Step by Step Example

This extension can create a concept cloud from text data captured in Modeler object.

This can be either output of IBM Text Analytics node (i.e. Text Link Analysis) or any other data column containing text concepts. This document will provide example for both cases. First let's get started with user interface.

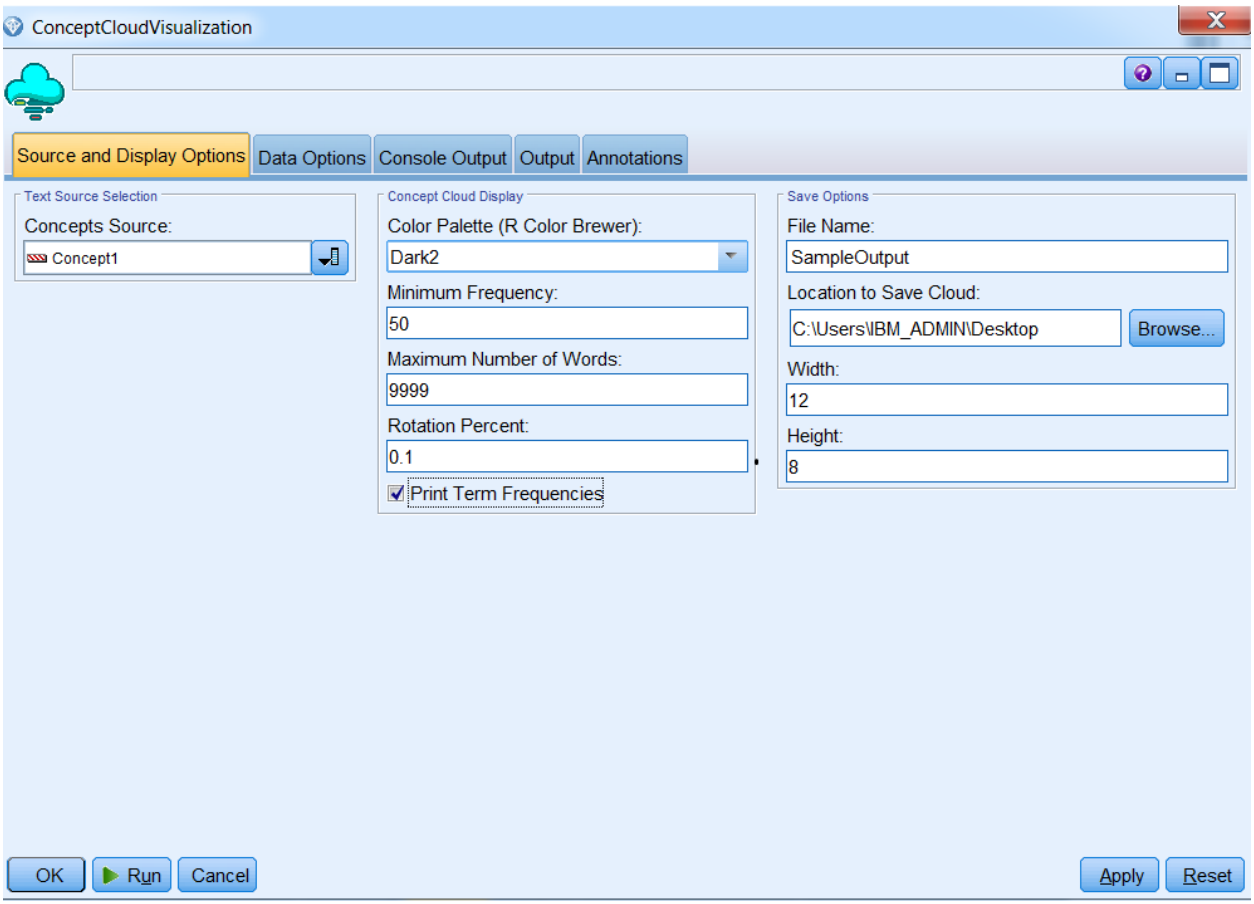
User Interface

First tab for this node ("Source and Display Options") allows to configure source of the text data to be visualized, style of the output cloud as well as save options of the output file.

First choose the source of the text data by choosing proper column in Concept Source field.

Adjust Concept Cloud Display options. Choose colors to be used based on the R Color Brewer package palettes. The minimum frequency of words to be used in the concept cloud, the maximum number of words to display, and the rotation percent of words can be set in this section. You can also print the words with their respective frequencies by checking the box in this section.

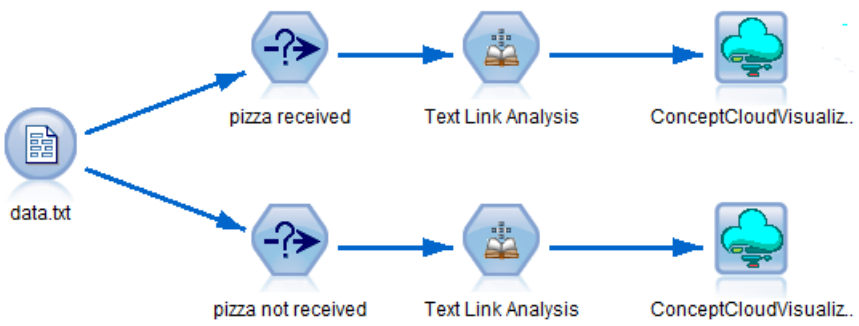
To save output concept file in PNG file, specify name of the file and its location in Save Options section. The width and height values in this section are in inches.



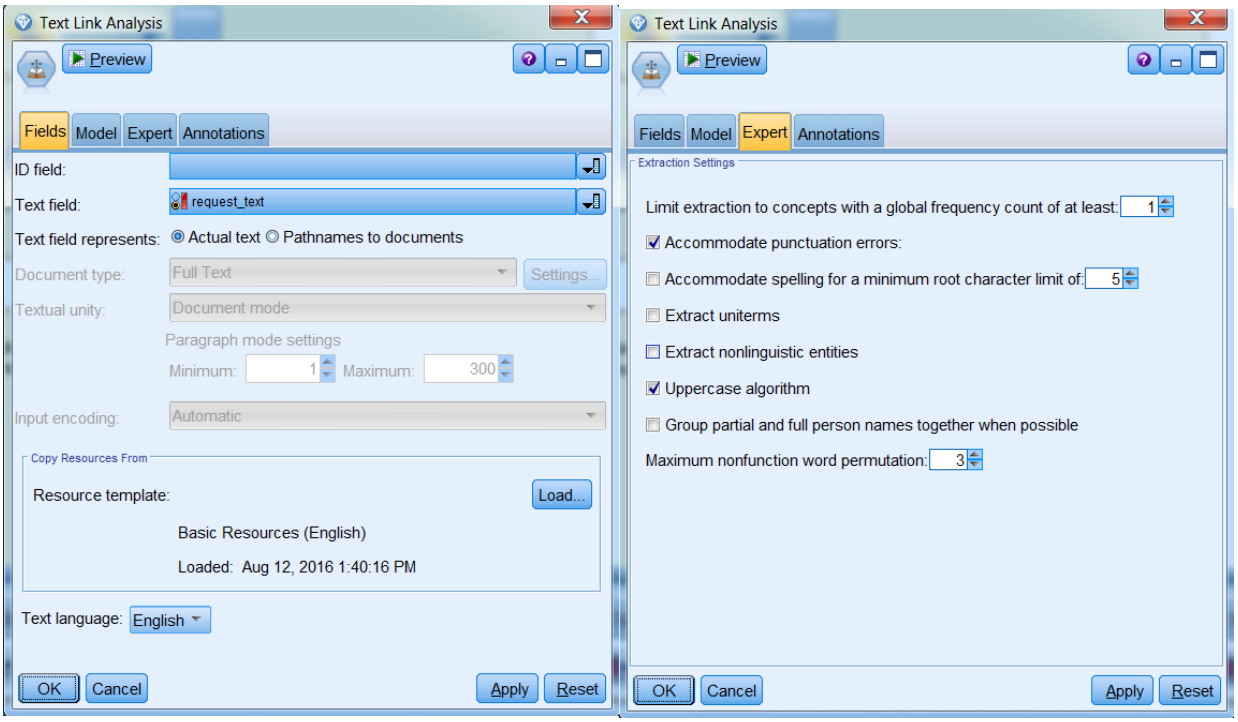
Text Link Analysis output visualization (Example 1 in sample stream)

In this example we will use dataset (data.txt) that is a group of postings from the Random Acts of Pizza sub-Reddit. Each observation in this dataset is the original post made asking for pizza donation and a column with True or False representing if original poster (OP) received a pizza.

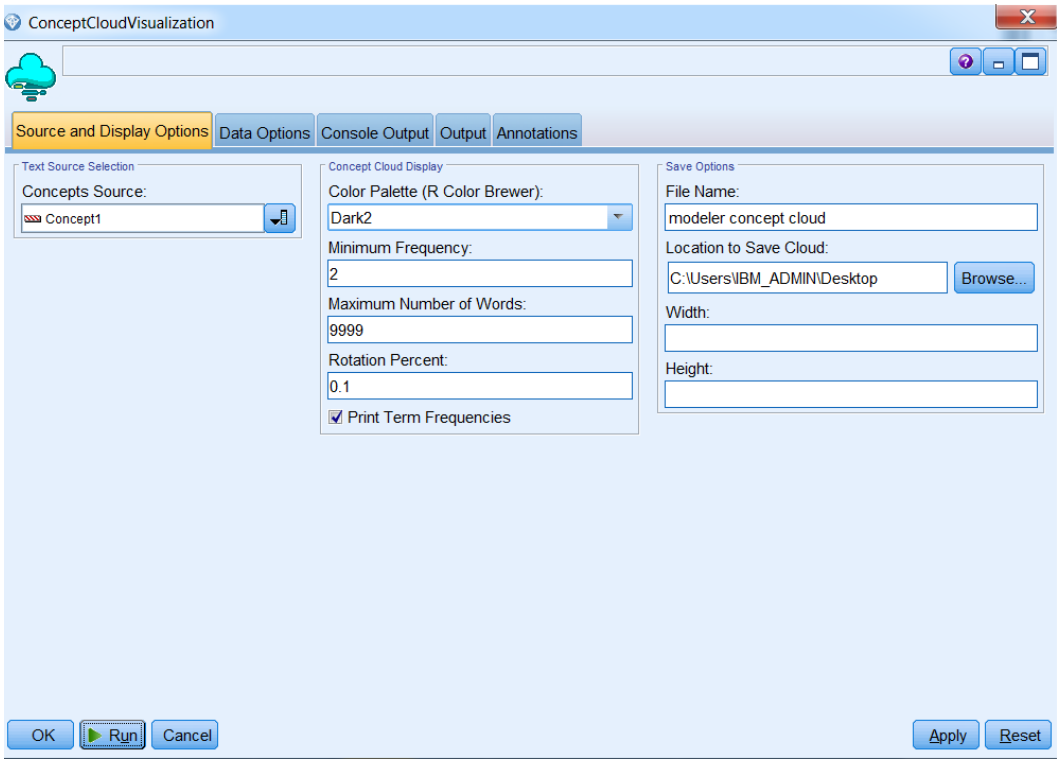
In preparation for visualization, we split the data into two streams (depending if OP received pizza or not) with the use of Select node, and then data is processed with Text Link Analysis (TLA) node to extract the concepts.



To explicitly favor compound concepts in extracted concepts list, TLA nodes were configured to use Basic Resources template and options to extract uniterms and nonlinguistic entities were deselected:

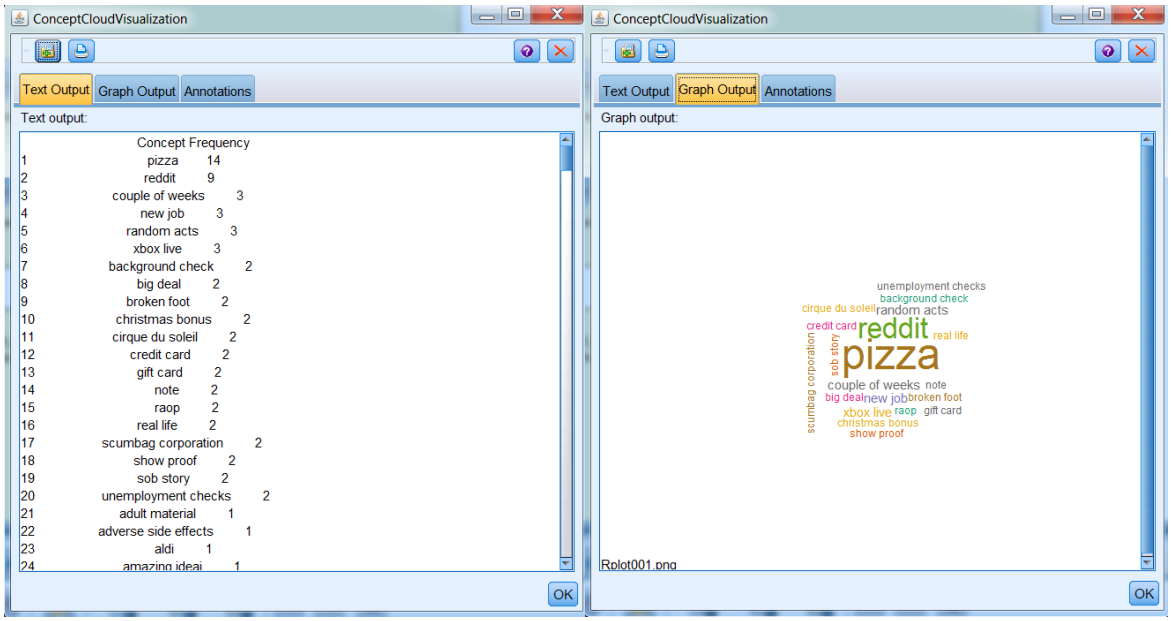


To visualize the output of the TLA nodes with the use of Concept Cloud Visualization nodes, connect it and configure to use concepts in Concept1 column:

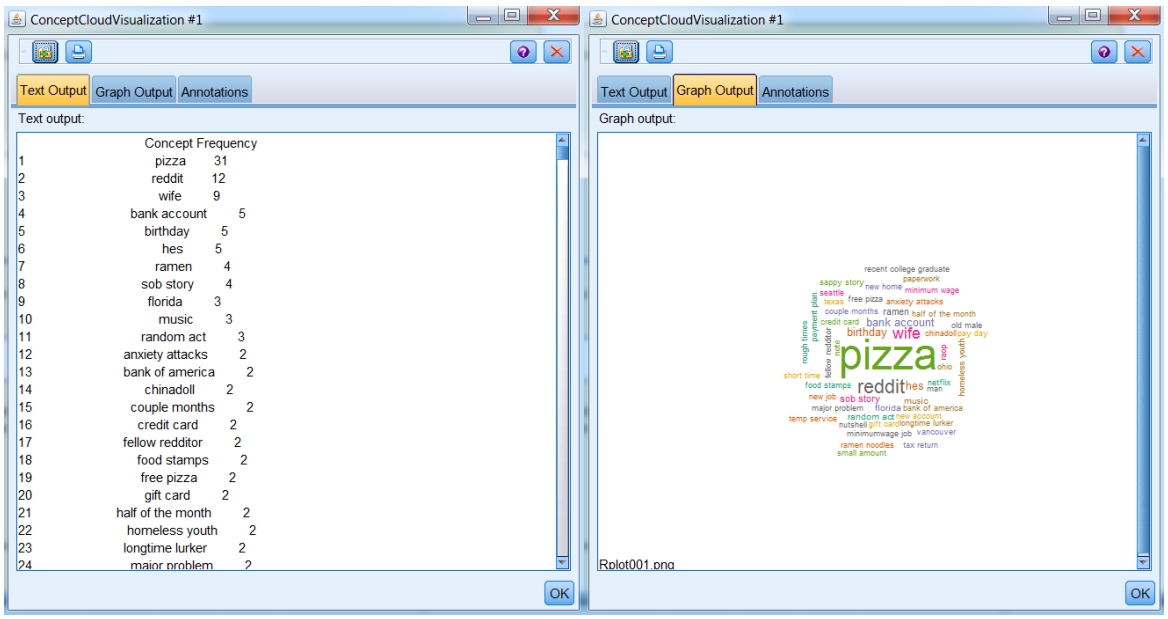


When run, this should produce following concept clouds:

1) For “pizza received”

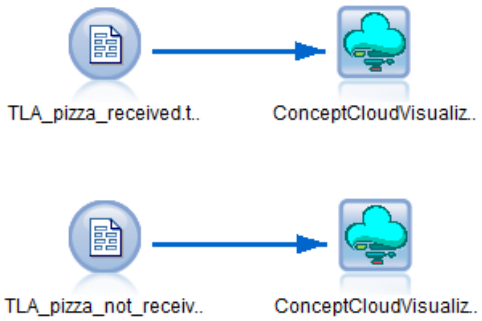


2) For “pizza not received”

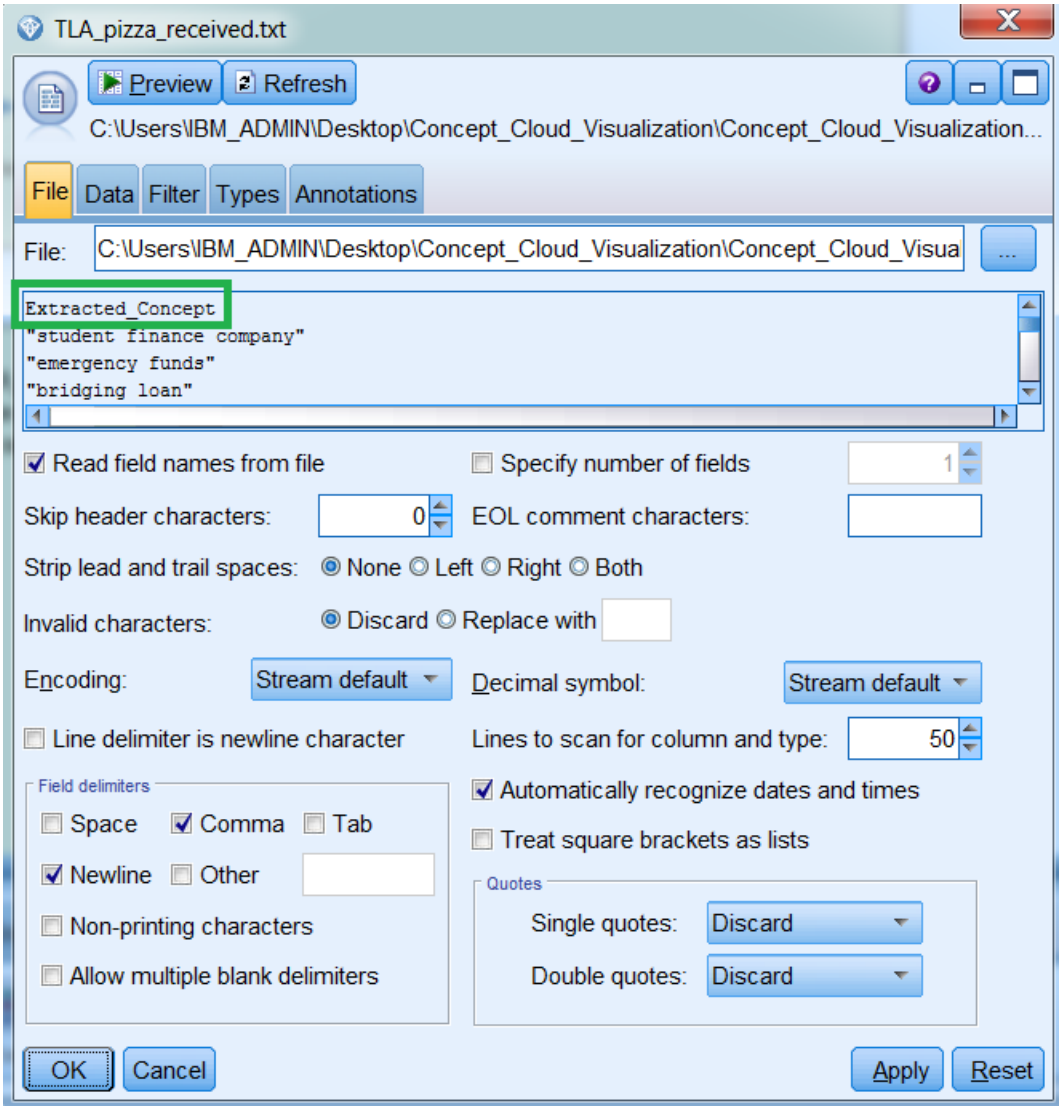


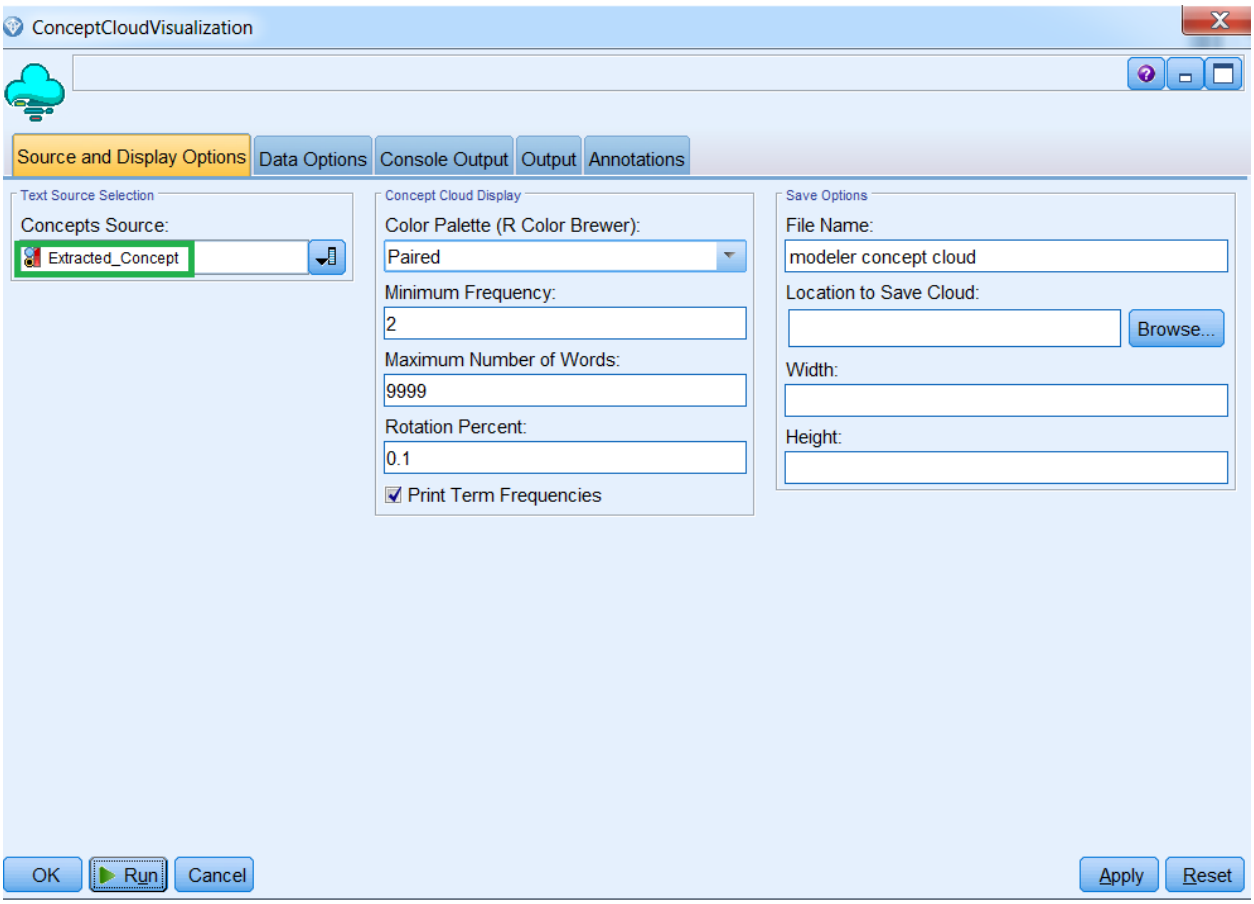
Text column output visualization (Example 2 in sample stream)

In this example we will use Concept Cloud Visualization node to create clouds for concepts that were captured in a text files.



By choosing Concepts Source to match the name of the column with our concepts, we ensure that data will be properly visualized:




ConceptCloudVisualization
Source and Display Options
Data Options
Console Output
Output
Annotations
Text Source Selection
Concepts Source:
Extracted_Concept
Concept Cloud Display
Color Palette (R Color Brewer):
Paired
Minimum Frequency:
2
Maximum Number of Words:
9999
Rotation Percent:
0.1
Print Term Frequencies
Save Options
File Name:
modeler concept cloud
Location to Save Cloud:
Browse...
Width:
Height:
OK
Run
Cancel
Apply
Reset