KPMG Data Analytics Virtual Internship

<u>Context</u>

Sprocket Central Pty Ltd , a medium size bikes & cycling accessories organization, needs help with its customer and transactions data. The organization has a large dataset relating to its customers, but their team is unsure how to effectively analyze it to help optimize its marketing strategy.

The client provided KPMG with 3 datasets:
Customer Demographic
Customer Addresses
Transactions data in the past 3 months

<u>Data cleaning</u>

The data cleaning process primarily involved:
-   checking the data structure
-   converting strings to categorical variables
-   formatting dates correctly
-   verifying for missing values and duplicated rows
-   ensuring the correct spelling of category names for categorical variables, and removing unnecessary columns.
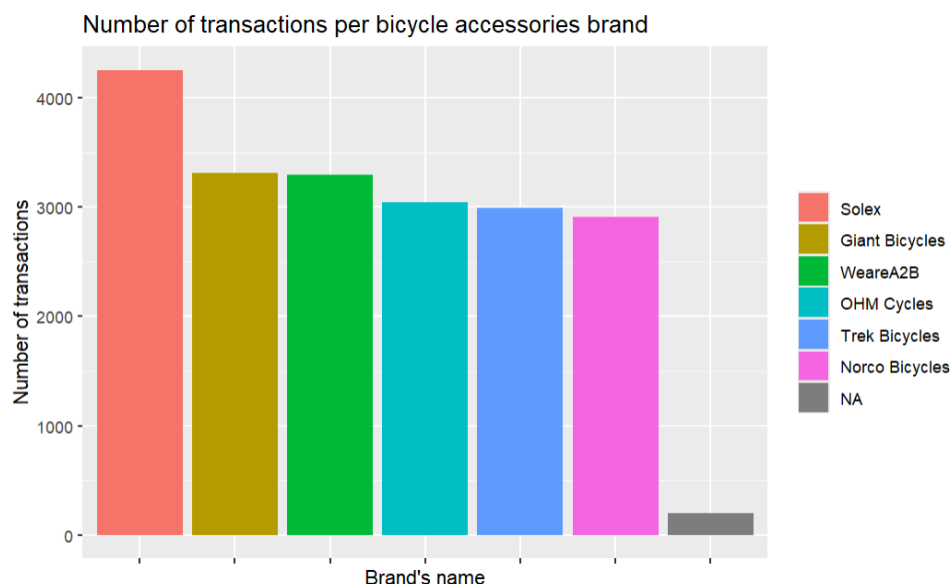
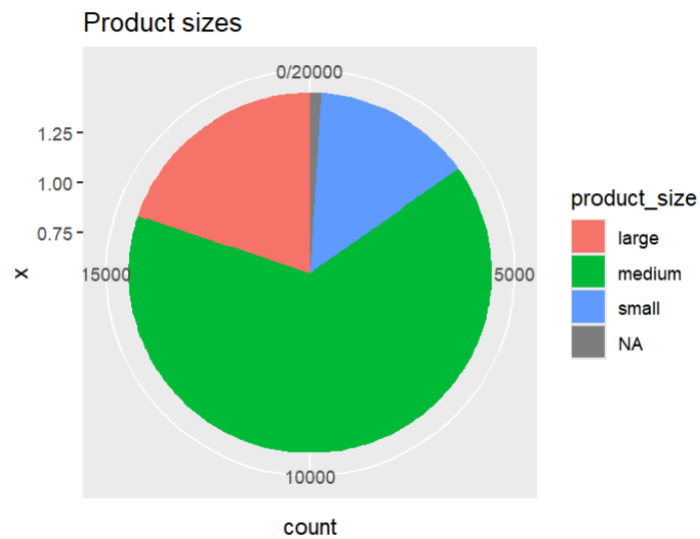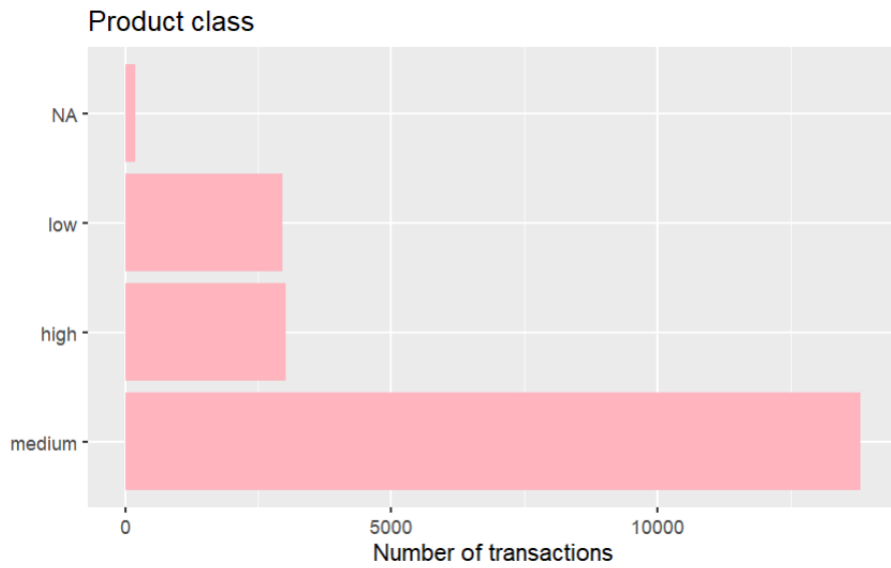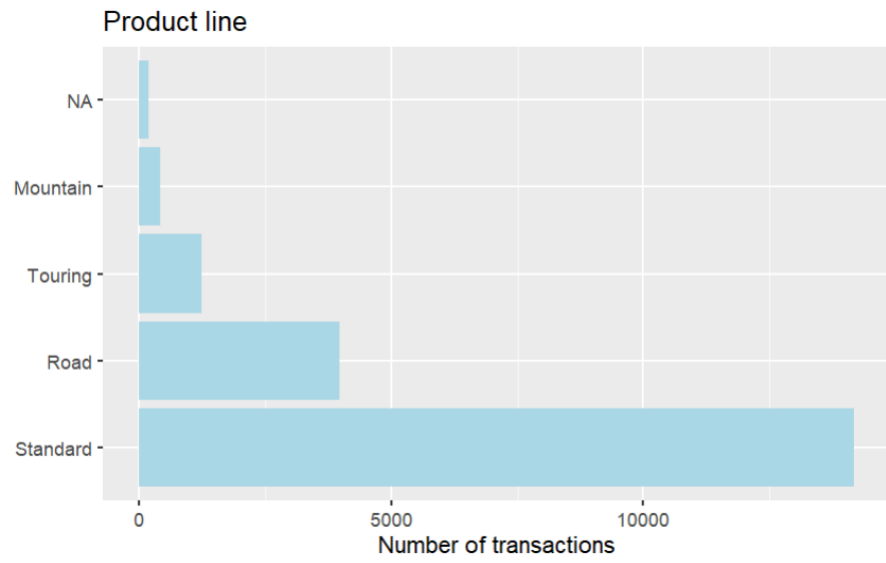**<u>Data exploration</u>**

1) Purchasing habits

Sprocket customers predominantly prefer accessories from the Solex brand, but they also enjoy purchasing from other brands.
Preferred product line by customers : Standard
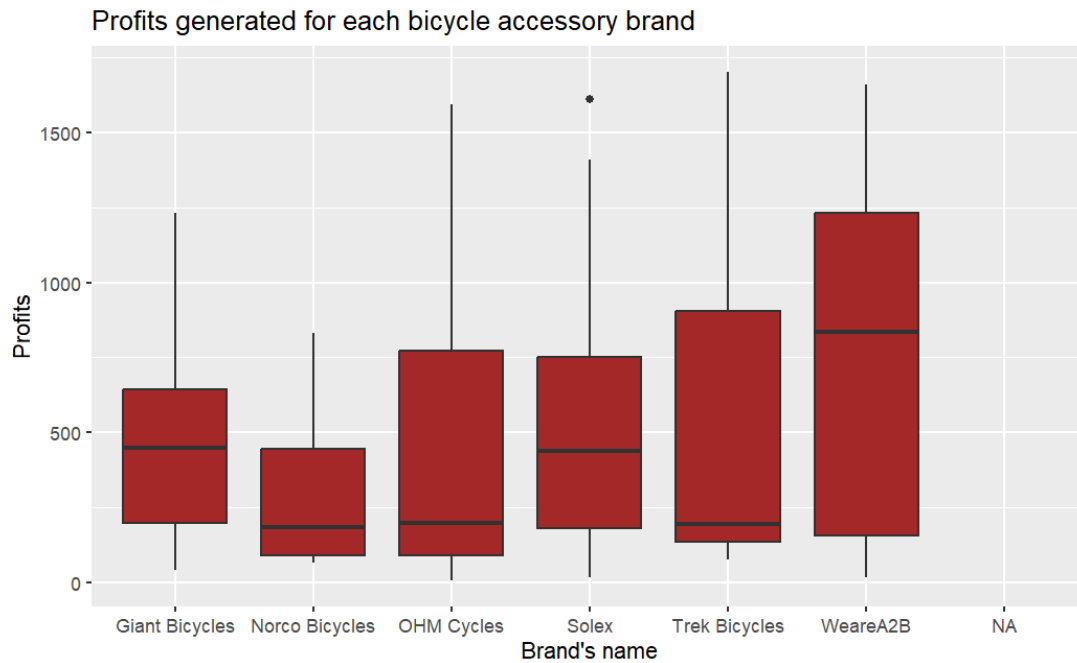Preferred product class by customers : Medium
Preferred product size by customers : Medium

Product line

Product class

Product sizes

2) Profits

The bicycle accessory brands that have generated the most profits are: WeareA2B, Trek Bicycles, OHM Cycles and Solex.

Profits generated for each bicycle accessory brand

3) Customer characteristics

What doesn't influence the quantity of purchases: gender and whether the customer owns a car or not.
Customers who made the most purchases in the last 3 years:
- Job industry category: Manufacturing
- Wealth segment: Mass customer
- Age group: 40-50; 50-60
- State: NSW (New South Wales)