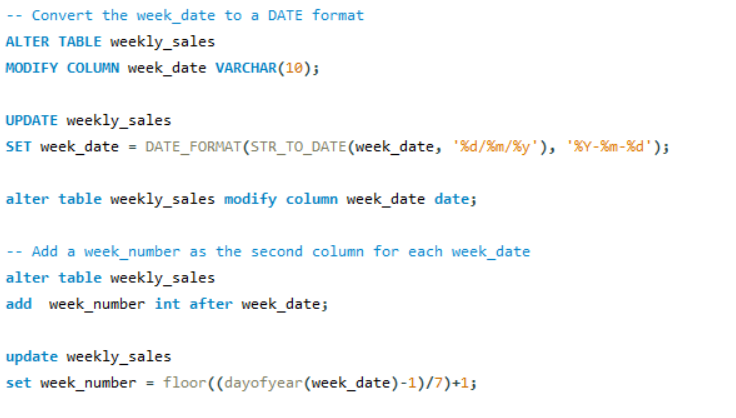
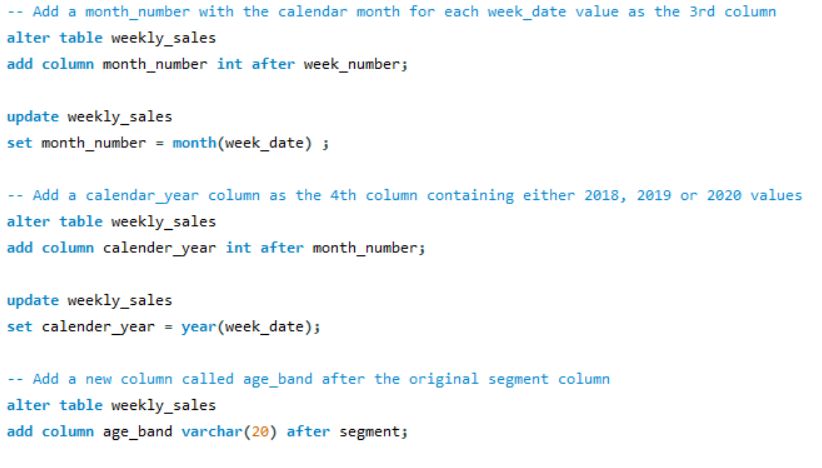
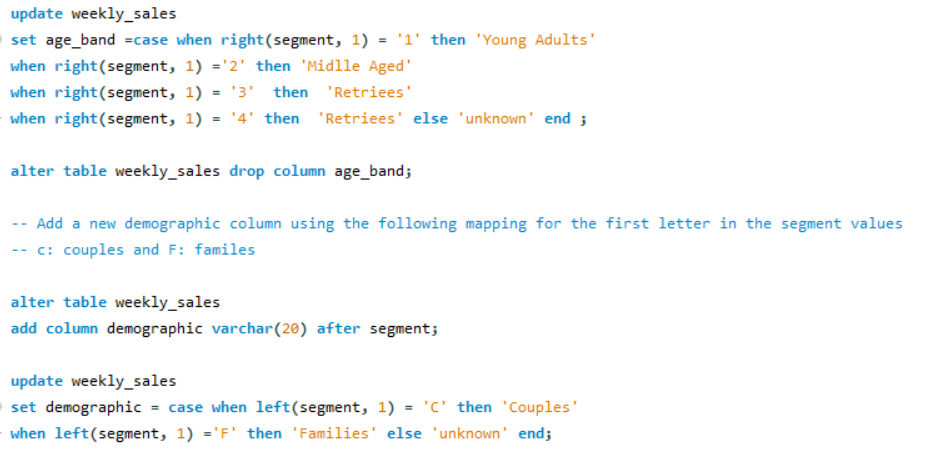
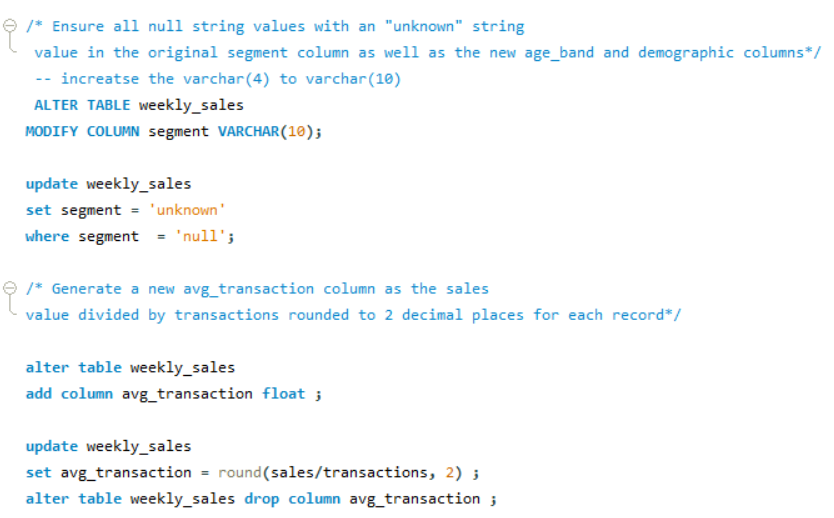
Danny Ma – DATA MART \_ 8 WEEK CHALLENGE

1. DATA CLEANING

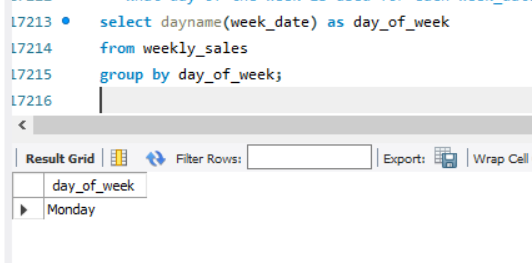




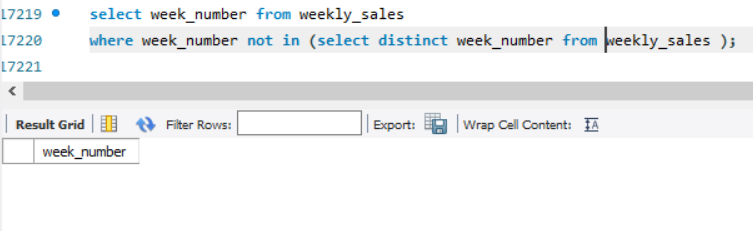




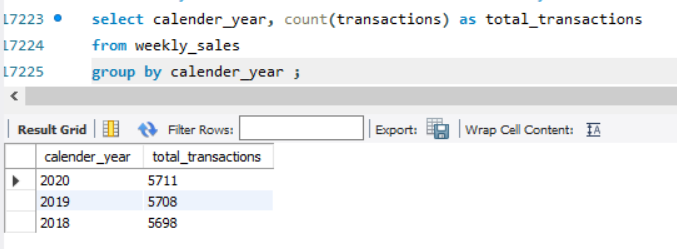
1. DATA EXPLORATION
2. What day of the week is used for each week\_date value?



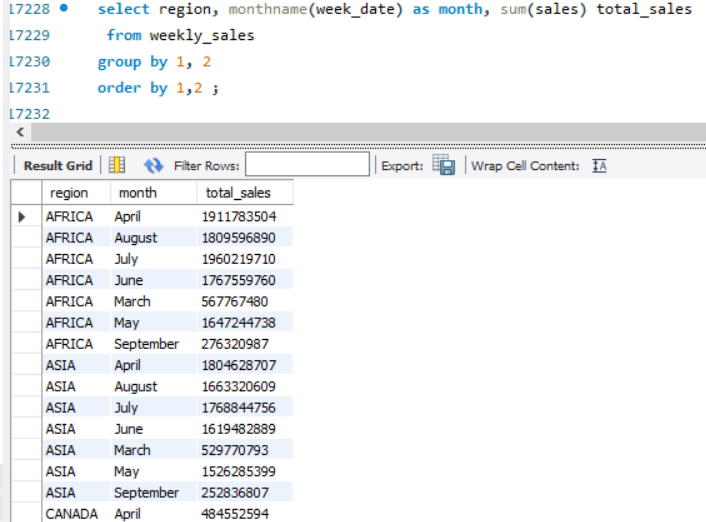
1. What range of week numbers are missing from the dataset?



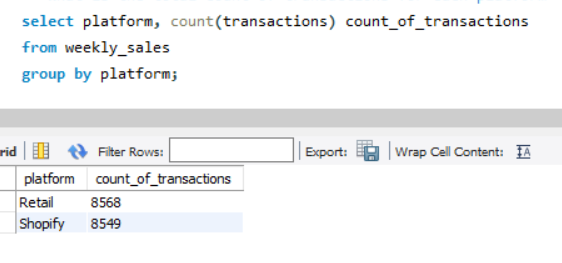
1. How many total transactions were there for each year in the dataset?



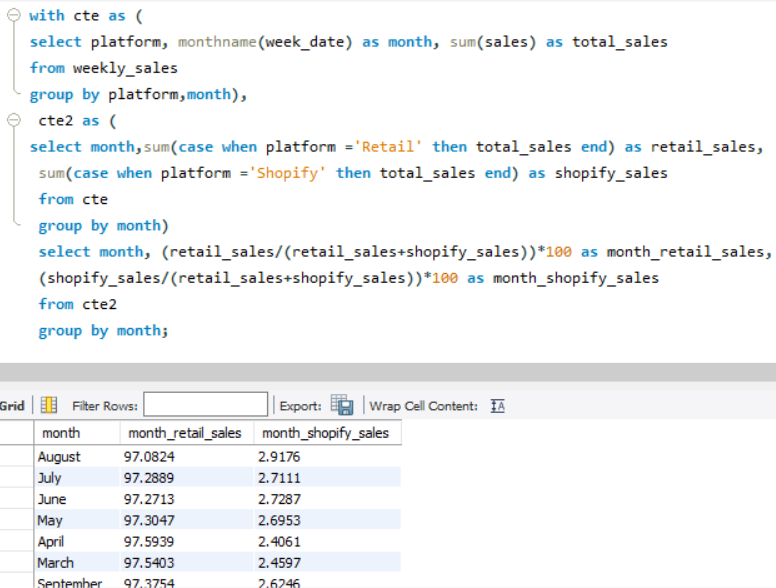
4.What is the total sales for each region for each month?



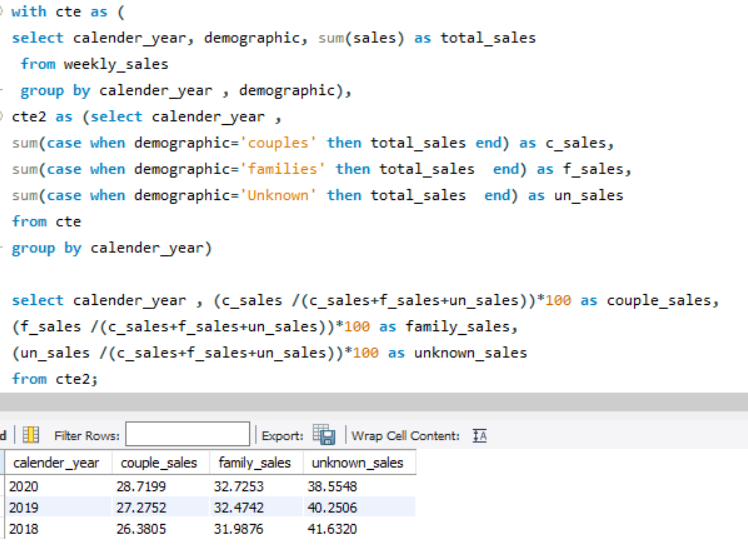
5.What is the total count of transactions for each platform?



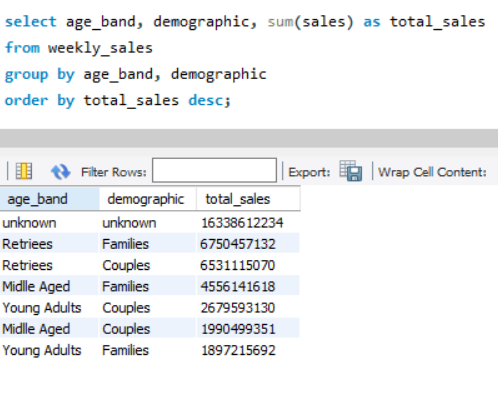
6.What is the percentage of sales for Retail vs Shopify for each month?



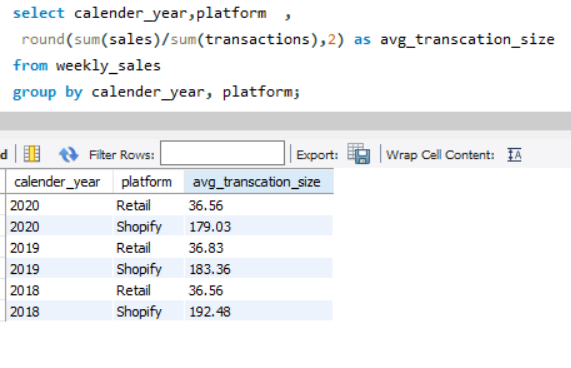
7.What is the percentage of sales by demographic for each year in the dataset?



8. Which age\_band and demographic values contribute the most to Retail sales?



9.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?



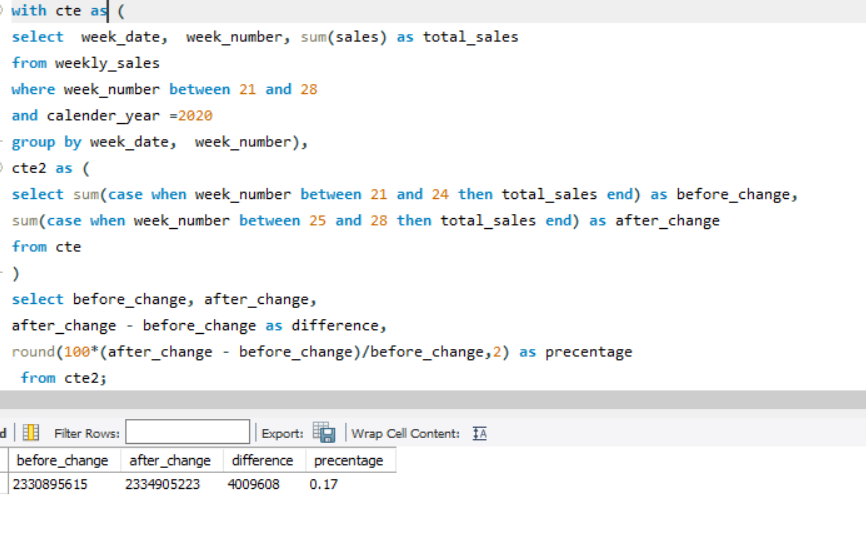
This technique is usually used when we inspect an important event and want to inspect the impact before and after a certain point in time.

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

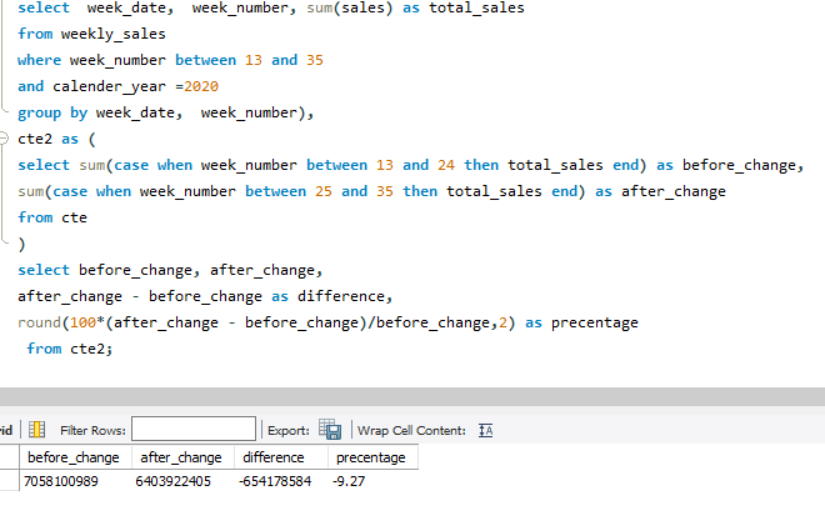
We would include all week\_date values for 2020-06-15 as the start of the period **after** the change and the previous week\_date values would be **before**

