

Quality Threshold clustering - user manual

Index

Introduction	2
Prerequisites.....	2
1.Server connection.....	2
2.Clustering database table.....	4
2.1 Graphical view	6
2.2 Tabular view.....	7
2.3 Saving centroids	8
3. Loading from file.....	8

1.Introduction

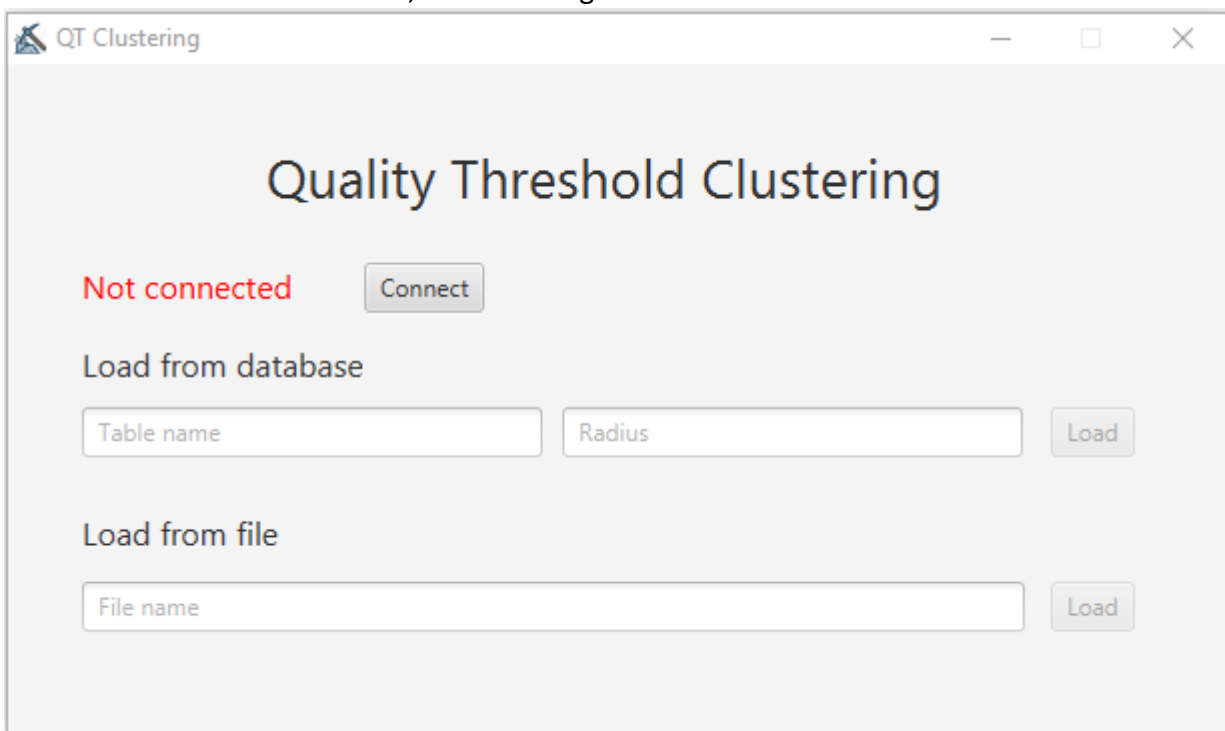
Quality Threshold clustering is an application which can make tables' clustering of a database through the quality threshold algorithm. The application needs a connection to a server, which connects to the database, execute the algorithm and returns the results, which can be viewed tabularly or graphically.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



QT Clustering

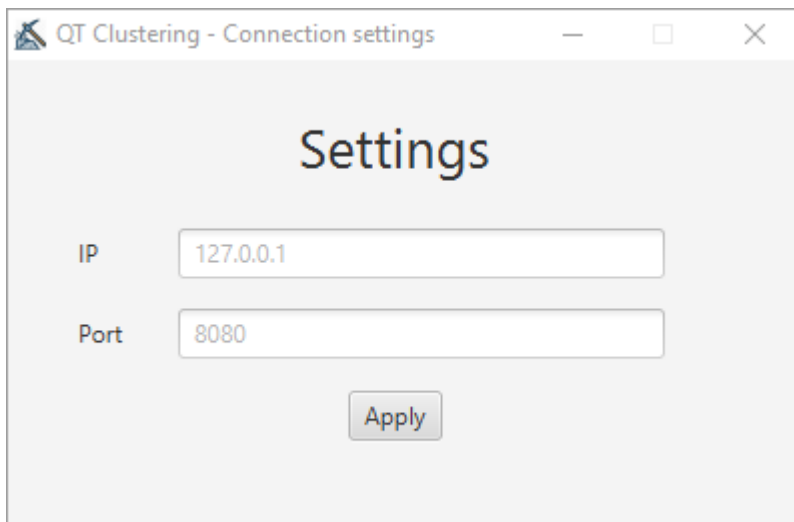
Quality Threshold Clustering

Not connected

Load from database

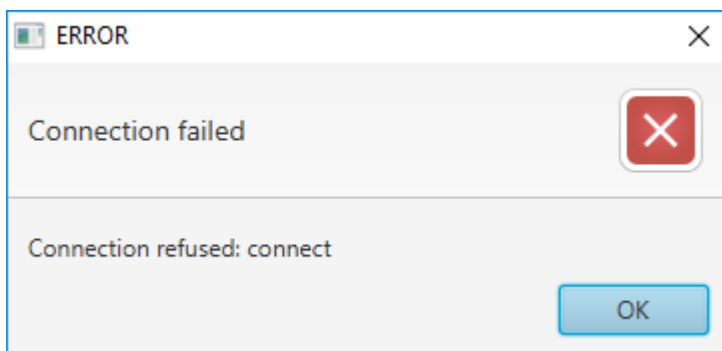
Load from file

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

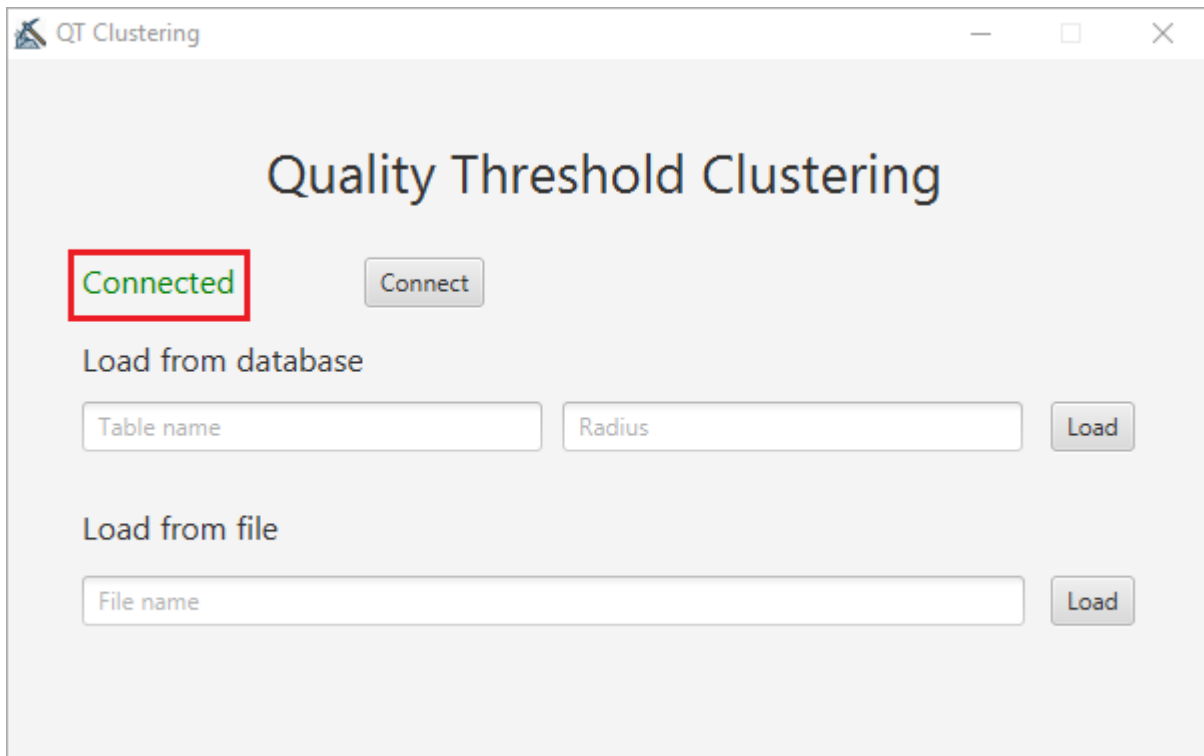


In it, has to be written the IP and the port of the server running the server application. By clicking 'Apply' the application will try to connect. If the IP and fields are empty, will be used the default prompt text (127.0.0.1:8080).

If the connection fails will be shown an error alert:



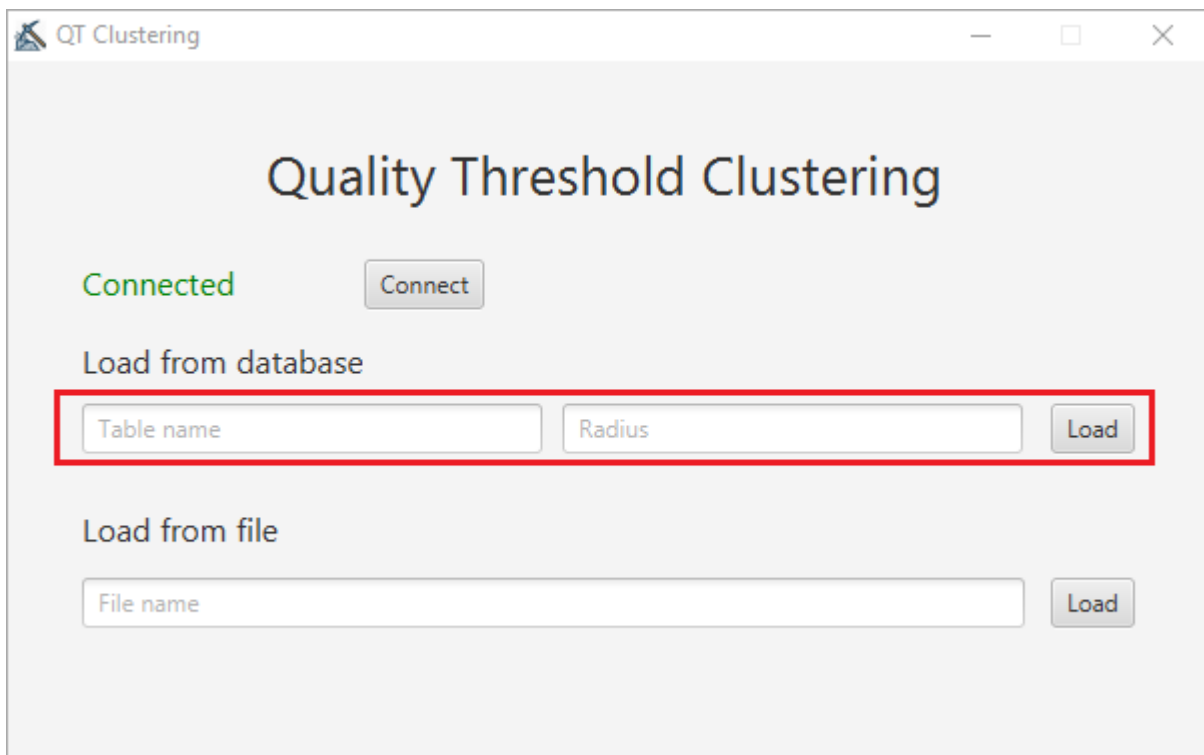
Otherwise the window will be closed and the “Not connected” will be changed in “connected” as shown.



The screenshot shows a window titled "QT Clustering" with standard window controls (minimize, maximize, close). The main heading is "Quality Threshold Clustering". Below this, the word "Connected" is displayed in green text and is highlighted with a red rectangular box. To the right of "Connected" is a "Connect" button. Under the heading "Load from database", there are two input fields: "Table name" and "Radius", followed by a "Load" button. Below this section, under the heading "Load from file", there is a "File name" input field followed by a "Load" button.

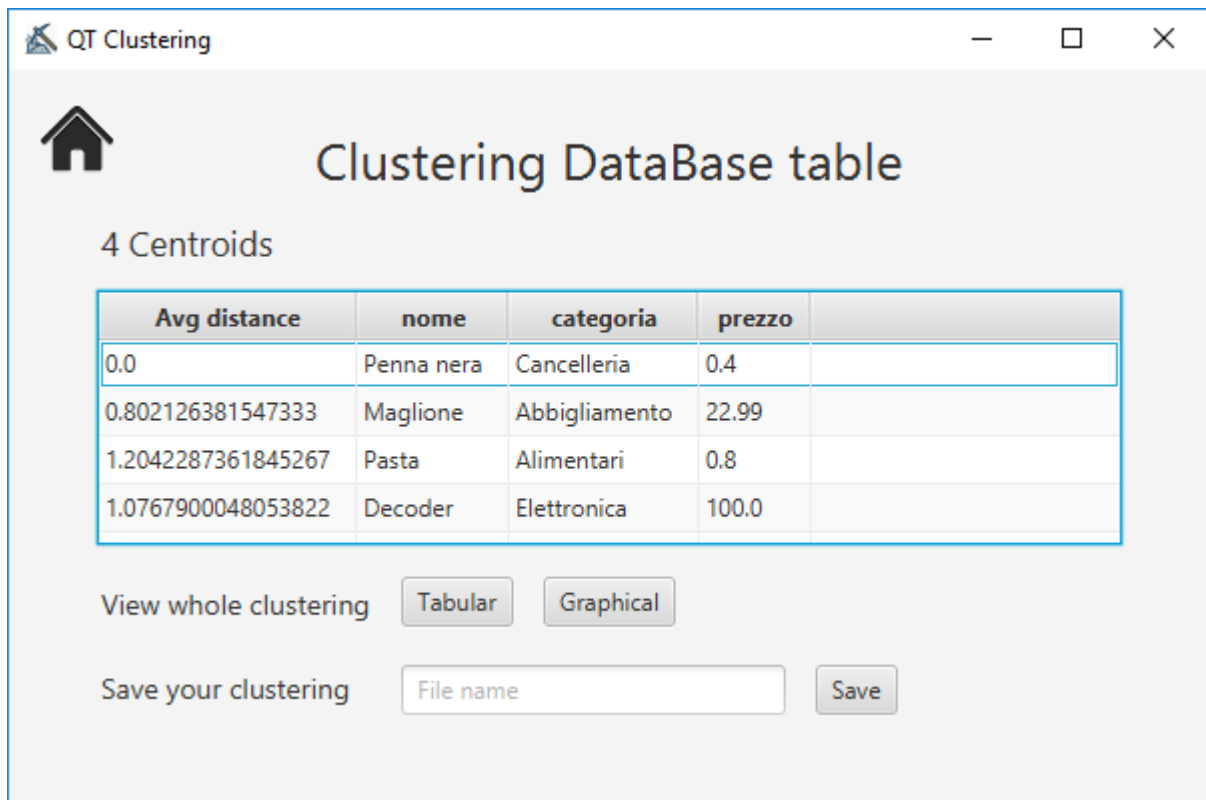
4. Clustering database table

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



This screenshot is similar to the one above, showing the "QT Clustering" window. The "Connected" status is still present. However, a red rectangular box highlights the "Load from database" section, which includes the "Table name" and "Radius" input fields and the "Load" button. The "Load from file" section remains visible below.

An example:



The screenshot shows a window titled "QT Clustering" with a home icon in the top left. The main heading is "Clustering DataBase table". Below it, it says "4 Centroids". A table displays the following data:

Avg distance	nome	categoria	prezzo	
0.0	Penna nera	Cancelleria	0.4	
0.802126381547333	Maglione	Abbigliamento	22.99	
1.2042287361845267	Pasta	Alimentari	0.8	
1.0767900048053822	Decoder	Elettronica	100.0	

Below the table, there are two buttons: "Tabular" (selected) and "Graphical". At the bottom, there is a "Save your clustering" section with a "File name" input field and a "Save" button.

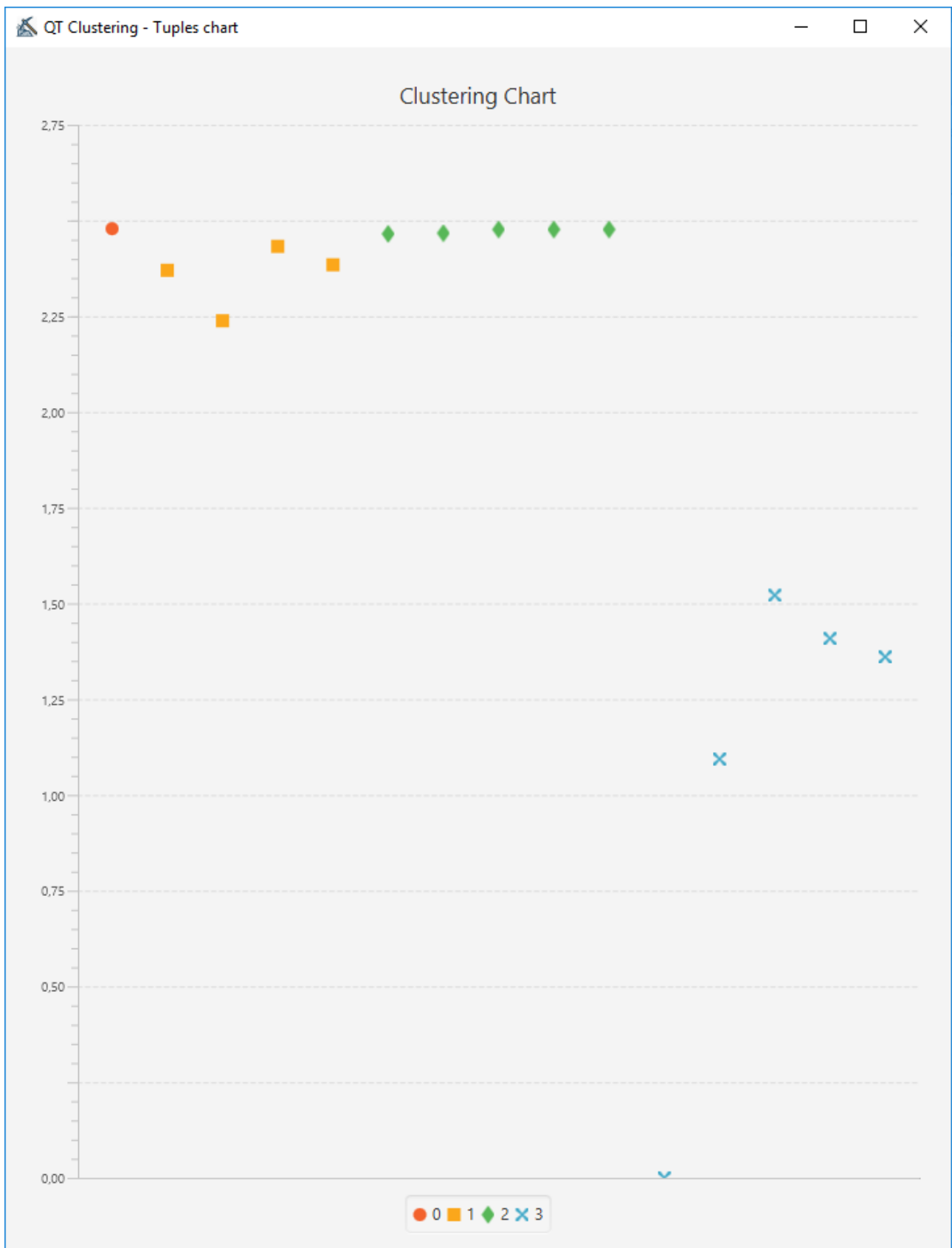
The **first column** represents the average distance from the clusters' tuples from the centroid.

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

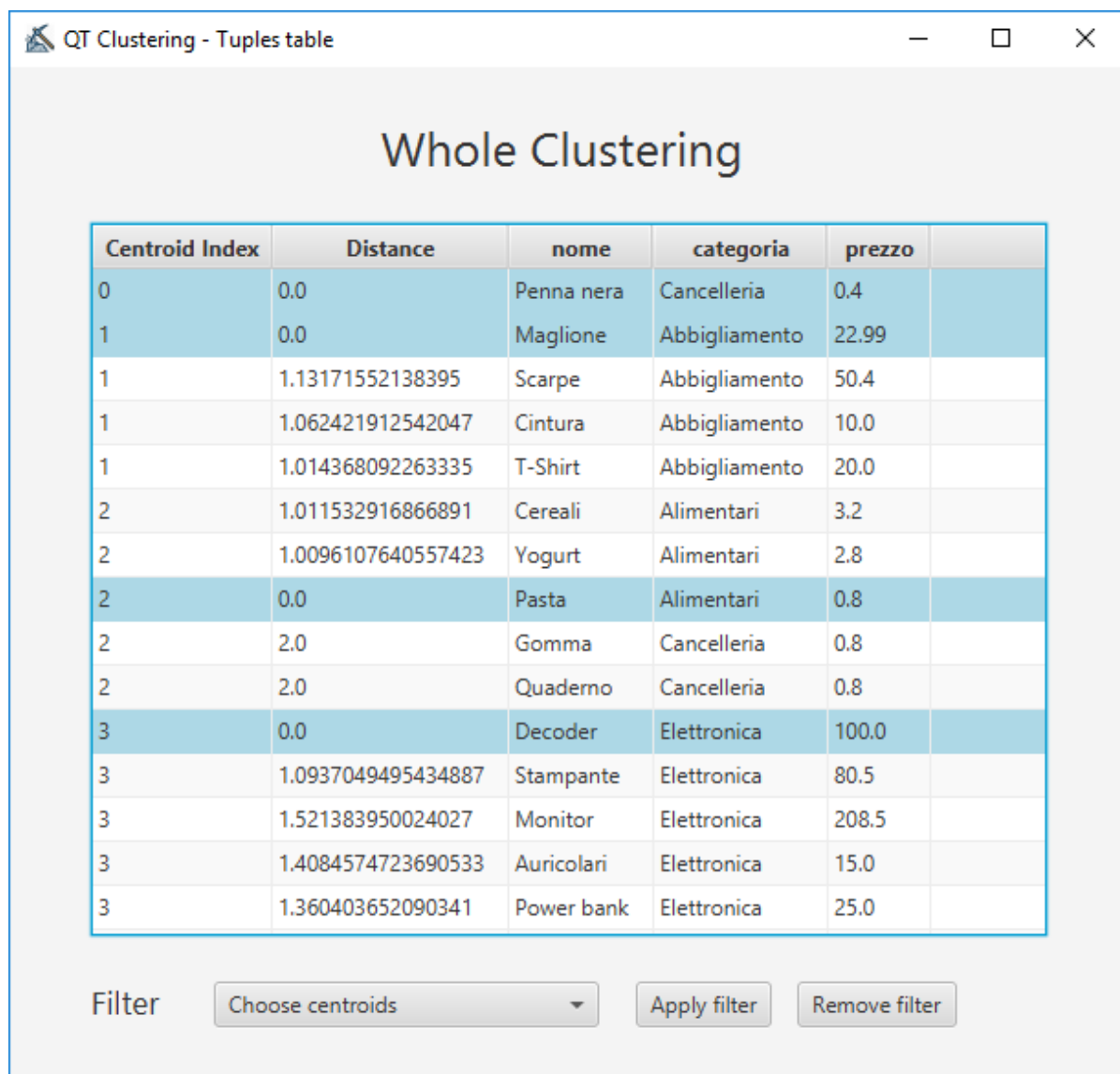
By clicking "Tabular" button, a new window containing whole tuples in a table will be displayed.

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.



The screenshot shows a window titled "QT Clustering - Tuples table" with a standard macOS-style title bar (minimize, maximize, close buttons). The main content area is titled "Whole Clustering" and contains a table with 6 columns: "Centroid Index", "Distance", "nome", "categoria", "prezzo", and an empty column. The table lists 20 tuples, each belonging to a cluster (0, 1, 2, or 3). Centroids are highlighted in light blue. Below the table, there is a "Filter" section with a dropdown menu set to "Choose centroids", and two buttons: "Apply filter" and "Remove filter".

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

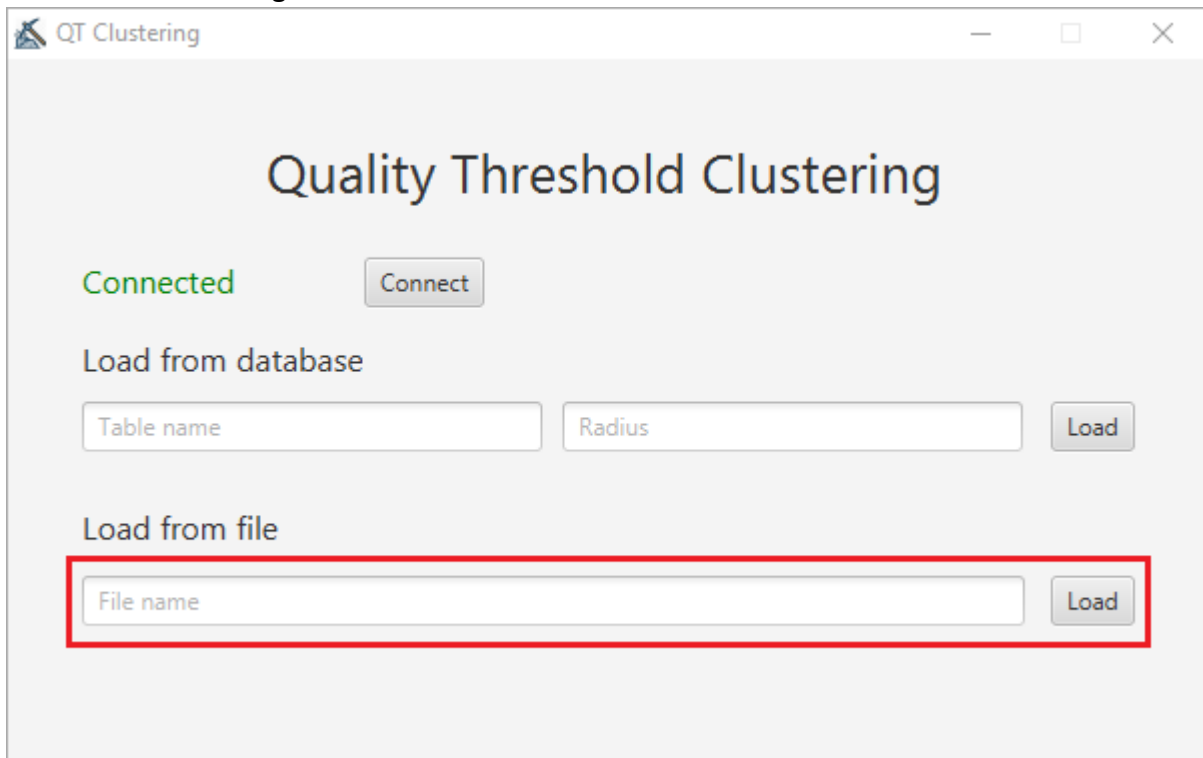
One or more cluster can be filter and the others not shown. Multiple select needs to click multiple times on dropdown menu.

4.3 Saving centroids

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

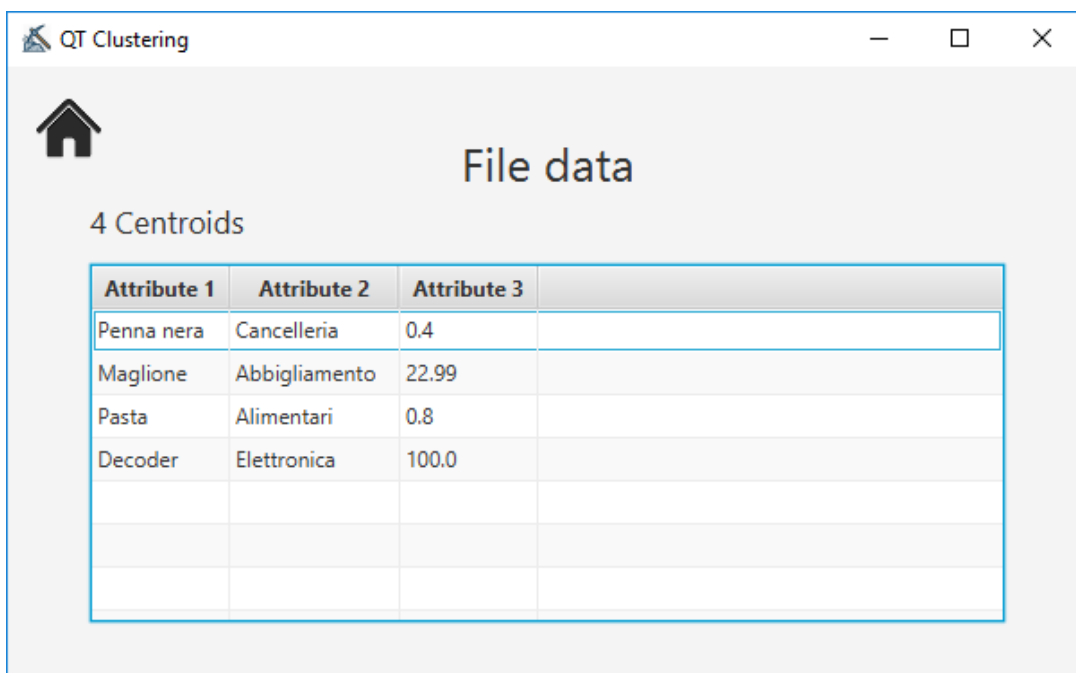
5. Loading from file

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.



The screenshot shows the 'QT Clustering' application window. The title bar says 'QT Clustering'. The main content area has a title 'Quality Threshold Clustering'. Below the title, there is a green status indicator 'Connected' and a 'Connect' button. There are two sections for loading data: 'Load from database' and 'Load from file'. The 'Load from database' section has a 'Table name' input field, a 'Radius' input field, and a 'Load' button. The 'Load from file' section has a 'File name' input field and a 'Load' button. The 'File name' input field and its 'Load' button are highlighted with a red rectangle.

Example:



The screenshot shows the 'QT Clustering' application window. The title bar says 'QT Clustering'. The main content area has a home icon (house) and a title 'File data'. Below the title, it says '4 Centroids'. There is a table with 4 columns: 'Attribute 1', 'Attribute 2', 'Attribute 3', and an empty column. The table contains 4 rows of data:

Attribute 1	Attribute 2	Attribute 3	
Penna nera	Cancelleria	0.4	
Maglione	Abbigliamento	22.99	
Pasta	Alimentari	0.8	
Decoder	Elettronica	100.0	

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