

The goal of this study is to train a model in order to predict housing prices. The dataset used in this case study is found in <u>https://www.kaggle.com/datasets/yasserh/housing-prices-dataset/data</u> and has 13 features and 545 samples. This dataset contains information on certain factors like house area, bedrooms, furnished, nearness to main road, etc, aiming to predict housing prices in the Northeast states of USA.

The dataset contains no missing values and includes several categorical features. Categorical features contain multiple levels, and the data was transformed to corresponding numeric codes, as detailed below:

mainroad:

- No (0)
- Yes (1)

guestroom:

- No (0)
- Yes (1)

basement:

- No (0)
- Yes (1)

hotwaterheating:

- No (0)
- Yes (1)

airconditioning:

- No (0)
- Yes (1)

prefarea:

- No (0)
- Yes (1)

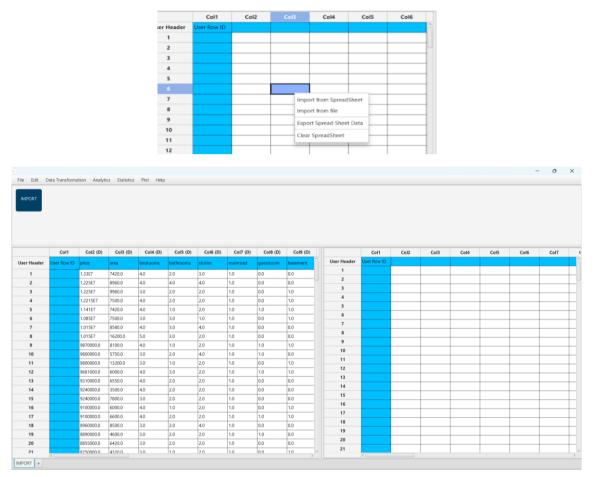
furnishingstatus:

• unfurnished (0)

- semi furnished (1)
- furnished (2)
- Southwest (3)

# Step 1: Import data from file

Right click on the input spreadsheet and choose the option "Import from file". Then navigate through your files to load the one with the housing price data.



# Step 2: Manipulate data

In order to use the data for training we have to exclude any columns that do not contain features. In our dataset there are no such columns. Therefore, we will include all columns in the training. We follow these steps to execute this:

- On the menu click on "Data Transformation"  $\rightarrow$  "Data Manipulation"  $\rightarrow$  "Select Column(s)"
- Select all columns.

IMPORT	Normalizers	•				
IMPORT	Data Manipulatio	Remove	Column(s)	_		
_	Split		olumn(s)			
	Variable Selectio	n 🕨 Matrix 1	ranspose			
		Sort by	Column			
		Fill Miss	ing Column(s)	Values		
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)
User Header	User Row ID	price	area	bedrooms	bathrooms	stories
1		1.33E7	7420.0	4.0	2.0	3.0
2		1.225E7	8960.0	4.0	4.0	4.0
3		1.225E7	9960.0	3.0	2.0	2.0
4		1.2215E7	7500.0	4.0	2.0	2.0
5		1.141E7	7420.0	4.0	1.0	2.0
6		1.085E7	7500.0	3.0	3.0	1.0
7		1.015E7	8580.0	4.0	3.0	4.0
8		1.015E7	16200.0	5.0	3.0	2.0
9		9870000.0	8100.0	4.0 3.0	1.0	2.0
10		9800000.0 9800000.0	5750.0 13200.0	3.0	2.0	4.0 2.0
12		9681000.0	6000.0	4.0	3.0	2.0
12		9310000.0	6550.0	4.0	2.0	2.0
14		9240000.0	3500.0	4.0	2.0	2.0
15		9240000.0	7800.0	3.0	2.0	2.0
16		9100000.0	6000.0	4.0	1.0	2.0
17		9100000.0	6600.0	4.0	2.0	2.0
18		8960000.0	8500.0	3.0	2.0	4.0
19		8890000.0	4600.0	3.0	2.0	2.0
20		8855000.0	6420.0	3.0	2.0	2.0
21		8750000.0	4320.0	3.0	1.0	2.0

The data will appear in the output spreadsheet.

## 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 to create 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".

and the second second		1																
IMPORT	TRAIN_TEST_SPL	T																
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	1	Col1	Col2	Col3	Col4	Col5	Col6	Col7	
Jser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basemen	User Header	User Row ID							
1		13300000.000	7420.0000000	4 0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7	1								
		0000	7420.0000000	4.000000						2								
2		0000	8960.0000000	4.0000000	4.0000000	4.0000000	1.0000000	0E-7	0E-7	3								-
3		12250000.000 0000	9960.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.00000C	4								-
4		12215000.000	7500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.000000	6		-	-					-
5		11410000.000	7420.0000000	4 0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.000000	7								
		0000	7500.0000000	4.0000000	-	-				8								
6					3.0000000	1.0000000	1.0000000	0E-7	1.000000	9								
7		10150000.000 0000	8580.0000000	4.0000000	3.0000000	4.0000000	1.0000000	0E-7	0E-7	10								
8		10150000.000 0000	16200.000000	5.0000000	3.0000000	2.0000000	1.0000000	0E-7	0E-7	11								-
9		9870000.0000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.000000	13								-
10		000 9800000.0000 000			2.0000000	4.0000000	1.0000000	1.0000000	0E-7	14								
		000	13200.000000	3.0000000		0.000				15								
11					1.0000000	2.0000000	1.0000000	0E-7	1.00000C	16			1					
12		9681000.0000 000			3.0000000	2.0000000	1.0000000	1.0000000	1.000000	17								
13		9310000.0000 000	6550.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	18		_						-
14		9240000.0000			2.0000000	2.0000000	1.0000000	0E-7	0E-7	20								-
		000		1		1		10000										

Split the dataset by choosing from the top ribbon: "Data Transformation"  $\rightarrow$  "Split"  $\rightarrow$  "Random Partitioning". Then choose the "Training set percentage" and the column for the sampling as shown below:

Ľ			-Stone						
	Variable Selectio		Partitioning						
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	
User Header 1	User Row ID	price 13300000.000 0000	area 7420.0000000	bedrooms 4.0000000	2.0000000	stories 3.0000000	mainroad 1.0000000	guestroom 0E-7	Random Partitioning
2		12250000.000	8960.0000000	4.0000000	4.0000000	4.0000000	1.0000000	0E-7	Training Set Percentage 75
3		12250000.000 0000	9960.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	
4		12215000.000 0000	7500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	Time-based RNG Seed 244416648
5		11410000.000 0000	7420.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	Stratified sampling Col3 area
6		10850000.000	7500.0000000	3.0000000	3.0000000	1.0000000	1.0000000	0E-7	
7		0000	8580.0000000 16200.000000	4.0000000	3.0000000	4.0000000	1.0000000	0E-7	Execute Cancel
8	_				3.0000000	2.0000000	1.0000000	0E-7	
9		9870000.0000 000 9800000.0000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	
10		9800000.0000	13200.000000	3.0000000	1.0000000	2.0000000	1.0000000	0E-7	
12		9681000.0000			3.0000000	2.0000000	1.0000000	1.0000000	
13		9310000.0000			2.0000000	2.0000000	1.0000000	0E-7	
14		9240000.0000	3500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	
15			7800 0000000		2 0000000	2.0000000	1.0000000	0F-7	

The results will appear on the output spreadsheet.

File Edit Dat	ta Transforma	tion Analytic	s Statistics	Plot Help														_
IMPORT	TRAIN_TEST_SPLI																	
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basem	User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestro
1		0000	7420.0000000	4.0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7	1			7420.0000000	4.0000000	2.0000000	3.0000000	1.0000000	0E-7
2		0000	8960.0000000		4.0000000	4.0000000	1.0000000	0E-7	0E-7	2		12250000.000 0000	8960.0000000		4.0000000	4.0000000	1.0000000	0E-7
3		12250000.000 0000		3.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.0000	3		12250000.000 0000	9960.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7
4		0000	7500.0000000		2.000000	2.0000000	1.0000000	0E-7	1.0000	4		0000	7420.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.00000
5		11410000.000 0000 10850000.000			1.0000000	2.000000	1.0000000	1.0000000	1.0000	5			7500.0000000		3.0000000	1.0000000	1.0000000	0E-7
6		10850000.000 0000 10150000.000			3.000000	1.0000000	1.0000000	0E-7	1.0000	6			8580.0000000		3.0000000	4.0000000	1.0000000	0E-7
7		10150000.000 0000 10150000.000	8580.0000000	4.0000000	3.000000	4.0000000	1.0000000	0E-7	0E-7	7			16200.000000 0 13200.000000		3.0000000	2.000000	1.0000000	0E-7
8					3.0000000	2.000000	1.0000000	0E-7	0E-7	8					1.0000000	2.0000000	1.0000000	0E-7
9		9870000.0000 000 9800000.0000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000	9		000	6000.0000000	4.0000000	3.0000000	2.000000	1.0000000	1.00000
10		9800000.0000 000 9800000.0000	5750.0000000	3.0000000	2.0000000	4.0000000	1.0000000	1.0000000	0E-7	10		000	6550.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7
11		0600	0		1.0000000	2.0000000	1.0000000	0E-7	1.0000	11		000 9240000.0000	3500.0000000	4.0000000	2.0000000	2.000000	1.0000000	0E-7
12		000 9310000.0000 000	6000.0000000		3.0000000	2.0000000	1.0000000	1.0000000	1.0000	12		000	7800.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7
13		000 9240000.0000			2.0000000	2.0000000	1.0000000	0E-7	0E-7	13		000 8960000.0000	6600.0000000 8500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	1.00000
14			7800.0000000		2.0000000	2.0000000	1.0000000	0E-7	0E-7	14			8500.0000000		2.0000000	4.0000000	1.0000000	0E-7

#### 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 data 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".

	-																	
IMPORT	TRAIN_TEST_SPL	T NORMALIZE_TRA	N_SET															
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	1		Col1	Col2	Col3	Col4	Col5	Col6	Col7
ser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	â	User Header	User Row ID						
1		13300000.000	7420.0000000	4.0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7		1							
-		0000	1420.0000000	4.000000						- 41	2							
2			8960.0000000		4.0000000	4.0000000	1.0000000	0E-7	0E-7	6	3							
3		12250000.000 0000	9960.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.0000000	¢	4							
4		11410000.000	7420.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	(	6							
5		10850000.000	7500.0000000	3.0000000	3.0000000	1.0000000	1.0000000	0E-7	1.0000000		7							
		0000	8580.0000000							£	8							
6	_	0000	8580.0000000	4.0000000	3.0000000	4.0000000	1.0000000	0E-7	0E-7		9							
7		10150000.000	16200.000000 0	5.0000000	3.0000000	2.0000000	1.0000000	0E-7	0E-7	6	10							_
8		9800000.0000 000	13200.000000	3.0000000	1.0000000	2.0000000	1.0000000	0E-7	1.0000000	1	11							
9		9681000.0000	6000.0000000	4.0000000	3.0000000	2.0000000	1.0000000	1.0000000	1.0000000		12							
10		000 9310000.0000 000	CEE0 0000000	4 0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7		14							
10		000	6550.0000000	4.0000000				122.22	22.2	-	15							
11		9240000.0000 000	3500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7		16							
12		9240000.0000	7800.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	(	17							
13		9100000.0000			2.0000000	2.0000000	1.0000000	1.0000000	1.0000000		18							
			8500.0000000		2.0000000	4.0000000	1.0000000	0E-7	0E-7		19 20							-
14		000				14.00000000	1.00000000	UE-1	05-7	1	20				-			

Normalize the data using Z-score by browsing: "Data Transformation"  $\rightarrow$  "Normalizers"  $\rightarrow$  "Z-Score". Then select all columns except "price" and click "Execute".

IMPORT De	ormalizers ata Manipulatic Ilit iriable Selectio	<ul> <li>Z Score</li> <li>Min-Max</li> <li>A</li> </ul>									
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Cc		Included Columns
ser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	base		Col3 area
1			7420.0000000		2.0000000	3.0000000	1.0000000	0E-7	0E-7	>>	Col4 bedrooms
2		12250000.000 0000			4.0000000	4.0000000	1.0000000	0E-7	0E-7		
3		12250000.000	9960.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.000	>	Col5 bathrooms
4		11410000.000			1.0000000	2.0000000	1.0000000	1.0000000	1.000		Col6 stories
5		10850000.000	7500.0000000	3.0000000	3.0000000	1.0000000	1.0000000	0E-7	1.000	<	Col7 mainroad
6		10150000.000			3.0000000	4.0000000	1.0000000	0E-7	0E-7		Col8 guestroom
7		10150000.000	16200.000000	5.0000000	3.0000000	2.0000000	1.0000000	0E-7	0E-7		Col9 basement
8		9800000.0000			1.0000000	2.0000000	1.0000000	0E-7	1.000	< <	
9		9681000.0000			3.0000000	2.0000000	1.0000000	1.0000000	1.000		Col10 hotwaterheating
		000 9310000.0000	0000.0000000	4.0000000			100000		0.00		Coll1 airconditioning
10					2.0000000	2.0000000	1.0000000	0E-7	0E-7		
11		9240000.0000 000	3500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	Execute	Cancel
12		9240000.0000 000			2.0000000	2.0000000	1.0000000	0E-7	0E-7		
13		9100000.0000 000			2.0000000	2.0000000	1.0000000	1.0000000	1.000		
14		8960000.0000 000	8500.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E-7	OE-7		
15		8890000.0000	4600 0000000	3.0000000	2 0000000	2 0000000	1.0000000	1 0000000	0F-7		

The results will appear on the output spreadsheet.

MPORT	TRAIN_TEST_SPLI	T NORWALDE_TRA	IN 587															
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	4		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (I
er Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	ĥ	User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad
1		13300000.000	7420.0000000	4.0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7				13300000.000 0000	1.0188414	1.4513175	1.3206150	1.3606716	0.401915
		0000	1420.0000000	4.0000000						- 1	2		12250000.000	1.7183192	1.4513175	5.1243580	2.5081245	0.401915
2		12250000.000 0000			4.0000000	4.0000000	1.0000000	0E-7	0E-7	(			0000 12250000.000 0000					
3			9960.0000000		2.0000000	2.0000000	1.0000000	0E-7	1.0000000	C	3		0000	2.1725255	0.0644276	1.3206150	0.2132186	0.401915
4		11410000.000 0000	7420.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	(	4		11410000.000 0000	1.0188414	1.4513175	-0.5812566	0.2132186	0.401915
5		10850000.000	7500.0000000		3.0000000	1.0000000	1.0000000	0E-7	1.0000000	-	5		10850000.000	1.0551779	0.0644276	3.2224865	-0.9342343	0.401915
			8580.0000000							÷	6		10150000.000	1.5457208	1.4513175	3.2224865	2.5081245	0.401915
6		0000	8580.0000000	4.0000000	3.0000000	4.0000000	1.0000000	0E-7	0E-7	-	7		0000 10150000.000 0000	E 0067720	2.8382074	3.2224865	0.2132186	0.401915
7		0000	16200.000000 0	5.0000000	3.0000000	2.0000000	1.0000000	0E-7	0E-7	(			0000	5.0067750				
8		9800000.0000 000	13200.000000	3.0000000	1.0000000	2.0000000	1.0000000	0E-7	1.0000000	C	8		9800000.0000 000	3.6441540	0.0644276	-0.5812566	0.2132186	0.401915
9		9681000.0000	6000.0000000	4.0000000	3.0000000	2.0000000	1.0000000	1.0000000	1.0000000		9		9681000.0000 000	0.3738684	1.4513175	3.2224865	0.2132186	0.401915
10		000 9310000.0000 000			2.0000000	2.0000000	1.0000000	0E-7	0E-7		10		9310000.0000 000	0.6236819	1.4513175	1.3206150	0.2132186	0.401915
										÷	11		9240000.0000	-0.7616474	1.4513175	1.3206150	0.2132186	0.401915
11		9240000.0000 000		4.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	1			000 9240000.0000 000					
12		000	7800.0000000		2.0000000	2.0000000	1.0000000	0E-7	0E-7	(	12		000	1.1914398	0.0644276	1.3206150	0.2132186	0.401915
13		9100000.0000 000	6600.0000000	4.0000000	2.0000000	2.0000000	1.0000000	1.0000000	1.0000000	c	13		9100000.0000 000	0.6463922	1.4513175	1.3206150	0.2132186	0.401915
14		8960000.0000			2.0000000	4.0000000	1.0000000	0E-7	0E-7	(	14		8960000.0000 000	1.5093843	0.0644276	1.3206150	2.5081245	0.401915
		000	4600.0000000							-	15		8890000.0000	0.000000	0.0644276	1.3206150	0.2132186	0.401915

## 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 data 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".

	ata Transforma			ны нер														_
MPORT	TRAIN_TEST_SPU1																	
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	Col1(		Col1	Col2	Col3	Col4	Col5	Col6	Co
ser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	hotw.	User Header	User Row ID						
1		12215000.000	7500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.0000000	0E-7	1							
2		9870000.0000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	0E-7	3							+
3		000 9800000.0000 000	5750.0000000	3.0000000	2.0000000	4.0000000	1.0000000	1.0000000	0E-7	0E-7	4							
4		9100000.0000	6000.0000000	4.0000000	1.0000000	2.0000000	1.0000000	0E-7	1.0000000	0E-7	5							
5		8855000.0000	6420.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	7						-	-
6		8575000.0000	8800.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	8							
7		000 8120000.0000 000	69.40.0000000	5.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	0E-7	9							
		000 8080940.0000 000	6640.0000000	5.000000		-					10							-
8		000	7000.0000000	3.0000000	2.000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	12							
9		7840000.0000 000	6360.0000000	3.000000	2.000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	13							
10		7700000.0000 000	6000.0000000	4.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	14						<u> </u>	+
11		7560000.0000 000	6000.0000000	3.0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7	0E-7	16							+
12		7525000.0000	6000.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	17							
13		7490000.0000	6600.0000000	3.0000000	1.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	18							+
14		7350000.0000	6000.0000000	3.0000000	1.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	20							-
			11440.000000		1.0000000	2 0000000	1.0000000	0F-7	1.0000000	0F-7 ~	21							-

Normalize the test set using the existing normalizer of the training set by browsing: "Analytics"  $\rightarrow$  "Existing Model Utilization"  $\rightarrow$  "Model (from Tab:) NORMALIZE\_TRAIN\_SET".

IMPORT	TRAIN_TEST_SPLI		ation	> > >				Existing Model Execution  Model (from Tab: )NORMALIZE_TR	
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Type Z Score Normalizer Model	
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	gu Description	
		12215000.000 0000	7500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	OE	
2		9870000.0000 000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.6 Model Input	
3		9800000.0000 000	5750.0000000	3.0000000	2.0000000	4.0000000	1.0000000	<sup>1.0</sup> Header -> Datatype	
4		9100000.0000 000	6000.0000000	4.0000000	1.0000000	2.0000000	1.0000000	oe area -> Double	
5		8855000.0000	6420.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E bedrooms -> Double	
6		000 8575000.0000	8800.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E stories -> Double	
7		000 8120000.0000			1.0000000	2.0000000	1.0000000	1. mainroad -> Double	
8		000 8080940.0000			2.0000000	4.0000000	1.0000000	guestroom -> Double	
-		000 7840000.0000						hetweterheating - Double	
9		000 7700000.0000	6360.0000000		2.000000	4.0000000	1.0000000	- Double	
10		000	6000.0000000	4.0000000	2.000000	4.0000000	1.0000000	OE Transfer Column(s) to Output	
11		7560000.0000 000	6000.0000000	3.0000000	2.000000	3.0000000	1.0000000	0E	
12		7525000.0000 000	6000.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E Execute Cancel	
13		7490000.0000 000	6600.0000000	3.0000000	1.0000000	4.0000000	1.0000000	0E	
14		7350000.0000 000	6000.0000000	3.0000000	1.0000000	2.0000000	1.0000000	0E	
15		7343000.0000	11440.000000	4 0000000	1 0000000	2.0000000	1.0000000	06	

IMPORT	TRAIN_TEST_SPLT	T COMMUNICATION																
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	Col1(		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	questroom	basement	hotw	User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad
1		12215000.000	7500.0000000	4.0000000	2.0000000	2.0000000	1.0000000	0E-7	1.0000000	0E-7	1			1.0331779	1.4513175	1.3206150	0.2132186	0.401915
2	-	0070000 0000	8100.0000000	4.0000000	1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	0E-7	2		9870000.0000 000	1.3277017	1.4513175	-0.5812566	0.2132186	0.401915
3		0000000 0000	5750.0000000	3.0000000	2.0000000	4.0000000	1.0000000	1.0000000	0E-7	0E-7	3		9800000.0000 000	0.2603169	0.0644276	1.3206150	2.5081245	0.40191
4		9100000.0000 000	6000.0000000	4.0000000	1.0000000	2.0000000	1.0000000	0E-7	1.0000000	0E-7	4		9100000.0000 000	0.3738684	1.4513175	-0.5812566	0.2132186	0.40191
5			6420.0000000	3.0000000	2.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	5		8855000.0000 000		0.0644276	1.3206150	0.2132186	0.40191
6		8575000.0000 000	8800.00000000		2.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	6		8575000.0000 000	1.6456462	0.0644276	1.3206150	0.2132186	0.40191
7		8120000.0000 000			1.0000000	2.0000000	1.0000000	1.0000000	1.0000000	0E-7	7		8120000.0000 000	0.7554018	2.8382074	-0.5812566	0.2132186	0.40191
8		8080940.0000 000	7000.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	8		8080940.0000 000 7840000.0000	0.8280748	0.0644276	1.3206150	2.5081245	0.40191
9		000	6360.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	9		7840000.0000 000 7700000.0000	0.5373827	0.0644276	1.3206150	2.5081245	0.40191
10		000	6000.0000000	4.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	10		7700000.0000 000 7560000.0000	0.3738684	1.4513175	1.3206150	2.5081245	0.40191
11		7560000.0000 000	6000.0000000	3.0000000	2.0000000	3.0000000	1.0000000	0E-7	0E-7	0E-7	11		7560000.0000 000 7535000.0000	0.3738684	0.0644276	1.3206150	1.3606716	0.40191
12		7525000.0000 000	6000.0000000	3.0000000	2.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	12		7525000.0000 000 7490000.0000	0.3738684	0.0644276	1.3206150	2.5081245	0.40191
13			6600.0000000		1.0000000	4.0000000	1.0000000	0E-7	0E-7	0E-7	13		7490000.0000 000 7350000.0000	0.6463922	0.0644276	-0.5812566	2.5081245	0.40191
14			6000.0000000		1.0000000	2.0000000	1.0000000	0E-7	0E-7	0E-7	14		7350000.0000 000 7343000.0000		0.0644276	-0.5812566	0.2132186	0.40191
15		7343000.0000	11440.000000	4.0000000	1.0000000	2.0000000	1.0000000	0E-7	1.0000000	0F-7 ~	15		7343000.0000	2.8447509	1.4513175	-0.5812566	0.2132186	0.40191

#### 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".

MPORT	TRAIN_TEST_SPLI	T CRIWLUE, IR		1270 <b>  19</b> 210														
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)	Col1		Col1	Col2	Col3	Col4	Col5	Col6	Co
er Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basement	hotw A	User Header	User Row ID						
1		13300000.000	1.0188414	1.4513175	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	-0.24	1							
2		12250000.000	1.7183192	1.4513175	5.1243580	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.24	2							
3		12250000.000		0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.24	4							-
4		11410000.000	1.0188414	1.4513175	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.24	5							
5		0000 10850000.000 0000	1.0100414	0.0644276	3.2224865	-0.9342343	0.4019153	-0.4771480	1.3405094	-0.24	6							
		0000	1.0551779								8							-
6		0000		1.4513175	3.2224865	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.24	9							
7		0000 9800000.0000	5.0067730	2.8382074	3.2224865	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.24	10							-
8		000	3.0441540	0.0644276	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.24	12							+
9		9681000.0000 000		1.4513175	3.2224865	0.2132186	0.4019153	2.0906616	1.3405094	4.091	13							
10		9310000.0000 000	0.0230619	1.4513175	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.24	14							
11		9240000.0000 000	-0.7616474	1.4513175	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	4.09	15							+
12		9240000.0000	1.1914398	0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.24	17							
13		9100000.0000 000	0.6463922	1.4513175	1.3206150	0.2132186	0.4019153	2.0906616	1.3405094	-0.24	18							
14		8960000.0000		0.0644276	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.24	19 20							-
		000 8890000.0000		0.0644276	1 3206150	0.2132186	0.4019153	2 0906616	-0.7441611	-0.24 ~	21							+

Choose the most important features using the Regression Analysis by browsing: "Data Transformation"  $\rightarrow$  "Variable Selection"  $\rightarrow$  "Regression Analysis". Then choose the "price" column as the intercept column, the Significance level ( $\alpha$ ) as 0.05 and include all columns.

IMPORT . D	lormalizers Pata Manipulatic plit	NORWALIZE_TR	AIN_SETEATURE_SE	ector, regressor						
V	ariable Selection	Best Firs						Regression Analysis Model		
		_	ion Analysis					Significance Level (α)	0.05	
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)		Select Intercept Column	Col8 charges	
User Header	User Row ID	price 13300000.000 0000	area 1.0188414	bedrooms 1.4513175	bathrooms 1.3206150	stories 1.3606716	0	Excluded Columns	Included Columns	
2		12250000.000 0000	1.7183192	1.4513175	5.1243580	2.5081245	0		Col2 age	
3		12250000.000 0000	2.1725255	0.0644276	1.3206150	0.2132186	0	>>	Col3 sex	
4		11410000.000 0000	1.0188414	1.4513175	-0.5812566	0.2132186	0	>	Col4 bmi	
5		10850000.000 0000	1.0551779	0.0644276	3.2224865	-0.9342343	0		Col5 children Col6 smoker	
6		10150000.000 0000	1.5457208	1.4513175	3.2224865	2.5081245	0	<		
7		10150000.000 0000	5.0067730	2.8382074	3.2224865	0.2132186	0			
8		9800000.0000 000	5.0441540	0.0644276	-0.5812566	0.2132186	0	<<		
9		9681000.0000 000		1.4513175	3.2224865	0.2132186	0			
10		9310000.0000 000	0.0250819	1.4513175	1.3206150	0.2132186	0	Execute	Cancel	
11		9240000.0000 000	-0.7616474	1.4513175	1.3206150	0.2132186	0			
12		9240000.0000 000	1.1914596	0.0644276	1.3206150	0.2132186	0			
13		9100000.0000 000	0.0403922	1.4513175	1.3206150	0.2132186	0			
14		8960000.0000 000	1.5093843	0.0644276	1.3206150	2.5081245	0			
15	<	8890000.0000	-0.2620204	0.0644276	1 3206150	0 2132186	n			

The results will appear on the output spreadsheet.

	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)		Col1	Col2 (S)	Col3 (S)	Col4 (S)	ColS (S)	Col6 (S)	Col7 (S)	Col8 (S)	Col9	Col10	Col11	Col12	Col13
88		6090000.0000	1.4512458	0.0644276	-0.5812566	1.3606716	User Header	User Row ID	a											
		000 6083000.0000					1		Regression Statistics											
89		6083000.0000 000	2.0180953	0.0644276	-0.5812566	-0.9342343	2			0.8247151										
90		6020000.0000 000	0.7826541	0.0644276	1.3206150	-0.9342343	4		A distant B	0.6801550										-
91			-0.6708061	1.4513175	-0.5812566	0.2132186	5		Standard Error	1079890.2152										
92		5950000.0000	0.5646351	0.0644276	-0.5812566	-0.9342345	6		Observations											
93		000 5950000.0000	0.0271500	0.0644276	-0.5812566	-0.934234	7			Degrees of	Sum of	Mean Square	Europtichic	Significance F						
		000 5950000.0000 000	0.0371303				9			Freedom 12.0000000	Squares 98202543121	81835452601		0E-7						
94		000	0.9329964	0.0644276	-0.5812566	0.2132186	10		Residual	296.0000000	46180049926	11661628769	70.1149700	01-1						
95		5950000.0000 000	0.9734208	1.4513175	1.3206150	0.2132186					6118.3000000 14438259304									
96		5950000.0000 000	0.6123268	0.0644276	1.3206150	2.5081245	11		Total	408.0000000	84121.500000 0									
97		5943000.0000 000	4 72 42 402	0.0644276	-0.5812566	-0.9342343	12			Coefficients			P-value	Lower 95.0%	Upper 95.0% 4899078.7708					
		000	4.7542492			_	13			4794101.4669 927				242	612 680708.16949					
98		5880000.0000 000	0.9007478	0.0644276	-0.5812566	-0.9342343	14			560208.79134 86			0E-7	49	24 199124.11361					
99		5880000.0000 000	0.6009716	0.0644276	1.3206150	1.3606716	15						0.2351862	04	51 638007.92443					
100		5873000.0000	2.8538350	0.0644276	-0.5812566	1.3606716	16			517787.31358 80 280090.26150			0E-7	57	04 515680.15809				L	
101		000 5866000.0000 000	0 1711701	0.0644276	-0.5812566	-0.9342343	17			389080.36150 93				05	81 260911.58916					-
		000	-0.1711791			_	18			147382.87864 95 150890 32531			0.0110783	6	84 266344.91936					
102		5810000.0000 000	0.2957450	1.4513175	-0.5812566	2.5081245	19			150890.32531 04 146045.07738			0.0105534	3	64 268912.95156					
103		5810000.0000 000	0.0105034	0.0644276	-0.5812566	1.3606716	20				62497.246238 0 54770.467323		0.0199461	7	90 302554.02098					-
104		5810000.0000 000	-0.1711791	0.0644276	-0.5812566	1.3606716	21		hotwaterheati ng airconditionin				0.0004189	1	34 508485.33792					-
105		5803000.0000		0.0644276	-0.5812566	-0.9342343	22		9	66 238011.54284	3 59231.196101		0E-7 0.0000702	37	94 354458.45192					
106		000 5775000.0000 000	0.2720604	0.0644276	1.3206150	2.5081245	23			286881.67317	z 57734.062936		0.0000010	173378.08689	58 400385.25946					-
		000	0.5730084						furnishingstøt	144245.96206				33947.390497	03 254544.53363					
107		5740000.0000 000 5652500.0000	-0.2438522	1.4513175	-0.5812566	0.2132186	25		us	41	3	2.5710542	0.0105032		08					

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

- price (p-value = 0.0)
- area (p-value = 0.0)
- bathrooms (p-value = 0.0)
- stories (p-value = 0.0)
- mainroad (p-value = 0.011078305445895668)
- guestroom (p-value = 0.010553358931198279)
- basement (p-value = 0. 01994612978933753)
- hotwaterheating (p-value = 4.1889240314438027E-4)
- airconditioning (p-value = 0.0)
- parking (p-value = 7.015384208333902E-5)
- prefarea (p-value = 1.0032449728224924E-6)
- furnishingstatus (p-value = 0.010503202918605318)

#### 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".

IMPORT	TRAIN_TEST_SPUT	NCRIMILIZE JRA	N SET	ECTON, NEW BROW																
		NCRMALIZE, TE	it, st	LECION (TAIN, SET																
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9		Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	Col9	Col
lser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basen	User Header	User Row ID									
1		13300000.000		1.4513175	1.3206150	1.3606716	0.4019153	-0.4771480	-0.744	2										_
2		12250000.000	1.7183192	1.4513175	5.1243580	2.5081245	0.4019153	-0.4771480	-0.744	4										
3		12250000.000 0000	2.1725255	0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	1.340	6 7										
4		11410000.000 0000	1.0188414	1.4513175	-0.5812566	0.2132186	0.4019153	2.0906616	1.340	8										
5		10850000.000 0000 10150000.000	1.0551779	0.0644276	3.2224865	-0.9342343	0.4019153	-0.4771480	1.340	10										
6		10150000.000 0000 10150000.000		1.4513175	3.2224865	2.5081245	0.4019153	-0.4771480	-0.744	11										
7		0000 9800000.0000	5.0067730	2.8382074	3.2224865	0.2132186	0.4019153	-0.4771480	-0.744	13 14										
8		9681000.0000	0.3738684	1.4513175	3.2224865	0.2132186	0.4019153	2.0906616	1.340	15 16										
10		9310000.0000	0.6236819	1.4513175	1.3206150	0.2132180	0.4019153	-0.4771480	-0.744	17 18										
11		9240000 0000		1.4513175	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744	19 20										
12		9240000.0000		0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744	21 22										
13		9100000.0000 000		1.4513175	1.3206150	0.2132186	0.4019153	2.0906616	1.340	23 24										
			1 5002042							25										

IMPORI INNIN\_TEST\_SPLIT NORMALIZE\_TRAIN\_SET NORMALIZE\_TEST\_SET PERTURE\_SELECTION\_REDRESSION PERTURE\_SELECTION\_TRAIN\_SET +

Manipulate the data by choosing the columns that correspond to the significant features (from the previous step) by browsing: "Data Transformation"  $\rightarrow$  "Data Manipulation"  $\rightarrow$  "Select Column(s)".

File Edit	Data Transforma	tion Analytic	cs Statistics	Plot Hel	р		
	Normalizers	•					
IMPORT	Data Manipulatio	Remove	Column(s)				
_	Split		olumn(s)				
	Variable Selectio	Matrix	ranspose				
		Sort by					
		Fill Missi	ing Column(s) Va	lues			
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad
		13300000.000	1.0188414	1.4513175	1.3206150	1.3606716	0.4019153
2		12250000.000	1.7183192	1.4513175	5.1243580	2.5081245	0.4019153
3		12250000.000	2.1725255	0.0644276	1.3206150	0.2132186	0.4019153
4		0000 11410000.000 0000	10100414	1.4513175	-0.5812566	0.2132186	0.4019153
		0000 10850000.000					
5		0000	1.0551779	0.0644276	3.2224865	-0.9342343	0.4019153
6		10150000.000 0000	1.5457208	1.4513175	3.2224865	2.5081245	0.4019153
7		10150000.000	5.0067730	2.8382074	3.2224865	0.2132186	0.4019153
8		9800000.0000	3.6441540	0.0644276	-0.5812566	0.2132186	0.4019153
9		000 9681000.0000 000		1.4513175	3.2224865	0.2132186	0.4019153
		000	0.5730004				
10		9310000.0000 000	0.6236819	1.4513175	1.3206150	0.2132186	0.4019153
11		9240000.0000 000	-0.7616474	1.4513175	1.3206150	0.2132186	0.4019153
12		9240000.0000 000	1.1914398	0.0644276	1.3206150	0.2132186	0.4019153
13		9100000.0000	0.6463922	1.4513175	1.3206150	0.2132186	0.4019153
14		000 8960000.0000		0.0644276	1 2206150	2 5001245	0.4010152
	AIN_TEST_SPLIT						
TR	and_rest_spen	TROTTINALIZE_IT	NUN NOR	mmenec_1co	PERIOR	_state non_k	CONCODION

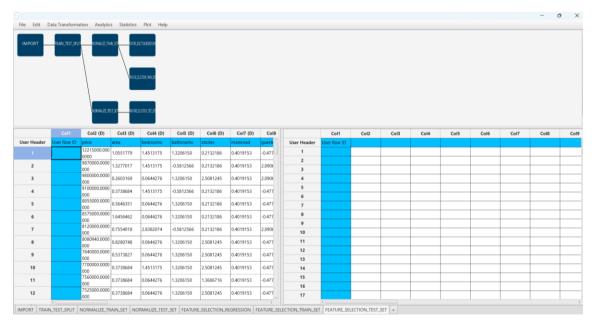
The results will appear on the output spreadsheet.

File Edit D	ata Transforma	tion Analytic	s Statistics	Plot Help													- 0	×
	TRAIN_TEST_SPLI	NORMALEE, FRA NORMALIEE, FRA		Internation														
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Coli
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestroom	basen	User Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroom	basem
1		13300000.000 0000	1.0188414	1.4513175	1.3206150	1.3606716	0.4019153	-0.4771480	-0.744	1		13300000.000 0000	1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441
2		12250000.000	1.7183192	1.4513175	5.1243580	2.5081245	0.4019153	-0.4771480	-0.744	2		0000	1./103192	5.1243580	2.5081245	0.4019153	-0.4771480	-0.7441
3		12250000.000	2.1725255	0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	1.340	3		12250000.000	2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480	1.34050
4		11410000.000 0000		1.4513175	-0.5812566	0.2132186	0.4019153	2.0906616	1.340	4		11410000.000	1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616	1.34050
5		10850000.000	1.0551779	0.0644276	3.2224865	-0.9342343	0.4019153	-0.4771480	1.340	5		10950000.000	1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	1.34050
6		10150000.000		1.4513175	3.2224865	2.5081245	0.4019153	-0.4771480	-0.744	6		10150000.000	1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480	-0.7441
7		10150000.000	5.0067730	2.8382074	3.2224865	0.2132186	0.4019153	-0.4771480	-0.744	7		10150000.000	5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480	-0.7441
8		9800000.0000 000		0.0644276	-0.5812566	0.2132186	0.4019153	-0.4771480	1.340	8		9800000.0000	3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480	1.34050
9		9681000.0000	0.3738684	1.4513175	3.2224865	0.2132186	0.4019153	2.0906616	1.340	9		9681000.0000		3.2224865	0.2132186	0.4019153	2.0906616	1.34050
10		9310000.0000 000	0.6236819	1.4513175	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744	10		9310000.0000	0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441
11		9240000.0000 000	-0.7616474	1.4513175	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744	11		9240000.0000	-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441
12		9240000.0000 000	1.1914398	0.0644276	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744	12		9240000.0000	1.1914398	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441
13		9100000.0000		1.4513175	1.3206150	0.2132186	0.4019153	2.0906616	1.340	13		9100000.0000 000		1.3206150	0.2132186	0.4019153	2.0906616	1.34050
			1 5002042									8960000.0000			-			+

#### 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 (from the step 6) by browsing: "Data Transformation"  $\rightarrow$  "Data Manipulation"  $\rightarrow$  "Select Column(s)".

le Edit D	ata Transforma	tion Analytic	s Statistics	Plot Help				
N	ormalizers	-						
	ata Manipulatio	n 🕨 Remove	Column(s)					
Sp	olit	Select Co	olumn(s)					
Vi	ariable Selection	n 🔸 Matrix Tr	ranspose					
		Sort by C	Column					
		Fill Missi	ng Column(s) st.gt	Values 94:00(187)#				
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Coli
ser Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestr
		12215000.000 0000	1.0551779	1.4513175	1.3206150	0.2132186	0.4019153	-0.477
2		9870000.0000 000	1.3277017	1.4513175	-0.5812566	0.2132186	0.4019153	2.090€
3		9800000.0000 000	0.2603169	0.0644276	1.3206150	2.5081245	0.4019153	2.090€
4		9100000.0000 000	0.3738684	1.4513175	-0.5812566	0.2132186	0.4019153	-0.477
5		8855000.0000 000	0.5646351	0.0644276	1.3206150	0.2132186	0.4019153	-0.477
· ·		8575000.0000		0.0511076	1.3206150	0.2132186	0.4019153	-0.477
6		000	1.6456462	0.0644276	1.5200150	0.2132100		_
		000 8120000.0000 000	0.7554018	2.8382074	-0.5812566	0.2132186	0.4019153	2.090€
6		000 8120000.0000 000 8080940.0000 000					0.4019153	2.090€ -0.477
6 7		000 8120000.0000 000 8080940.0000 000 7840000.0000 000	0.7554018	2.8382074	-0.5812566	0.2132186		
6 7 8		000 8120000.0000 000 8080940.0000 000 7840000.0000 000 7700000.0000 000	0.7554018	2.8382074 0.0644276	-0.5812566 1.3206150	0.2132186	0.4019153	-0.477
6 7 8 9		000 8120000.0000 000 8080940.0000 000 7840000.0000 000 7700000.0000	0.7554018 0.8280748 0.5373827	2.8382074 0.0644276 0.0644276	-0.5812566 1.3206150 1.3206150	0.2132186 2.5081245 2.5081245	0.4019153	-0.477

Excluded Columns -		Included Columns
Col4 bedrooms	>>	Col2 price
		Col3 area
	>	Col5 bathrooms
		Col6 stories
	<	Col7 mainroad
		Col8 guestroom
	<<	Col9 basement
		Col10 hotwaterheating
		Coll11 airconditioning

File Edit D		c															-	o ×
IMPORT	TRAIN, TEST, SPUT	NCRIMILEE, FRA	N, ST BAR S	ECTOR (FARESON EECTOR (FARESON EECTOR (FARESON)														
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9
User Header	User Row ID	price	area	bedrooms	bathrooms	stories	mainroad	guestr	User Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroom	basement	hotwate
		12215000.000 0000		1.4513175	1.3206150	0.2132186	0.4019153	-0.477	1		12215000.000	1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438
2		9870000.0000 000	1.3277017	1.4513175	-0.5812566	0.2132186	0.4019153	2.090€	2		9870000.0000	1.3277017	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438
3		9800000.0000 000	0.2603169	0.0644276	1.3206150	2.5081245	0.4019153	2.0906	3		000 9800000.0000 000	0.3603160	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.2438
4		9100000.0000	0.3738684	1.4513175	-0.5812566	0.2132186	0.4019153	-0.477			000 9100000.0000 000	0.2003103						
5		8855000.0000	0.5646351	0.0644276	1.3206150	0.2132186	0.4019153	-0.477	4		000	0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438
6		000 8575000.0000 000		0.0644276	1.3206150	0.2132186	0.4019153	-0.477	5		8855000.0000 000	0.5646351	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438
		000 8120000.0000	1.0430402					-	6		8575000.0000 000		1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438
7				2.8382074	-0.5812566	0.2132186	0.4019153	2.090€	7		8120000.0000 000	0.7554018	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438
8		8080940.0000 000		0.0644276	1.3206150	2.5081245	0.4019153	-0.477	8		8080940.0000 000	0.8280748	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438
		7840000.0000 000	0.5373827	0.0644276	1.3206150	2.5081245	0.4019153	-0.477	9		7840000.0000	0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438
9						2.5081245	0.4019153	-0.477	10		000 7700000.0000 000	0 3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438
9 10		7700000.0000	0.3738684	1.4513175	1.3206150													
		7560000.0000	0.3738684	0.0644276	1.3206150	1.3606716	0.4019153	-0.477			000 7560000.0000							
10		7700000.0000 000 7560000.0000 000 7525000.0000 000	0.3738684				0.4019153	-0.477	11		000 7560000.0000 000 7525000.0000	0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438

## 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".

| N_TEST_SPUT | NORMALIZE THE  | WIER  | EDDA (TRAN, EE   | RAIN, MODELLÁN   
   
  |  
   
  |  
  |  
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   |   |
|-------------|--|---|--
--
---
--
---
--
---
--
---
--
---
--
--|--
---
---
---	---
---	---
Col1	Col2 (D)
   
  | Col6 (D)   
   
  | Col7 (D)   
  | Col8 (D)   
   
  |  
  |  
   | Col1  
  | Col2   
  | Col3  | Col4   
  | Col5  | Col6   
  | Col7  | Col8  | 8   |
|             |  |   | bathrooms  | stories  
   
  | mainroad   
   
  | guestroom  
  | basement   
   
  |  
  | User Header  
   | User Row ID   
  |  
  |   |  
  |   |  
  |   |   |   |
| 13          | 3300000.000  | 1.0188414   | 1.3206150  | 1.3606716  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  | · .  
  | 1  
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 12          | 2250000.000  | 1 7183192   | 5 1243580  | 2 5081245  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 00          | 2250000.000  | 11100102  |  |  
   
  |  
   
  |  
  |  
   
  |  
  | 4  
   |   
  |  
  |   |  
  |   |  
  |   |   | -   |
| 00          | 000  | 2.1725255   | 1.3206150  | 0.2132186  
   
  | 0.4019153  
   
  | -0.4771480   
  | 1.3405094  
   
  |  
  | 5  
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 11          | 1410000.000  | 1.0188414   | -0.5812566   | 0.2132186  
   
  | 0.4019153  
   
  | 2.0906616  
  | 1.3405094  
   
  |  
  | 6  
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 10          | 0850000.000  | 1.0551779   | 3.2224865  | -0.9342343   
   
  | 0.4019153  
   
  | -0.4771480   
  | 1.3405094  
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   |   |
| 10          | 0150000.000  | 1.5457208   | 3.2224865  | 2.5081245  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 10          | 0150000.000  | 5 0067730   | 3 2224865  | 0.2132186  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  | 10   
   |   
  |  
  |   |  
  |   |  
  |   |   | -   |
| 00          | 800000.0000  |   |  |  
   
  |  
   
  |  
  |  
   
  |  
  | 11   
   |   
  |  
  |   |  
  |   |  
  |   |   | _   |
| 00          | 00   | 3.6441540   |  |  
   
  |  
   
  |  
  | 1.3405094  
   
  |  
  | 12   
   |   
  |  
  |   |  
  |   |  
  |   |   |   |
| 90          | 00   | 0.3738684   | 3.2224865  | 0.2132186  
   
  | 0.4019153  
   
  | 2.0906616  
  | 1.3405094  
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   |   |
| 93          | 310000.0000<br>00  | 0.6236819   | 1.3206150  | 0.2132186  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  |  
   |   
  |  
  |   |  
  |   |  
  |   |   | -   |
| 97          | 240000.0000  | -0.7616474  | 1.3206150  | 0.2132186  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  | 16   
   |   
  |  
  |   |  
  |   |  
  |   |   | -   |
|             | 240000.0000  |   |  | -  
   
  | 0.4019153  
   
  | -0.4771480   
  | -0.7441611   
   
  |  
  | 17   
   |   
  |  
  |   |  
  |   |  
  |   |   | -   |
| •           | Coll P<br>Row ID P<br>1<br>1<br>1<br>1<br>0<br>1<br>1<br>1<br>0<br>1<br>1<br>0<br>1<br>1<br>0<br>1<br>1<br>0<br>0<br>1<br>1<br>0<br>0<br>1<br>0<br>1<br>0<br>0<br>9<br>9<br>9<br>9 | Coll Col2 (D)<br>Food D Price<br>Col2 Col2 (D)<br>Food D Price<br>Col2 Col2 (D)<br>Food D Price<br>Col2 Col2 (D)<br>Food D Price<br>Col2 Col2 (D)<br>Col2 | Coll         Coll & Coll | Coll         Coll <th< td=""><td>Cold         Cold (D)         <thcold (d)<="" th="">         Cold (D)         <thc< td=""><td>Cold         Cold 2 (D)         Cold 2 (D)<td>Cold         Cold         <th< td=""><td>Cold         Cold (D)         <thcold (d)<="" th="">         Cold (D)         <thc< td=""><th>Coll         Coll (D)         <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll></th></thc<></thcold></td></th<></td></td></thc<></thcold></td></th<> | Cold         Cold (D)         Cold (D) <thcold (d)<="" th="">         Cold (D)         <thc< td=""><td>Cold         Cold 2 (D)         Cold 2 (D)<td>Cold         Cold         <th< td=""><td>Cold         Cold (D)         <thcold (d)<="" th="">         Cold (D)         <thc< td=""><th>Coll         Coll (D)         <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll></th></thc<></thcold></td></th<></td></td></thc<></thcold> | Cold         Cold 2 (D)         Cold 2 (D) <td>Cold         Cold         <th< td=""><td>Cold         Cold (D)         <thcold (d)<="" th="">         Cold (D)         <thc< td=""><th>Coll         Coll (D)         <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll></th></thc<></thcold></td></th<></td> | Cold         Cold <th< td=""><td>Cold         Cold (D)         <thcold (d)<="" th="">         Cold (D)         <thc< td=""><th>Coll         Coll (D)         <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll></th></thc<></thcold></td></th<> | Cold         Cold (D)         Cold (D) <thcold (d)<="" th="">         Cold (D)         <thc< td=""><th>Coll         Coll (D)         <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll></th></thc<></thcold> | Coll         Coll (D)         Coll (D) <thcoll (d)<="" th="">         Coll (D)         <thc< th=""><td>Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)&lt;</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<></td></thc<></thcoll> | Coll         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Multicity           Coll         Coll 2(0)         Coll 2(0)         Coll 2(0)         Coll 4(0)         Coll 5(0)         Coll 6(0)         Coll 7(0)         Coll 8(0)         Formation           Nov 10         price         area         bathocoms         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         mainwald         Coll 7(0)         Coll 8(0)         Formation         Stories         Coll 8(0)         Coll 8(0)< | Coll         Coll <th< td=""><td>Coll         Coll         <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<></td></th<> | Coll         Coll <th< td=""><td>Cold (C)         Cold (C)</td><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<></td></th<> | Cold (C)         Cold (C) | Coll         Coll <th< td=""><td>Coll         Coll         <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<></td></th<> | Coll         Coll <th< td=""><td>Coll         Coll         <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<></td></th<> | Coll         Coll <th< td=""><td>Coll (0)         Coll (0)</td><td>Coll         Coll         <th< td=""></th<></td></th<> | Coll (0)         Coll (0) | Coll         Coll <th< td=""></th<> |

Use the k Nearest Neighbors (kNN) method to train and fit the model by browsing: "Analytics"  $\rightarrow$  "Regression"  $\rightarrow$  "k Nearest Neighbors (kNN)" and set the "Target Column" as the column corresponding to "price" and the "Number of Neighbors" to 5.

		_							
File Edit	Data Transformat	ion Analytic	s Statistics	Plot Help					
		Regressi		k-Near	est Neighbors	(kNN)			
IMPORT	TRAIN_TEST_SPLIT	Classific	ation	Fully C	onnected Neur	al Network			
		Clusterin	ng	Radial	Basis Function	Network			
		Anomah	y Detection	Linear	SGD				
		Existing	Model Utilizati	ion XGBoo	st				
				Rando	m Forest				
				-					
		NORMALIZE TE		EECTION TEST SET				A	
		WURIWALIZE_IE	51,30	0.000,00,0				🕀 kNN Regression Model	×
		_		_					
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Target Column Col2 price	-
	User Row ID	price	area	bathrooms	stories	mainroad	guestroom	raiger ooldinin eol2 price	÷.,
1		13300000.000 0000	1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480		_
2		12250000.000 0000	1.7183192	5.1243580	2.5081245	0.4019153	-0.4771480	Number of Neighbors 5	
3		12250000.000 0000	2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480		_
4		11410000.000 0000	1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616		
5		10850000.000 0000	1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	Execute Cancel	
6		10150000.000 0000	1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480		
7		10150000.000 0000	5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480		
8		9800000.0000 000	3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480		
9		9681000.0000 000	0.3738684	3.2224865	0.2132186	0.4019153	2.0906616		
10		9310000.0000 000	0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480		
11		9240000.0000 000	-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480		
12	< (	9240000.0000	1.1914398	1.3206150	0.2132186	0.4019153	-0.4771480		
IMPORT TRAI	N_TEST_SPLIT	ORMALIZE_TF	RAIN_SET NO	RMALIZE_TEST	SET FEATUR	E_SELECTION_R	EGRESSION		

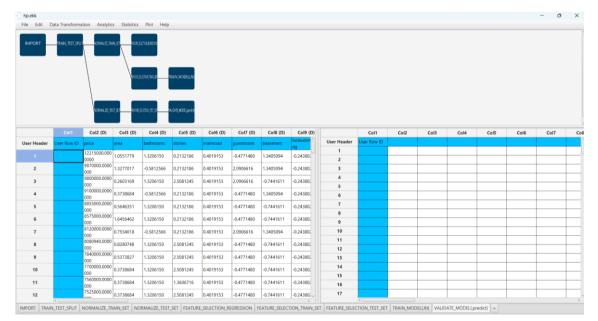
The predictions will appear on the output spreadsheet.

MPORT	TRAIN, TEST_SPLI	NCRIMULEE, TRAI		170(83815) 170(74/37	RAIN, MODELL HU												
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)		Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (S)	Col7 (D)	Col8 (S)
	User Row ID		area	bathrooms	stories	mainroad	guestroom	basement	User Header	User Row ID		kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN
1		13300000.000 0000	1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	1			13188929.411		0E-7	Entry 48	0.0741518	Entry 27
2		12250000.000		5.1243580	2.5081245	0.4019153	-0.4771480	-0.7441611	2		12250000.000	12229889.412		0E-7	Entry 30	0.7453612	Entry 14
3		12250000.000	2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	3		12250000.000	12205901.173		0E-7	Entry 105	0.5148770	Entry 213
4		11410000.000	1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	4		11410000.000	11363133.239		0E-7	Entry 311	0.3809051	Entry 109
5		10850000.000	1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	1.3405094	5		10850000.000	10826633.754 9854	Entry 5	0E-7	Entry 24	0.5940605	Entry 93
6		10150000.000	1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480	-0.7441611	6			10145159.949		0E-7	Entry 1	0.6919589	Entry 48
7		10150000.000	5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480	-0.7441611	7			10119807.531 6795		0E-7	Entry 304	0.7793816	Entry 361
8		9800000.0000	3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	8		9800000.0000	9781644.4849		0E-7	Entry 49	0.4922288	Entry 65
9		9681000.0000	0.3738684	3.2224865	0.2132186	0.4019153	2.0906616	1.3405094	9		9681000.0000 000	9664271.3162 595	Entry 9	0E-7	Entry 124	0.9579445	Entry 148
10		9310000.0000 000	0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	10		9310000.0000 000	9292345.9693	Entry 10	0E-7	Entry 114	0.4948162	Entry 27
11		9240000.0000 000	-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	11		9240000.0000 000	9205858.5426 032	Entry 11	0E-7	Entry 208	0.4765667	Entry 242
		9240000.0000	1 101 1000	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	12		9240000.0000	9198353.0763	Entry 12	0E-7	Entry 172	0.3732271	Entry 231

# 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 browse: "Analytics"  $\rightarrow$  "Existing Model Utilization". Then choose Model "(from Tab:) TRAIN\_MODEL (.fit)" and transfer the "price" column in the output.

IMPORT	TRAIN_TEST_SPL	Regressi Classific Clusterin Anomal	on ation 19 7 Detection Model Utilizat	Plot Help	RAIN_MODEL(St)			
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	(
User Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroom	bas
		12215000.000 0000	1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3
2		9870000.0000 000	1.3277017	-0.5812566	0.2132186	0.4019153	2.0906616	1.3
3		9800000.0000	0.2603169	1.3206150	2.5081245	0.4019153	2.0906616	-0.7
4		000 9100000.0000	0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3
-		000	0.5750004	*0.3812300	0.2132100	0.4019133	*0.4771400	1.3
5		8855000.0000 000	0.5646351	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7
6		8575000.0000 000	1.6456462	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7
7		8120000.0000	0.7554018	-0.5812566	0.2132186	0.4019153	2.0906616	1.3
8		000 8080940.0000	0.8280748	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7
9		000 7840000.0000 000	0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.
		7700000.0000	0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.
10					1			-
10 11		000 7560000.0000 000	0.3738684	1.3206150	1.3606716	0.4019153	-0.4771480	-0.

Model	(from Tab: )TRAIN	N_MODEL( 👻
Туре	kNN Model	
escription		
odel Input		
Header -> Datatype area -> Double bedrooms -> Double bathrooms -> Double stories -> Double mainroad -> Double		
guestroom -> Double basement -> Double hotwaterheating -> Do insertitisation -> Do Transfer Column(s) to Excluded Columns	Output	Included Columns -
Assement -> Double hotwaterheating -> Do liseanditioning -> Do Transfer Column(s) to	Output	Col2 price
Assement -> Double hotwaterheating -> Do Transfer Column(s) to Excluded Columns	Output	Col2 price
Assement -> Double hotwaterheating -> Do Transfer Column(s) to Excluded Columns Col3 area	Output	Col2 price
Desement -> Double     notwaterheating -> Dr     Transfer Column(s) to     Excluded Columns     Col3 area     Col4 bathrooms	Output	Col2 price
assement -> Double notwaterheating -> Do Transfer Column(s) to Excluded Column Col3 area Col4 bathrooms Col5 stories	Output	Col2 price
assement → Double hotwaterheating → Dr Transfer Column(s) to Excluded Columns Col3 area Col4 bathrooms Col5 stories Col6 mainroad	Output	Col2 price
asement -> Double howaterheating -> D Transfer Column(s) to Eccluded Columns Col3 area Col4 bathrooms Col5 stories Col5 maincad Col7 guestroom	Output	Col2 price
asement -> Double Ontwaterheating -> D Construction Excluded Columns Col3 area Col4 bathrooms Col5 stories Col6 maincoad Col7 guestoom Col8 basement	Output	Col2 price

IMPORT	ata Transforma TRAIN_TEST_SPU	tion Analytics		Plot Help													- 0	×
		NORMALIZE TE			TRAIN, MODELL fit)													
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)		Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (D)	Col7 (S)	Co
Jser Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroom	basement	hotwater	User Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1	Closest NN2	Distance from NN2	Closest NN3	Dista NN3
1		12215000.000	1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.24380;	1		8394729.1211	Entry 8	0.6174273	Entry 5	0.7637626	Entry 93	0.76
2		9870000.0000	1.3277017	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.24380;	2		6824551.4683	Entry 20	0.3507424	Entry 61	0.3732271	Entry 115	0.692
3		980000.0000	0 2603169	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.243802	3		6439216.7259	Entry 120	0.7641550	Entry 31	1.1056799	Entry 56	1.10
4		000 9100000.0000 000	0 2728694	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.243802	4		801 4813236.0342 003	Entry 224	0.3602715	Entry 70	0.3625968	Entry 176	0.36
		000 8855000.0000 000	0.5750004									003 9151949.6764 435	Endy 254					-
5		9575000.0000		1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.243802	5		435	Entry 10	0.0090941	Entry 114	0.4921283	Entry 27	0.51
6		000	1.6456462	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.24380;	6		5759676.3731 597		0.3437411	Entry 47	0.3783940	Entry 132	0.47
7		8120000.0000 000	0.7554018	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.243802	7		6462229.4133 875	Entry 311	0.1437566	Entry 109	0.1497027	Entry 4	0.33
8		8080940.0000 000		1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.243802	8		8305353.8342 013	Entry 14	0.1049318	Entry 30	0.1399091	Entry 35	0.34
		7840000.0000 000	0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.243802	9		7808748.8221	Entry 59	0.3333598	Entry 60	0.3334772	Entry 48	0.74
9		7700000 0000	0.3730604	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.243802	10		5409578.3681 111	Entry 75	0.3333333	Entry 96	0.6020464	Entry 286	0.67
9 10		000	0.3730084															
-		7700000.0000 000 7560000.0000 000	0.3738684	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	-0.24380;	11		6017173.8042 042	Entry 66	0.3333333	Entry 103	0.3379986	Entry 32	0.47

# 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".

hp.ekk																-	0	>
ile Edit D	ata Transforma	tion Analytic	s Statistics	Plot Help														
IMPORT	TRAIN, TEST, SPLI	NCRIMULES, FR		14'00, 1888-501 Extor, 1944, 59 Extor, 1975, 59	TRAIN_IMODELL_FA)	STATSTICS, ACCURACI												
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (D)	Col7 (S)	Col8 (D)	Col9 (S)		Col1	Col2	Col3	Col4	Col5	Col6	Col7	_
User Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1	n Closest NN2	Distance from NN2	Closest NN3	Distance from NN3	Closest NN4	User Header	User Row ID							
1		8394729.1211 079	Entry 8	0.6174273	Entry 5	0.7637626	Entry 93	0.7645004	Entry 65	1								-
2		6824551.4683 578	Entry 20	0.3507424	Entry 61	0.3732271	Entry 115	0.6929062	Entry 94	3								
3		6439216.7259	Entry 120	0.7641550	Entry 31	1.1056799	Entry 56	1.1067204	Entry 78	4								_
4		4813236.0342	Entry 234	0.3602715	Entry 70	0.3625968	Entry 176	0.3648357	Entry 192	5								-
5		9151949.6764 435	Entry 10	0.0090941	Entry 114	0.4921283	Entry 27	0.5114741	Entry 59	7								
6		5759676.3731	Entry 144	0.3437411	Entry 47	0.3783940	Entry 132	0.4718113	Entry 53	8								_
7		597 6462229.4133 875	Entry 311	0.1437566	Entry 109	0.1497027	Entry 4	0.3357936	Entry 18	10								
8		8305353.8342	Entry 14	0.1049318	Entry 30	0.1399091	Entry 35	0.3405947	Entry 41	11								
9		7808748.8221	Entry 59	0.3333598	Entry 60	0.3334772	Entry 48	0.7453560	Entry 1	12								
		5409578.3681	Entry 75	0.3333333	Entry 96	0.6020464	Entry 286	0.6793799	Entry 188	14								
10		111								15								
10		6017173.8042 042	Entry 66	0.3333333	Entry 103	0.3379986	Entry 32	0.4714045	Entry 150	16								

Calculate the statistical metrics for the regression by browsing: "Statistics"  $\rightarrow$  "Model Metrics"  $\rightarrow$  "Regression Metrics".

hp.ekk File Edit [	Data Transform	ation Analytic	s Statistics	Plot Help				
IMPORT	TRAIN_TEST_SPL		Domain -	APD			_	
			Probabilit	y Distribution F		egression Metri assification Me		
				e Statistics ce Intervals is Testing	•			
		NORMALIZE, TE		ases Number Genera Experiments	itor	STATISTICS ACCURAC	es	
	Col1	Col2 (D)	Col3 (5)	Col4 (D)	Col5 (S)	Col6 (D)	Col7 (S)	Col8
User Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1	<sup>n</sup> Closest NN2	Distance from NN2	<sup>1</sup> Closest NN3	Distance NN3
		8394729.1211 079	Entry 8	0.6174273	Entry 5	0.7637626	Entry 93	0.76450
2		6824551.4683	Entry 20	0.3507424	Entry 61	0.3732271	Entry 115	0.69290
3		6439216.7259	Entry 120	0.7641550	Entry 31	1.1056799	Entry 56	1.10672
4		4813236.0342	Entry 234	0.3602715	Entry 70	0.3625968	Entry 176	0.36483
5		9151949.6764	Entry 10	0.0090941	Entry 114	0.4921283	Entry 27	0.51147
6		5759676.3731	Entry 144	0.3437411	Entry 47	0.3783940	Entry 132	0.47181
7		6462229.4133	Entry 311	0.1437566	Entry 109	0.1497027	Entry 4	0.33579
8		8305353.8342	Entry 14	0.1049318	Entry 30	0.1399091	Entry 35	0.34059
		7808748.8221	Entry 50	0.3333598	Entry 60	0.3334772	Entry 48	0.74535
9								
9 10		5409578.3681 111	Entry 75	0.3333333	Entry 96	0.6020464	Entry 286	0.67937
			Entry 75 Entry 66	-		0.6020464	Entry 286 Entry 32	0.67937

The results will appear on the output spreadsheet.

le Edit D	ata Transforma	tion Analytic	s Statistics	Plot Help														
IMPORT	TRAIN_TEST_SPLF	T CONALLE, TA	914.2		RAIN, MODELL FR	SPITSTICS ACORACE												
	Col1	Col2 (D)	Col3 (S)	Col4 (D)	Col5 (S)	Col6 (D)	Col7 (S)	Col8 (D)	Col9 (S)		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6	Col7	-
ser Header	User Row ID	kNN Prediction	Closest NN1	Distance from NN1		Distance from NN2		Distance from NN3	Closest NN4	User Header	User Row ID	Mean Squared	Root Mean	Mean Absolute Error	R Squared			
1		8394729.1211 079	Entry 8	0.6174273	Entry 5		Entry 93	0.7645004	Entry 65	1		14108899052 94.4163000	1187808.8673 244	853056.98213 28	0.5852034			
2		6824551.4683	Entry 20	0.3507424	Entry 61	0.3732271	Entry 115	0.6929062	Entry 94	2								
3		6439216.7259	Entry 120	0.7641550	Entry 31	1.1056799	Entry 56	1.1067204	Entry 78	3								-
4		4813236.0342	Entry 234	0.3602715	Entry 70	0.3625968	Entry 176	0.3648357	Entry 192	5								
5		9151949.6764	Entry 10	0.0090941	Entry 114	0.4921283	Entry 27	0.5114741	Entry 59	6								
6		5759676.3731	Entry 144	0.3437411	Entry 47	0.3783940	Entry 132	0.4718113	Entry 53	7 8								-
7		6462229.4133	Entry 311	0.1437566	Entry 109	0.1497027	Entry 4	0.3357936	Entry 18	9								-
8		8305353.8342	Entry 14	0.1049318	Entry 30		Entry 35	0.3405947	Entry 41	10								
9		013 7808748.8221 759	Entry 50	0.3333598	-	0.3334772	-	0.7453560		11								-
		759 5409578.3681 111	chuy 39		Entry 60	-	Entry 48		Entry 1	13								
10		111 6017173.8042	Entry 75	0.3333333	Entry 96	0.6020464	Entry 286	0.6793799	Entry 188	14								
11		6017173.8042 042 7563332.4143	Entry 66	0.3333333	Entry 103	-	Entry 32	0.4714045	Entry 150	15								
				0E-7	Entry 14	0.3764257	Entry 30	0.3938957	Entry 68	16								

# Step 12: Reliability check of 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\_PRICE".

Import data into the input spreadsheet of the "EXCLUDE\_PRICE" tab from the output of the "FEATURE\_SELECTION\_TRAIN\_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from SpreadSheet".

