

8WeekSQLChallenge.com

CASE STUDY #5



DATA MART
fresh is best

DataWithDanny.com

Introduction

Data Mart is Danny's latest venture and after running international operations for his online supermarket that specialises in fresh produce - Danny is asking for your support to analyse his sales performance.

In June 2020 - large scale supply changes were made at Data Mart. All Data Mart products now use sustainable packaging methods in every single step from the farm all the way to the customer.

Danny needs your help to quantify the impact of this change on the sales performance for Data Mart and it's separate business areas.

The key business question he wants you to help him answer are the following:

- What was the quantifiable impact of the changes introduced in June 2020?
- Which platform, region, segment and customer types were the most impacted by this change?


- What can we do about future introduction of similar sustainability updates to the business to minimise impact on sales?

Available Data

For this case study there is only a single table: `data_mart.weekly_sales`

The **Entity Relationship Diagram** is shown below with the data types made clear, please note that there is only this one table - hence why it looks a little bit lonely!

data_mart.weekly_sales	
week_date	VARCHAR(7)
region	VARCHAR(13)
platform	VARCHAR(7)
segment	VARCHAR(4)
customer_type	VARCHAR(8)
transactions	INTEGER
sales	INTEGER



Column Dictionary

The columns are pretty self-explanatory based on the column names but here are some further details about the dataset:

1. Data Mart has international operations using a multi-**region** strategy
2. Data Mart has both, a retail and online **platform** in the form of a Shopify store front to serve their customers
3. Customer **segment** and **customer_type** data relates to personal age and demographics information that is shared with Data Mart
4. **transactions** is the count of unique purchases made through Data Mart and **sales** is the actual dollar amount of purchases

Each record in the dataset is related to a specific aggregated slice of the underlying sales data rolled up into a **week_date** value which represents the start of the sales week.

