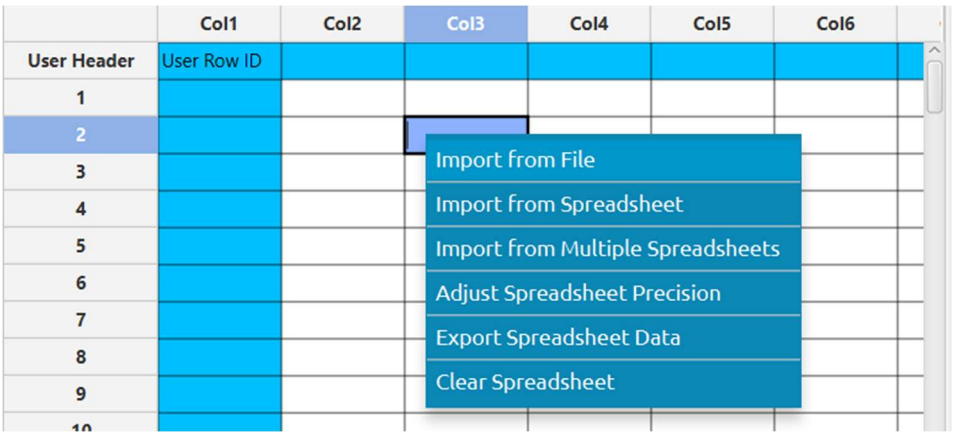
Prediction of Insurance Charges

The goal of this study is to train a model in order to predict insurance charges. The dataset used in this case study, which can be found in https://www.kaggle.com/datasets/thedevastator/prediction-of-insurance-charges-using-age-gender/data, has 8 features and 1338 samples. It contains detailed information about insurance customers including their age, sex, body mass index (BMI), number of children, smoking status and region. Having access to such valuable information allows analysts to get a better view into customer behavior and the factors that contribute to their insurance charges.

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 insurance charges data.

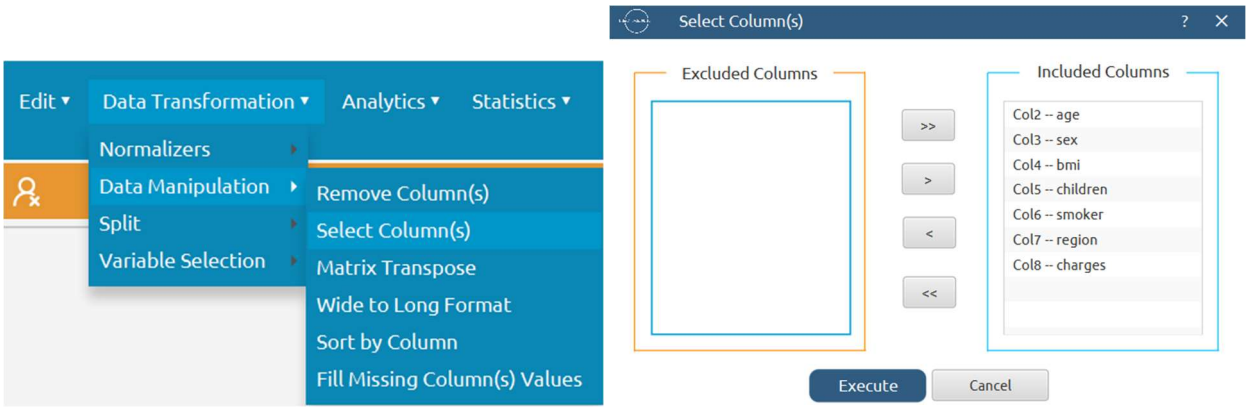


The data will appear on the left spreadsheet.

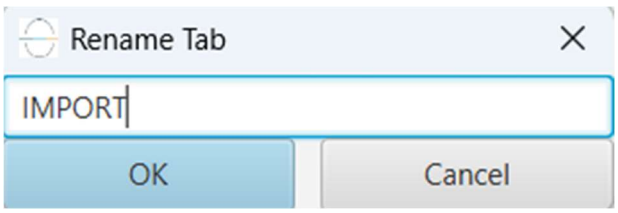
	Col1	Col2 (I)	Col3 (S)	Col4 (D)	Col5 (I)	Col6 (S)	Col7 (S)	Col8 (D)
User Header	User Row ID	age	sex	bmi	children	smoker	region	charges
1	0	19	female	27.9	0	yes	southwest	16884.924
2	1	18	male	33.77	1	no	southeast	1725.5523
3	2	28	male	33.0	3	no	southeast	4449.462
4	3	33	male	22.705	0	no	northwest	21984.47061
5	4	32	male	28.88	0	no	northwest	3866.8552
6	5	31	female	25.74	0	no	southeast	3756.6216
7	6	46	female	33.44	1	no	southeast	8240.5896
8	7	37	female	27.74	3	no	northwest	7281.5056
9	8	37	male	29.83	2	no	northeast	6406.4107
10	9	60	female	25.84	0	no	northwest	28923.13692
11	10	25	male	26.22	0	no	northeast	2721.3208
12	11	62	female	26.29	0	yes	southeast	27808.7251
13	12	23	male	34.4	0	no	southwest	1826.843
14	13	56	female	39.82	0	no	southeast	11090.7178
15	14	27	male	42.13	0	yes	southeast	39611.7577

Step 2: Manipulate data

In this dataset there are not any empty 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.