Using this analysis approach - answer the following questions:

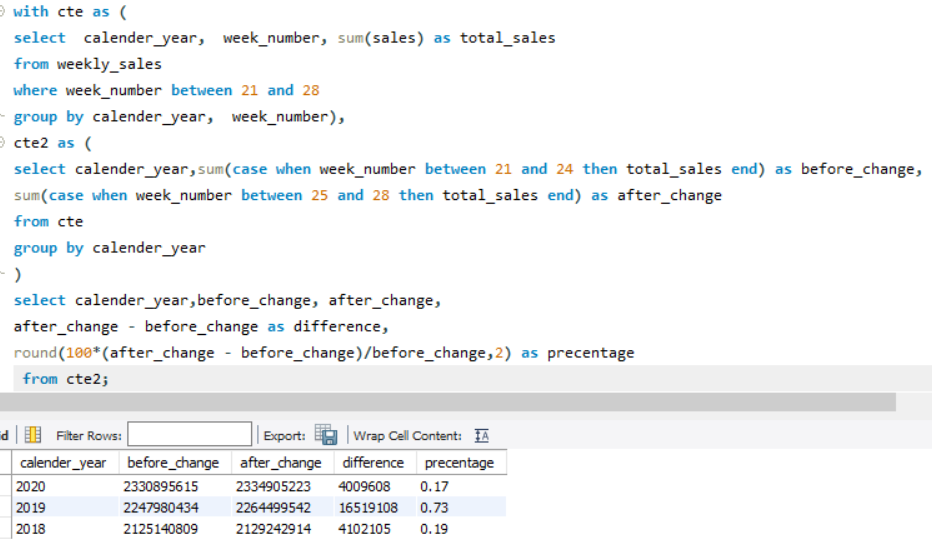
1.What is the total sales for the 4 weeks before and after 2020-06-15? What is the growth or reduction rate in actual values and percentage of sales?

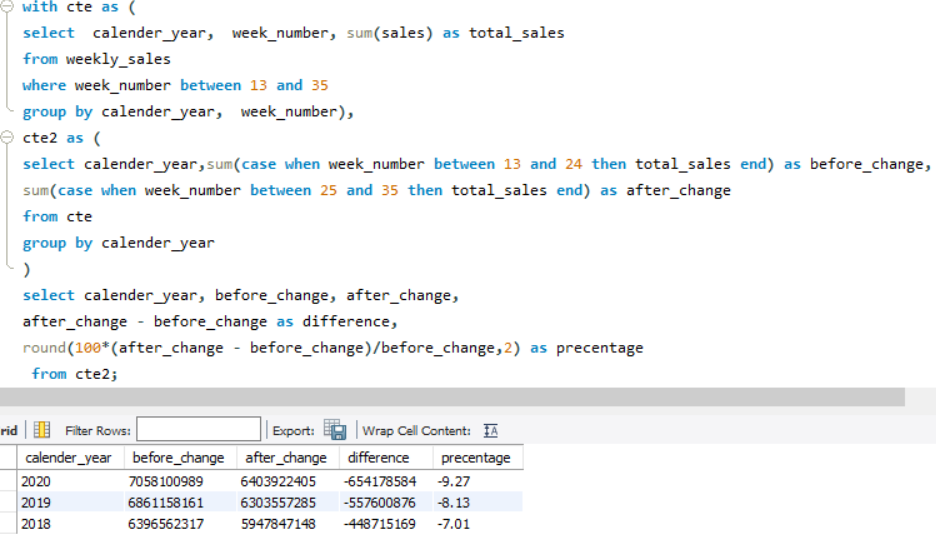


2. What about the entire 12 weeks before and after?



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





BONUS QUESTION

Which areas of the business have the highest negative impact in sales metrics performance in 2020 for the 12 week before and after period?

* region
* platform
* age\_band
* demographic
* customer\_type

