Task two- The client has selected store numbers **77, 86 and 88** as trial stores and want control stores to be established stores that are operational for the entire observation period. We would want to match trial stores to control stores that are similar to the trial store prior to the trial period of Feb 2019 in terms of : -

1. Monthly overall sales revenue
2. Monthly number of customers
3. Monthly number of transactions per customer

Step 1: Filter Pre-Trial Period (before Feb 2019)

```
CREATE OR REPLACE TABLE `river-hold-450804-s3.product_sales.pre_trial_data` AS

SELECT *

FROM `river-hold-450804-s3.product_sales.data`

WHERE DATE < '2019-02-01'; (A new table is created preview table)
```

Step 2: Compute Monthly Metrics per Store

```
CREATE OR REPLACE TABLE `river-hold-450804-s3.product_sales.monthly_store_metrics` AS

SELECT

 STORE_NBR,

 FORMAT_DATE('%Y-%m', DATE) AS month,

 SUM(TOT_SALES) AS monthly_revenue,

 COUNT(DISTINCT LYLTY_CARD_NBR) AS monthly_unique_customers,

 COUNT(DISTINCT TXN_ID) AS monthly_transactions,

 ROUND(COUNT(DISTINCT TXN_ID)*1.0 / COUNT(DISTINCT LYLTY_CARD_NBR), 2) AS
txn_per_customer

FROM `river-hold-450804-s3.product_sales.pre_trial_data`

GROUP BY STORE_NBR, month;
```

This gives one row per **store + month** with:

- Total revenue
- Unique customers
- Number of transactions
- Transactions per customer(A new table is created preview table)

Step 3: Separate Trial vs Candidate Control Stores

```
-- Trial stores 77, 86, 88

CREATE OR REPLACE TABLE `river-hold-450804-s3.product_sales.trial_store_metrics` AS

SELECT *
```

```
FROM `river-hold-450804-s3.product_sales.monthly_store_metrics`
WHERE STORE_NBR IN (77, 86, 88);

-- All other stores as candidates
CREATE OR REPLACE TABLE `river-hold-450804-s3.product_sales.control_store_candidates` AS
SELECT *
FROM `river-hold-450804-s3.product_sales.monthly_store_metrics`
WHERE STORE_NBR NOT IN (77, 86, 88);
```

(A table is created and contains only these 3 stores information.)

Step 4: Compare Trial Stores with Candidate Stores

```
CREATE OR REPLACE TABLE `river-hold-450804-s3.product_sales.store_similarity` AS
SELECT
  t.STORE_NBR AS trial_store,
  c.STORE_NBR AS candidate_store,
  AVG(ABS(t.monthly_revenue - c.monthly_revenue)) AS avg_revenue_diff,
  AVG(ABS(t.monthly_unique_customers - c.monthly_unique_customers)) AS avg_customer_diff,
  AVG(ABS(t.txn_per_customer - c.txn_per_customer)) AS avg_txn_per_cust_diff
FROM `river-hold-450804-s3.product_sales.trial_store_metrics` t
JOIN `river-hold-450804-s3.product_sales.control_store_candidates` c
  ON t.month = c.month
GROUP BY trial_store, candidate_store;
```

This calculates **how different each candidate store is** compared to each trial store.

(table similarity is created which contains details of 77 store like avg_revenue_diff,avg_customer_diff,etc).

Step 5: Pick the Best Control Store

```
SELECT *
FROM (
  SELECT
    *,
    ROW_NUMBER() OVER(
```

```
    PARTITION BY trial_store
    ORDER BY (avg_revenue_diff + avg_customer_diff + avg_txn_per_cust_diff) ASC
  ) AS rank
 FROM `river-hold-450804-s3.product_sales.store_similarity`
)
WHERE rank = 1;
```