

# Activity 03: Data Warehouse Optimization

---

In this activity, you will work with your team to optimize the performance of a query to make it return results in the shortest time possible.

## Team Challenge

WWI wants you to prove your ability to optimize queries using their data. They have provided the following query:

```
SELECT
  MIN(AvgPrice) as MinCustomerAvgPrice
  ,MAX(AvgPrice) as MaxCustomerAvgPrice
  ,MIN(AvgTotalAmount) as MinCustomerAvgTotalAmount
  ,MAX(AvgTotalAmount) as MaxCustomerAvgTotalAmount
  ,MIN(AvgProfitAmount) as MinAvgProfitAmount
  ,MAX(AvgProfitAmount) as MaxAvgProfitAmount
FROM
  (
    SELECT
      FS.CustomerID
      ,MIN(FS.Quantity) as MinQuantity
      ,MAX(FS.Quantity) as MaxQuantity
      ,AVG(FS.Price) as AvgPrice
      ,AVG(FS.TotalAmount) as AvgTotalAmount
      ,AVG(FS.ProfitAmount) as AvgProfitAmount
      ,COUNT(DISTINCT FS.StoreId) as DistinctStores
    FROM
      wwi_perf.Sale_Heap FS
    GROUP BY
      FS.CustomerId
  ) T
```

You are not allowed to modify this table (in fact, **you should not modify this table** as it may affect the other labs). However, they have asked if you can **create a new table** and optimize the query such that it runs much faster than the above one. Working with your team, just how fast can you make this query return results?

You must **not scale up the SQL Pool**! Make optimizations to improve query performance while remaining in the bounds of the current pool size. Scaling up the SQL Pool can cause resource problems in your region when running this lab with multiple participants.

Share your best query time with your learning adviser who will report it back for comparison with the results from other teams.

NOTE: You can experiment in your individual environments, however when selecting the environment in which to produce your "official" results you should pick the one with the largest SQL Pool size among those available in your Table Group.