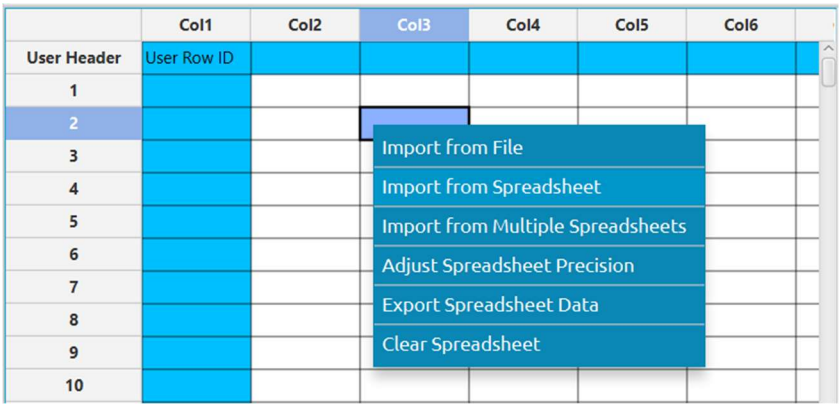
All of the data will appear in the output (right) spreadsheet. This tab can be renamed "IMPORT" by right-clicking on it and choosing the "Rename" option.



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.

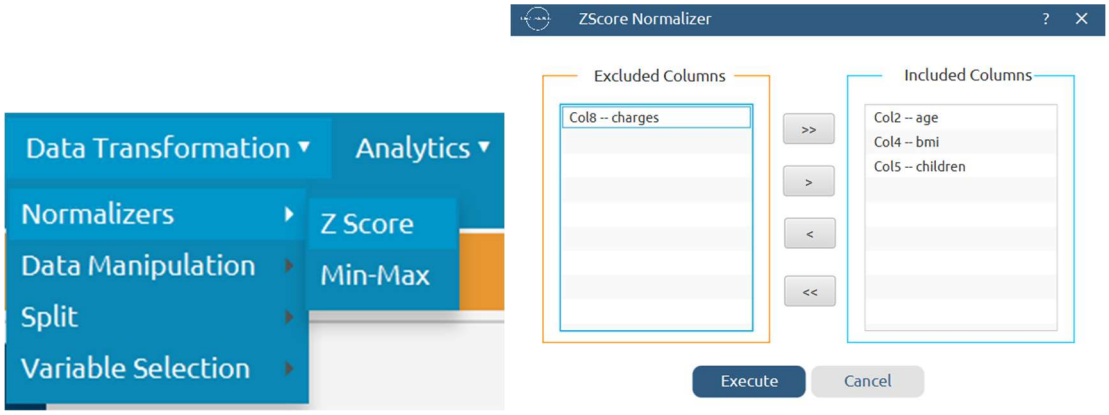
Step 4: Normalize the training set

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 (S)	Col4 (D)	Col5 (I)	Col6 (S)	Col7 (S)	Col8 (D)
User Header	User Row ID	age	sex	bmi	children	smoker	region	charges
1	0	19	female	27.9	0	yes	southwest	16884.924
2	1	18	male	33.77	1	no	southeast	1725.5523
3	2	28	male	33.0	3	no	southeast	4449.462
4	3	33	male	22.705	0	no	northwest	21984.47061
5	4	32	male	28.88	0	no	northwest	3866.8552
6	5	31	female	25.74	0	no	southeast	3756.6216
7	6	46	female	33.44	1	no	southeast	8240.5896
8	8	37	male	29.83	2	no	northeast	6406.4107
9	9	60	female	25.84	0	no	northwest	28923.13692
10	10	25	male	26.22	0	no	northeast	2721.3208
11	12	23	male	34.4	0	no	southwest	1826.843
12	13	56	female	39.82	0	no	southeast	11090.7178
13	14	27	male	42.13	0	yes	southeast	39611.7577
14	15	19	male	24.6	1	no	southwest	1837.237
15	16	52	female	30.78	1	no	northeast	10797.3362

Normalize the data using Z-score: *Data Transformation → Normalizers → Z Score* and select all columns except the "charges" target column.



The results will appear on the output spreadsheet.

	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (D)	Col6 (S)	Col7 (S)	Col8 (D)
User Header	User Row ID	age	sex	bmi	children	smoker	region	charges
1	0	-1.4420522	female	-0.4483371	-0.8940884	yes	southwest	16884.924
2	1	-1.5137975	male	0.5084180	-0.0720511	no	southeast	1725.5523
3	2	-0.7963444	male	0.3829152	1.5720236	no	southeast	4449.462
4	3	-0.4376178	male	-1.2950735	-0.8940884	no	northwest	21984.47061
5	4	-0.5093631	male	-0.2886063	-0.8940884	no	northwest	3866.8552
6	5	-0.5811084	female	-0.8003969	-0.8940884	no	southeast	3756.6216
7	6	0.4950712	female	0.4546311	-0.0720511	no	southeast	8240.5896
8	8	-0.1506366	male	-0.1337652	0.7499863	no	northeast	6406.4107
9	9	1.4995056	female	-0.7840978	-0.8940884	no	northwest	28923.13692
10	10	-1.0115803	male	-0.7221614	-0.8940884	no	northeast	2721.3208
11	12	-1.1550709	male	0.6111021	-0.8940884	no	southwest	1826.843
12	13	1.2125244	female	1.4945114	-0.8940884	no	southeast	11090.7178
13	14	-0.8680897	male	1.8710198	-0.8940884	yes	southeast	39611.7577
14	15	-1.4420522	male	-0.9862063	-0.0720511	no	southwest	1837.237
15	16	0.9255431	female	0.0210760	-0.0720511	no	northeast	10797.3362

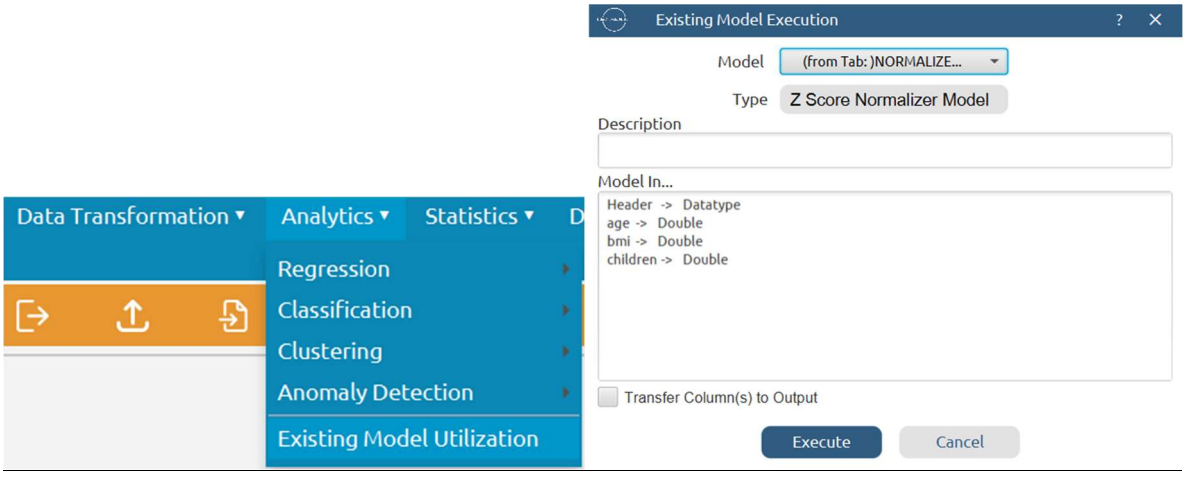
Step 5: Normalize the test set

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 (S)	Col4 (D)	Col5 (I)	Col6 (S)	Col7 (S)	Col8 (D)
User Header	User Row ID	age	sex	bmi	children	smoker	region	charges
1	7	37	female	27.74	3	no	northwest	7281.5056
2	11	62	female	26.29	0	yes	southeast	27808.7251
3	17	23	male	23.845	0	no	northeast	2395.17155
4	24	37	male	28.025	2	no	northwest	6203.90175
5	25	59	female	27.72	3	no	southeast	14001.1338
6	29	31	male	36.3	2	yes	southwest	38711.0
7	30	22	male	35.6	0	yes	southwest	35585.576
8	33	63	male	28.31	0	no	northwest	13770.0979
9	37	26	male	20.8	0	no	southwest	2302.3
10	38	35	male	36.67	1	yes	northeast	39774.2763
11	41	31	female	36.63	2	no	southeast	4949.7587
12	42	41	male	21.78	1	no	southeast	6272.4772
13	44	38	male	37.05	1	no	northeast	6079.6715
14	47	28	female	34.77	0	no	northwest	3556.9223
15	49	36	male	35.2	1	yes	southeast	38709.176

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.

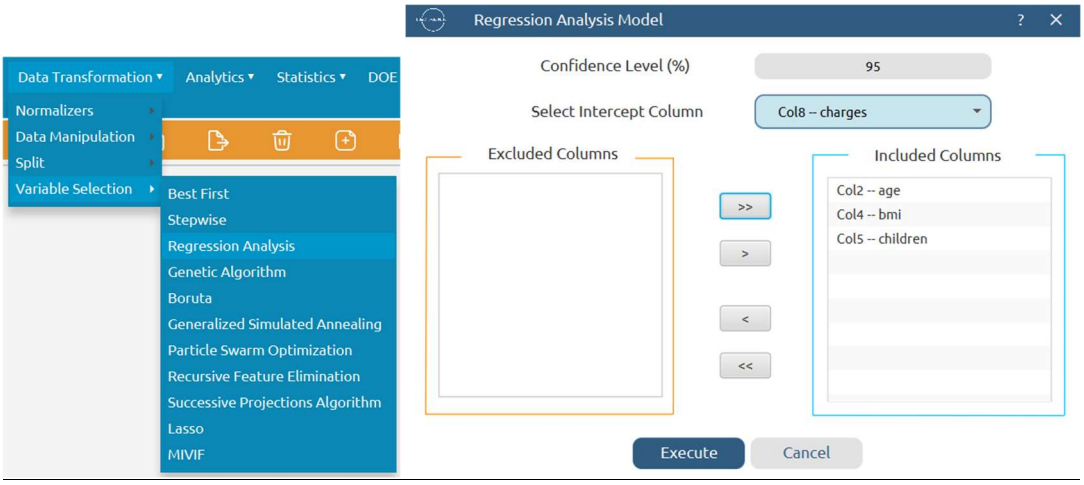
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (D)	Col6 (S)	Col7 (S)	Col8 (D)
User Header	User Row ID	age	sex	bmi	children	smoker	region	charges
1	7	-0.1506366	female	-0.4744156	1.5720236	no	northwest	7281.5056
2	11	1.6429962	female	-0.7107521	-0.8940884	yes	southeast	27808.7251
3	17	-1.1550709	male	-1.1092642	-0.8940884	no	northeast	2395.17155
4	24	-0.1506366	male	-0.4279633	0.7499863	no	northwest	6203.90175
5	25	1.4277603	female	-0.4776754	1.5720236	no	southeast	14001.1338
6	29	-0.5811084	male	0.9207843	0.7499863	yes	southwest	38711.0
7	30	-1.2268163	male	0.8066909	-0.8940884	yes	southwest	35585.576
8	33	1.7147415	male	-0.3815109	-0.8940884	no	northwest	13770.0979
9	37	-0.9398350	male	-1.6055707	-0.8940884	no	southwest	2302.3
10	38	-0.2941272	male	0.9810909	-0.0720511	yes	northeast	39774.2763
11	41	-0.5811084	female	0.9745713	0.7499863	no	southeast	4949.7587
12	42	0.1363447	male	-1.4458399	-0.0720511	no	southeast	6272.4772
13	44	-0.0788913	male	1.0430273	-0.0720511	no	northeast	6079.6715
14	47	-0.7963444	female	0.6714086	-0.8940884	no	northwest	3556.9223
15	49	-0.2223819	male	0.7414946	-0.0720511	yes	southeast	38709.176

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 "charges" 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	0.3488613				
3		R Square	0.1217042				
4		Adjusted R Square	0.1190693				
5		Standard Error	11066.9970067				
6		Observations	1004				
7							
8			Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	Significance F
9		Regression	3	16971659382.634127	5657219794.211375	46.1895219	0E-7
10		Residual	1000	122478422747.18384	122478422.7471838		
11		Total	1003	139450082129.81796			

	Degrees of Freedom	Sum of Squares	Mean Square	F-statistic	Significance F	
Regression	3	16971659382.634127	5657219794.211375	46.1895219	0E-7	
Residual	1000	122478422747.18384	122478422.7471838			
Total	1003	139450082129.81796				
	Coefficients	Standard Error	t-statistic	P-value	Lower 95.0%	Upper 95.0%
charges	12823.9241044	349.2713285	36.7162233	0E-7	12138.5353269	13509.3128820
age	3426.9942947	352.6739791	9.7171736	0E-7	2734.9283628	4119.0602266
bmi	1840.6905865	351.9047307	5.2306503	2E-7	1150.1341809	2531.2469921
children	398.9991025	350.3202892	1.1389552	0.2549946	-288.4480916	1086.4462966

The significant numerical features according to the p-value are the following:

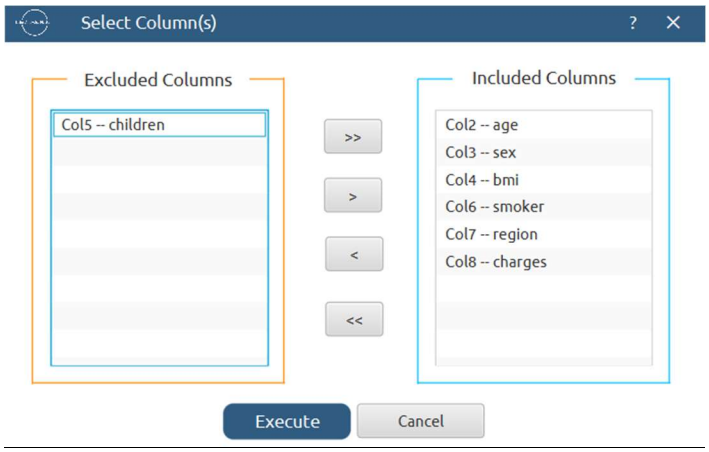
- age (p-value = 0E-7)
- bmi (p-value = 2E-7)

Step 7: Feature selection: Train set

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

Import data into the input spreadsheet of the "FEATURE_SELECTION_TRAIN_SET" 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 by choosing the columns that correspond to the significant features (as indicated by the previous step) by browsing: *Data Transformation → Data Manipulation → Select Column(s)*



The results will appear on the output spreadsheet.

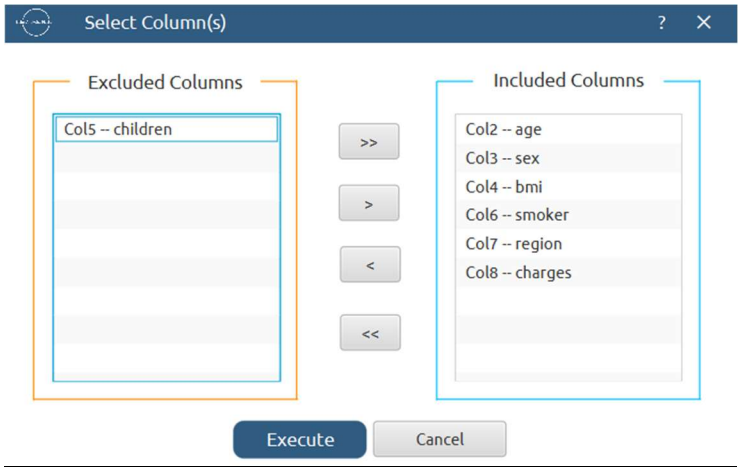
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (S)	Col7 (D)
User Header	User Row ID	age	sex	bmi	smoker	region	charges
1	0	-1.4420522	female	-0.4483371	yes	southwest	16884.924
2	1	-1.5137975	male	0.5084180	no	southeast	1725.5523
3	2	-0.7963444	male	0.3829152	no	southeast	4449.462
4	3	-0.4376178	male	-1.2950735	no	northwest	21984.47061
5	4	-0.5093631	male	-0.2886063	no	northwest	3866.8552
6	5	-0.5811084	female	-0.8003969	no	southeast	3756.6216
7	6	0.4950712	female	0.4546311	no	southeast	8240.5896
8	8	-0.1506366	male	-0.1337652	no	northeast	6406.4107
9	9	1.4995056	female	-0.7840978	no	northwest	28923.13692
10	10	-1.0115803	male	-0.7221614	no	northeast	2721.3208
11	12	-1.1550709	male	0.6111021	no	southwest	1826.843
12	13	1.2125244	female	1.4945114	no	southeast	11090.7178
13	14	-0.8680897	male	1.8710198	yes	southeast	39611.7577
14	15	-1.4420522	male	-0.9862063	no	southwest	1837.237
15	16	0.9255431	female	0.0210760	no	northeast	10797.3362

Step 8: Feature selection: Test set

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

Import data into the input spreadsheet of the "FEATURE_SELECTION_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 by choosing the columns that correspond to the significant features by browsing: *Data Transformation → Data Manipulation → Select Column(s)*



The results will appear on the output spreadsheet.

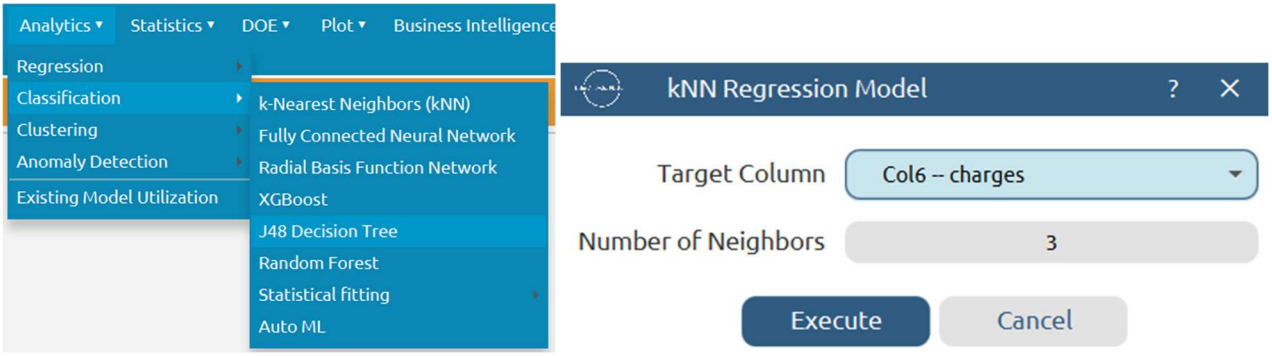
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (S)	Col7 (D)
User Header	User Row ID	age	sex	bmi	smoker	region	charges
1	7	-0.1506366	female	-0.4744156	no	northwest	7281.5056
2	11	1.6429962	female	-0.7107521	yes	southeast	27808.7251
3	17	-1.1550709	male	-1.1092642	no	northeast	2395.17155
4	24	-0.1506366	male	-0.4279633	no	northwest	6203.90175
5	25	1.4277603	female	-0.4776754	no	southeast	14001.1338
6	29	-0.5811084	male	0.9207843	yes	southwest	38711.0
7	30	-1.2268163	male	0.8066909	yes	southwest	35585.576
8	33	1.7147415	male	-0.3815109	no	northwest	13770.0979
9	37	-0.9398350	male	-1.6055707	no	southwest	2302.3
10	38	-0.2941272	male	0.9810909	yes	northeast	39774.2763
11	41	-0.5811084	female	0.9745713	no	southeast	4949.7587
12	42	0.1363447	male	-1.4458399	no	southeast	6272.4772
13	44	-0.0788913	male	1.0430273	no	northeast	6079.6715
14	47	-0.7963444	female	0.6714086	no	northwest	3556.9223
15	49	-0.2223819	male	0.7414946	yes	southeast	38709.176

Step 9: 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 "FEATURE_SELECTION_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.

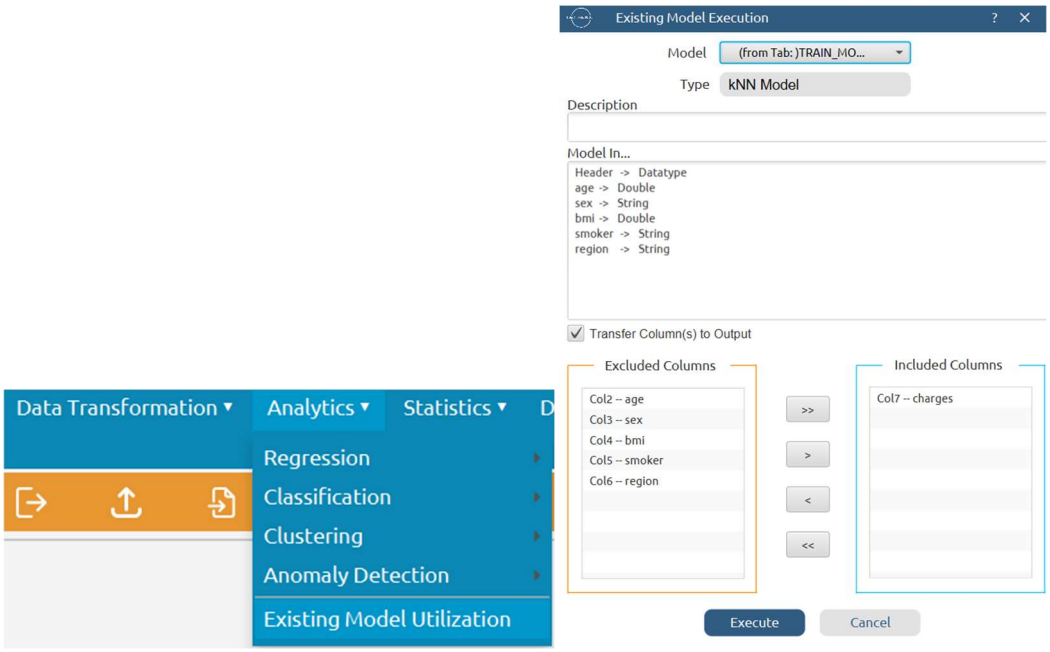
	Col1	Col2 (D)	Col3 (D)	Col4 (I)	Col5 (D)	Col6 (I)	Col7 (D)	Col8 (I)	Col9 (D)
User Header	User Row ID	charges	kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Distance from NN3
1	0	16884.924	17022.6151402	0	0.0	126	0.0107614	1120	0.1281714
2	1	1725.5523	1571.6301042	1	0.0	663	0.0029594	22	0.0088781
3	2	4449.462	4312.8733941	2	0.0	470	0.0234821	511	0.0280691
4	3	21984.47061	21578.9276968	3	0.0	451	0.0756509	1324	0.0971680
5	4	3866.8552	3900.1701985	4	0.0	434	0.0230517	1081	0.0281141
6	5	3756.6216	3982.6767119	5	0.0	406	0.0580555	583	0.0602843
7	6	8240.5896	8425.8563796	6	0.0	1280	0.0435789	1057	0.0494279
8	8	6406.4107	6442.5648388	8	0.0	300	0.0650781	1116	0.0871066
9	9	28923.13692	28287.8625453	9	0.0	912	0.0316497	287	0.0660138
10	10	2721.3208	2815.8390552	10	0.0	809	0.0102233	902	0.0355386
11	12	1826.843	1910.0465221	12	0.0	761	0.0215227	880	0.0242569
12	13	11090.7178	11143.1747372	13	0.0	676	0.0343799	386	0.0481612
13	14	39611.7577	39579.9634870	14	0.0	377	0.0842073	82	0.1629011
14	15	1837.237	1830.1849295	15	0.0	791	0.0807103	248	0.0995426
15	16	10797.3362	13740.6195180	16	0.0	1098	0.0025558	696	0.0463125

Step 10: 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 "FEATURE_SELECTION_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 "charges" to be transferred to the output spreadsheet.



The predictions will appear on the output spreadsheet.

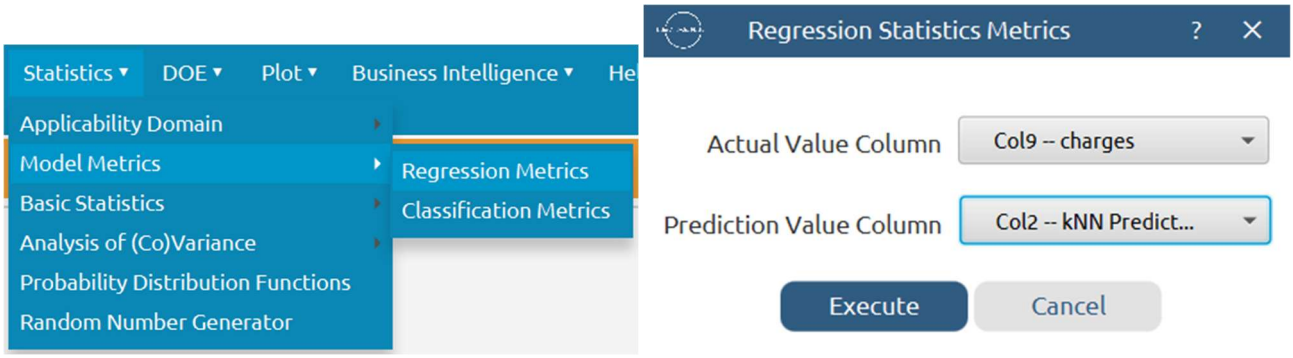
	Col1	Col2 (D)	Col3 (I)	Col4 (D)	Col5 (I)	Col6 (D)	Col7 (I)	Col8 (D)	Col9 (D)
User Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Distance from NN3	charges
1	7	6596.6066173	831	0.0555471	1319	0.0579665	1219	0.0699170	7281.5056
2	11	26694.8903115	664	0.0988558	843	0.1441625	655	0.2190168	27808.7251
3	17	2508.5472209	1114	0.0178908	809	0.0690730	1077	0.0731158	2395.17155
4	24	6488.6025722	176	0.0223320	1063	0.0266036	1173	0.0397056	6203.90175
5	25	12882.7131338	1330	0.0687599	789	0.0880708	48	0.0885324	14001.1338
6	29	40989.4574382	19	0.0345888	609	0.0458380	34	0.0652729	38711.0
7	30	36144.6677790	1291	0.0678820	1288	0.1110943	373	0.1133044	35585.576
8	33	13435.2223111	337	0.0298438	341	0.0508826	624	0.0878905	13770.0979
9	37	3817.3889501	737	0.0780199	1245	0.1037152	1002	0.1413690	2302.3
10	38	39712.6704161	947	0.0794113	1249	0.1046056	185	0.1422414	39774.2763
11	41	4249.9764768	632	0.0525942	1127	0.0893900	307	0.0914041	4949.7587
12	42	6422.6759918	408	0.0675914	872	0.0856667	1221	0.0885324	6272.4772
13	44	10314.5469700	325	0.0903761	1318	0.0968224	988	0.1079078	6079.6715
14	47	4244.4713748	1271	0.0656666	1190	0.0844632	440	0.0861099	3556.9223
15	49	37771.0069870	53	0.0207156	1118	0.0668749	158	0.1307366	38709.176

Step 11: 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		34589542.578254	5881.2874932	3328.8180017	0.7967932

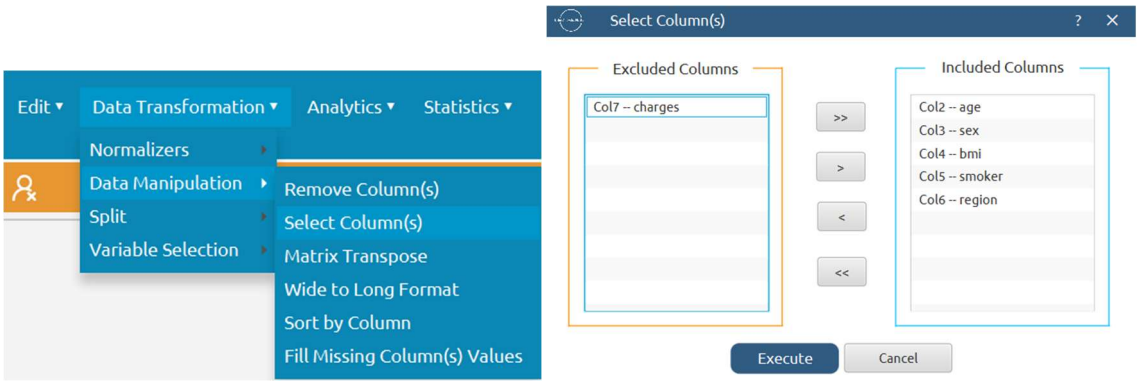
Step 12: Reliability check for each record of the test set

Step 12.a: Create the domain

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

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

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

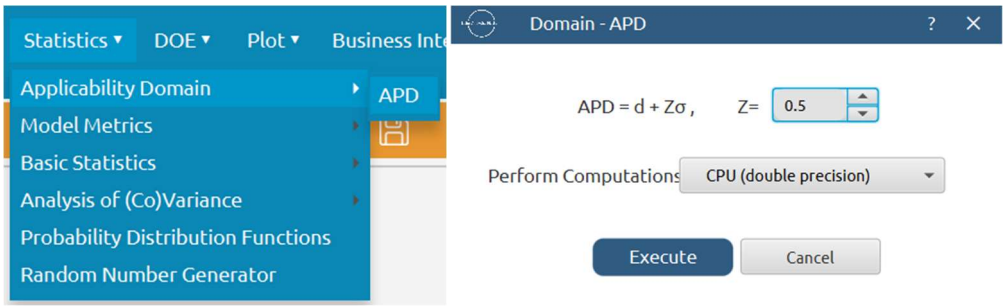


The results will appear on the output spreadsheet.

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_CHARGES" 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.

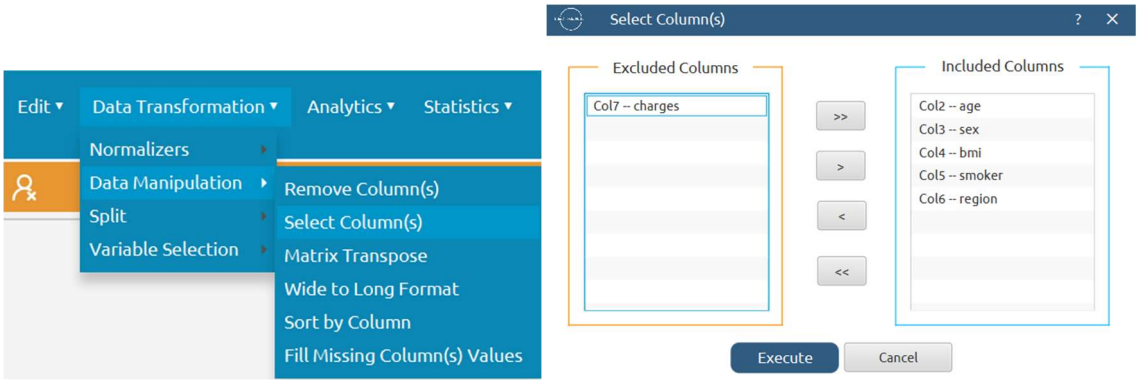
	Col1	Col2 (D)	Col3 (D)	Col4 (S)
User Header	User Row ID	Domain	APD	Prediction
1	0	0.0	1.3124962	reliable
2	1	0.0	1.3124962	reliable
3	2	0.0	1.3124962	reliable
4	3	0.0	1.3124962	reliable
5	4	0.0	1.3124962	reliable
6	5	0.0	1.3124962	reliable
7	6	0.0	1.3124962	reliable
8	8	0.0	1.3124962	reliable
9	9	0.0	1.3124962	reliable
10	10	0.0	1.3124962	reliable
11	12	0.0	1.3124962	reliable
12	13	0.0	1.3124962	reliable
13	14	0.0	1.3124962	reliable
14	15	0.0	1.3124962	reliable
15	16	0.0	1.3124962	reliable

Step 12.b: Check the test set reliability

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

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

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

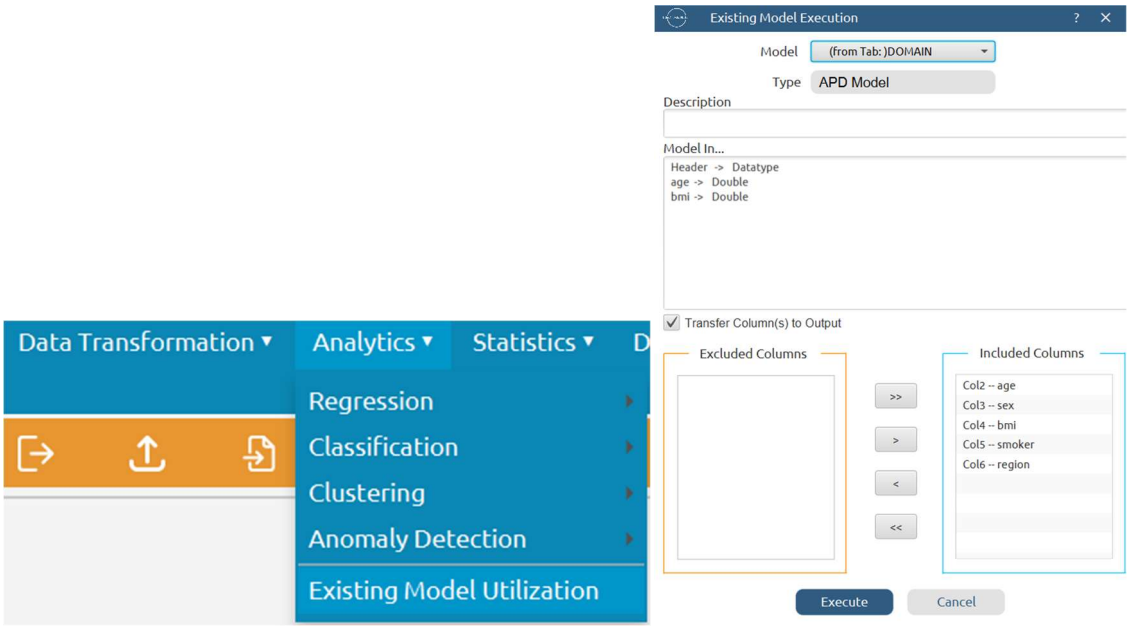


The results will appear on the output spreadsheet.

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_CHARGES_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 (S)	Col7 (D)	Col8 (S)	Col9 (S)
User Header	User Row ID	Domain	APD	Prediction	age	sex	bmi	smoker	region
1	7	0.0733972	1.3124962	reliable	-0.1506366	female	-0.4744156	no	northwest
2	11	0.0660112	1.3124962	reliable	1.6429962	female	-0.7107521	yes	southeast
3	17	0.0619364	1.3124962	reliable	-1.1550709	male	-1.1092642	no	northeast
4	24	0.0717453	1.3124962	reliable	-0.1506366	male	-0.4279633	no	northwest
5	25	0.0358579	1.3124962	reliable	1.4277603	female	-0.4776754	no	southeast
6	29	0.1434999	1.3124962	reliable	-0.5811084	male	0.9207843	yes	southwest
7	30	0.0726468	1.3124962	reliable	-1.2268163	male	0.8066909	yes	southwest
8	33	0.0929047	1.3124962	reliable	1.7147415	male	-0.3815109	no	northwest
9	37	0.0717453	1.3124962	reliable	-0.9398350	male	-1.6055707	no	southwest
10	38	0.1299828	1.3124962	reliable	-0.2941272	male	0.9810909	yes	northeast
11	41	0.1104235	1.3124962	reliable	-0.5811084	female	0.9745713	no	southeast
12	42	0.0040748	1.3124962	reliable	0.1363447	male	-1.4458399	no	southeast
13	44	0.1108336	1.3124962	reliable	-0.0788913	male	1.0430273	no	northeast
14	47	0.0937244	1.3124962	reliable	-0.7963444	female	0.6714086	no	northwest
15	49	0.0969430	1.3124962	reliable	-0.2223819	male	0.7414946	yes	southeast

Final Isalos Workflow

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