Quality Threshold clustering - user manual

Index

1.Introduction

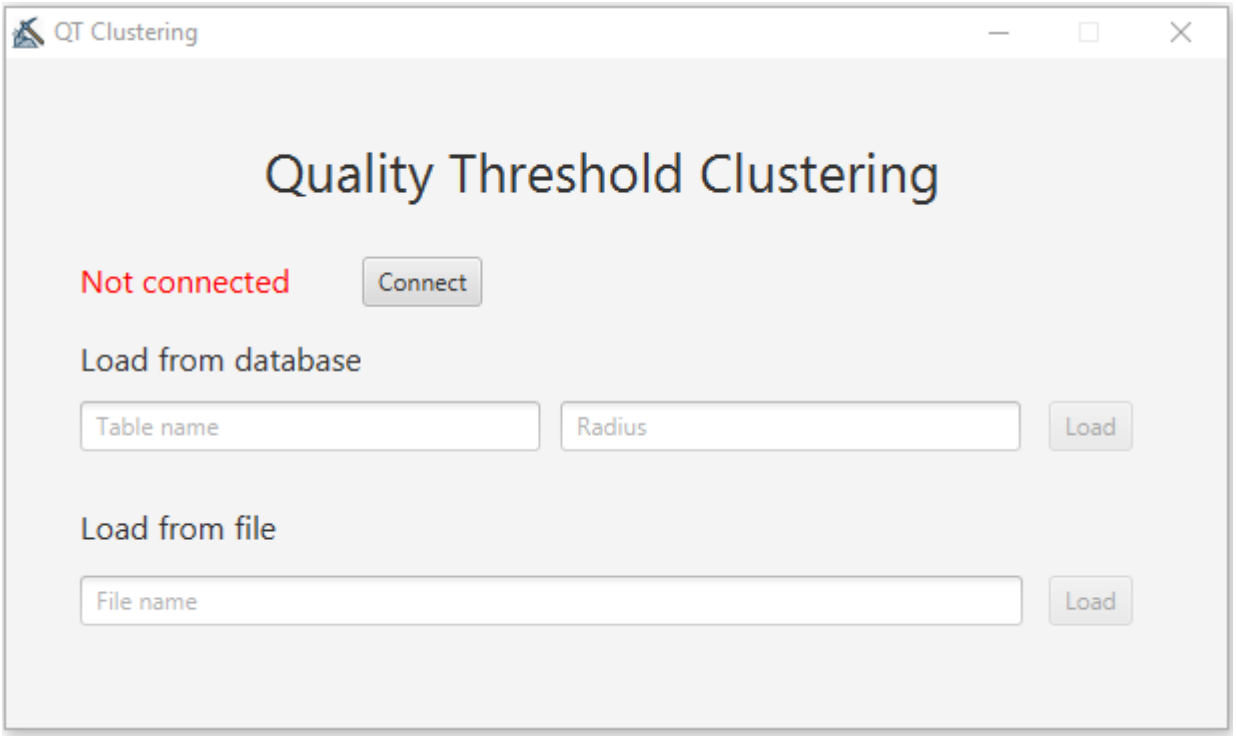
Quality Threshold clustering is an application which can make tables' clustering of a database through the quality threshold algorithm. The application needs a connection to a server, which connects to the database, execute the algorithm and returns the results, which can be viewed tabularly or graphically. The results can be saved on the server and following loaded.

2.Prerequisites

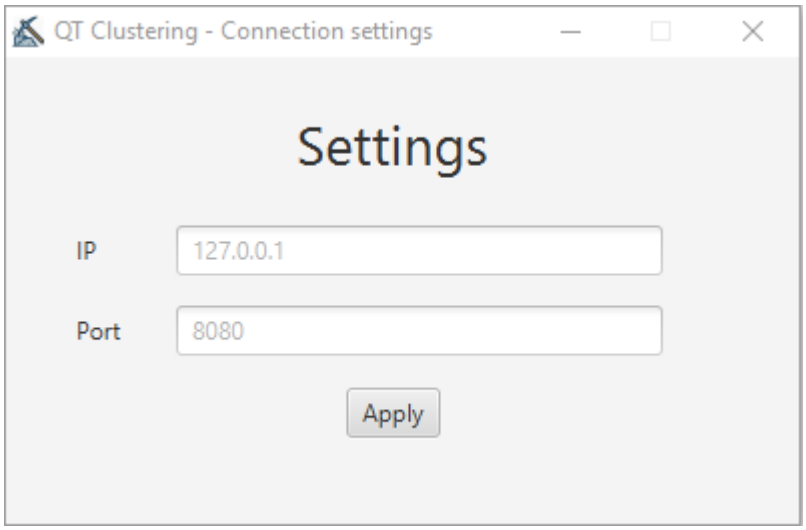
The application, being written in java can be executed on any operating system where is installed the Java Virtual Machine. It can be downloaded at: https://www.java.com/it/download/.

3.Server connection

When the software is executed, the following screen will be shown:

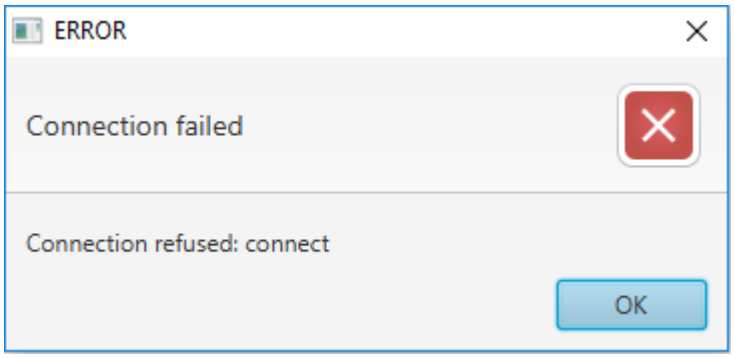


In this state can't be fullfilled any operations except of the connecton with the server by pressing the 'connect' button, then will be opened the following window:

