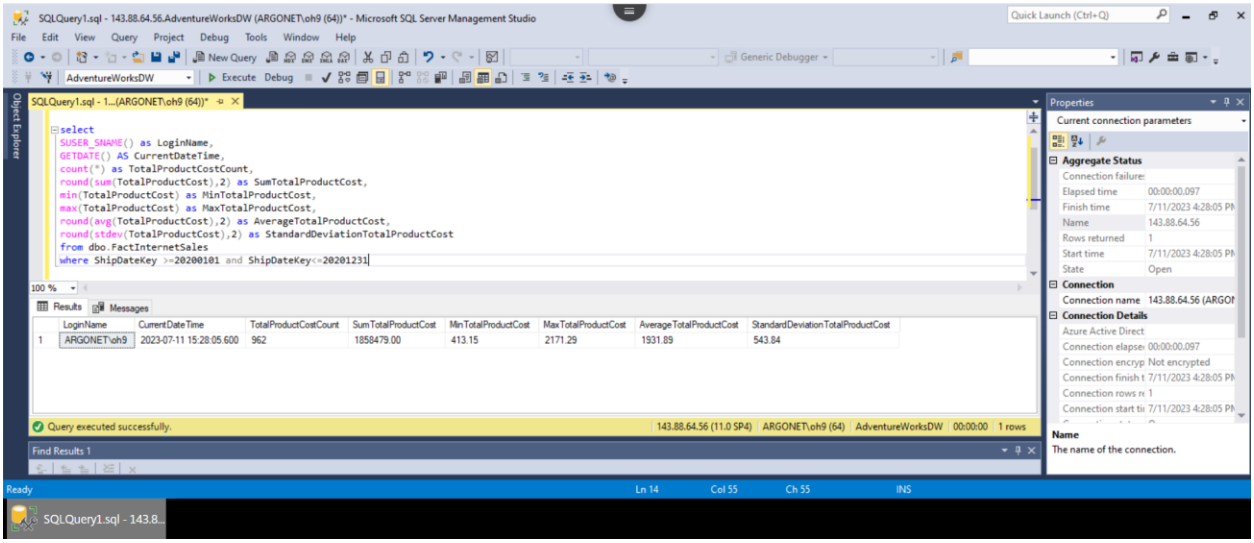
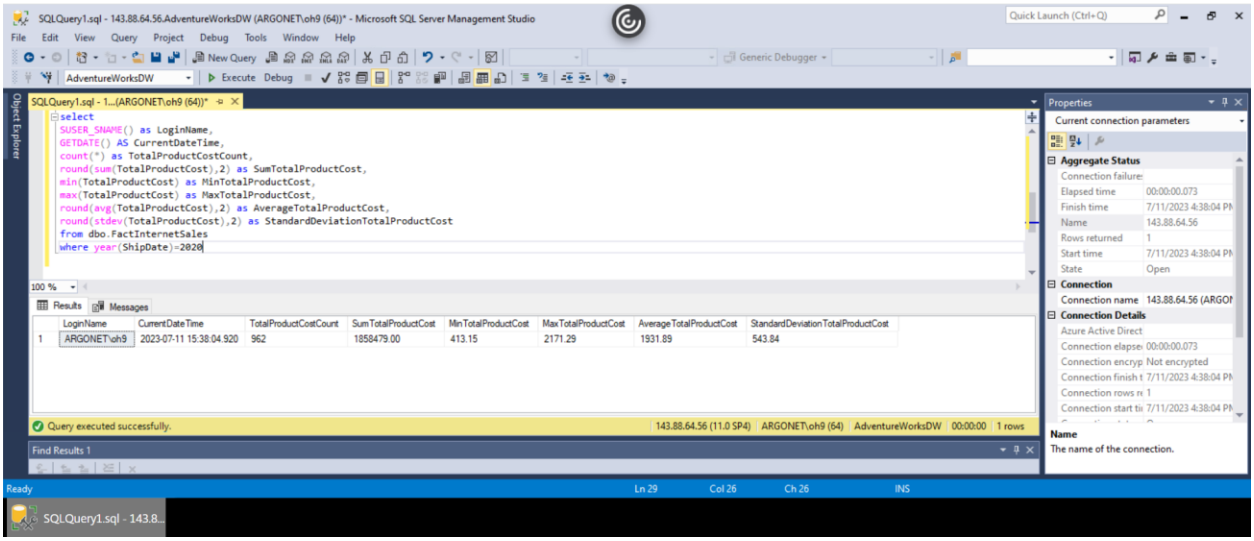
Author: Oscar Hernandez Mata

Summary: I utilized **Microsoft SQL Server Management Studio** to perform comprehensive data analysis on the dbo.FactInternetSales for the 2020 calendar year by including aggregate metrics calculation, monthly and regional sales analysis, and high-value sales day identification. Next, I executed queries for count, sum, min, max, average, and standard deviation of TotalProductCost for 2020 and 2021, using joins to integrate date and currency dimensions. Finally, I used **Power BI** for bar and pie chart visualizations to display Internet Sales by Fiscal Year for Bikes and TotalProductCost by product subcategory, ensuring data accuracy and actionable insights to support strategic decision-making.

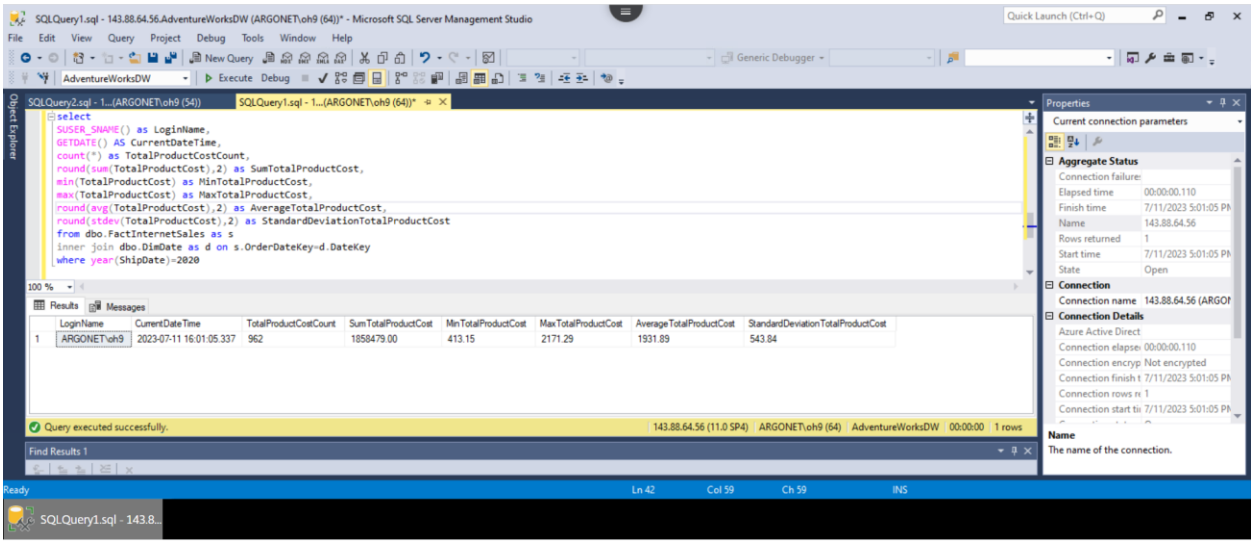
I selected the count, sum, min, max, average and standard deviation of the TotalProductCost in dbo.FactInternetSales for the 2020 calendar year using the ShipDateKey. Then I rounded the sum, average and standard deviation to 2 decimal points.



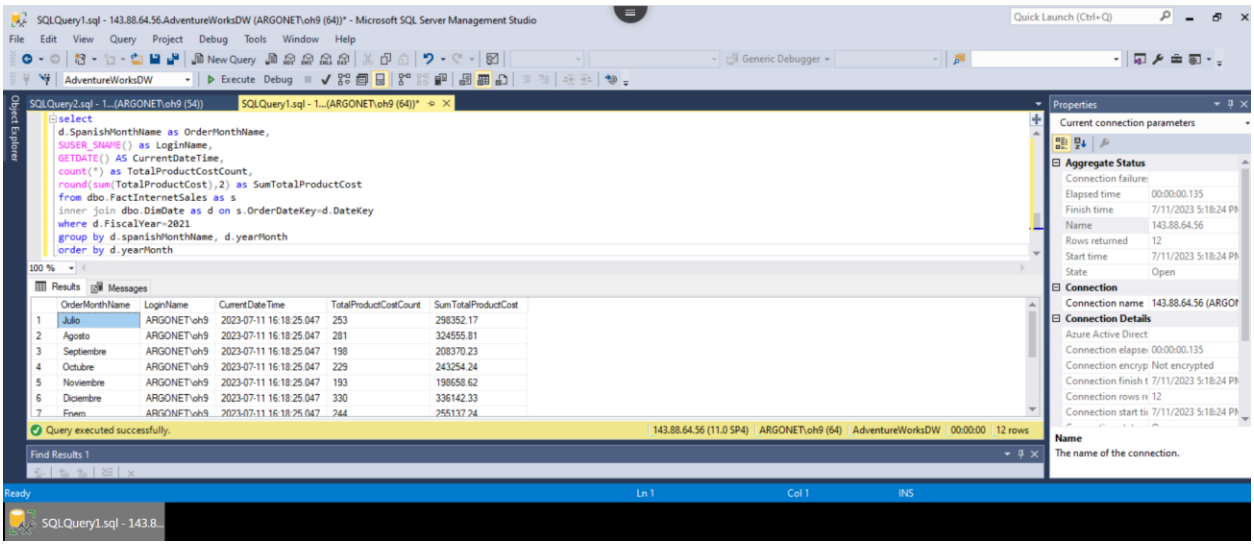
In the next step I selected the count, sum, min, max, average and standard deviation of the TotalProductCost in dbo.FactInternetSales for the 2020 calendar year using the ShipDate. Similarly, I rounded the sum, average and standard deviation to 2 decimal points.



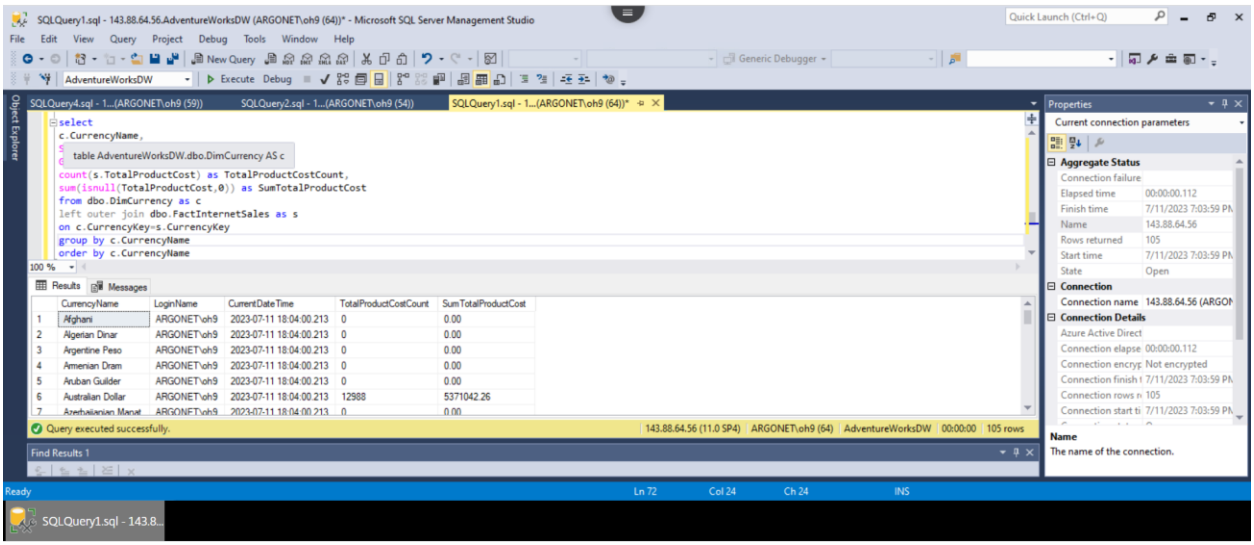
Next, I selected the count, sum, min, max, average and standard deviation of the TotalProductCost in dbo.FactInternetSales for ShipDates in the 2020 fiscal year using a join to the Date dimension. In addition, I rounded the sum, average and standard deviation to 2 decimal points.



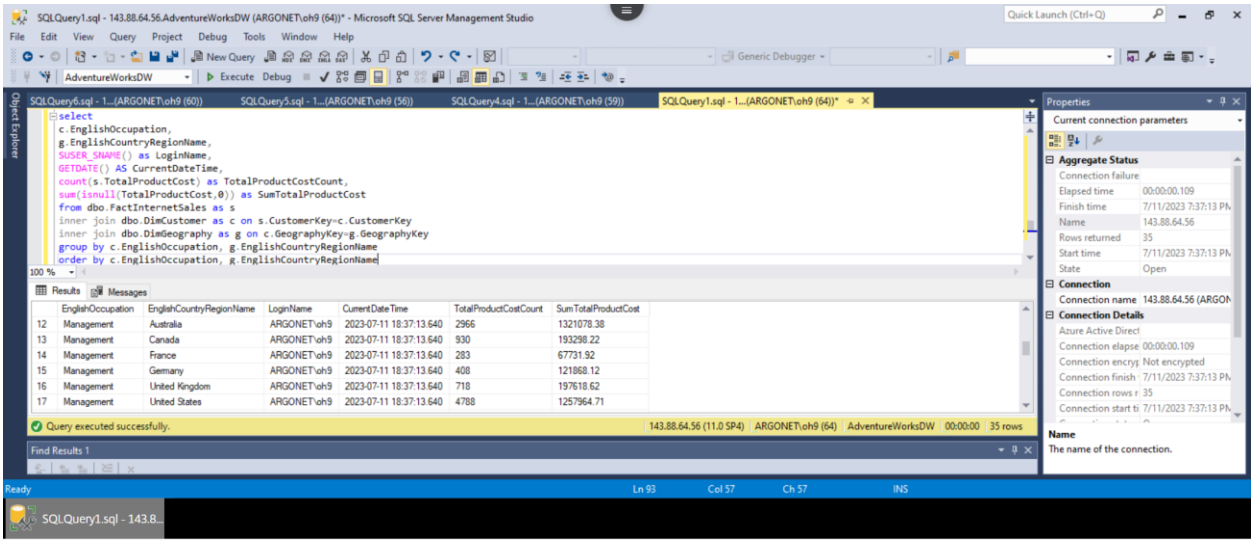
This time I selected the count, and sum of the TotalProductCost in dbo.FactInternetSales per month for the 2021 fiscal year using a join to the Date dimension grouping by the SpanishMonthName and ordering by the yearMonth. Then I rounded the sum to 2 decimal points.



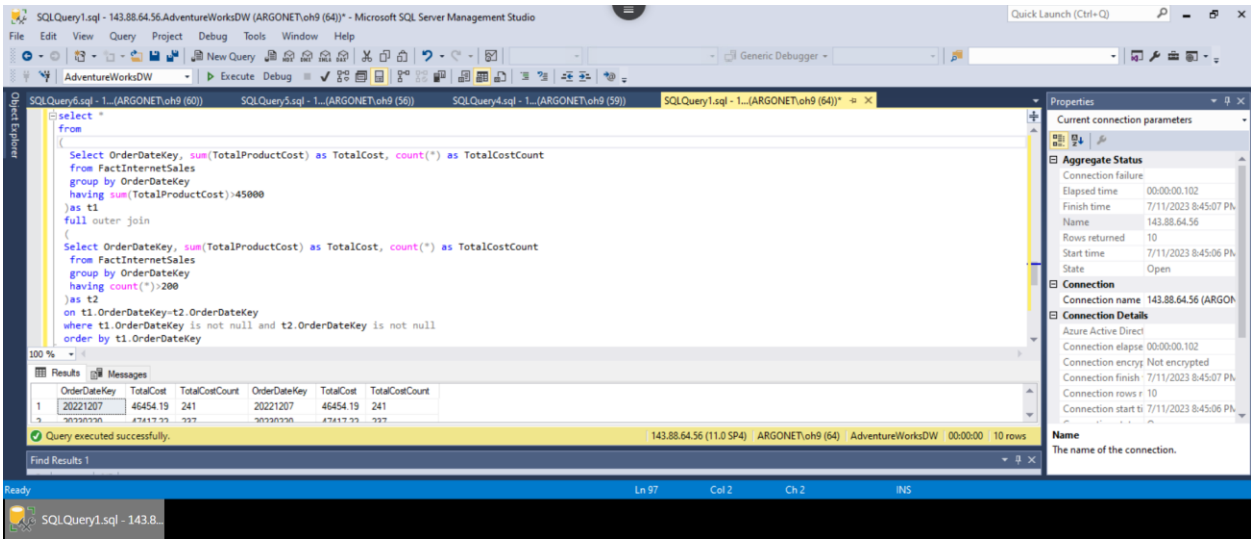
In this step I selected all of the currencies from the currency dimension table. I used a left outer join to the FactInternetSales table to get the count and sum of the TotalProductCost. Then I used the ISNULL function to prevent any NULLs in the output.



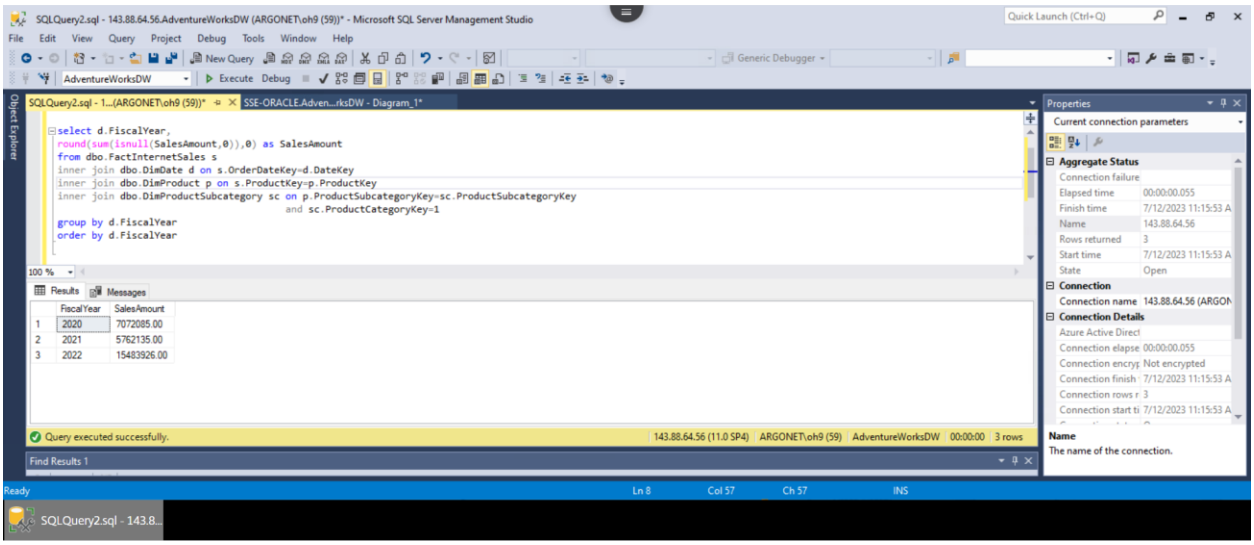
Subsequently I selected the count, and sum of the TotalProductCost in dbo.FactInternetSales per EnglishCountryRegionName and EnglishOccupation grouping by the EnglishCountryRegionName and EnglishOccupation and Ordering by the EnglishCountryRegionName and EnglishOccupation.



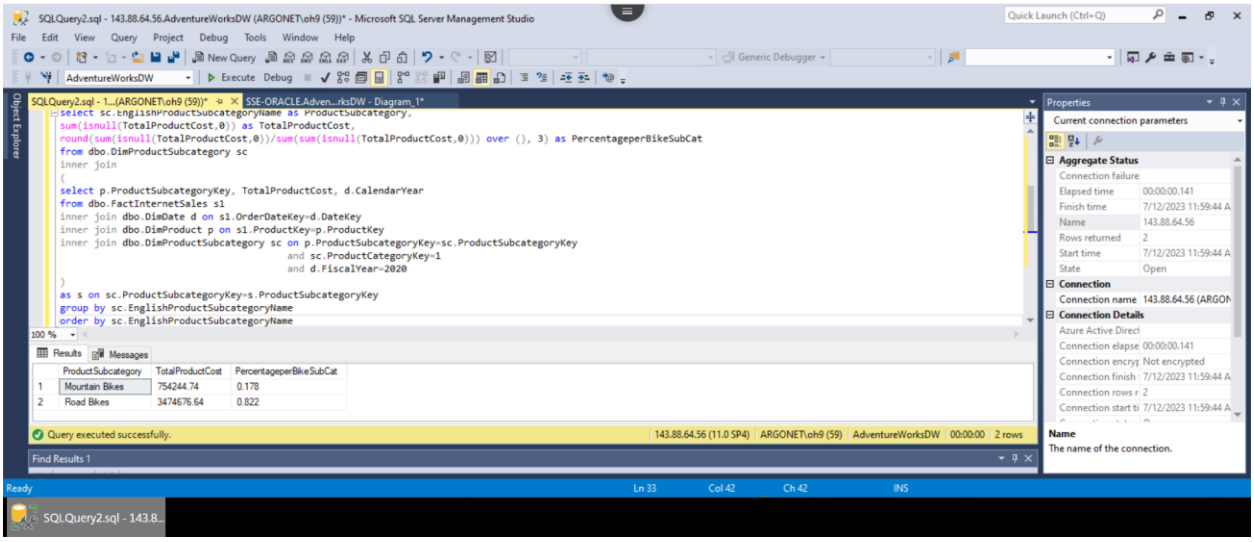
The next task was to find the days where the sum of the total product cost exceeds 45,000 and count exceeds 200. For that I used a full outer join to join 2 subqueries and the HAVING clause.

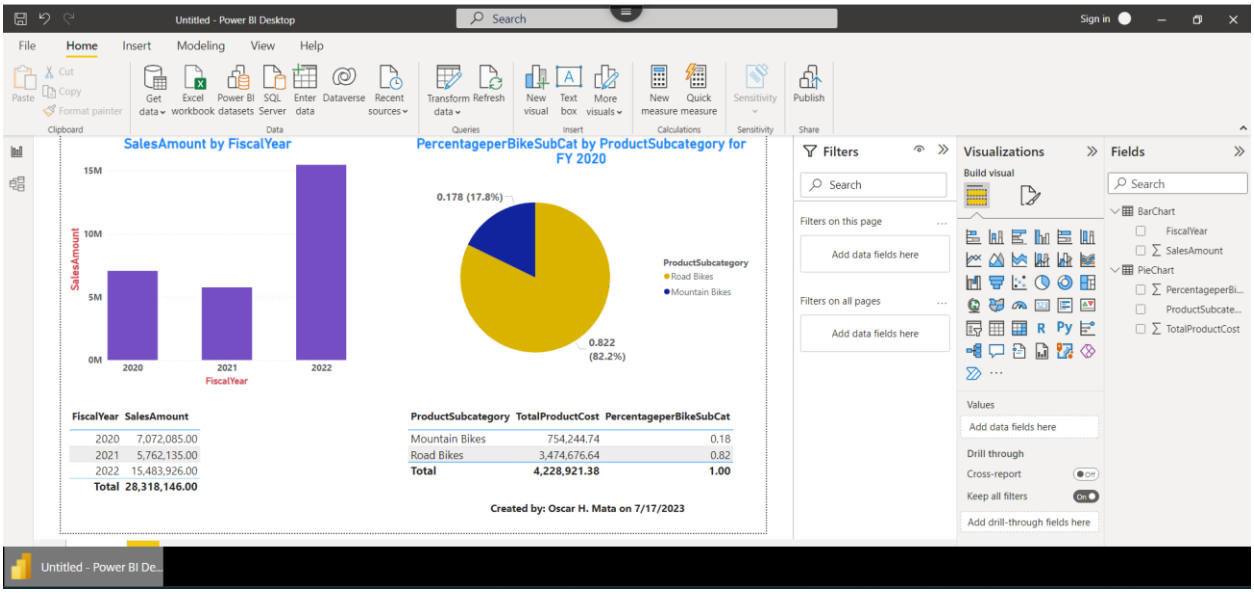


In order to visualize results I wrote a query for a **bar chart** using Internet Sales by Fiscal Year for Bikes. I displayed the fiscal year on the horizontal axis and the total product cost on the vertical axis (See graphics below).



Finally, I wrote a query for a **pie chart** using TotalProductCost and percentage per product subcategory for fiscal year 2020 (See graphics below).





Graphics conducted using Power BI (bar chart on the left and pie chart on the right).

Does the total for the subcategories for fiscal year 2020 equal the amount for fiscal year 2020 in the previous query? I concluded that the subcategories for fiscal year 2020 do match the amount for fiscal year 2020 in Sales Amount. 1341121.04+5730963.67=7072084.71, which rounded is 7072085 (See snip below).