University of Minho
School of Engineering

ISLab

Machine Learning and Decision-Making

ADI @ LEI/3º, MiEI/4º - 2º Semestre

Filipe Gonçalves, Inês Alves, Cesar Analide

Part III – March 2022

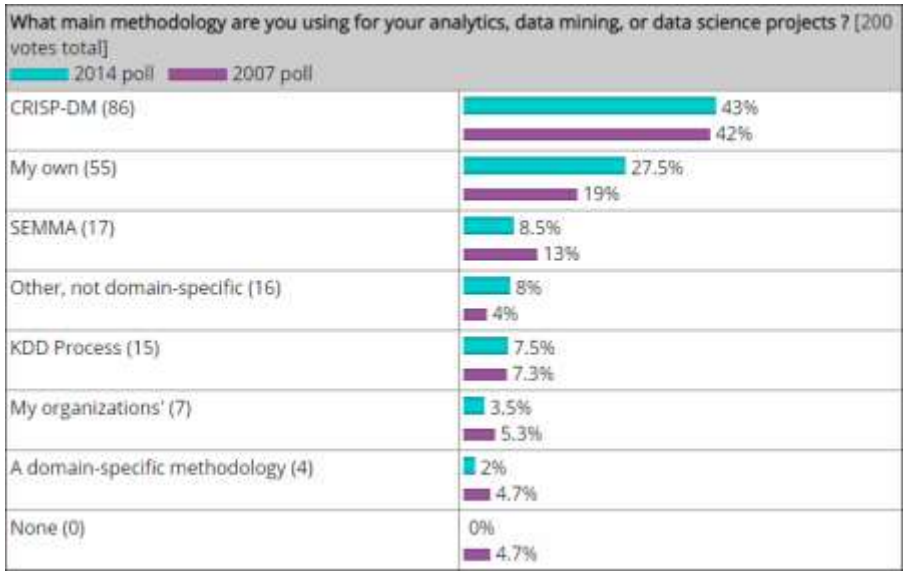
Contents

- Methodologies
- Knime
  - Good Habits
  - Metanodes
  - Data Ingestion
  - Data Partitioning/Segregation
- Data Quality and Exploration
- Hands On

Methodologies

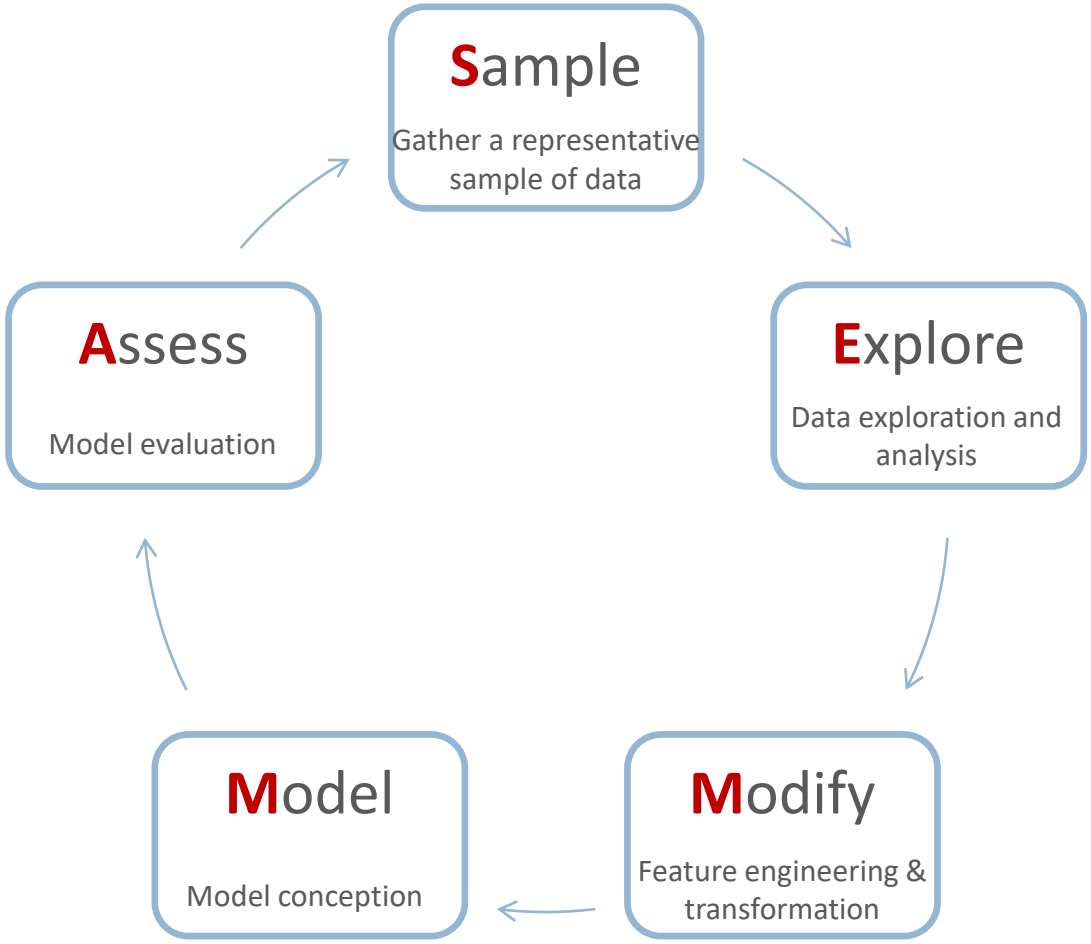
Why standard methodologies?

- Allows projects to be replicated
- Aid project planning and management
- Encourage best practices and help to obtain better results



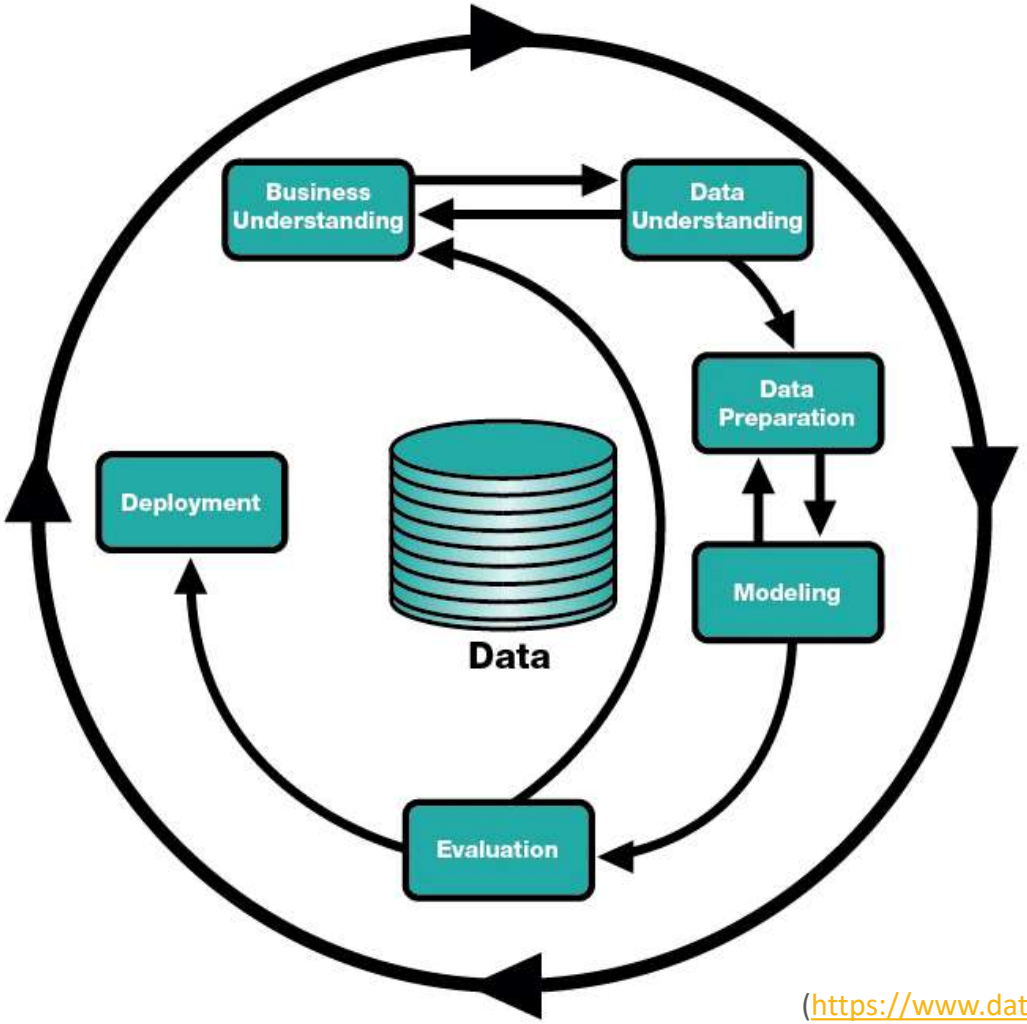
(https://www.kdnuggets.com/2014/10/crisp-dm-top-methodology-analytics-data-mining-data-science-projects.html)

SEMMA


Sample
Gather a representative sample of data
Explore
Data exploration and analysis
Modify
Feature engineering & transformation
Model
Model conception
Assess
Model evaluation

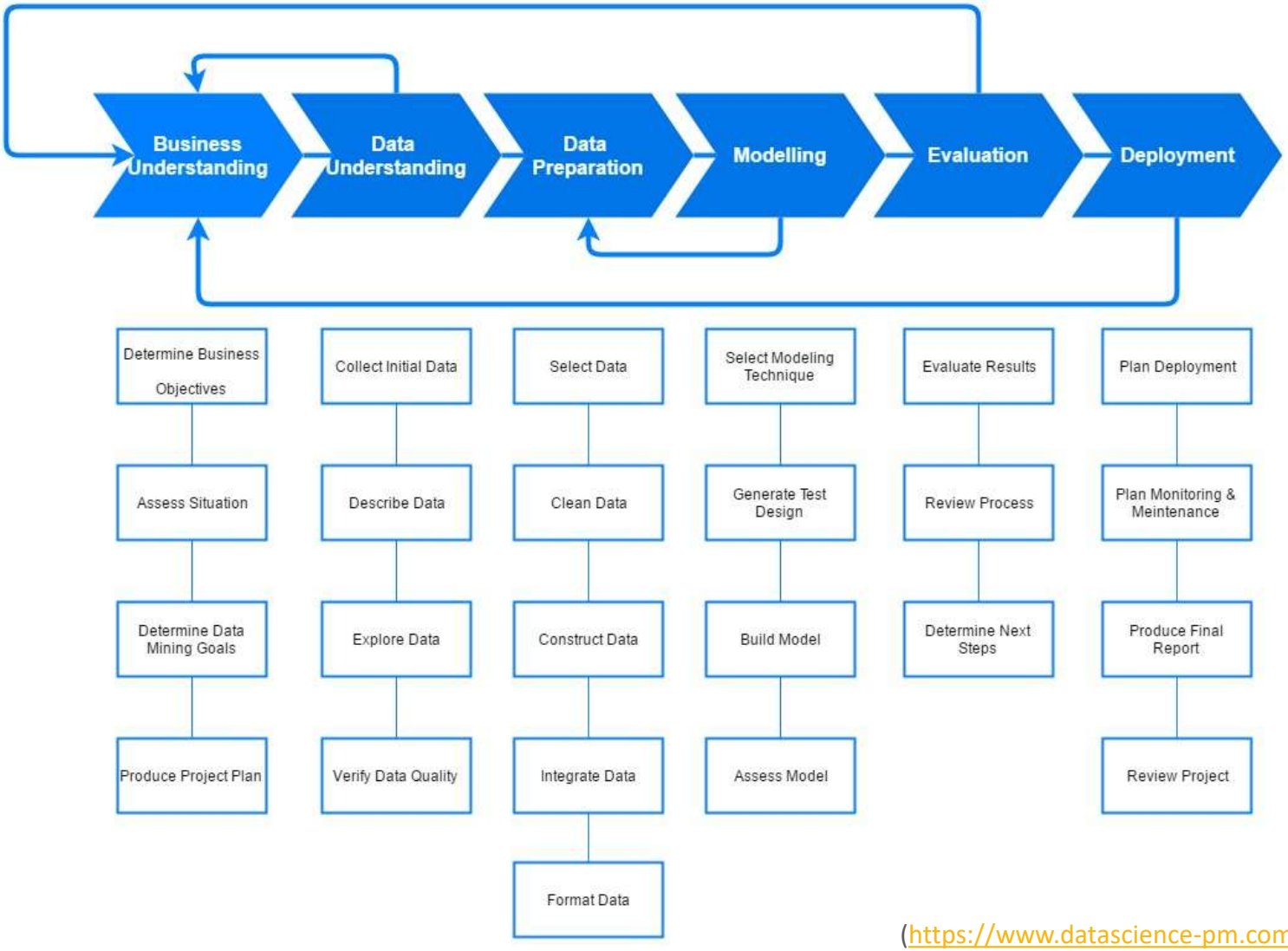
CRISP-DM

CRISP-DM stands for Cross Industry Standard Process for Data Mining



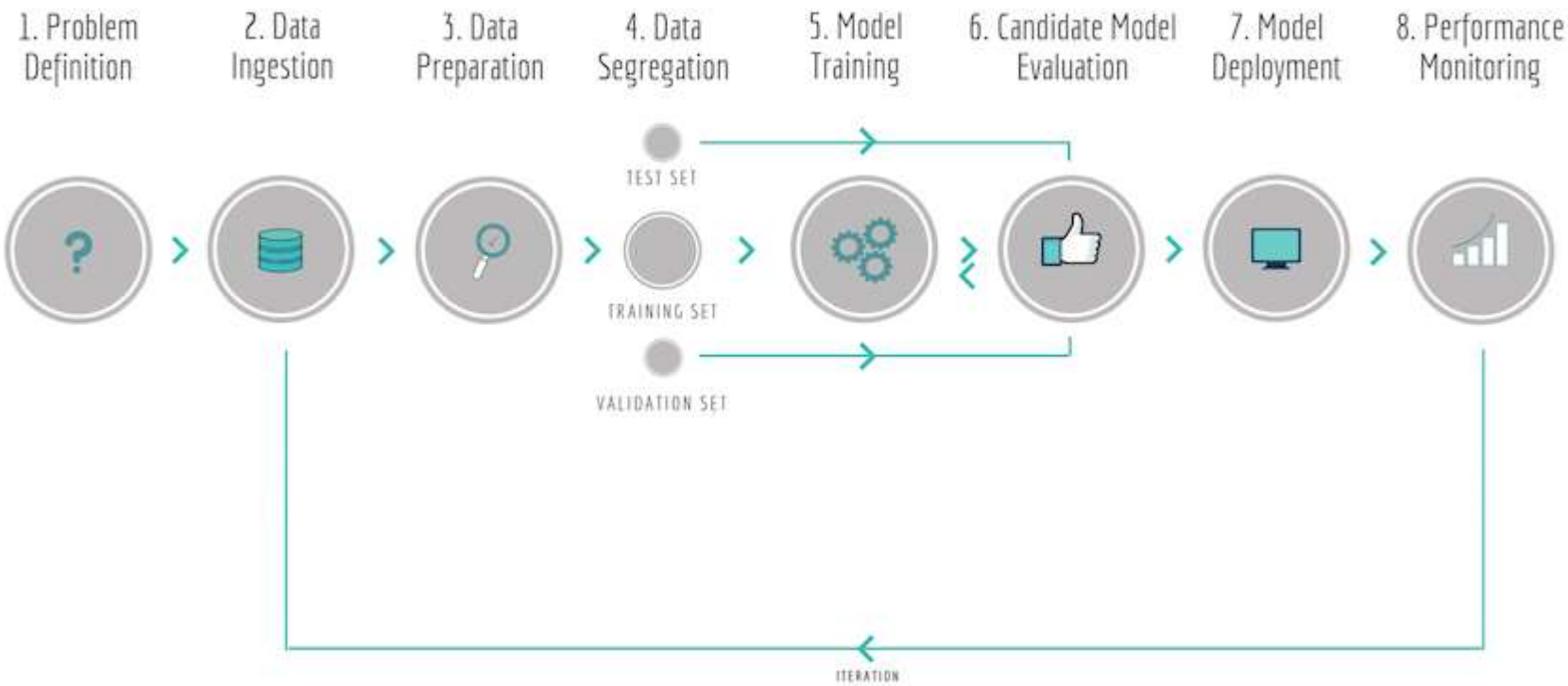
(https://www.datascience-pm.com/crisp-dm-2/)

CRISP-DM



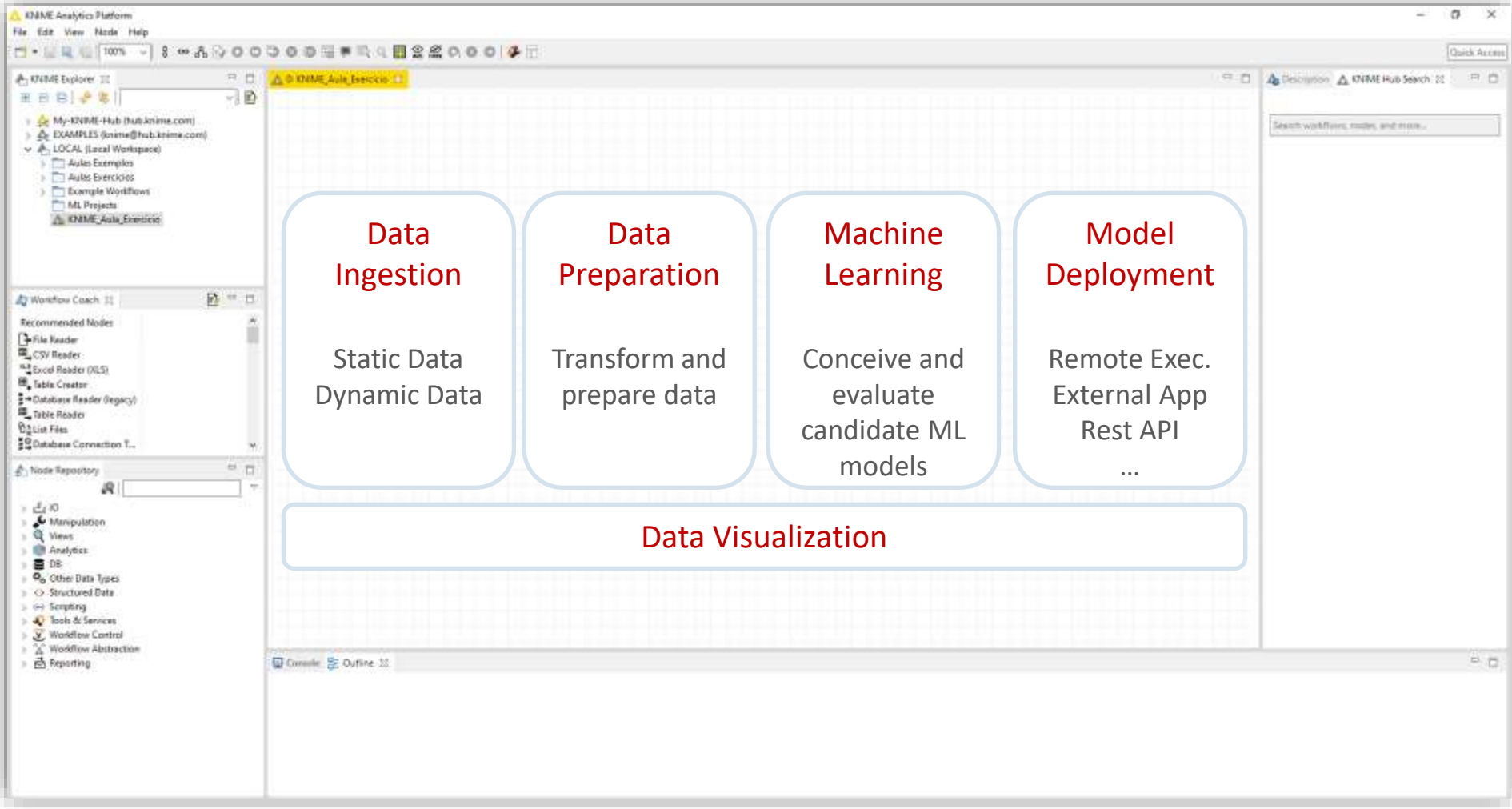
(https://www.datascience-pm.com/crisp-dm-2/)

A Machine Learning Pipeline



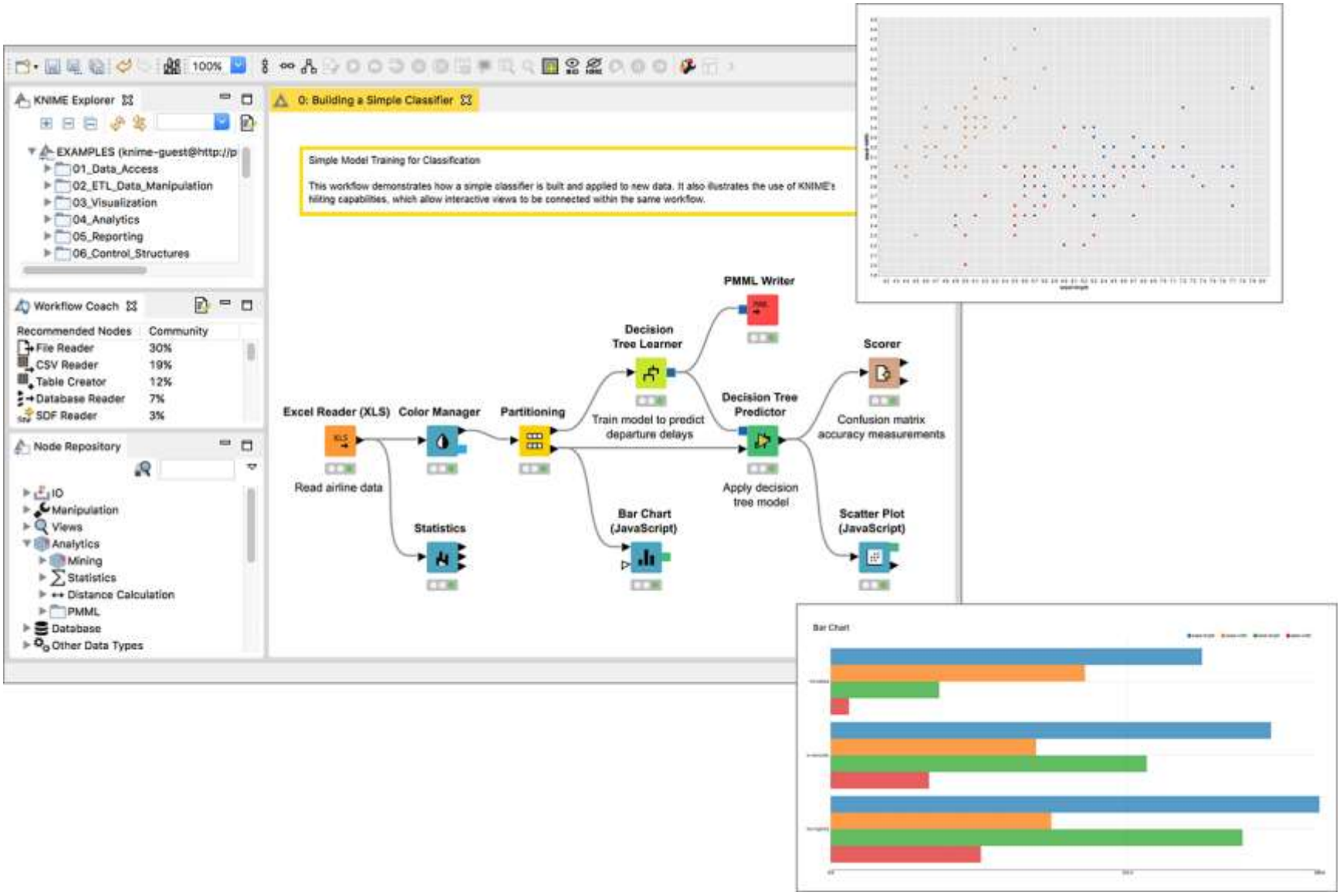
(https://towardsdatascience.com/architecting-a-machine-learning-pipeline-a847f094d1c7)

Generic Workflow Structure @ Knime



Open for Innovation ®

KNIME

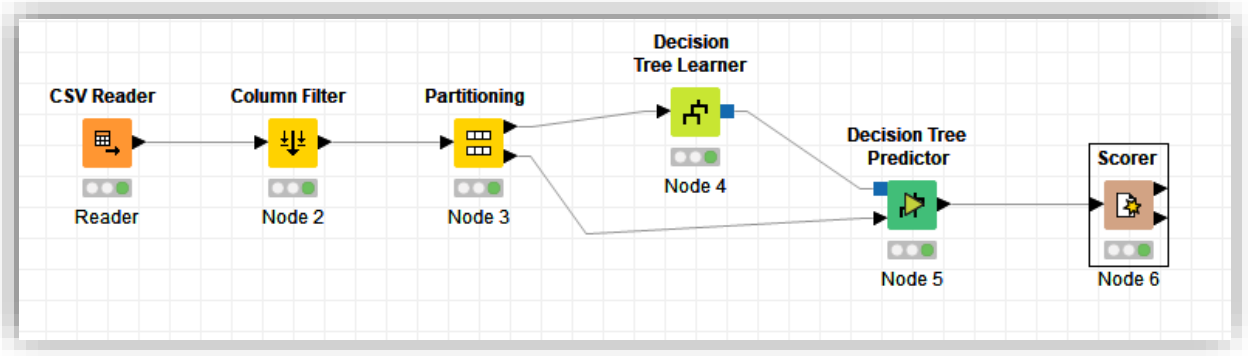


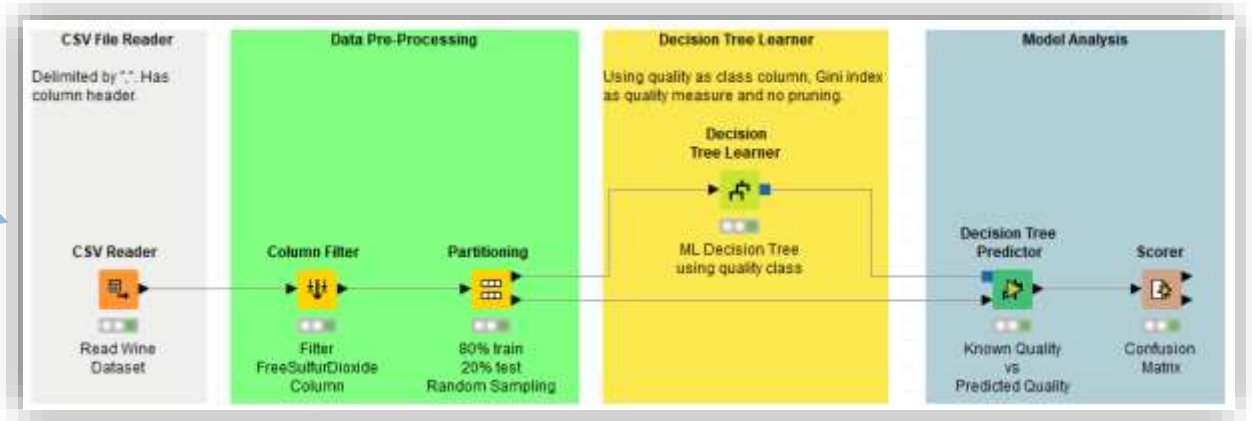
Good Habits!

- Rename nodes
- Add annotations

and…

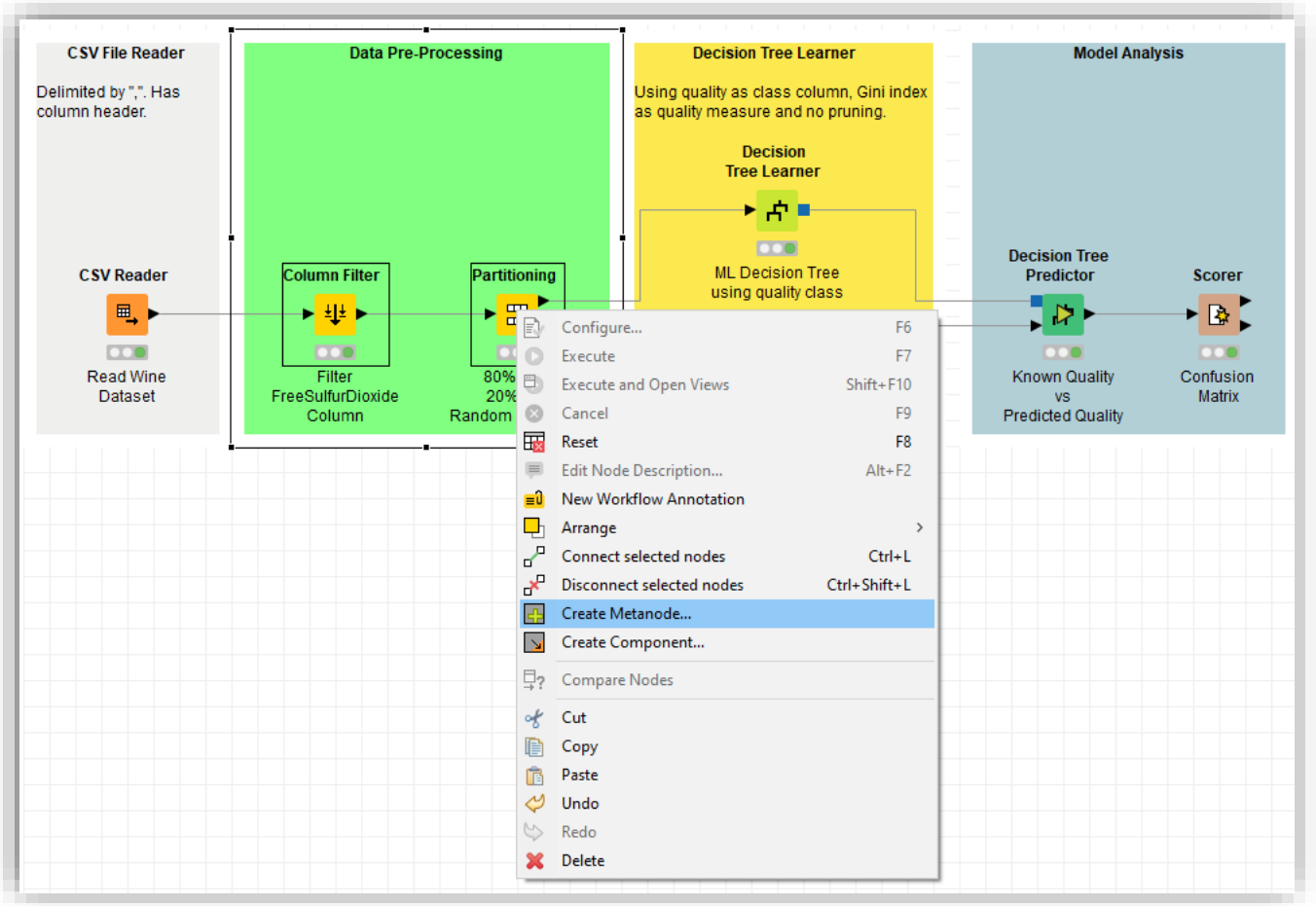
- Use Metanodes!





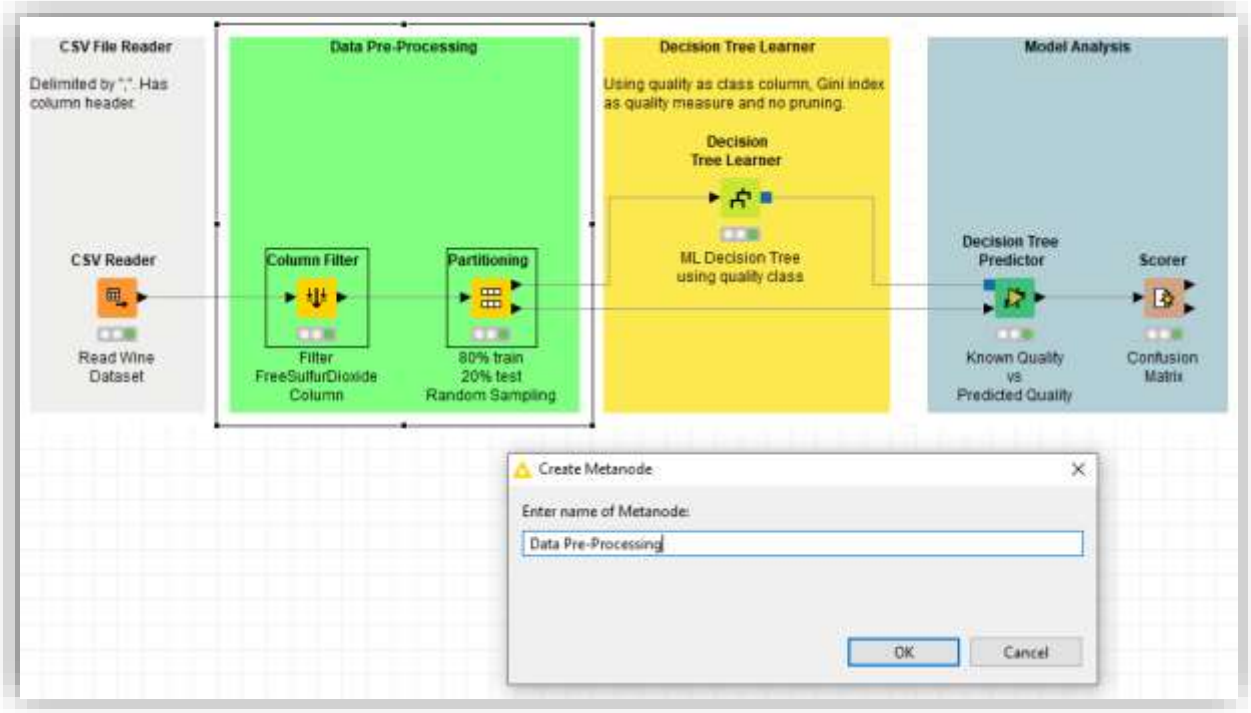
Metanodes

A Metanode is a node with other nodes inside! Use Metanodes for a Tidy Workflow!

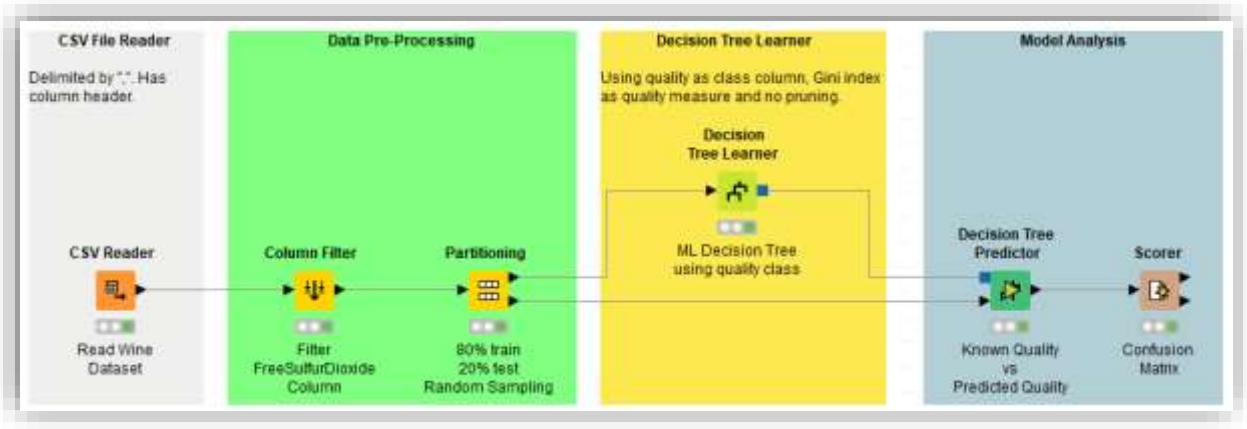


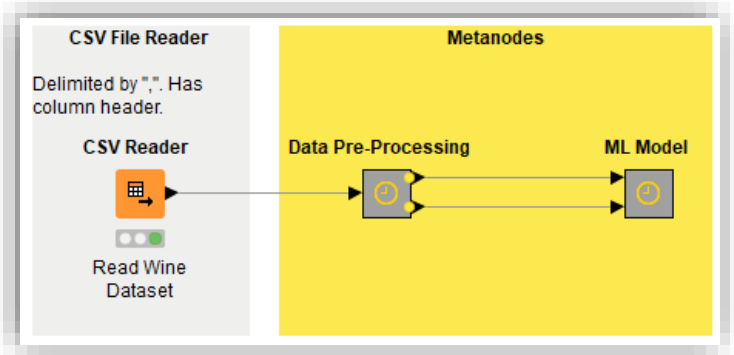
Metanodes

A Metanode is a node with other nodes inside! Use Metanodes for a Tidy Workflow!

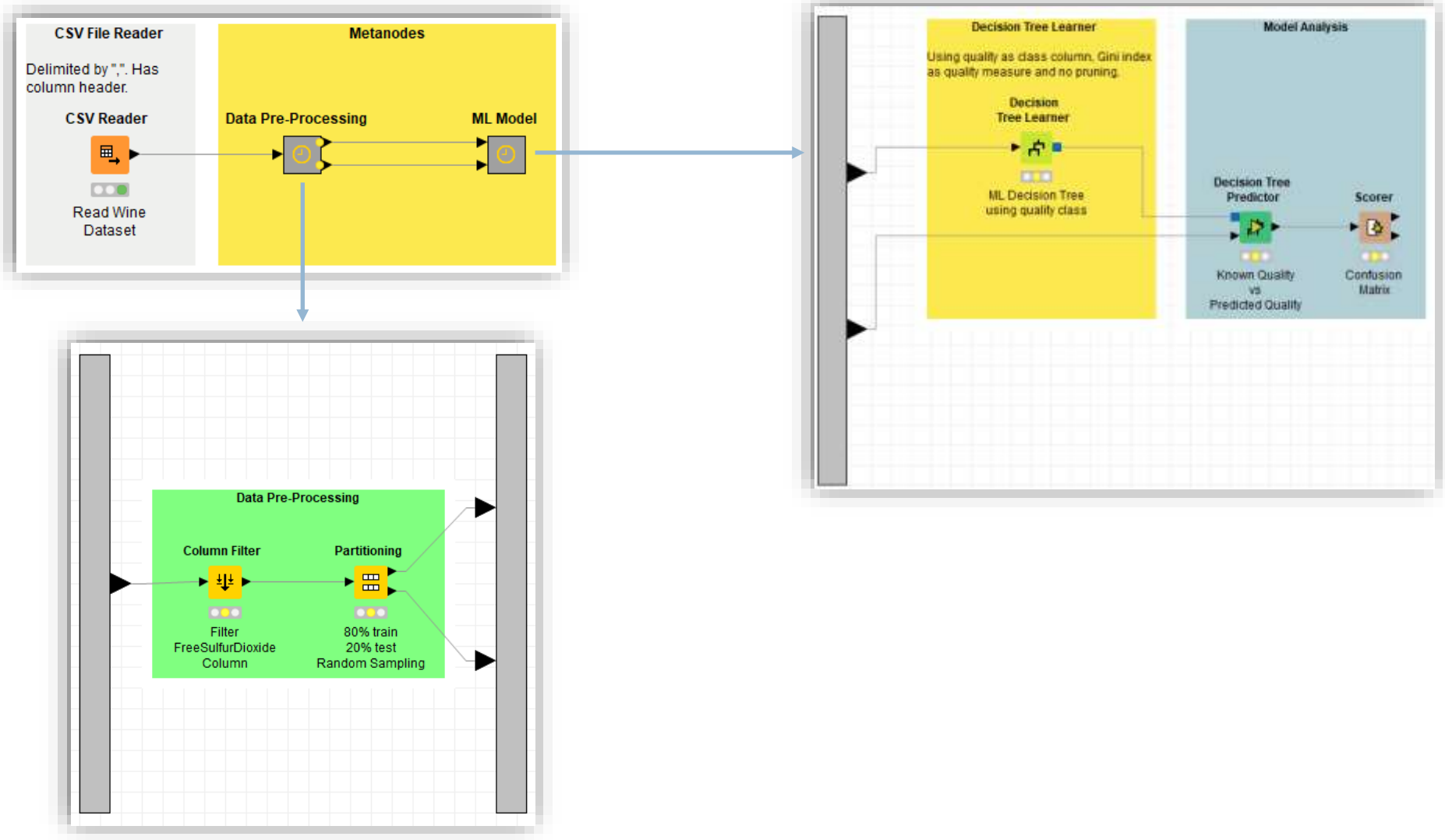


Metanodes

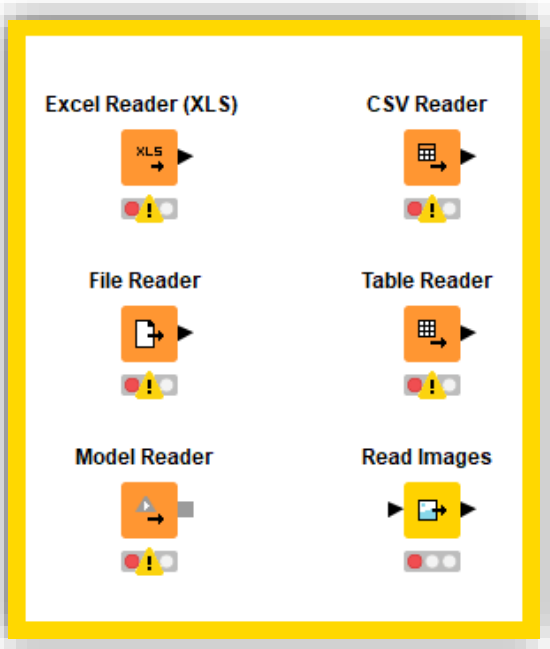

CSV File Reader
Delimited by ",". Has column header.
CSV Reader
Read Wine Dataset
Data Pre-Processing
Column Filter
Filter FreeSulfurDioxide Column
Partitioning
80% train 20% test Random Sampling
Decision Tree Learner
Using quality as class column, Gini index as quality measure and no pruning.
Decision Tree Learner
ML Decision Tree using quality class
Model Analysis
Decision Tree Predictor
Known Quality vs Predicted Quality
Scorer
Confusion Matrix


CSV File Reader
Delimited by ",". Has column header.
CSV Reader
Read Wine Dataset
Metanodes
Data Pre-Processing
ML Model

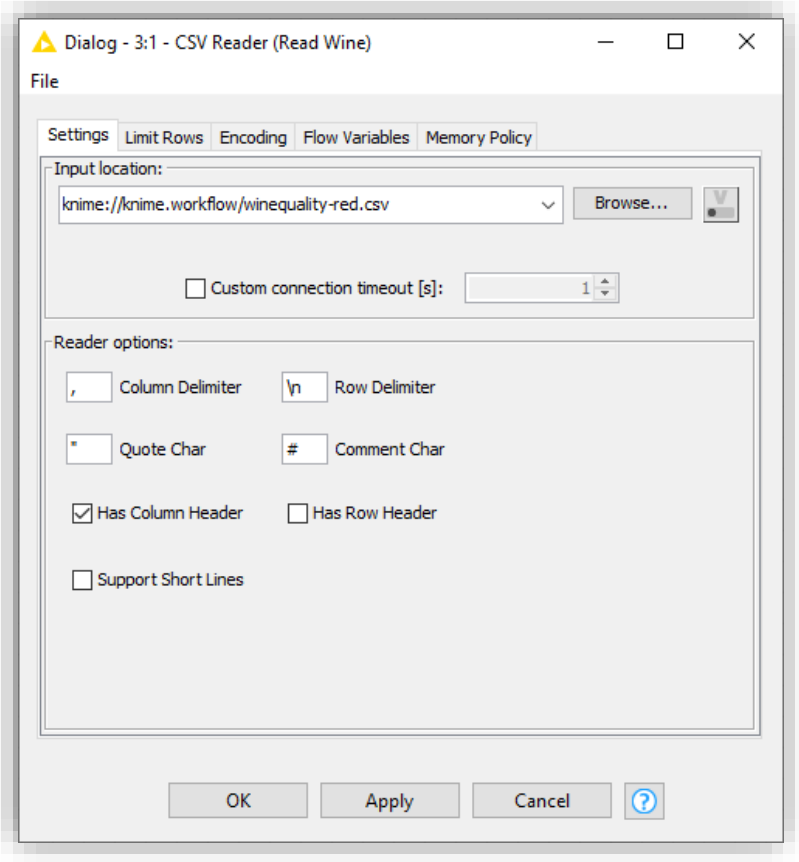
Metanodes


CSV File Reader
Delimited by ",". Has column header.
CSV Reader
Read Wine Dataset
Metanodes
Data Pre-Processing
ML Model
Decision Tree Learner
Using quality as class column, Gini index as quality measure and no pruning.
Decision Tree Learner
ML Decision Tree using quality class
Model Analysis
Decision Tree Predictor
Scorer
Known Quality vs Predicted Quality
Confusion Matrix
Data Pre-Processing
Column Filter
Partitioning
Filter FreeSulfurDioxide Column
80% train 20% test Random Sampling

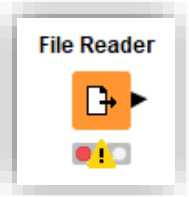
Data Readers

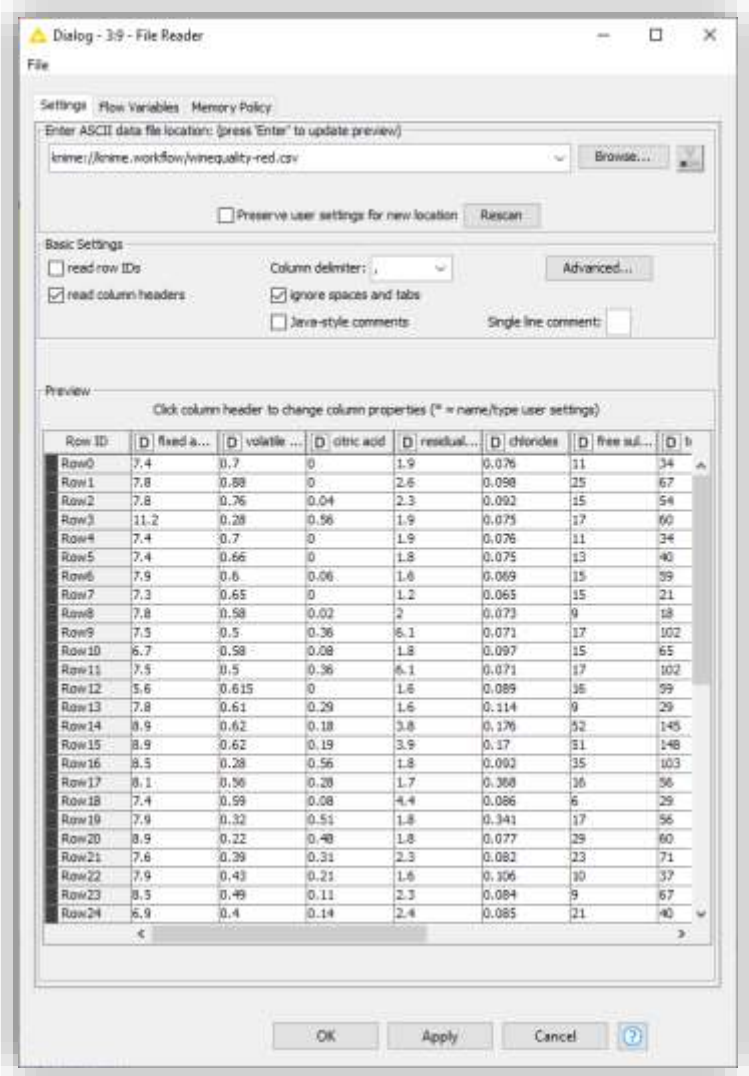


Data Readers

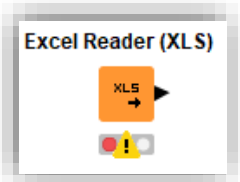


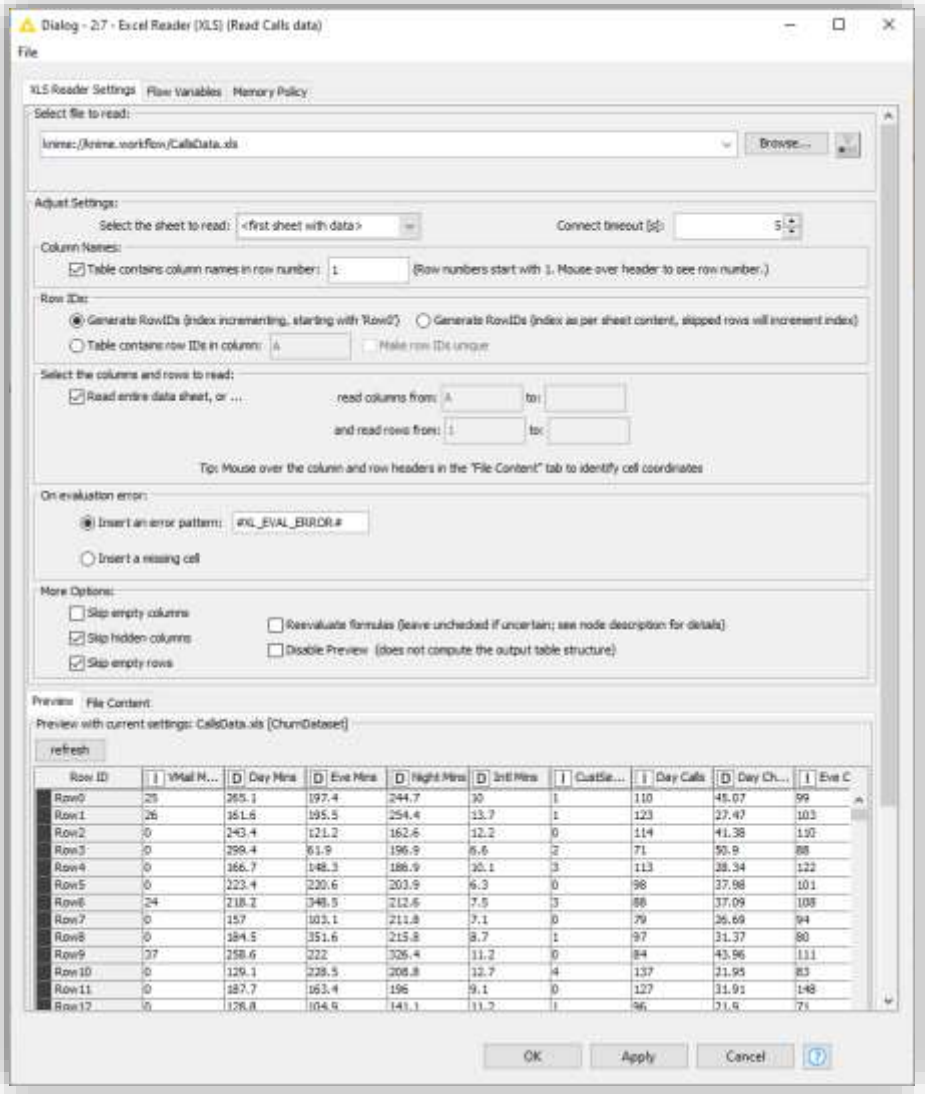
Data Readers



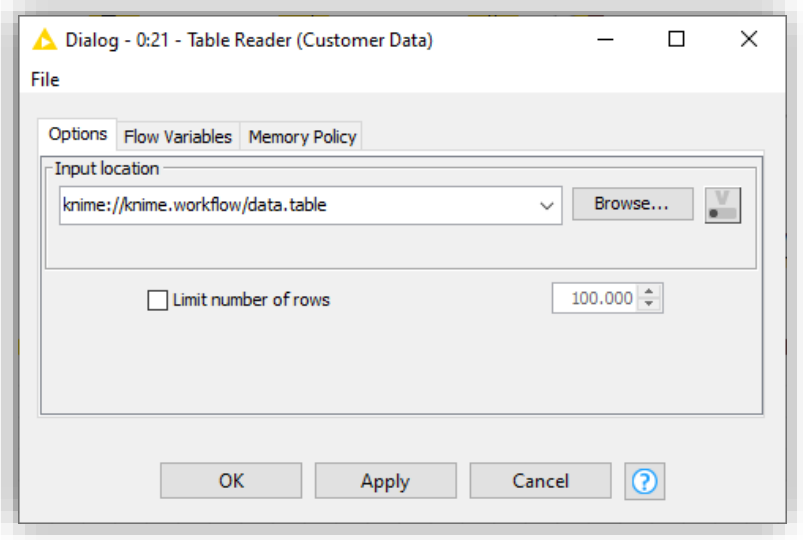


Data Readers





Data Readers



Loaded Data

File Table - 3:1 - CSV Reader (Read Wine)

File Hilite Navigation View

Table "winequality-red.csv" - Rows: 1599 | Spec - Columns: 12 | Properties | Flow Variables

Row ID	D fixed a...	D volatile ...	D citric acid	D residual...	D chlorides	D free sul...	D total su...	D density	D pH	D sulphates	D alcohol	S quality
Row0	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row1	7.8	0.88	0	2.6	0.098	25	67	0.997	3.2	0.68	9.8	=5
Row2	7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	=5
Row3	11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	=6
Row4	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row5	7.4	0.66	0	1.8	0.075	13	40	0.998	3.51	0.56	9.4	=5
Row6	7.9	0.6	0.06	1.6	0.069	15	59	0.996	3.3	0.46	9.4	=5
Row7	7.3	0.65	0	1.2	0.065	15	21	0.995	3.39	0.47	10	=7
Row8	7.8	0.58	0.02	2	0.073	9	18	0.997	3.36	0.57	9.5	=7
Row9	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row10	6.7	0.58	0.08	1.8	0.097	15	65	0.996	3.28	0.54	9.2	=5
Row11	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row12	5.6	0.615	0	1.6	0.089	16	59	0.994	3.58	0.52	9.9	=5
Row13	7.8	0.61	0.29	1.6	0.114	9	29	0.997	3.26	1.56	9.1	=5
Row14	8.9	0.62	0.18	3.8	0.176	52	145	0.999	3.16	0.88	9.2	=5
Row15	8.9	0.62	0.19	3.9	0.17	51	148	0.999	3.17	0.93	9.2	=5
Row16	8.5	0.28	0.56	1.8	0.092	35	103	0.997	3.3	0.75	10.5	=7
Row17	8.1	0.56	0.28	1.7	0.368	16	56	0.997	3.11	1.28	9.3	=5
Row18	7.4	0.59	0.08	4.4	0.086	6	29	0.997	3.38	0.5	9	=4
Row19	7.9	0.32	0.51	1.8	0.341	17	56	0.997	3.04	1.08	9.2	=6
Row20	8.9	0.22	0.48	1.8	0.077	29	60	0.997	3.39	0.53	9.4	=6
Row21	7.6	0.39	0.31	2.3	0.082	23	71	0.998	3.52	0.65	9.7	=5
Row22	7.9	0.43	0.21	1.6	0.106	10	37	0.997	3.17	0.91	9.5	=5
Row23	8.5	0.49	0.11	2.3	0.084	9	67	0.997	3.17	0.53	9.4	=5
Row24	6.9	0.4	0.14	2.4	0.085	21	40	0.997	3.43	0.63	9.7	=6

Loaded Data

Input Features/Input Vector

Target/Class/Label

File Table - 3:1 - CSV Reader (Read Wine)

File Hilite Navigation View

Table "winequality-red.csv" - Rows: 1599 | Spec - Columns: 12 | Properties | Flow Variables

Row ID	D fixed a...	D volatile ...	D citric acid	D residual...	D chlorides	D free sul...	D total su...	D density	D pH	D sulphates	D alcohol	S quality
Row0	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row1	7.8	0.88	0	2.6	0.098	25	67	0.997	3.2	0.68	9.8	=5
Row2	7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	=5
Row3	11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	=6
Row4	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row5	7.4	0.66	0	1.8	0.075	13	40	0.998	3.51	0.56	9.4	=5
Row6	7.9	0.6	0.06	1.6	0.069	15	59	0.996	3.3	0.46	9.4	=5
Row7	7.3	0.65	0	1.2	0.065	15	21	0.995	3.39	0.47	10	=7
Row8	7.8	0.58	0.02	2	0.073	9	18	0.997	3.36	0.57	9.5	=7
Row9	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row10	6.7	0.58	0.08	1.8	0.097	15	65	0.996	3.28	0.54	9.2	=5
Row11	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row12	5.6	0.615	0	1.6	0.089	16	59	0.994	3.58	0.52	9.9	=5
Row13	7.8	0.61	0.29	1.6	0.114	9	29	0.997	3.26	1.56	9.1	=5
Row14	8.9	0.62	0.18	3.8	0.176	52	145	0.999	3.16	0.88	9.2	=5
Row15	8.9	0.62	0.19	3.9	0.17	51	148	0.999	3.17	0.93	9.2	=5
Row16	8.5	0.28	0.56	1.8	0.092	35	103	0.997	3.3	0.75	10.5	=7
Row17	8.1	0.56	0.28	1.7	0.368	16	56	0.997	3.11	1.28	9.3	=5
Row18	7.4	0.59	0.08	4.4	0.086	6	29	0.997	3.38	0.5	9	=4
Row19	7.9	0.32	0.51	1.8	0.341	17	56	0.997	3.04	1.08	9.2	=6
Row20	8.9	0.22	0.48	1.8	0.077	29	60	0.997	3.39	0.53	9.4	=6
Row21	7.6	0.39	0.31	2.3	0.082	23	71	0.998	3.52	0.65	9.7	=5
Row22	7.9	0.43	0.21	1.6	0.106	10	37	0.997	3.17	0.91	9.5	=5
Row23	8.5	0.49	0.11	2.3	0.084	9	67	0.997	3.17	0.53	9.4	=5
Row24	6.9	0.4	0.14	2.4	0.085	21	40	0.997	3.43	0.63	9.7	=6

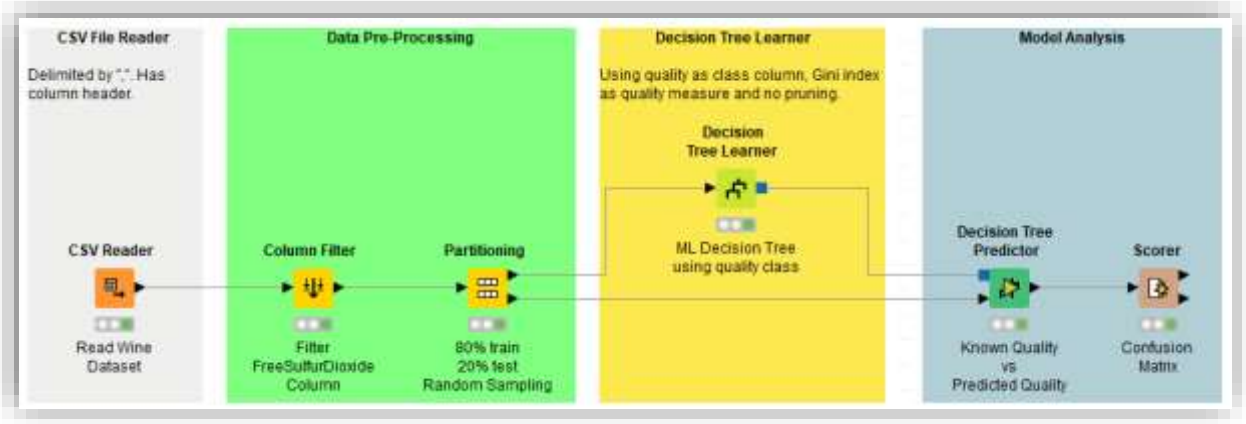
Data Partitioning/Segregation

File Table - 3:1 - CSV Reader (Read Wine)

File Hilite Navigation View

Table "winequality-red.csv" - Rows: 1599 | Spec - Columns: 12 | Properties | Flow Variables

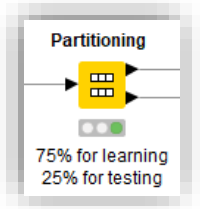
Row ID	fixed a...	volatile ...	citric acid	residual...	chlorides	free sul...	total su...	density	pH	sulphates	alcohol	quality
Row0	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row1	7.8	0.88	0	2.6	0.098	25	67	0.997	3.2	0.68	9.8	=5
Row2	7.8	0.76	0.04	2.3	0.092	15	54	0.997	3.26	0.65	9.8	=5
Row3	11.2	0.28	0.56	1.9	0.075	17	60	0.998	3.16	0.58	9.8	=6
Row4	7.4	0.7	0	1.9	0.076	11	34	0.998	3.51	0.56	9.4	=5
Row5	7.4	0.66	0	1.8	0.075	13	40	0.998	3.51	0.56	9.4	=5
Row6	7.9	0.6	0.06	1.6	0.069	15	59	0.996	3.3	0.46	9.4	=5
Row7	7.3	0.65	0	1.2	0.065	15	21	0.995	3.39	0.47	10	=7
Row8	7.8	0.58	0.02	2	0.073	9	18	0.997	3.36	0.57	9.5	=7
Row9	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row10	6.7	0.58	0.08	1.8	0.097	15	65	0.996	3.28	0.54	9.2	=5
Row11	7.5	0.5	0.36	6.1	0.071	17	102	0.998	3.35	0.8	10.5	=5
Row12	5.6	0.615	0	1.6	0.089	16	59	0.994	3.58	0.52	9.9	=5
Row13	7.8	0.61	0.29	1.6	0.114	9	29	0.997	3.26	1.56	9.1	=5
Row14	8.9	0.62	0.18	3.8	0.176	52	145	0.999	3.16	0.88	9.2	=5
Row15	8.9	0.62	0.19	3.9	0.17	51	148	0.999	3.17	0.93	9.2	=5
Row16	8.5	0.28	0.56	1.8	0.092	35	103	0.997	3.3	0.75	10.5	=7
Row17	8.1	0.56	0.28	1.7	0.368	16	56	0.997	3.11	1.28	9.3	=5
Row18	7.4	0.59	0.08	4.4	0.086	6	29	0.997	3.38	0.5	9	=4
Row19	7.9	0.32	0.51	1.8	0.341	17	56	0.997	3.04	1.08	9.2	=6
Row20	8.9	0.22	0.48	1.8	0.077	29	60	0.997	3.39	0.53	9.4	=6
Row21	7.6	0.39	0.31	2.3	0.082	23	71	0.998	3.52	0.65	9.7	=5
Row22	7.9	0.43	0.21	1.6	0.106	10	37	0.997	3.17	0.91	9.5	=5
Row23	8.5	0.49	0.11	2.3	0.084	9	67	0.997	3.17	0.53	9.4	=5
Row24	6.9	0.4	0.14	2.4	0.085	21	40	0.997	3.43	0.63	9.7	=6



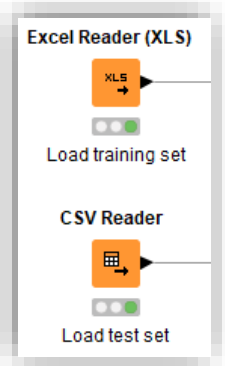
Data Partitioning/Segregation

Building a supervised ML model (data-driven):

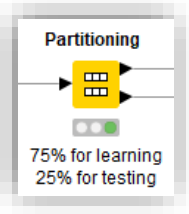
- Use a training set
  - To train the model with - used for learning
- Use a test set
  - To test the model with - used to evaluate the model on unseen data (unbiased evaluation)
- Whenever possibly, use a validation set as well
  - Provides an unbiased evaluation of a model fit on the training set while tuning the model's hyperparameters

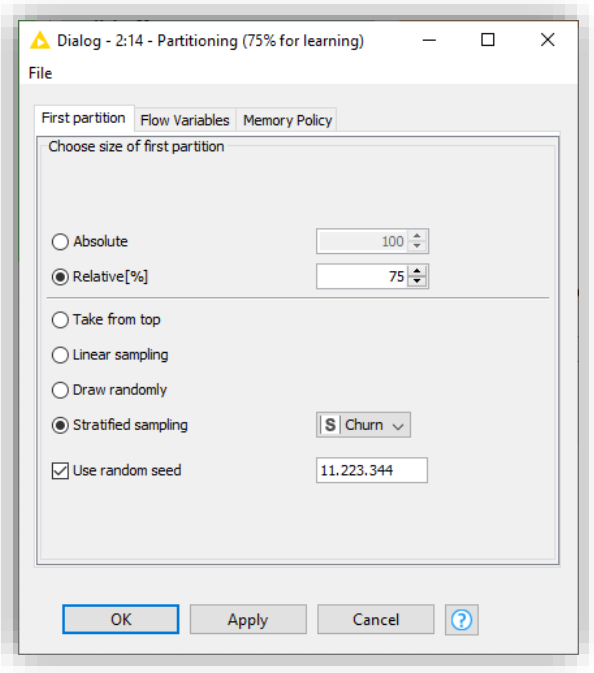


Or



Data Partitioning/Segregation





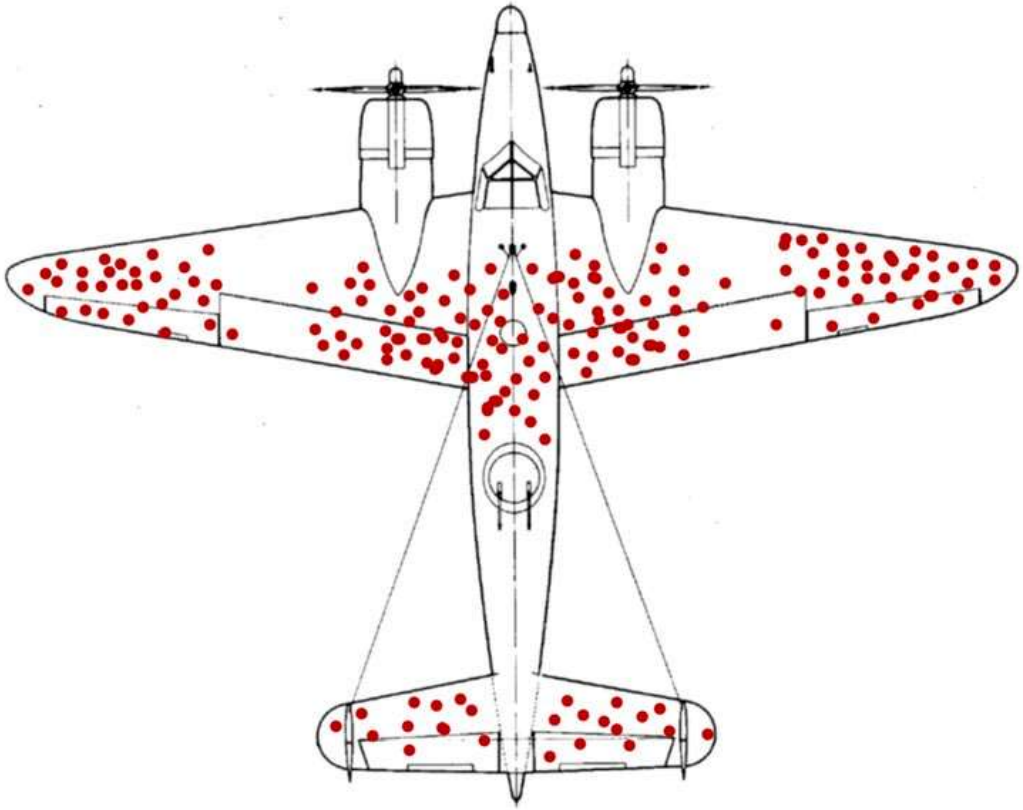
Data Quality

Think clearly…

During WWII, the US Navy tried to determine where they needed to armor their aircraft to ensure they came back home. They ran an analysis of where planes had been shot up.

Everybody told that, obviously, the places that needed to be up-armored are the wingtips, the central body, and the elevators. That's where the planes were all getting shot up!

Abraham Wald, a statistician, disagreed.

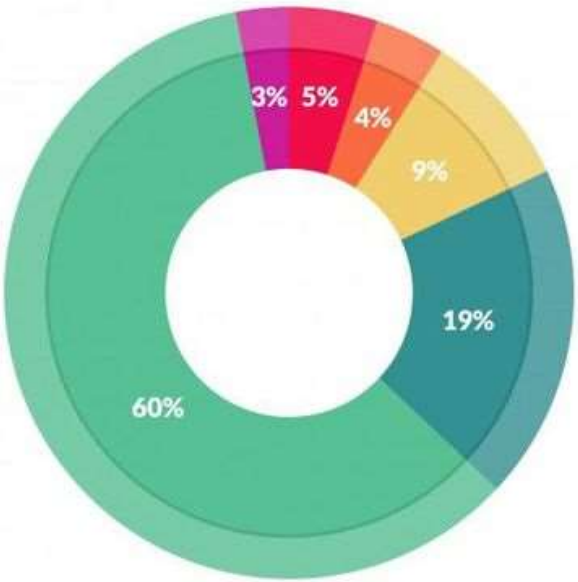


Why?

Data Quality

Indeed… Cleaning and manipulating data may be considered as the:

- Most Time-Consuming task
- Least Enjoyable task (by some!)



What data scientists spend the most time doing

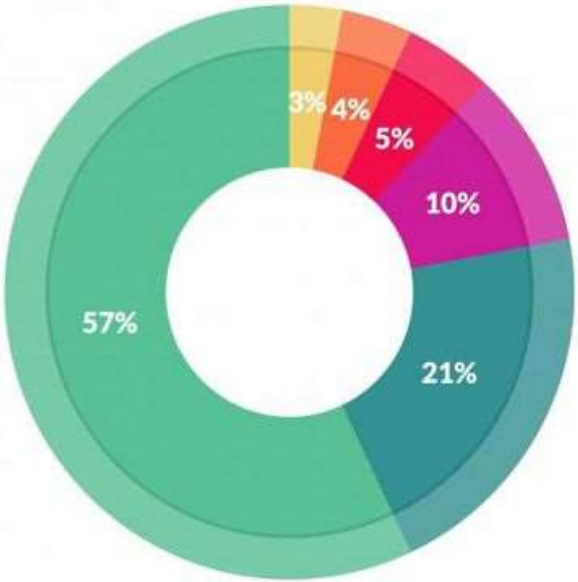
- *Building training sets: 3%*
- *Cleaning and organizing data: 60%*
- *Collecting data sets; 19%*
- *Mining data for patterns: 9%*
- *Refining algorithms: 4%*
- *Other: 5%*

(https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#1594bda36f63)

Data Quality

Indeed… Cleaning and manipulating data may be considered as the:

- Most Time-Consuming task
- Least Enjoyable task (by some!)



What's the least enjoyable part of data science?

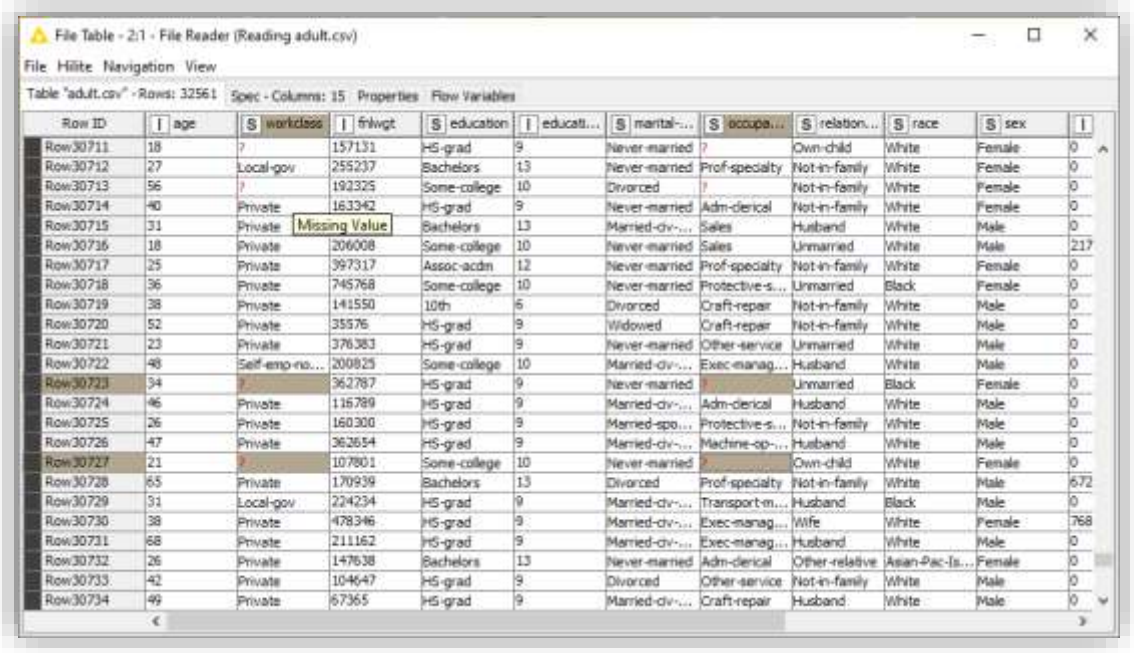
- *Building training sets: 10%*
- *Cleaning and organizing data: 57%*
- *Collecting data sets: 21%*
- *Mining data for patterns: 3%*
- *Refining algorithms: 4%*
- *Other: 5%*

(https://www.forbes.com/sites/gilpress/2016/03/23/data-preparation-most-time-consuming-least-enjoyable-data-science-task-survey-says/#1594bda36f63)

Data Quality

A few problems… How to solve them?

- Missing values
  - Information that is not available because it wasn't collected or because it consisted of sensitive information
  - Features that are not applicable in all cases
- Duplicated Records
  - Same (or similar) data collected from different sources

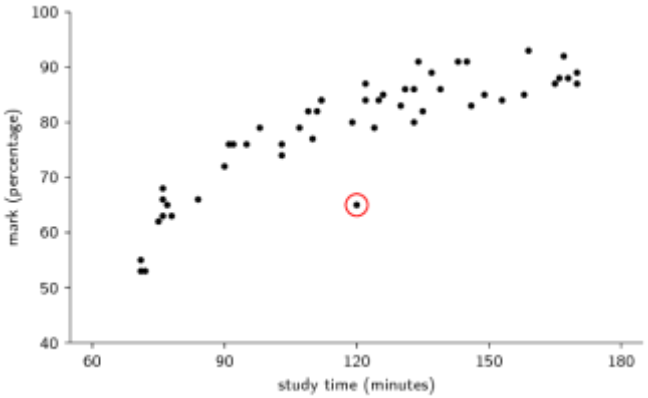


Row ID	age	workclass	fnlwgt	education	educati...	marital-...	occupa...	relation...	race	sex	
Row30711	18	?	157131	HS-grad	9	Never-married	?	Own-child	White	Female	0
Row30712	27	Local-gov	255237	Bachelors	13	Never-married	Prof-specialty	Not-in-family	White	Female	0
Row30713	56	?	192325	Some-college	10	Divorced	?	Not-in-family	White	Female	0
Row30714	40	Private	163342	HS-grad	9	Never-married	Adm-clerical	Not-in-family	White	Female	0
Row30715	31	Private	Missing Value	Bachelors	13	Married-civ-...	Sales	Husband	White	Male	0
Row30716	18	Private	206008	Some-college	10	Never-married	Sales	Unmarried	White	Male	217
Row30717	25	Private	397317	Assoc-acdn	12	Never-married	Prof-specialty	Not-in-family	White	Female	0
Row30718	36	Private	745768	Some-college	10	Never-married	Protective-s...	Unmarried	Black	Female	0
Row30719	38	Private	141550	10th	6	Divorced	Craft-repair	Not-in-family	White	Male	0
Row30720	52	Private	35576	HS-grad	9	Widowed	Craft-repair	Not-in-family	White	Male	0
Row30721	23	Private	376383	HS-grad	9	Never-married	Other-service	Unmarried	White	Male	0
Row30722	48	Self-emp-no...	200825	Some-college	10	Married-civ-...	Exec-manag...	Husband	White	Male	0
Row30723	34	?	362787	HS-grad	9	Never-married	?	Unmarried	Black	Female	0
Row30724	46	Private	116789	HS-grad	9	Married-civ-...	Adm-clerical	Husband	White	Male	0
Row30725	26	Private	160300	HS-grad	9	Married-spo...	Protective-s...	Not-in-family	White	Male	0
Row30726	47	Private	363654	HS-grad	9	Married-civ-...	Machine-op-...	Husband	White	Male	0
Row30727	21	?	107801	Some-college	10	Never-married	?	Own-child	White	Female	0
Row30728	65	Private	170939	Bachelors	13	Divorced	Prof-specialty	Not-in-family	White	Male	672
Row30729	31	Local-gov	224234	HS-grad	9	Married-civ-...	Transport-m...	Husband	Black	Male	0
Row30730	38	Private	478346	HS-grad	9	Married-civ-...	Exec-manag...	Wife	White	Female	768
Row30731	68	Private	211162	HS-grad	9	Married-civ-...	Exec-manag...	Husband	White	Male	0
Row30732	26	Private	147638	Bachelors	13	Never-married	Adm-clerical	Other-relative	Asian-Pac-Is...	Female	0
Row30733	42	Private	104647	HS-grad	9	Divorced	Other-service	Not-in-family	White	Male	0
Row30734	49	Private	67365	HS-grad	9	Married-civ-...	Craft-repair	Husband	White	Male	0

Data Quality

A few problems… How to solve them?

- Noise
  - Modifications to the original records (data that is corrupted or distorted) due to technological limitations, sensor error or even human error
- Outliers
  - A data point that differs significantly from other observations



Data Exploration

Why?

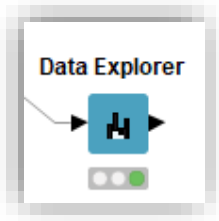
- Understand the data and its characteristics
- Evaluate its quality
- Find patterns and relevant information

Data Exploration

How?

- Central Tendency: average, mode, median…
- Statistical dispersion: variance, standard deviation, interquartile range…
- Probability distribution: Gaussian, Uniform, Exponential…
- Correlation/Dependence: between pairs of features, with the dependent feature…
- Data viz: tables, charts, boxplots, scatter plots, histograms, …

Data Explorer Node



Data Explorer View

Numeric | Nominal | Data Preview

Search:

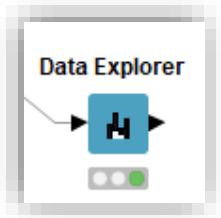
Column	Exclude Column	Minimum	Maximum	Mean	Standard Deviation	Variance	Skewness	Kurtosis	Overall Sum	No. zeros	No. missings
CustomerKey		11000	27336	17559.847	5576.039	31092215.201	0.333	-1.566	266330201	0	0
WebActivity		0	5	0.999	1.520	2.310	1.395	0.687	15159	9159	0
SentimentRating		0	5	1.851	1.620	2.624	0.482	-0.958	28073	4173	0
EstimatedYearlyIncome		10000	170000	57718.072	32091.910	1029890707.928	0.796	0.617	875410000	0	0
NumberOfContracts		0	4	1.465	1.145	1.311	0.430	-0.457	22227	3711	0
Age		29	100	48.203	11.300	127.694	0.571	-0.182	731101	0	0
Target		0	1	0.487	0.500	0.250	0.053	-1.997	7383	7784	0
Available401K		0	1	0.696	0.460	0.211	-0.854	-1.270	10562	4605	0
CustomerValueSegment		1	3	2.097	0.689	0.475	-0.129	-0.898	31809	0	0
ChurnScore		0	1	0.269	0.332	0.110	1.254	0.296	4078.300	5299	0
CallActivity		1	5	3.237	1.262	1.594	-0.302	-0.915	49094	0	0

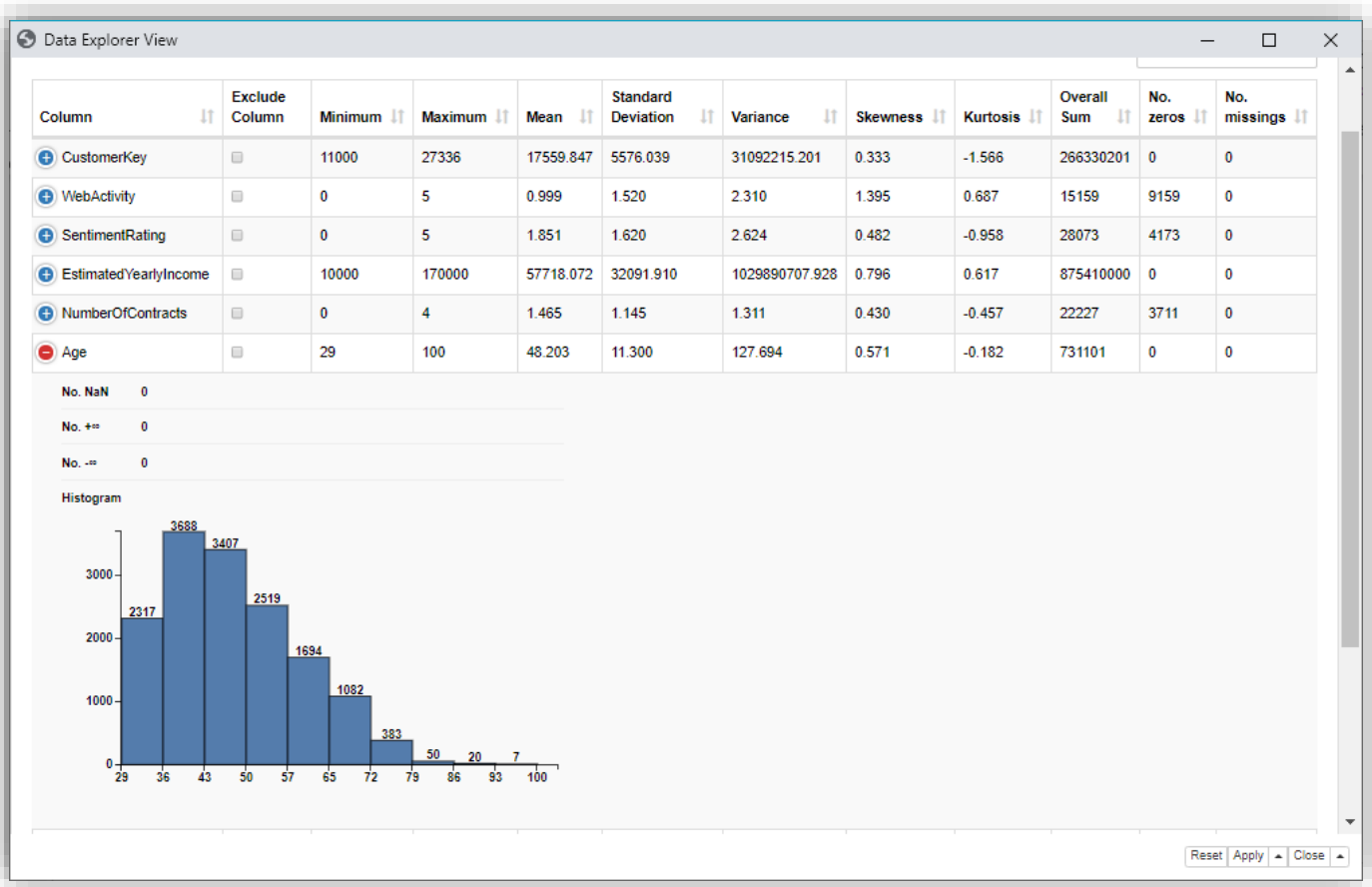
Showing 1 to 11 of 11 entries

Reset | Apply | Close

Install KNIME JavaScript Views (Labs) extension

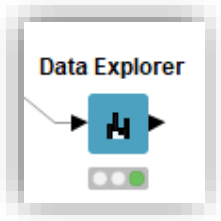
Data Explorer Node

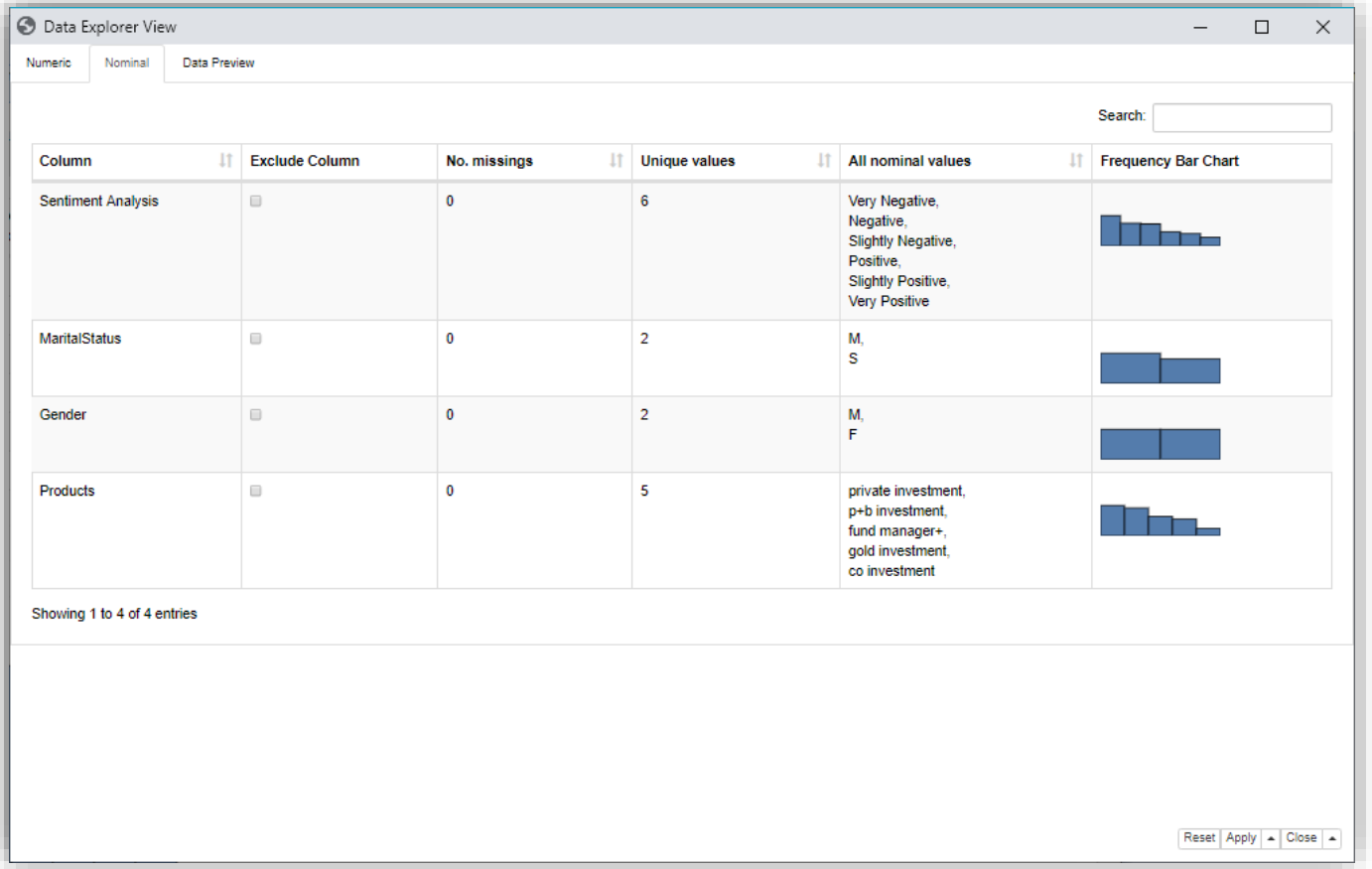




Install KNIME JavaScript Views (Labs) extension

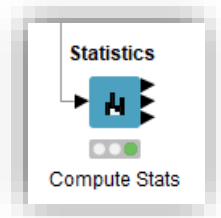
Data Explorer Node





Install KNIME JavaScript Views (Labs) extension

Statistics Node



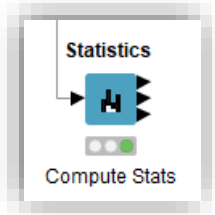
Statistics Table - 0:726 - Statistics (Compute Stats)

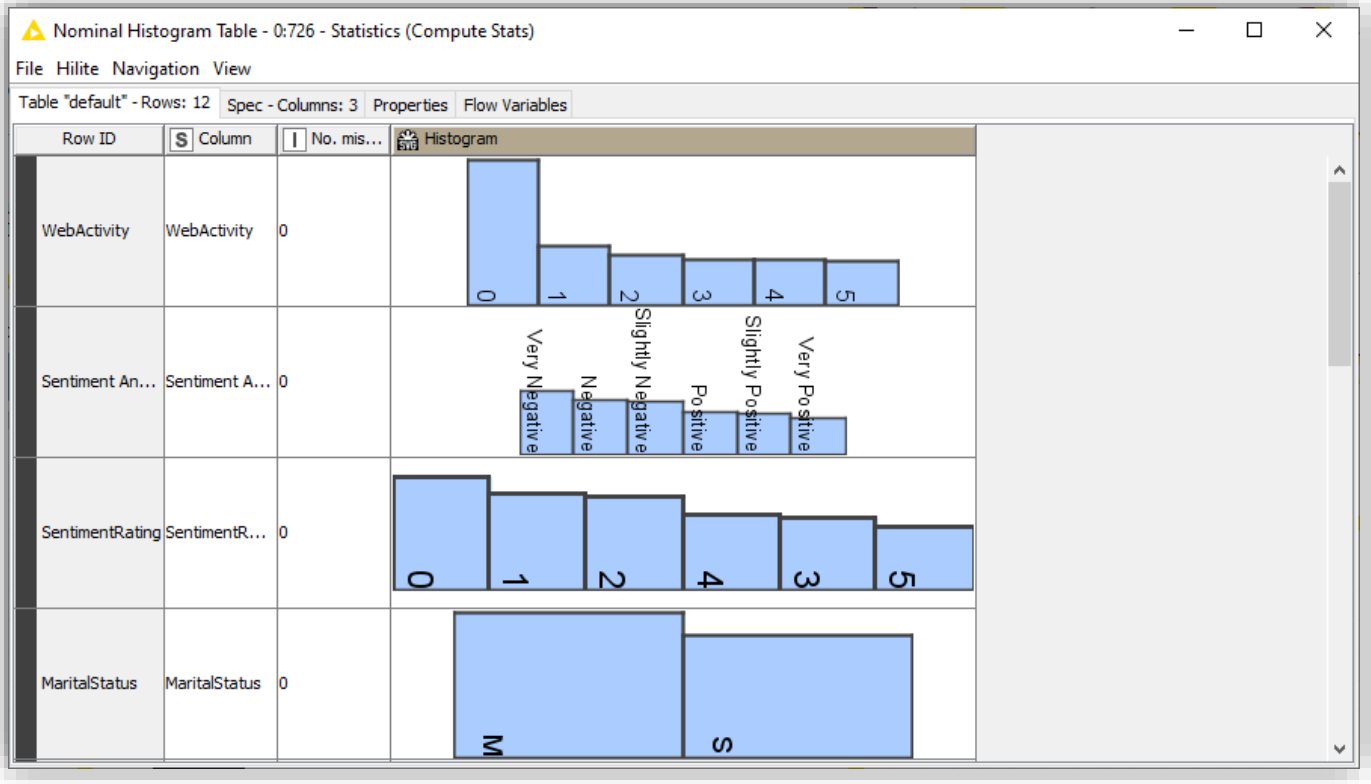
File Hilite Navigation View

Table "default" - Rows: 11 | Spec - Columns: 16 | Properties | Flow Variables

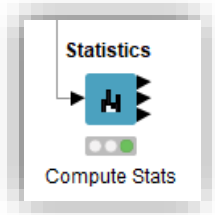
Row ID	Column	Min	Max	Mean	Std. deviation	Variance	Skewness	Kurtosis	Overall sum	No. missings	Nc
EstimatedYea...	EstimatedYe...	10,000	170,000	57,718.072	32,091.91	1,029,890,7...	0.796	0.617	875,410,000	0	0
NumberOfCo...	NumberOfC...	0	4	1.465	1.145	1.311	0.43	-0.457	22,227	0	0
Age	Age	29	100	48.203	11.3	127.694	0.571	-0.182	731,101	0	0
Target	Target	0	1	0.487	0.5	0.25	0.053	-1.997	7,383	0	0
Available401K	Available401K	0	1	0.696	0.46	0.211	-0.854	-1.27	10,562	0	0

Statistics Node





Statistics Node



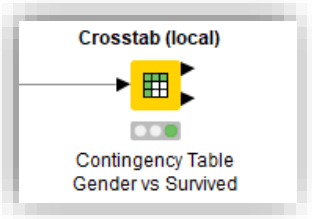
Occurrences Table - 0:726 - Statistics (Compute Stats)

File Hilite Navigation View

Table "default" - Rows: 68 | Spec - Columns: 36 | Properties | Flow Variables

Row ID	I WebActivity	I Count (WebActivity)	D Relative Frequency (WebActivity)	S Sentiment Analysis	I Count (Sentiment Analysis)	D Relati
Row0	0	9159	0.604	Very Negative	4173	0.275
Row1	1	1983	0.131	Negative	3122	0.206
Row2	2	1366	0.09	Slightly Negative	3023	0.199
Row3	3	963	0.063	Positive	1960	0.129
Row4	4	925	0.061	Slightly Positive	1690	0.111
Row5	5	771	0.051	Very Positive	1199	0.079
Row6	?	?	?	?	?	?
Row7	?	?	?	?	?	?
Row8	?	?	?	?	?	?
Row9	?	?	?	?	?	?
Row10	?	?	?	?	?	?
Row11	?	?	?	?	?	?
Row12	?	?	?	?	?	?
Row13	?	?	?	?	?	?
Row14	?	?	?	?	?	?
Row15	?	?	?	?	?	?
Row16	?	?	?	?	?	?
Row17	?	?	?	?	?	?
Row18	?	?	?	?	?	?
Row19	?	?	?	?	?	?
Row20	?	?	?	?	?	?
Row21	?	?	?	?	?	?

Contingency Tables



Cross Tabulation of Survived by Sex

Frequency	female	male	Total
0	81	468	549
1	233	109	342
Total	314	577	891

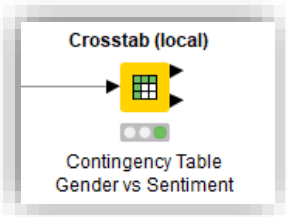
☑ Frequency
☐ Expected
☐ Deviation
☐ Percent
☐ Row Percent
☐ Column Percent
☐ Cell Chi-Square

Max rows: 10
Max columns: 10

Statistics for Table of Survived by Sex

Statistic	DF	Value	Prob
Chi-Square	1	263,0506	3,71E-59
Fisher's Exact Test (2-tail)			6,46E-60

Contingency Tables



Cross tabulation - 0:732 - Crosstab (local)

File

Frequency Percent	F	M	Total
Negative	1.585	1.537	3.122
	10,4503%	10,1338%	20,5842%
Positive	941	1.019	1.960
	6,2043%	6,7185%	12,9228%
Slightly Negative	1.501	1.522	3.023
	9,8965%	10,0349%	19,9314%
Slightly Positive	861	829	1.690
	5,6768%	5,4658%	11,1426%
Very Negative	2.054	2.119	4.173
	13,5426%	13,9711%	27,5137%
Very Positive	639	560	1.199
	4,2131%	3,6922%	7,9053%
Total	7.581	7.586	15.167
	49,9835%	50,0165%	100%

☑ Frequency
☐ Expected
☐ Deviation
☑ Percent
☐ Row Percent
☐ Column Percent
☐ Cell Chi-Square

Max rows: 10

Max columns: 10

Statistics for Table of Sentiment Analysis by Gender

Statistic	DF	Value	Prob
Chi-Square	5	10,8099	0,0553

Hands On


KNIME Analytics Platform
File Edit View Node Help
KNIME Explorer
My-KNIME-Hub (hub.knime.com)
EXAMPLES (knime@hub.knime.com)
LOCAL (Local Workspace)
Aulas Exemplos
Aulas Exercícios
Example Workflows
ML Projects
KNIME_Aula_Exercicio
Workflow Coach
Recommended Nodes
File Reader
CSV Reader
Excel Reader (XLS)
Table Creator
Database Reader (legacy)
Table Reader
List Files
Node Repository
IO
Manipulation
Views
Analytics
DB
Other Data Types
Structured Data
Scripting
Tools & Services
Workflow Control
Workflow Abstraction
Reporting
KNIME_Aula_Exercicio
KNIME Hub Search
Console
Outline
HANDS ON