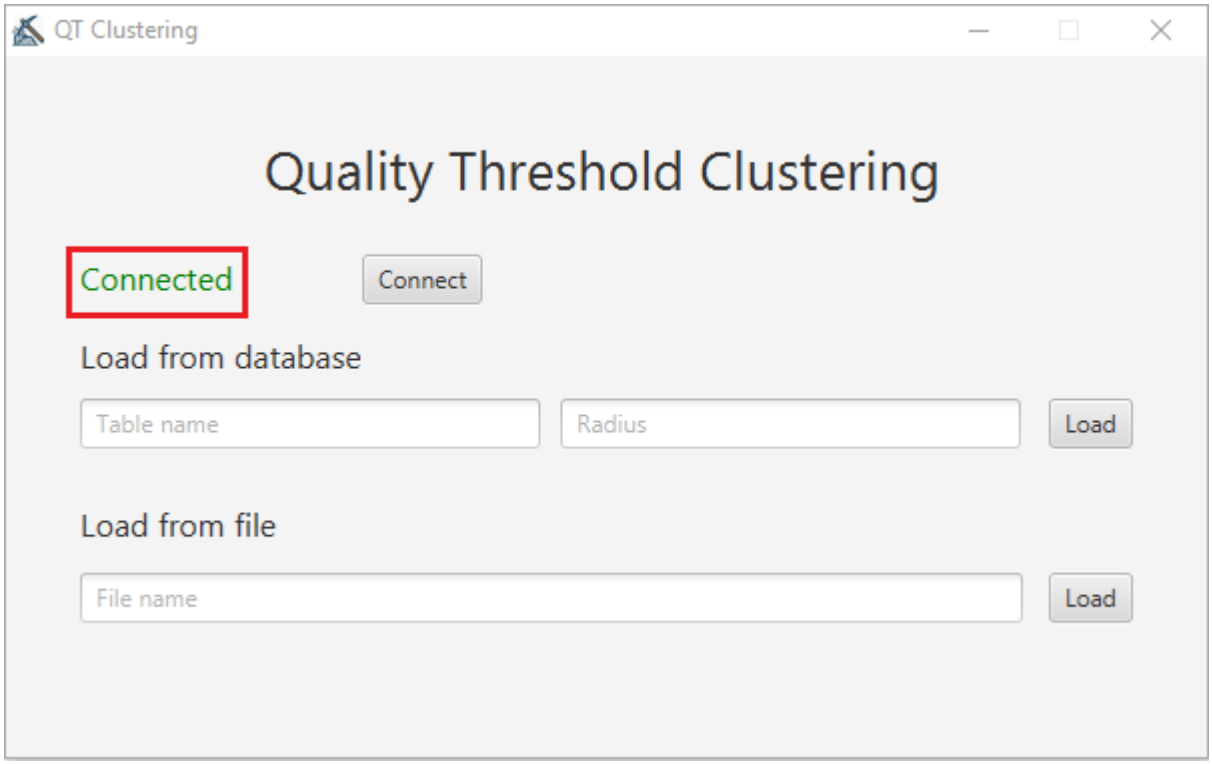


In it, has to be written the IP and the port of the server running the server application. By clicking ‘Apply’ the application will try to connect. If the IP and fields are empty, will be used the default prompt text (127.0.0.1:8080).

If the connection fails will be shown an error alert:

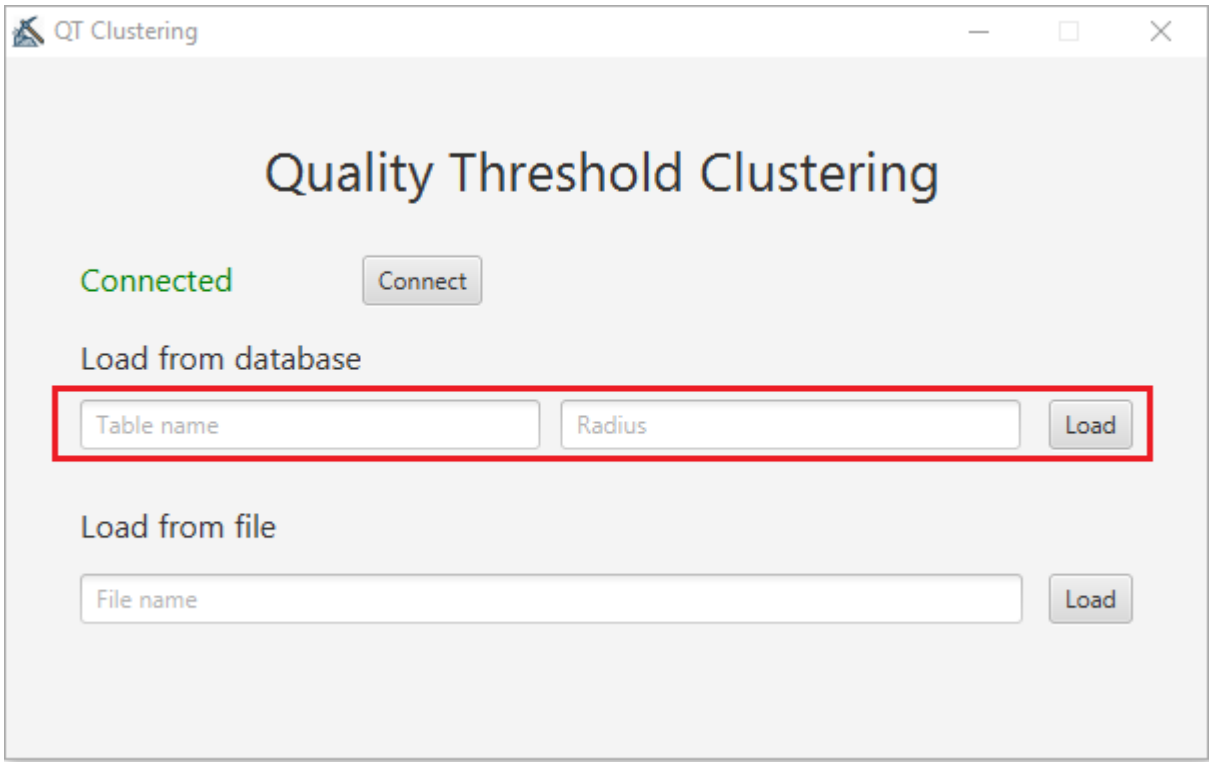


Otherwise the window will be closed and the “Not connected” will be changed in “connected” as shown.

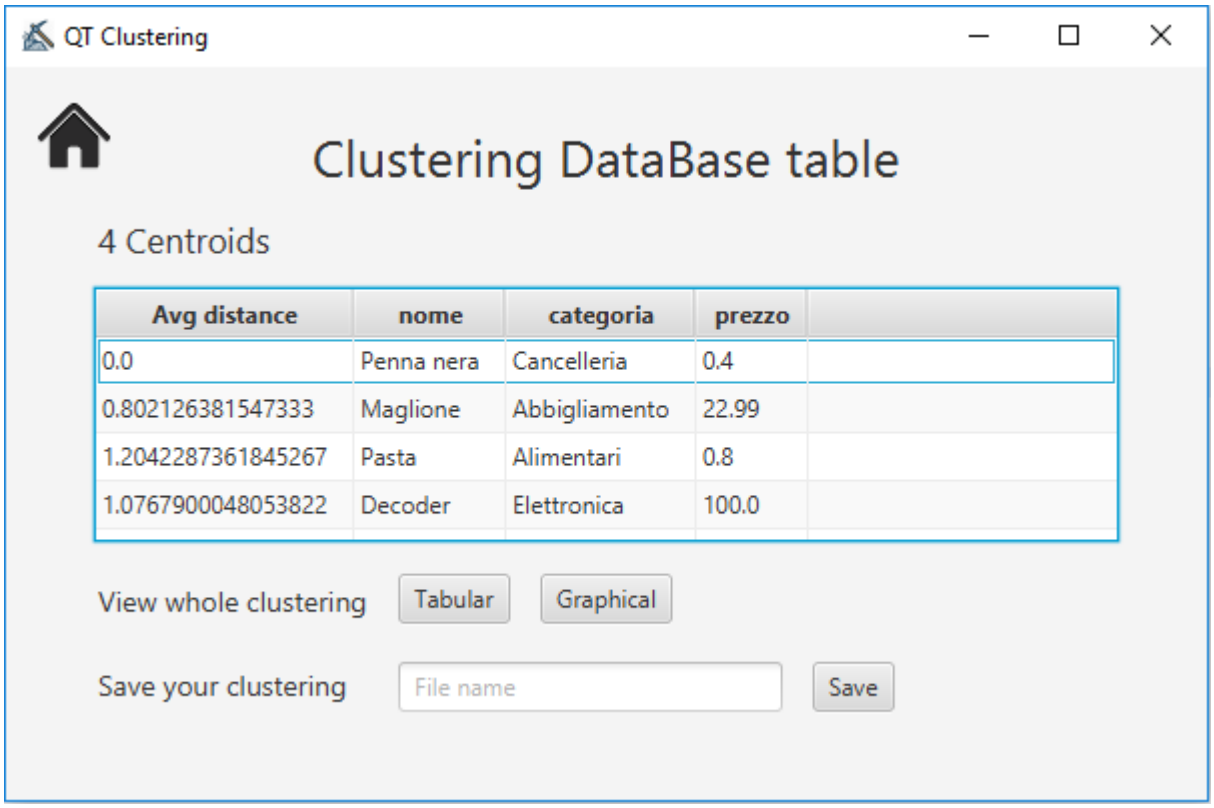


4.Clustering database table

To get the clustering from a DB table, it’s required the table name and the maximum radius of each cluster.



An example:



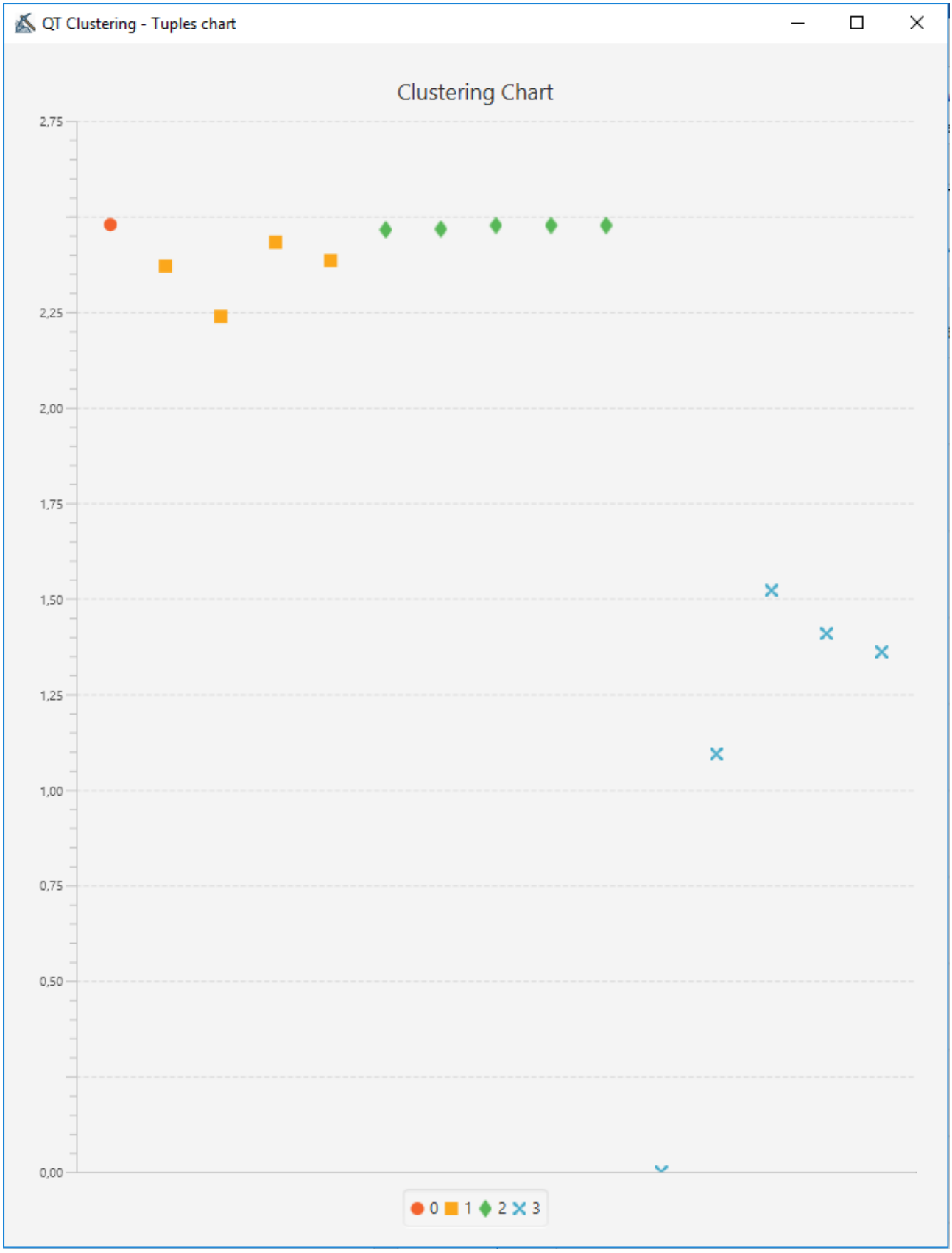
The first column represents the average distance from the clusters' tuples from the centroid.

