Data 415 Final Assignment

**SSIS Integration process**

**I Created Separate Database for SSIS.**

**A screenshot of a computer

Description automatically generated**

**3. Upload to your SQL Server as CSV or Excel.**

**A screenshot of a computer

Description automatically generated**

**A screenshot of a computer

Description automatically generated**

**4. If uploaded as CSV, ensure the datatypes for the columns have been assigned correctly.**

**A screenshot of a computer

Description automatically generated**

**5. You can save the CSV as Excel and upload it to SQL Server**

**A screenshot of a computer

Description automatically generated**

**6. Choose your data file from SQL.**

**A screenshot of a computer

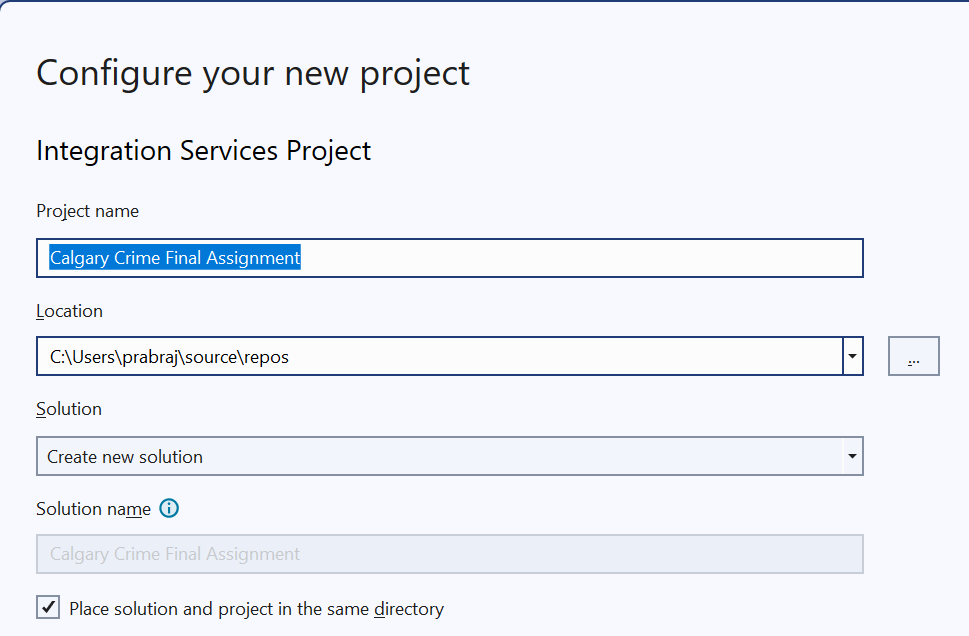
Description automatically generated**

**7. Create a new Table called Calgary\_Crime\_Disorder\_FinalAssignment with columns [Sector] [Community Name] [Category] [Crime Count] [Resident Count] and [Location]. Ensure the datatypes are correct.**

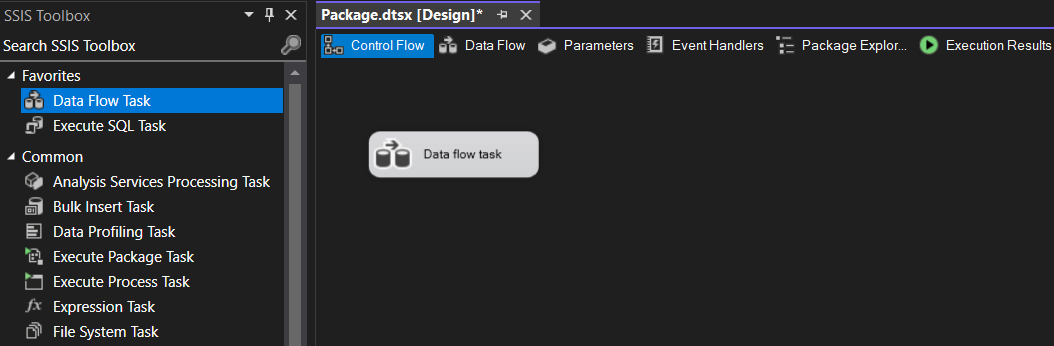
**A screenshot of a computer

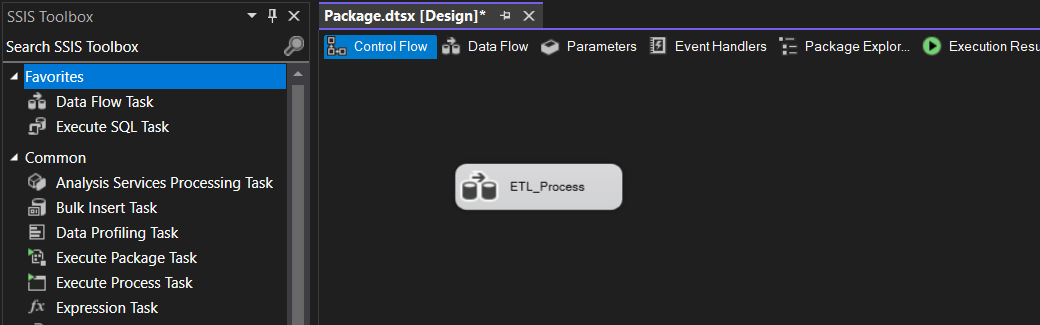
Description automatically generated**

**8. Create an Integrated Project**

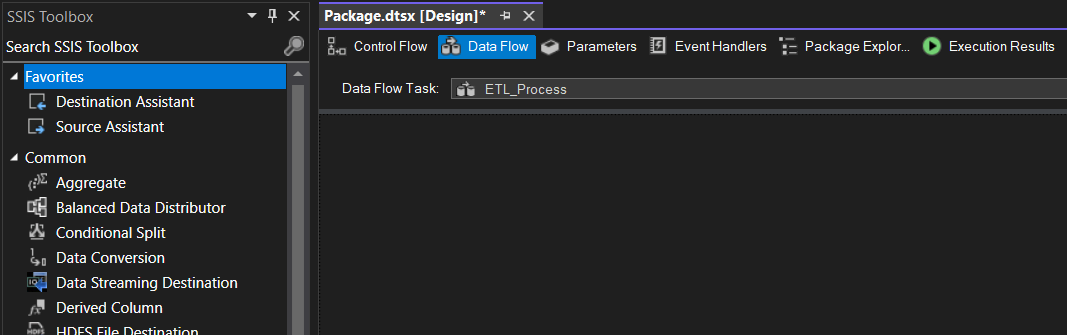
****

**Drag the data flow task and rename it, <ELT\_Process>**

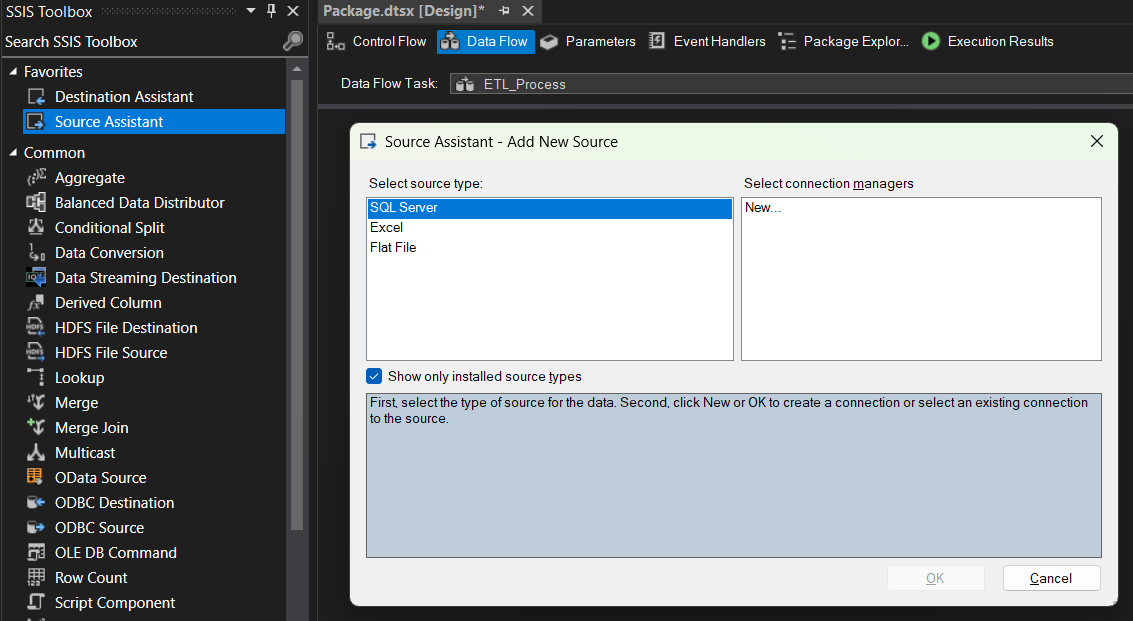
****

****

**Double-click on task ETL\_Process … And open dataflow task.**

****

**Drag Source Assistant to the work area.**

****

**Source Assistant Connection Box:**

**A screenshot of a computer

Description automatically generated**

**Selecting the Main Table.**

**A screenshot of a computer

Description automatically generated**

**Testing the connection,**

**A computer screen shot of a connection manager

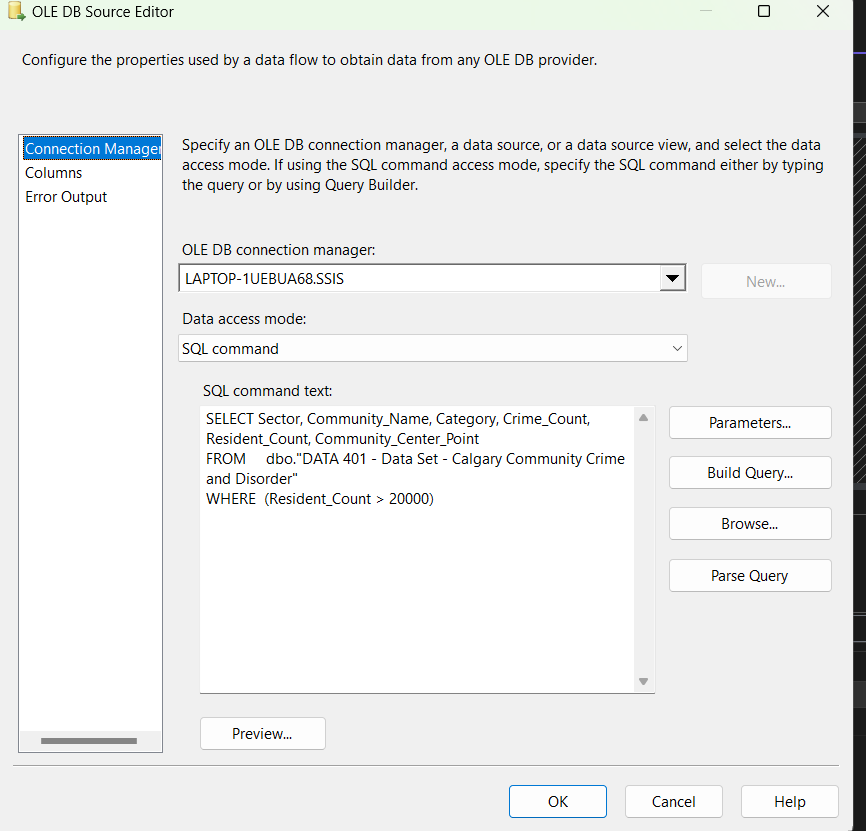
Description automatically generated**

**Previewing the table for the source Assistant.**

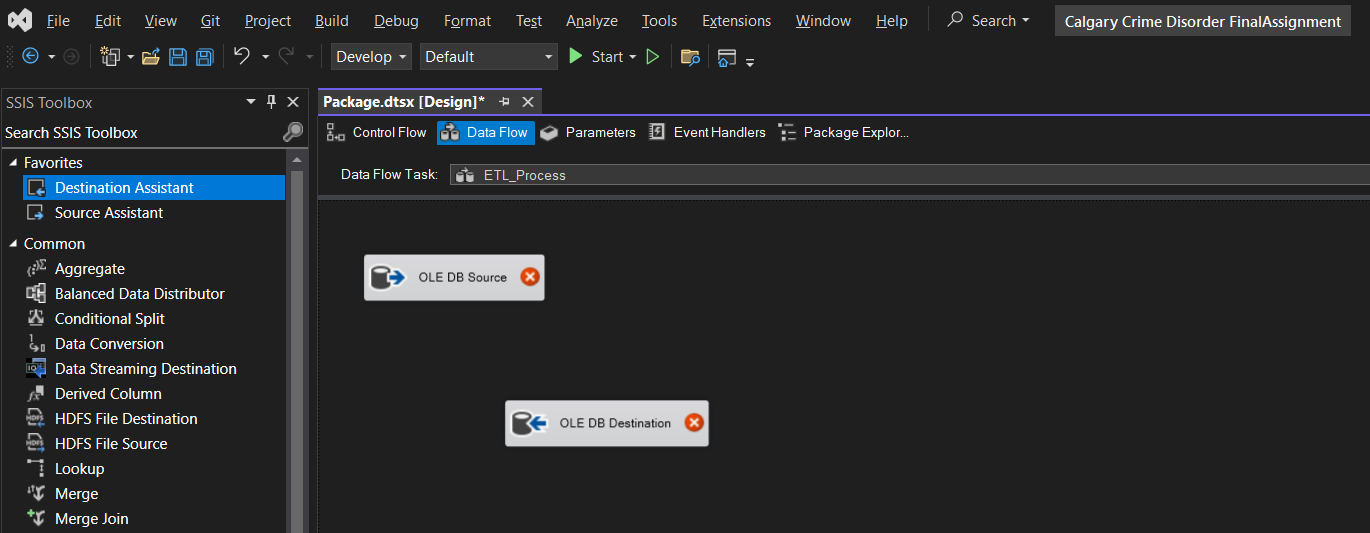
**A screenshot of a computer

Description automatically generated**

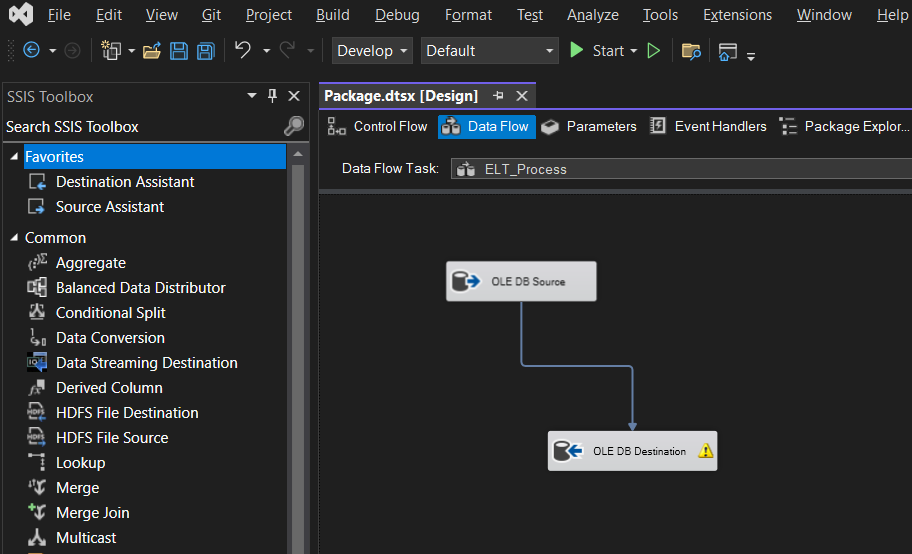
**16. Click on data access mode, select SQL command, and write the following SQL query.  
SELECT [Sector], [Community Name], [Category], [Crime Count], [Resident Count], [Community Centre Point] FROM [dbo]. [YYC\_Calgary\_Crime\_and\_Disorder\_Exercise] WHERE [Resident Count] > 20000**

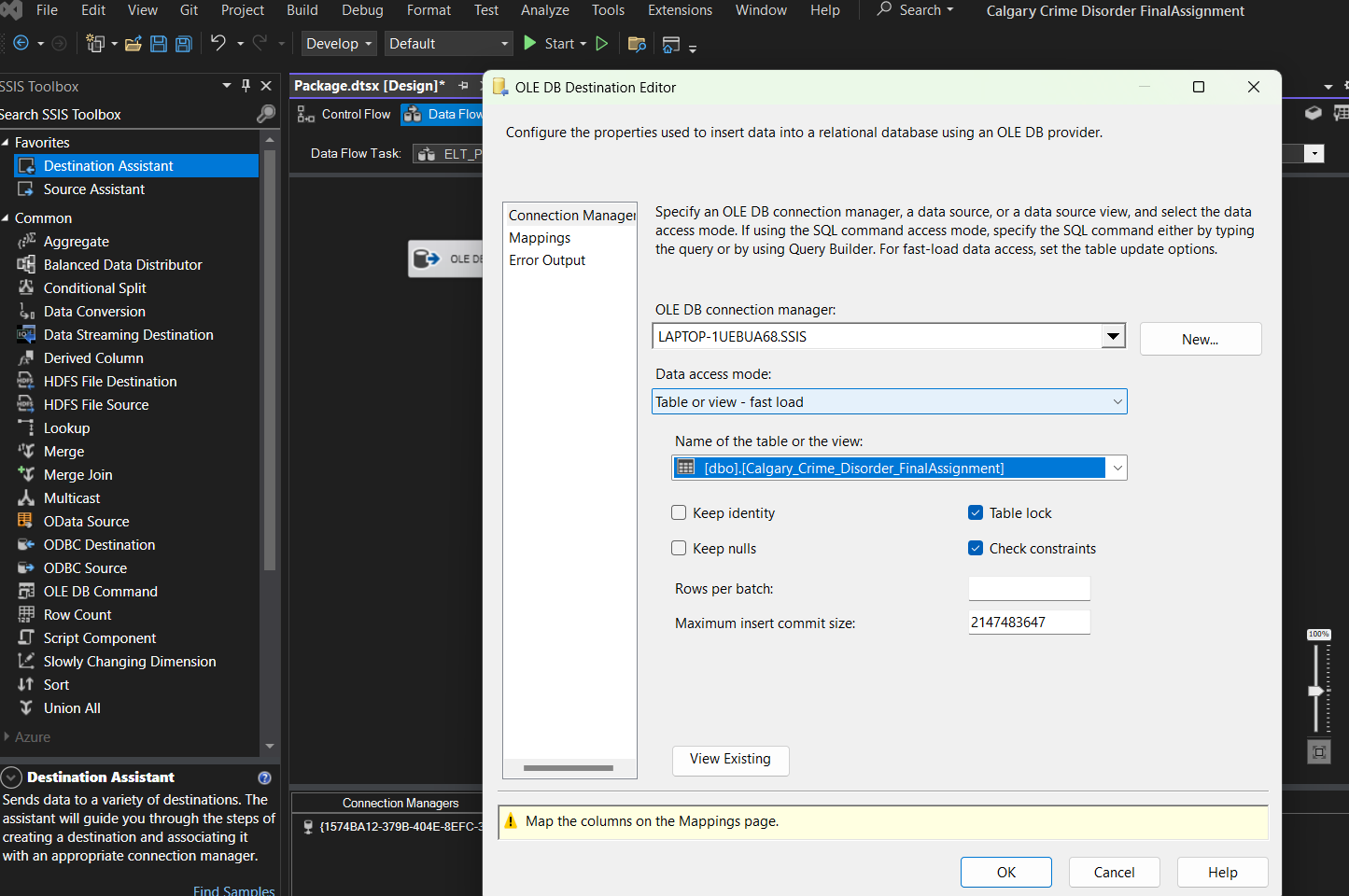
****

**17. Drag Destination Assistant to the work area.**

****

**Connecting Source to Destination.**

****

**18. Double click on OLE DB Destination and select the destination table,  
Calgary\_Crime\_Disorder\_FinalAssignment.  
  
**

**19. Click on Mapping and map the input column to the output column.**

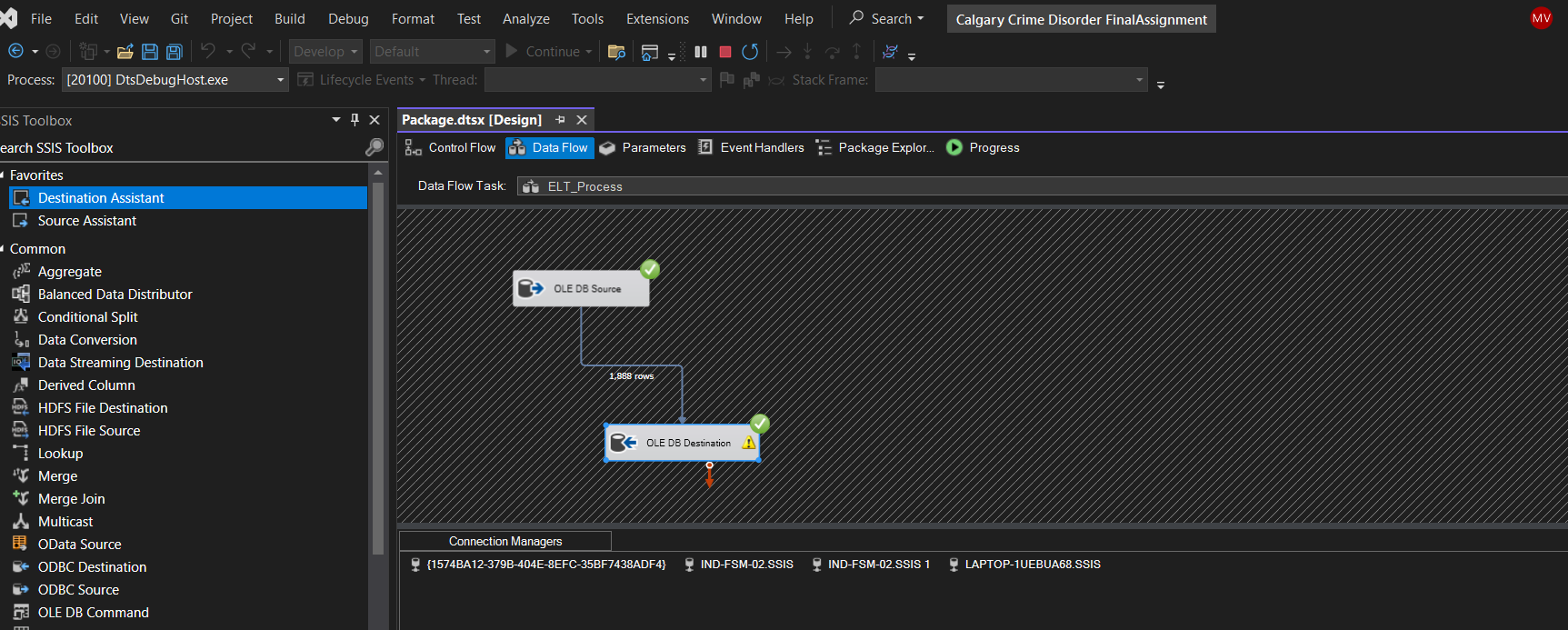
**20. Connect the Columns, Community Centre Point to the Location.**

**A screenshot of a computer

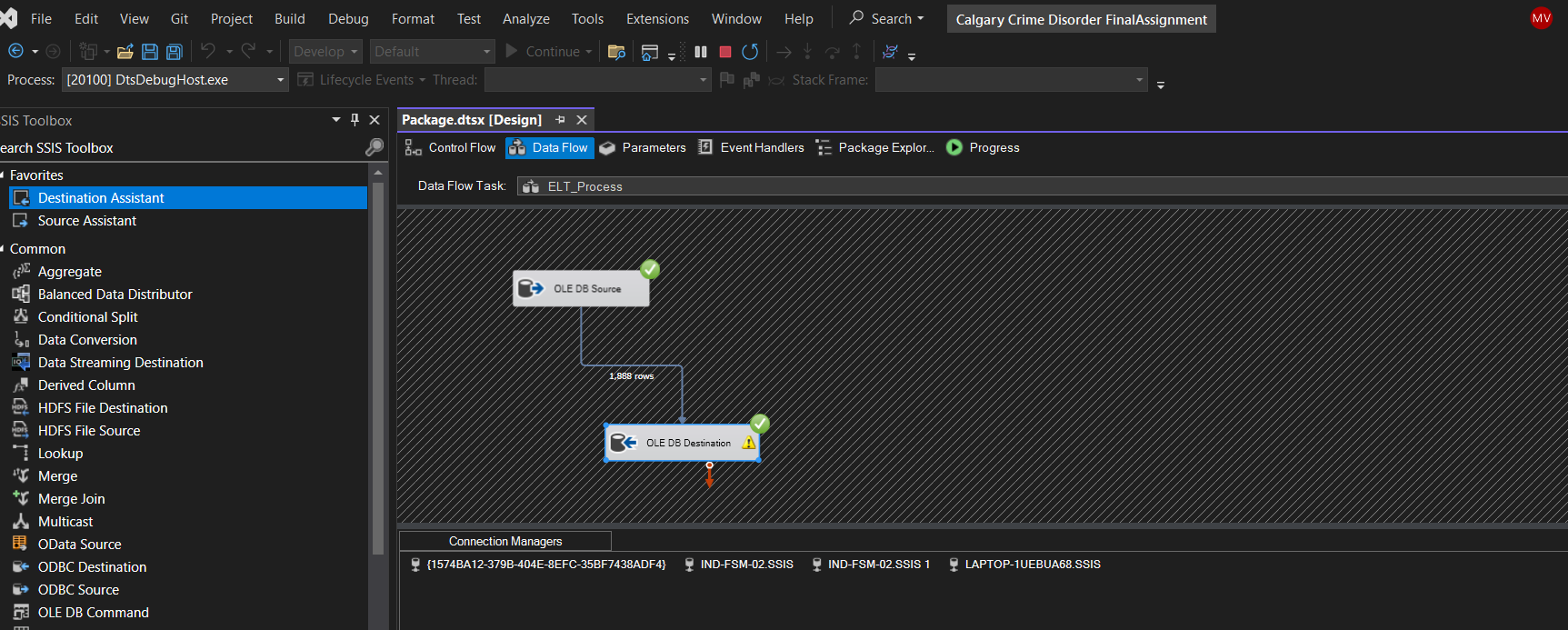
Description automatically generated**

**The X mark goes off in the destination folder,**

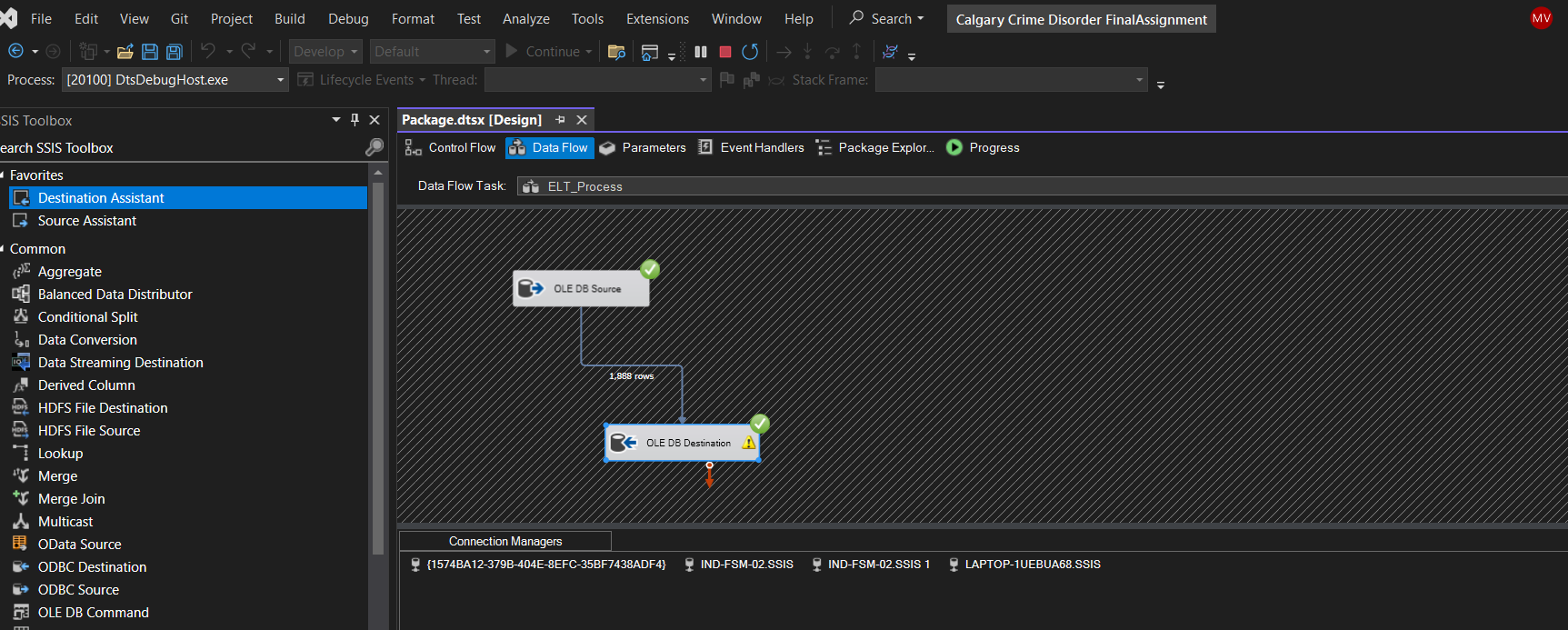
**21. Click on Start and execute the project**

****

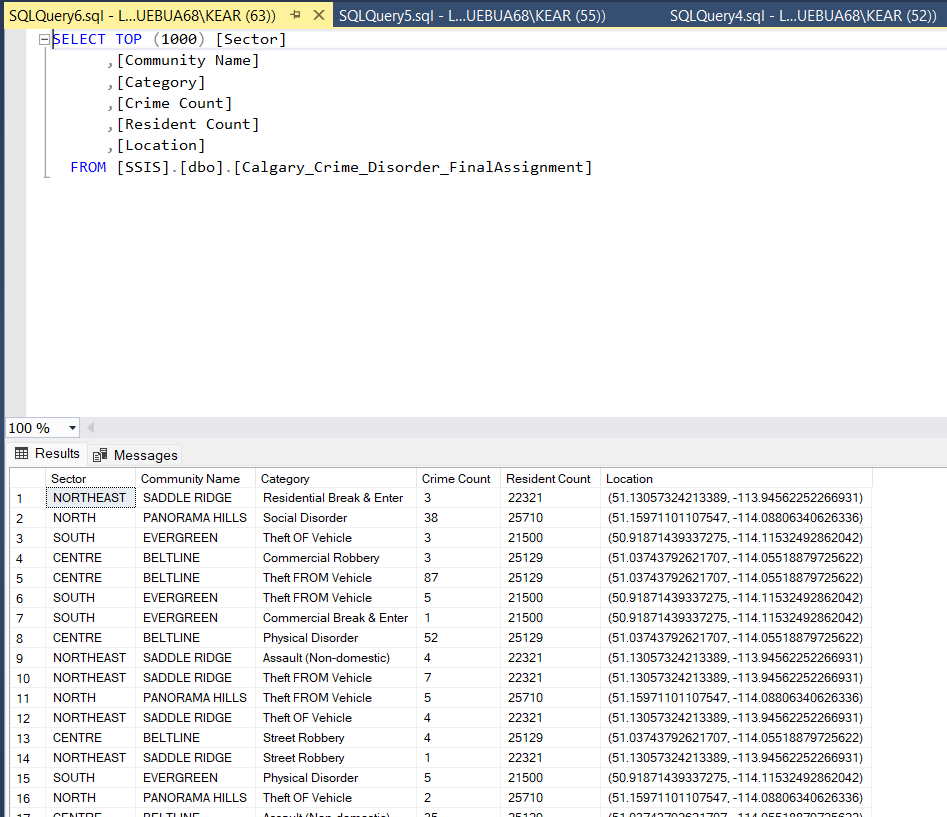
**22. You should get the green check mark for executing the project successfully.**

****

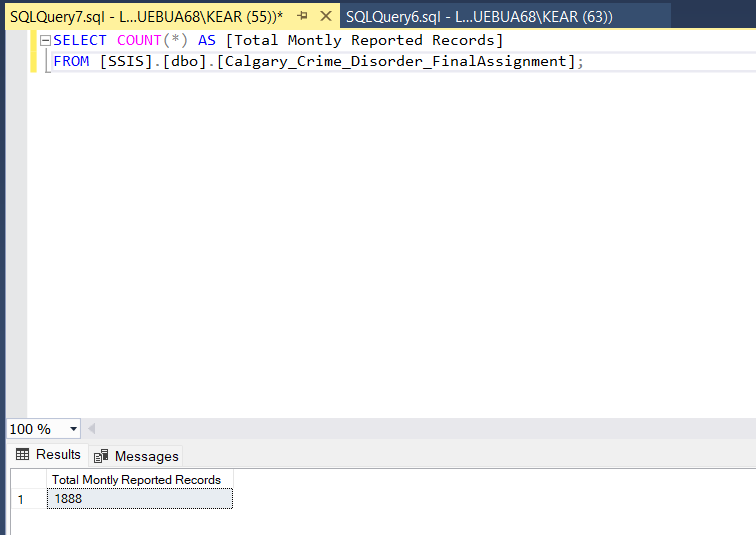
**23. Package execution completed with success.**

****

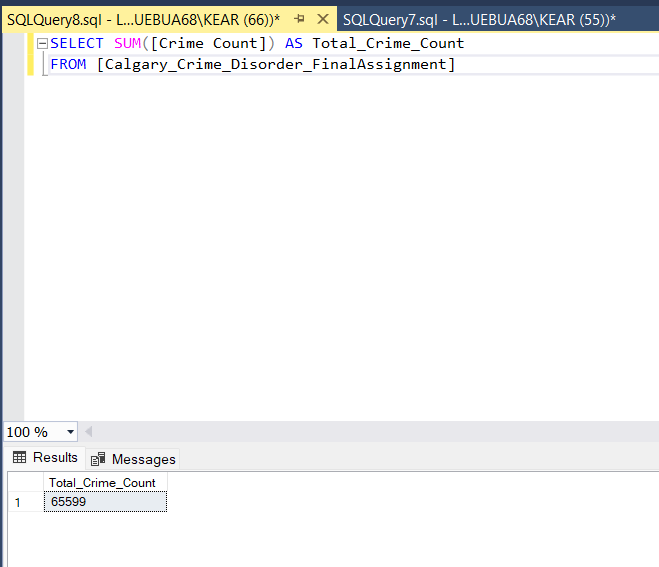
**24. Now you can see the queried data should be in your SQL server in the new Calgary\_Crime\_Disorder\_FinalAssignment Table.**

****

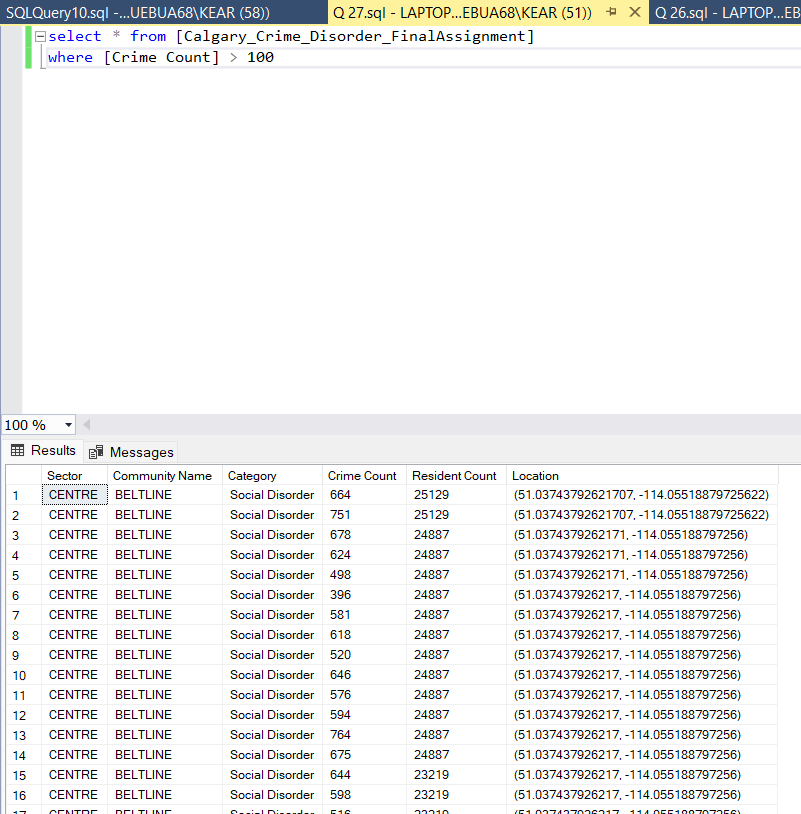
**25. What was the total number of monthly reported records for these communities with resident counts greater than 20,000?**

****

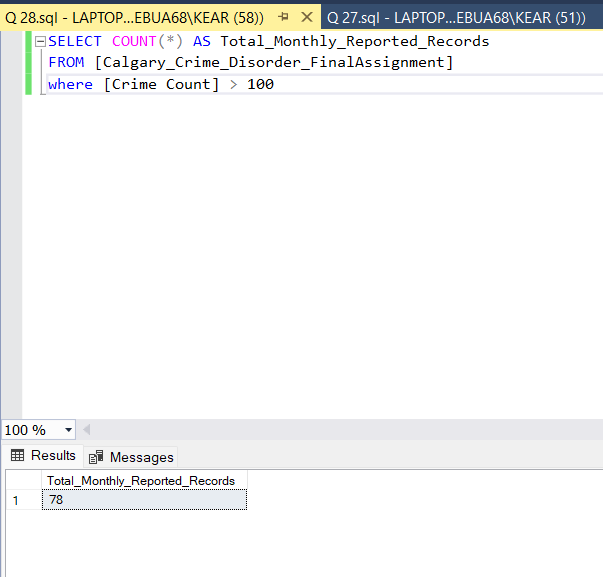
**26. What is the total crime count from all reported monthly records in this query from Section 25?**

****

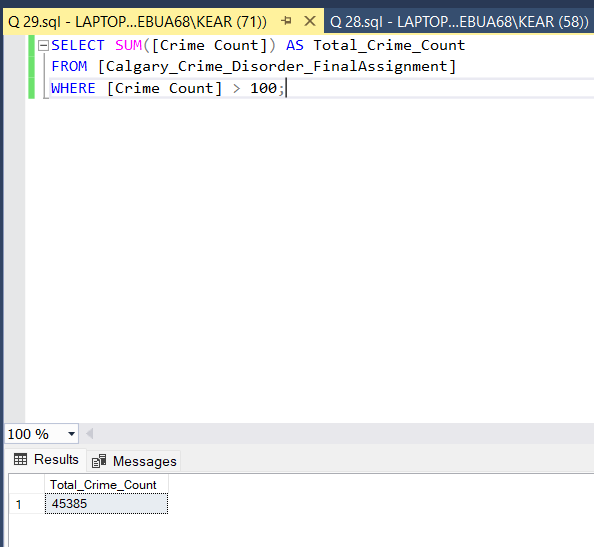
**27. Build a Query from the table in Section 24 where the Crime count is greater than 100.**

****

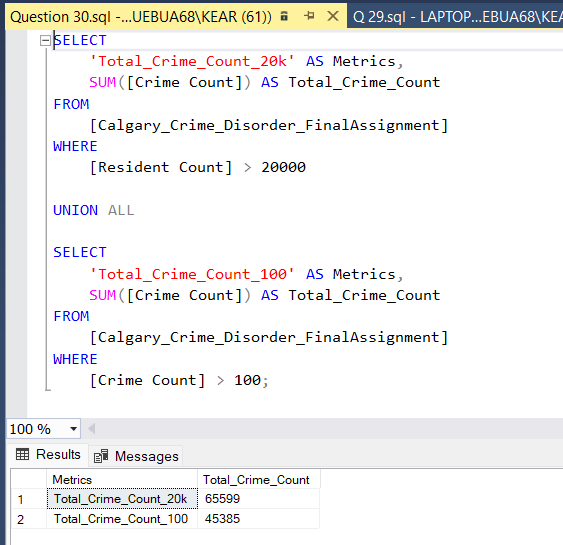
**28. What is the total number of monthly reported records for the crime count greater than 100 from Question 27?**

****

**29. What is the total crime count for this query in question 28?**

****

**30. How does the total crime count in question 29 for communities with a resident count greater than 20,000 compared to the total crime count for a query with the crime count greater than 100 from the whole Calgary dataset?**

****

**Calculate the Percentage Difference:**

**Comparison of Total Crime Counts**

1. Total Crime Count for Communities with a Resident Count Greater Than 20,000: **65,599**2. Total Crime Count for Records Where the Crime Count is Greater Than 100: **45,385**

**Percentage Difference Calculation**

**Percentage Difference= {[(65,599−45,385​)] / [45,385]} ×100 ≈ 44.3%**

The total crime counts for communities with a resident count greater than 20,000 is approximately 44.3% higher compared to the total crime count for records where the crime count is greater than 100. The total crime count for these communities is about 1.44 times (or 44%) higher than the total crime count for the other dataset.