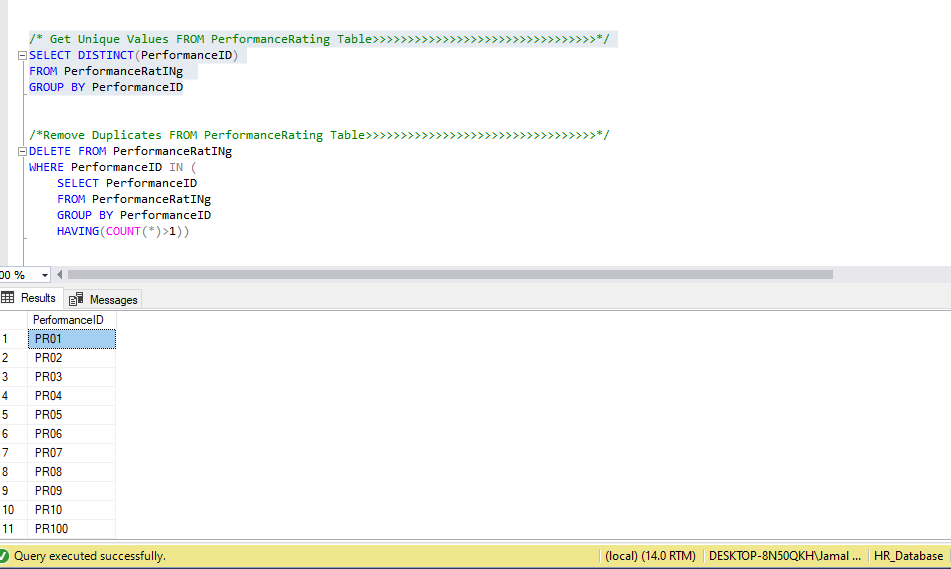
Data Preprocessing Using SQL

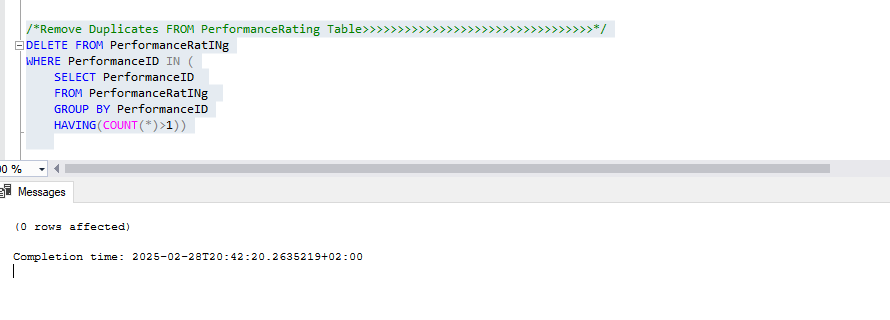
Codes & Results

Performance Rating table

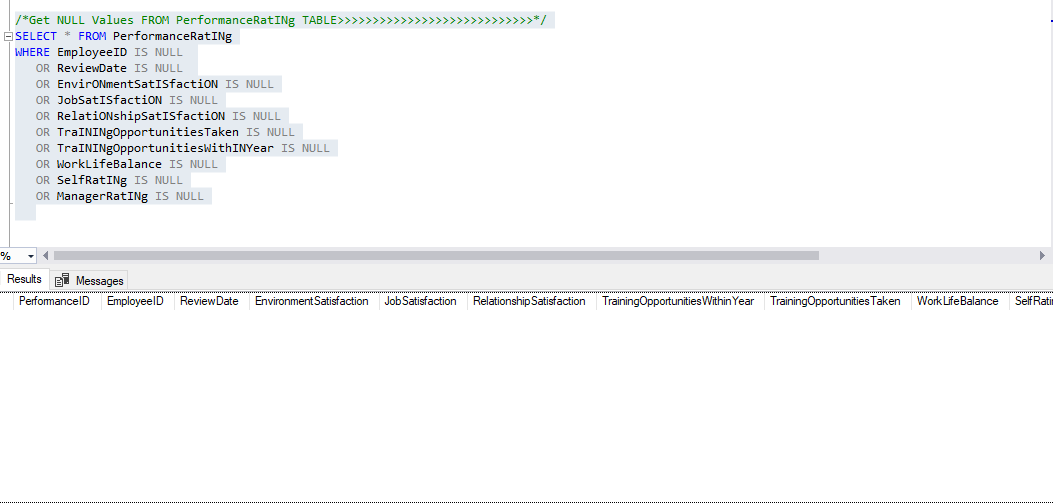
\*\*Get unique values from Performance Rating table:



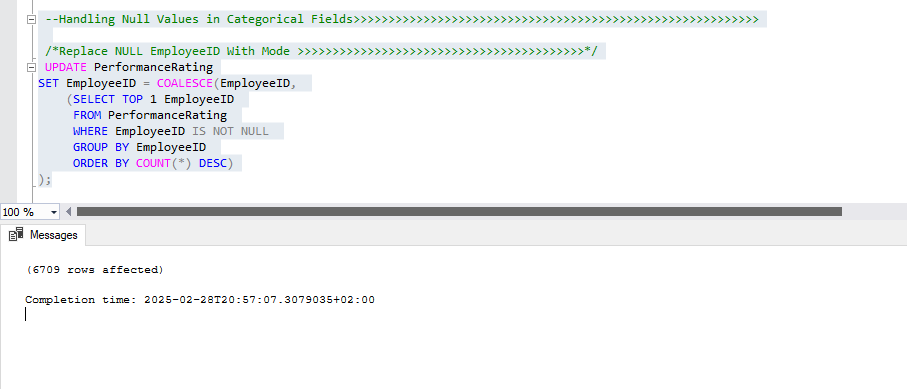
\*\*Remove duplicates from Performance Rating table:

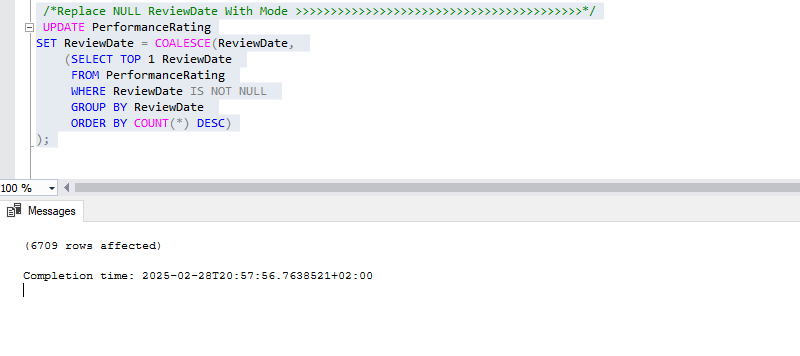


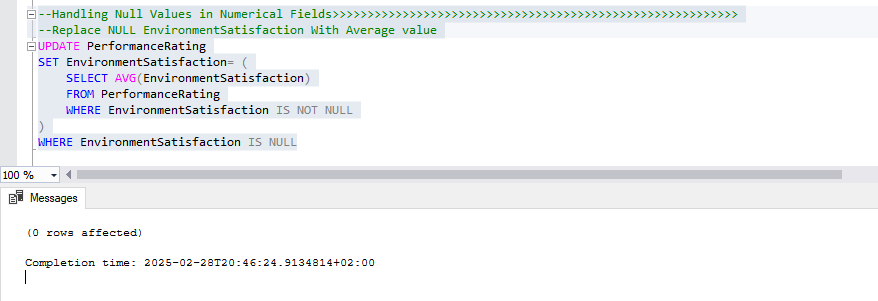
\*\*Get null values from Performance Rating table:

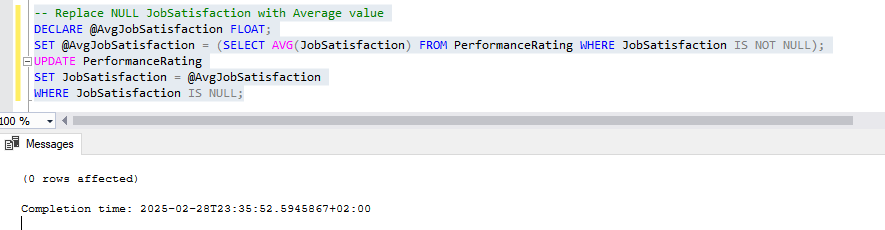
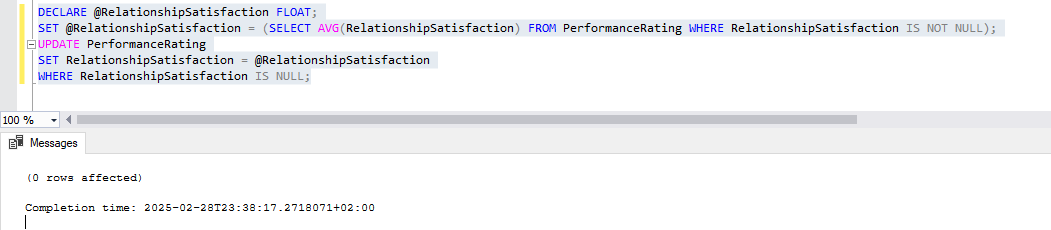
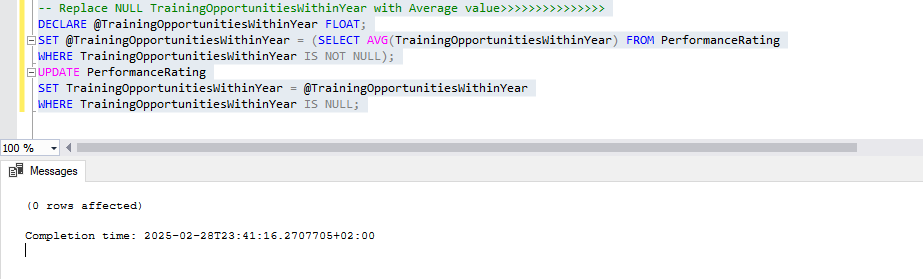
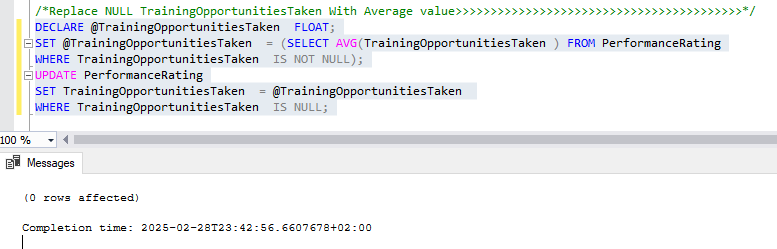
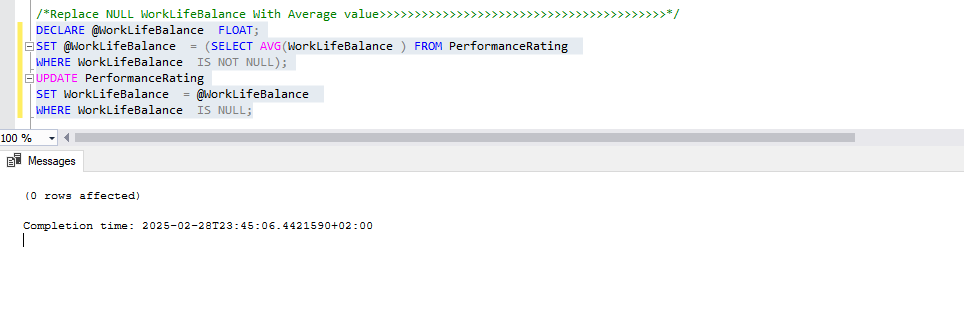
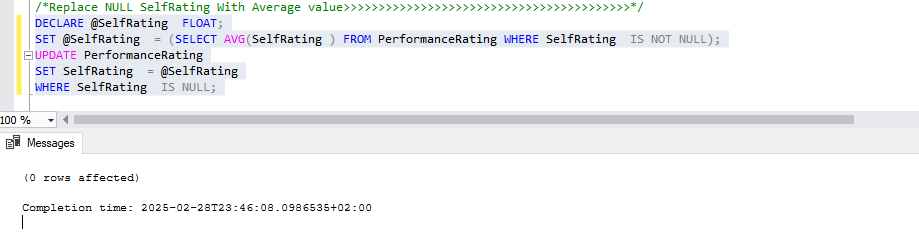
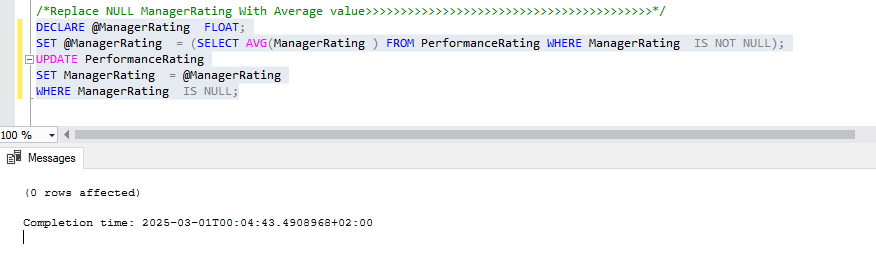


