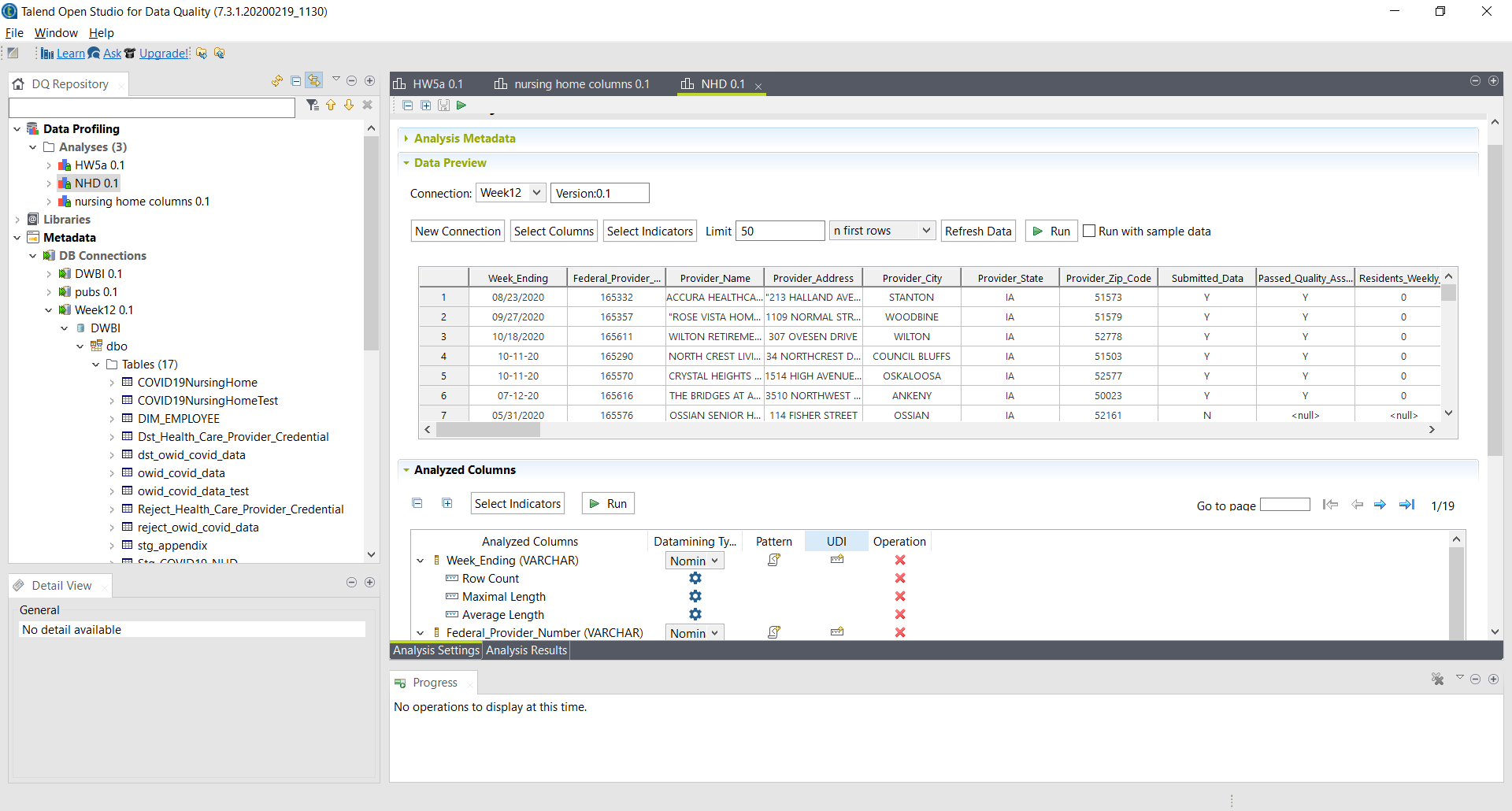
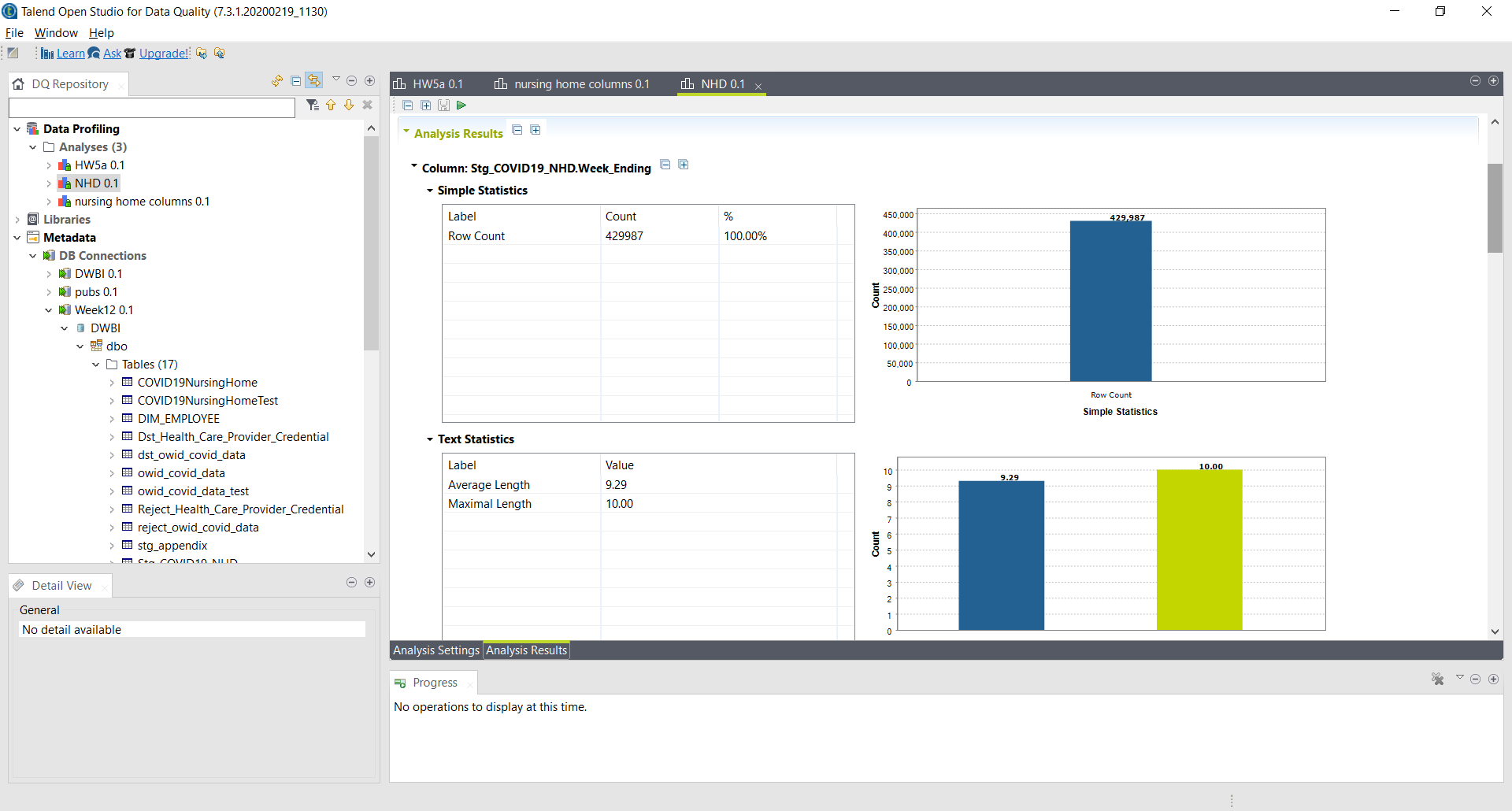
HW9b – Talend Profiling

Smit Patil

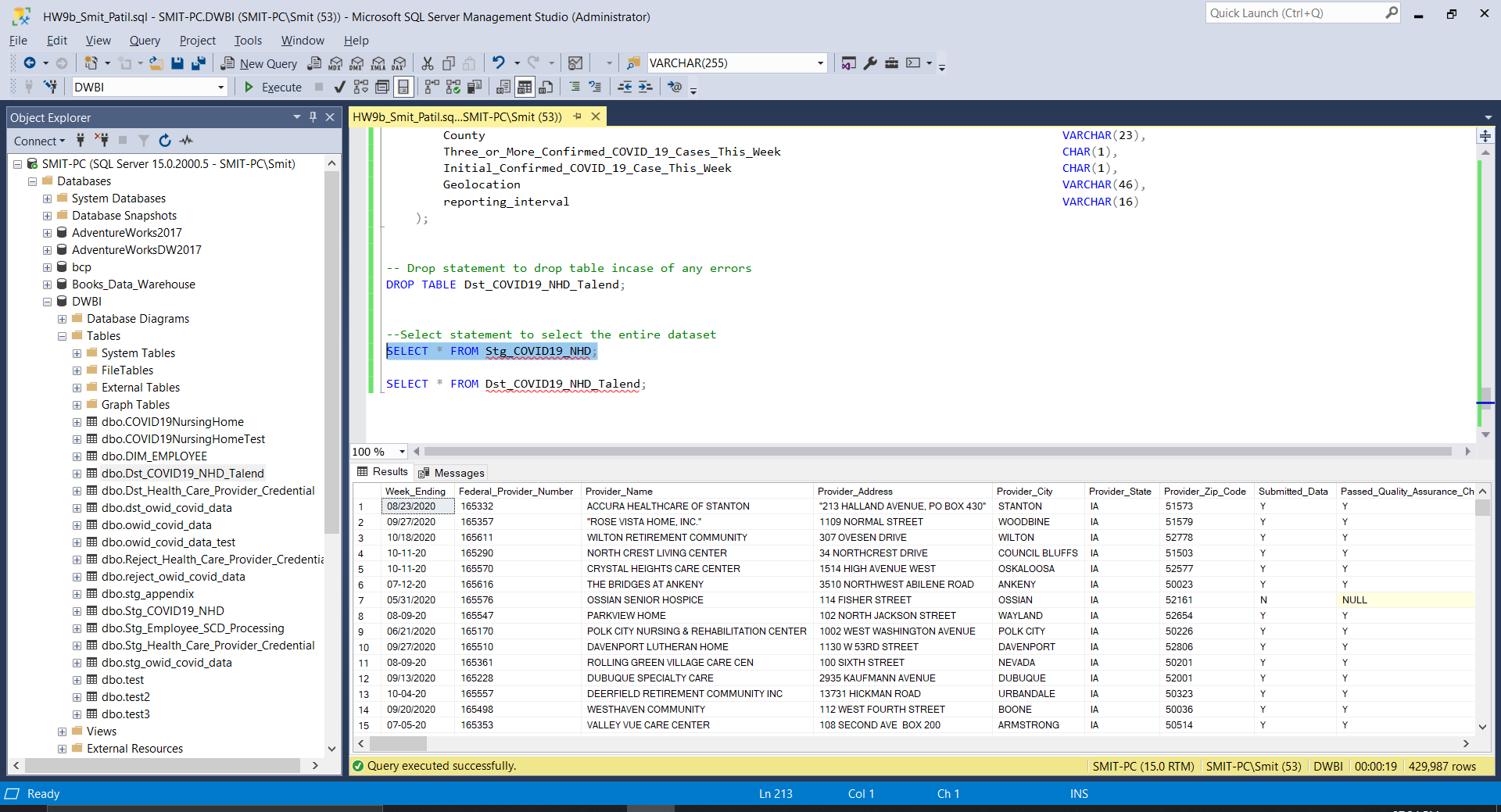
1. Profiled the stagged data in Talend and calculated the length for the CHAR and VARCHAR datatypes.



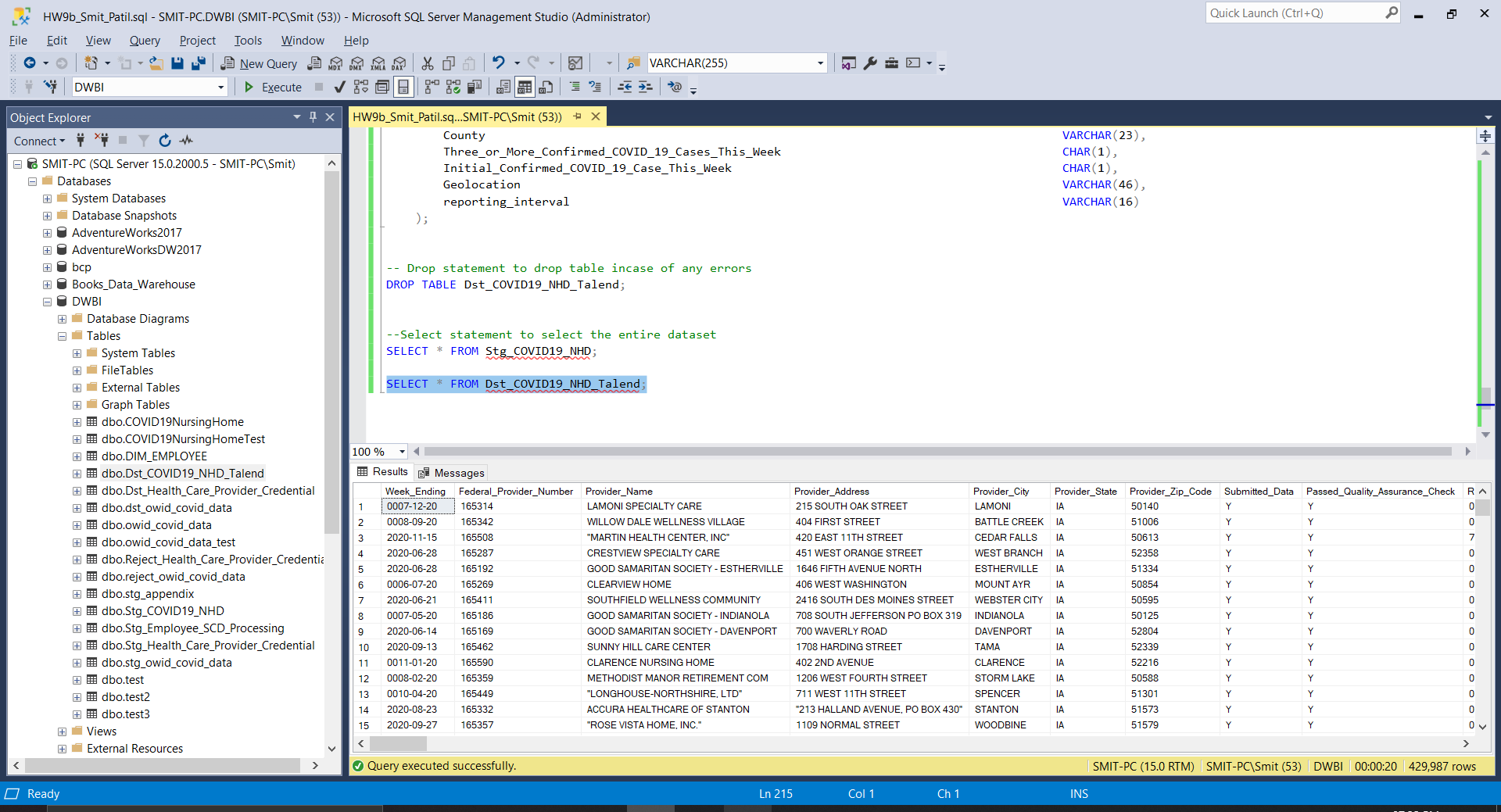
1. Use Simple Statistics and Text Statistics to calculate Maximum, Minimum, Average length of each column.



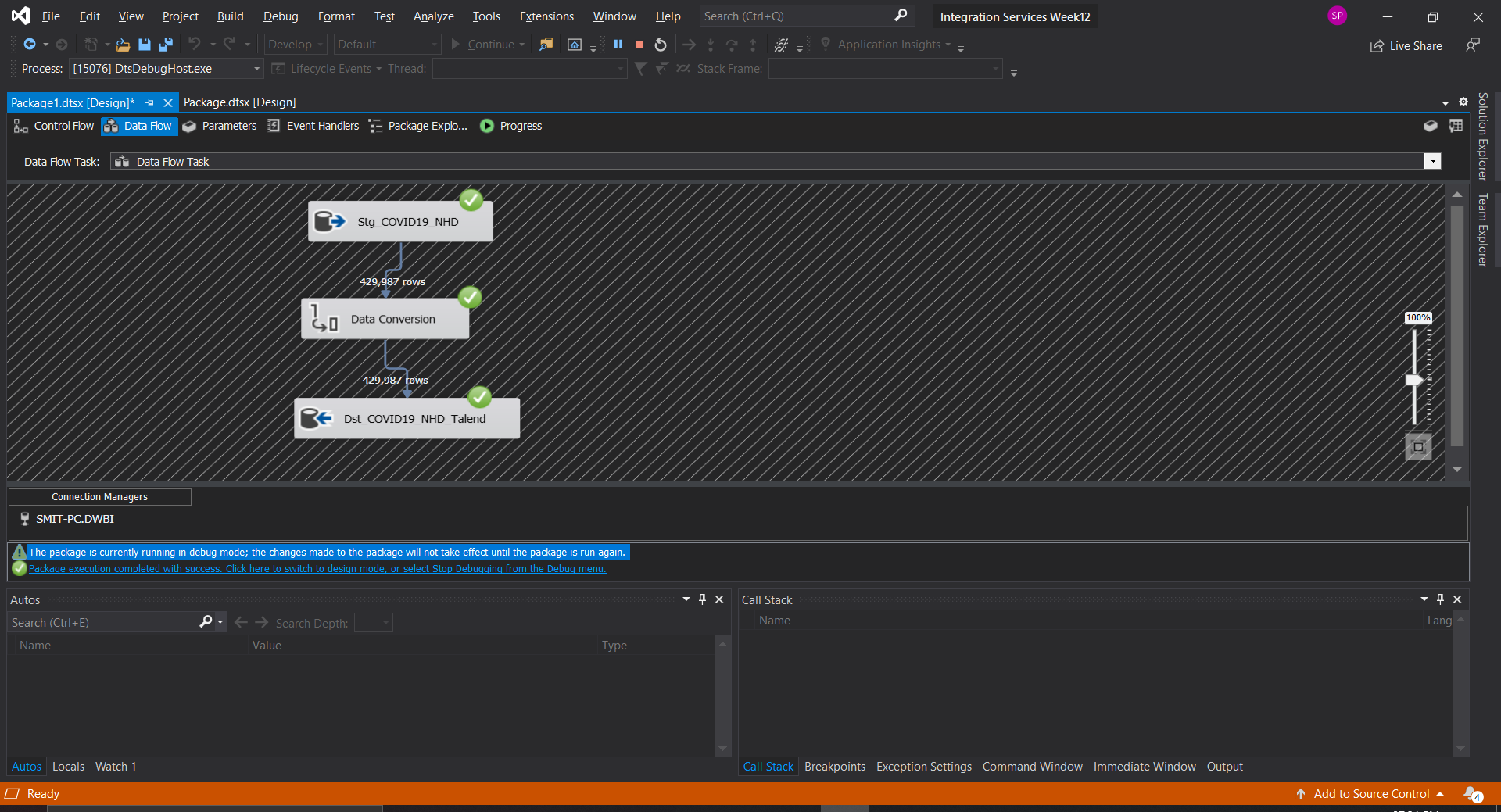
1. Staging table output displaying all the data loaded successfully in the staging table.



1. Destination Table output with all column with proper datatypes loaded successfully.



1. Successful running of the SSIS package to load data into the Destination Table



1. Stacked bar chart with Total Confirmed cases vs. Total Deaths

Chart, bar chart

Description automatically generated

1. United States Heat Map of Total Deaths across the country

Map

Description automatically generated

1. The critical difference with profiling data manually vs. using Talend Data Studio to profile data was time saved by using the tool. Manual profiling can also lead to more number of errors due to human errors. Using a profiling tool helps mitigate such issues and can be used confidently by a developer without any manual errors. Apart from the time saved, Talend Data Studio also has a host of features that can be implemented to achieve even more complexity.