Bee Cycle

A SQL Analysis of a Cycling Company

Lara Carter
Last Updated: Feb 2024

Table of Contents

Project Summary

Project Summary

About: Opening in July 2017, Bee Cycle is a fictitious cycling company that sells a range of cycling products in stores across the globe.

Problem: Since the COVID-19 pandemic, Bee Cycle has started to suffer a decline in profits.

Objective: Bee Cycle has asked you, a data analyst, to provide data driven insights to help inform Bee Cycle's marketing strategy, with the aim of increasing company profits.

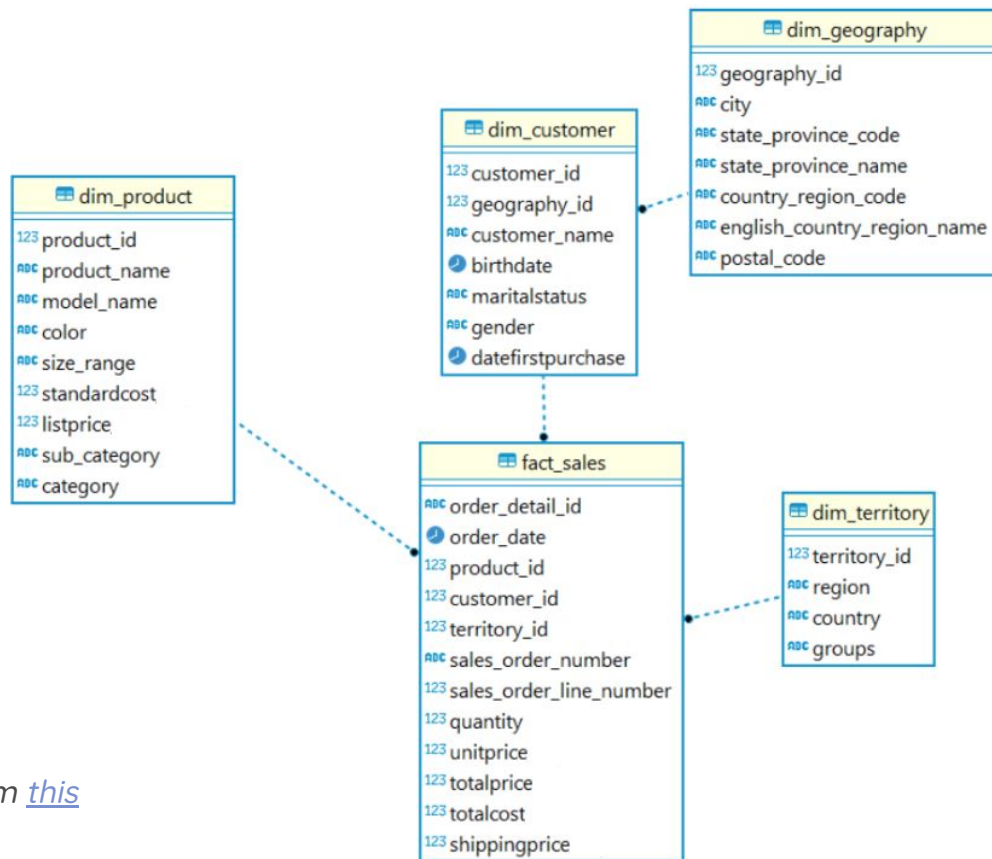| Skills	Tools
➢ Sorting and Filtering ➢ Data Cleaning and Wrangling ➢ Data Quality Assurance ➢ Exploratory Data Analysis ➢ Queries, Subqueries, and Common Table Expressions ➢ Joining Data from Relational Datasets	**BigQuery**
➢ Data Visualisations	**Tableau**

Data

Data

Bee Cycle provided you with **5 datasets*** to analyse.

Relations between datasets are summarised in this entity relationship diagram:

A data dictionary is also available to view here.

All data is fictitious and was retrieved from this public repository.

dim_product
- 123 product_id
- ABC product_name
- ABC model_name
- ABC color
- ABC size_range
- 123 standardcost
- 123 listprice
- ABC sub_category
- ABC category

dim_customer
- 123 customer_id
- 123 geography_id
- ABC customer_name
- 🕐 birthdate
- ABC maritalstatus
- ABC gender
- 🕐 datefirstpurchase

dim_geography
- 123 geography_id
- ABC city
- ABC state_province_code
- ABC state_province_name
- ABC country_region_code
- ABC english_country_region_name
- ABC postal_code

fact_sales
- ABC order_detail_id
- 🕐 order_date
- 123 product_id
- 123 customer_id
- 123 territory_id
- ABC sales_order_number
- 123 sales_order_line_number
- 123 quantity
- 123 unitprice
- 123 totalprice
- 123 totalcost
- 123 shippingprice

dim_territory
- 123 territory_id
- ABC region
- ABC country
- ABC groups

Evaluation of the Data

❖ The dataset is **comprehensive** in that it includes information on a range of variables including locations, demographics, product colours and sizes, and transactions since store opening.
  ➢ However, **additional granularity** such as the time an order was placed, whether purchases were in-store or online, or more detailed product features may provide further insights to help answer the business questions.

❖ The data can be considered **reliable**, in that it is first-party data that is complete and consistent and generally free from outliers, duplicates or invalid values.
  ➢ However, the data is **not wholly up to date**, as it only provides sales data until June 2022. More current data could therefore further enhance reliability.

Project Process

Project Process

1   **Upload data (5 CSV files) to BigQuery**

2   **Initial inspection of the data**
- ➢ Assessed each variable, checked their data types, and number of rows/columns in each table
- ➢ Assessed table relations and primary and foreign keys
- ➢ Created a data dictionary

Project Process

3   **Data cleaning, wrangling and analysis using SQL queries**
  - ➢ Sorted and filtered the data
  - ➢ Checked and handled null, outlying and impossible values, as well as any spelling errors or misfielded values
  - ➢ Used queries such as aggregate, window, string, conversion, conditional, and date and time functions, as well as join statements
  - ➢ Used subqueries and common table expressions to answer complex questions

4   **Create data visualisations in Tableau**

Analysis

Key Questions

We will focus our analysis on the following questions:

1) Which products are most profitable?
2) Which months are most profitable?
3) Which regions are most profitable?
4) What type of customer generates the most profit?
5) Which customers spend the most money?
6) a. Which cities are most profitable?
   b. What are the most popular products within these cities?
   c. In what colours?

Data Overview

1360

customers

119

products available

6

countries with stores

5954

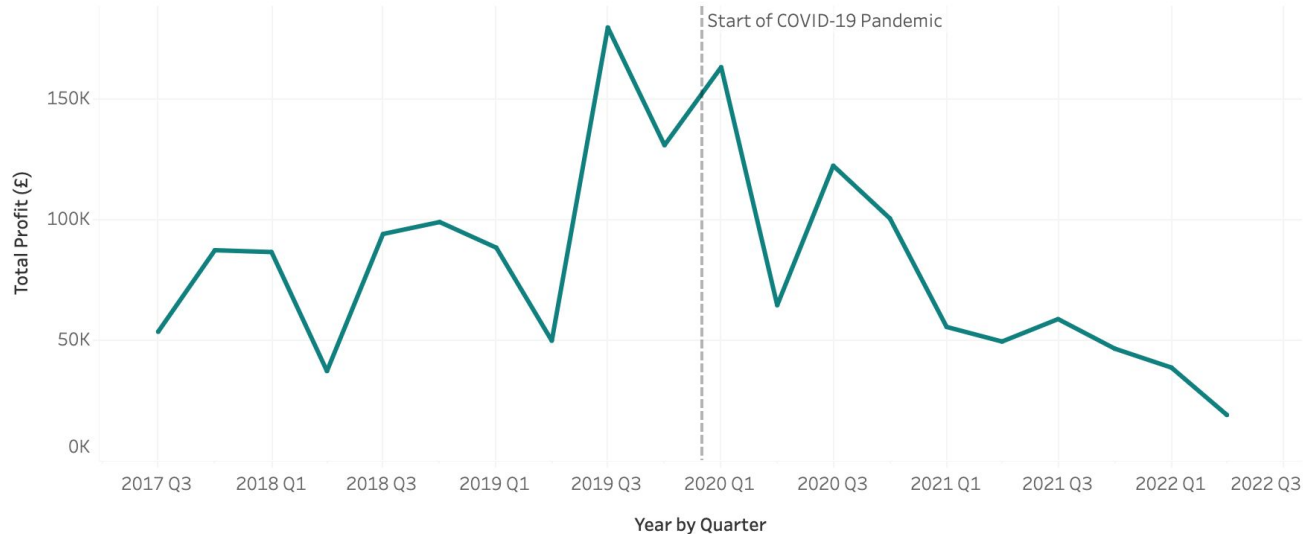total orders

£3,862,950.25

total revenue

£1,621,608.21

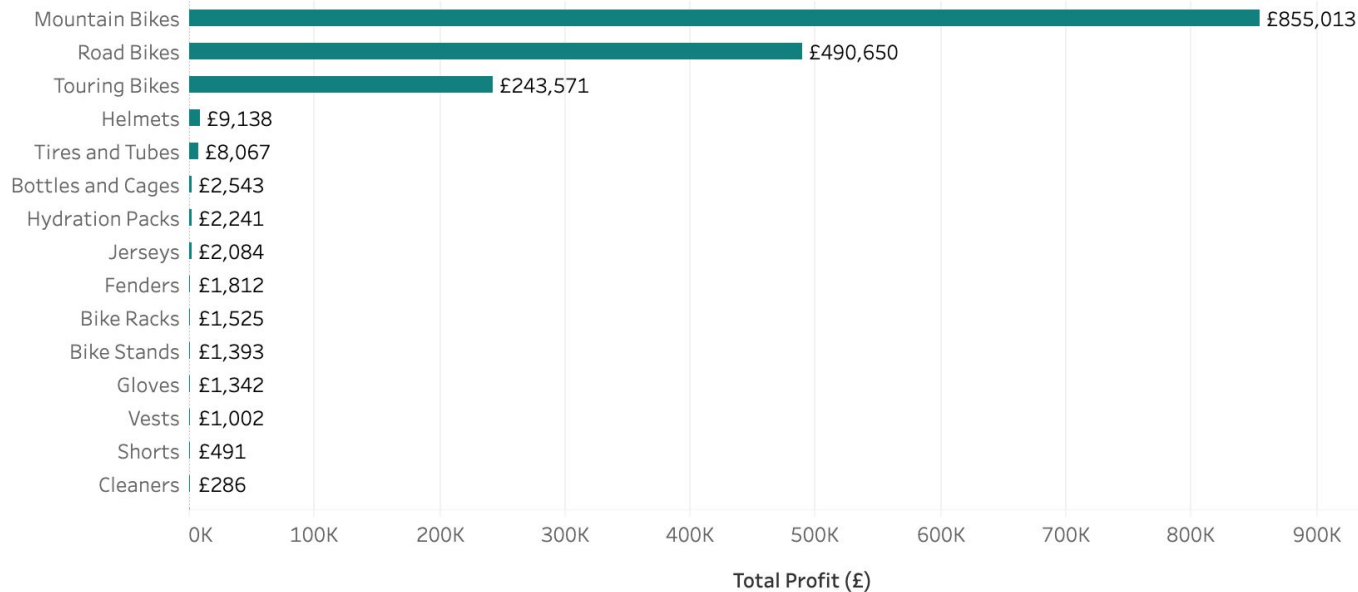total profits

Data Overview

Total Profit by Year

There was a sharp decline in profits shortly after the COVID-19 pandemic began in Dec 2019.

Profits briefly increased but have **continued to decline again since the end of Q3 2020**.
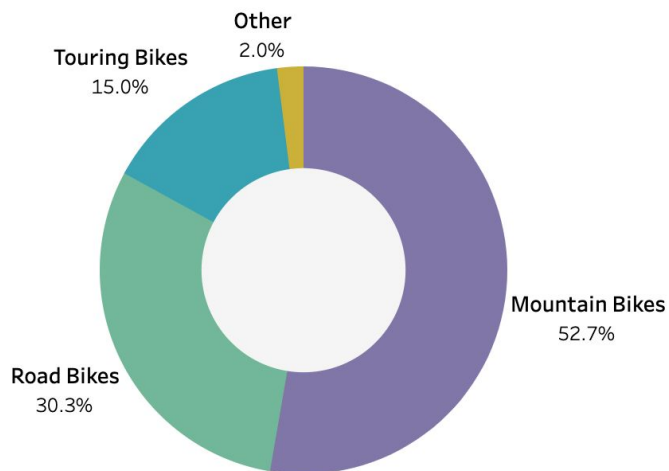
Most Profitable Products

Total Profit by Product

Product	Total Profit
Mountain Bikes	£855,013
Road Bikes	£490,650
Touring Bikes	£243,571
Helmets	£9,138
Tires and Tubes	£8,067
Bottles and Cages	£2,543
Hydration Packs	£2,241
Jerseys	£2,084
Fenders	£1,812
Bike Racks	£1,525
Bike Stands	£1,393
Gloves	£1,342
Vests	£1,002
Shorts	£491
Cleaners	£286

Total Profit (£)

0K 100K 200K 300K 400K 500K 600K 700K 800K 900K

Most Profitable Products

Percentage of Total Profits by Product

Other
2.0%

Touring Bikes
15.0%

Mountain Bikes
52.7%

Road Bikes
30.3%
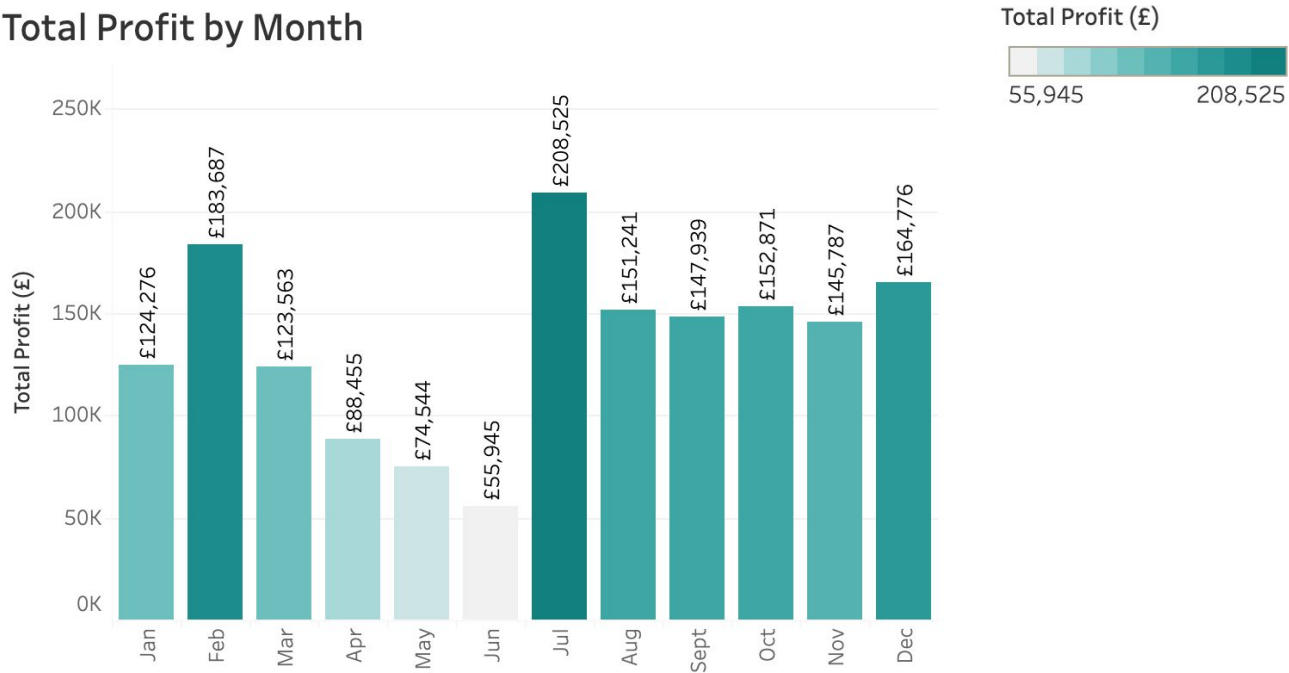
As seen from the current and previous slide, the most profitable products are:

1) **Mountain Bikes**
2) **Road Bikes**
3) **Touring Bikes**

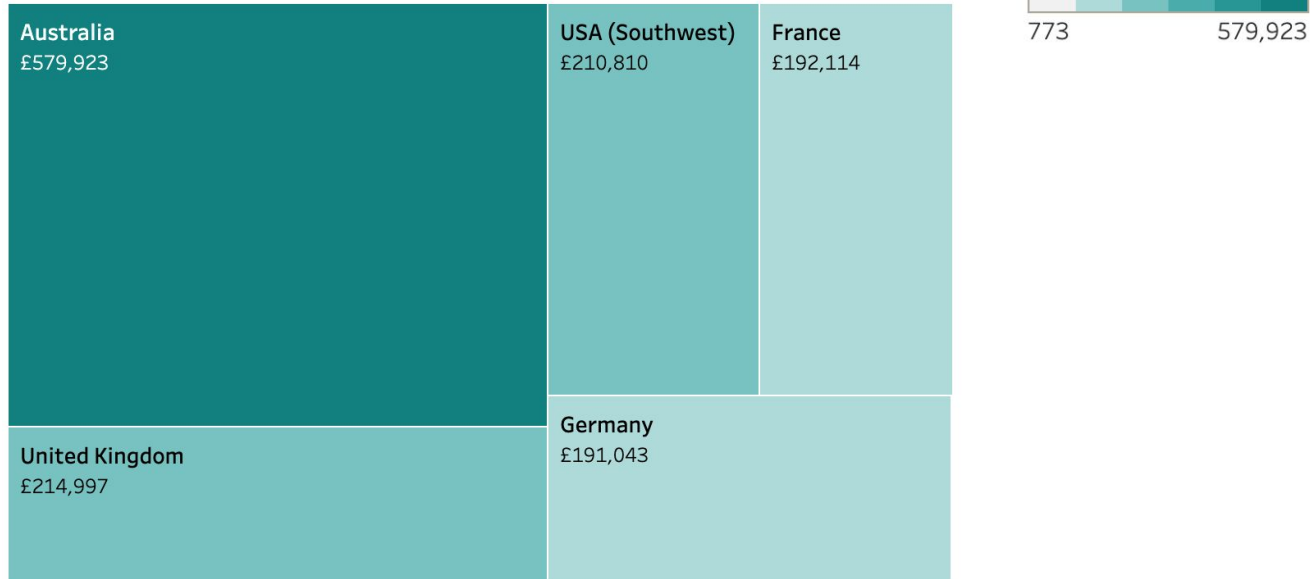# Most Profitable Months

Total Profit by Month
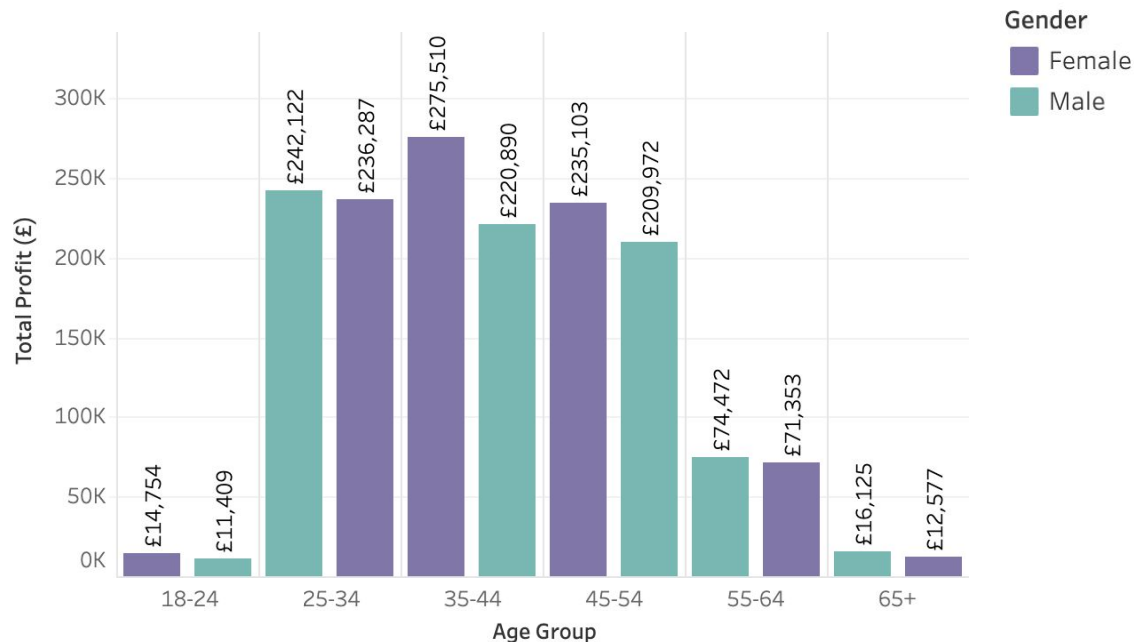
Total Profit (£)
55,945 — 208,525

Month	Total Profit (£)
Jan	£124,276
Feb	£183,687
Mar	£123,563
Apr	£88,455
May	£74,544
Jun	£55,945
Jul	£208,525
Aug	£151,241
Sept	£147,939
Oct	£152,871
Nov	£145,787
Dec	£164,776

Top 5 Most Profitable Regions

Most Profitable Customer Type

Total Profit by Age and Gender

Gender
- Female
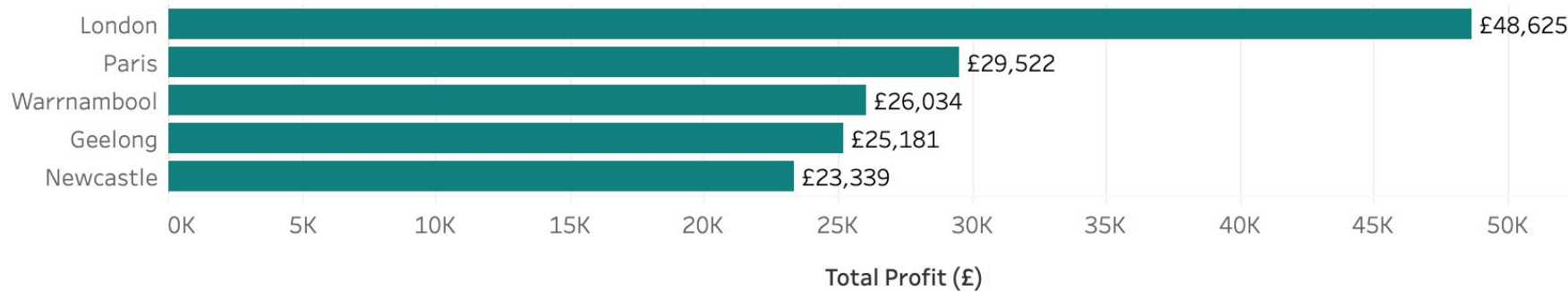- Male

Customers generating the most profit are **men/women aged 25-54.**

The single most profitable customer type is **women aged 35-44.**

Top 10 Highest Spending Customers

Name	Total Spent
Nichole Nara	£9307
Kaitlyn Henderson	£9306
Margaret He	£9288
Randall Dominguez	£9286
Adriana Gonzalez	£9270
Rosa Hu	£9251
Brandi Gill	£9237
Brad She	£9221
Francisco Sara	£9215
Kate Anand	£7610

Top 5 Most Profitable Cities

Top 5 Most Profitable Cities

City	Total Profit (£)
London	£48,625
Paris	£29,522
Warrnambool	£26,034
Geelong	£25,181
Newcastle	£23,339

Total Profit (£)

Top 2 Most Popular Products in Most Profitable Cities

Popular Products and Colours in Most Profitable Cities
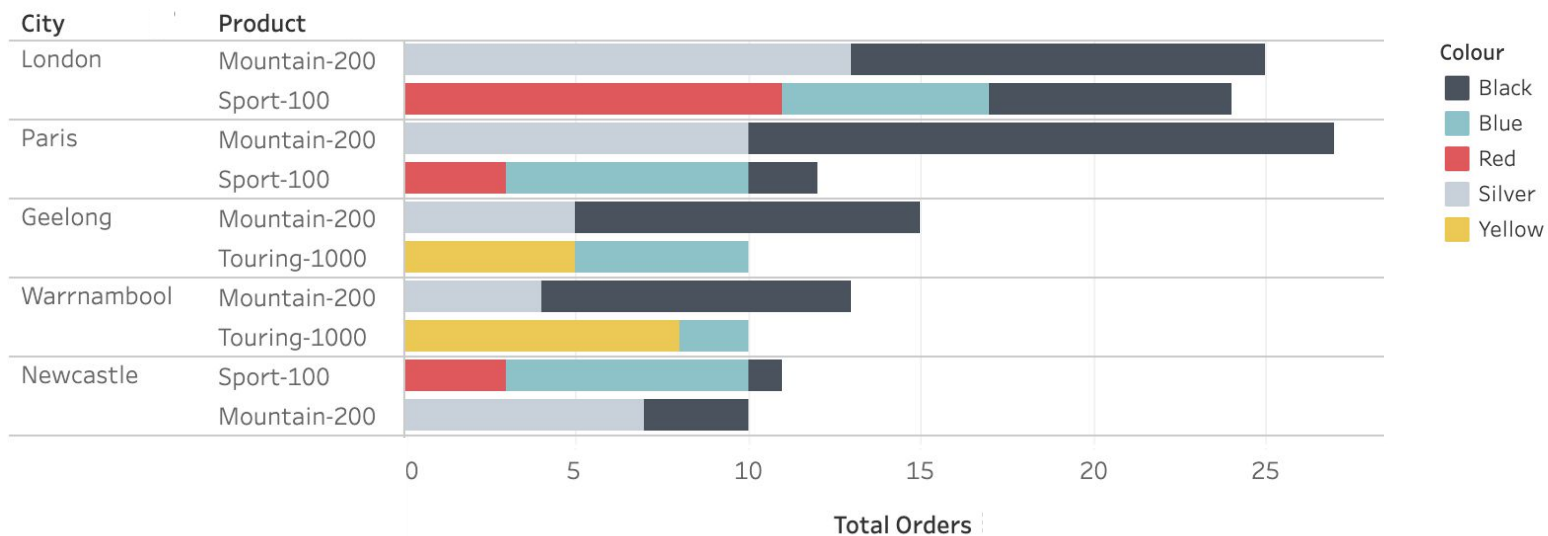
Actions

Recommendations

1 Focus marketing on promoting products that generate the most profit: **mountain bikes, road bikes, and touring bikes**.

2 Schedule marketing campaigns during **July** (the most profitable month). Consider sponsoring cycling events and competitions that take place during July such as the **Tour de France**.

3 Increase marketing in regions that generate the most profit: **Australia, UK, Southwest USA, France and Germany**.

Recommendations

4 Target marketing towards customer profiles that generate the most profit: **men/women aged 25-54** and specifically **women aged 35-44**.

5 Consider implementing a **loyalty scheme** that provides rewards and discounts to high spenders. This will maintain good relations and encourage spending.

6 Develop billboards in the top 5 most profitable cities: **London, Paris, Warrnambool, Geelong, and Newcastle**.

Ensure billboards display the top 2 most popular products for that particular city and in the most popular colours.

Next Steps

➢ Monitor and evaluate the effectiveness of implemented recommendations. For example, assess whether profits have increased since loyalty schemes were introduced.

➢ Gather additional data sources, such as more up-to-date sales data, customer feedback, product features, and competitor information, to enrich the analysis and gain deeper insights into customer behavior and market dynamics.

Please click [here](#) to view SQL queries on GitHub