\*\*Handling null values in Performance Rating table:

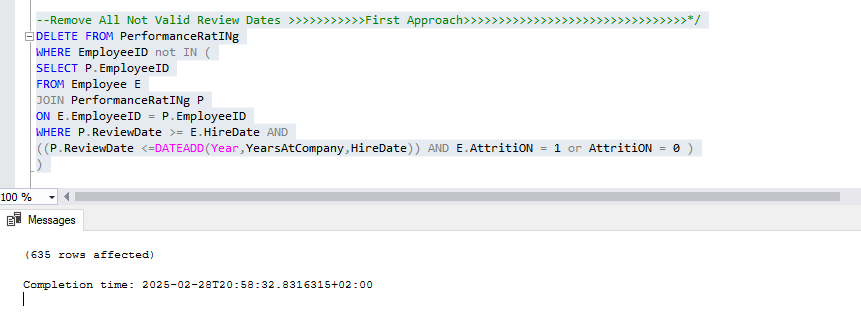


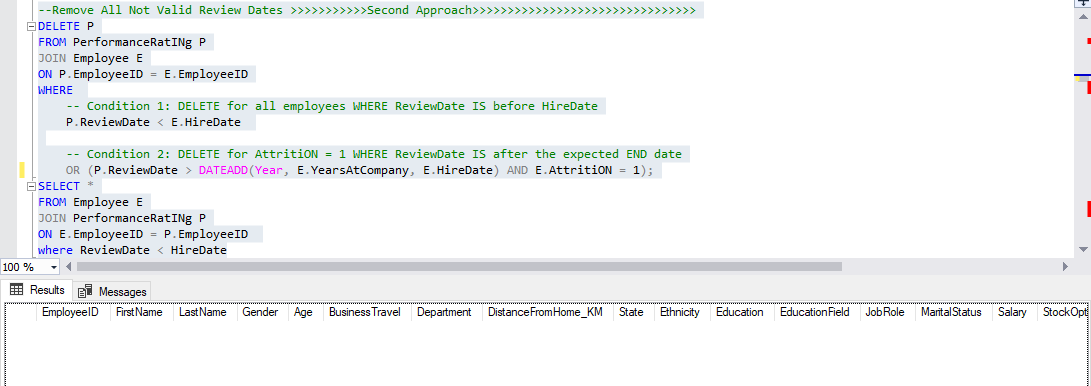






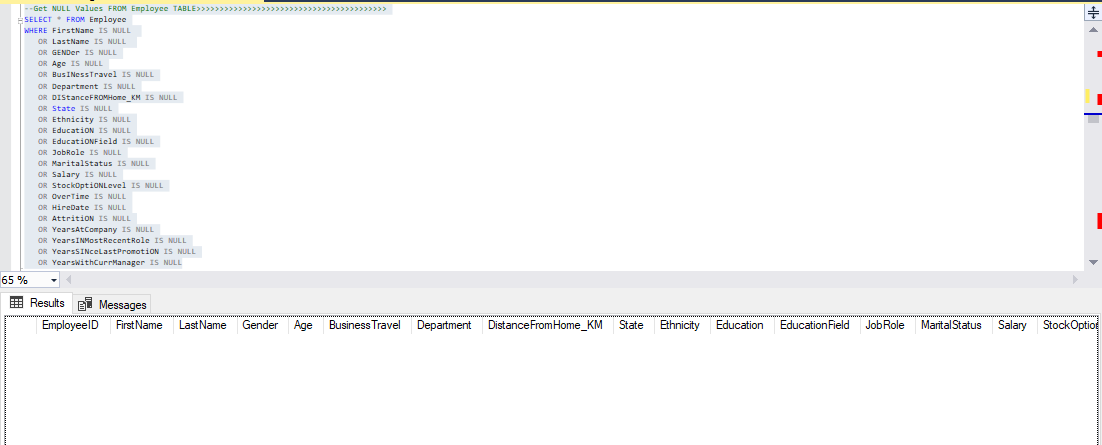
\*\*Remove invalid Review Dates:



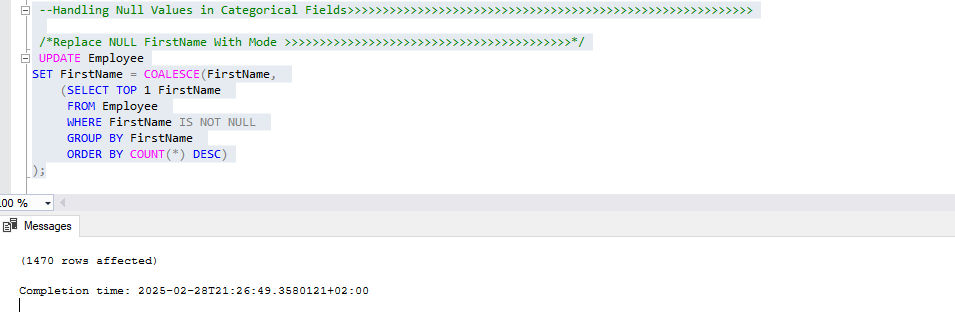


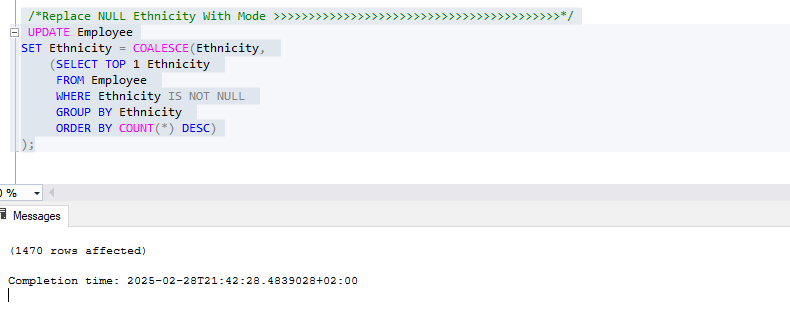
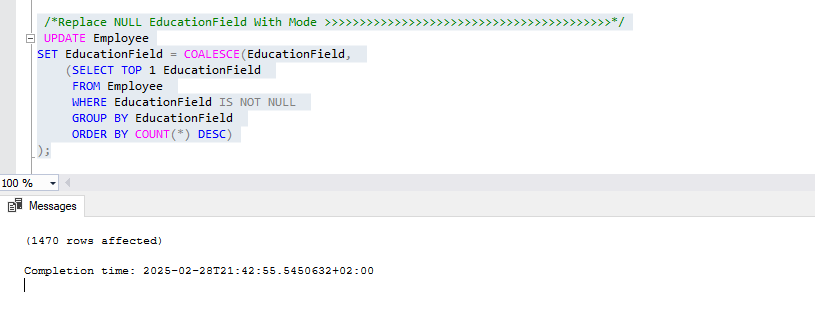
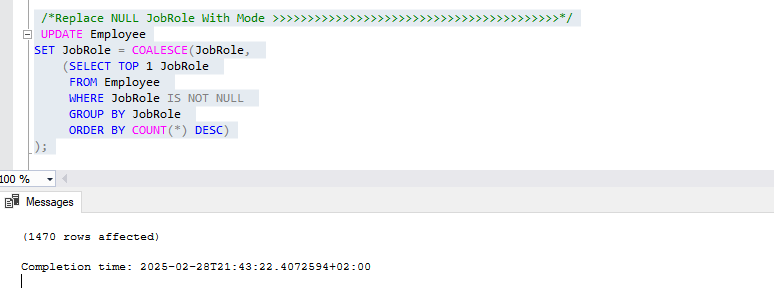
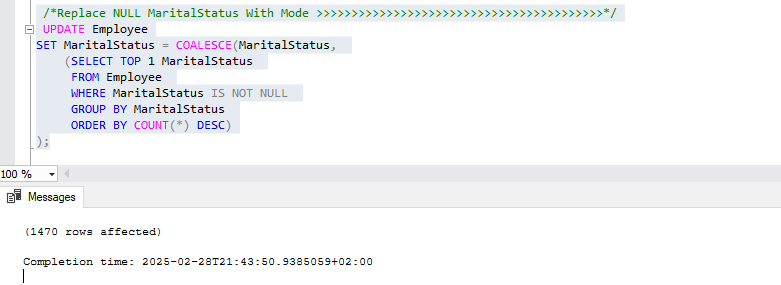
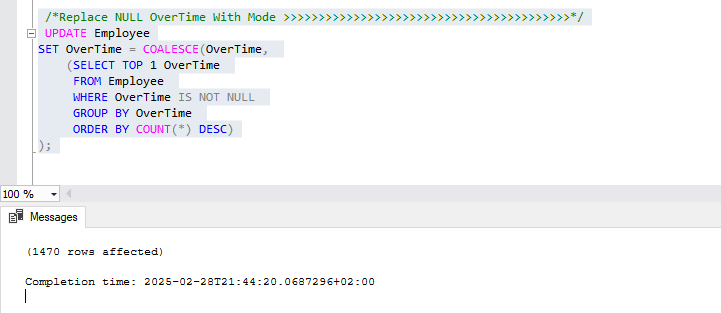
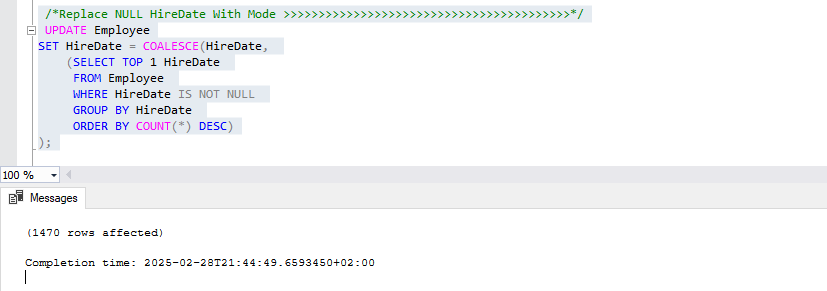
Employee table

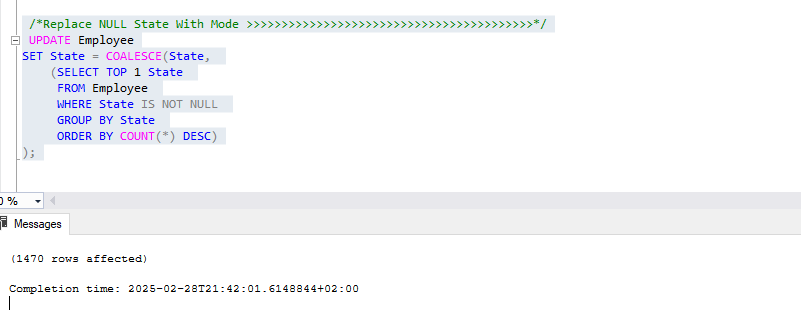
\*\*Get null values from Employee table:

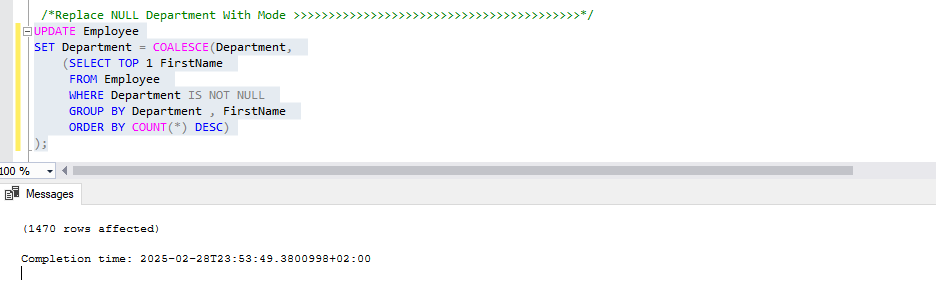


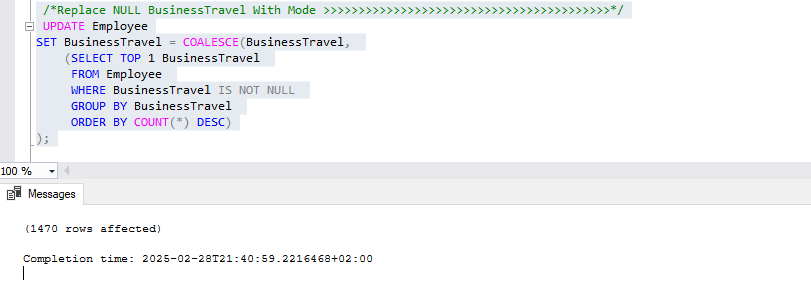
\*\*Replace null values in categorical fields with mode:

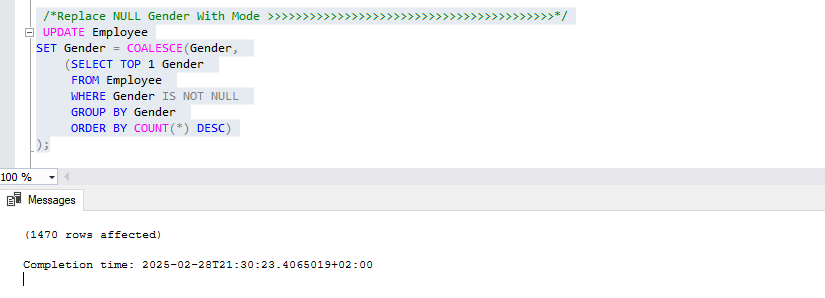


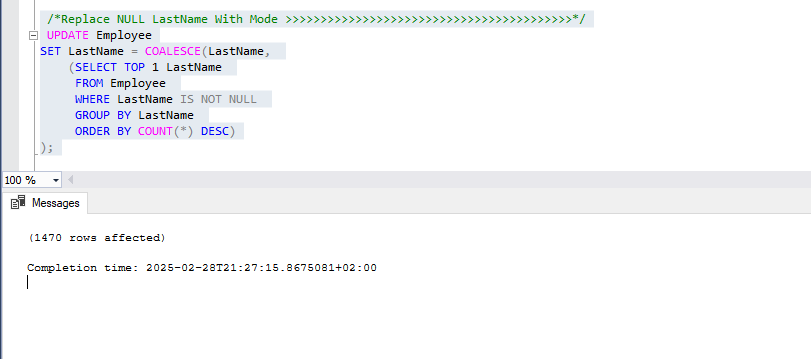




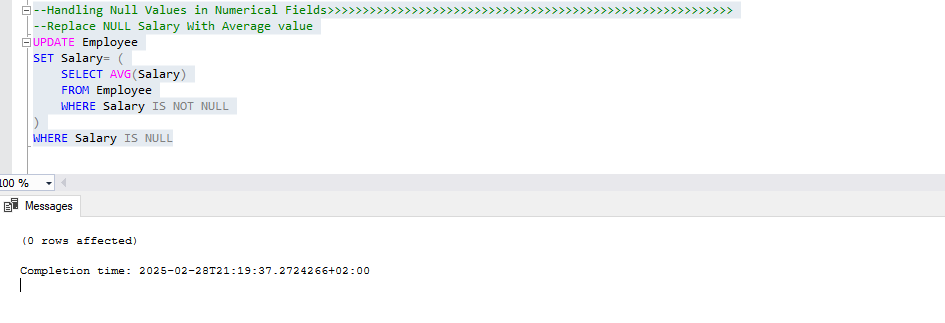


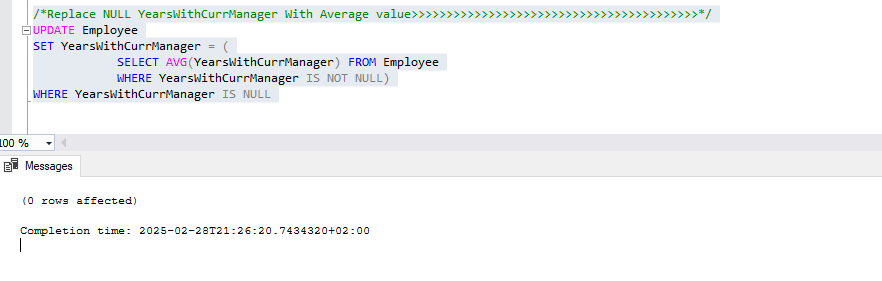


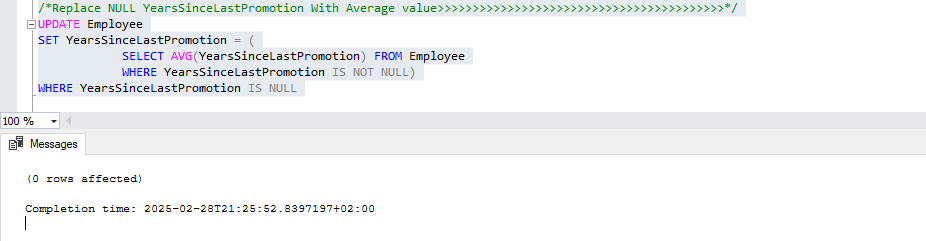


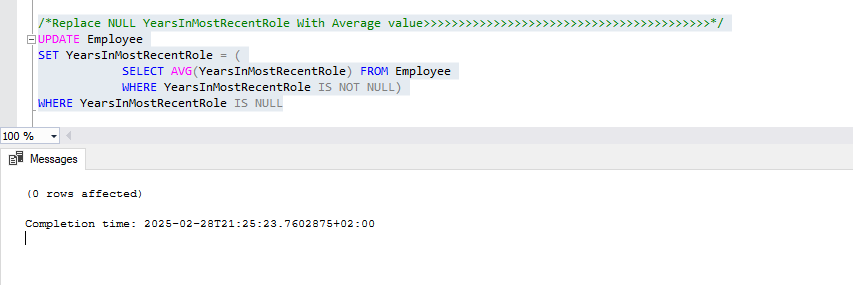


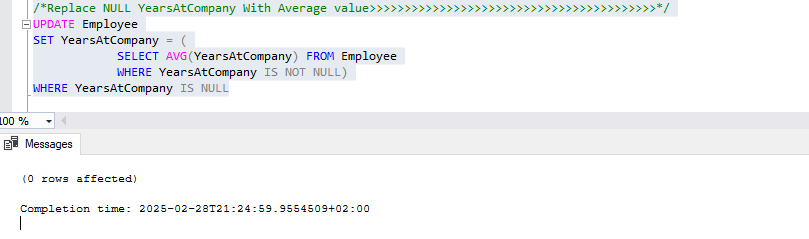
\*\*Replace null values in numerical fields with Average:

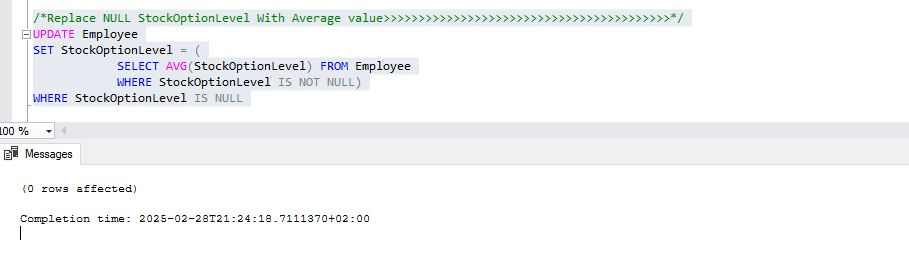


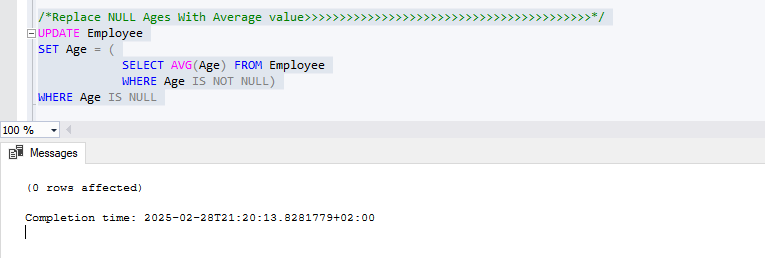
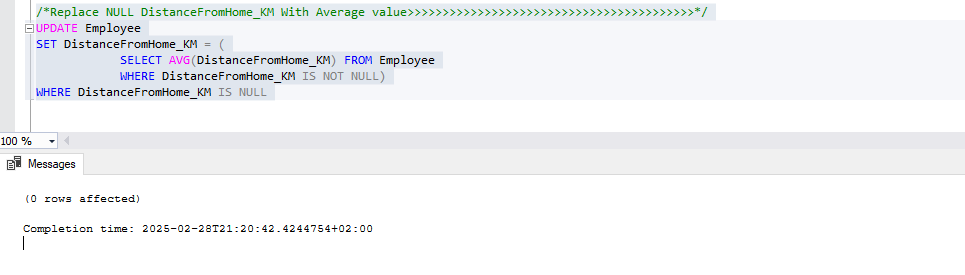
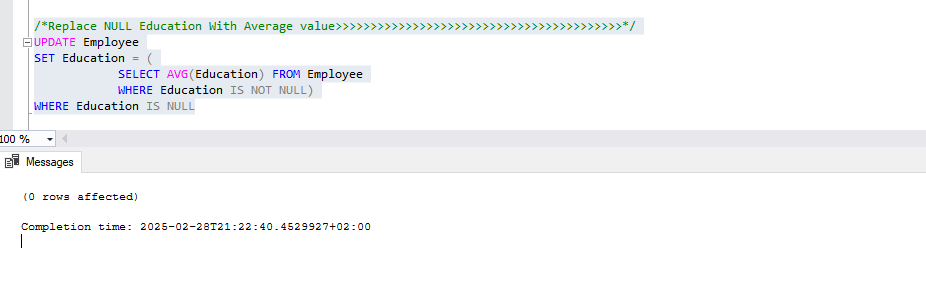
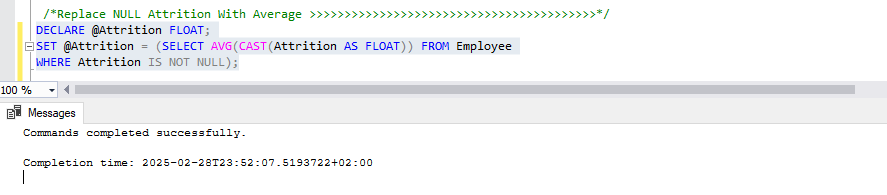




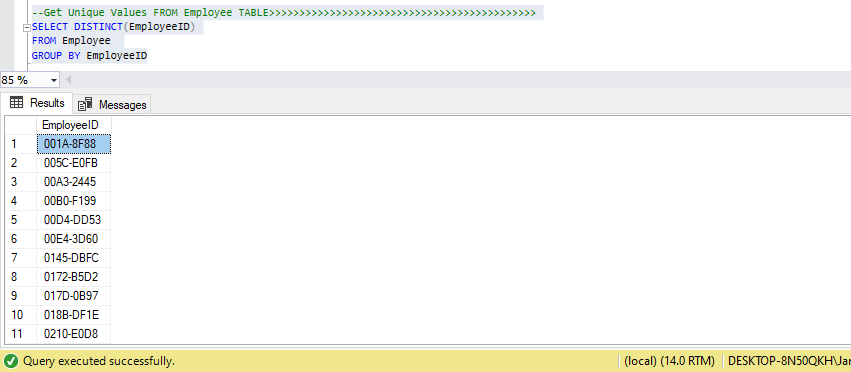


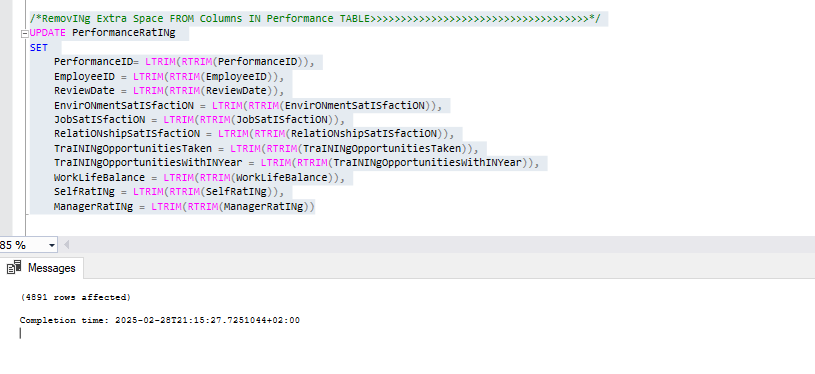




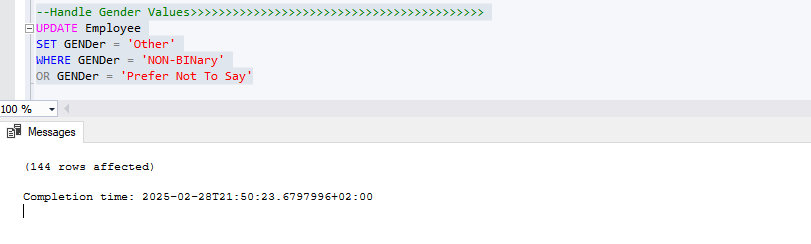


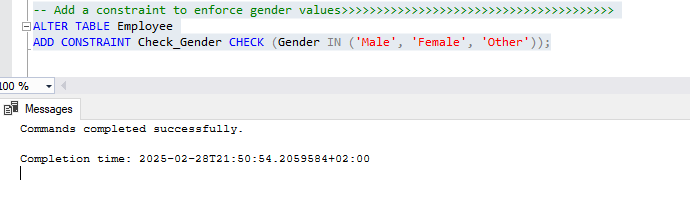
\*\*Get unique values from Employee table:



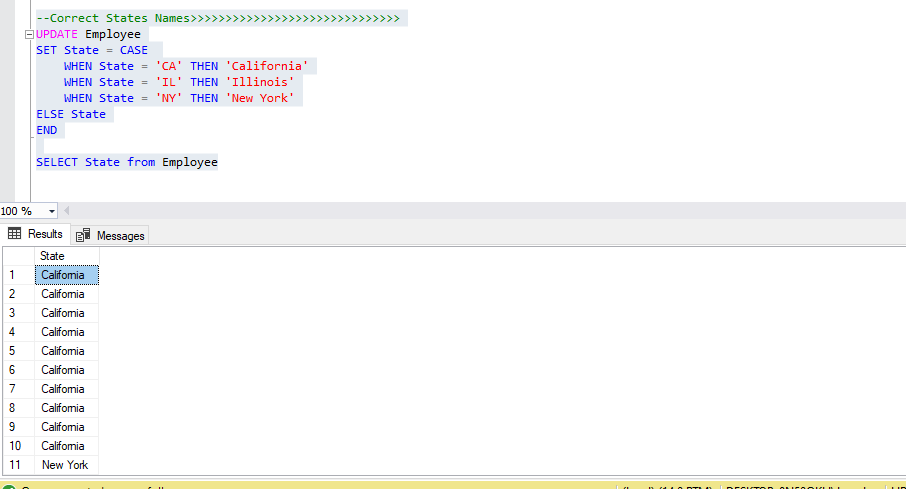


\*\*Modifying Gender column in Employee table:

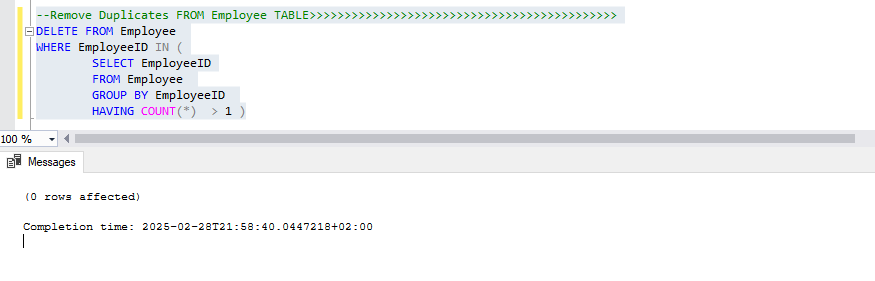




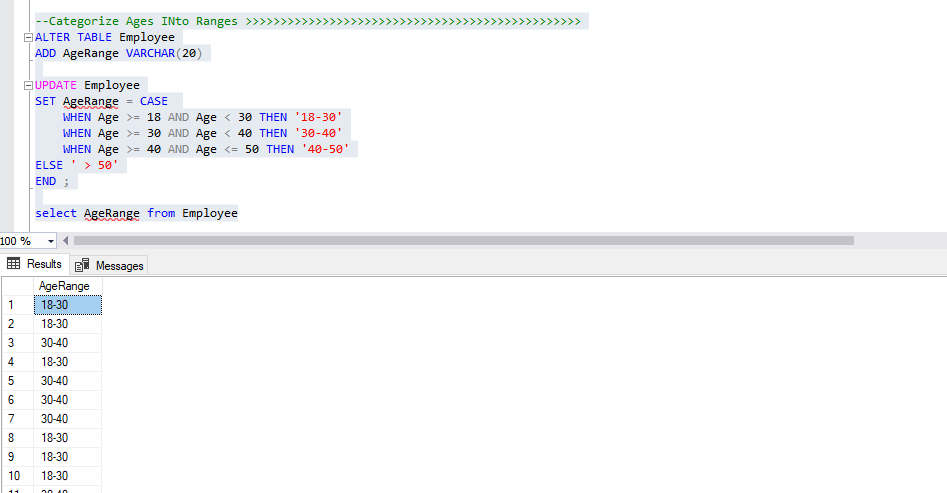
\*\*Correct states names:

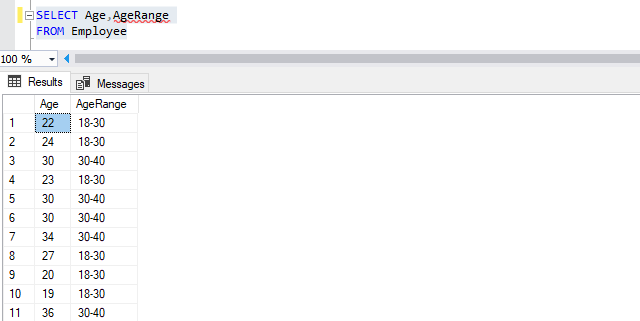


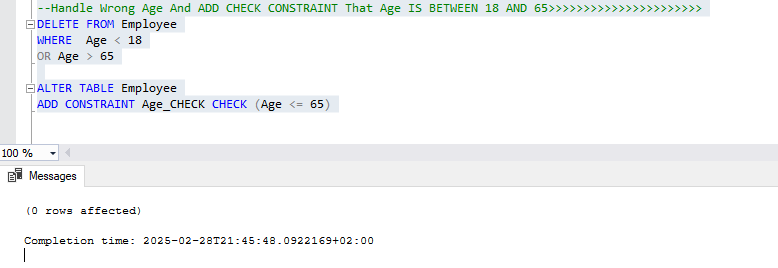
\*\*Remove duplicates:



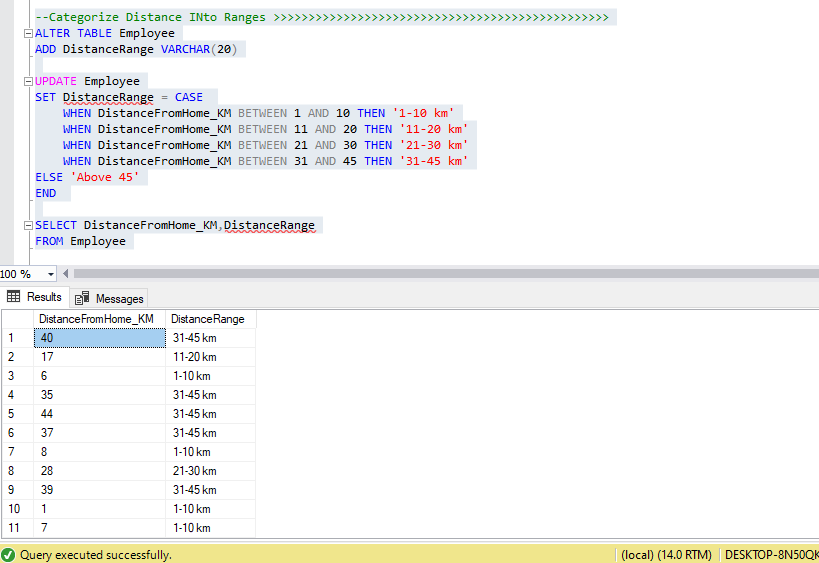
\*\*Categories Age into ranges & add age constraint:



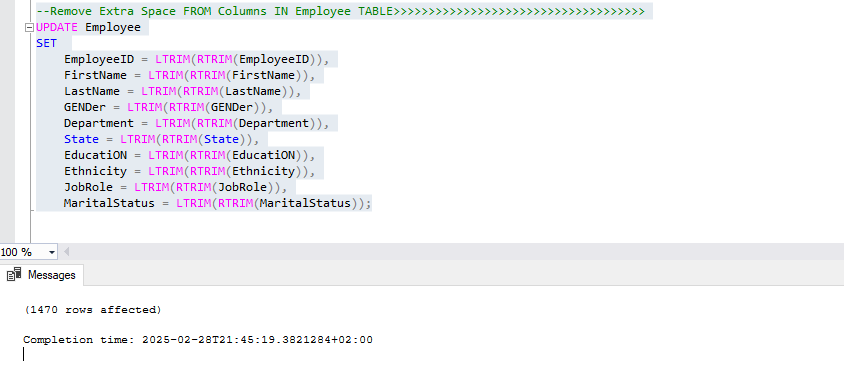




\*\*Categories Distance into ranges:

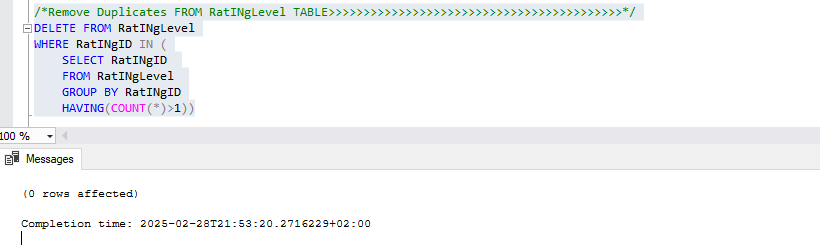


\*\*Remove extra spaces from columns:



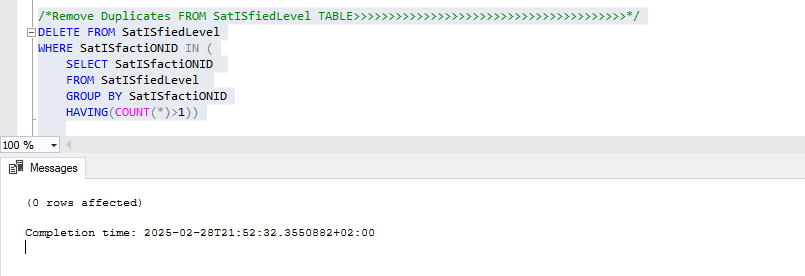
Rating Level table

\*\* Remove duplicates:



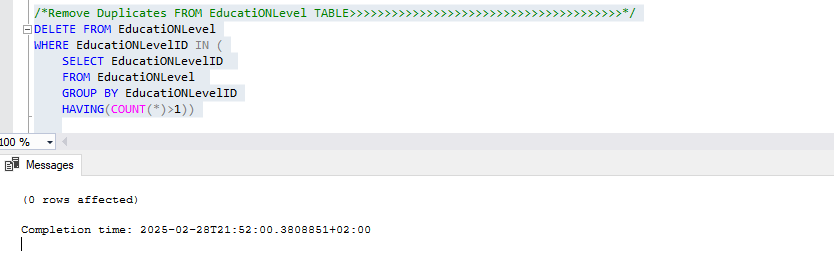
Satisfaction Level table

\*\* Remove duplicates:

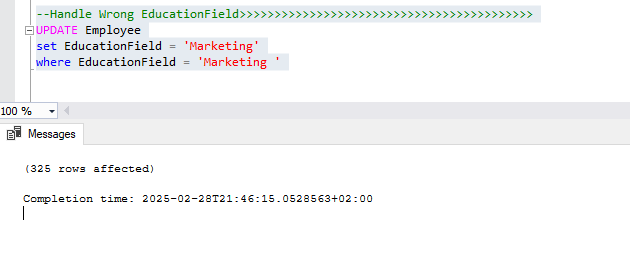


Education Level table

\*\* Remove duplicates:



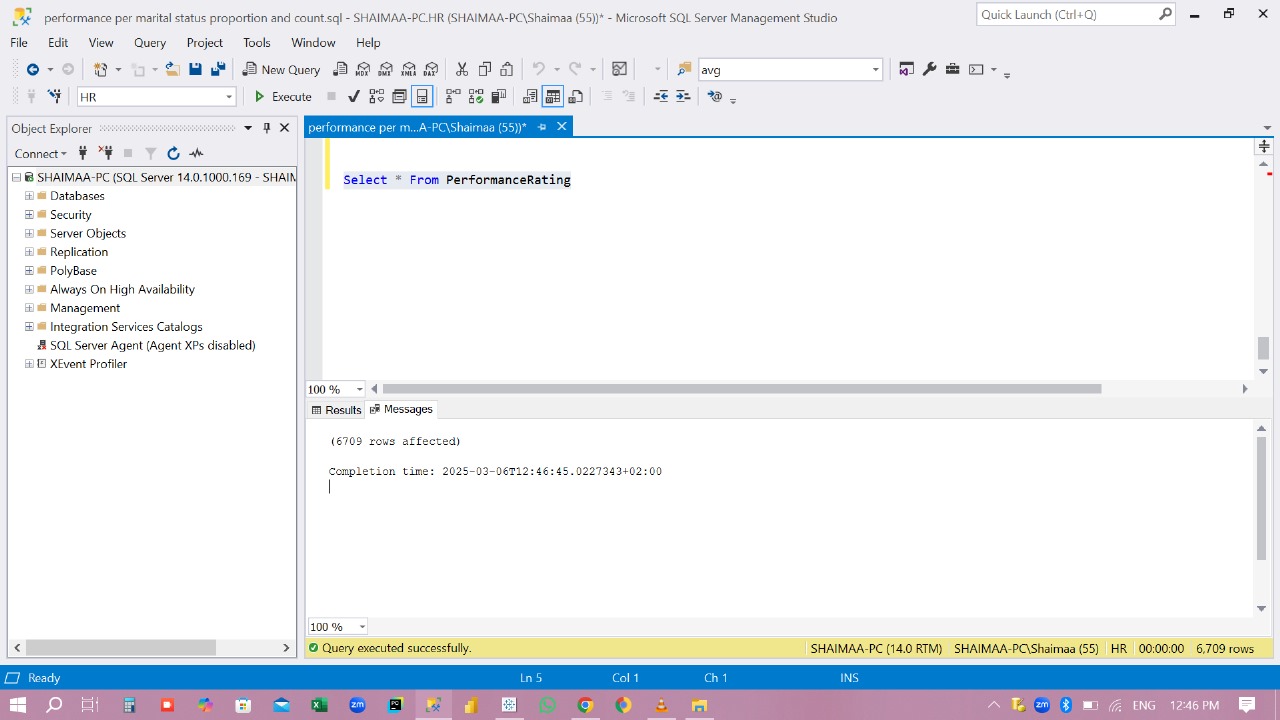
\*\* Handle inconsistent data format by removing extra space in some values:



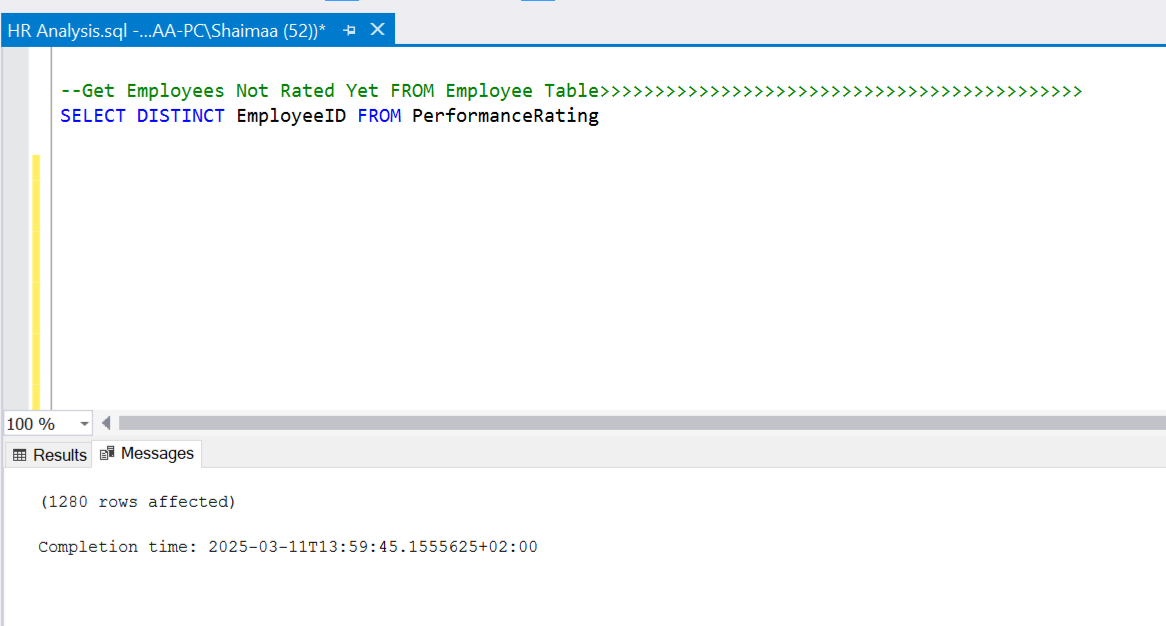
\*\*Additional explorations:

Case 1:

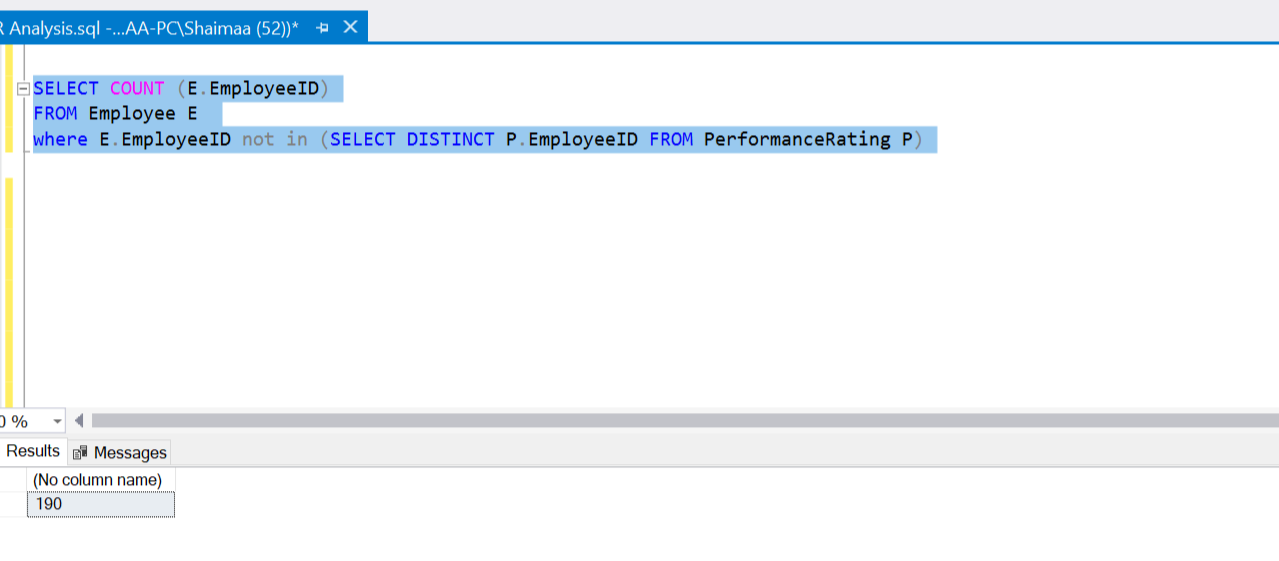
🡪Number of rows in PerformanceRating table:



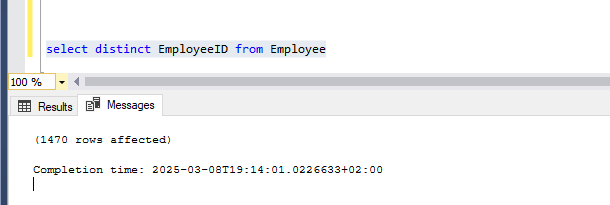
🡪 Rated Employees: (Count Unique Employee ID compared to from Employee table)



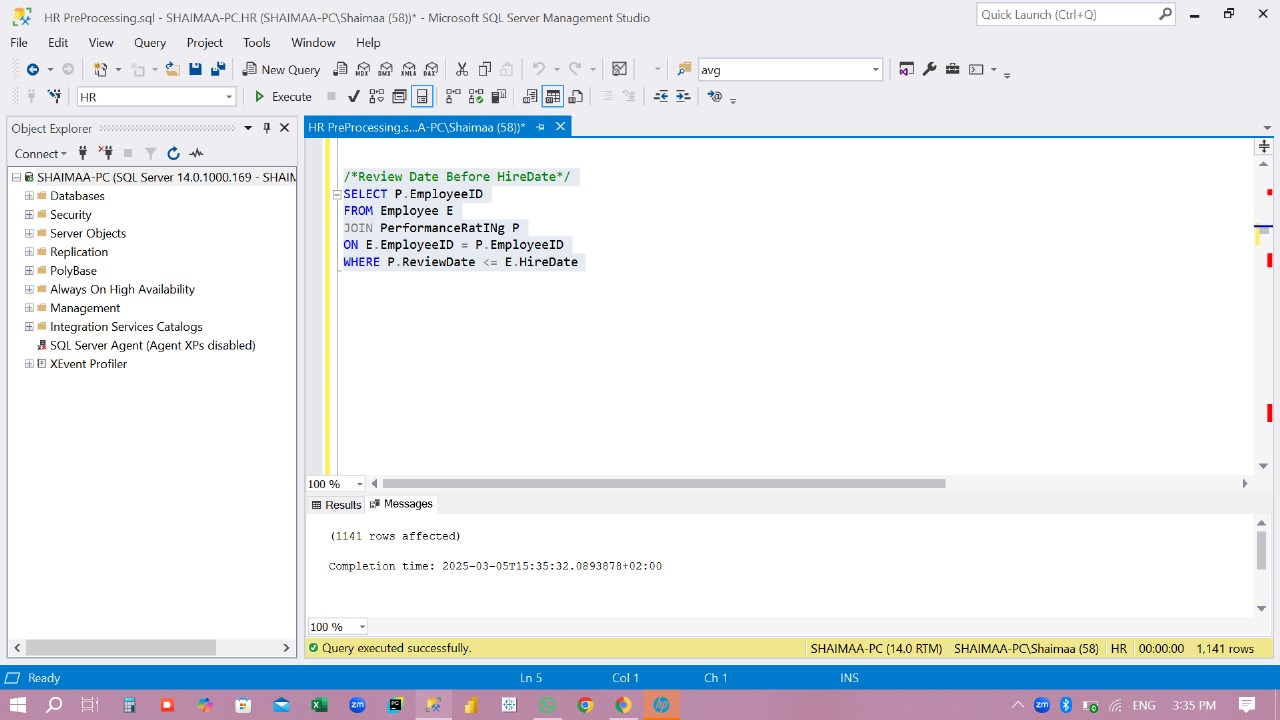
**→**UnRated Employee :

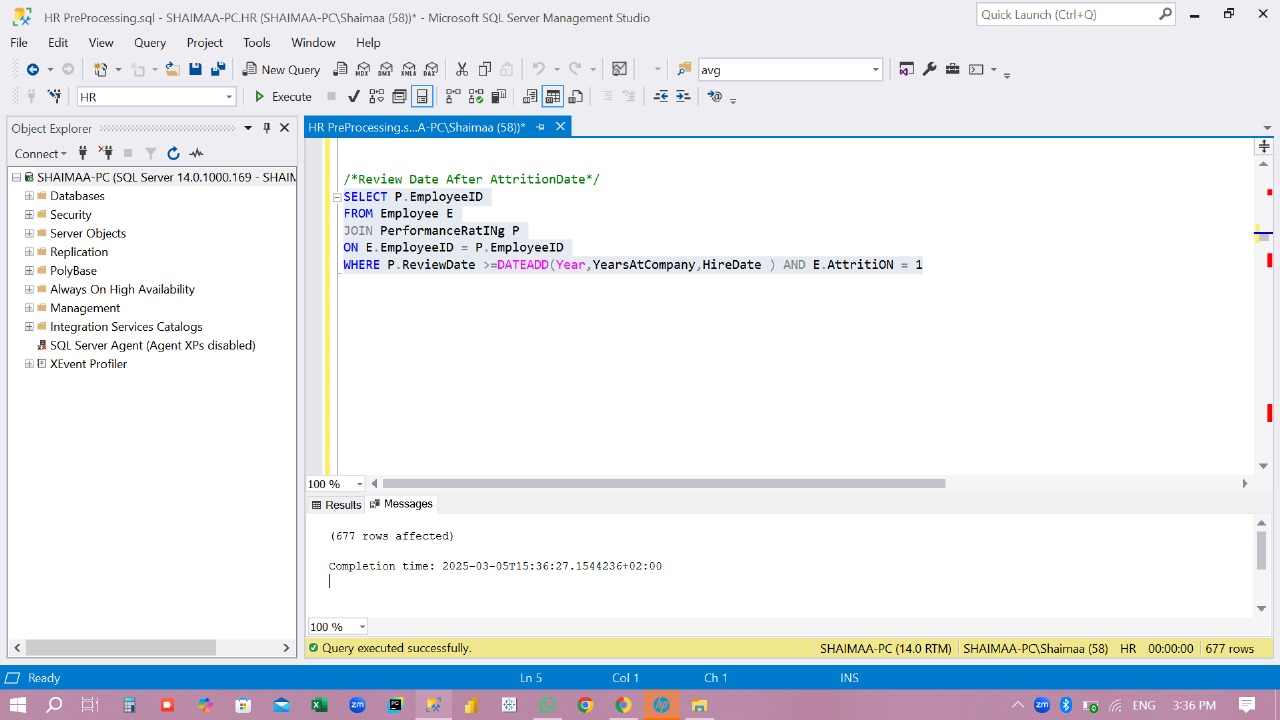


**→**Total Employees :

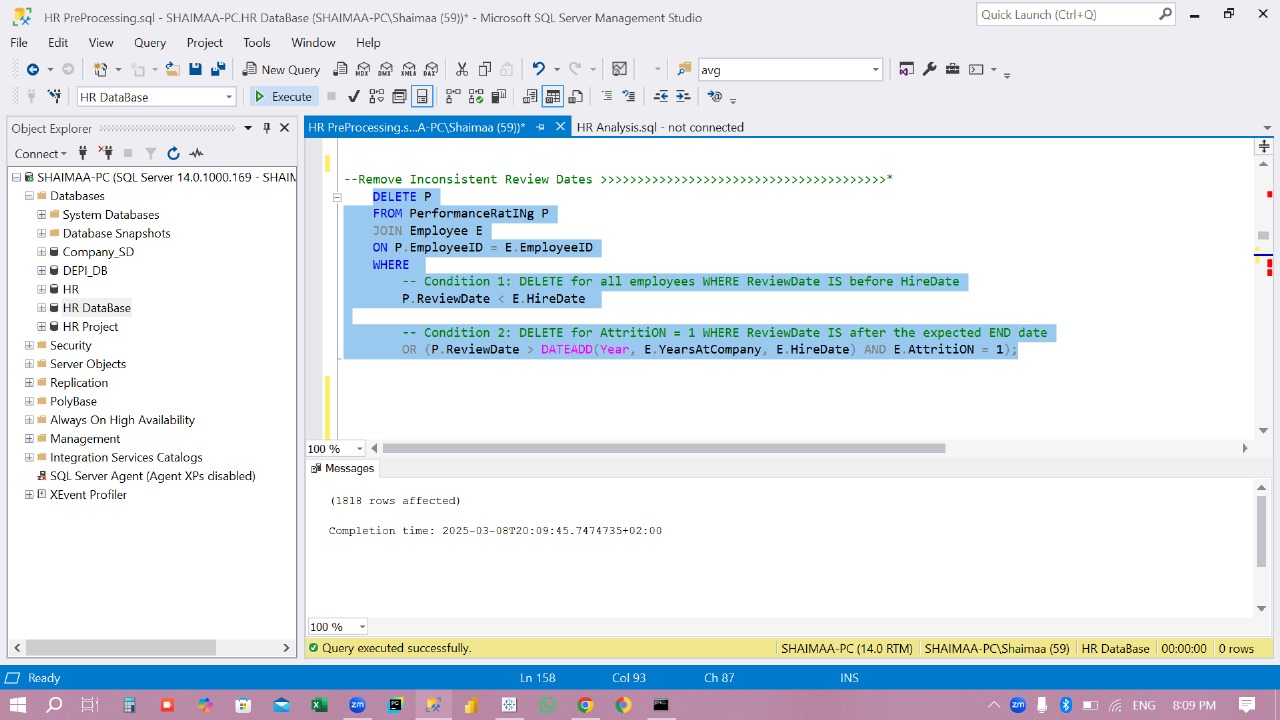


🡪Find inconsistent review dates (before hire date or after attrition date for resigned employees):

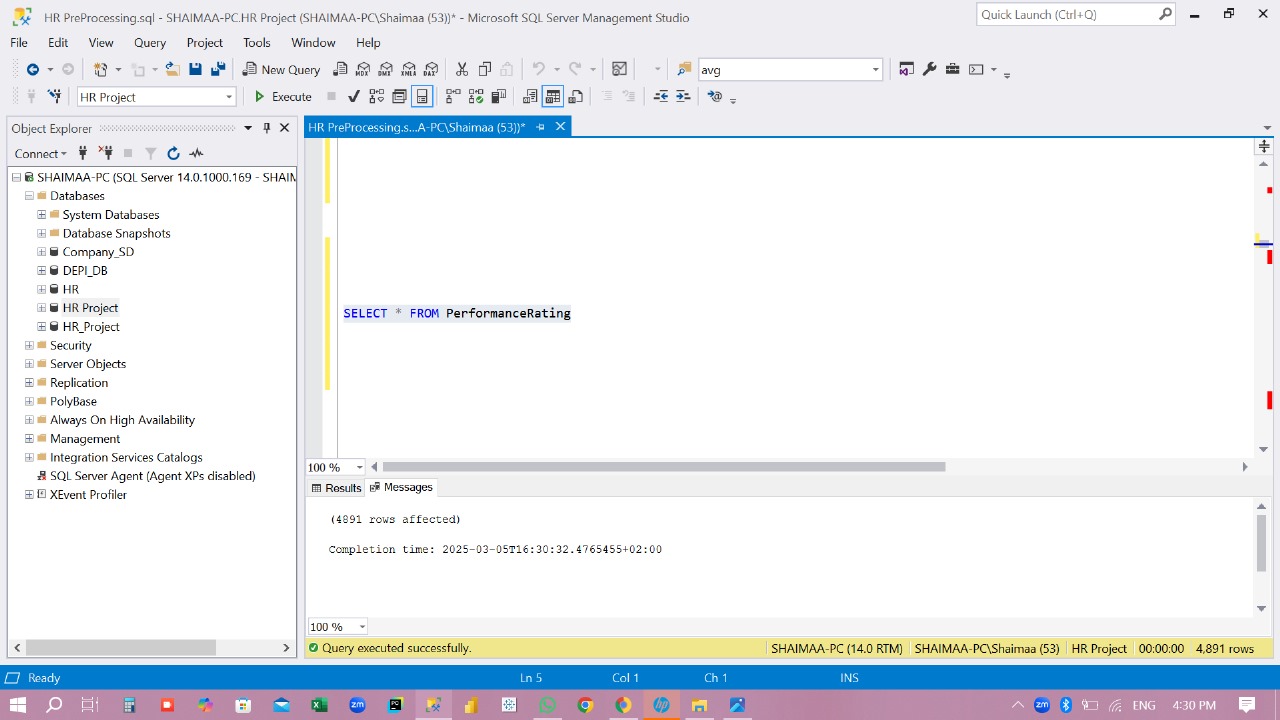




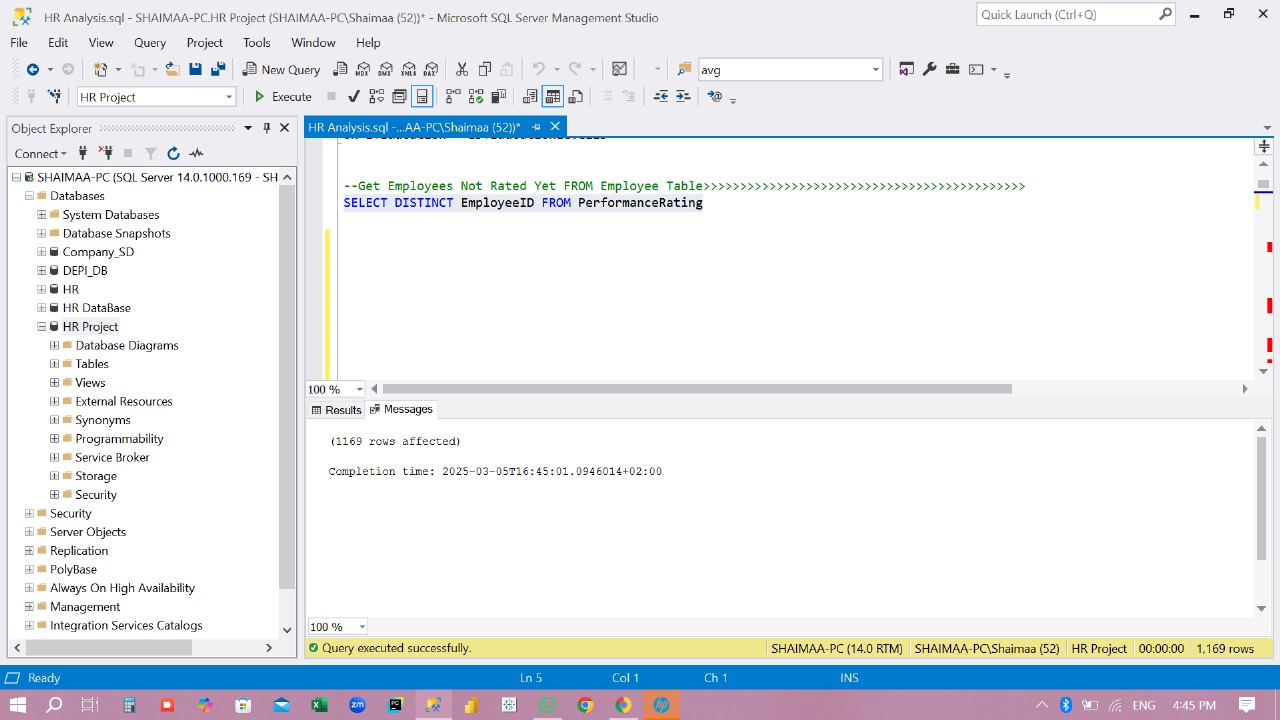
🡪Remove inconsistent review dates:



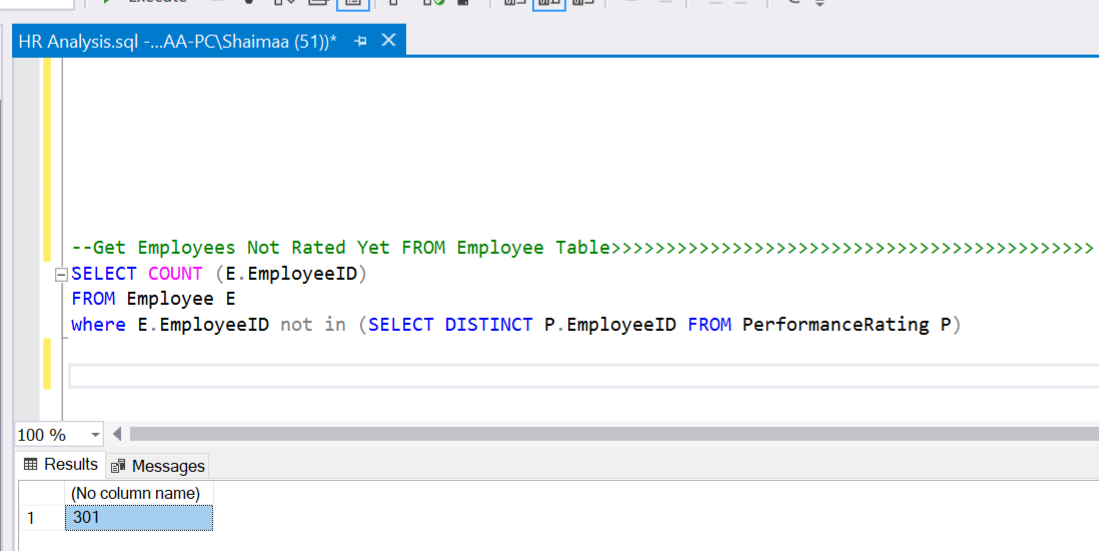
🡪Final corrected number of rows:



Indicate Count Of Rated Employees after drop inconsistent Rates:

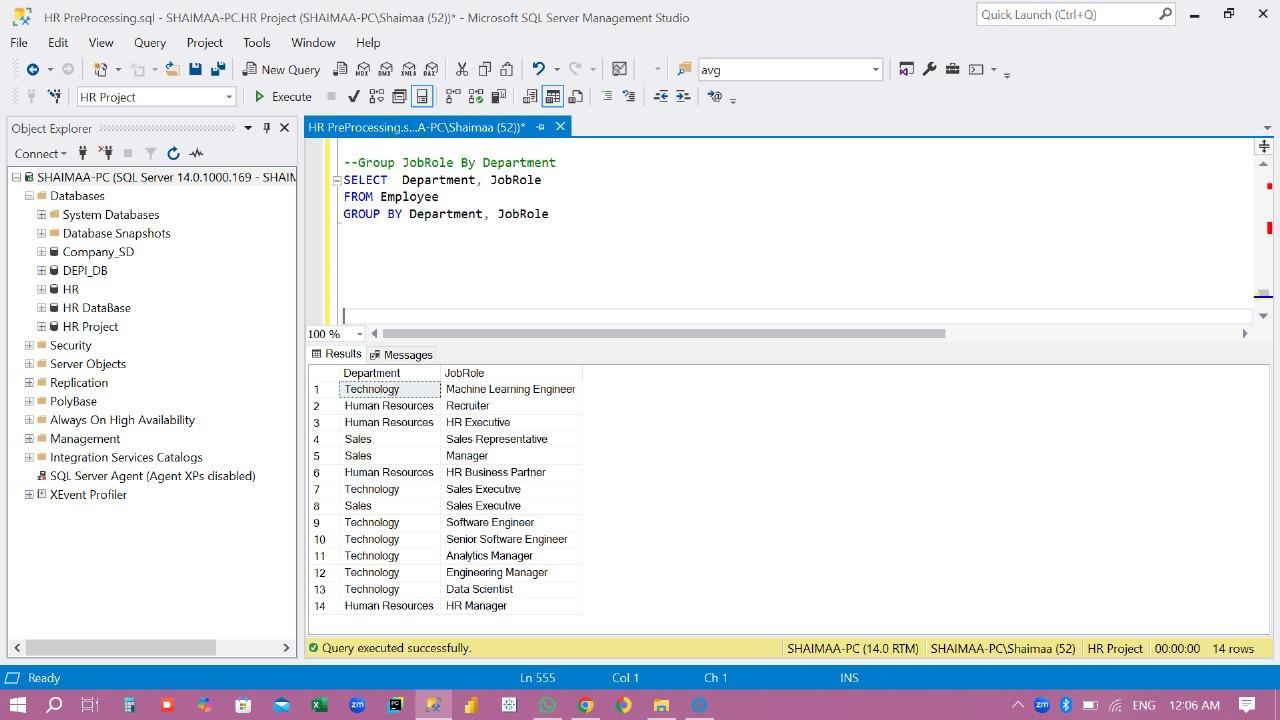


Find UnRated Employees After Drop Inconsistent Rates:



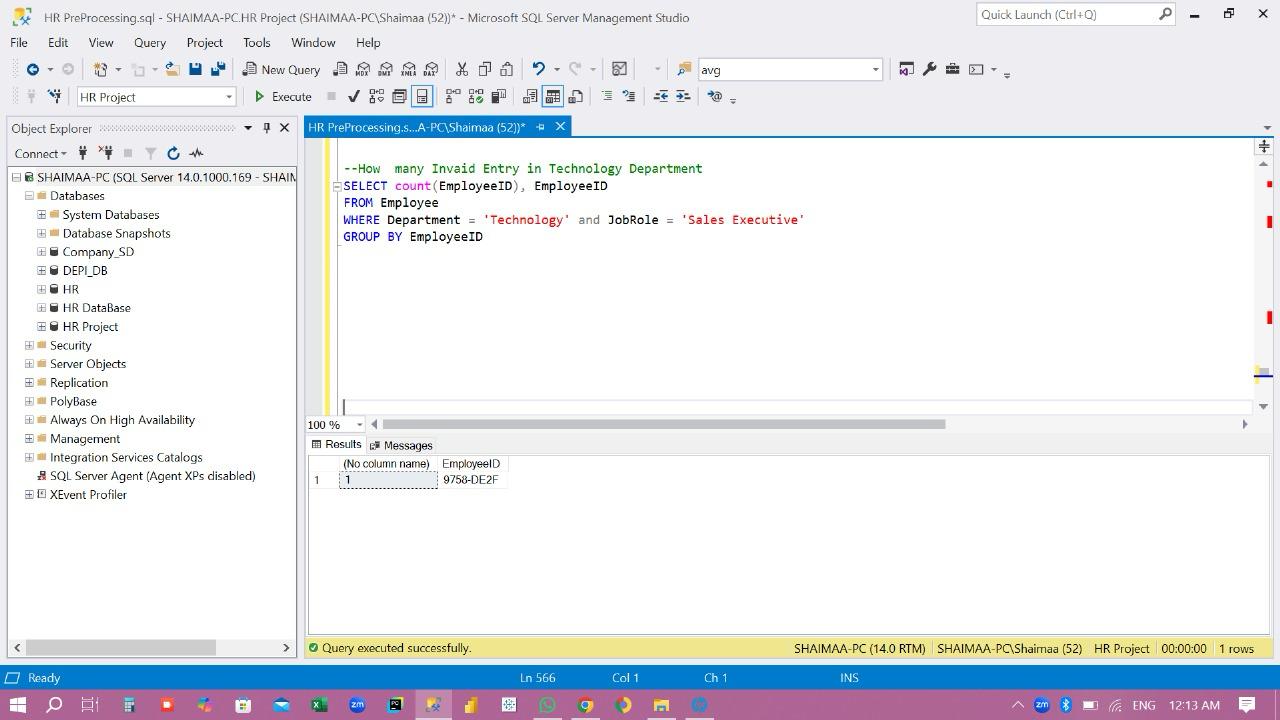
Case 2:

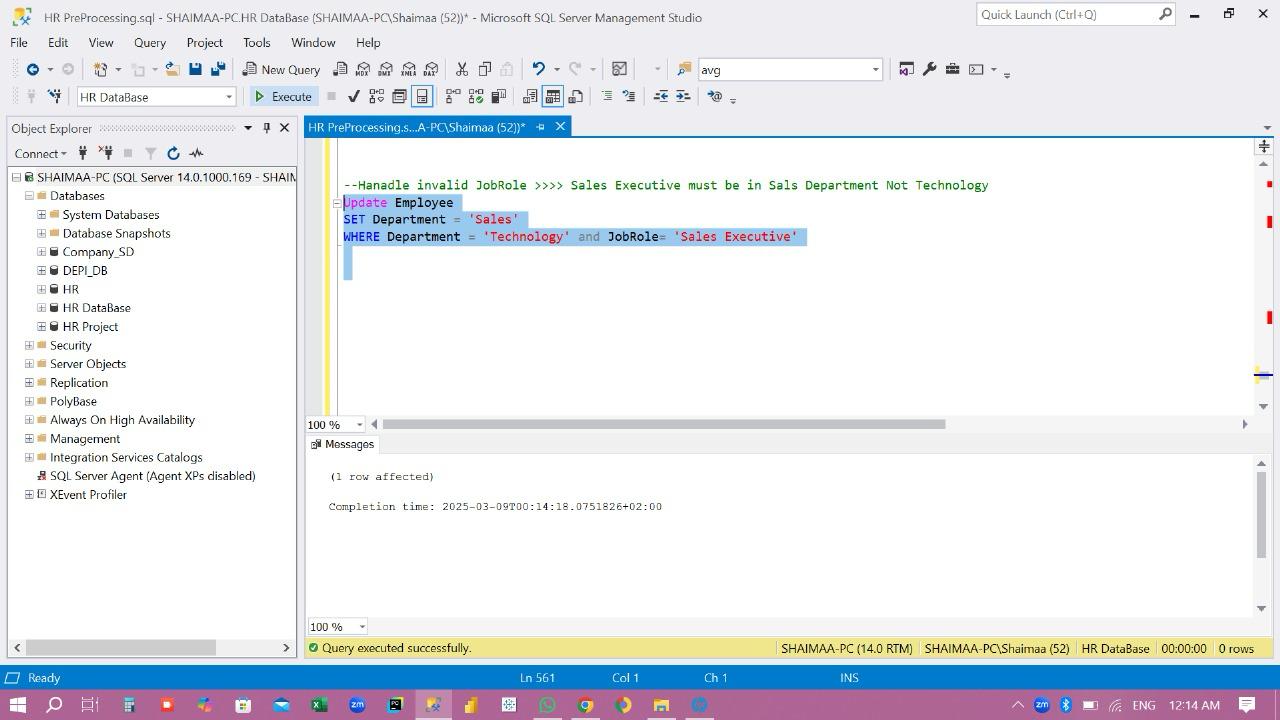
🡪Find different job roles per department:



🡪How many inconsistent entries of sales executive in Technology department:

🡪Handle mistake:





🡪Check again: