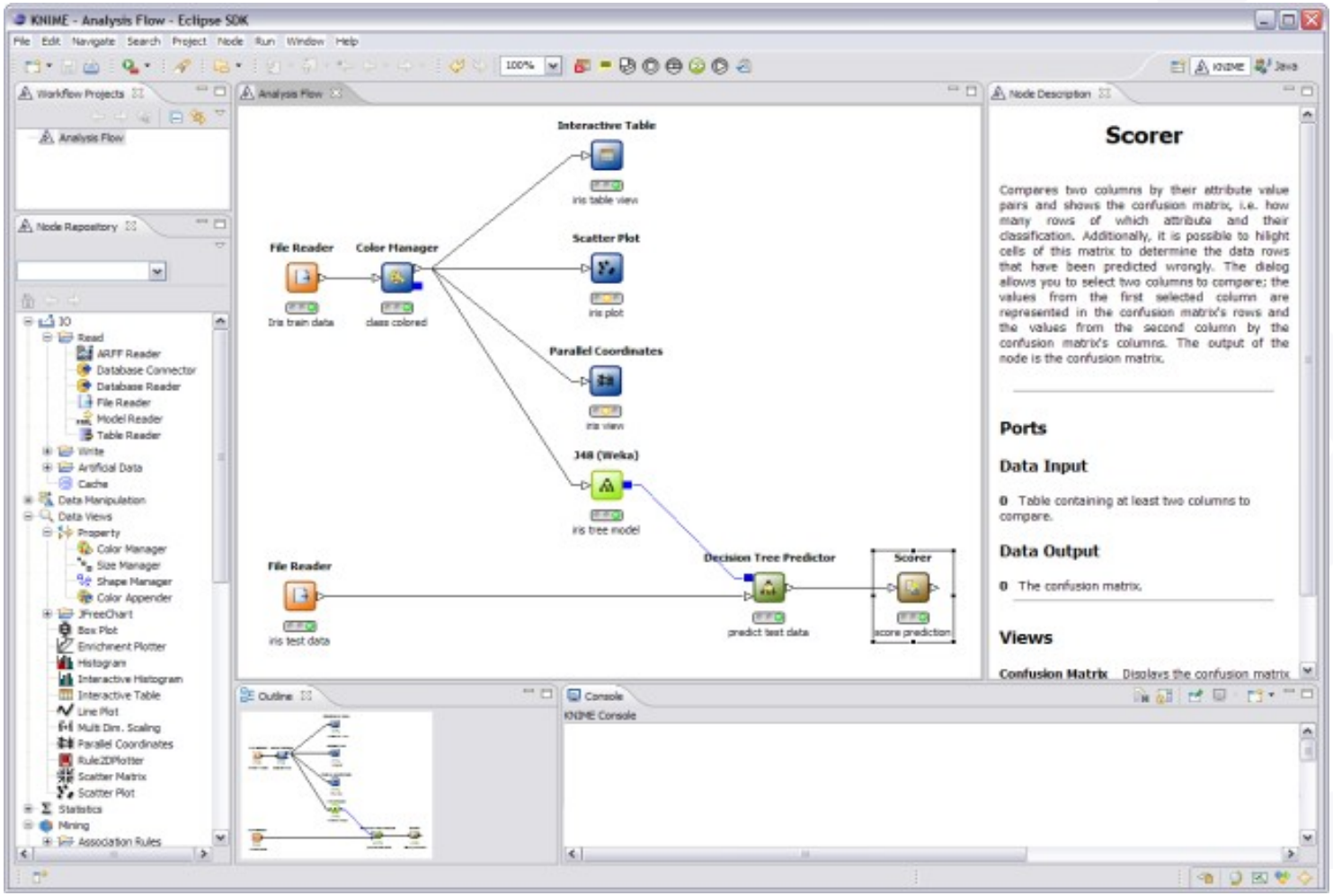
OME & KNIME

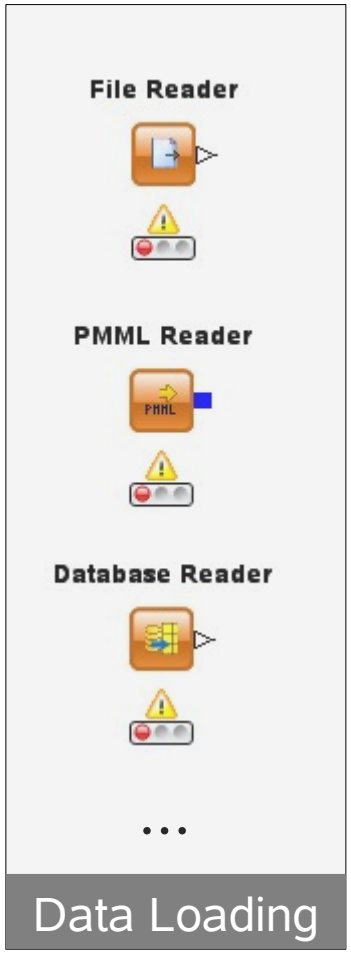
100615 OME metting.ppt

- Modular open source data analysis environment
- Used in pharmaceutical companies, life science, research
- > 3,000 active users
- Development was started 2004
- Current Version: KNIME 2.1.2

The KNIME platform



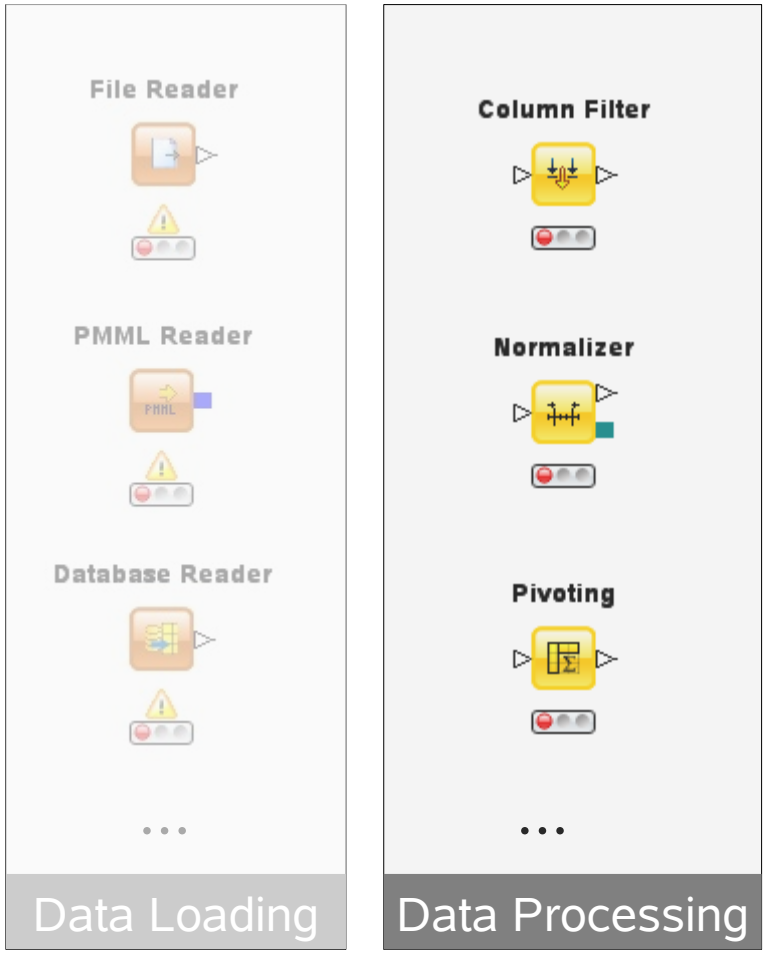
KNIME



KNIME loads and integrates data from diverse data sources:

- Different data bases
- Various file formats (CSV, XML, SDF, etc.)
- Images (**using Bio-Formats**)

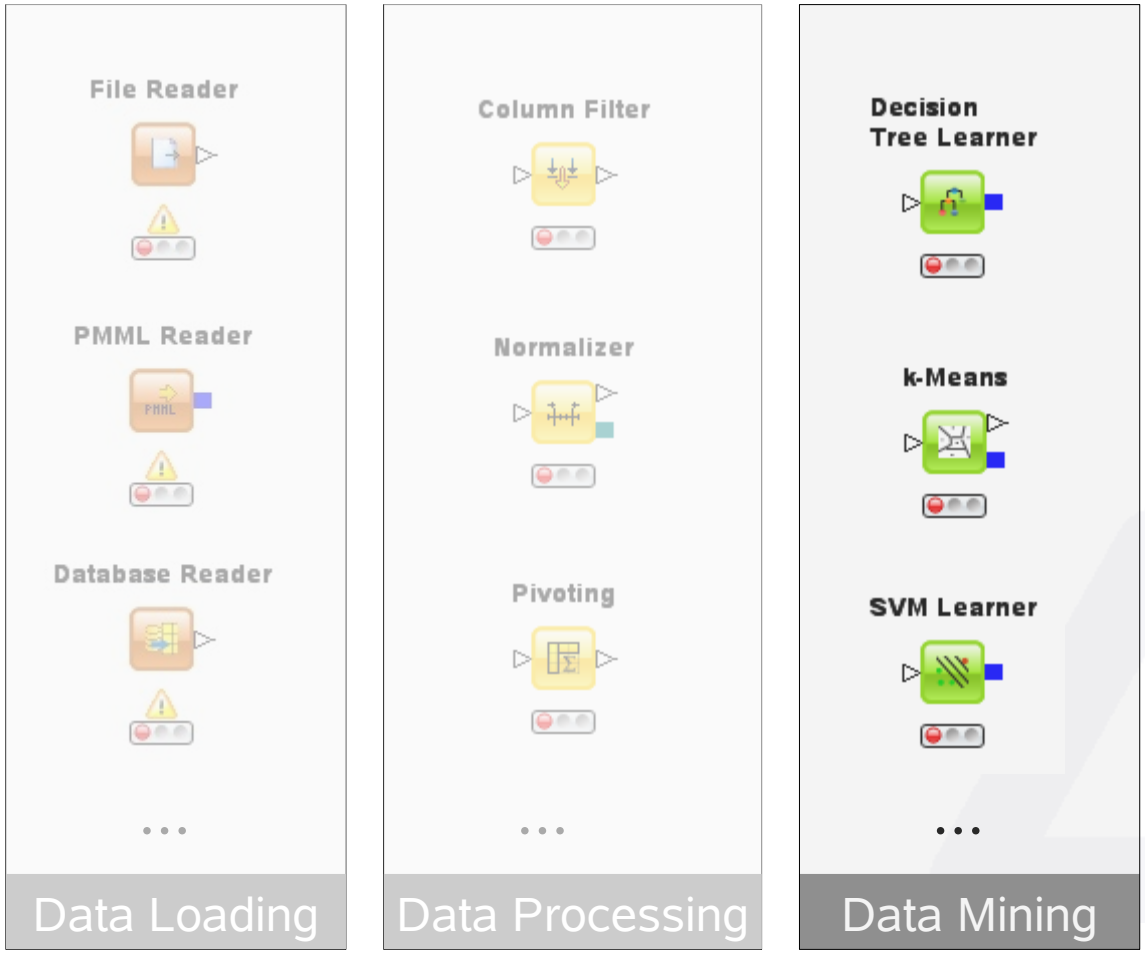
KNIME



KNIME provides huge repository of modules for easy-to-use, modular

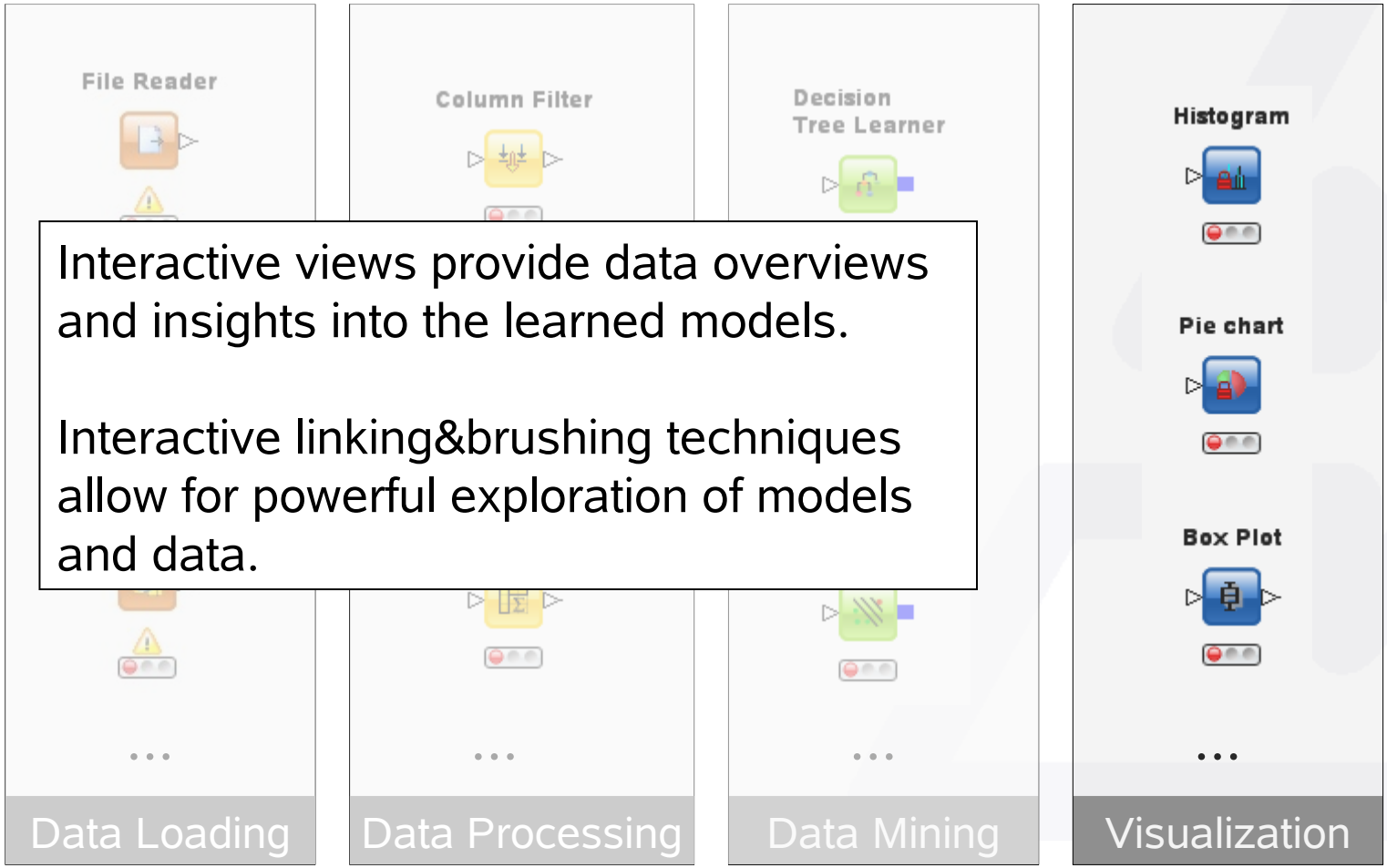
- Data preprocessing
- Data fusion
- Data transformation

KNIME



In addition to standard data mining techniques, KNIME adds cutting edge data analysis algorithms. (…thanks to its academic roots)

KNIME



Interactive views provide data overviews and insights into the learned models.

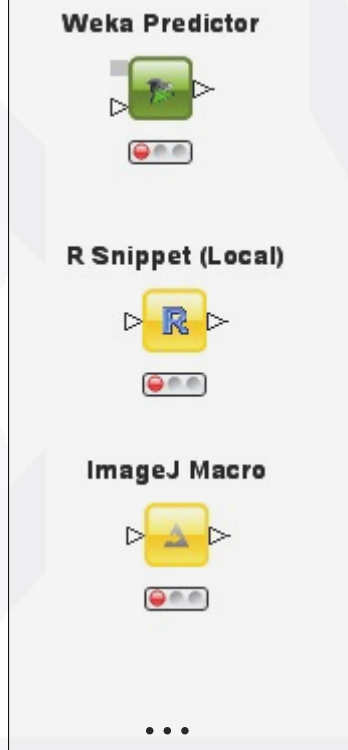
Interactive linking&brushing techniques allow for powerful exploration of models and data.

KNIME

Due to its open API and “node-in-a-sandbox”-approach additional (also external) tools are easily integrated, e.g.

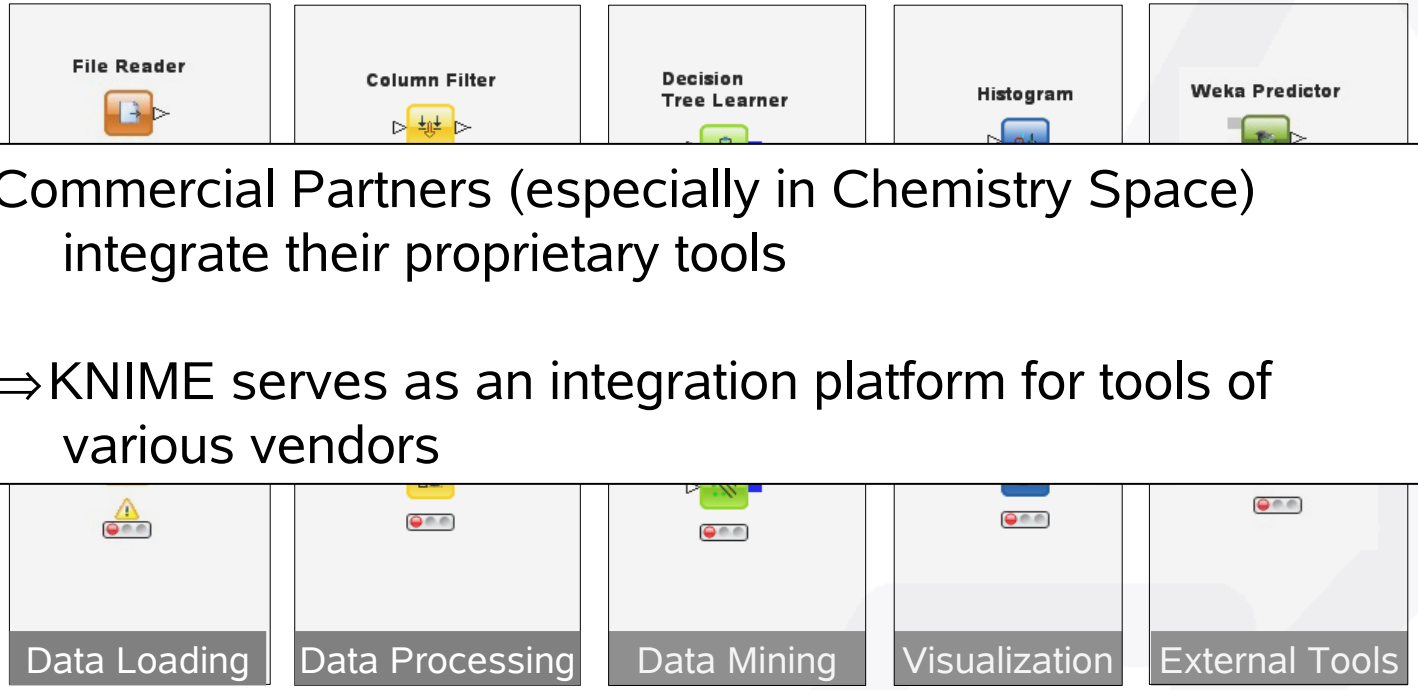
- Access to the statistics tool R
- Complete integration of the machine learning library WEKA
- Application area specific integration, e.g. CDK (Chemical Development Kit)
- **Inclusion of almost all ImageJ functionalities**

KNIME is Eclipse-based: Integrating other Eclipse projects such as BIRT, DTP, etc. provides even more functionality



Data Loading | Data Processing | Data Mining | Visualization | External Tools

KNIME

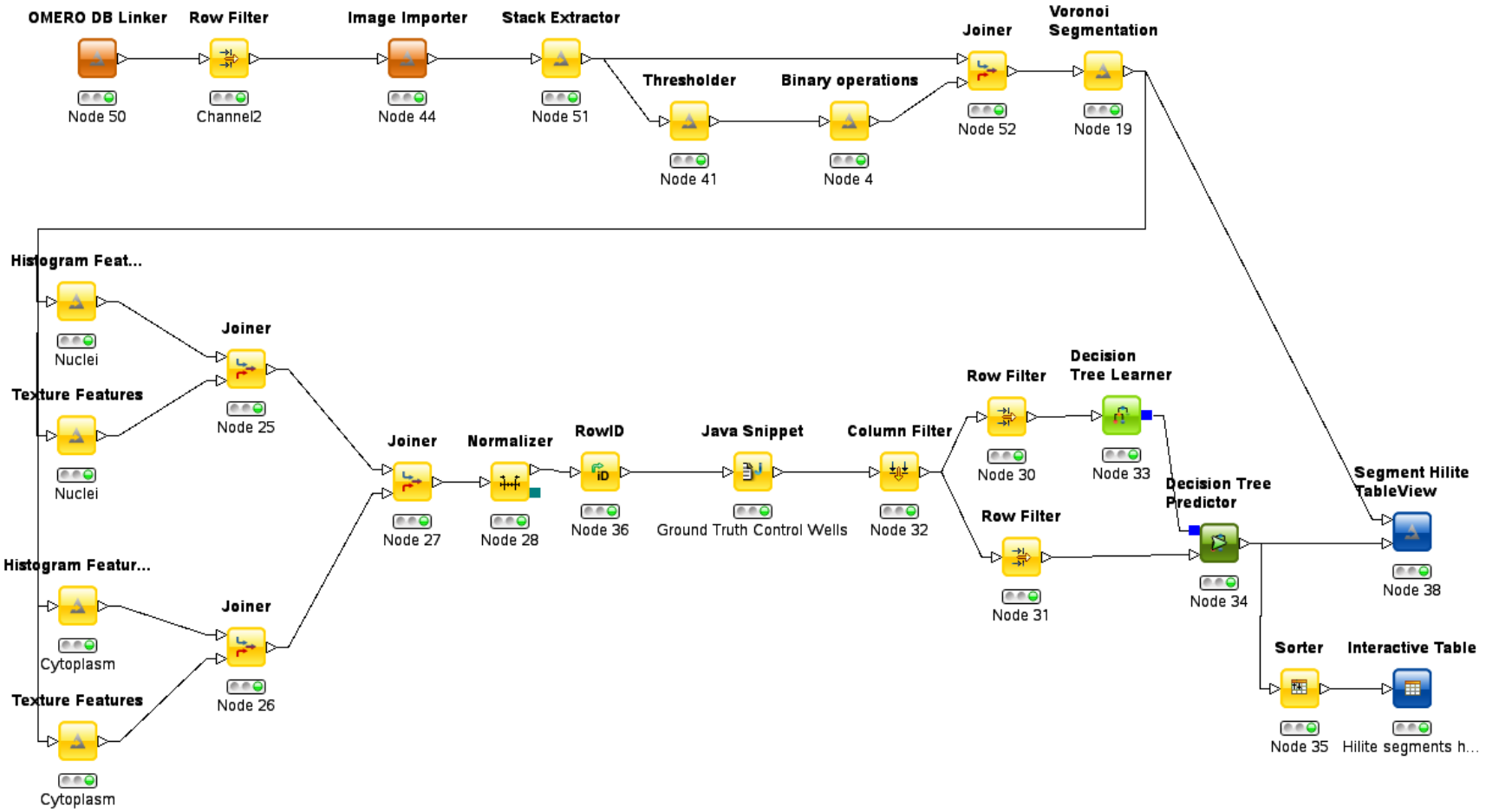


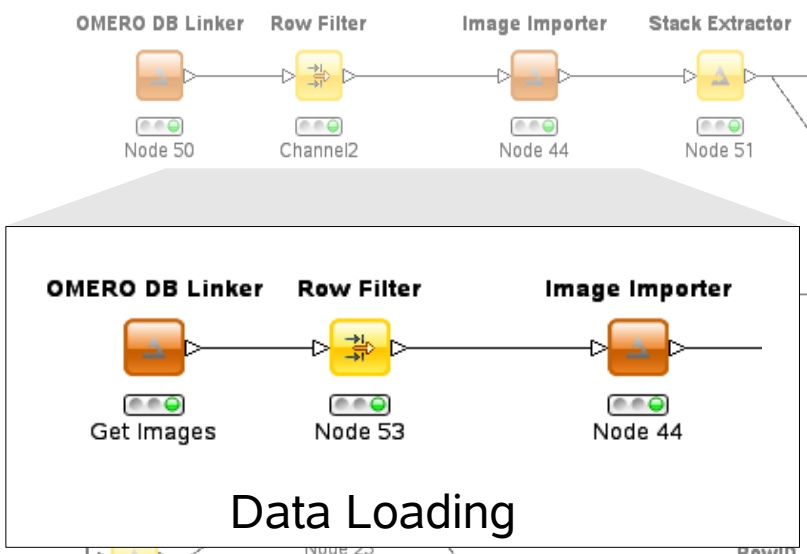
Commercial Partners (especially in Chemistry Space) integrate their proprietary tools

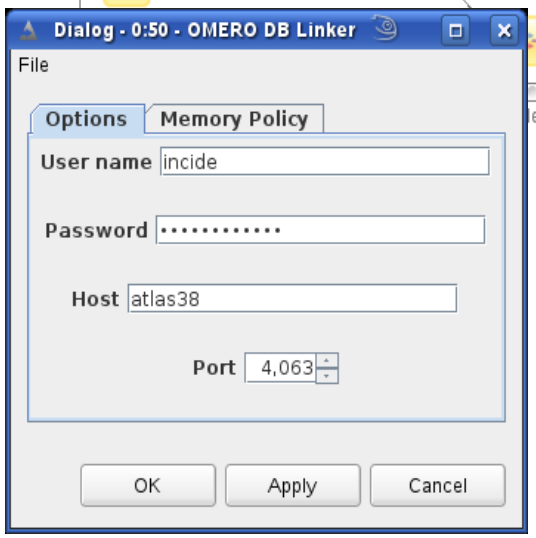
⇒KNIME serves as an integration platform for tools of various vendors

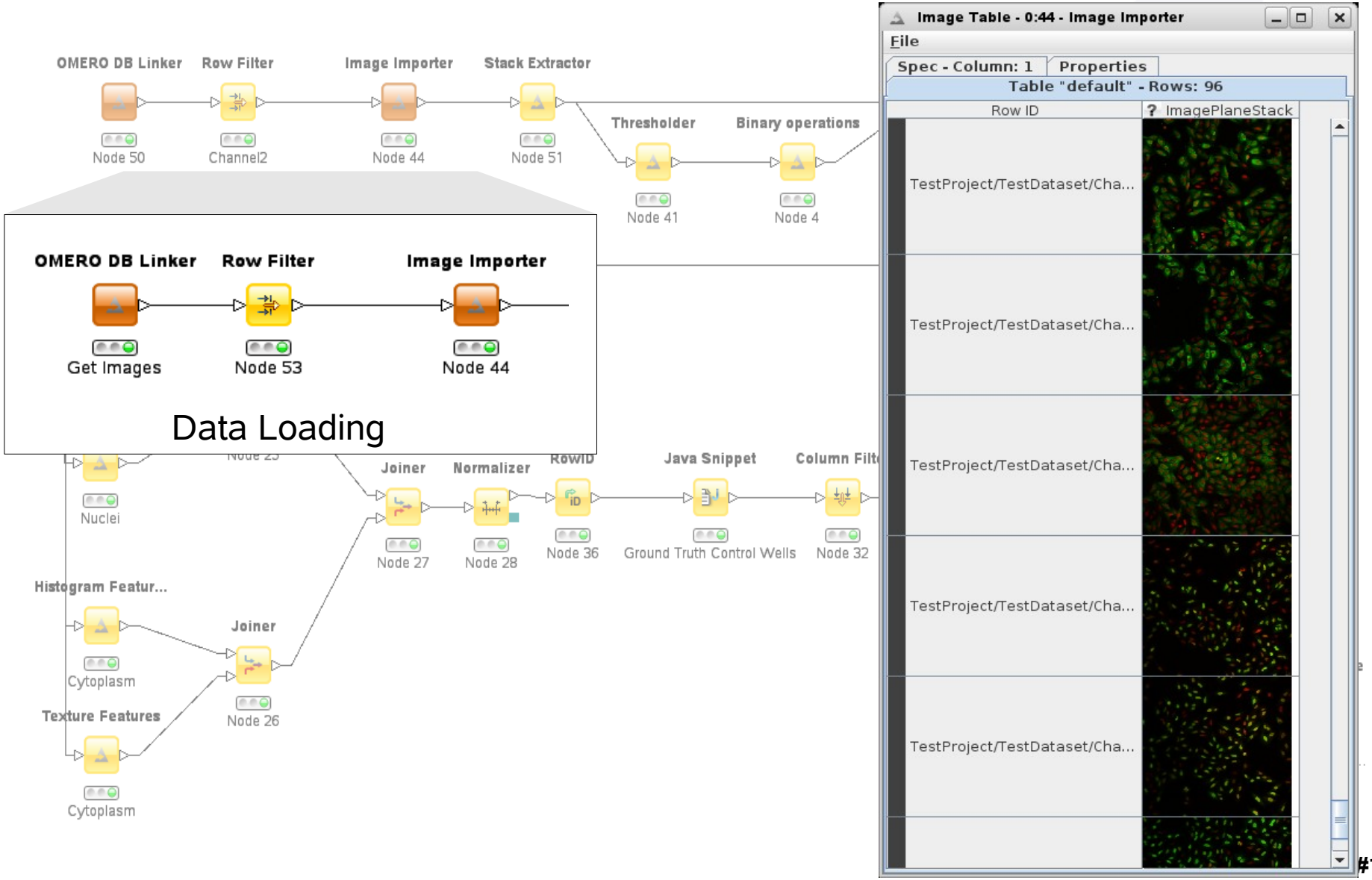
3rd Party Tools

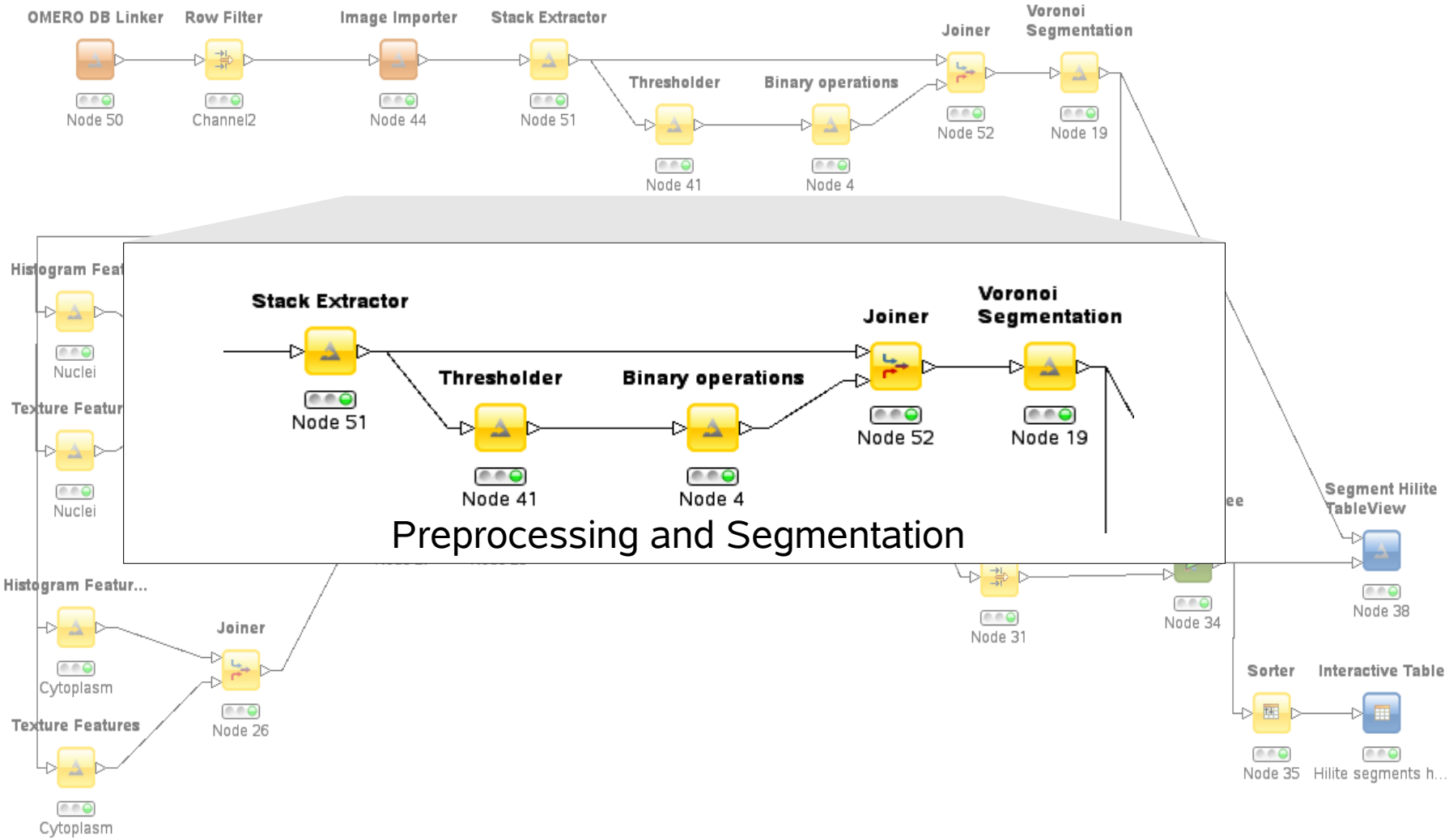
An Exemplary Workflow



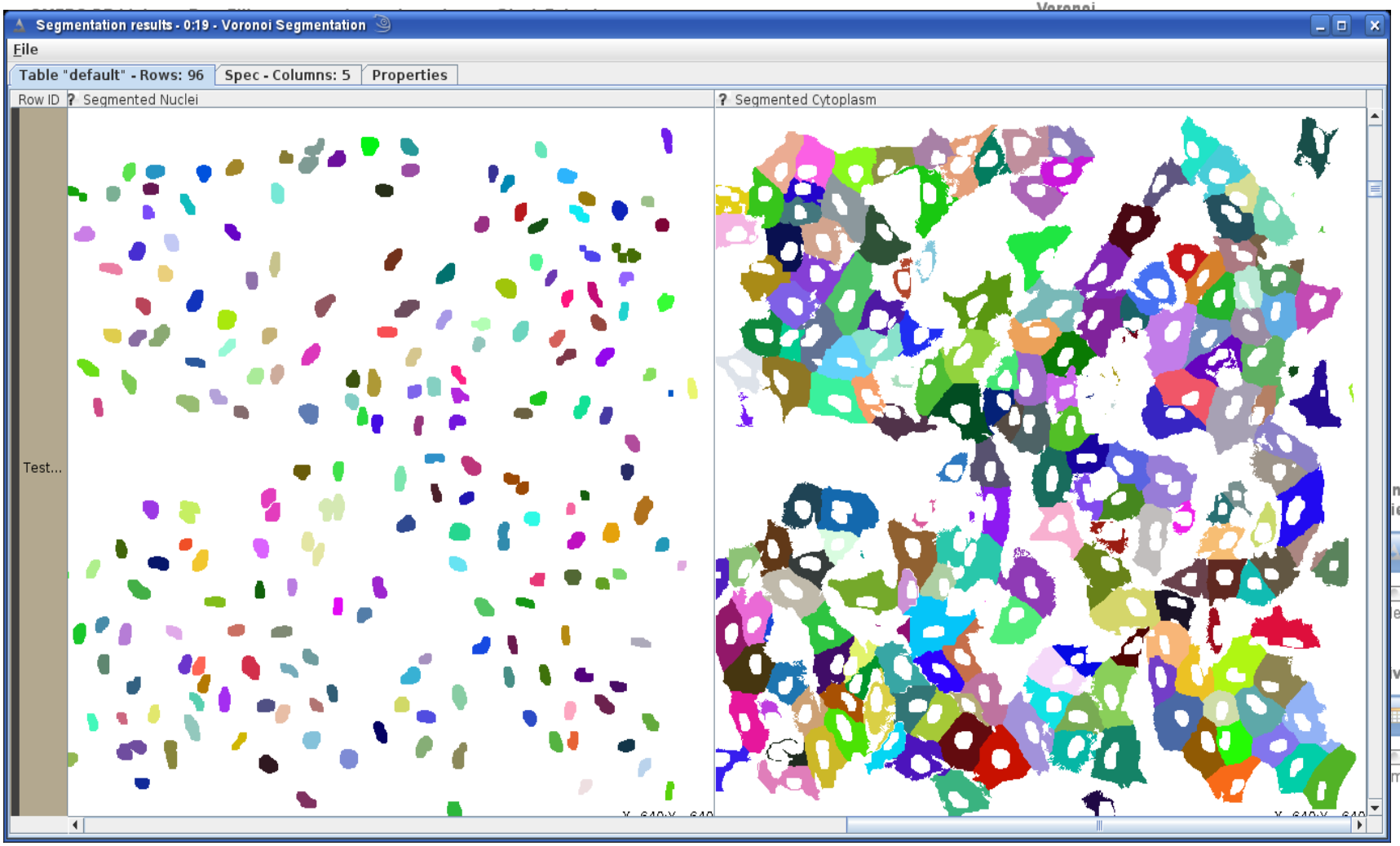

OMERO DB Linker
Row Filter
Image Importer
Get Images
Node 53
Node 44
Data Loading

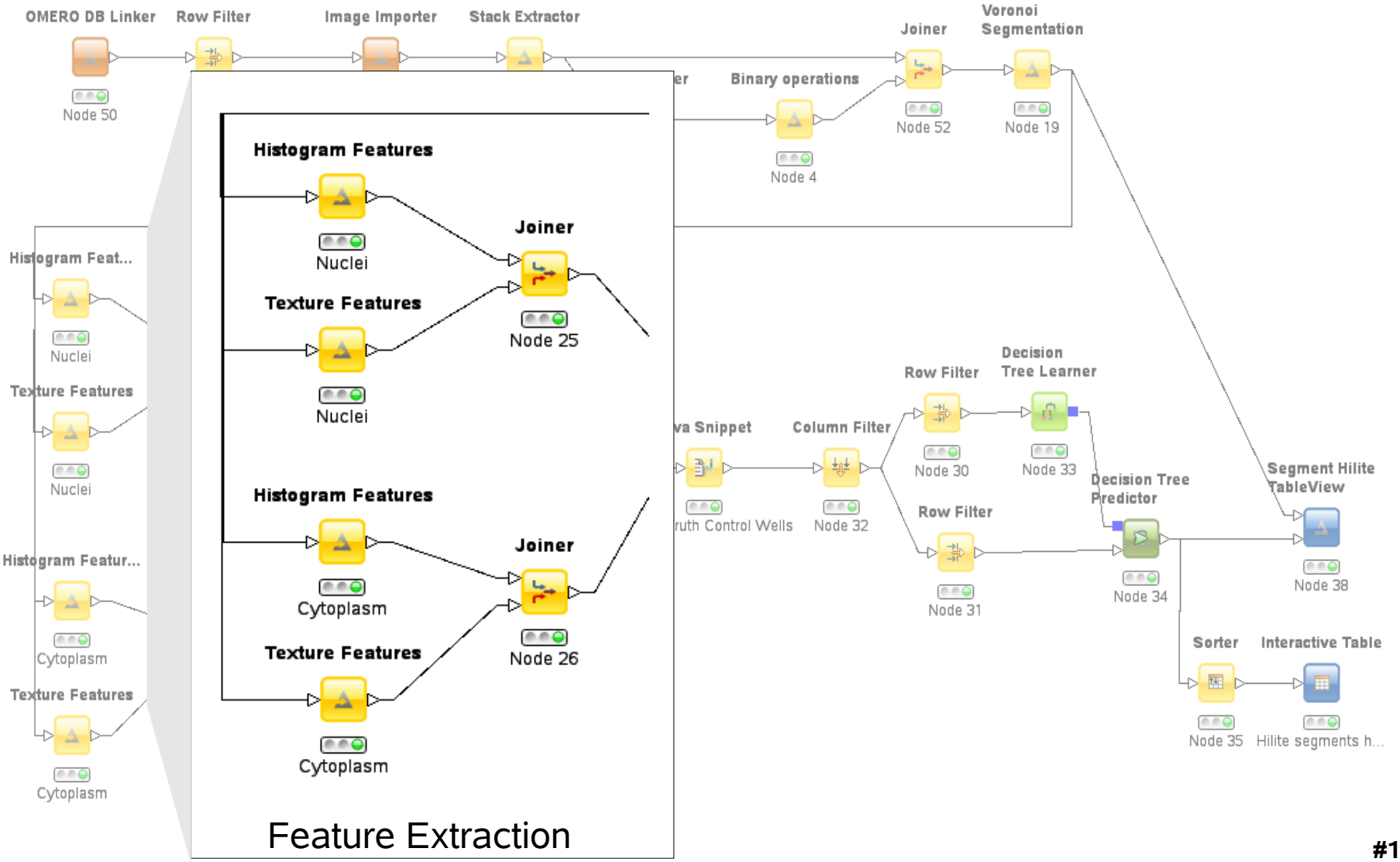

Dialog - 0:50 - OMERO DB Linker
File
Options
Memory Policy
User name incide
Password ••••••••••••
Host atlas38
Port 4,063
OK
Apply
Cancel


OMERO DB Linker
Row Filter
Image Importer
Stack Extractor
Thresholder
Binary operations
Node 50
Channel2
Node 44
Node 51
Node 41
Node 4
OMERO DB Linker
Row Filter
Image Importer
Get Images
Node 53
Node 44
Data Loading
Nuclei
Joiner
Normalizer
RowID
Java Snippet
Column Filt
Node 27
Node 28
Node 36
Ground Truth Control Wells
Node 32
Histogram Featur...
Joiner
Cytoplasm
Node 26
Texture Features
Cytoplasm
Image Table - 0:44 - Image Importer
File
Spec - Column: 1
Properties
Table "default" - Rows: 96
Row ID
ImagePlaneStack
TestProject/TestDataset/Cha...
TestProject/TestDataset/Cha...
TestProject/TestDataset/Cha...
TestProject/TestDataset/Cha...
TestProject/TestDataset/Cha...



Preprocessing and Segmentation


Segmentation results - 0:19 - Voronoi Segmentation
File
Table "default" - Rows: 96
Spec - Columns: 5
Properties
Row ID
Segmented Nuclei
Segmented Cytoplasm
Test...



Feature Extraction

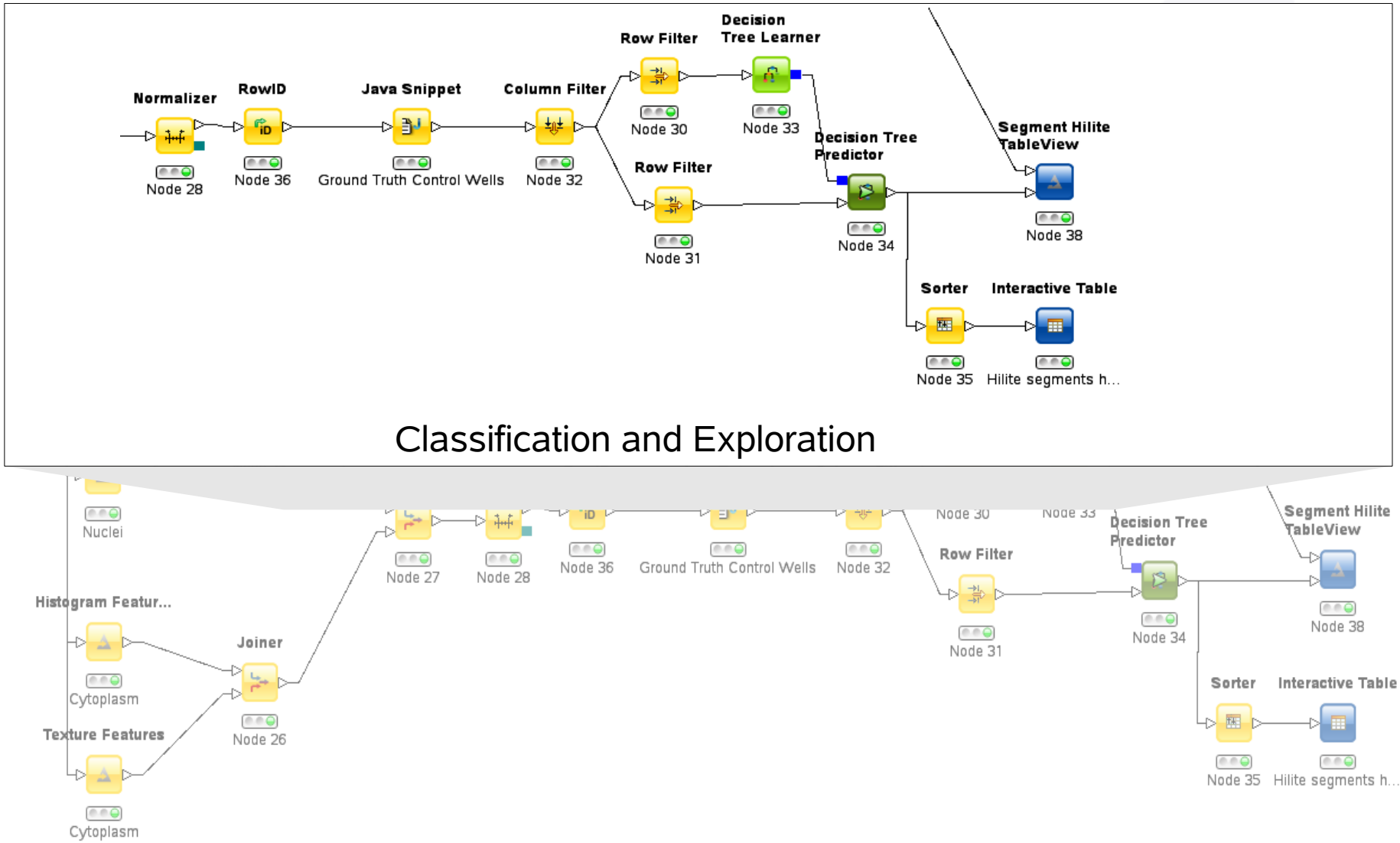
Joined table - 0:27 - Joiner

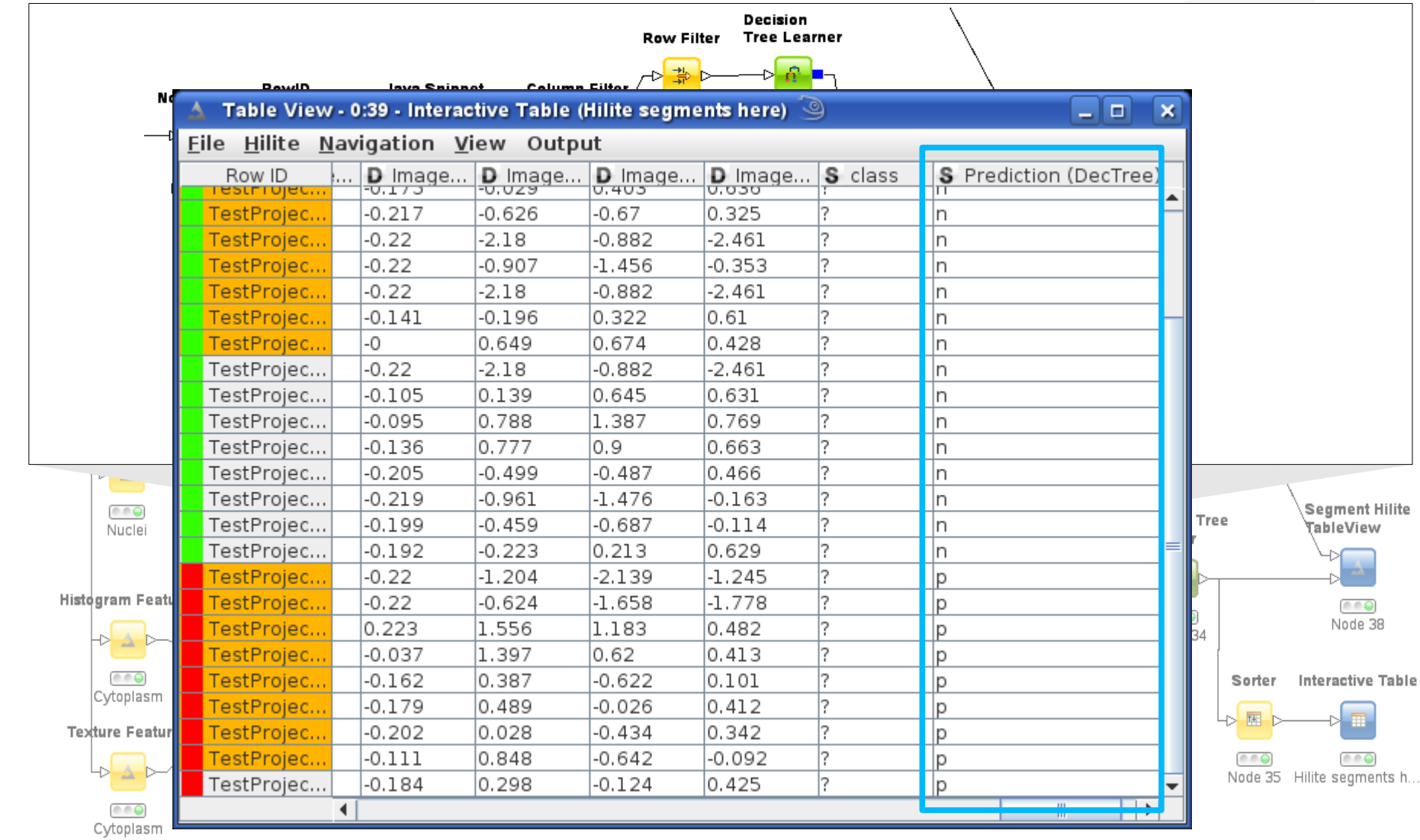
File

Table "default" - Rows: 14977 | Spec - Columns: 42 | Properties

Row ID	ImagePlane [Z0;C1;T0] mean	Imag...	Imag...	Image...	Image...	Image...	Image...	Image...	Image...	Image
TestProjec...	11.087	0.536	0.287	0.033	0.014	2.931	4	15	0.336	0.329
TestProjec...	152.807	1.745	3.045	0.92	1.026	3.761	90	162	0.403	3.279
TestProjec...	83.12	1.99	3.961	0.199	0.675	5.437	42	112	0.03	4.267
TestProjec...	24.301	0.961	0.924	0.02	0.078	4.149	13	34	0.172	0.4
TestProjec...	81.57	1.942	3.771	0.294	0.702	5.339	53	111	0.029	2.833
TestProjec...	71.199	1.449	2.1	0.226	0.266	4.897	38	87	0.091	1.444
TestProjec...	83.786	2.257	5.093	0.136	0.97	5.844	33	121	0.022	3.766
TestProjec...	115.124	6.167	38.036	154.365	830.606	4.12	49	132	0.266	5.056
TestProjec...	76.815	1.397	1.953	0.044	0.169	5.066	54	95	0.064	1.612
TestProjec...	10.062	0.44	0.194	0.024	0.015	2.058	7	12	1	0
TestProjec...	54.376	1.307	1.709	0.215	0.225	4.579	30	66	0.108	1.031
TestProjec...	34.12	0.877	0.77	-0.128	0.078	4.098	25	50	0.231	0.371
TestProjec...	33.46	0.909	0.827	-0.104	0.109	4.06	24	48	0.291	0.346
TestProjec...	6.311	0.318	0.101	0.009	0.004	1.659	4	8	0.276	0.371
TestProjec...	41.25	1.138	1.294	0.224	0.135	4.642	11	54	0.128	0.661
TestProjec...	17.336	0.58	0.337	0.023	0.013	3.367	9	23	0.449	0.24
TestProjec...	17.558	0.594	0.353	0.015	0.016	3.225	12	23	0.461	0.21
TestProjec...	15.786	0.689	0.475	0.006	0.045	3.189	10	21	0.703	0.15
TestProjec...	116.058	2.408	5.796	-0.451	1.447	5.509	58	152	0.032	5.028
TestProjec...	93.94	3.226	10.405	11.624	23.178	5.167	36	112	0.059	6.541
TestProjec...	97.219	3.631	13.184	22.322	58.305	5.598	40	123	0.029	4.46
TestProjec...	13.361	0.683	0.466	-0.009	0.042	2.845	10	17	0.346	0.242
TestProjec...	99.012	1.552	2.41	0.029	0.271	5.423	60	125	0.043	1.842
TestProjec...	35.791	1.166	1.359	0.086	0.141	4.399	23	46	0.146	0.402
TestProjec...	8.568	0.493	0.243	-0.008	0.016	2.554	5	17	0.833	0.081
TestProjec...	47.746	1.19	1.416	-0.119	0.143	4.78	30	66	0.105	0.614
TestProjec...	30.151	0.898	0.806	-0.118	0.069	3.998	21	48	0.237	0.45
TestProjec...	164.083	5.061	25.612	77.39	343.68	5.028	95	189	0.041	6.156
TestProjec...	114.401	3.323	11.044	0.756	3.483	6.299	26	168	0.013	13.897
TestProjec...	14.329	0.506	0.256	0.007	0.012	2.85	8	19	0.727	0.097
TestProjec...	3.736	0.295	0.087	-0.006	0.003	1.387	2	5	1	0

Feature Extraction


Normalizer
RowID
Java Snippet
Column Filter
Row Filter
Decision
Tree Learner
Node 28
Node 36
Ground Truth Control Wells
Node 32
Node 30
Node 33
Row Filter
Decision Tree
Predictor
Segment Hilite
TableView
Node 31
Node 34
Node 38
Sorter
Interactive Table
Node 35
Hilite segments h...
Classification and Exploration
Nuclei
Histogram Featur...
Joiner
Cytoplasm
Texture Features
Node 26
Cytoplasm
Node 27
Node 28
Node 36
Ground Truth Control Wells
Node 32
Node 30
Node 33
Row Filter
Decision Tree
Predictor
Segment Hilite
TableView
Node 31
Node 34
Node 38
Sorter
Interactive Table
Node 35
Hilite segments h...


Row Filter
Decision
Tree Learner
Table View - 0:39 - Interactive Table (Hilite segments here)
File Hilite Navigation View Output
Row ID | D Image... | D Image... | D Image... | D Image... | S class | S Prediction (DecTree)
TestProjec... | -0.217 | -0.626 | -0.67 | 0.325 | ? | n
TestProjec... | -0.22 | -2.18 | -0.882 | -2.461 | ? | n
TestProjec... | -0.22 | -0.907 | -1.456 | -0.353 | ? | n
TestProjec... | -0.22 | -2.18 | -0.882 | -2.461 | ? | n
TestProjec... | -0.141 | -0.196 | 0.322 | 0.61 | ? | n
TestProjec... | -0 | 0.649 | 0.674 | 0.428 | ? | n
TestProjec... | -0.22 | -2.18 | -0.882 | -2.461 | ? | n
TestProjec... | -0.105 | 0.139 | 0.645 | 0.631 | ? | n
TestProjec... | -0.095 | 0.788 | 1.387 | 0.769 | ? | n
TestProjec... | -0.136 | 0.777 | 0.9 | 0.663 | ? | n
TestProjec... | -0.205 | -0.499 | -0.487 | 0.466 | ? | n
TestProjec... | -0.219 | -0.961 | -1.476 | -0.163 | ? | n
TestProjec... | -0.199 | -0.459 | -0.687 | -0.114 | ? | n
TestProjec... | -0.192 | -0.223 | 0.213 | 0.629 | ? | n
TestProjec... | -0.22 | -1.204 | -2.139 | -1.245 | ? | p
TestProjec... | -0.22 | -0.624 | -1.658 | -1.778 | ? | p
TestProjec... | 0.223 | 1.556 | 1.183 | 0.482 | ? | p
TestProjec... | -0.037 | 1.397 | 0.62 | 0.413 | ? | p
TestProjec... | -0.162 | 0.387 | -0.622 | 0.101 | ? | p
TestProjec... | -0.179 | 0.489 | -0.026 | 0.412 | ? | p
TestProjec... | -0.202 | 0.028 | -0.434 | 0.342 | ? | p
TestProjec... | -0.111 | 0.848 | -0.642 | -0.092 | ? | p
TestProjec... | -0.184 | 0.298 | -0.124 | 0.425 | ? | p
Nuclei
Histogram Featu
Cytoplasm
Texture Featur
Cytoplasm
Segment Hilite
TableView
Node 38
Sorter
Interactive Table
Node 35
Hilite segments h...

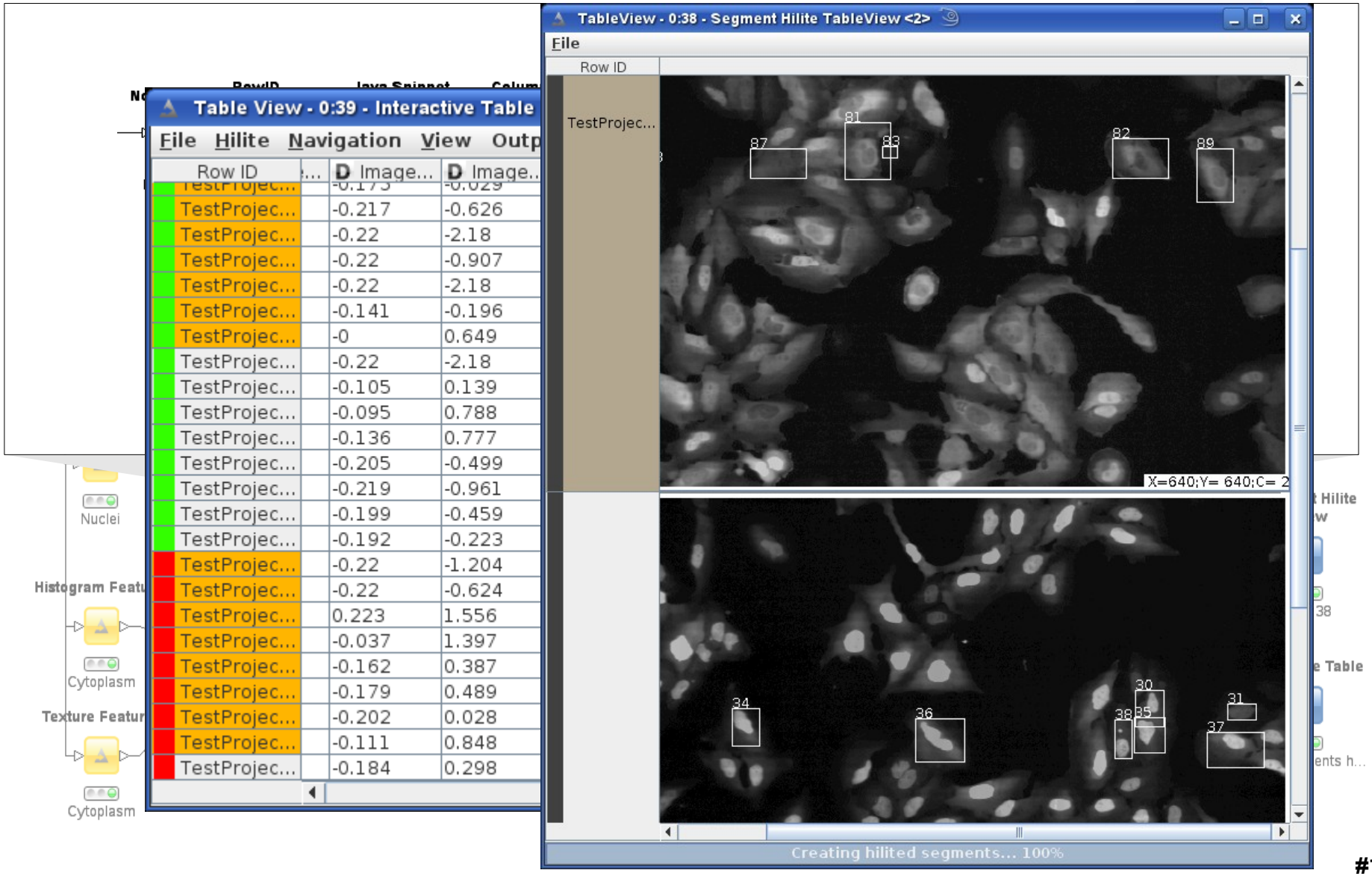
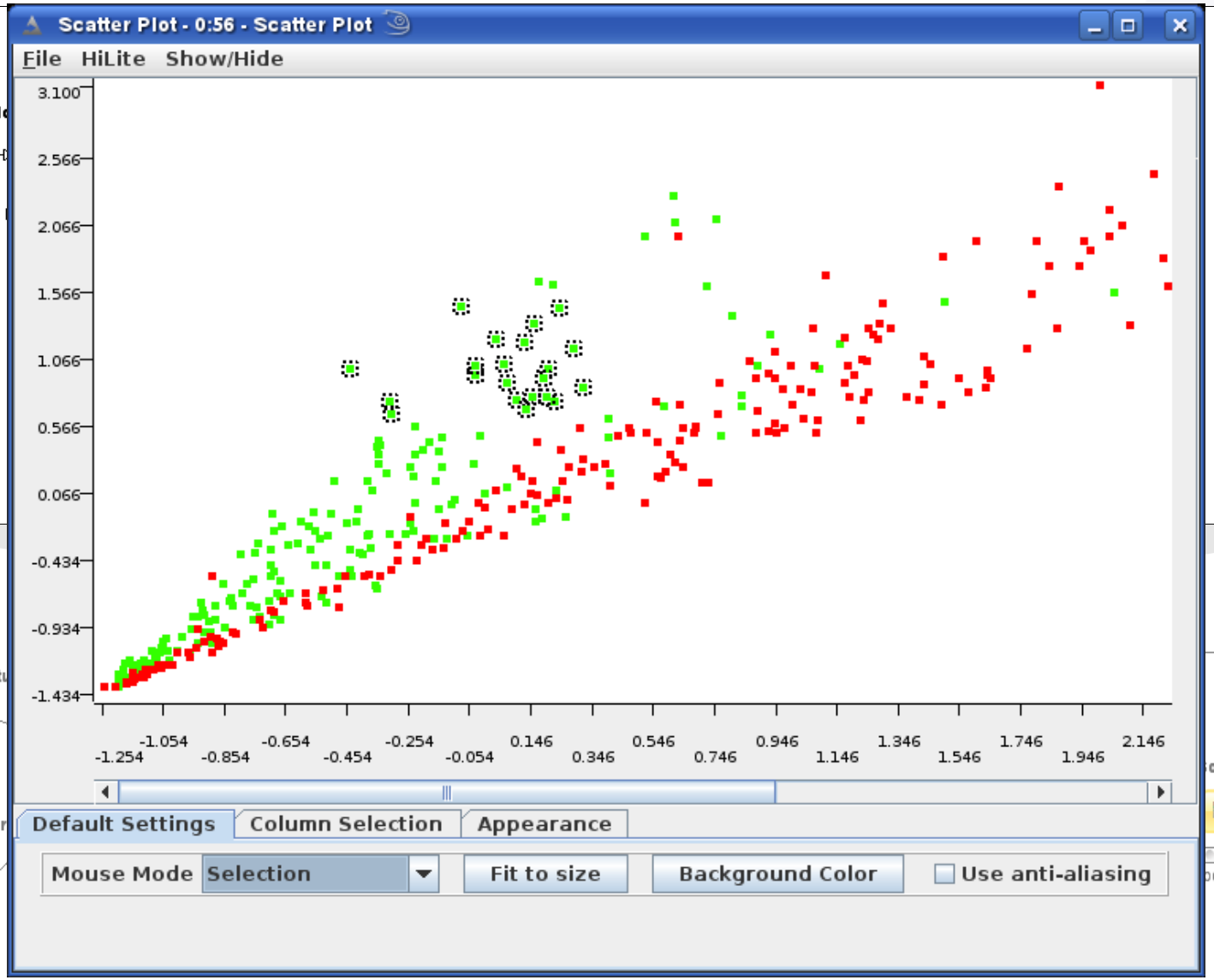


Table View - 0:39 - Interactive Table
File Hilite Navigation View Outp
Row ID
D Image...
D Image...
TestProjec... -0.217 -0.626
TestProjec... -0.22 -2.18
TestProjec... -0.22 -0.907
TestProjec... -0.22 -2.18
TestProjec... -0.141 -0.196
TestProjec... -0 0.649
TestProjec... -0.22 -2.18
TestProjec... -0.105 0.139
TestProjec... -0.095 0.788
TestProjec... -0.136 0.777
TestProjec... -0.205 -0.499
TestProjec... -0.219 -0.961
TestProjec... -0.199 -0.459
TestProjec... -0.192 -0.223
TestProjec... -0.22 -1.204
TestProjec... -0.22 -0.624
TestProjec... 0.223 1.556
TestProjec... -0.037 1.397
TestProjec... -0.162 0.387
TestProjec... -0.179 0.489
TestProjec... -0.202 0.028
TestProjec... -0.111 0.848
TestProjec... -0.184 0.298
TableView - 0:38 - Segment Hilite TableView <2>
File
Row ID
TestProjec...
81
87
83
82
89
X=640;Y= 640;C= 2
34
36
30
38
35
31
37
Creating hilited segments... 100%
Nuclei
Histogram Featu
Cytoplasm
Texture Featur
Cytoplasm
Hilite
38
Table


Scatter Plot - 0:56 - Scatter Plot
File HiLite Show/Hide
3.100
2.566
2.066
1.566
1.066
0.566
0.066
-0.434
-0.934
-1.434
-1.254 -1.054 -0.854 -0.654 -0.454 -0.254 -0.054 0.146 0.346 0.546 0.746 0.946 1.146 1.346 1.546 1.746 1.946 2.146
Default Settings
Column Selection
Appearance
Mouse Mode
Selection
Fit to size
Background Color
Use anti-aliasing

Some technical remarks

- KNIME takes care of caching and swapping the data rows → able to process huge data tables
- Images are treated as BLOBS
- Extension through the eclipse plug-in/update mechanism

Resources

http://www.knime.org - the home of KNIME

http://www.inf.uni-konstanz.de/sighca - WIKI related to High Content Analytics, also providing the associated plugins

http://labs.knime.org - additional experimental plugins