



MEAL KIT DELIVERY ANALYSIS

CREATED BY: SANG LU (SANDRA)

PERFECT QUALITY



INTRODUCTION TO COMPANY

HelloFresh is a popular meal kit delivery service that operates in various countries and has gained widespread recognition for its convenience and fresh, high-quality ingredients. Founded in 2011, the company aims to simplify the cooking experience for busy individuals and families by providing pre-portioned ingredients and easy-to-follow recipes delivered right to their doorstep.

Customers can choose from a diverse range of weekly recipes, catering to different dietary preferences and restrictions, such as vegetarian, vegan, gluten-free, and more. With HelloFresh, users can enjoy the benefits of home-cooked meals without the hassle of grocery shopping and meal planning.



**HELLO
FRESH**

QUESTIONS

- ▶ For each week of 2020 in the US, how many boxes did we deliver, and what revenue (in Euros) did those deliveries generate?
- ▶ How many delivery weeks did the years 2019 and 2020 have?
- ▶ Fetch the customer IDs of all HelloFresh customers in the US who received their box for free on January 15, 2021.
- ▶ What proportion of US-based customers chose to have a “Vegetarian Box” delivered in the past 7 days? Bonus points if you make the range dynamic.
- ▶ On average, how loyal are iOS users in the US compared to Android users? (Loyalty can be defined as the total count of boxes the customer has received ever since they joined HelloFresh)
- ▶ How many customers have ordered more than one type of product since they joined HelloFresh?

TABLE SCHEMA

| Table | boxes_shipped | customers | products |
|--------|---|---|---|
| Schema | box_id: int, delivery_date: string, customer_id: int, product_id: int, revenue_eur: float, country: string, delivery_week: string | customer_id: int, country: string, device_type: string, customer_since: string | product_id: int, country: string, product_name: string, product_family: string |

ADDITIONAL INFORMATION

- ▶ All tables might have additional columns, which are not necessary in this example.
- ▶ You may not need to use all tables in every task.
- ▶ All tables are clean and have no duplicate records, and contain data for 2016 onward.
- ▶ A given customer, in a given country, may have between 0 and 1 delivered boxes for a given week.
- ▶ Product_name can be a ‘Classic Box’, ‘Family Box’, or ‘Vegetarian Box’. Each product_name is assigned a product_id.
- ▶ The “device_type” can be “iOS” or “Android”.
- ▶ Customer IDs are unique within each country, but not necessarily across countries.
- ▶ Dates are stored in a string, with format YYYY-MM-DD.
- ▶ boxes_shipped.delivery_week follows the format YYYY-WXX
e.g: 2021-W01 = the first week of 2021, followed by week 2021-W02, and so on).

QUESTION 1 - For each week of 2020 in the US, how many boxes did we deliver, and what revenue (in Euros) did those deliveries generate?

```
SELECT delivery_week,  
       COUNT(box_id) AS NumberOfBoxesDelivered,  
       SUM(revenue_eur) AS TotalRevenueGenerated  
FROM boxes_shipped  
WHERE YEAR(delivery_date) = 2020 AND country = 'US'  
GROUP BY 1  
ORDER BY 1;
```

QUESTION 2 - How many delivery weeks did the years 2019 and 2020 have?

```
SELECT YEAR(delivery_date) AS year,  
       COUNT(DISTINCT delivery_week) AS NumberOfDelivery_Weeks  
FROM boxes_shipped  
WHERE YEAR(delivery_date) = 2019  
      OR YEAR(delivery_date) = 2020  
GROUP BY 1;
```

snappify.com

QUESTION 3 - Fetch the customer IDs of all HelloFresh customers in the US who received their box for free on January 15, 2021.

```
SELECT DISTINCT customer_id
FROM boxes_shipped
WHERE (
    revenue_eur = 0
    OR revenue_eur IS NULL
)
AND country = 'US'
AND DATE_FORMAT(delivery_date, '%Y-%m-%d') = '2021-01-15';
```

QUESTION 4 - What proportion of US-based customers chose to have a “Vegetarian Box” delivered in the past 7 days? Bonus points if you make the range dynamic.

```
CREATE FUNCTION DynamicRange(N INT)
RETURNS FLOAT(2) AS
BEGIN
    WITH past_N_days_boxes_shipped AS (
        SELECT customer_id AS US_cusotmer_id,
        (
            CASE
                WHEN product_id = (
                    SELECT product_id
                    FROM products
                    WHERE product_name = 'Vegetarian Box'
                )
                AND DATEDIFF(CURDATE(), DATE(delivery_date)) BETWEEN 0 AND N THEN customer_id
                ELSE NULL
            END
        ) AS US_Vegetarians_id_past7days
        FROM boxes_shipped
        WHERE country = 'US'
    )
    SELECT ROUND(
        COUNT(DISTINCT US_Vegetarians_id_past7days) / COUNT(DISTINCT US_cusotmer_id) * 100,
        2
    ) AS %_US_Vegetarians_past7days
    FROM past_N_days_boxes_shipped;
END;
```

QUESTION 5 - On average, how loyal are iOS users in the US compared to Android users?

```
SELECT device_type,
       IFNULL(COUNT(box_id) / COUNT(x.customer_id), 0) AS avg_US_customer_loyalty
FROM boxes_shipped AS x
      JOIN customers AS c ON x.customer_id = c.customer_id
      AND x.country = c.country
      AND c.country = 'US'
GROUP BY 1;
```

QUESTION 6 - How many customers have ordered more than one type of product since they joined HelloFresh?

```
SELECT COUNT(*) AS customers_OrderMoreThanOneProjects
FROM boxes_shipped
GROUP BY customer_id,
         country
HAVING COUNT(DISTINCT product_id) > 1;
```

snappyf.com

Thank you!