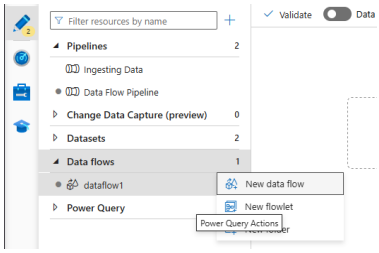
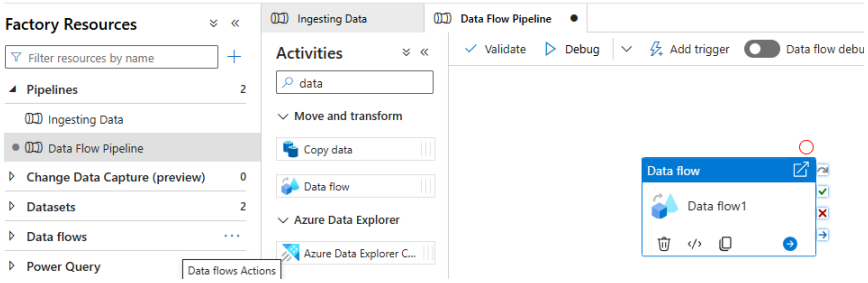
Ansh Ranjan
ADF Case Study
Exercise 3: Data Transformation & Loading for Covid-19 Analysis

TASK 1: Use ADF's Mapping Data Flow

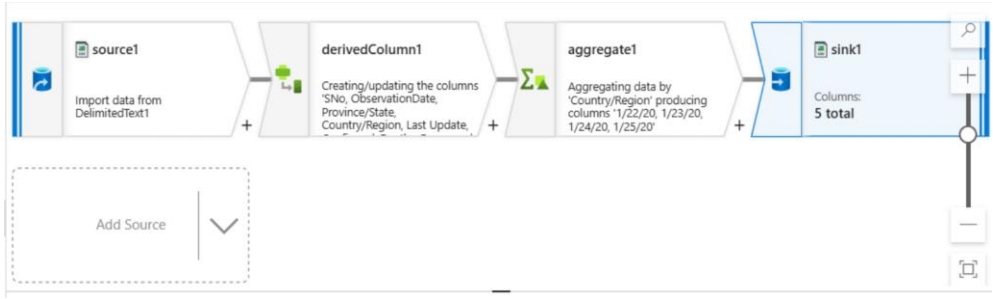
1. Go to Author tab > Data Flow > New Data Flow



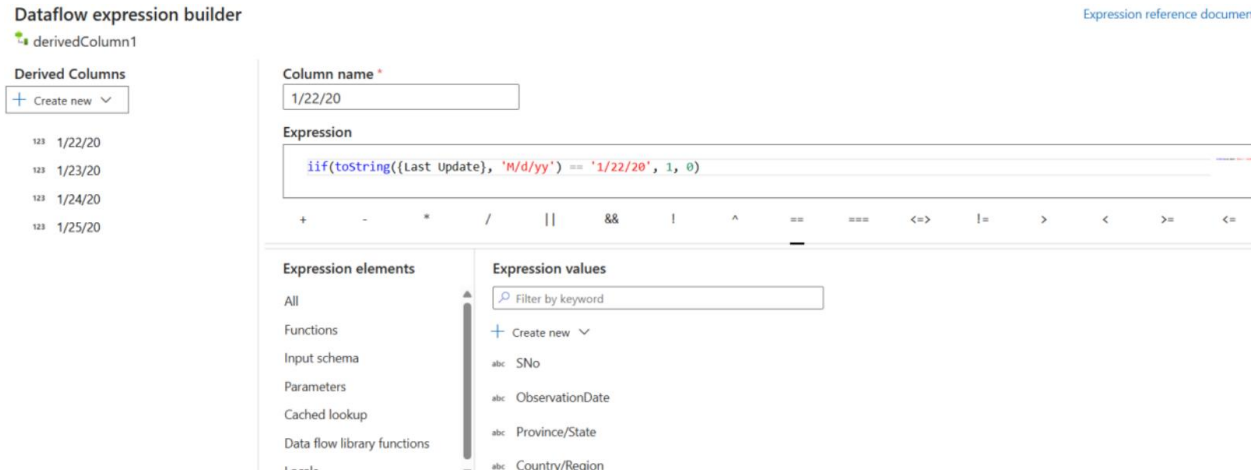
2. Go to Author tab > Pipeline > New Pipeline > and grab the Data Flow activity



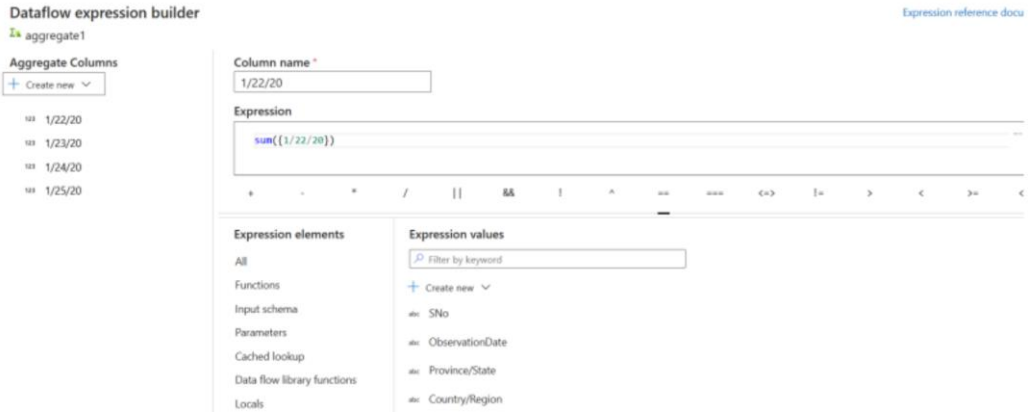
3. Designing the Data Flow. To use adf's data flow, for aggregate purpose first we must use derived columns, then aggregate and at last sink in storage which is accessible in databricks notebooks.



4. In derived columns, make columns for date name for which you are going to do analytical process.

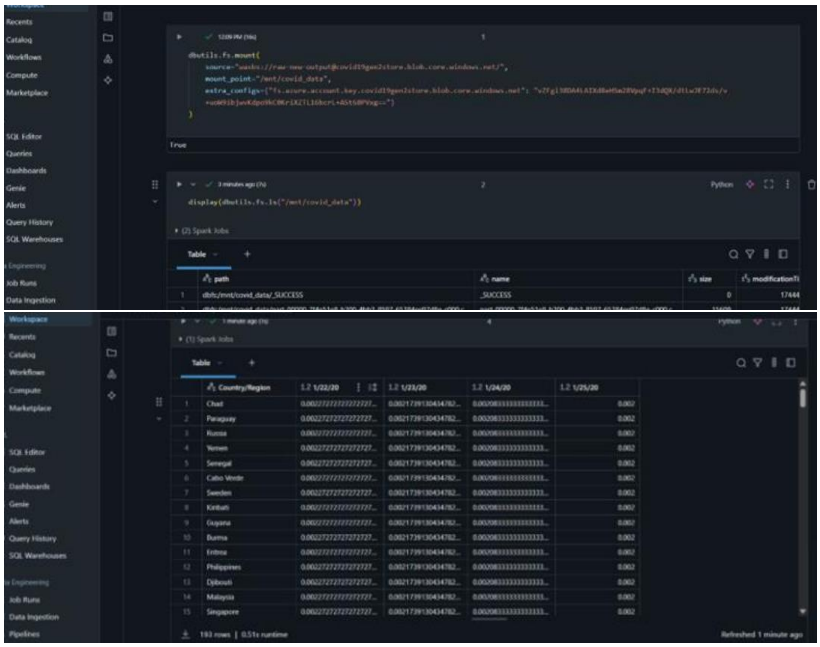


5. After creating columns count the no of cases by date and grouping by each region.



TASK 2: Load the transformed data into Azure Databricks for further analysis

1. Launch a Databricks workspace and Launch it
2. Create a new notebook



3. Creating a visualization based on number of cases per day grouping