1. Data Cleansing Steps

Graphical user interface, text, application, email

Description automatically generated

### 2. Data Exploration

1. What day of the week is used for each week\_date value?

Graphical user interface, text, application

Description automatically generated

1. What range of week numbers are missing from the dataset?

Graphical user interface, text, application

Description automatically generated

1. How many total transactions were there for each year in the dataset?

Graphical user interface, application

Description automatically generated

1. What is the total sales for each region for each month?

Graphical user interface, text

Description automatically generated

1. What is the total count of transactions for each platform?

Graphical user interface, text, application

Description automatically generated

1. What is the percentage of sales for Retail vs Shopify for each

month?

Graphical user interface, text, application

Description automatically generated

Table

Description automatically generated

Table

Description automatically generated

1. What is the percentage of sales by demographic for each year in the dataset?

Graphical user interface, text, application

Description automatically generated

1. Which age\_band and demographic values contribute the most to

Retail sales?

Graphical user interface, application

Description automatically generated

1. Can we use the avg\_transaction column to find the average transaction size for each year for Retail vs Shopify? If not - how would you calculate it instead?

Graphical user interface

Description automatically generated with medium confidence

**3. Before & After Analysis**

**This technique is usually used when we inspect an important event and want to inspect the impact before and after a certain point in time.**

**Taking the week\_date value of 2020-06-15 as the baseline week where the Data Mart sustainable packaging changes came into effect.**

**We would include all week\_date values for 2020-06-15 as the start of the period after the change and the previous week\_date values would be before**

**Using this analysis approach - answer the following questions:**

1. **What is the total sales for the 4 weeks before and after 2020-06-15? What is the growth or reduction rate in actual values and percentage of sales?**

Graphical user interface, text, application

Description automatically generated

1. **What about the entire 12 weeks before and after?**

Graphical user interface, text, application

Description automatically generated

1. **How do the sale metrics for these 2 periods before and after compare with the previous years in 2018 and 2019?**

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated