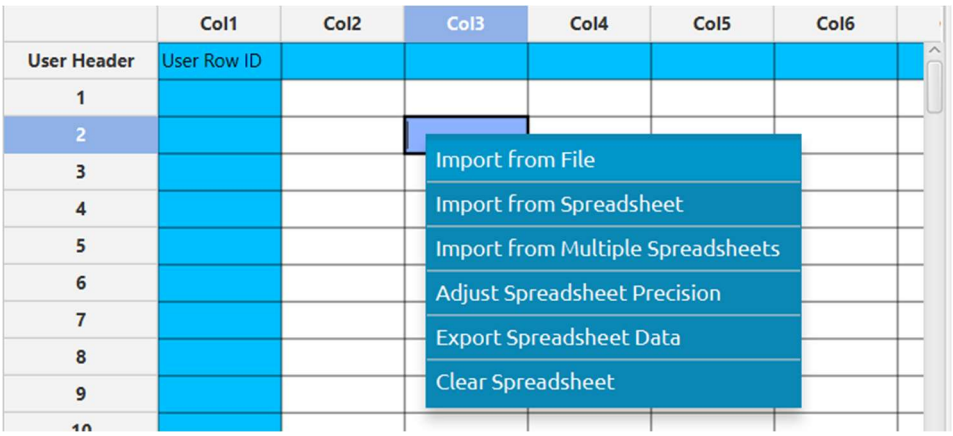
Student Performance (Multiple Linear Regression)

This dataset contains academic, demographic, and behavioural attributes from 10000 student records, aiming to analyse the factors influencing academic performance. Key variables include study hours, extracurricular activities, sleep hours, previous scores and other predictors. The data, which can be found in https://www.kaggle.com/datasets/nikhil7280/student-performance-multiple-linear-regression/data, is used to explore how different factors correlate with student performance outcomes, providing insights for educational strategies and policy development.

Isalos version used: 2.0.6

Step 1: Import data from file

Right click on the input spreadsheet (left) and choose the option “Import from File”. Then navigate through your files to load the one with the student performance scores data.

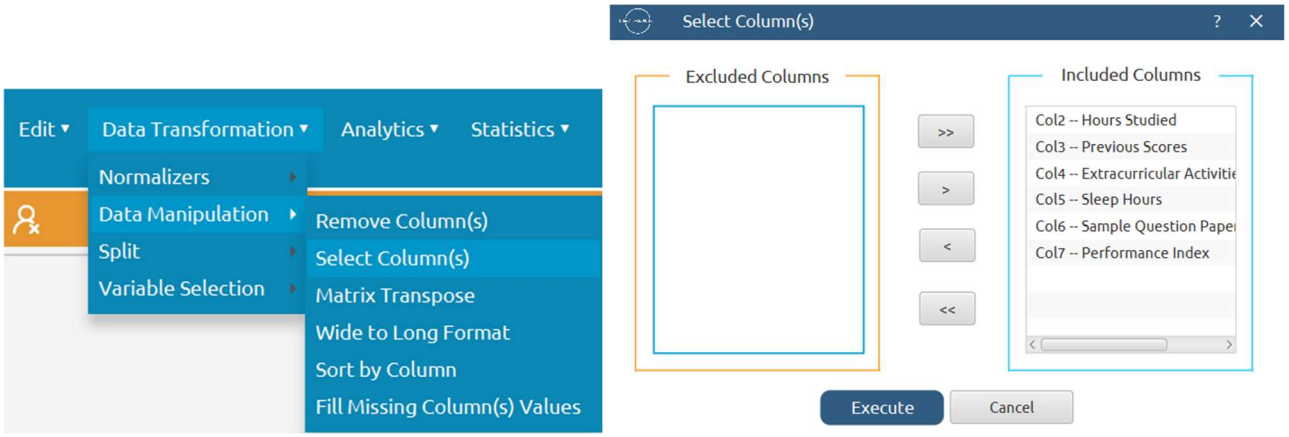


The data will appear on the left spreadsheet.

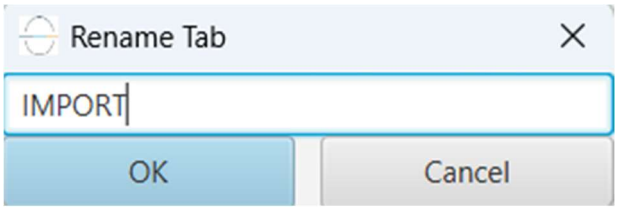
	Col1	Col2 (I)	Col3 (I)	Col4 (S)	Col5 (I)	Col6 (I)	Col7 (D)
User Header	User Row ID	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question Papers Practiced	Performance Index
1		7	99	Yes	9	1	91.0
2		4	82	No	4	2	65.0
3		8	51	Yes	7	2	45.0
4		5	52	Yes	5	2	36.0
5		7	75	No	8	5	66.0
6		3	78	No	9	6	61.0
7		7	73	Yes	5	6	63.0
8		8	45	Yes	4	6	42.0
9		5	77	No	8	2	61.0
10		4	89	No	4	0	69.0
11		8	91	No	4	5	84.0
12		8	79	No	6	2	73.0
13		3	47	No	9	2	27.0
14		6	47	No	4	2	33.0
15		5	79	No	7	8	68.0
16		2	72	No	4	3	43.0
17		8	73	Yes	8	4	67.0
18		6	83	Yes	7	2	70.0
19		2	54	Yes	4	9	30.0
20		5	75	No	7	0	63.0

Step 2: Manipulate data

In this dataset there are not any missing values, so we can select all the columns to be used. On the menu click on *Data Transformation → Data Manipulation → Select Column(s)* and select all columns.



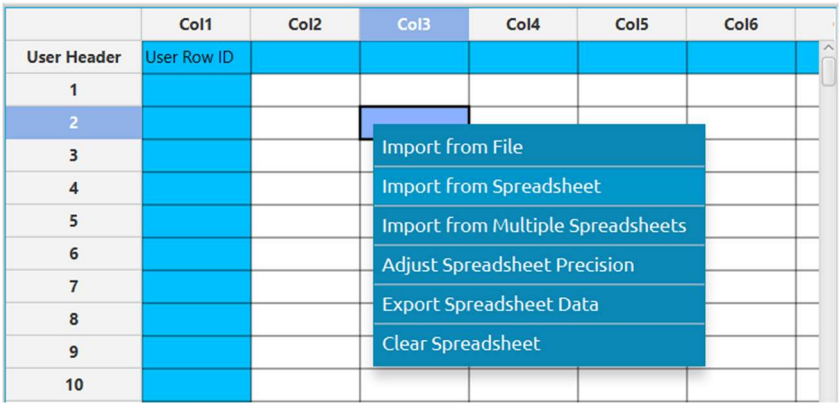
All of the data will appear in the output (right) spreadsheet. This tab can be renamed “IMPORT” by right-clicking on it and choosing the “Rename” option.



Step 3: Split data

Create a new tab by pressing the “+” button on the bottom of the page with the name “TRAIN_TEST_SPLIT” which we will use for splitting the train and test set.

Import data into the input spreadsheet of the “TRAIN_TEST_SPLIT” tab from the output of the “IMPORT” tab by right-clicking on the input spreadsheet and then choosing “Import from Spreadsheet”.



Split the dataset by choosing *Data Transformation → Split → Random Partitioning*. Then choose the “Training set percentage” and the column for the sampling as shown below:

The results will be two separate spreadsheets, “TRAIN_TEST_SPLIT: Training Set” and “TRAIN_TEST_SPLIT: Test Set”, which will be available to import into the next tabs.

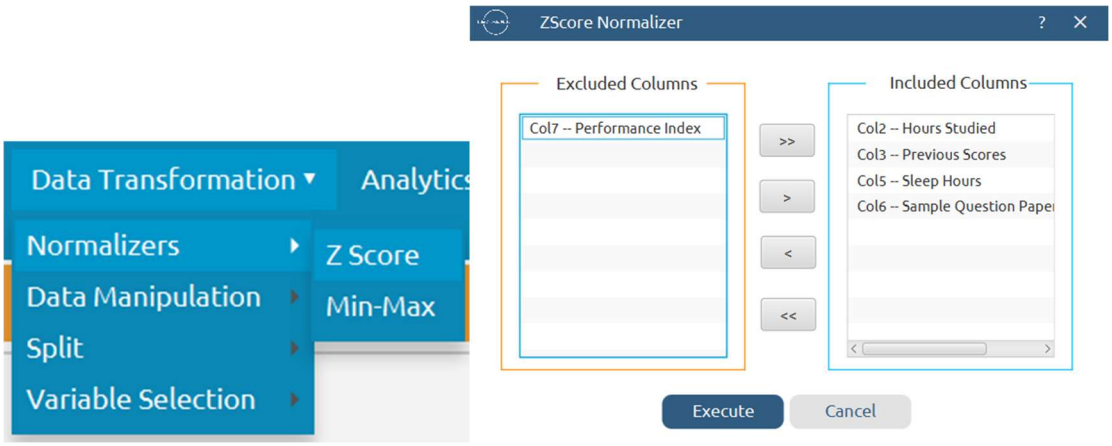
Step 4: Normalize the training set

Create a new tab by pressing the “+” button on the bottom of the page with the name “NORMALIZE_TRAIN_SET”.

Import into the input spreadsheet of the “NORMALIZE_TRAIN_SET” tab the train set from the output of the “TRAIN_TEST_SPLIT” tab by right-clicking on the input spreadsheet and then choosing “Import from Spreadsheet”. From the available Select input tab options choose “TRAIN_TEST_SPLIT: Training Set”.

	Col1	Col2 (I)	Col3 (I)	Col4 (S)	Col5 (I)	Col6 (I)	Col7 (D)
User Header	User Row ID	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question Papers Practiced	Performance Index
1		8	51	Yes	7	2	45.0
2		5	52	Yes	5	2	36.0
3		7	75	No	8	5	66.0
4		3	78	No	9	6	61.0
5		7	73	Yes	5	6	63.0
6		5	77	No	8	2	61.0
7		4	89	No	4	0	69.0
8		8	91	No	4	5	84.0
9		3	47	No	9	2	27.0
10		6	47	No	4	2	33.0
11		5	79	No	7	8	68.0
12		8	73	Yes	8	4	67.0
13		6	83	Yes	7	2	70.0
14		2	54	Yes	4	9	30.0
15		5	75	No	7	0	63.0
16		6	96	No	9	0	85.0
17		9	74	Yes	7	6	73.0
18		3	61	No	6	3	35.0
19		4	79	No	8	9	66.0
20		3	94	Yes	6	5	74.0

Normalize the data using Z-score: *Data Transformation → Normalizers → Z Score* and select all columns except the “Performance Index” target column.



The results will appear on the output spreadsheet.

	Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question Papers Practiced	Performance Index
1		1.1658115	-1.0607181	Yes	0.2725774	-0.9026160	45.0
2		0.0032552	-1.0032100	Yes	-0.9063905	-0.9026160	36.0
3		0.7782927	0.3194768	No	0.8620613	0.1407367	66.0
4		-0.7717824	0.4920011	No	1.4515452	0.4885209	61.0
5		0.7782927	0.2044605	Yes	-0.9063905	0.4885209	63.0
6		0.0032552	0.4344930	No	0.8620613	-0.9026160	61.0
7		-0.3842636	1.1245905	No	-1.4958744	-1.5981845	69.0
8		1.1658115	1.2396067	No	-1.4958744	0.1407367	84.0
9		-0.7717824	-1.2907506	No	1.4515452	-0.9026160	27.0
10		0.3907739	-1.2907506	No	-1.4958744	-0.9026160	33.0
11		0.0032552	0.5495093	No	0.2725774	1.1840894	68.0
12		1.1658115	0.2044605	Yes	0.8620613	-0.2070476	67.0
13		0.3907739	0.7795417	Yes	0.2725774	-0.9026160	70.0
14		-1.1593012	-0.8881938	Yes	-1.4958744	1.5318737	30.0
15		0.0032552	0.3194768	No	0.2725774	-1.5981845	63.0
16		0.3907739	1.5271473	No	1.4515452	-1.5981845	85.0
17		1.5533303	0.2619687	Yes	0.2725774	0.4885209	73.0
18		-0.7717824	-0.4856369	No	-0.3169066	-0.5548318	35.0
19		-0.3842636	0.5495093	No	0.8620613	1.5318737	66.0
20		-0.7717824	1.4121311	Yes	-0.3169066	0.1407367	74.0

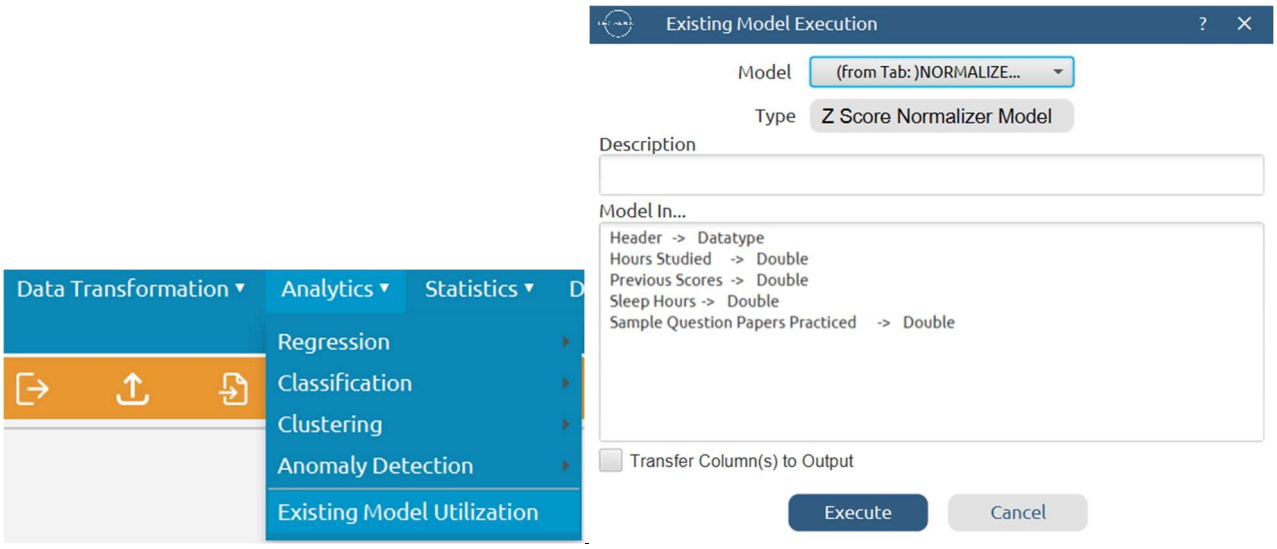
Step 5: Normalize the test set

Create a new tab by pressing the "+" button on the bottom of the page with the name "NORMALIZE_TEST_SET".

Import into the input spreadsheet of the "NORMALIZE_TEST_SET" tab the test set from the output of the "TRAIN_TEST_SPLIT" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet". From the available Select input tab options choose "TRAIN_TEST_SPLIT: Test Set".

	Col1	Col2 (I)	Col3 (I)	Col4 (S)	Col5 (I)	Col6 (I)	Col7 (D)
User Header	User Row ID	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question Papers Practiced	Performance Index
1		7	99	Yes	9	1	91.0
2		4	82	No	4	2	65.0
3		8	45	Yes	4	6	42.0
4		8	79	No	6	2	73.0
5		2	72	No	4	3	43.0
6		1	99	Yes	4	3	71.0
7		1	85	No	5	6	57.0
8		7	62	Yes	7	4	49.0
9		9	84	Yes	6	6	83.0
10		5	70	Yes	6	9	58.0
11		9	72	No	8	2	68.0
12		2	63	Yes	6	0	39.0
13		4	73	Yes	7	0	58.0
14		3	76	Yes	4	3	54.0
15		6	62	Yes	9	0	52.0
16		4	93	No	8	3	78.0
17		2	70	Yes	5	8	47.0
18		8	92	Yes	4	7	87.0
19		5	88	No	5	2	71.0
20		8	60	No	4	7	54.0

Normalize the test set using the existing normalizer of the training set: *Analytics → Existing Model Utilization → Model (from Tab:) NORMALIZE_TRAIN_SET*



The results will appear on the output spreadsheet.

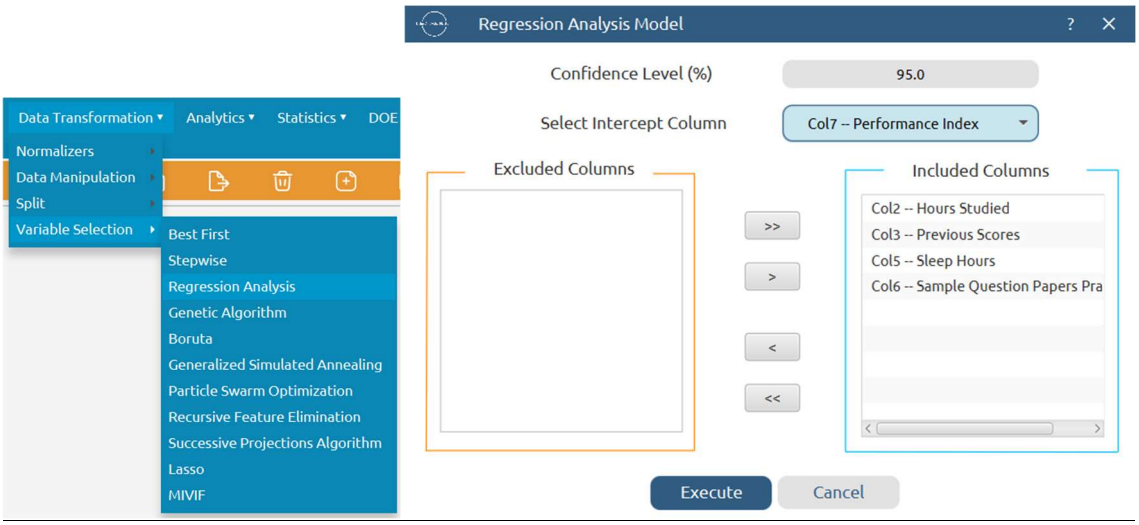
	Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question Papers Practiced	Performance Index
1		0.7782927	1.6996717	Yes	1.4515452	-1.2504003	91.0
2		-0.3842636	0.7220336	No	-1.4958744	-0.9026160	65.0
3		1.1658115	-1.4057668	Yes	-1.4958744	0.4885209	42.0
4		1.1658115	0.5495093	No	-0.3169066	-0.9026160	73.0
5		-1.1593012	0.1469524	No	-1.4958744	-0.5548318	43.0
6		-1.5468200	1.6996717	Yes	-1.4958744	-0.5548318	71.0
7		-1.5468200	0.8945580	No	-0.9063905	0.4885209	57.0
8		0.7782927	-0.4281288	Yes	0.2725774	-0.2070476	49.0
9		1.5533303	0.8370499	Yes	-0.3169066	0.4885209	83.0
10		0.0032552	0.0319362	Yes	-0.3169066	1.5318737	58.0
11		1.5533303	0.1469524	No	0.8620613	-0.9026160	68.0
12		-1.1593012	-0.3706207	Yes	-0.3169066	-1.5981845	39.0
13		-0.3842636	0.2044605	Yes	0.2725774	-1.5981845	58.0
14		-0.7717824	0.3769849	Yes	-1.4958744	-0.5548318	54.0
15		0.3907739	-0.4281288	Yes	1.4515452	-1.5981845	52.0
16		-0.3842636	1.3546229	No	0.8620613	-0.5548318	78.0
17		-1.1593012	0.0319362	Yes	-0.9063905	1.1840894	47.0
18		1.1658115	1.2971148	Yes	-1.4958744	0.8363052	87.0
19		0.0032552	1.0670823	No	-0.9063905	-0.9026160	71.0
20		1.1658115	-0.5431450	No	-1.4958744	0.8363052	54.0

Step 6: Feature selection

Create a new tab by pressing the "+" button on the bottom of the page with the name "FEATURE_SELECTION_REGRESSION".

Import data into the input spreadsheet of the "FEATURE_SELECTION_REGRESSION" tab from the output of the "NORMALIZE_TRAIN_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Then do regression analysis with the "Performance Index" column as the intercept: *Data Transformation → Variable Selection → Regression Analysis*



The results will appear on the right spreadsheet.

	Col1	Col2 (S)	Col3 (S)	Col4 (S)	Col5 (S)	Col6 (S)	Col7 (S)
User Header	User Row ID						
1		Regression Statistics					
2		Multiple R	0.9941838				
3		R Square	0.9884015				
4		Adjusted R Square	0.9883953				
5		Standard Error	2.0690397				
6		Observations	7500				
7							
8			Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	Significance F
9		Regression	4	2734257.6801934	683564.4200484	159676.7915708	0.0
10		Residual	7495	32085.5352732	4.2809253		
11		Total	7499	2766343.2154667			

	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	Significance F	
Regression	4	2734257.6801934	683564.4200484	159676.7915708	0.0	
Residual	7495	32085.5352732	4.2809253			
Total	7499	2766343.2154667				
	Coefficients	Standard Error	t-statistic	P-value	Lower 95.0%	Upper 95.0%
Performance Index	55.2162667	0.0238912	2311.1537886	0.0	55.1694332	55.2631001
Hours Studied	7.3439405	0.0239031	307.2379113	0.0	7.2970837	7.3907973
Previous Scores	17.7149981	0.0238994	741.2322449	0.0	17.6681486	17.7618476
Sleep Hours	0.8064854	0.0238951	33.7510806	0E-7	0.7596443	0.8533265
Sample Question Papers Practiced	0.5670981	0.0238990	23.7289168	0E-7	0.5202493	0.6139470

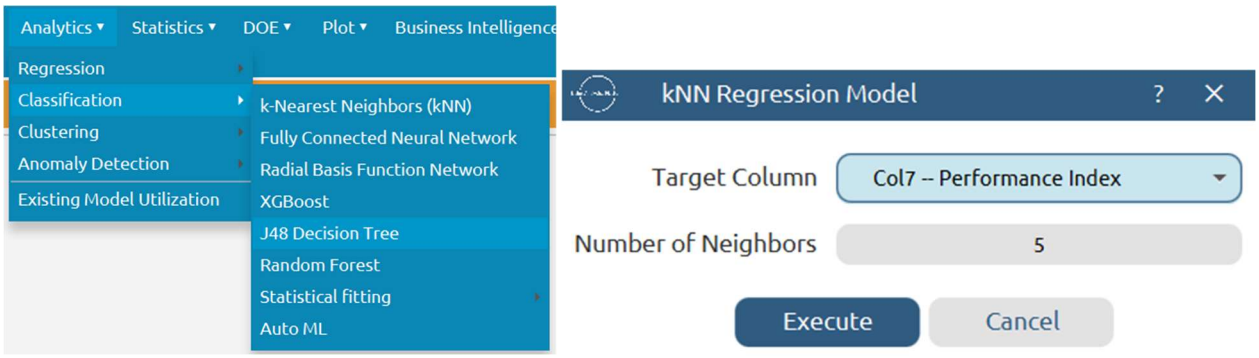
All the features are significant according to the p-value.

Step 7: Train the model

Create a new tab by pressing the "+" button on the bottom of the page with the name "TRAIN_MODEL(.fit)".

Import data into the input spreadsheet of the "TRAIN_MODEL(.fit)" tab from the output of the "NORMALIZE_TRAIN_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Use the k-Nearest Neighbors (kNN) method to train and fit the model: *Analytics → Regression → k-Nearest Neighbors (kNN)*



The predictions will appear on the output spreadsheet.

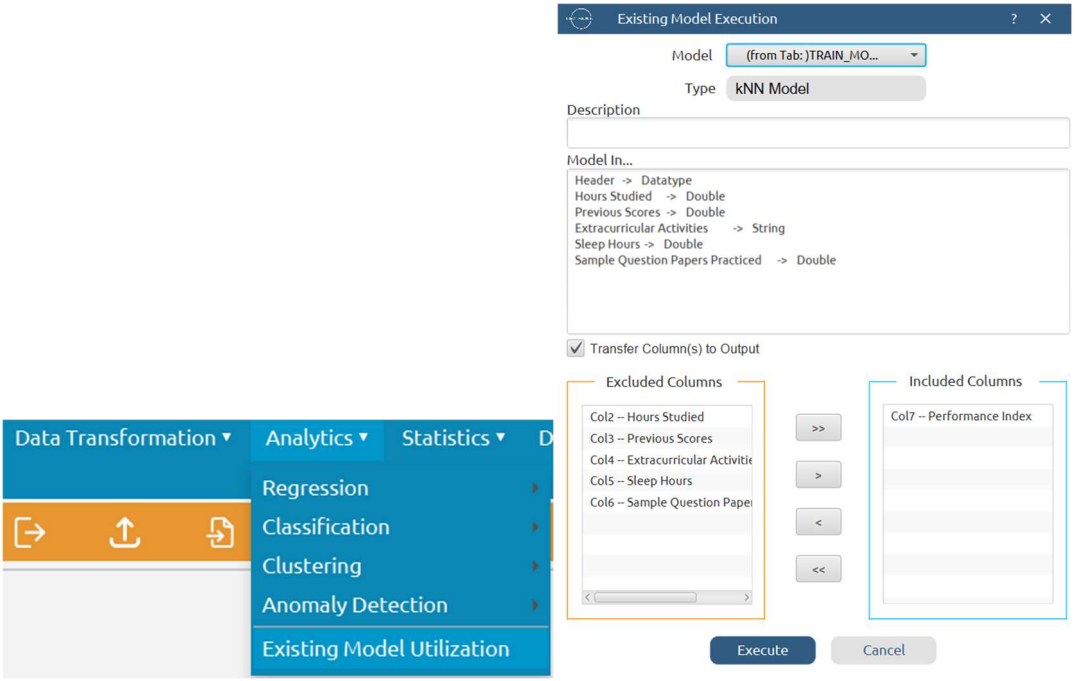
	Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (S)	Col7 (D)	Col8 (S)	Col9 (D)	Col10 (S)	Col11 (D)	Col12 (S)	Col13 (D)
User Header	User Row ID	Performance Index	kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Distance from NN3	Closest NN4	Distance from NN4	Closest NN5	Distance from NN5
1		45.0	44.9682252	Entry 1	0.0	Entry 1163	0.1221931	Entry 2282	0.1397409	Entry 1500	0.1422019	Entry 2152	0.1506238
2		36.0	36.1409791	Entry 2	0.0	Entry 271	0.0677966	Entry 425	0.0847458	Entry 5419	0.1016949	Entry 6237	0.1111111
3		66.0	66.0543745	Entry 3	0.0	Entry 3832	0.0508475	Entry 3314	0.0508475	Entry 3594	0.1016949	Entry 366	0.1111111
4		61.0	60.9880952	Entry 4	0.0	Entry 2634	0.0847458	Entry 5803	0.1111111	Entry 7476	0.1123964	Entry 791	0.1250000
5		63.0	63.1399800	Entry 5	0.0	Entry 1161	0.0338983	Entry 6383	0.1123964	Entry 7383	0.1123964	Entry 3939	0.1123964
6		61.0	61.0323083	Entry 6	0.0	Entry 924	0.1123964	Entry 3079	0.1186441	Entry 5023	0.1250000	Entry 515	0.1261439
7		69.0	68.7836561	Entry 7	0.0	Entry 5240	0.0169492	Entry 5657	0.0338983	Entry 6306	0.1123964	Entry 3141	0.1506238
8		84.0	84.0260285	Entry 8	0.0	Entry 1568	0.1111111	Entry 1871	0.1123964	Entry 5157	0.1123964	Entry 4966	0.1161670
9		27.0	27.4386584	Entry 1192	0.0	Entry 9	0.0	Entry 4473	0.1261439	Entry 3893	0.1295148	Entry 499	0.1397409
10		33.0	33.1516631	Entry 10	0.0	Entry 6399	0.0169492	Entry 3941	0.1261439	Entry 5970	0.1397409	Entry 3236	0.1422019
11		68.0	67.8947507	Entry 11	0.0	Entry 658	0.1016949	Entry 4488	0.1161670	Entry 7463	0.1261439	Entry 3526	0.1261439
12		67.0	67.1886330	Entry 12	0.0	Entry 3078	0.1016949	Entry 7396	0.1161670	Entry 2298	0.1186441	Entry 1025	0.1295148
13		70.0	70.0879286	Entry 13	0.0	Entry 5131	0.0169492	Entry 3808	0.1111111	Entry 819	0.1111111	Entry 6038	0.1397409
14		30.0	29.9752947	Entry 14	0.0	Entry 6951	0.1295148	Entry 6247	0.1706452	Entry 3324	0.1753032	Entry 3865	0.1844194
15		63.0	62.7730809	Entry 15	0.0	Entry 447	0.0338983	Entry 1605	0.0508475	Entry 5991	0.0677966	Entry 1727	0.1161670
16		85.0	84.9702741	Entry 16	0.0	Entry 2907	0.0847458	Entry 2181	0.1123964	Entry 4037	0.1295148	Entry 1563	0.1349462
17		73.0	72.5199378	Entry 4753	0.0	Entry 17	0.0	Entry 2493	0.1186441	Entry 5715	0.1221931	Entry 6134	0.1295148
18		35.0	38.0175316	Entry 4241	0.0	Entry 18	0.0	Entry 989	0.0338983	Entry 4167	0.0677966	Entry 7280	0.0847458
19		66.0	65.8745870	Entry 19	0.0	Entry 2597	0.0677966	Entry 78	0.0847458	Entry 1391	0.1186441	Entry 251	0.1186441
20		74.0	74.0038989	Entry 20	0.0	Entry 5140	0.0169492	Entry 843	0.0508475	Entry 211	0.0508475	Entry 4469	0.1221931

Step 8: Validate the model

Create a new tab by pressing the "+" button on the bottom of the page with the name "VALIDATE_MODEL(.predict)".

Import data into the input spreadsheet of the "VALIDATE_MODEL(.predict)" tab from the output of the "NORMALIZE_TEST_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

To validate the model: *Analytics → Existing Model Utilization→ Model (from Tab:) TRAIN_MODEL(.fit)*. Choose the column "Performance Index" to be transferred to the output spreadsheet.



The predictions will appear on the output spreadsheet.

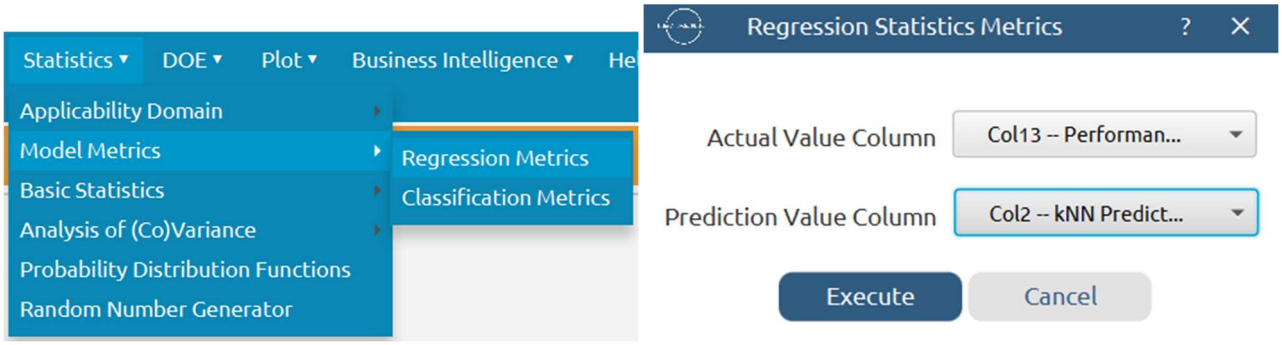
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (D)	Col7 (S)	Col8 (D)	Col9 (S)	Col10 (D)	Col11 (S)	Col12 (D)	Col13 (D)
User Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Distance from NN3	Closest NN4	Distance from NN4	Closest NN5	Distance from NN5	Performance Index
1		89.8674608	Entry 1322	0.0508475	Entry 191	0.1261439	Entry 1885	0.1355932	Entry 4487	0.1706452	Entry 3618	0.1748032	91.0
2		63.5932216	Entry 2390	0.0508475	Entry 2521	0.0677966	Entry 3141	0.1123964	Entry 2438	0.1123964	Entry 3762	0.1161670	65.0
3		40.1887821	Entry 2820	0.1111111	Entry 1050	0.1397409	Entry 1720	0.1681010	Entry 3883	0.1706452	Entry 7005	0.1753032	42.0
4		71.2655578	Entry 3818	0.0338983	Entry 2522	0.1261439	Entry 4261	0.1349462	Entry 4323	0.1706452	Entry 4960	0.1706452	73.0
5		47.4523066	Entry 1491	0.1016949	Entry 5638	0.1111111	Entry 2489	0.1123964	Entry 6804	0.1161670	Entry 4408	0.1221931	43.0
6		72.9518050	Entry 4222	0.0	Entry 6066	0.1221931	Entry 3637	0.1681010	Entry 5487	0.1706452	Entry 5340	0.1723410	71.0
7		61.2719098	Entry 5592	0.0169492	Entry 5667	0.0677966	Entry 1789	0.1111111	Entry 4979	0.1261439	Entry 5135	0.1301617	57.0
8		58.8462077	Entry 7048	0.0	Entry 6288	0.1295148	Entry 4337	0.1295148	Entry 3151	0.1349462	Entry 4411	0.1506238	49.0
9		80.7545137	Entry 5376	0.1123964	Entry 4686	0.1123964	Entry 7147	0.1397409	Entry 1451	0.1397409	Entry 1134	0.1506238	83.0
10		55.0377706	Entry 3552	0.0847458	Entry 7320	0.1186441	Entry 6228	0.1186441	Entry 4231	0.1397409	Entry 1773	0.1422019	58.0
11		69.9363620	Entry 5739	0.0508475	Entry 90	0.0677966	Entry 4668	0.0677966	Entry 6569	0.1221931	Entry 1926	0.1349462	68.0
12		35.9162903	Entry 6190	0.1016949	Entry 2051	0.1261439	Entry 2785	0.1397409	Entry 456	0.1397409	Entry 2541	0.1422019	39.0
13		55.9294838	Entry 4639	0.0	Entry 2437	0.0169492	Entry 6523	0.0508475	Entry 7093	0.1123964	Entry 6721	0.1355932	58.0
14		56.9453151	Entry 4744	0.0338983	Entry 4687	0.0508475	Entry 1992	0.1016949	Entry 5391	0.1111111	Entry 4256	0.1161670	54.0
15		48.3325532	Entry 1210	0.0677966	Entry 4967	0.0847458	Entry 4220	0.1161670	Entry 3755	0.1295148	Entry 2769	0.1295148	52.0
16		77.0682915	Entry 1229	0.0508475	Entry 520	0.1123964	Entry 1188	0.1221931	Entry 3132	0.1261439	Entry 2410	0.1301617	78.0
17		43.0588967	Entry 1348	0.0338983	Entry 4474	0.1186441	Entry 2269	0.1221931	Entry 3826	0.1295148	Entry 728	0.1295148	47.0
18		83.8857179	Entry 3006	0.0847458	Entry 5383	0.1123964	Entry 3640	0.1186441	Entry 6300	0.1221931	Entry 6813	0.1349462	87.0
19		70.3553138	Entry 3953	0.1016949	Entry 155	0.1123964	Entry 6625	0.1161670	Entry 6673	0.1261439	Entry 1165	0.1301617	71.0
20		50.2402206	Entry 636	0.0	Entry 1808	0.1016949	Entry 5889	0.1111111	Entry 4499	0.1349462	Entry 5846	0.1397409	54.0

Step 9: Statistics calculation

Create a new tab by pressing the "+" button on the bottom of the page with the name "STATISTICS_ACCURACIES".

Import data into the input spreadsheet of the "STATISTICS_ACCURACIES" tab from the output of the "VALIDATE_MODEL(.predict)" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Calculate the statistical metrics for the regression: *Statistics → Model Metrics → Regression Metrics*



The results will appear on the output spreadsheet.

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)
User Header	User Row ID	Mean Squared Error	Root Mean Squared Error	Mean Absolute Error	R Squared
1		8.7171405	2.9524804	2.3482714	0.9766679

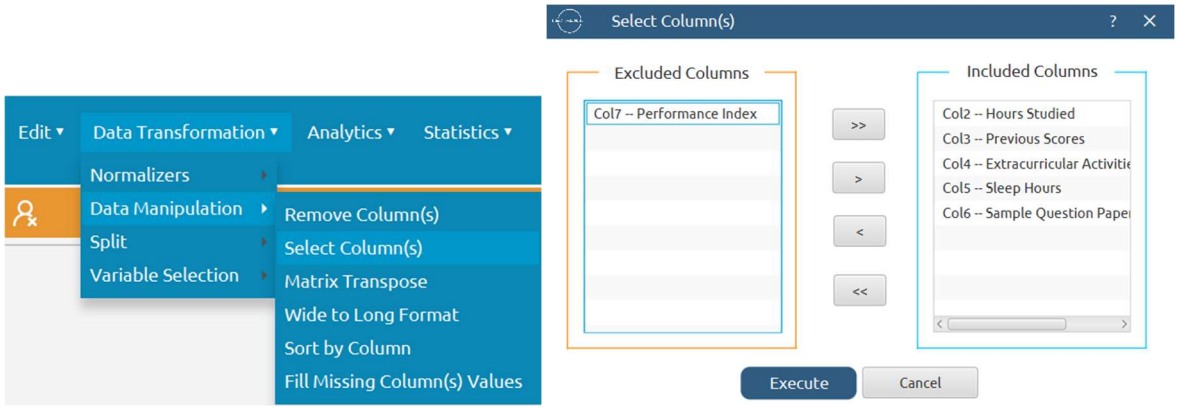
Step 10: Reliability check for each record of the test set

Step 10.a: Create the domain

Create a new tab by pressing the "+" button on the bottom of the page with the name "EXCLUDE_SCORES".

Import data into the input spreadsheet of the "EXCLUDE_SCORES" tab from the output of the "NORMALIZE_TRAIN_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Manipulate the data to exclude the target column "Performance Index": *Data Transformation → Data Manipulation → Select Column(s)*

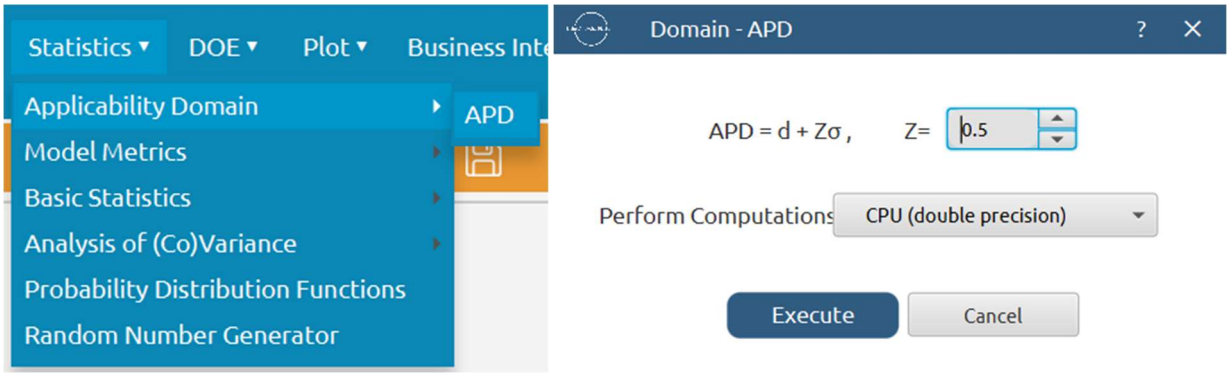


The results will appear on the output spreadsheet.

Create a new tab by pressing the "+" button on the bottom of the page with the name "DOMAIN".

Import data into the input spreadsheet of the "DOMAIN" tab from the output of the "EXCLUDE_SCORES" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Create the domain: *Statistics → Applicability Domain → APD*



The results will appear on the output spreadsheet.

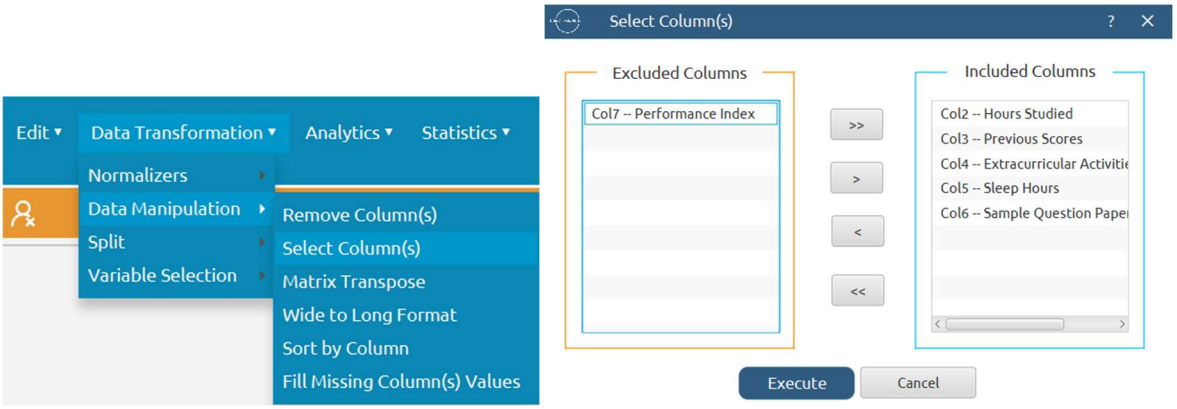
	Col1	Col2 (D)	Col3 (D)	Col4 (S)
User Header	User Row ID	Domain	APD	Prediction
1		0.0	2.2397111	reliable
2		0.0	2.2397111	reliable
3		0.0	2.2397111	reliable
4		0.0	2.2397111	reliable
5		0.0	2.2397111	reliable
6		0.0	2.2397111	reliable
7		0.0	2.2397111	reliable
8		0.0	2.2397111	reliable
9		0.0	2.2397111	reliable
10		0.0	2.2397111	reliable
11		0.0	2.2397111	reliable
12		0.0	2.2397111	reliable
13		0.0	2.2397111	reliable
14		0.0	2.2397111	reliable
15		0.0	2.2397111	reliable
16		0.0	2.2397111	reliable
17		0.0	2.2397111	reliable
18		0.0	2.2397111	reliable
19		0.0	2.2397111	reliable
20		0.0	2.2397111	reliable

Step 10.b: Check the test set reliability

Create a new tab by pressing the "+" button on the bottom of the page with the name "EXCLUDE_SCORES_TEST_SET".

Import data into the input spreadsheet of the "EXCLUDE_SCORES_TEST_SET" tab from the output of the "NORMALIZE_TEST_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Manipulate the data to exclude the target column "Performance Index": *Data Transformation → Data Manipulation → Select Column(s)*

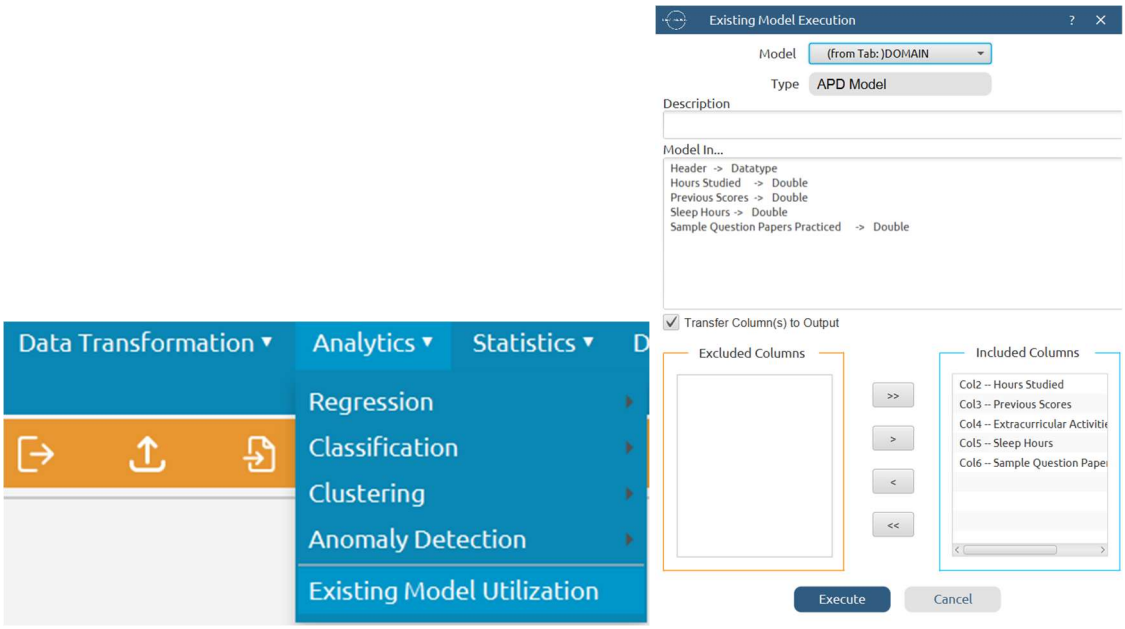


The results will appear on the output spreadsheet.

Create a new tab by pressing the "+" button on the bottom of the page with the name "RELIABILITY".

Import data into the input spreadsheet of the "RELIABILITY" tab from the output of the "EXCLUDE_SCORES_TEST_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from Spreadsheet".

Check the Reliability: *Analytics → Existing Model Utilization → Model (from Tab:) DOMAIN*



The results will appear on the output spreadsheet. There are no unreliable samples in the test set.

	Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (D)	Col7 (S)	Col8 (D)	Col9 (D)
User Header	User Row ID	Domain	APD	Prediction	Hours Studied	Previous Scores	Extracurricular Activities	Sleep Hours	Sample Question ...
1		0.1725244	2.2397111	reliable	0.7782927	1.6996717	Yes	1.4515452	-1.2504003
2		0.1150162	2.2397111	reliable	-0.3842636	0.7220336	No	-1.4958744	-0.9026160
3		0.0	2.2397111	reliable	1.1658115	-1.4057668	Yes	-1.4958744	0.4885209
4		0.1150162	2.2397111	reliable	1.1658115	0.5495093	No	-0.3169066	-0.9026160
5		0.3450487	2.2397111	reliable	-1.1593012	0.1469524	No	-1.4958744	-0.5548318
6		0.0	2.2397111	reliable	-1.5468200	1.6996717	Yes	-1.4958744	-0.5548318
7		0.0575081	2.2397111	reliable	-1.5468200	0.8945580	No	-0.9063905	0.4885209
8		0.0	2.2397111	reliable	0.7782927	-0.4281288	Yes	0.2725774	-0.2070476
9		0.0	2.2397111	reliable	1.5533303	0.8370499	Yes	-0.3169066	0.4885209
10		0.0575081	2.2397111	reliable	0.0032552	0.0319362	Yes	-0.3169066	1.5318737
11		0.1150162	2.2397111	reliable	1.5533303	0.1469524	No	0.8620613	-0.9026160
12		0.0	2.2397111	reliable	-1.1593012	-0.3706207	Yes	-0.3169066	-1.5981845
13		0.0	2.2397111	reliable	-0.3842636	0.2044605	Yes	0.2725774	-1.5981845
14		0.1150162	2.2397111	reliable	-0.7717824	0.3769849	Yes	-1.4958744	-0.5548318
15		0.1725244	2.2397111	reliable	0.3907739	-0.4281288	Yes	1.4515452	-1.5981845
16		0.0575081	2.2397111	reliable	-0.3842636	1.3546229	No	0.8620613	-0.5548318
17		0.1150162	2.2397111	reliable	-1.1593012	0.0319362	Yes	-0.9063905	1.1840894
18		0.2300325	2.2397111	reliable	1.1658115	1.2971148	Yes	-1.4958744	0.8363052
19		0.0575081	2.2397111	reliable	0.0032552	1.0670823	No	-0.9063905	-0.9026160
20		0.0	2.2397111	reliable	1.1658115	-0.5431450	No	-1.4958744	0.8363052

Final Isalos Workflow

Following the above-described steps, the final workflow on Isalos will look like this: