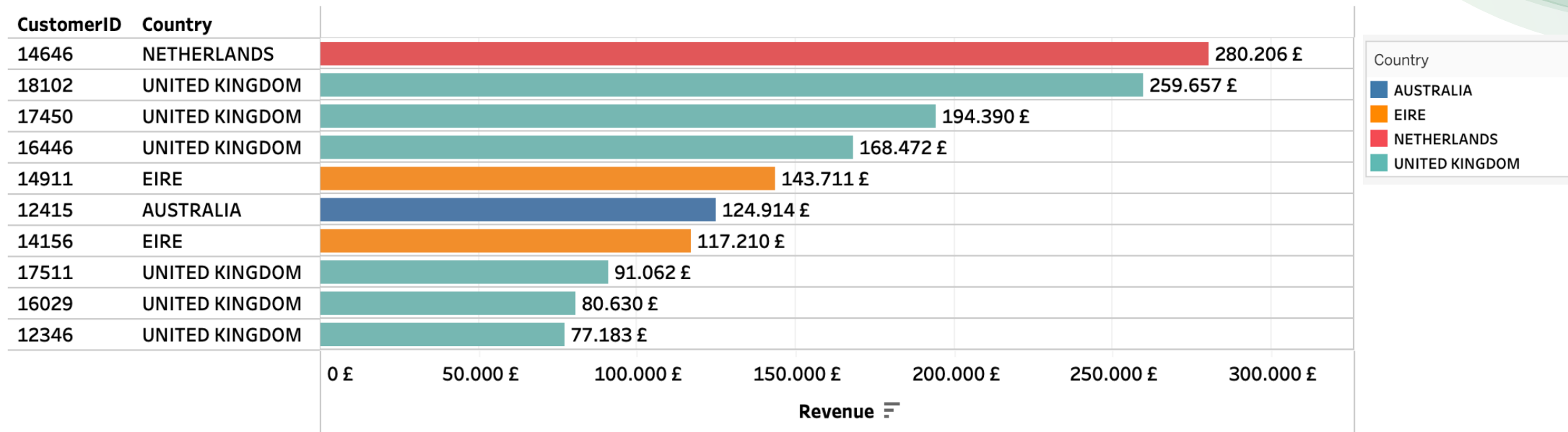- Meets ROCCC data quality principles.

3. Process – Data Cleaning & Preparation

- Removed nulls, duplicates, invalid entries.

- Classified transactions: Sales vs Returns.

- Standardized text fields (uppercase, trimmed).

- Created new columns:
  - Total Revenue
  - Year / Month / Weekday / Hour
  - Timeslot category

- Final cleaned dataset ready for análisis.

4. TOP 10 Customers – Revenue Contribution

TOP 10 Customers

CustomerID	Country	Revenue
14646	NETHERLANDS	280.206 £
18102	UNITED KINGDOM	259.657 £
17450	UNITED KINGDOM	194.390 £
16446	UNITED KINGDOM	168.472 £
14911	EIRE	143.711 £
12415	AUSTRALIA	124.914 £
14156	EIRE	117.210 £
17511	UNITED KINGDOM	91.062 £
16029	UNITED KINGDOM	80.630 £
12346	UNITED KINGDOM	77.183 £

Country

- AUSTRALIA
- EIRE
- NETHERLANDS
- UNITED KINGDOM

Revenue

- Top 10 customers = **18.6% of total revenue.**
- **6 out of 10** are based in the **United Kingdom.**
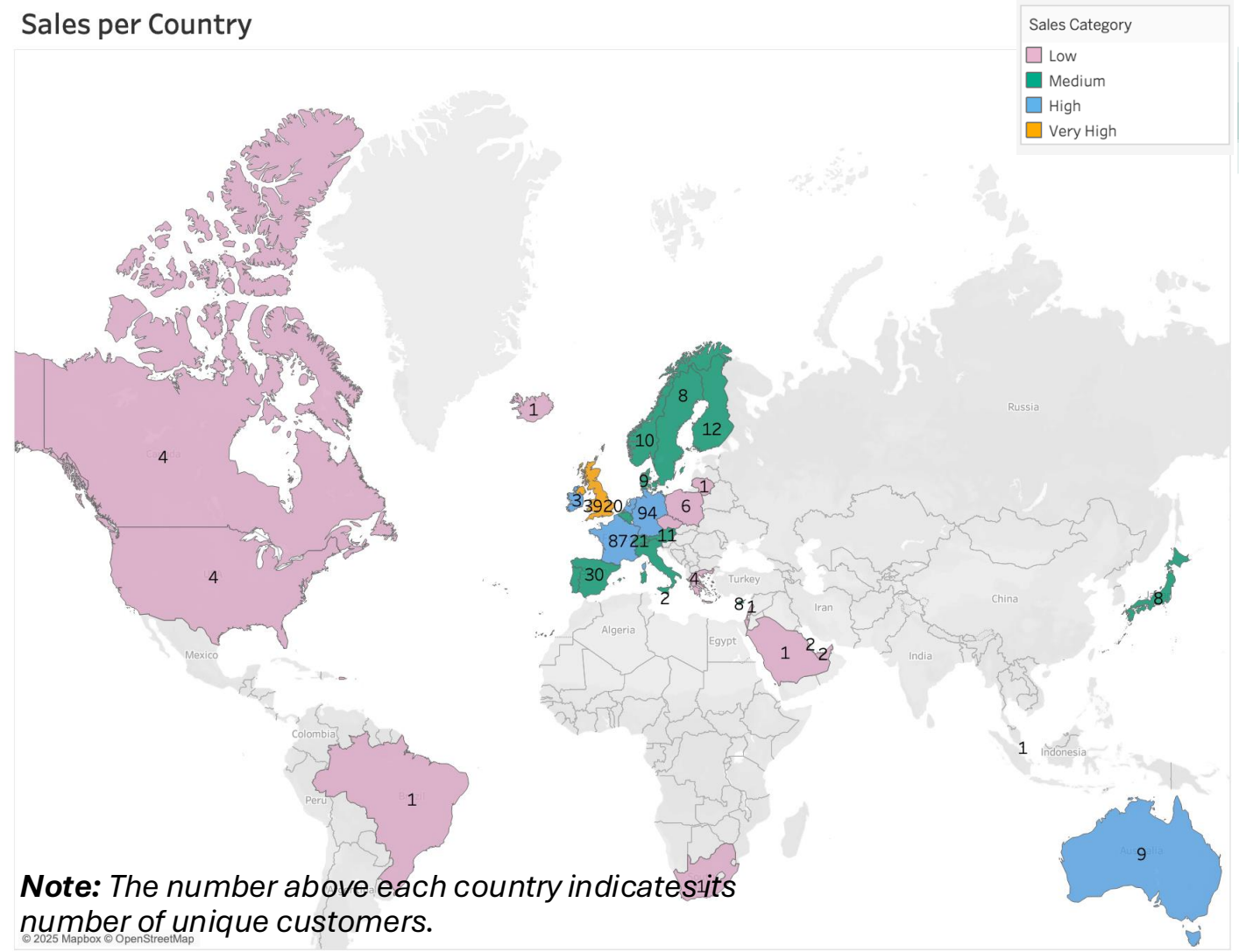- NL **Netherlands**, AU **Australia**, and IE **Ireland** also stand out.

5. Sales per Country – Geographic Distribution

- GB UK: most revenue and largest client base.

- NL Netherlands & IE Ireland: **top 2–3 in sales** with very few clients.

- 🌍 Fewer clients in DE Germany, FR France, ES Spain → lower average spend.

Sales per Country

Sales Category
- Low
- Medium
- High
- Very High

Note: *The number above each country indicates its number of unique customers.*

© 2025 Mapbox © OpenStreetMap
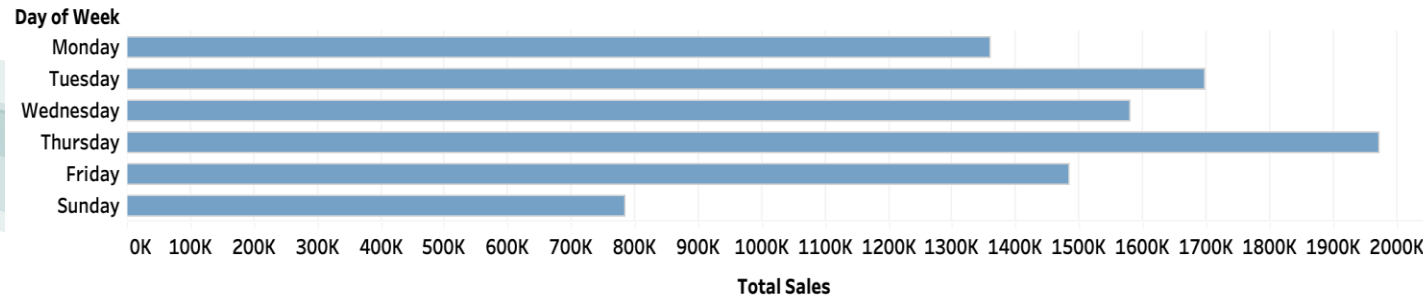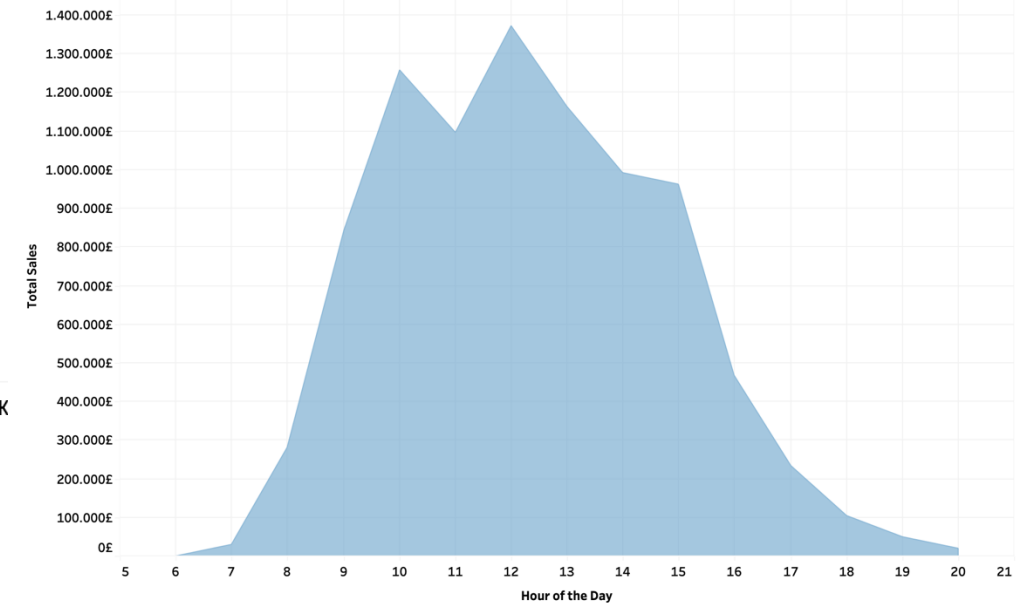
6. Sales Distribution Over Time

- Peak sales on Thursdays (23.1%) and Tuesdays (20.5%).
- 59% of purchases happen in the morning (10–13h).
- September–October = highest sales; February and April = lowest.

Sales per Month (abs)
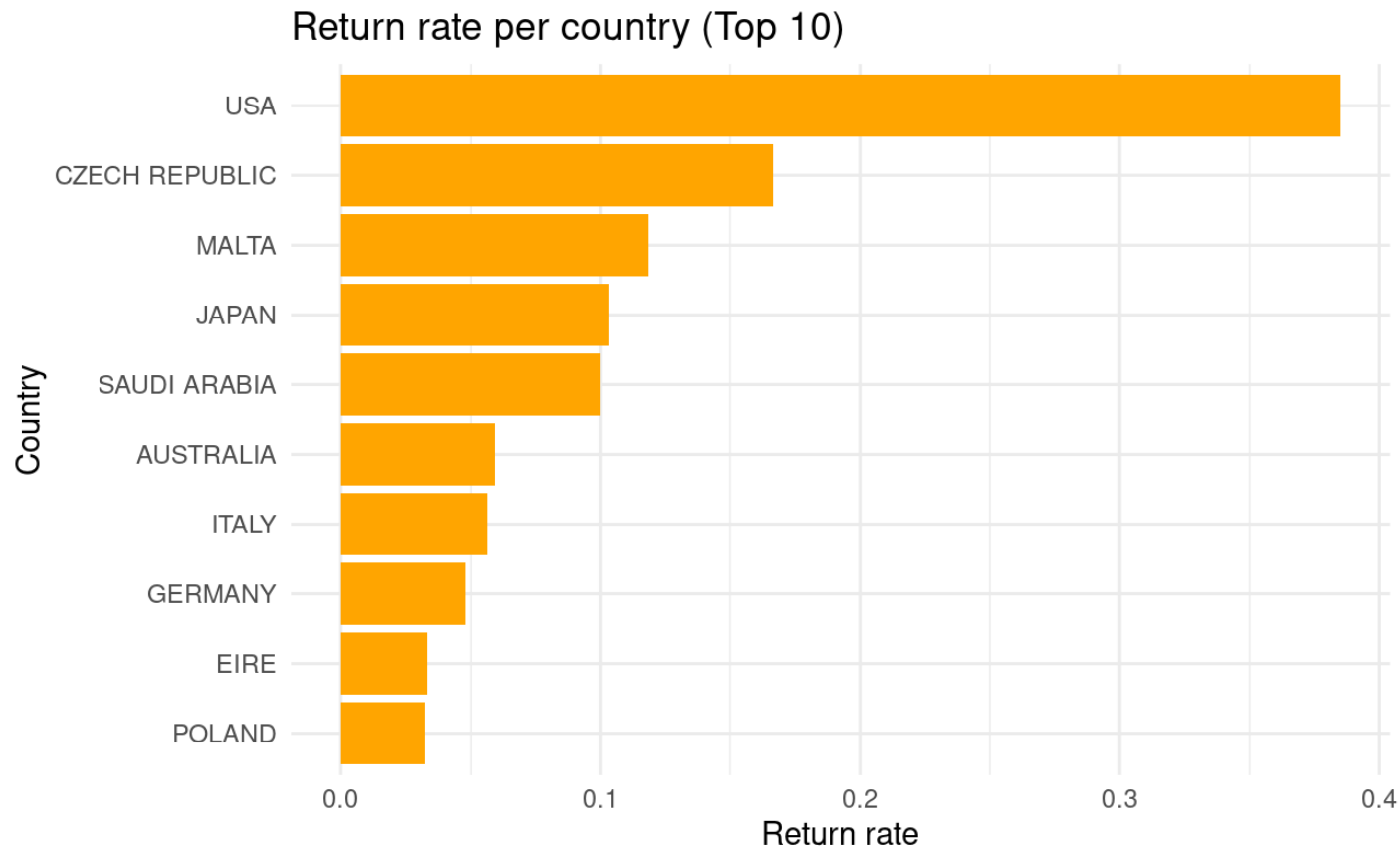

Sales per Day of Week
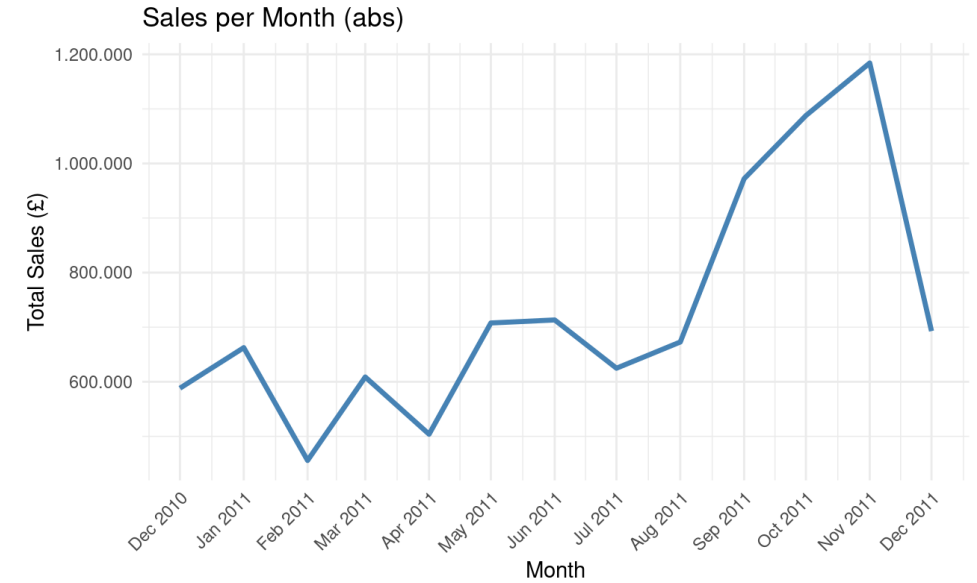
Sales per Hour of the Day

7. Return Rate by Country

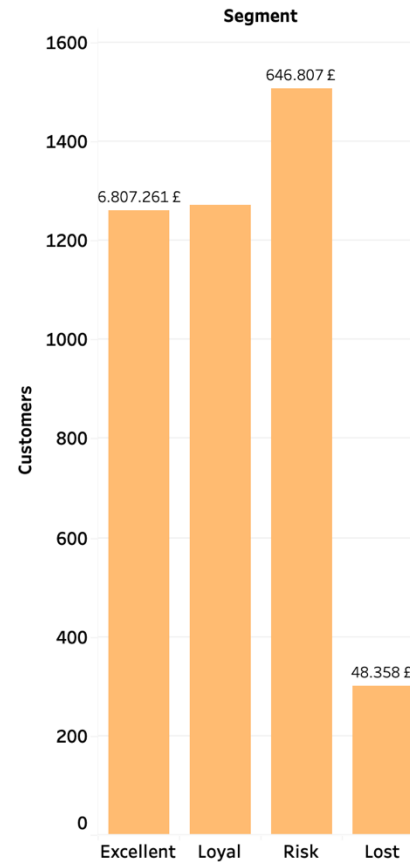Return rate per country (Top 10)

- Overall return rate: 2.2%.

- USA leads returns (84.6%) with rate of 38.5%.
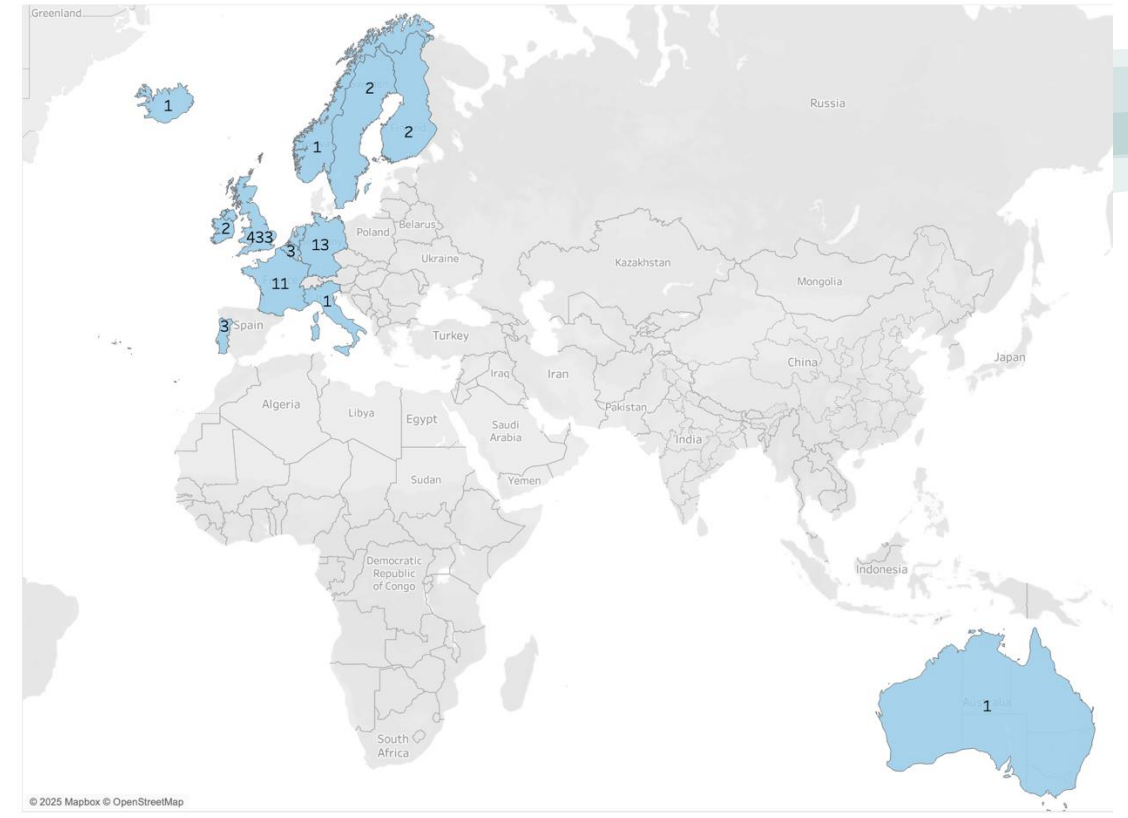
- 19 products were returned but never sold.

8. Customer Segmentation– RFM Analysis

- Top-tier customers (RFM = 12) make up 10.9% of clients.

- They generate 49.3% of total revenue.

- Mostly located in the UK (433), Germany (13), and France (11).

Customers Distribution by RFM Segment

Geographic Distribution of TOP RFM (SCORE 12/12)

9. Key Performance Indicators (KPIs)

TOTAL SALES
8.280.356 £

Numbers of Customers RFM 12/12
474

Number Unique of Customers
4371

Return Rate
2.20%

10. Dashboard Overview

TOTAL SALES
8.280.356 £

Number Unique of Customers
4371

Numbers of Customers RFM 12/12
474

Return Rate
2.20%

TOP 10 Customers

| CustomerID	Country	Revenue
14646	NETHERLANDS	280.206 £
18102	UNITED KINGDOM	259.657 £
17450	UNITED KINGDOM	194.390 £
16446	UNITED KINGDOM	168.472 £
14911	EIRE	143.711 £
12415	AUSTRALIA	124.914 £
14156	EIRE	117.210 £
17511	UNITED KINGDOM	91.062 £
16029	UNITED KINGDOM	80.630 £
12346	UNITED KINGDOM	77.183 £

Customers per Country

The total customers are displayed on a logarithmic scale to better visualize differences between countries with few and many customers.
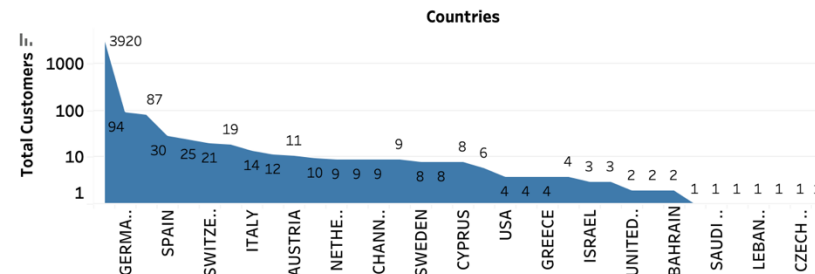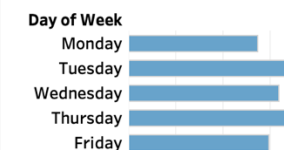
Customers Distribution by RFM Segment
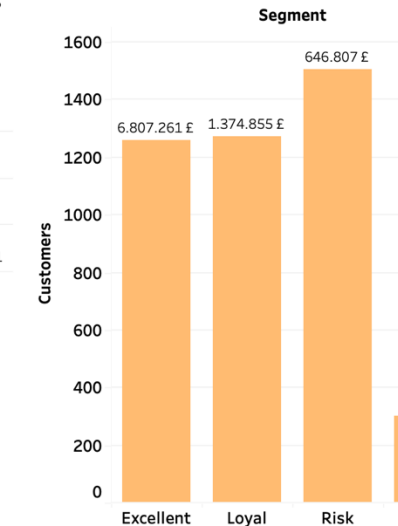
Sales per Country
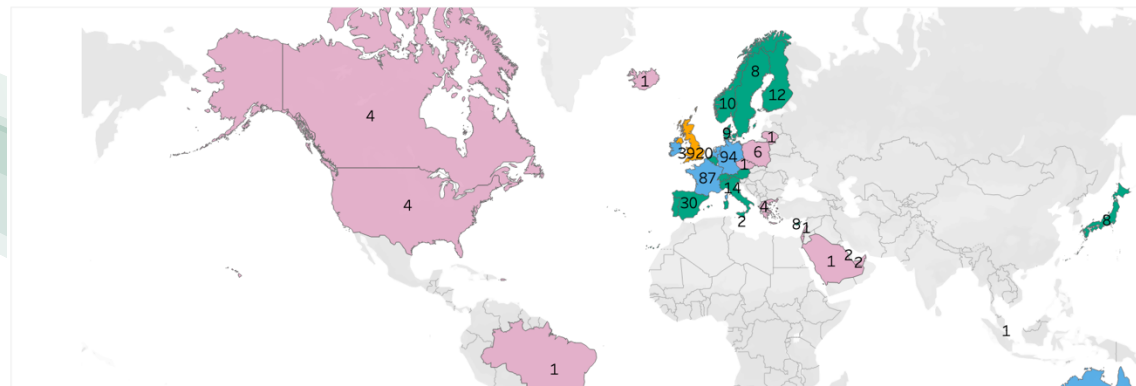
Sales Category: Low, Medium, High, Very High

Sales per Day of Week

Day of Week: Monday, Tuesday, Wednesday, Thursday, Friday

Sales per Hour of t

11. Recommendations & Next Steps

- Focus on retaining high-value customers (RFM 12/12).

- Investigate high return rate in the U.S.

- Reinforce midweek promotions (especially Tue–Thu).

- Target top-performing countries outside the UK.

- Monitor and optimize key time slots (10h–13h).

12. Tools & Methodology

- **SQL (BigQuery):** Data cleaning, transformation, deduplication, column creation.

- **R:** Data analysis, custom visualizations, time features.

- **Tableau:** Dashboards, interactive maps, bar & area charts.

- **PowerPoint:** Presentation design.

- **Dataset:** Online Retail, Kaggle
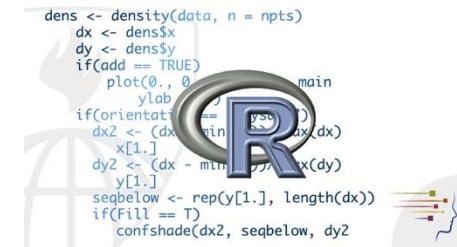
13. Conclusions

- Top 10 customers contribute significantly, representing 18.57% of total sales.

- United Kingdom dominates sales volume, but valuable customers also come from Netherlands, Ireland, and Australia.

- Sales peak on Thursdays and Tuesdays, with mornings being the busiest time of day.

- Return rate is low overall (2.2%), but USA shows unusually high return activity.

- RFM analysis reveals a small segment (~11%) of top customers driving nearly 50% of revenue.

- Geographic distribution highlights opportunities for targeted marketing in high-value regions.