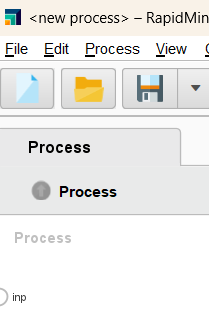
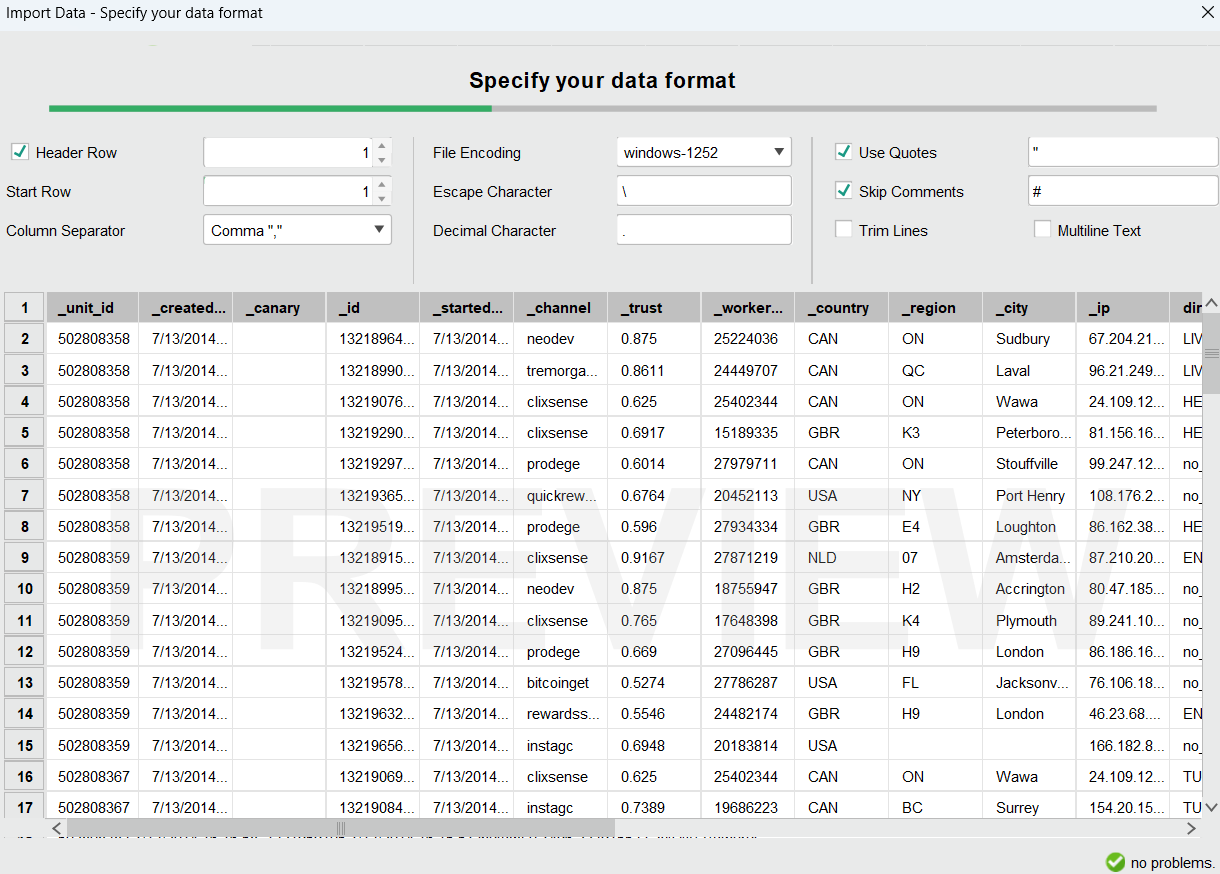
**MSDS655 – Business Intelligence - Lab 5**

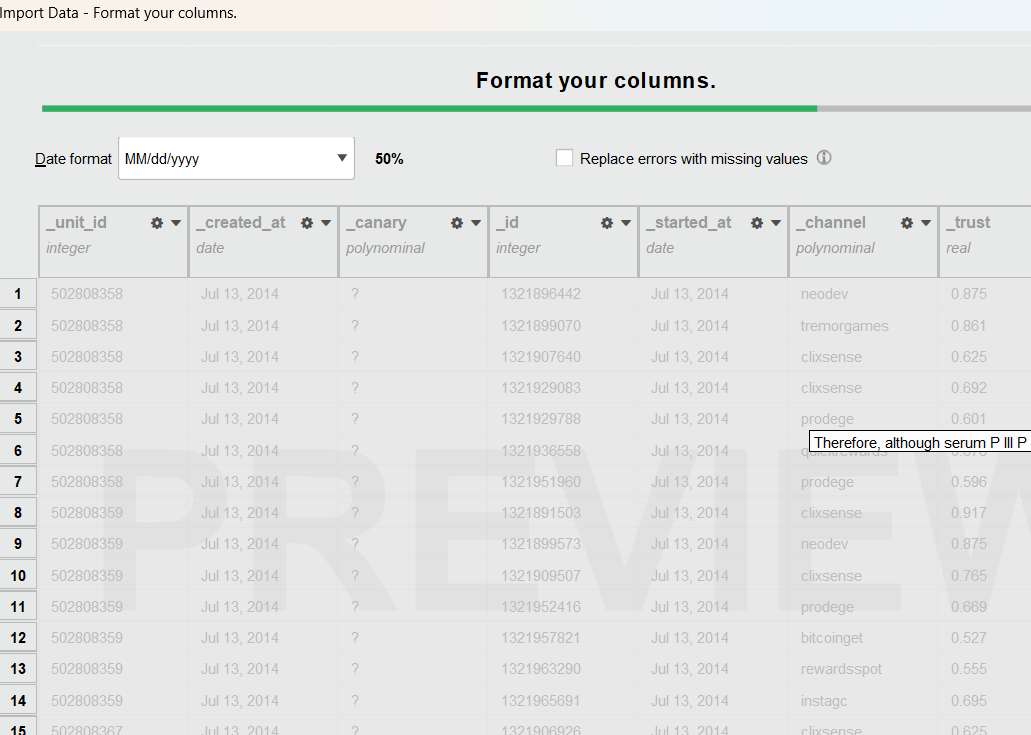
**Topic: Predictive Analytics using Text Analysis, Text Processing and Web Mining.**

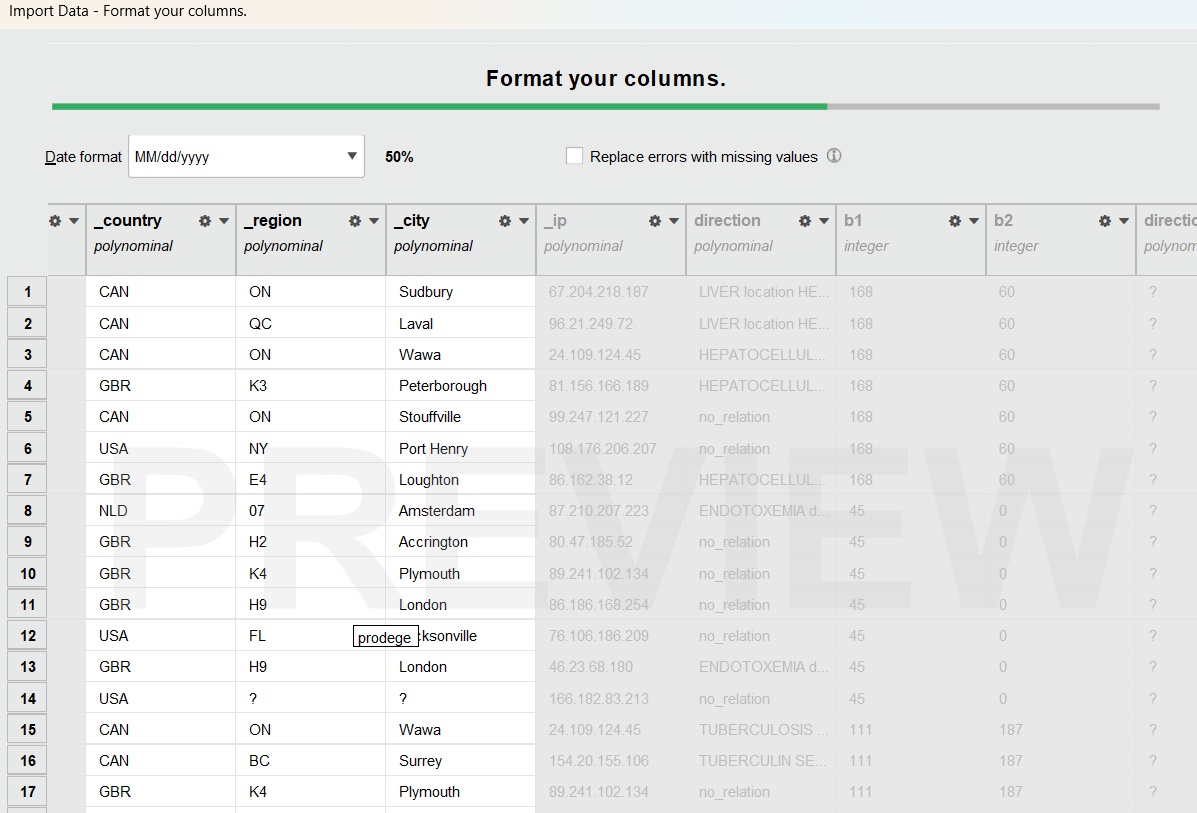
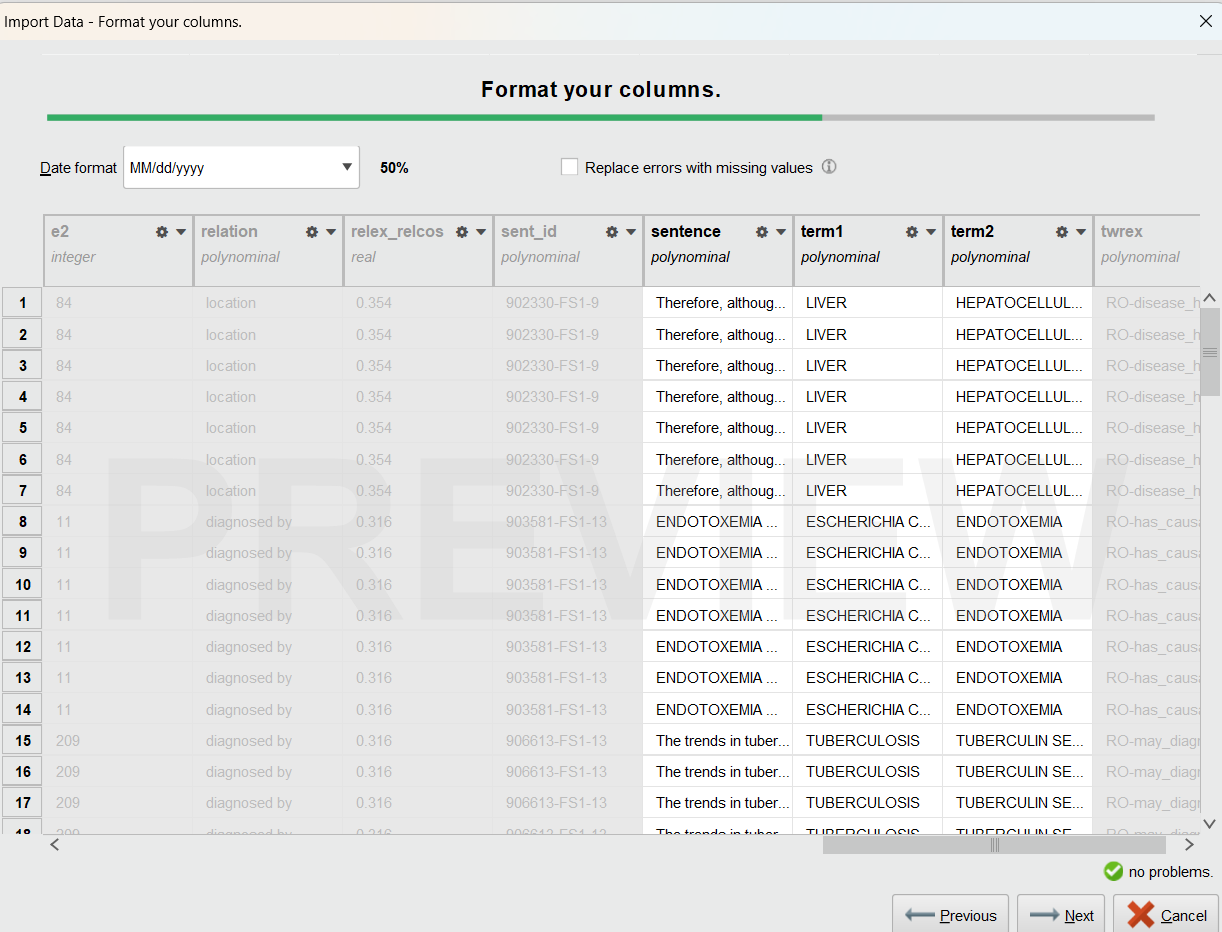
**Goal: The goal of this lab is import a medical text csv file into Rapid Miner. Then, you will apply filter and sort criteria to find information about epileptic seizures sorted by country. This data set is open source data provided by Kaggle for analytical research.**

Start new process  


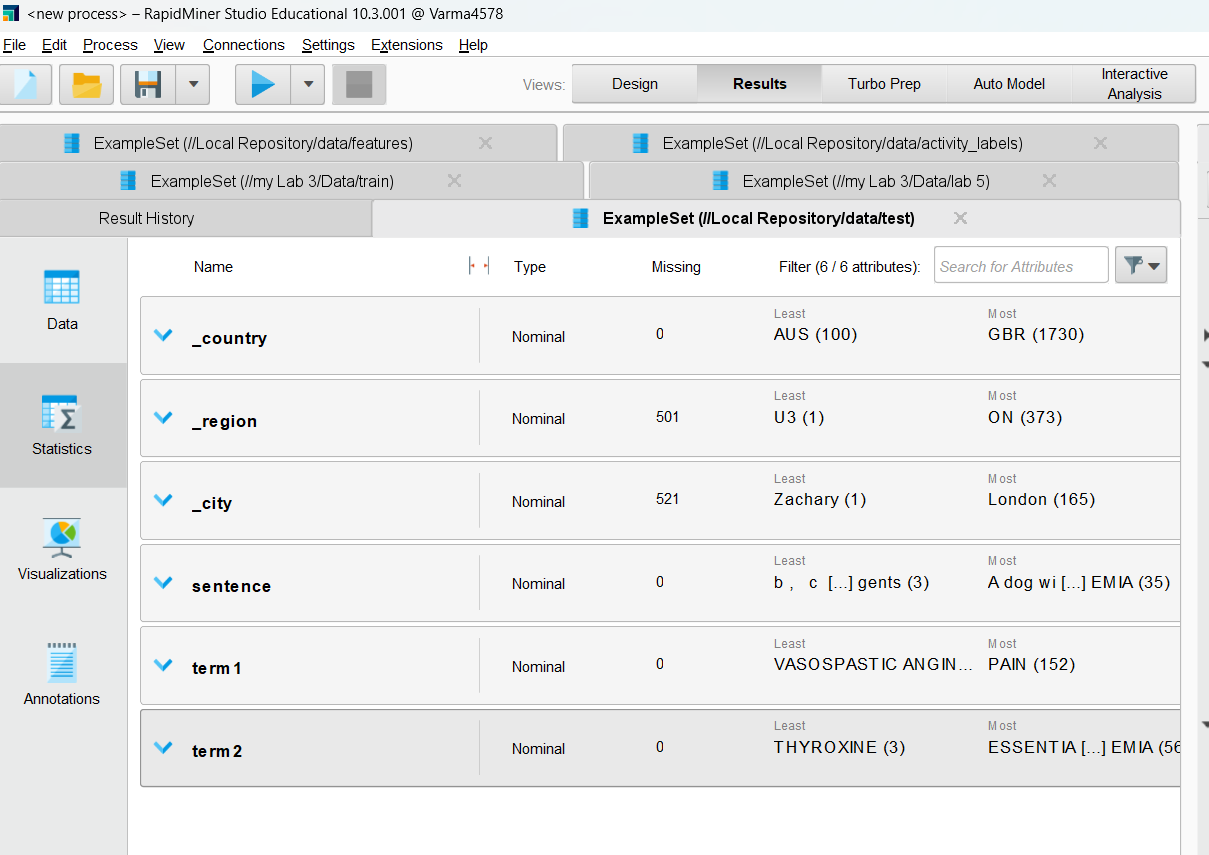
Import data



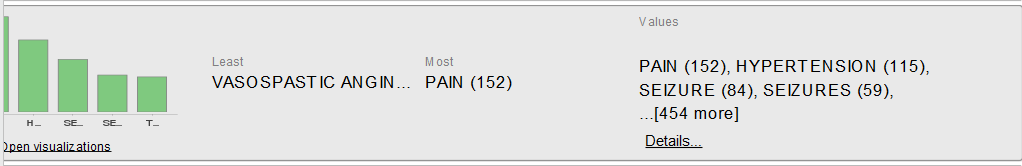
Exclude the coloumns which are not needed  


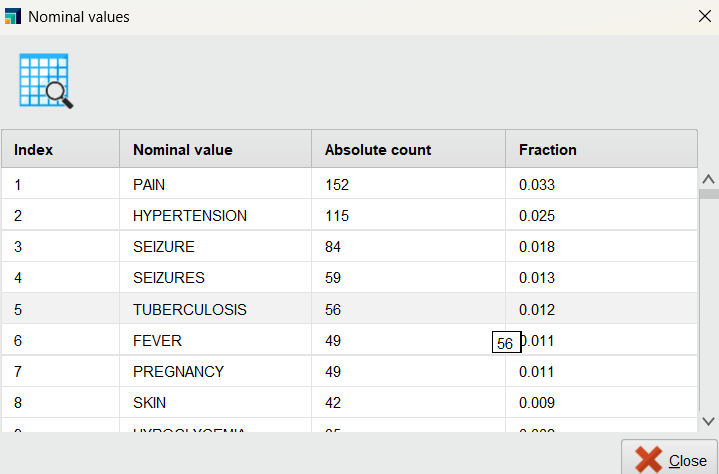
 

Click on finish button

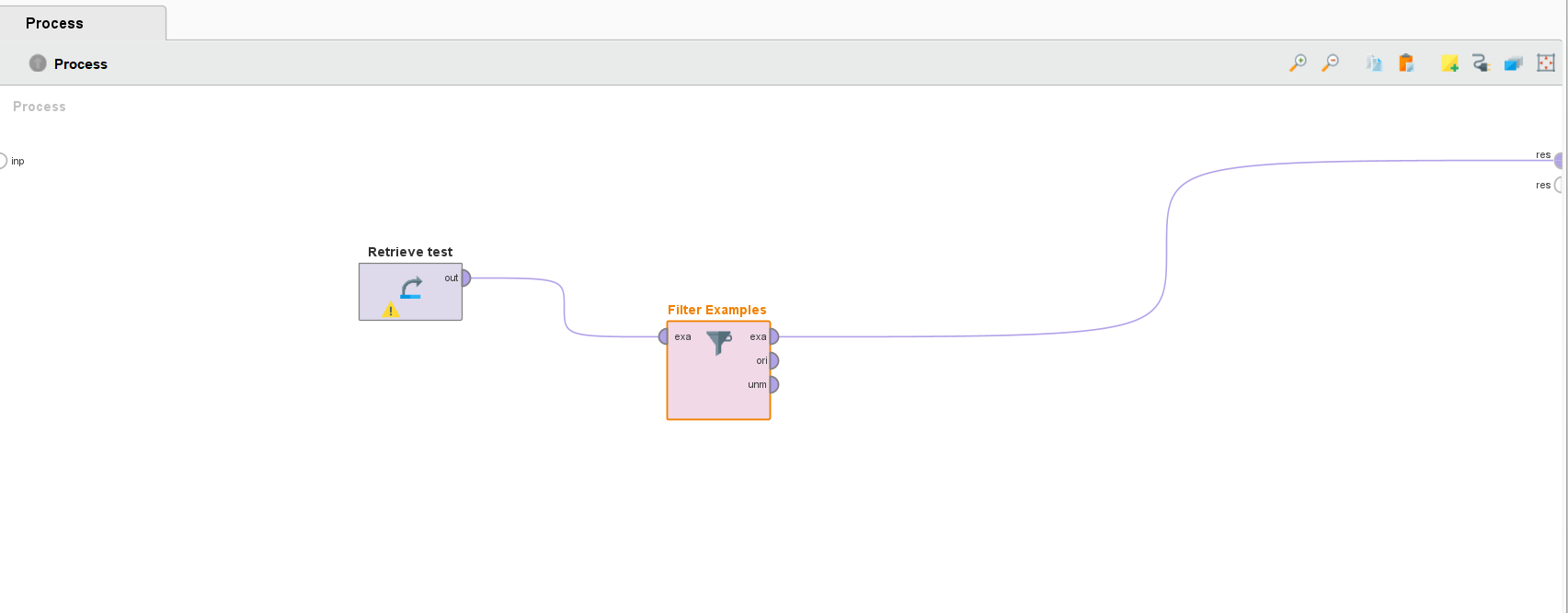
Now click on arrow to start the process then click for dstatistics to know the count information  


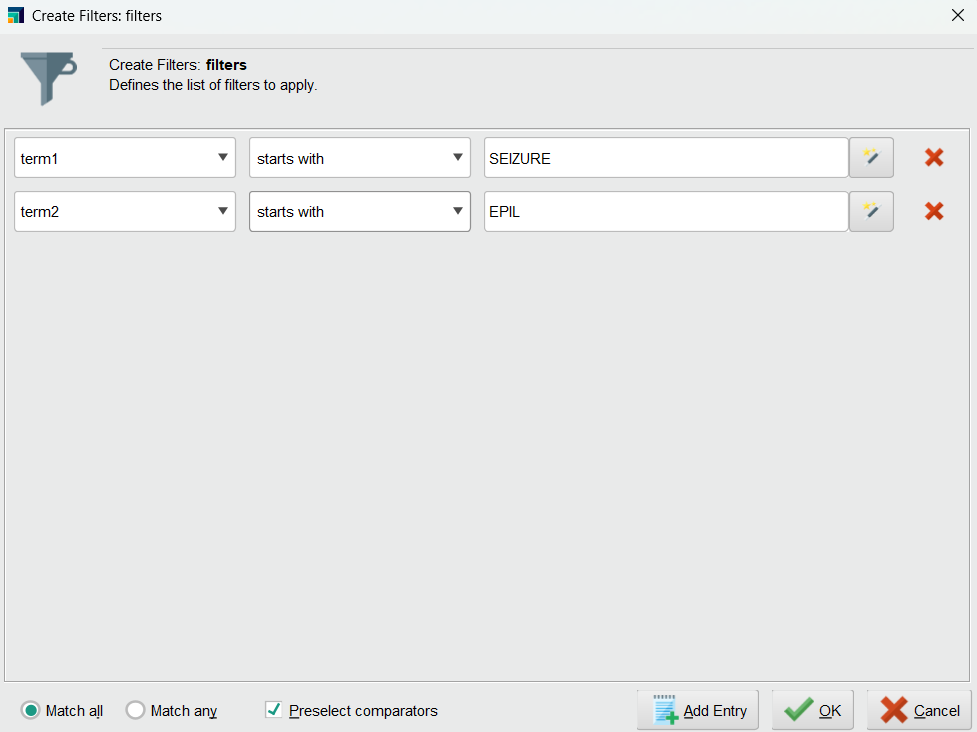
Click on arrow under term 1  

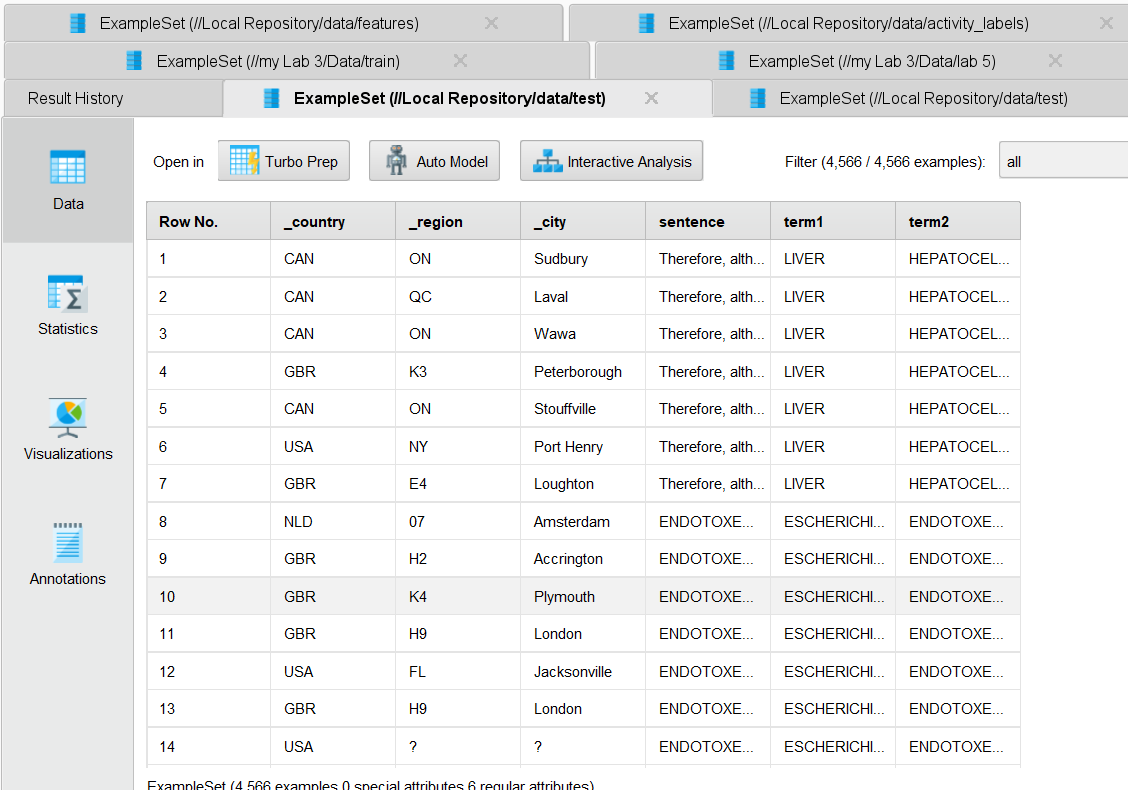

And then Details under values  


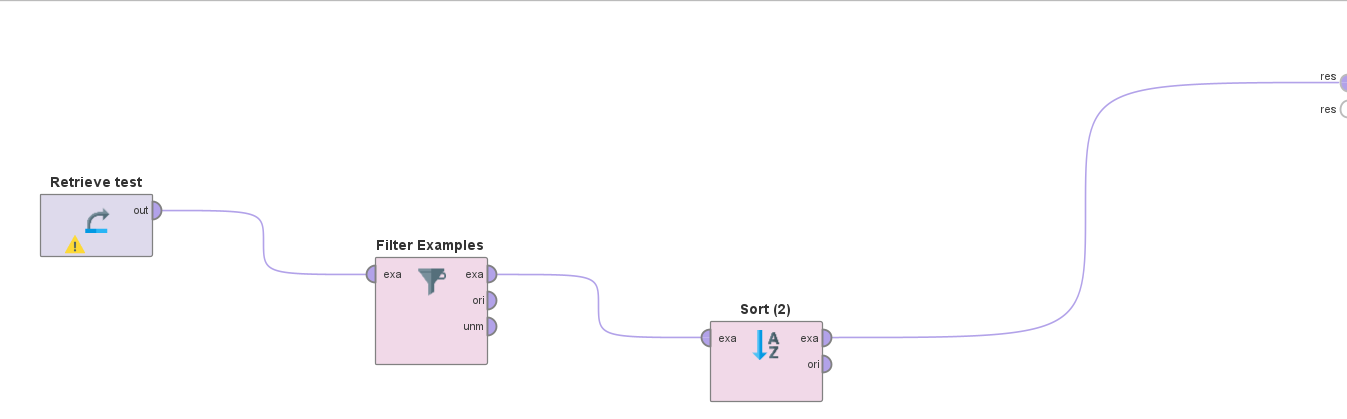


Based on this list, we can see that the most common terms (conditions) are PAIN, HYPERTENSION and SEIZURE and the plural SEIZURES.



Creating filter  


Results  




I can see Country is sorted in order  
