**Learning Objective:** Cleaning Data

**Context:** Real world data usually comes from various sources. If measures are not put in place to authenticate the data entered and stored, it may come with unwanted characters, improper fonts and cases, nulls, duplicates and so on. It is important that an analyst cleans data before proceeding to the analysis stage. In this exercise, the learner is going to remove duplicates and clean data.

**Instructions:**

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 Cleaning tool and connect it to the “U” output anchor of the Unique tool
5. In the Configuration Window, select only Country/Region under “Select Fields to Cleanse” and Title Case under “Modify Case”. Leave all other checkboxes in the Configuration Window unchecked.
6. Toggle between the input and output anchors to check in Country/Region in is the title case

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

**Answer:** 8

**End Goal:** Image of final visualization

