



Khiops Visualization Guide 10.1



■ Khiops

- Optimal data preparation based on discretization and value grouping
- Scoring models for classification and regression
- Correlation analysis between pairs of variables
- Automatic variable construction for multi-table relational mining



■ Khiops Visualization

- Analysis of Khiops results using an interactive visualization tool



■ Khiops Coclustering

- Correlation analysis of two or more variables using a hierarchical coclustering model
- Prediction models for cluster assignment



■ Khiops Covisualization

- Exploratory analysis of Khiops Coclustering results using an interactive visualization tool

This guide is about the Khiops Visualization component.

Abstract

Khiops Visualization is a visualization plug-in of the data preparation and scoring tool Khiops. Khiops Visualization allows visualizing all analysis results in an intuitive way, providing a quick and easy interpretation of the discovered patterns.

Summary

1. Presentation	3
2. Installation	3
3. First step	3
4. Functionalities.....	4
4.1. Screenshot	4
4.2. Table display	5
4.3. Zoom	6
4.4 Tree preparation.....	7
5. Technical limits	8

1. Presentation

Khiops Visualization allows to visualize and navigate in an intuitive way in the result reports of the Khiops software. This document presents only general features. It is strongly recommended to read the Khiops tutorial which presents all types of analysis and the associated visualization.

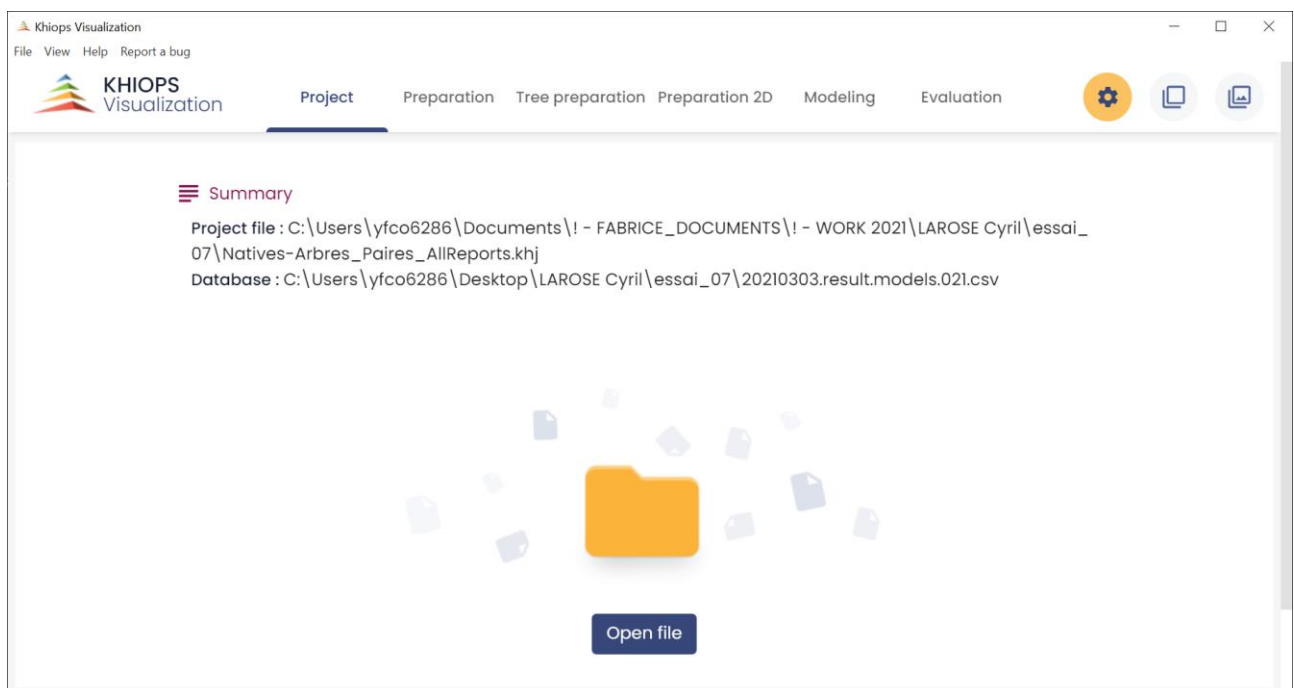
See the Khiops guide for more information on Khiops.

2. Installation

The installation of Khiops Visualization is supported by the installation of the **Khiops** tool available on the website www.khiops.com.

3. First step

The entry point of Khiops Visualization is the khv file generated by Khiops at the end of the analysis. A double-click on this file opens Khiops Visualization.



Khiops Visualization is composed of several tabbed panes. The first one is the project pane : it presents the report file and database locations.

Depending on the analysis type, the panes and their contents are not the same. Khiops Visualization can be composed with 5 panes:

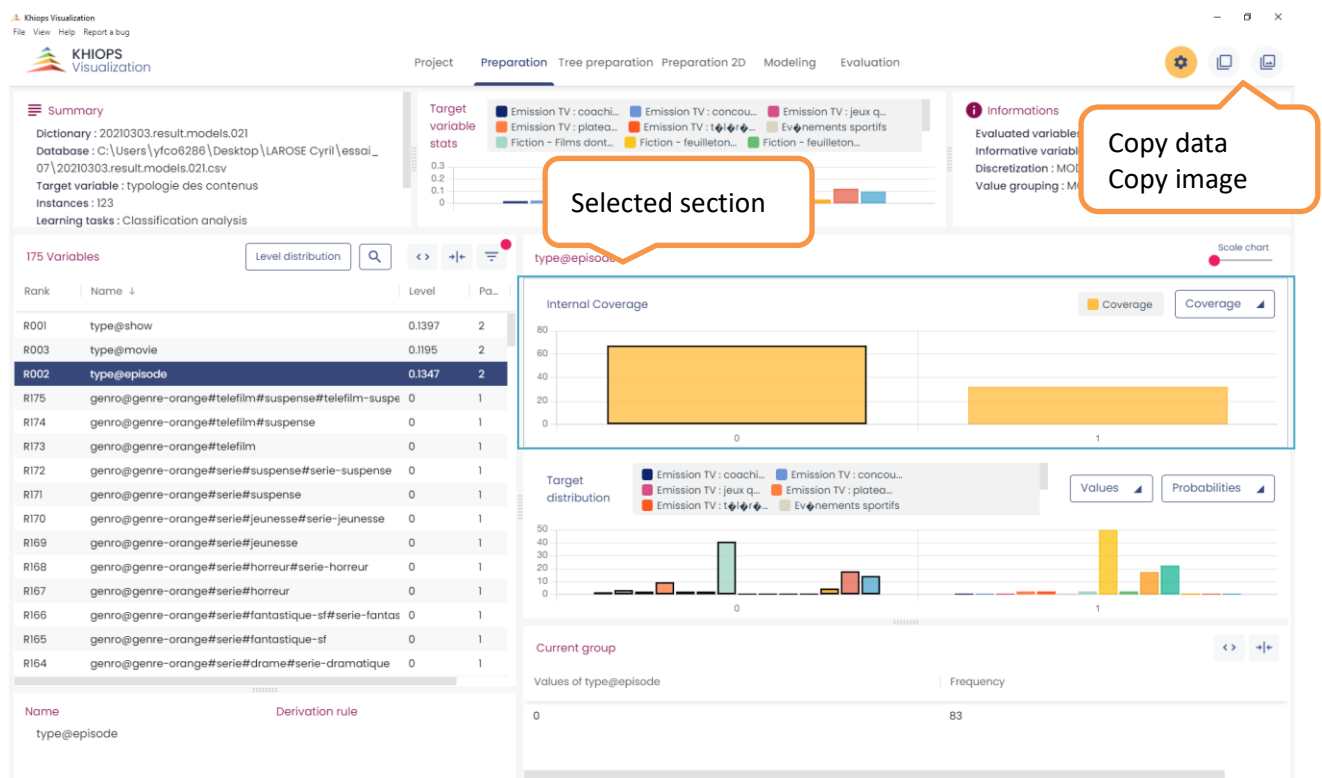
- Preparation: displays the Preparation report.
- Tree preparation : displays the preparation report for tree variables
- Preparation 2D: displays the 2D preparation report.
- Modeling: displays the modeling report.
- Evaluation: displays on one pane the test, train and evaluation reports.


The description of all these reports can be found in section 3.3. of Khiops Guide.


4. Functionalities

4.1. Screenshot

A useful feature available on all window sections is the screenshot. You can copy to clipboard a selected section by a simple mouse click on one of the top buttons. A section is selected when it is surrounded by blue lines.




Using  (or [Ctrl-C]) you get a picture in bitmap format, that you can import in any image editing tool.

Using  (or [Ctrl-D]) you get the raw data (a table in CSV format), that you can import in any text editing tool.

4.2. Table display



Clicking  opens a list of columns you can hide or display.
The red dot indicates that at least one column is hidden

175 Variables

Level distribution

type@episode

Rank	Name	Level	Pa...	Min
R001	type@show	0.1397	2	0
R002	type@episode	0.1347	2	0
R003	type@movie	0.1195	2	0
R004	categ@categories#emissions	0.1139	2	0
R005	genro@genre-orange#film	0.0761	2	0
R006	genro@genre-orange#court-metrage	0.0598	2	0
R007	categ@categories#fiction	0.0580	2	0
R008	genro@genre-orange#serie	0.0429	2	0
R009	categ@categories#emissions#divertissement	0.0416	2	0
R010	genro@genre-orange#emission	0.0409	2	0
R011	categ@categories#emissions#magazine	0.0393	2	0
R012	categ@categories#programme-d-information	0.0367	2	0
R013	genro@genre-orange#magazine	0.0275	2	0
R014	genro@genre-orange#court-metrage#drame	0.0134	2	0
R015	genro@genre-orange#court-metrage#drame#drame	0.0134	2	0

Name Derivation rule

- ☒ Rank
- ☒ Name
- ☒ Level
- ☒ Parts
- ☐ Values
- ☐ Type
- ☐ Mode
- ☐ Mode Coverage
- ☒ Min
- ☐ Max
- ☐ Mean
- ☐ Std Dev
- ☐ Missing Number
- ☐ Derivation Rule



allow to automatically redimension the columns of the table.



opens a search box.

1032 Variables

Level distribution

Rank | Name | Level ↓ | Pa... | Min

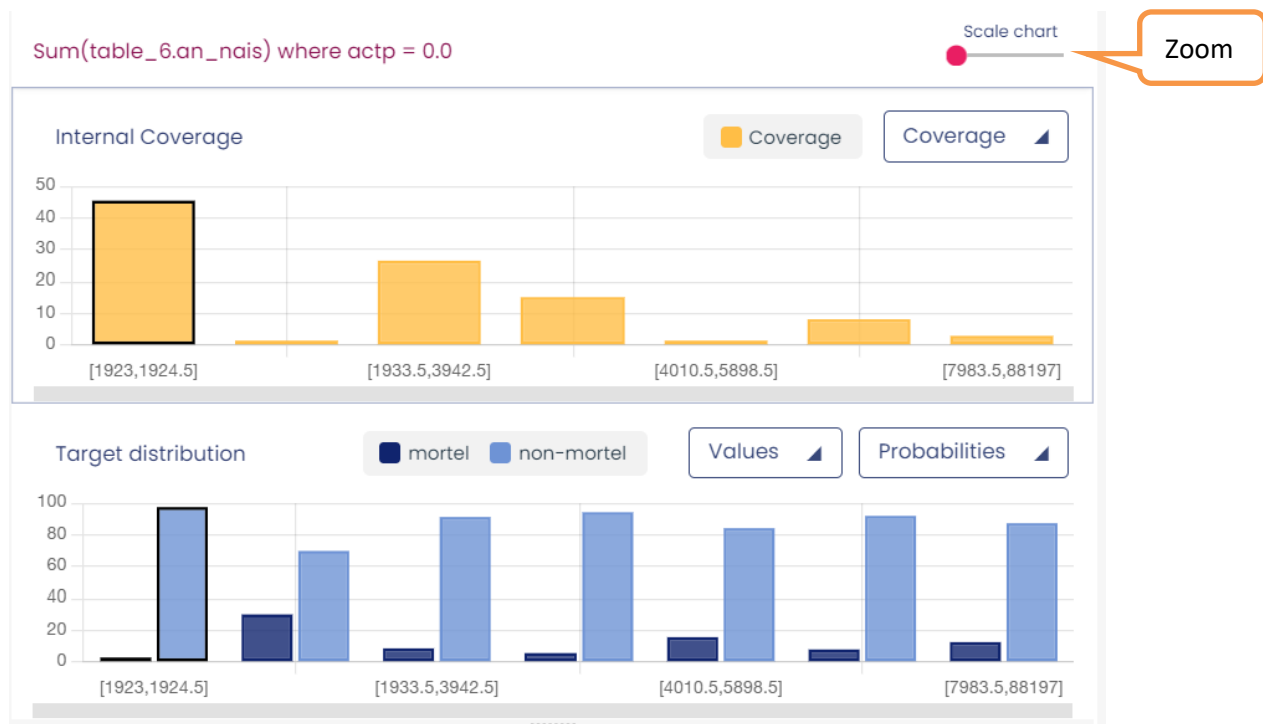
R0001	agg	0.0629	2	0
R0002	voie	0.0623	2	0
R0003	dep	0.0566	4	0
R0004	catr	0.0518	4	0
R0005	prl	0.0498	2	0
R0006	pr	0.0486	4	0
R0007	Mean(table_6.an_nais) where loep <> 0.0	0.0473	4	1919
R0008	Mode(table_6.secu) where actp = 0.0	0.0469	4	0
R0009	Mode(table_6.secu) where etatp = 0.0	0.0468	4	0
R0010	Median(table_6.an_nais) where loep <> 0.0	0.0465	4	1919
R0011	Mode(table_6.secu) where loep = 0.0	0.0463	4	0
R0012	Sum(table_6.an_nais) where loep <> 0.0	0.0458	3	1919
R0013	Max(table_6.an_nais) where loep <> 0.0	0.0455	5	1919
R0014	Mode(table_7.catv) where obs = 0.0	0.0452	5	0

1 to 500 of 1,032 | Page 1 of 3

Long tables are split into pages of at most 500 lines.

4.3. Zoom

The sliding bars allow zooming on the chart.



4.4 Tree preparation

This pane allows to explore each tree in an interactive manner. A typical screenshot is given below.

The upper left (orange rectangle) is like the other preparation panes and allows to select the tree variable and visualize the groups of leaves (leaves are considered as modalities of the tree variable and grouped accordingly) and the conditional class histograms.

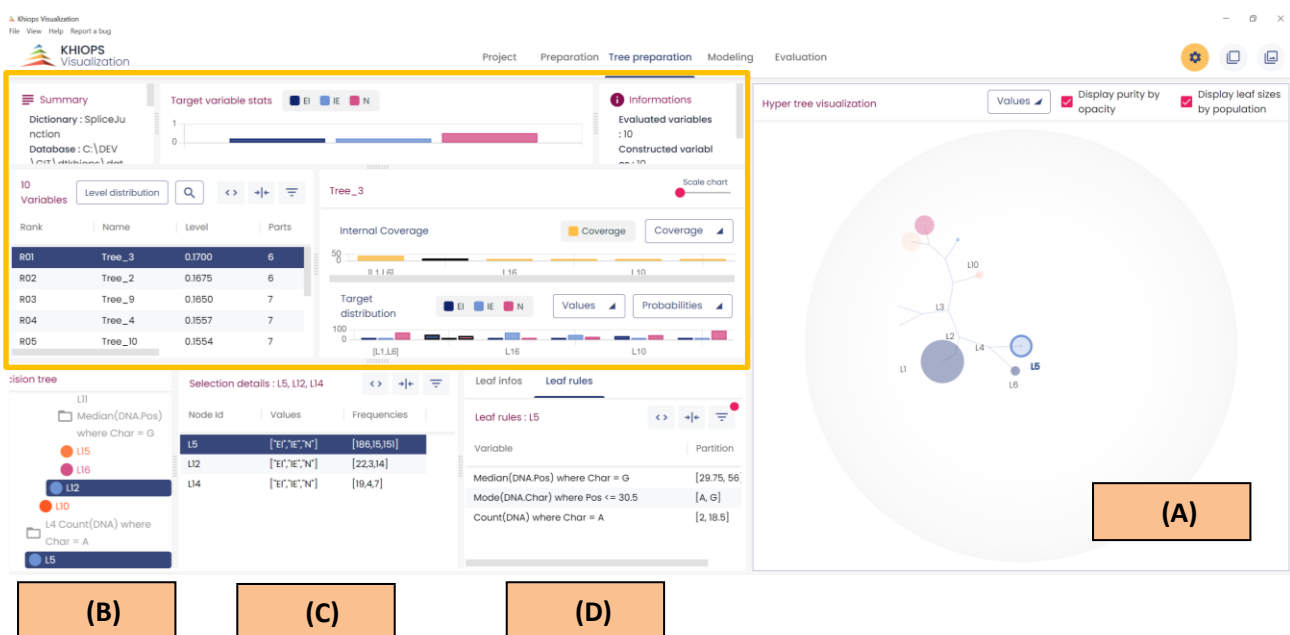
The right sub-pane (A) is a visualization of the tree structure in a hyperbolic space; dragging a leaf to the center magnifies the neighboring area of the tree.

Leaf area can be made proportional to its population, leaf opacity can be made proportional to its target class distribution purity (the more, the purer). The tree is also given as a hierarchy (B).

Sub-pane (C) describes the content of the selected group of leaves

A sub-pane at the bottom (D) gives the information related to the chosen leaf, with two sub-sub-panes: the leaf infos displays the histogram of the target values, and the leaf rules provides the sequence of tree rules (ordered from top to bottom) leading to it.

There are multiple selection modes. Basically, everything that is clickable in one panel selects what is relevant in the others. For example, a click on a leaf in the hypertree (A) select the other leaves of the same group in (A) and (B), the information on the related group in variable sub-pane (in orange rectangle) and in (C), and the information related to the selected leaf in (D).



5. Technical limits

In order to be responsive on very large files, some display limitations are taken into account in Khiops Visualization:

- Only the 10 000 most frequent values are displayed per group.