



8 Week SQL Challenge

Start your SQL learning journey today!



Data With Danny Case Studies Getting Started Resources About

Case Study #7 - Balanced Tree Clothing Co.

Danny Ma · July 2, 2021

8WEEKSQLCHALLENGE.COM
CASE STUDY #7



BALANCED TREE
CLOTHING COMPANY

DATAWITHDANNY.COM

Balanced Tree Clothing Company prides themselves on providing an optimised range of clothing and lifestyle wear for the modern adventurer!

Danny, the CEO of this trendy fashion company has asked you to assist the team's merchandising teams analyse their sales performance and generate a basic financial report to share with the wider business.

Available Data

For this case study there is a total of 4 datasets for this case study - however you will only need to utilise 2 main tables to solve all of the regular questions, and the additional 2 tables are used only for the bonus challenge question!

Product Details

`balanced_tree.product_details` includes all information about the entire range that Balanced Clothing sells in their store.

product_id	price	product_name	category_id	segment_id
c4a632	13	Navy Oversized Jeans - Womens	1	3
e83aa3	32	Black Straight Jeans - Womens	1	3
e31d39	10	Cream Relaxed Jeans - Womens	1	3
d5e9a6	23	Khaki Suit Jacket - Womens	1	4
72f5d4	19	Indigo Rain Jacket - Womens	1	4
9ec847	54	Grey Fashion Jacket - Womens	1	4
5d267b	40	White Tee Shirt - Mens	2	5
c8d436	10	Teal Button Up Shirt - Mens	2	5
2a2353	57	Blue Polo Shirt - Mens	2	5

product_id	price	product_name	category_id	segment_id
f084eb	36	Navy Solid Socks - Mens	2	6
b9a74d	17	White Striped Socks - Mens	2	6
2feb6b	29	Pink Fluro Polkadot Socks - Mens	2	6

Product Sales

`balanced_tree.sales` contains product level information for all the transactions made for Balanced Tree including quantity, price, percentage discount, member status, a transaction ID and also the transaction timestamp.

prod_id	qty	price	discount	member	txn_id	start_txn_time
c4a632	4	13	17	t	54f307	2021-02-13 01:59:43.296
5d267b	4	40	17	t	54f307	2021-02-13 01:59:43.296
b9a74d	4	17	17	t	54f307	2021-02-13 01:59:43.296
2feb6b	2	29	17	t	54f307	2021-02-13 01:59:43.296
c4a632	5	13	21	t	26cc98	2021-01-19 01:39:00.3456
e31d39	2	10	21	t	26cc98	2021-01-19 01:39:00.3456
72f5d4	3	19	21	t	26cc98	2021-01-19 01:39:00.3456
2a2353	3	57	21	t	26cc98	2021-01-19 01:39:00.3456
f084eb	3	36	21	t	26cc98	2021-01-19 01:39:00.3456
c4a632	1	13	21	f	ef648d	2021-01-27 02:18:17.1648

Product Hierarchy & Product Price

These tables are used only for the bonus question where we will use them to recreate the

`balanced_tree.product_details` table.

`balanced_tree.product_hierarchy`

id	parent_id	level_text	level_name
1		Womens	Category
2		Mens	Category
3	1	Jeans	Segment
4	1	Jacket	Segment
5	2	Shirt	Segment
6	2	Socks	Segment
7	3	Navy Oversized	Style
8	3	Black Straight	Style
9	3	Cream Relaxed	Style
10	4	Khaki Suit	Style
11	4	Indigo Rain	Style
12	4	Grey Fashion	Style
13	5	White Tee	Style
14	5	Teal Button Up	Style
15	5	Blue Polo	Style
16	6	Navy Solid	Style
17	6	White Striped	Style
18	6	Pink Fluro Polkadot	Style

balanced_tree.product_prices

id	product_id	price
7	c4a632	13
8	e83aa3	32
9	e31d39	10
10	d5e9a6	23
11	72f5d4	19
12	9ec847	54
13	5d267b	40
14	c8d436	10
15	2a2353	57
16	f084eb	36
17	b9a74d	17
18	2feb6b	29

Interactive SQL Instance

You can use the embedded DB Fiddle below to easily access these example datasets - this interactive session has everything you need to start solving these questions using SQL.

You can click on the **Edit on DB Fiddle** link on the top right hand corner of the embedded session below and it will take you to a fully functional SQL editor where you can write your own queries to analyse the data.

You can feel free to choose any SQL dialect you'd like to use, the existing Fiddle is using PostgreSQL 13 as default.

Serious SQL students will have access to the same relevant schema SQL and example solutions which they can use with their Docker setup from within the course player!

Schema SQL Query SQL Results Edit on DB Fiddle

```
CREATE SCHEMA balanced_tree;

CREATE TABLE balanced_tree.product_hierarchy (
    "id" INTEGER,
    "parent_id" INTEGER,
    "level_text" VARCHAR(19),
    "level_name" VARCHAR(8)
);

INSERT INTO balanced_tree.product_hierarchy
("id", "parent_id", "level_text", "level_name")
VALUES
('1', NULL, 'Womens', 'Category'),
('2', NULL, 'Mens', 'Category'),
('3', '1', 'Jeans', 'Segment'),
('4', '1', 'Jacket', 'Segment'),
('5', '2', 'Shirt', 'Segment'),
('6', '2', 'Socks', 'Segment'),
('7', '3', 'Navy Oversized', 'Style'),
('8', '3', 'Black Straight', 'Style'),
('9', '3', 'Cream Relaxed', 'Style'),
('10', '4', 'Khaki Suit', 'Style'),
('11', '4', 'Indigo Rain', 'Style'),
('12', '4', 'Grey Fashion', 'Style'),
('13', '5', 'White Tee', 'Style'),
('14', '5', 'Teal Button Up', 'Style'),
('15', '5', 'Blue Polo', 'Style'),
('16', '6', 'Navy Solid', 'Style'),
('17', '6', 'White Striped', 'Style'),
('18', '6', 'Pink Fluro Polkadot', 'Style');
```

Case Study Questions

The following questions can be considered key business questions and metrics that the Balanced Tree team requires for their monthly reports.

Each question can be answered using a single query - but as you are writing the SQL to solve each individual problem, keep in mind how you would generate all of these metrics in a single SQL script which the Balanced Tree team can run each month.

High Level Sales Analysis

1. What was the total quantity sold for all products?
2. What is the total generated revenue for all products before discounts?
3. What was the total discount amount for all products?

Transaction Analysis

1. How many unique transactions were there?
2. What is the average unique products purchased in each transaction?
3. What are the 25th, 50th and 75th percentile values for the revenue per transaction?
4. What is the average discount value per transaction?
5. What is the percentage split of all transactions for members vs non-members?
6. What is the average revenue for member transactions and non-member transactions?

Product Analysis

1. What are the top 3 products by total revenue before discount?
2. What is the total quantity, revenue and discount for each segment?
3. What is the top selling product for each segment?
4. What is the total quantity, revenue and discount for each category?
5. What is the top selling product for each category?
6. What is the percentage split of revenue by product for each segment?
7. What is the percentage split of revenue by segment for each category?
8. What is the percentage split of total revenue by category?
9. What is the total transaction “penetration” for each product? (hint: penetration = number of transactions where at least 1 quantity of a product was purchased divided by total number of transactions)
10. What is the most common combination of at least 1 quantity of any 3 products in a 1 single transaction?

Reporting Challenge

Write a single SQL script that combines all of the previous questions into a scheduled report that the Balanced Tree team can run at the beginning of each month to calculate the previous month's values.

Imagine that the Chief Financial Officer (which is also Danny) has asked for all of these questions at the end of every month.

He first wants you to generate the data for January only - but then he also wants you to demonstrate that you can easily run the same analysis for February without many changes (if at all).

Feel free to split up your final outputs into as many tables as you need - but be sure to explicitly reference which table outputs relate to which question for full marks :)

Bonus Challenge

Use a single SQL query to transform the `product_hierarchy` and `product_prices` datasets to the `product_details` table.

Hint: you may want to consider using a recursive CTE to solve this problem!

Conclusion

Sales, transactions and product exposure is always going to be a main objective for many data analysts and data scientists when working within a company that sells some type of product - Spoiler alert: nearly all companies will sell products!

Being able to navigate your way around a product hierarchy and understand the different levels of the structures as well as being able to join these details to sales related datasets will be super valuable for anyone wanting to work within a financial, customer or exploratory analytics capacity.

Hopefully these questions helped provide some exposure to the type of analysis we perform daily in these sorts of roles!

Ready for the next 8 Week SQL challenge case study? Click on the banner below to get started with case study #8!

[8WEEKSQLCHALLENGE.COM](https://8weeksqlchallenge.com/case-study-7/)

CASE STUDY #8



FRESH SEGMENTS
EXTRACT MAXIMUM VALUE

DATAWITHDANNY.COM

Official Solutions

If you'd like to see the official code solutions and explanations for this case study and a whole lot more, please consider joining me for the [Serious SQL course](#) - you'll get access

to all course materials and I'm on hand to answer all of your additional SQL questions directly!

Serious SQL is priced at \$49USD and \$29 for students and includes access to all written course content, community events as well as live and recorded SQL training videos!

Please send an email to support@datawithdanny.com from your educational email or include your enrolment details or student identification for a speedy response!

Community Solutions

This section will be updated in the future with any community member solutions with a link to their respective GitHub repos!

Final Thoughts

The 8 Week SQL Challenge is proudly brought to you by me - Danny Ma and the [Data With Danny](#) virtual data apprenticeship program.

Students or anyone undertaking further studies are eligible for a \$20USD student discount off the price of Serious SQL please send an email to support@datawithdanny.com from your education email or include information about your enrolment for a fast response!

We have a large student community active on the official DWD Discord server with regular live events, trainings and workshops available to all Data With Danny students, plus early discounted access to all future paid courses.

There are also opportunities for 1:1 mentoring, resume reviews, interview training and more from myself or others in the DWD Mentor Team.

From your friendly data mentor, Danny :)

All 8 Week SQL Challenge Case Studies

All of the 8 Week SQL Challenge case studies can be found below:

- [Case Study #1 - Danny's Diner](#)

- Case Study #2 - Pizza Runner
- Case Study #3 - Foodie-Fi
- Case Study #4 - Data Bank
- Case Study #5 - Data Mart
- Case Study #6 - Clique Bait
- Case Study #7 - Balanced Tree Clothing Co.
- Case Study #8 - Fresh Segments

Share: [Twitter](#), [Facebook](#)

