**Learning Objective:** Cleaning and preparing data for analysis

**Context:** In this exercise, the student is going to identify/remove duplicates using the Unique Tool and modify the case of values in a column using the Data Cleansing Tool. Real world data usually comes in dirty. If measures are not put in place to authenticate the data as it is entered and stored, it may come with unwanted characters, improper fonts, missing values, and duplicates. It is important that an analyst cleans the data before proceeding to the analysis stage.

#### **Steps to be executed by the student:**

1. Click on Preparation on the Tool Palette
2. Find and drag the Unique Tool tool and connect it to the Input Data tool
3. In the Configuration Window, select Customer ID
4. Find and drag the Data Cleansing tool and connect it to the “U” output anchor of the Unique tool from the step above.
5. In the Configuration Window, select only Country/Region under “Select Fields to Cleanse” and Title Case under “Modify Case”. Leave other checkboxes in the Configuration Window unchecked.
6. Toggle between the input and output anchors to check in Country/Region changed from lower case to title case

**Exercise Question:** How many duplicate values are in the dataset?

**Answer:** 8

**End Goal:** Image of final visualization

