UniCom User Guide

Daniel Otal Rodríguez. March 2016

UniCom (stands for Universal Comparator) is a software which gives the user the ability to compare objects of any domain. The program can read massive datasets provided by the user and treat its information to compare multiple objects with a reference object chosen by the user. Results are presented through a radar chart and the percentage of similarity between both items, calculated by the shared area on the chart. In order to *UniCom* to operate needs a CSV file, where the dataset must be stored.

What is a CSV file?

Comma-separated values (CSV) files are a widely used practice of storing data in plain text. It is especially common for moving and converting data between spreadsheet programs that run on incompatible formats. There are many different specifications, but the most popular one (RFC 4180) indicates that each line in a CSV file is a data record and each record consists of one or more fields separated by commas. Normally the format is best used to represent sequences of records in which each record has an identical list of fields, as a single relation in a relational database. *UniCom's* CSV is not very different from the standard one, but it also has its own rules.

Formatting the UniCom CSV File

There are three types of data records in a CSV file formatted for *UniCom*: labels, reference object and compare object. Taking this into consideration, the requirements for the *Unicom* CSV File are:

- There must be at least one data record of every type and they must be stored in the correct order. The first line of the file encompasses the names of the labels for every data field. Second line corresponds to the reference object. The following lines of the file are reserved for the compare objects, as there is no limit of entries of comparable items. As a result, *UniCom's* CSV file should have a minimum of three lines.

- Label type of data records are alphanumerical. While reference and compare objects data fields are only numerical, where the dot is used as the decimal mark.

- There must be the same number of data fields for each data record. In case there is no value for a reference or compare object data record, the value of the data field will be considered zero.

- Field delimiters are semicolons (;), instead of commas. They must be located only between data fields without exception.

- A line break marks the end of a data record.

- The file must have the extension '.csv' in order for the program to recognize the file.

CSV Example

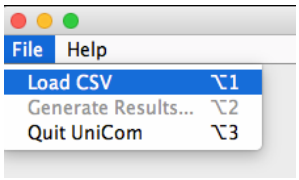
```
Year;Horse Power;Price;Weight
2004;200;12000;1200
2006;160;15000;900
2005;150;18000;950
```

File in plain text 'cars.csv'

The basic example above runs perfectly on UniCom if it is saved in a CSV file and loaded into the program. This dataset has four data records with four data fields. First line is the label data record, which gives a name for every column if we picture the file as a spreadsheet. Second line is the reference object and the following two lines are the compare objects. More entries can be added to the file if the user edits the file by hand with a text editor.

Steps to Load a CSV File

Go to 'File' and select 'Load CSV'. Use the file browser to select the CSV file. If you want to load another CSV file, the process can be performed again.



Generate Results…

Underneath the 'File' menu there is an option 'Generate Results…'. It generates a CSV file (following UniCom format) with the dataset which is currently loaded in the program, but adding new data fields for the calculated areas in the chart and the similarity results of every entry. This may help to process the data with external programs.