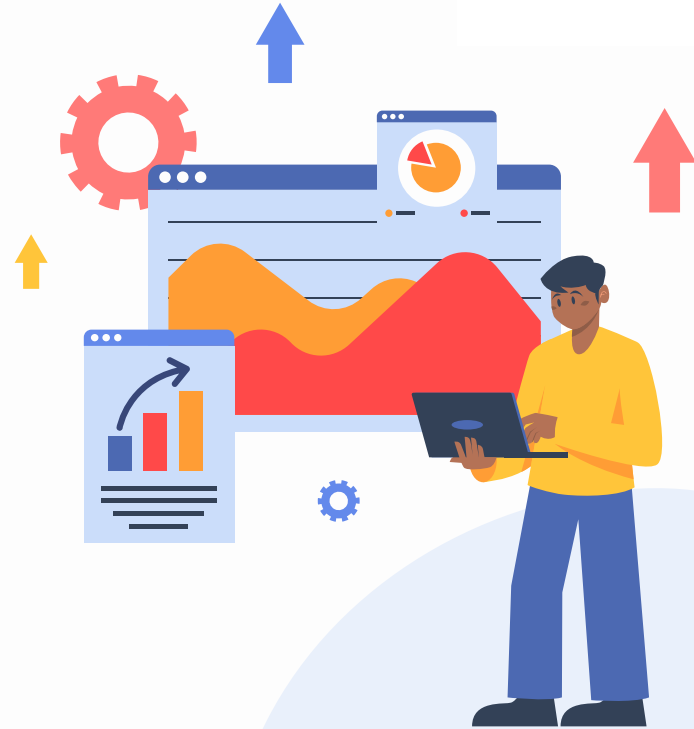


Data Analysis - Case Study



RAPP 

Case Study

Client

The client is an online retailer seeking to adopt a data-driven approach to understand their business and customers better.

Resources

Two Datasets are provided, being Transactional data & Customer data.



Objectives & Overview

Understand Customer Behaviour & Making the Data Accessible

Analyse transactional and demographic data to uncover patterns and insights.

Ensuring the data is accessible for continuous analysts for business decisions by data aggregation and analysis.



01

Product Analyst

Understand the products and segment customers based on demographics and purchasing behaviour.

02

Evaluate Sale-Trends

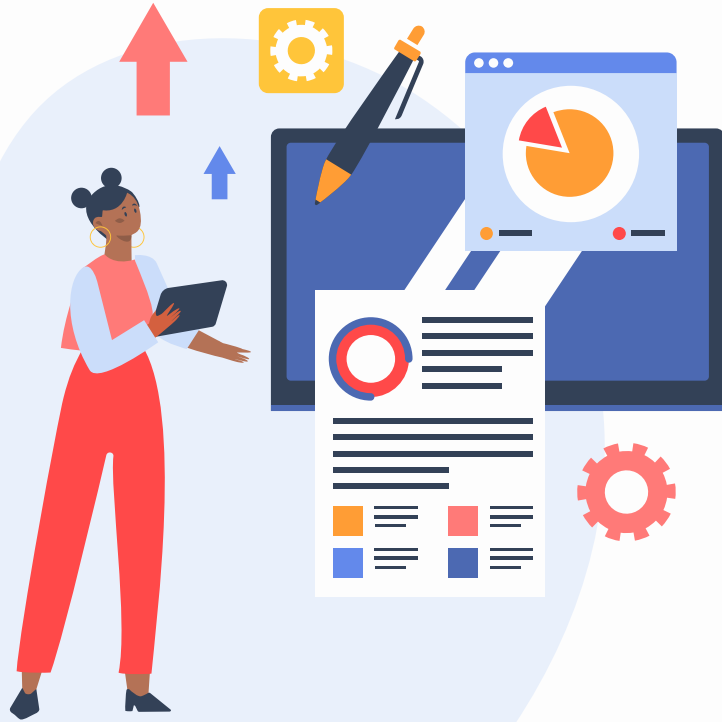
Examine sales trends over time to identify seasonality and sales peaks.

03

Data-driven business decisions.

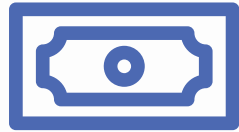
Offer data-driven strategies to improve commercial performance and customer satisfaction.





Data Understanding

Datasets Used



Transactions

Details of purchases including product ID, purchase date, value, and currency.



Customers

Demographic details including age group, gender, and membership length.



Datasets

Customers

Buyer_id	A unique identifier for each buyer (customer), which links to the 'transaction_data'
gender	The gender of the customer.
membership_length	The length of time (in years) that the customer has been a member.
age_group	The age group of the customer (e.g., 18-25, 26-35).

Transactions

Purchase_id	A unique identifier for each transaction.
Buyer_id	A unique identifier for each buyer (customer).
country	The country where the transaction took place.
product_id	The country where the transaction took place.
department	The department to which the product belongs (e.g., Electronics, Clothing, Accessories).
category	The category of the product within the department.
purchase_date	The date when the transaction occurred.
currency	The currency used for the transaction.
value_of_item	The monetary value of the item purchased.



Data Overview

Transactions

A total of **217,977 transactions** recorded.

Involving **129,113** unique **customers**.

Date Range

Sales data spans from **01/08/2019** to **07/08/2019**.

Customer Demographics

Membership duration ranges from **0 - 14 years**.

Age groups vary from **18 to 60+**.

Gender based purchasing.

Product Categorisation

9 departments are represented.

These departments are further **divided into 15 categories**. There are **207,459** amount different products.

Geographical Distribution

Data includes transactions from **6 distinct countries**.

Transactions from **other regions** are grouped into a single category labeled "RoW", **Rest of the World**.