The other columns are the fields of the centroids' tuples.

By clicking the **home button**, it will go back to the main window.

4.1 Graphical view

By clicking “Graphical” button, a new window containing the tuples distances chart will be displayed



Each dot, is a tuple, and it's Y axis position, is the distance between itself and the centroid of the biggest cluster, so the latter is in the 0.0 position. The X axis as the only objective to prevent the overlapping of the dots.

The tuples of each cluster are displayed in different colours and shapes.

4.2 Tabular view

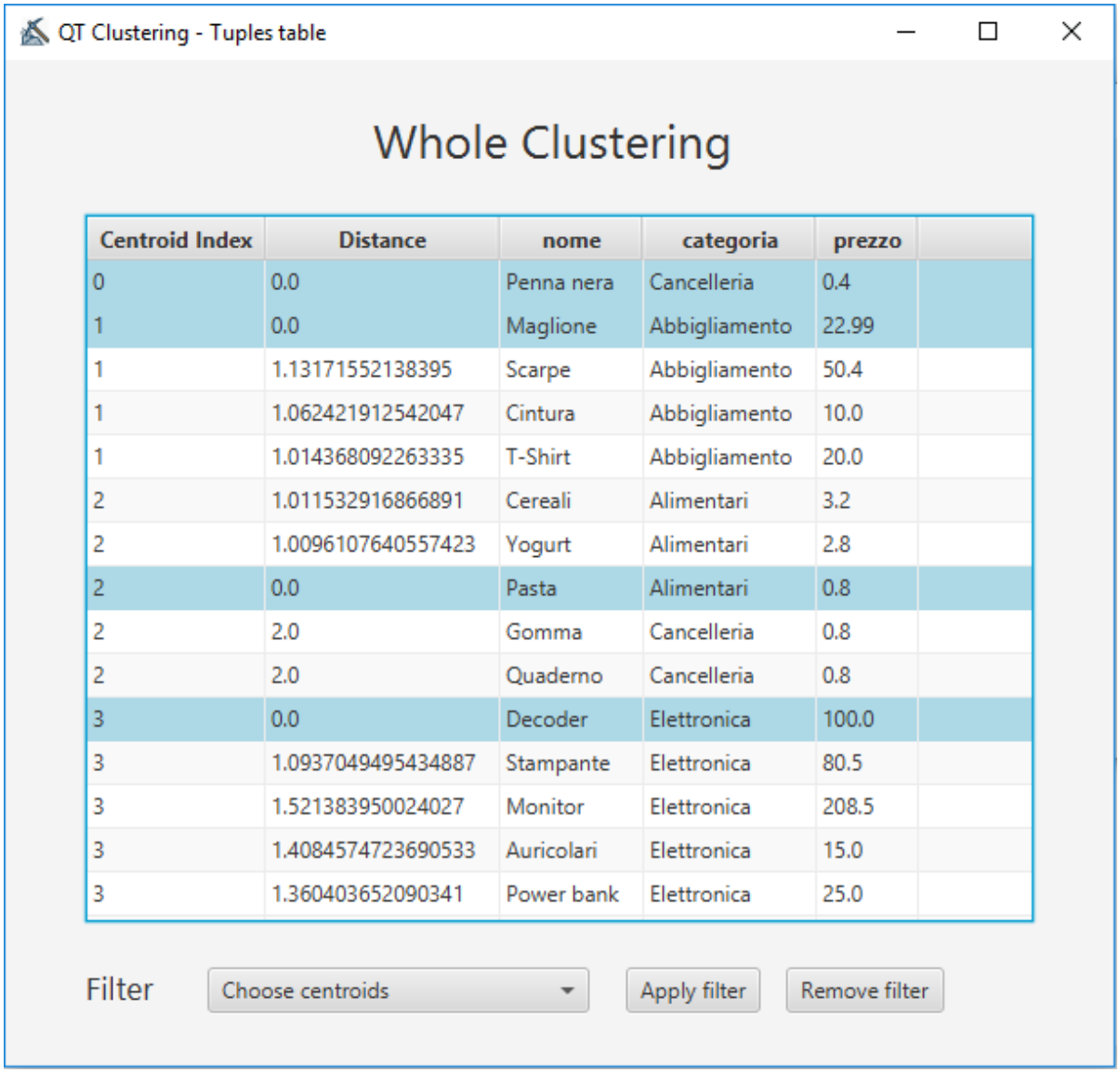
By clicking "Tabular" button, a new window containing whole tuples in a table will be displayed.

The first column represents the cluster which the tuple belongs.

The second column is the distance between tuple and its centroid.

The others are the values of the tuples.

The centroids are coloured in sky blue.



Centroid Index	Distance	nome	categoria	prezzo
0	0.0	Penna nera	Cancelleria	0.4
1	0.0	Maglione	Abbigliamento	22.99
1	1.13171552138395	Scarpe	Abbigliamento	50.4
1	1.062421912542047	Cintura	Abbigliamento	10.0
1	1.014368092263335	T-Shirt	Abbigliamento	20.0
2	1.011532916866891	Cereali	Alimentari	3.2
2	1.0096107640557423	Yogurt	Alimentari	2.8
2	0.0	Pasta	Alimentari	0.8
2	2.0	Gomma	Cancelleria	0.8
2	2.0	Quaderno	Cancelleria	0.8
3	0.0	Decoder	Elettronica	100.0
3	1.0937049495434887	Stampante	Elettronica	80.5
3	1.521383950024027	Monitor	Elettronica	208.5
3	1.4084574723690533	Auricolari	Elettronica	15.0
3	1.360403652090341	Power bank	Elettronica	25.0

Filter

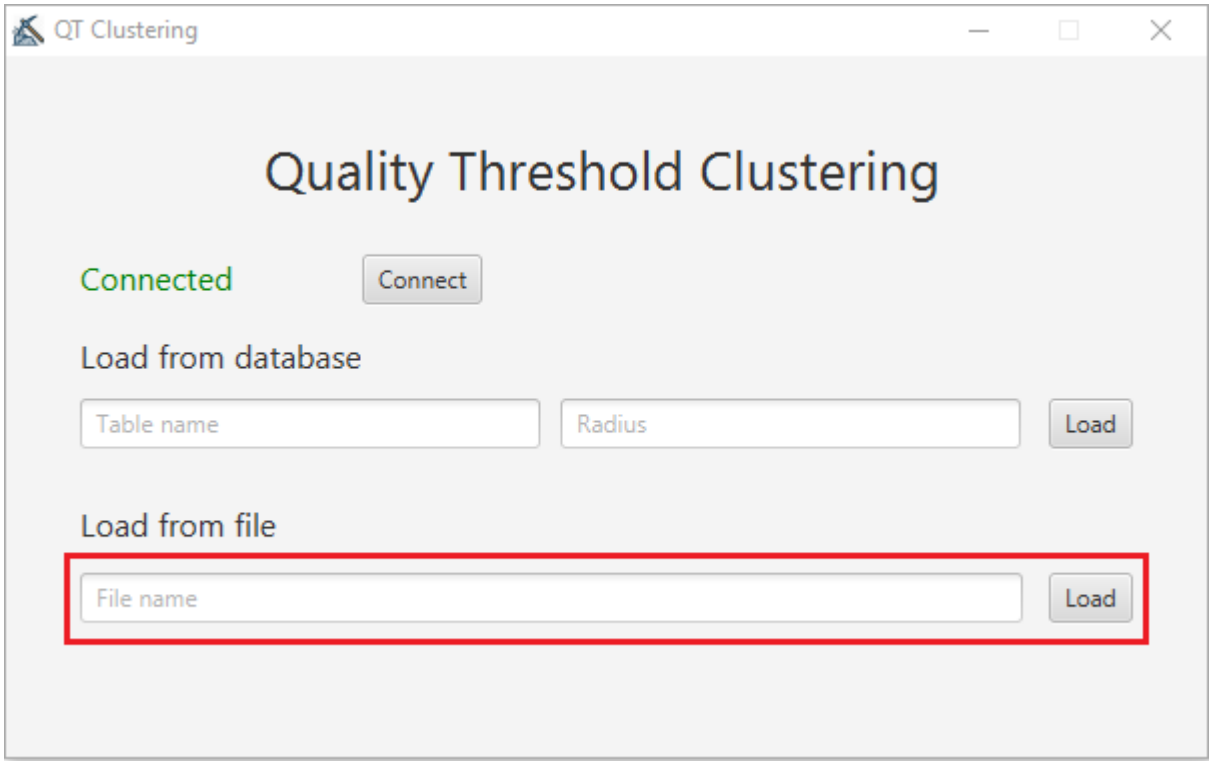
One or more cluster can be filter and the others not shown. Multiple select needs to click multiple times on dropdown menu.

4.3 Saving centroids

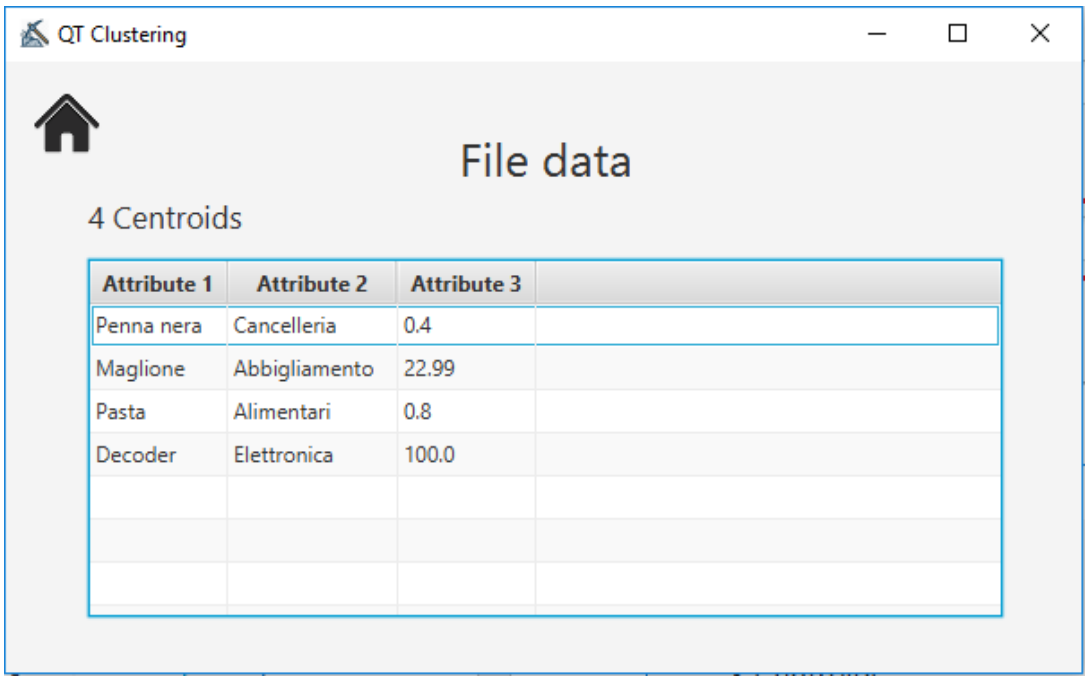
The centroids can be saved on a file on the server, typing is name and clicking “save”. An alert notify the user that the save is successfully completed or not.

5. Loading from file

After saving the centroid it can be loaded at any time, typing in the relative field the name of the saved file and clicking load.



Example:



By clicking the **home button**, it will go back to the main window.