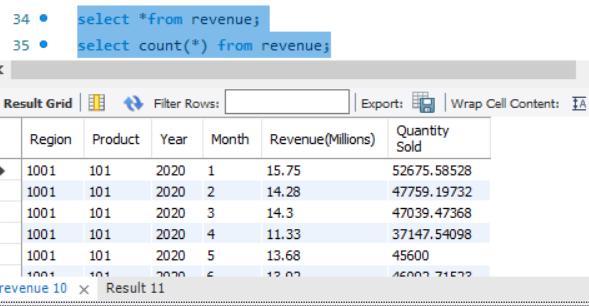
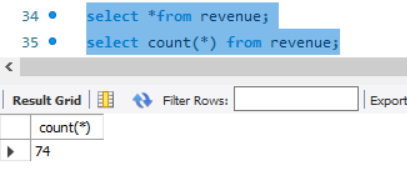
**DATA ENGINEER ASSIGNMENT – VISUAL BI Solutions**

->

First step is to create tables for data shared. Three tables were created namely revenue, region\_master, product\_master

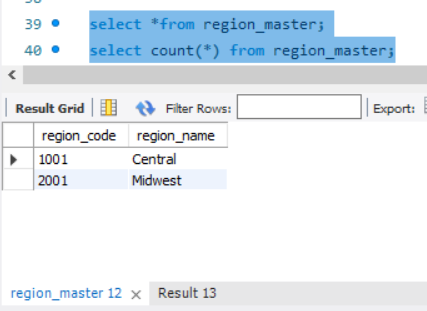
**Revenue:**

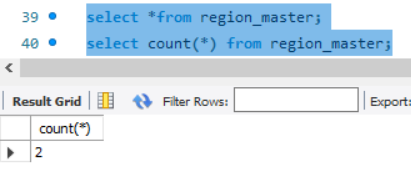




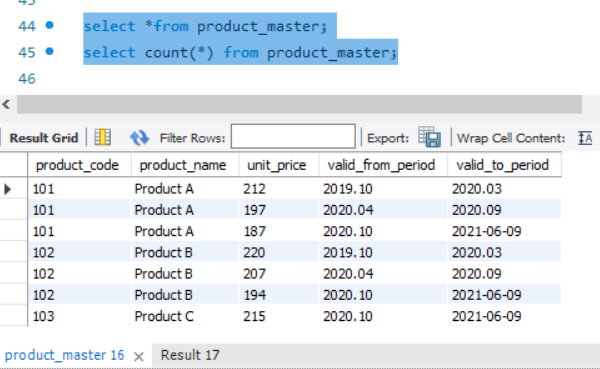
There were totally 74 records in the revenue table

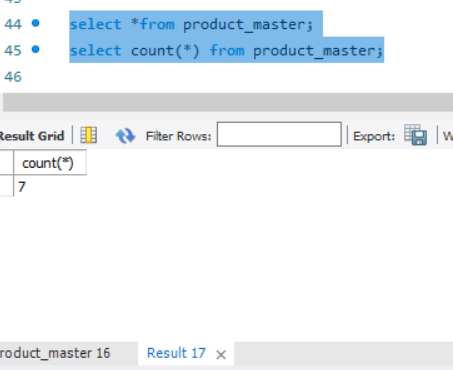
**Region Master:**





**Product Master:**



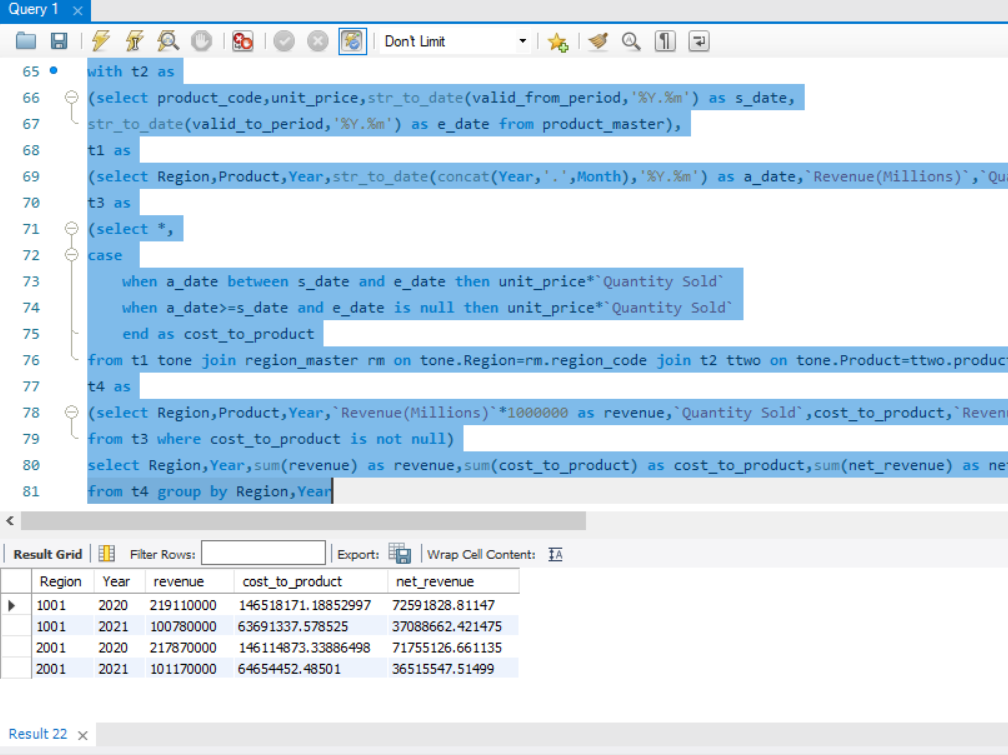


Please be noted that I have used current date in product master table while inserting values into table when the date field is 9999.12

**Question-1:**

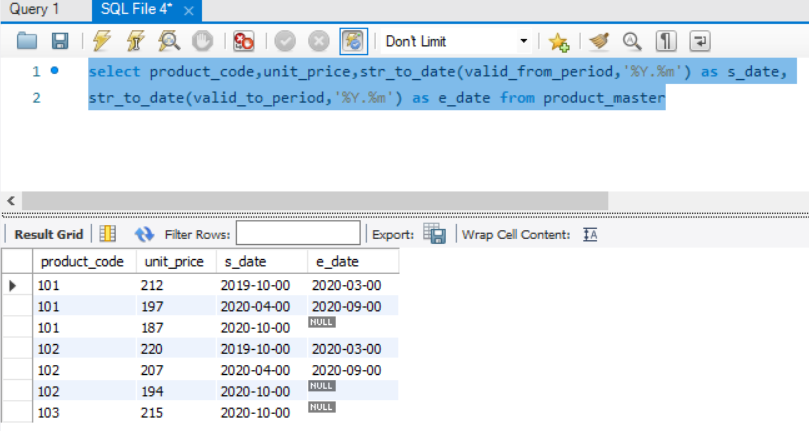
**Query to display Revenue, Cost to Product and Net Revenue for each region for year 2020 and 2021?**

**Query:**

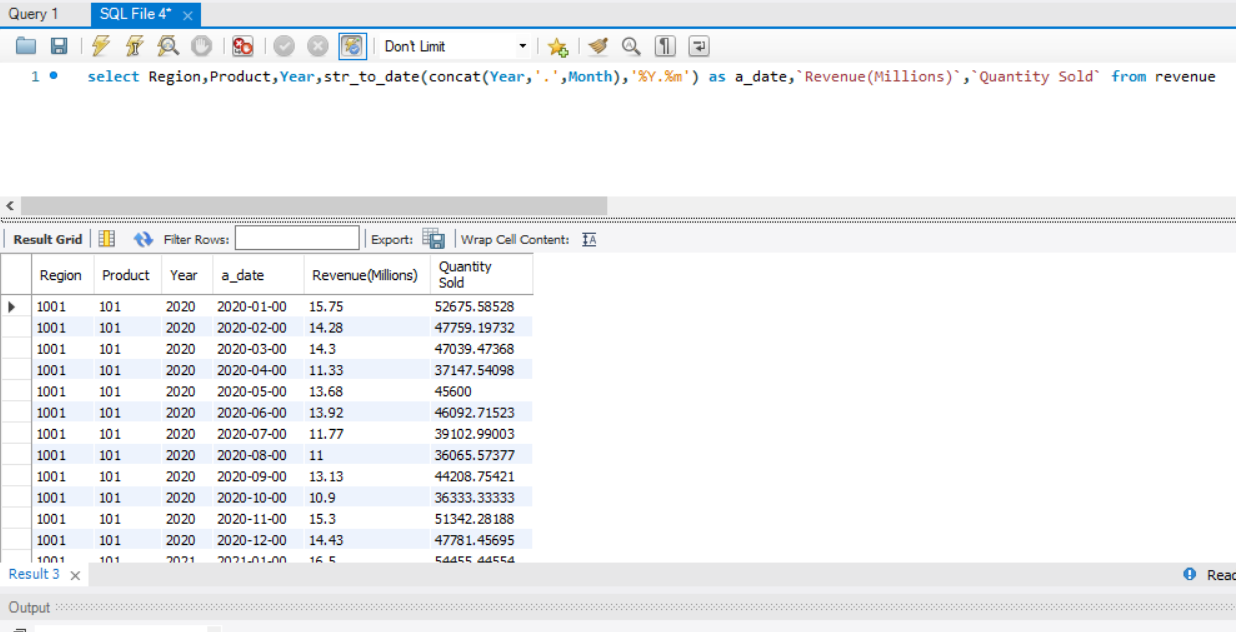


**Analysis and Approach:**

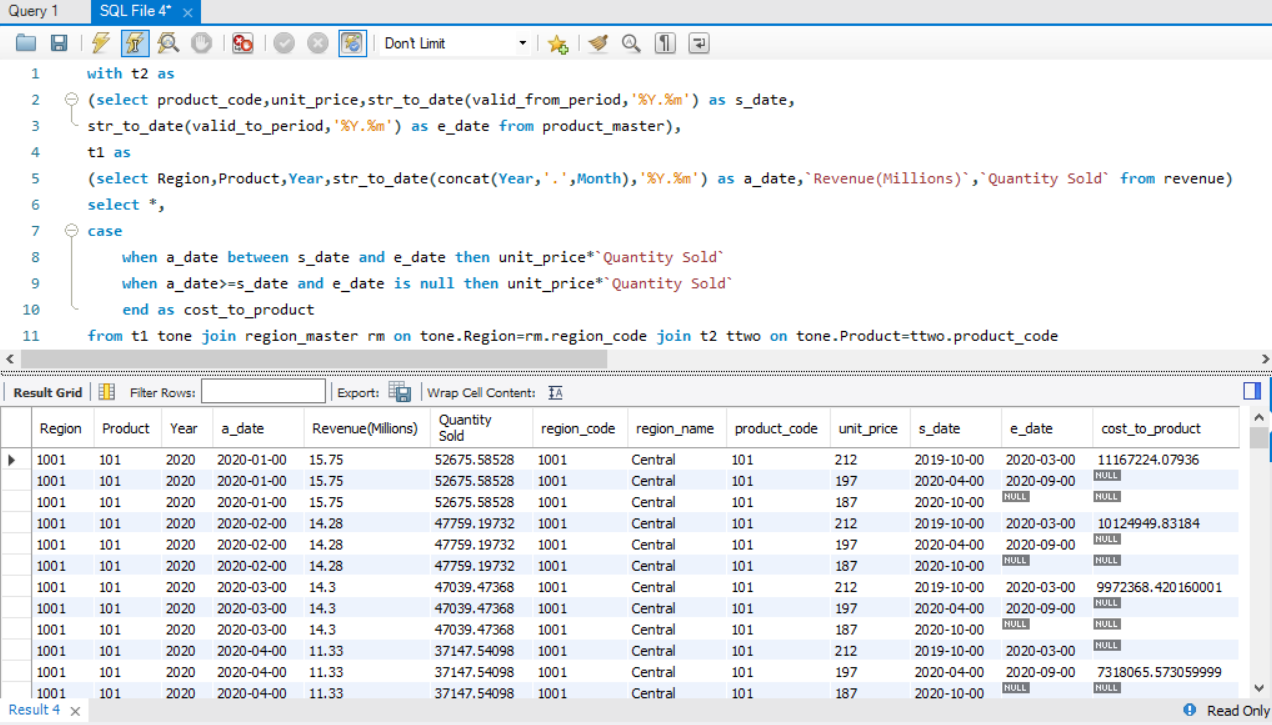
**Step1: Create CTE t2 for product\_master table and converting given dates to specified format**



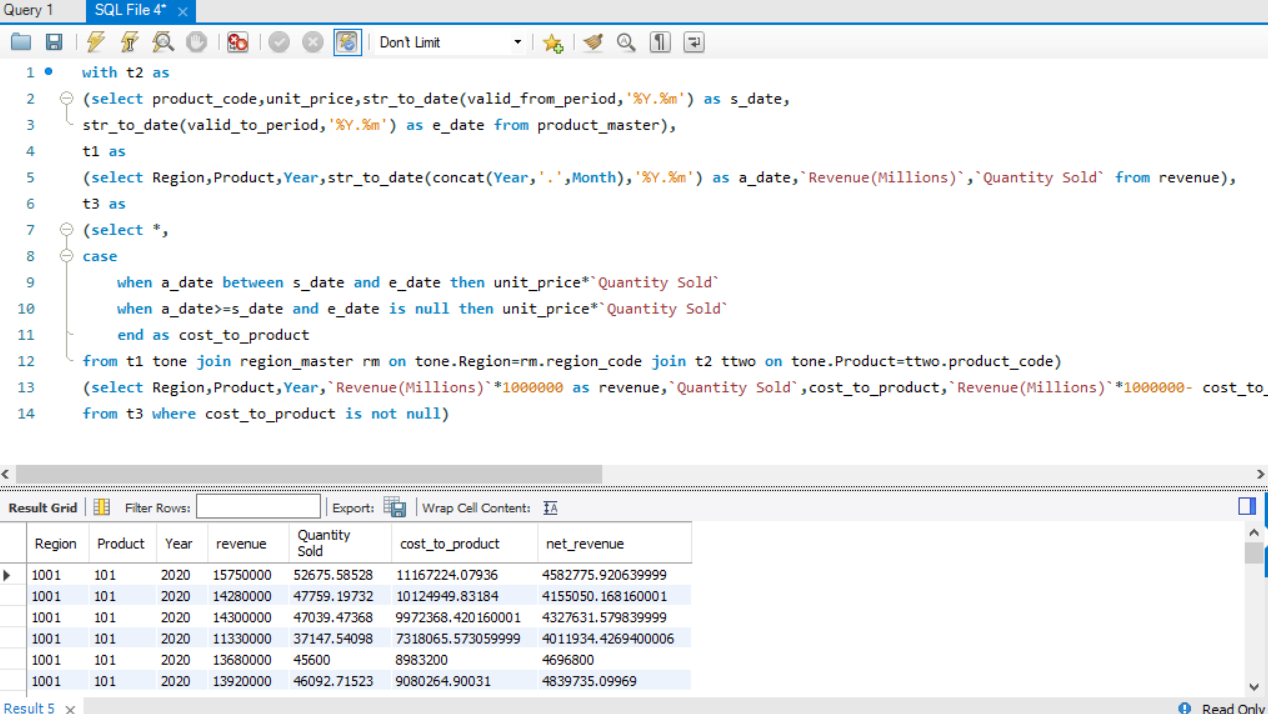
**Step2: Create CTE t1 from revenue table with dates formatting in the correct manner to compare while calculating cost to product for unit price**



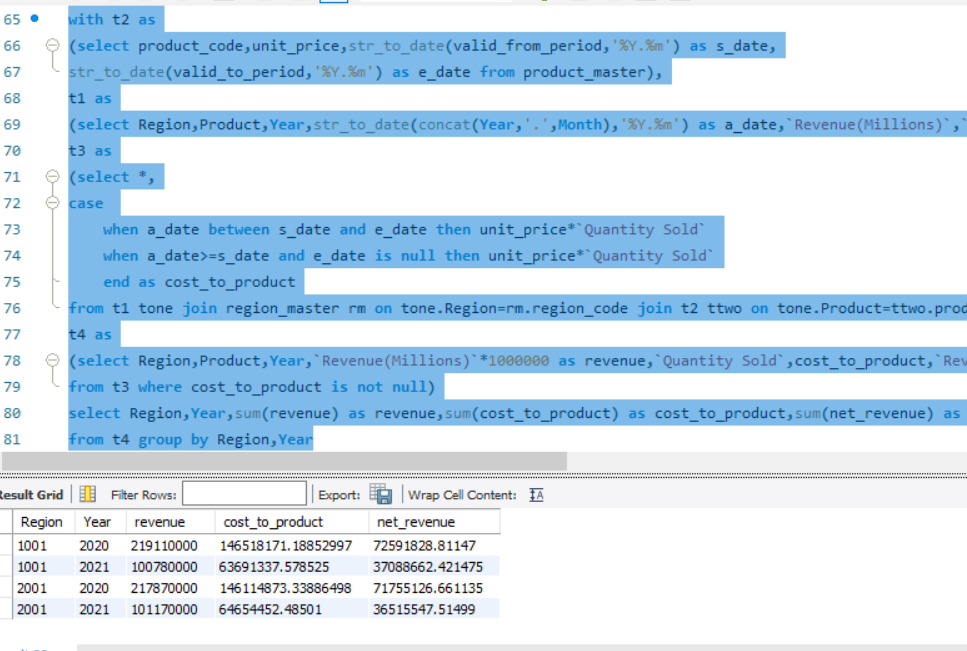
**Step3: Create CTE t3 which calculates cost to product. Here I validated whether actual date in revenue table is between start and end date in product master table for unit price**



**Step4: Create CTE t4 which extracts required reporting information and excluding not null values for cost to product**



**Step 5: From t4 CTE just need to aggregate revenue, cost to product and net revenue by region and year**

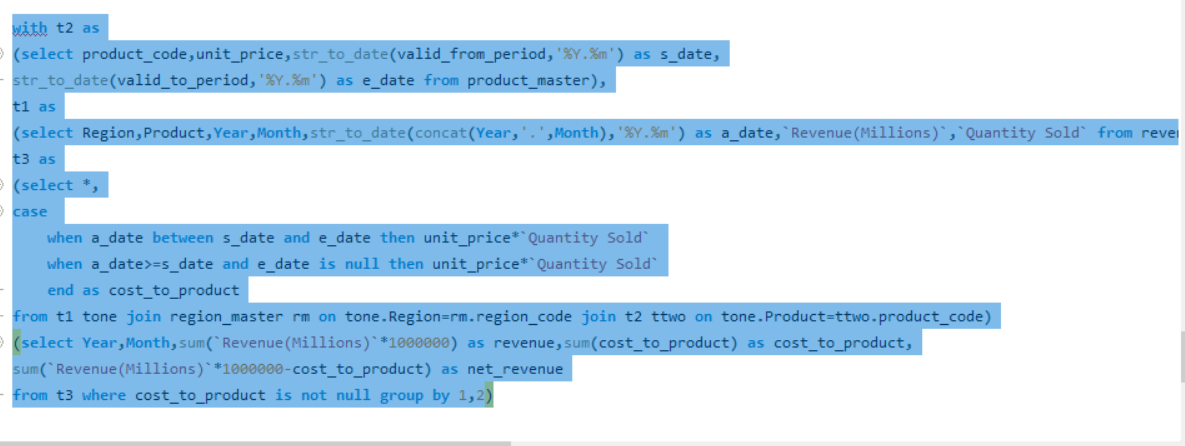


**Question-2: Query to display company’s YTD revenue, YTD cost\_to\_product and YTD net revenue for each month for both the years**

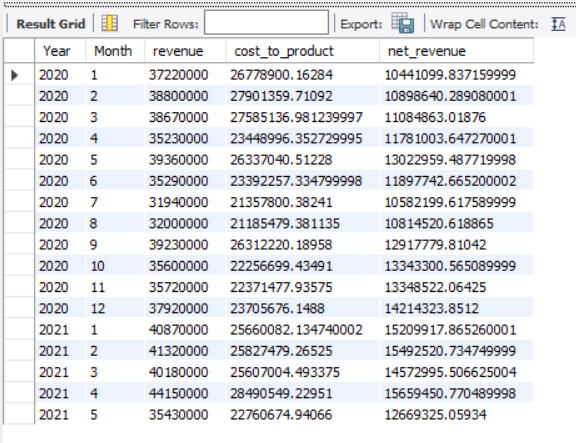
**Not Sure about YTD Year to Date:**

**First Approach considering Trend Chart which basically needs group by Year Month and Revenue:**

**Query:**



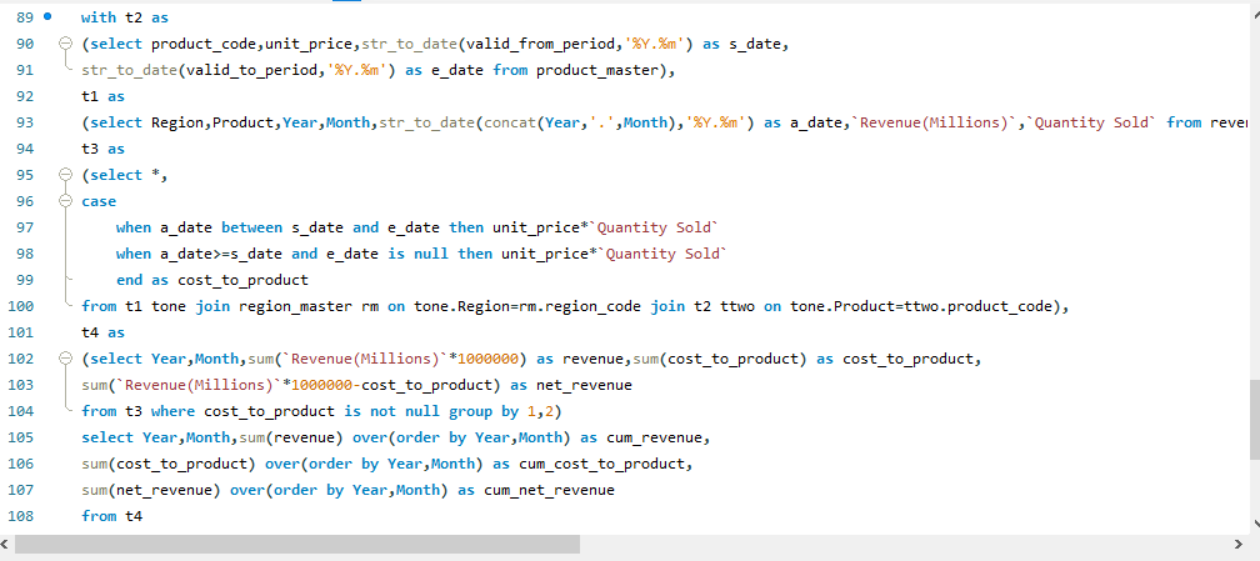
**Result Set:**



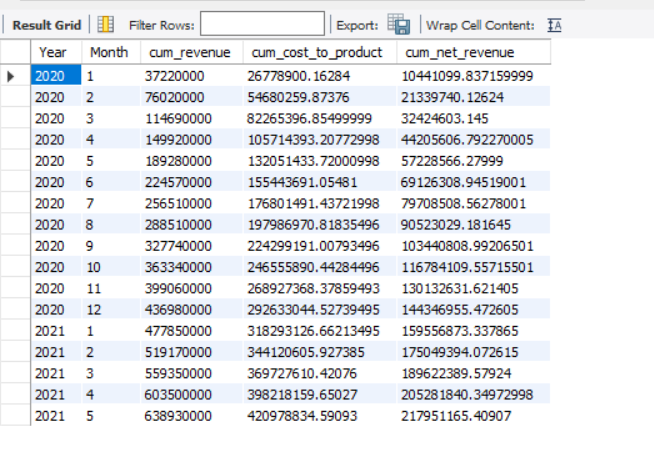
**Second Approach:**

**Considering Cumulative Sum which is Year to Date or up to date Revenue**

**Query:**



**Result set:**



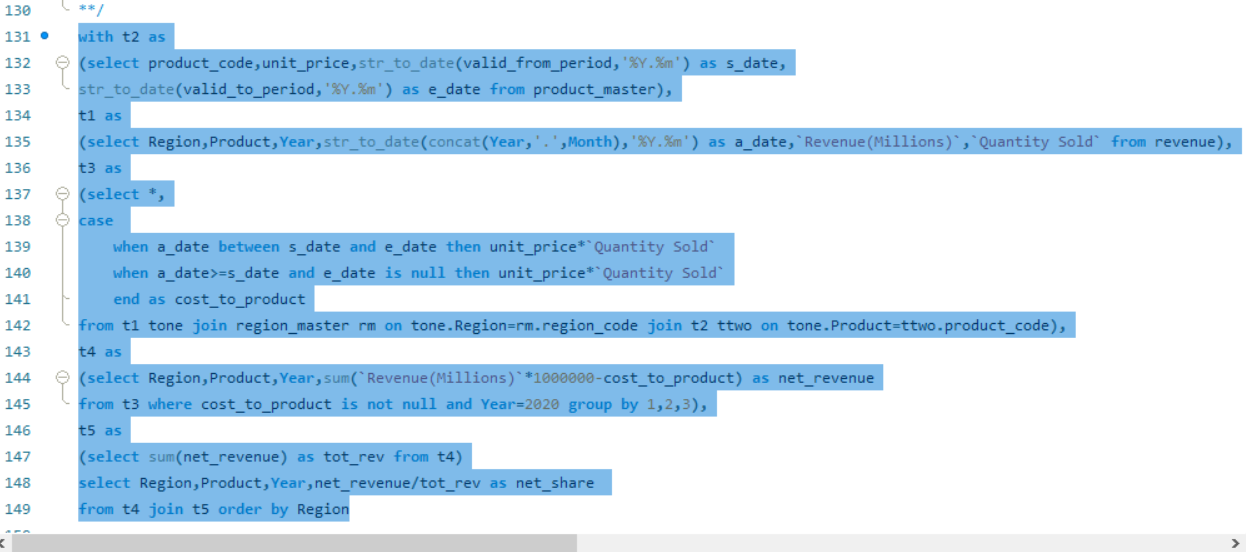
**Analysis and Approach:** Most of the approach is similar to that of Question-1 with adding few fields in the query to report in better view. Based on my understanding YTD Revenue, Cost to Product and Net Revenue for each month for both years is calculating the cumulative sum(second approach) but considering Trend Chart (First Approach)

Four CTE’s created in Q-1 were used with very minor changes like adding attributes based on requirement and key change is in Fourth CTE is aggregated by Year Month and calculated cumulative sum of revenue, cost to product and net revenue using Window functions

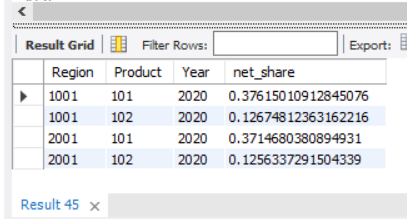
**Question-3:**

**Query to display Net Share of each Product in Revenue for each Region for Year 2020**

**Query:**



**Result Set:**



**Analysis and Approach:**

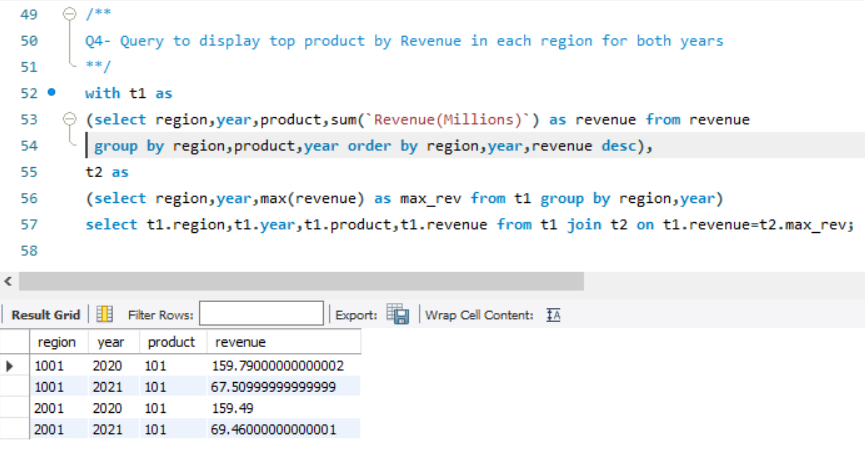
Firstly, based on my understanding Net Share is Net Revenue of each product in Year 2020 across each region divided by Sum Total of Net Revenue for all product across two regions in the Year 2020. If we need more granularity to check net share then needs to include group by region as well while calculating total revenue

Most of the CTE's created in Question-1 were useful with very minor changes in field names reporting and including them in the aggregations

**Question-4:**

**Query to display top product by Revenue in each Region for both Years?**

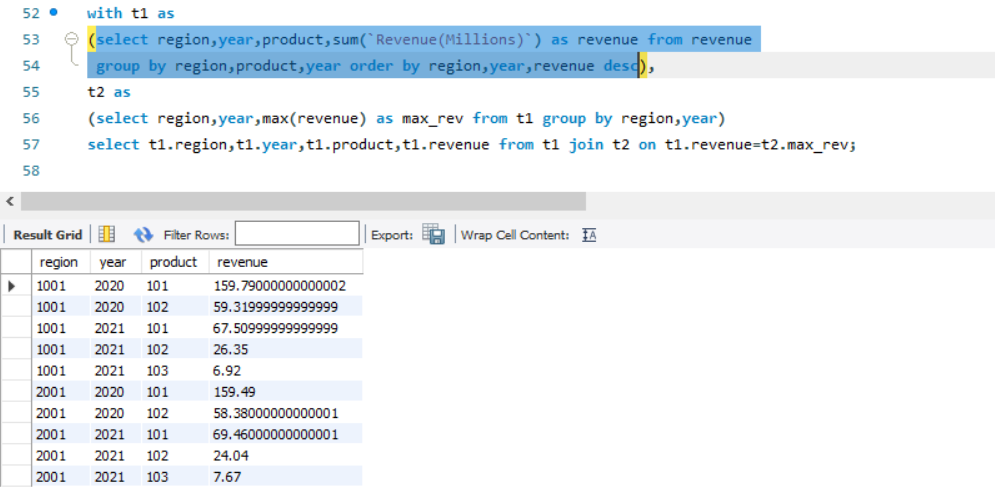
**Query:**



**Analysis and Approach:**

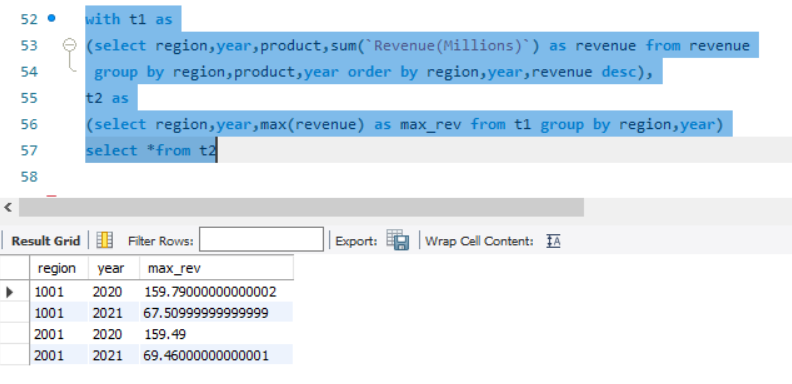
**Step1: Create CTE t1 which reports total revenue generated by each product across each region in each year**

The highlighted one in blue represents query for the above:

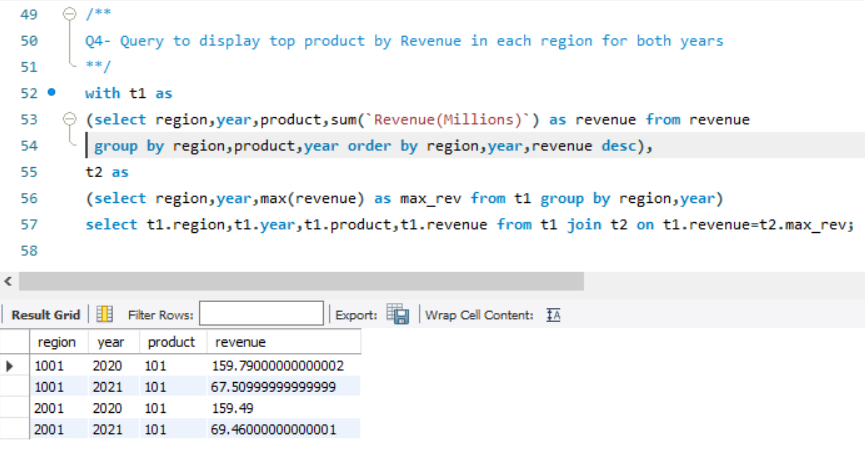


If we observe that it reports total revenue generated by different products across each region in each year. But our problem is to find the Product which generated maximum revenue across each region in different year.

**Step2: Create another CTE t2 and aggregate by region and year and calculate maximum revenue across each region and year**



**Step3: Join CTE t1 and CTE t2 on revenue as a result will be reporting only Top Products across each region in different year**



**Product 101 is the one which is generating maximum revenue in both regions and in both years**

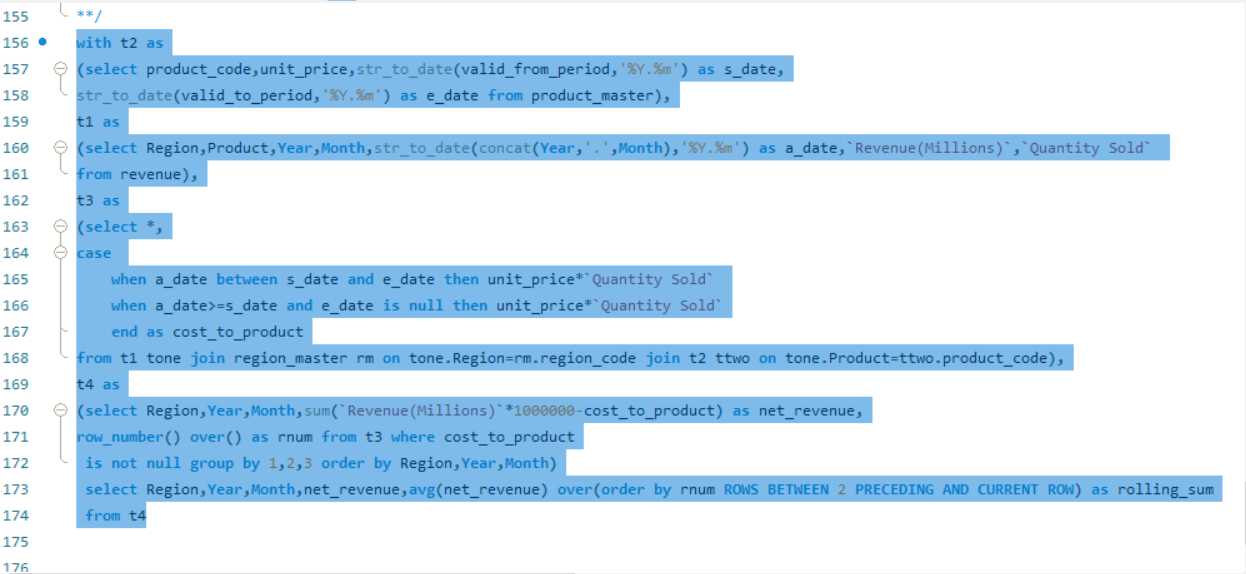
**Question-5:**

**Query to Display Average Rolling previous 3 Months Net Revenue for each Month and Region**

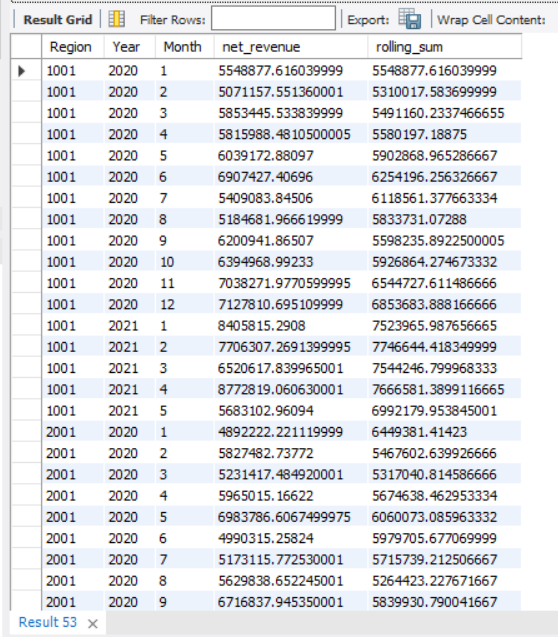
**Confused with respect to granularity to include Year or not based on past three months.**

**Based on my Approach considered last three months for 2021 January as 2020 November and 2020 December and 2021 January. The query mentioned below works for that however if we need to aggregate at month for each region without considering the year granularity the approach would be different with group by function excluding year**

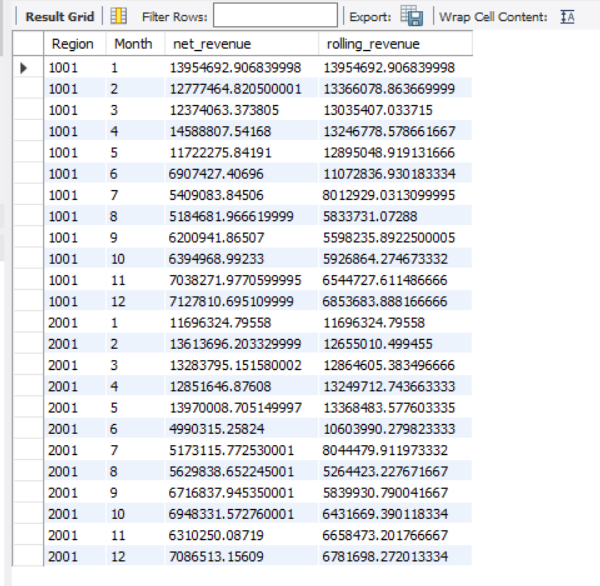
**Query:**

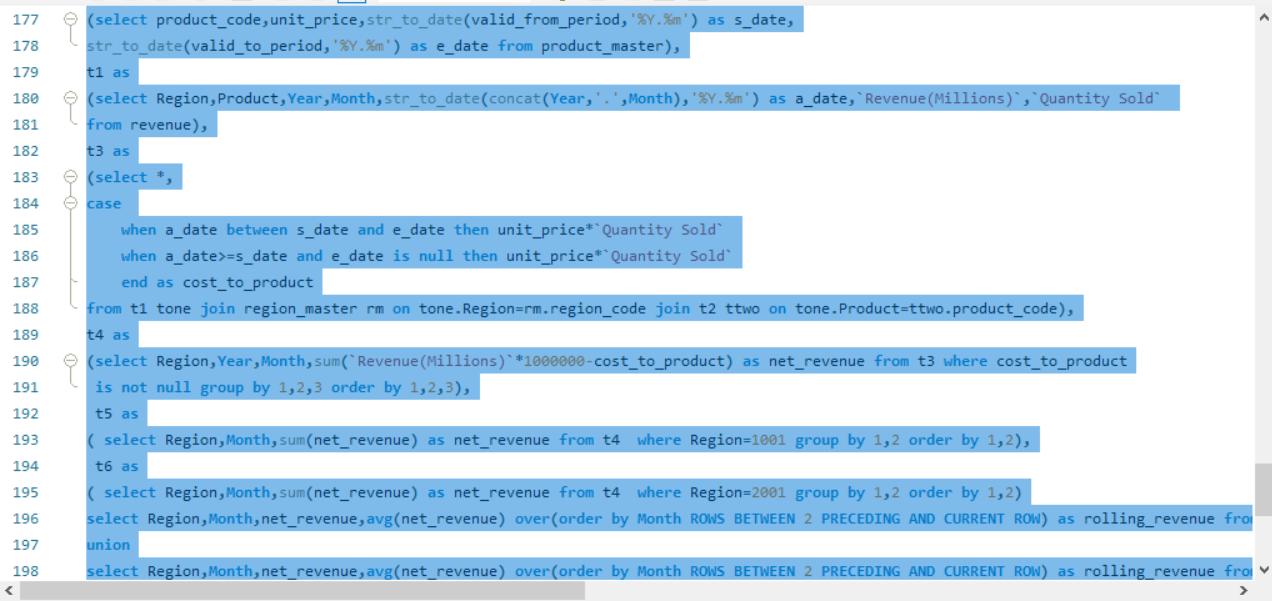


**Result set:**



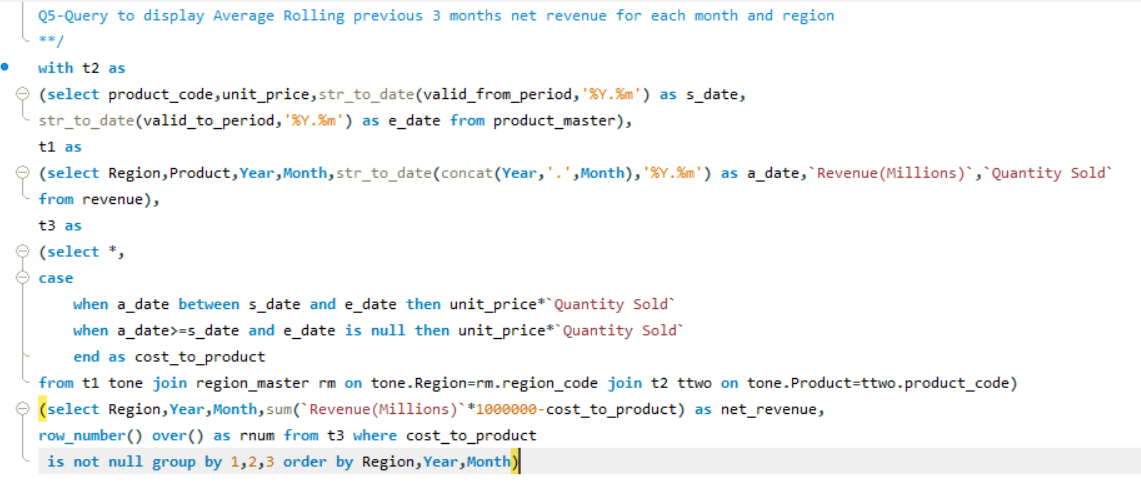
**Second Approach: excluding Year granularity**

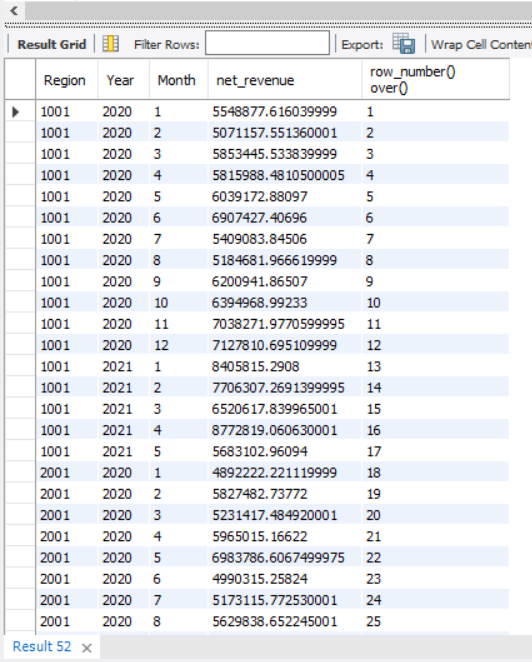




**Analysis and Approach:**

Most of CTE’s created in Question-2 were useful and in the first step calculated Region, Year, Month Net Revenue and include row number function to calculate rolling sum





Using this output as CTE and calculated the average rolling sum using windowing analytical functions as shown in the actual query and result set in First Approach

Second Approach:

**From this output the approach is same by using Union Concept filtering Region 2001 and 2002 and then calculate Avg rolling revenue and then union on these two**