Methodology



Data Cleaning & Preparation

Data Engineering

Data Merging & Currency
Conversion



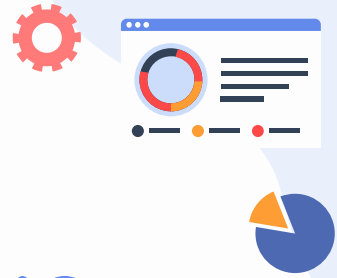
Descriptive Statistics

Summarise Data & Uncover Key
Insights



Detailed Analyst

Segmentation, Basket & Timer
Series Analyst



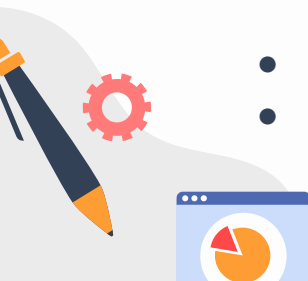
Extract, Transform, Load



01 Data Cleaning

- Ensuring no missing values or duplicates
- Applying consistent data types across datasets.
- Formatting datasets correctly

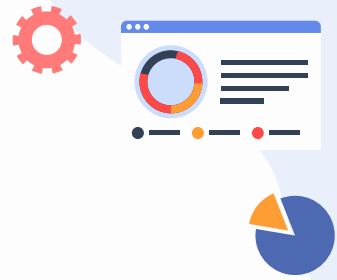
02 Data Engineering

- Apply conversion rates for consistency
 - Identify key columns for merging
 - Handling Missing Values
- 

03 Data Aggregation and Analysis

Aggregated detailed transactional data into data cubes for efficient summarisation and analysis by the tables main attributes.

- Organised the data by important attributes, including date, country, department, and product_id.



Questions I am trying to answer

which are the best-selling products, by region and department?

What key insights & observations can be drawn from the data?

Understand high and low performing demographics.

What extra information would be helpful

**What
Data-driven
business
decisions can be
made?**





Understanding Customer Behaviour

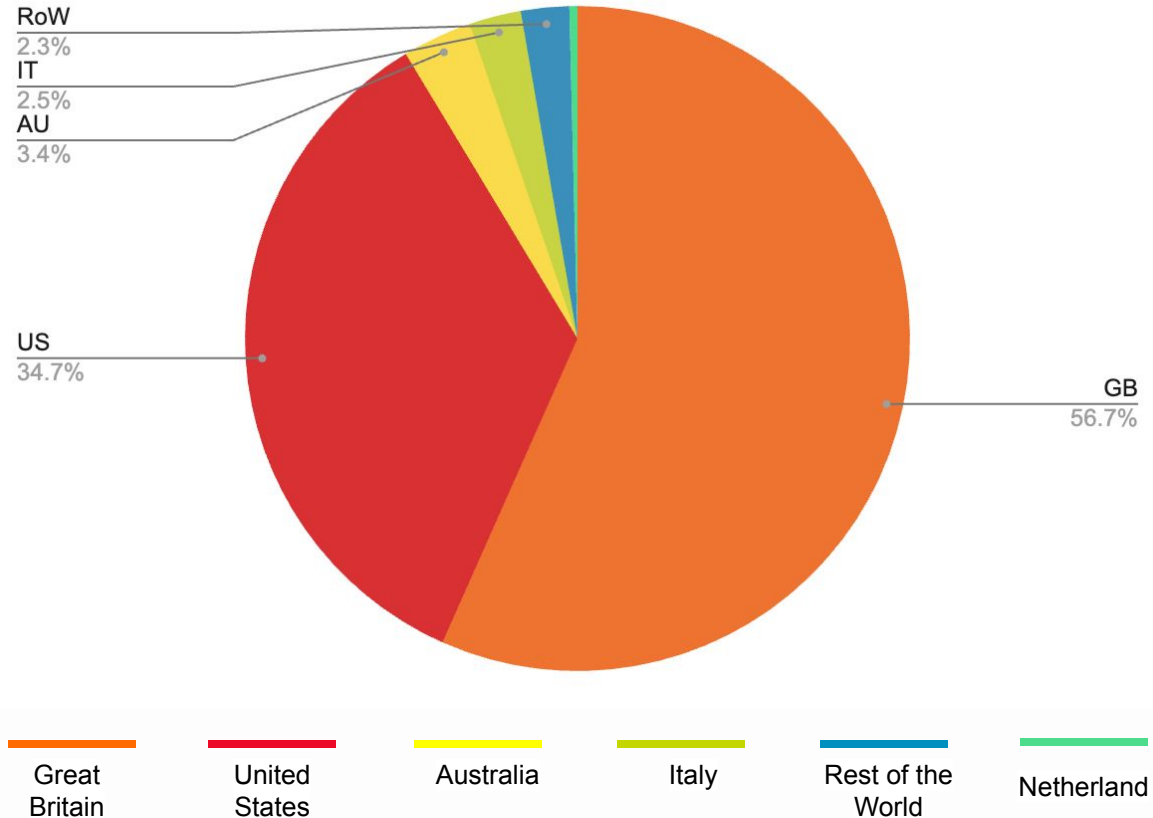
What affects the revenue?

Where is the revenue coming from?

Transaction Analyst

Great Britain brings in the most revenue within the time period.

Compared to rest of the countries, GB and US encompasses the amount of sales.



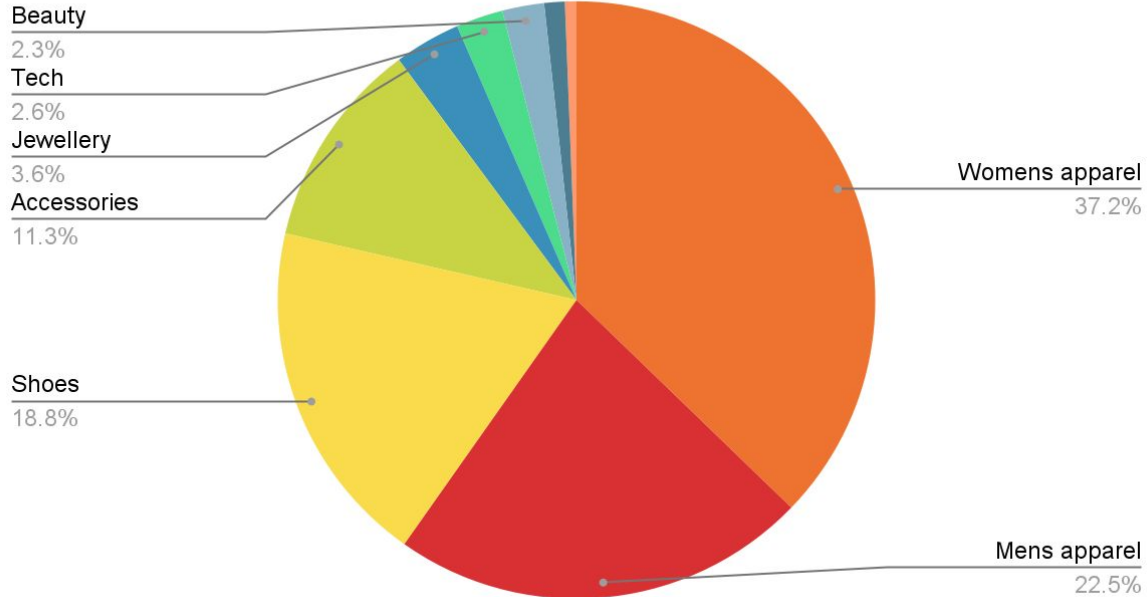
What are the top selling Departments

Transaction Analyst

Womens apparel brings in the most revenue.

Clothing takes 75%+ of the revenue.

Revenue by department

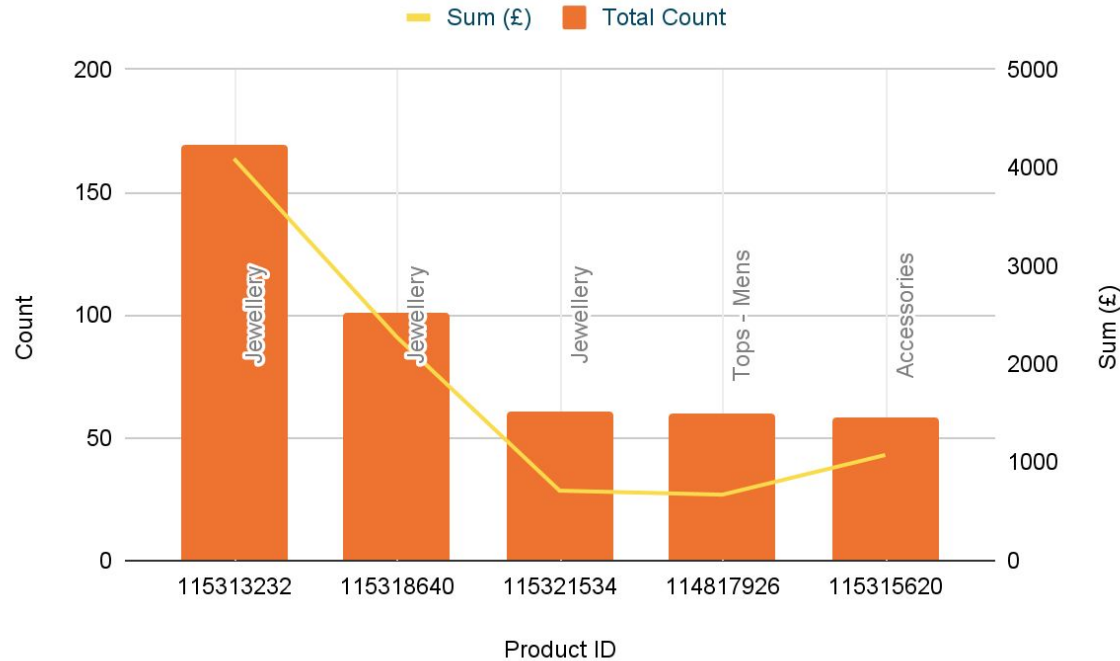


What are the top selling products?

Transaction Analyst

Jewellery product is the top selling product.

Purchases can be done in combination with basket analysis.



Does Pricing affect sale?

Transaction Analyst

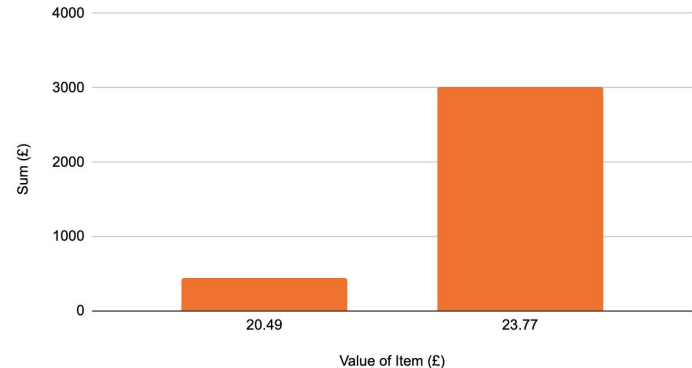
There is a difference in price for the same products within the same region.

Which price performs better, the higher or lower price?

Product id	Country	Currency	Value of Item(£)	Amount Sold	Sum (£)
115313232	AU	USD	31.56	9	284.04
115313232	GB	USD	23.77	1	23.77
115313232	GB	USD	31.56	7	220.92
115313232	RoW	USD	31.56	3	94.68
115313232	US	USD	20.49	22	450.78
115313232	US	USD	23.77	127	3018.79

Country	Value of Item(£)	Discounted
US	20.49	13.80%
US	23.77	0

Discounted vs Non-Discounted



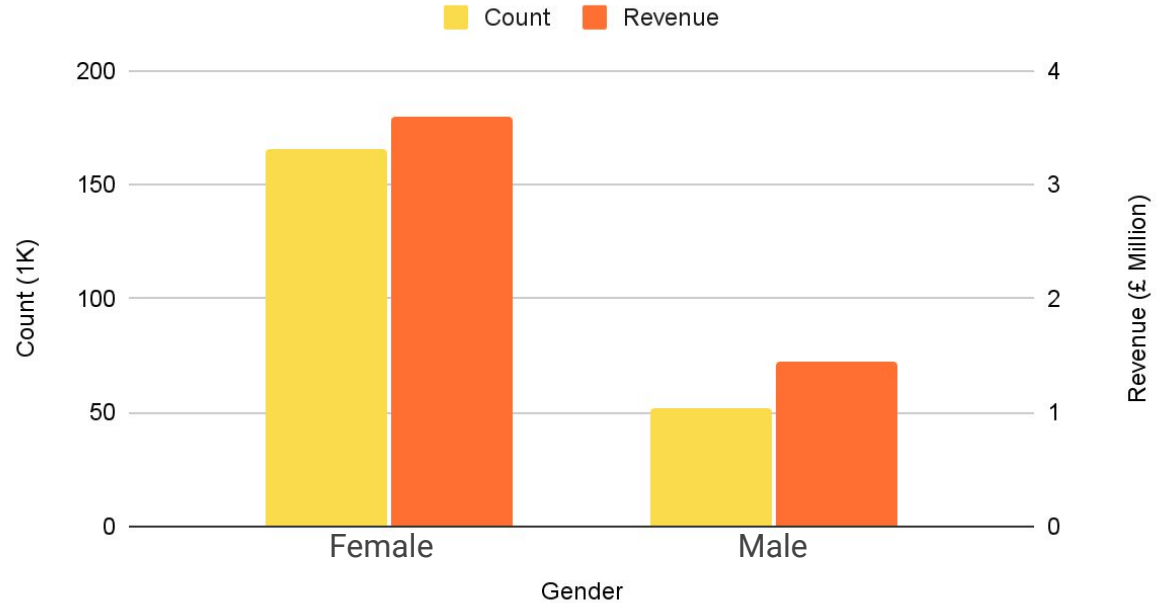
Gender Contribution



Demographic Analyst

Women contribute almost 4x more than Men to the revenue and count.

Count and Revenue



Age Group Contribution

Count and Revenue



Demographic Analyst

30-40 age group are the biggest contributors to revenue as there is the most count of them.



Gender contribution

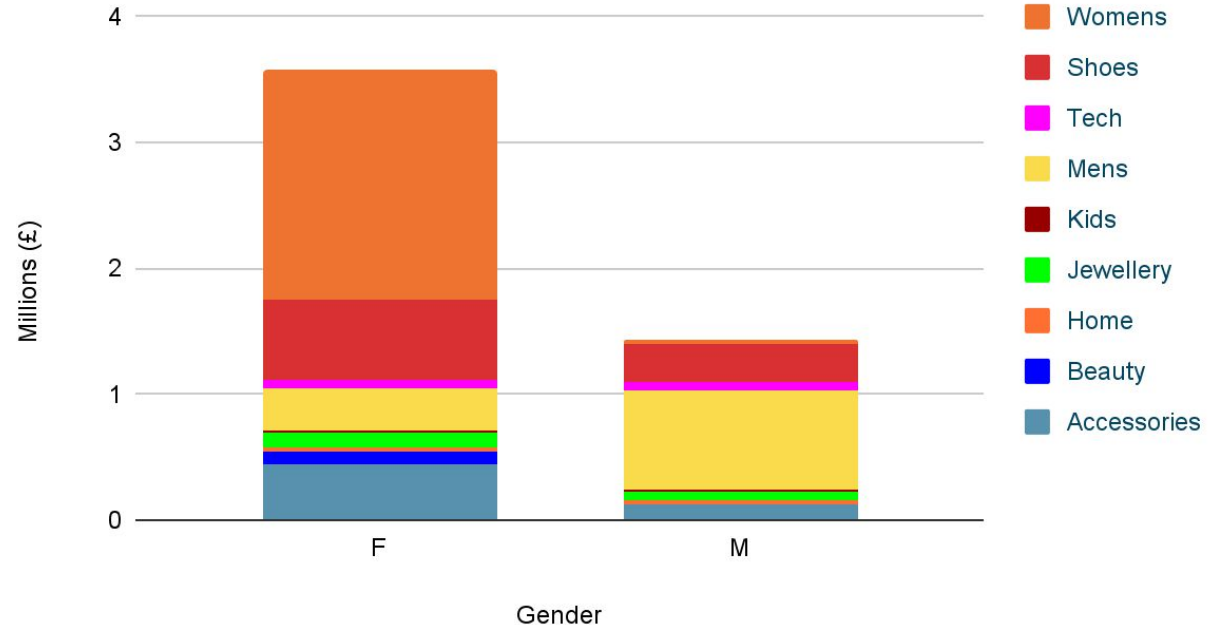


Customer Segmentation

With Females being the biggest contributors.

You can see that clothing is the largest purchases.

Departments between Gender



How different age groups consume.

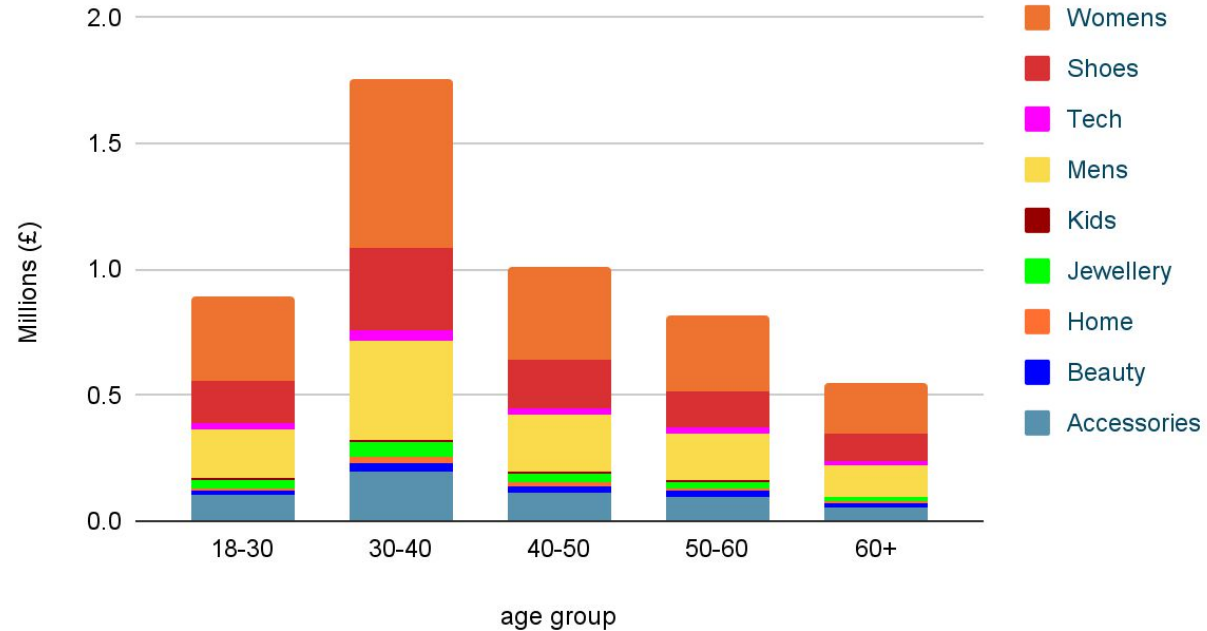


Customer Segmentation

With 30-40 age group being the biggest contributors.

Womens Apparel also follows being the biggest contributor throughout the age groups.

Departments between age group



How do the Countries consume?

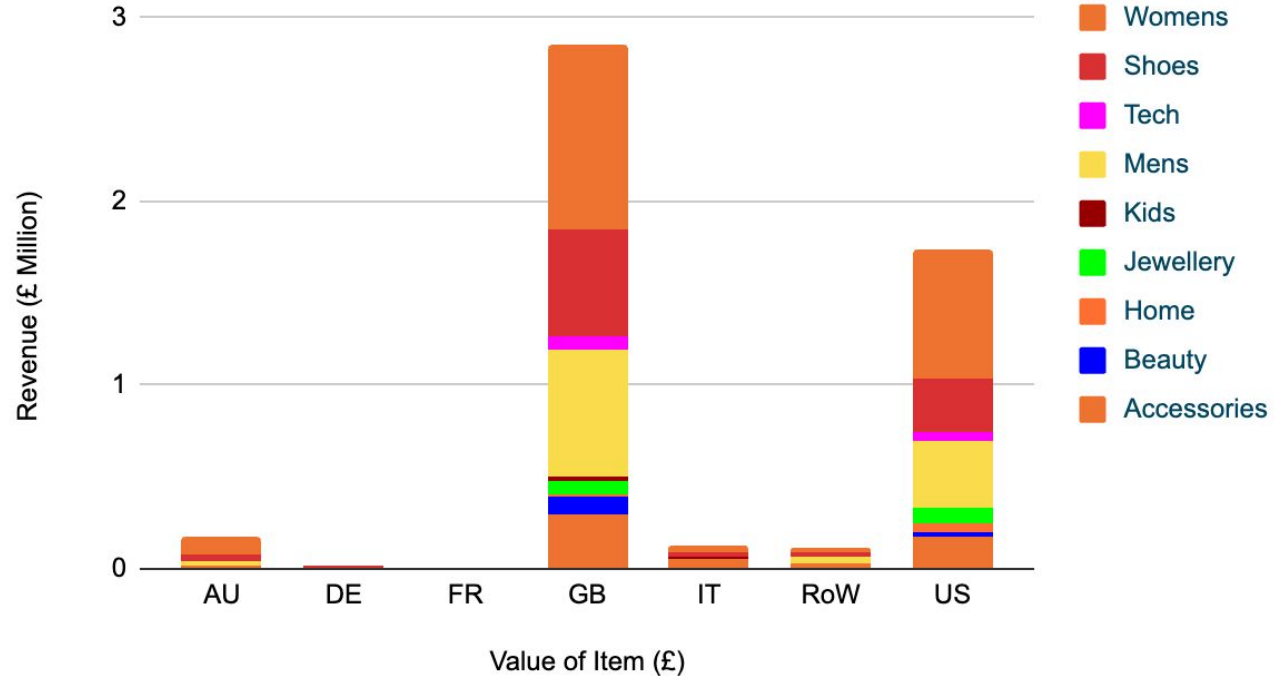


Customer Segmentation

Womens apparel is bought at most throughout all.

Some departments are not bought in various countries.

Departments between Countries



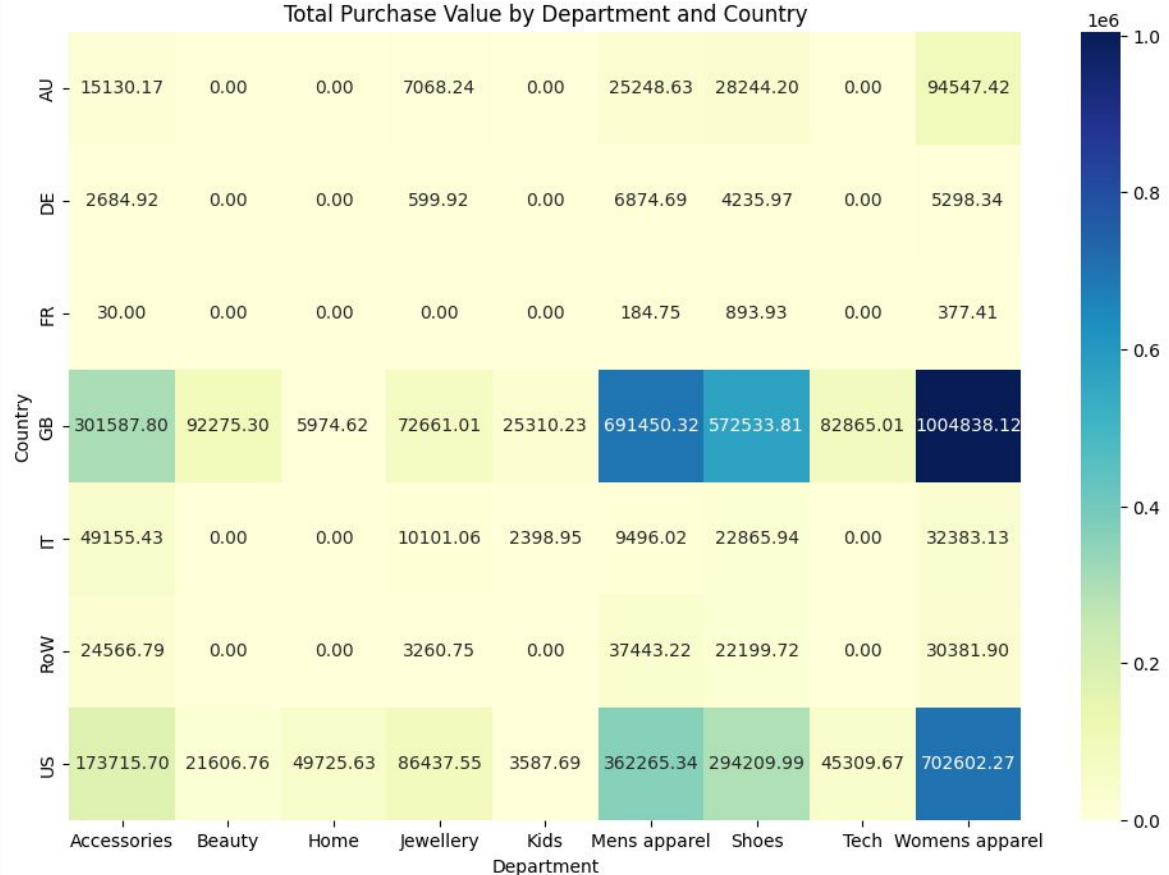
How do the Countries consume?



Customer Segmentation

Womens apparel is bought at most throughout all.

Some departments are not bought in various countries.



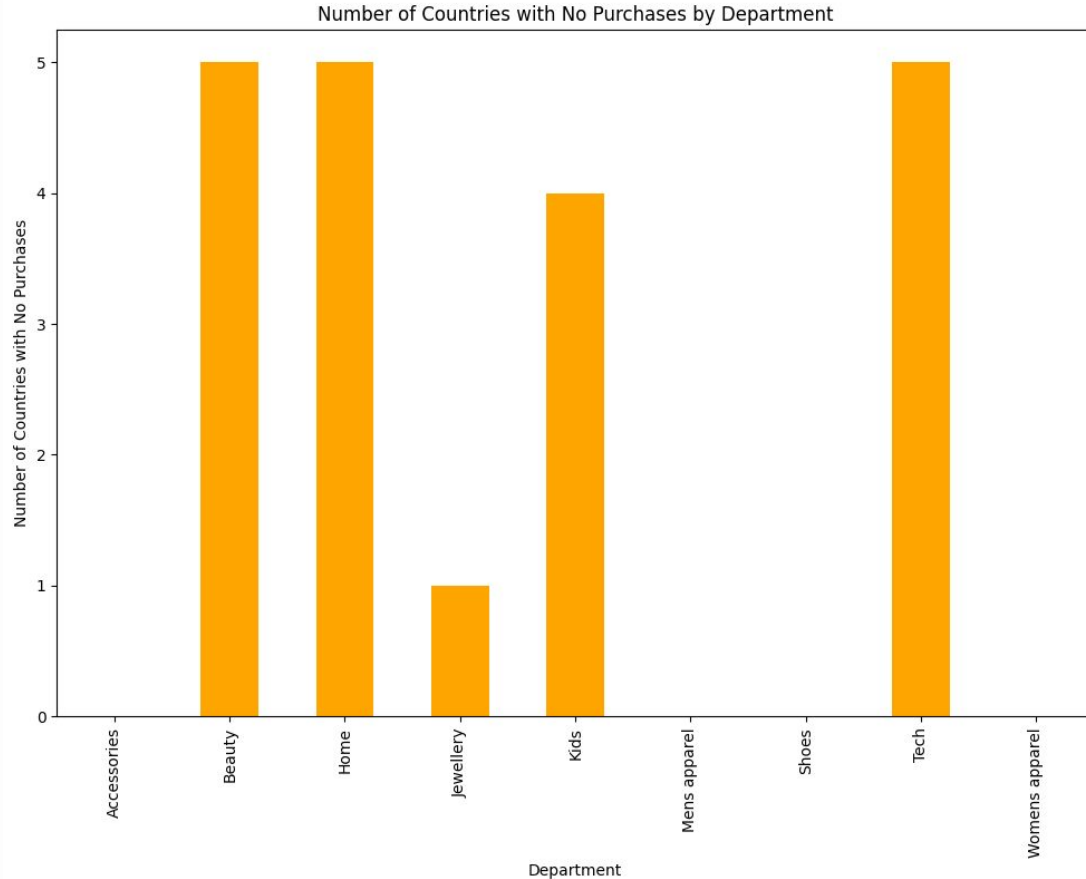
How do the Countries consume?



Customer Segmentation

Womens apparel is bought at most throughout all.

Some departments are not bought in various countries.



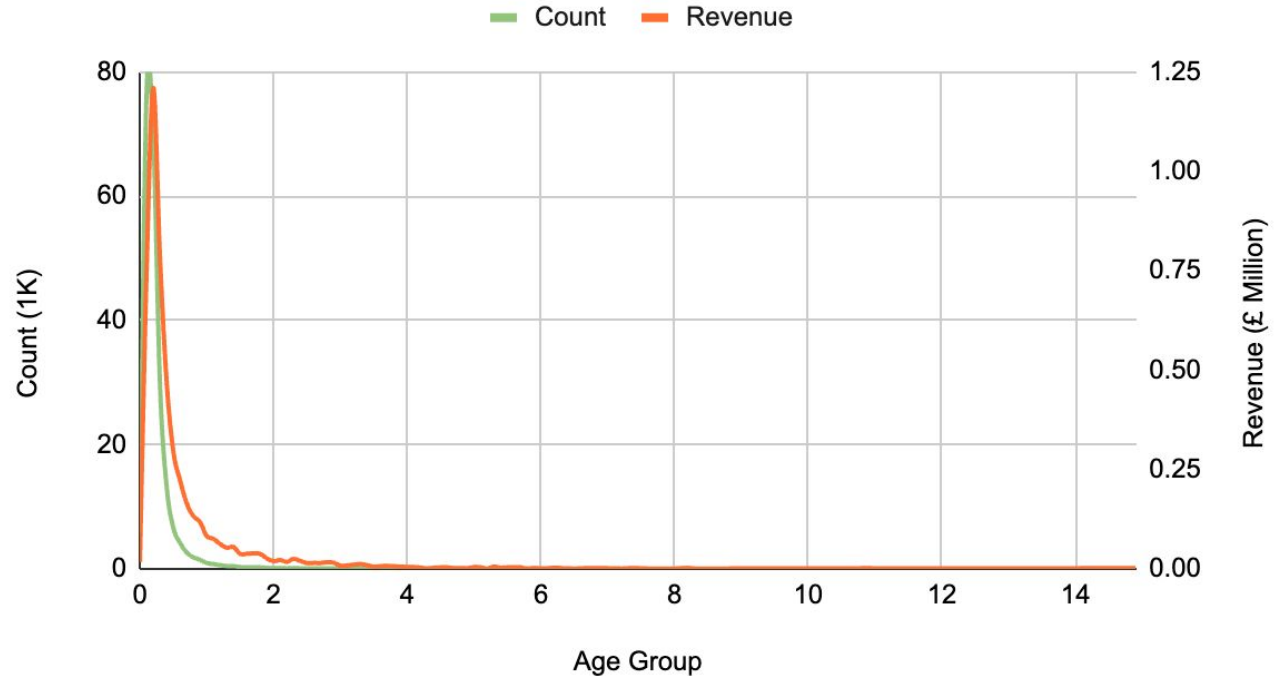
When is the most spending done?



Time-Series Analyst

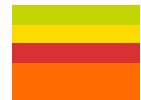
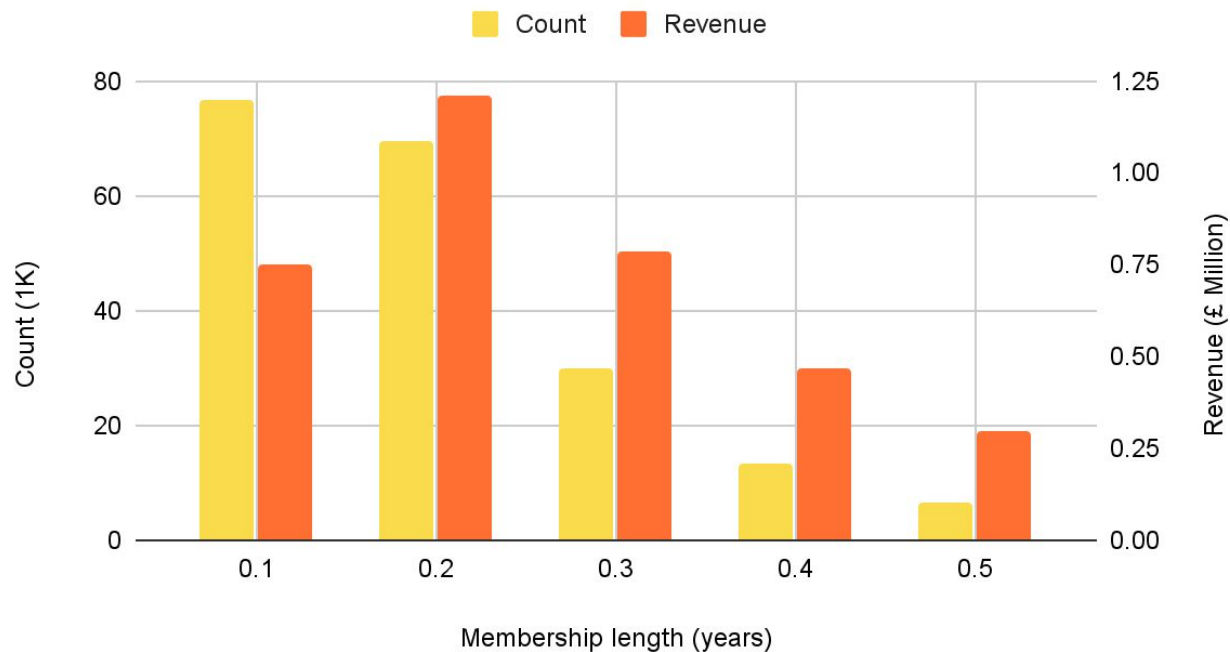
Majority of the sales come from new customers.

Count and Revenue



When is the most spending done?

Count and Revenue



Time-Series Analyst

With new customers dominating the revenue contribution.

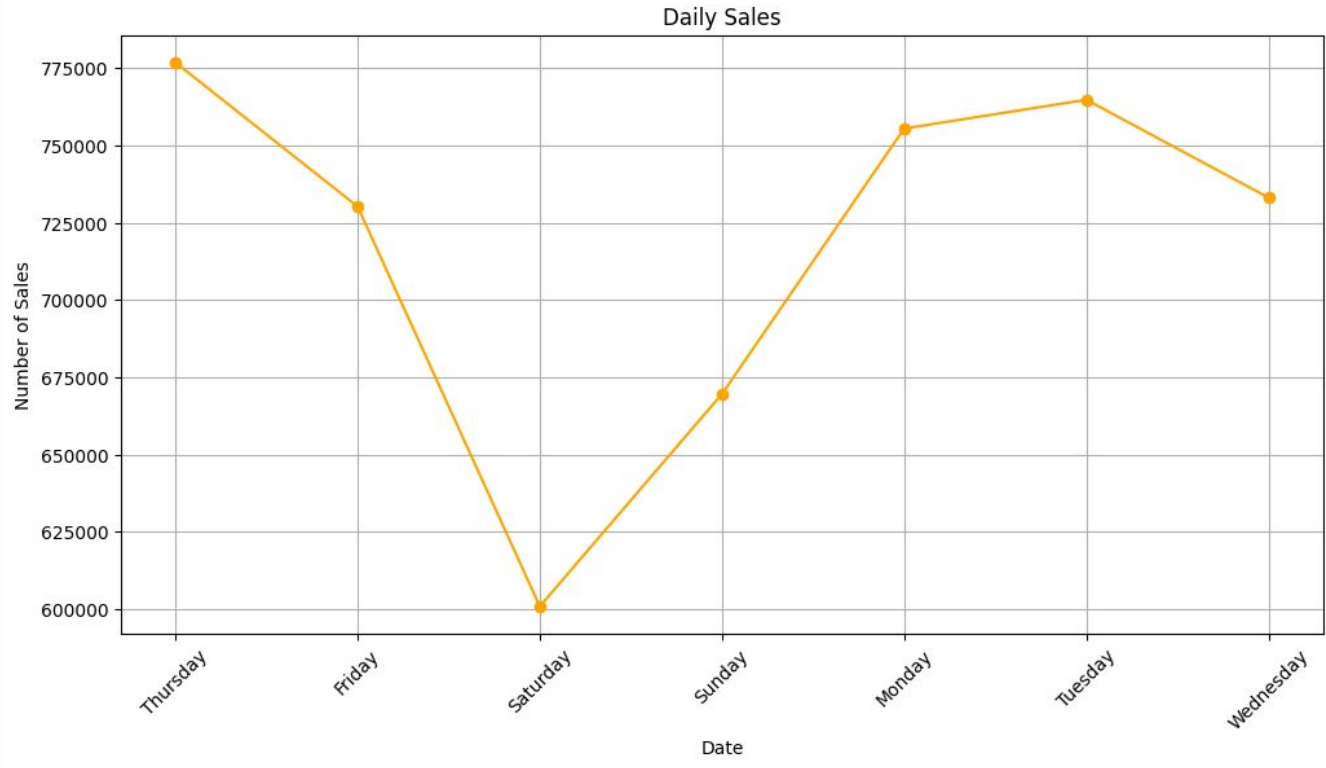


When is the most spending done?



Time-Series Analyst

The weekdays generate the most amount of sales compared to the weekend





Data Driven Decisions

How I can use the data to give a data driven business idea.



Product Portfolio Management

1

2

Regional Marketing Strategies



Topics

Customer Segmentation and Personalisation

3

4

Product Development and Sourcing

Operational Efficiency

5

6

Customer Experience Enhancement



Optimising Sales Understanding



Additional Datasets that can be used

- Seasonal Data
- Website Analytics Data
- Customer Feedback and Reviews
- Inventory and Supply Chain Data



Creating a Visual Report

- Clarity and Simplicity
- Audience Understanding
- Consistency
- Actionable Insights
- Interactivity



Further Analysis to drive Future Performance

- Cohort Analysis
- Basket Analysis
- Predictive Modeling
- A/B testing



Thanks!

Do you have any questions?