hp.ekk ile Edit D	ata Transforma	tion Analytic	s Statistics	Plot Help	,											-	0	
IMPORT	TRAIN_TEST_SPUT	NJRIMUZEJR			TRAIN, MODEL, SK.													
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	î	Col1	Col2	Col3	Col4	Col5	Col6	Col7	Col8	
ser Header	User Row ID		area	bathrooms	stories	mainroad	guestroom	basement 🦳	User Header	User Row ID								
1		13300000.000	1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	1									
2		0000 12250000.000 0000	1 7192102	5.1243580	2.5081245	0.4019153	-0.4771480	-0.7441611	2									
		0000 12250000.000 0000	1.7103132						4									-
3		0000	2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	5									_
4		11410000.000 0000	1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	6									_
5		10850000.000	1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	1.3405094	7									
6		10150000.000	1 5457208	3.2224865	2.5081245	0.4019153	-0.4771480	-0.7441611	8									
		0000 10150000.000 0000						_	9								ļ	
7		0000	5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480	-0.7441611	10									-
8		9800000.0000 000	3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	12									-
9		9681000.0000 000	0.3738684	3.2224865	0.2132186	0.4019153	2.0906616	1.3405094	13									-
		9310000.0000	0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	14									
10				-			-0.4771480	-0.7441611	15									_
		9240000.0000	0.7616474	1 2206150														
10 11 12		9240000.0000 000 9240000.0000	-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	16									_

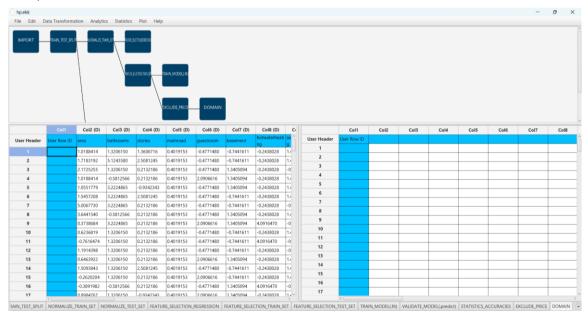
Manipulate the data to exclude the column that corresponds to the "price" by browsing: "Data Transformation"  $\rightarrow$  "Data Manipulation"  $\rightarrow$  "Select Columns". Then select all the columns except the "price".

File Edit	Data Transforma	ation Analytic	s Statistics	Plot Help			
	Vormalizers	•					
	Data Manipulatio		Column(s)				
	Split	Hemore					
		Select Co	olumn(s)				
\ \	/ariable Selectio	n 🕨 Matrix Ti	anspose				
		Sort by C	olumn		RAIN_MODEL(.fit)		
		Fill Missi	ng Column(s) ۱		XCLUDE_PRICE		
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D
User Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroom
1		13300000.000 0000	1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480
2		12250000.000 0000	1.7183192	5.1243580	2.5081245	0.4019153	-0.4771480
3		12250000.000 0000	2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480
4		11410000.000 0000	1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616
5		10850000.000 0000	1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480
6		10150000.000 0000	1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480
6 7 8		10150000.000 0000	5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480
		9800000.0000	3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480
8		000	5.0441540				
8		9681000.0000 000	0.3738684	3.2224865	0.2132186	0.4019153	2.0906616
		9681000.0000 000 9310000.0000 000			0.2132186	0.4019153	2.0906616
9		9681000.0000 000 9310000.0000	0.3738684	3.2224865			

Excluded Columns		Included Columns
Col2 price	>>	Col3 area
		Col4 bathrooms
	>	Col5 stories
		Col6 mainroad
	<	Col7 guestroom
		Col8 basement
	<<	Col9 hotwaterheating
		Col10 airconditioning
		Call1 parking

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\_PRICE" tab by right-clicking on the input spreadsheet and then choosing "Import from SpreadSheet".



Create the domain of applicability by browsing: "Statistics"  $\rightarrow$  "Domain APD".

hp.ekk							
File Edit	Data Transforma	tion Analyti	ics Statistics	Plot Help			
			Domain -	APD			
IMPORT	TRAIN_TEST_SPLI	NORMALIZE_TR	Nodel Me	trics	•		
			Probability	y Distribution F	unctions		
			Descriptiv	e Statistics			
			Confidenc	e Intervals	•		
			Hypothesi	s Testing			
			Weight Ca	-			
				lumber Genera	tor	_	
				Experiments		DOMAIN	
			Design of	experiments			
	_	1					
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Coli
User Header	User Row ID	area	bathrooms	stories	mainroad	guestroom	basem
1		1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480	-0.744
2		1.7183192	5.1243580	2.5081245	0.4019153	-0.4771480	-0.744
3		2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405
4		1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405
5		1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	1.3405
6		1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480	-0.744
7		5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480	-0.744
8		3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405
9		0.3738684	3.2224865	0.2132186	0.4019153	2.0906616	1.3405
10		0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744
11		-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744
12		1.1914398	1.3206150	0.2132186	0.4019153	-0.4771480	-0.744
13		0.6463922	1.3206150	0.2132186	0.4019153	2.0906616	1.3405
14		1.5093843	1.3206150	2.5081245	0.4019153	-0.4771480	-0.744
15		-0.2620204	1.3206150	0.2132186	0.4019153	2.0906616	-0.744
16		-0.3891982	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405
17		0.8984767	1 3206150	-0.9342343	0.4019153	2 0906616	1 3405

RAIN\_TEST\_SPLIT NORMALIZE\_TRAIN\_SET NORMALIZE\_TEST\_SET FEATURE\_SELECTION\_REGRESSION FEATURI

IMPORT	Data Transforma	ation Analyti		200,8850	IRAIN, MODELLAD													
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	EXCLUDE_PRICE	Col6 (D)	Col7 (D)	Col8 (D)	C		Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5	Col6	Col7	Col
ser Header	User Row ID	area	bathrooms	stories	mainroad	guestroom	basement	hotwaterheat	i air	User Header	User Row ID	Domain	APD	Prediction				
1		1.0188414	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	-0.2438028	9	1		0E-7	3.9052670	reliable				
2		1.7183192	5.1243580	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	12	2		0E-7	3.9052670	reliable				
3		2.1725255	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	-0	3		0E-7	3.9052670	reliable				
4		1.0188414	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	14	4		0E-7	3.9052670	reliable				
5		1.0551779	3.2224865	-0.9342343	0.4019153	-0.4771480	1.3405094	-0.2438028	12	5		0E-7	3.9052670	reliable				
6		1.5457208	3.2224865	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	14	6		0E-7	3.9052670	reliable				
7		5.0067730	3.2224865	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	-0	7		0E-7	3.9052670	reliable				
8		3.6441540	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	14	8		0E-7	3.9052670	reliable				
9		0.3738684	3.2224865	0.2132186	0.4019153	2.0906616	1.3405094	4.0916470	-0	9		0E-7	3.9052670	reliable				
10		0.6236819	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1/	10		0E-7	3.9052670	reliable				
11		-0.7616474	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	4.0916470	-0	11		0E-7	3.9052670	reliable				
12		1,1914398	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	-0	12		0E-7	3.9052670	reliable				
13		0.6463922	1.3206150	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	1/	13		0E-7	3.9052670	reliable				
		1.5093843	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	12	14		0E-7	3.9052670	reliable				
14		-0.2620204	1.3206150	0.2132186	0.4019153	2.0906616	-0.7441611	-0.2438028	14	15		0E-7	3.9052670	reliable				
14							1.3405094	4.0916470	-0	16		0E-7	3.9052670	reliable				
14 15 16		-0.3891982	-0.5812566	0.2132186	0.4019153	-0.4771480												

#### 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\_PRICE\_TEST\_SET".

Import data into the input spreadsheet of the "EXCLUDE\_PRICE\_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".

Filter the data to exclude the column that corresponds to the "price" by browsing: "Data Transformation"  $\rightarrow$  "Data Manipulation"  $\rightarrow$  "Select Columns". Then select all the columns except "price".

ile Edit	Data Transforma	ation Analytic	s Statistics	Plot Help			
	Normalizers	-					
IMPORT	Data Manipulatio	n  Remove	Column(s)				
	Split	Select Co	olumn(s)				
	Variable Selectio						
		1					
		Sort by (	Jolumn		RAIN_MODEL(.fit)		
		NORMALIZE, TE	ng Column(s) ' ST_SETRAUR:		XCLUDE_PRICE	DOMAIN	
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7
User Header	User Row ID	price	area	bathrooms	stories	mainroad	guestroo
		12215000.000	1.0551779	1.3206150	0.2132186	0.4019153	-0.47714
2		9870000.0000 000	1.3277017	-0.5812566	0.2132186	0.4019153	2.090661
3		9800000.0000	0.2603169	1.3206150	2.5081245	0.4019153	2.090661
4		9100000.0000	0.3738684	-0.5812566	0.2132186	0.4019153	-0.47714
5		8855000.0000 000	0.5646351	1.3206150	0.2132186	0.4019153	-0.47714
6		8575000.0000 000	1.6456462	1.3206150	0.2132186	0.4019153	-0.47714
7		8120000.0000 000	0.7554018	-0.5812566	0.2132186	0.4019153	2.090661
8		8080940.0000 000	0.8280748	1.3206150	2.5081245	0.4019153	-0.47714
		7840000.0000 000	0.5373827	1.3206150	2.5081245	0.4019153	-0.47714
9		7700000.0000	0.3738684	1.3206150	2.5081245	0.4019153	-0.47714
9 10		000					
-		000 7560000.0000 000	0.3738684	1.3206150	1.3606716	0.4019153	-0.47714

Excluded Columns	1	Included Columns	-
Col2 price	>>	Col3 area	î
		Col4 bathrooms	
	>	Col5 stories	
		Col6 mainroad	
	<	Col7 guestroom	
		Col8 basement	I
	<<	Col9 hotwaterheating	
		Col10 airconditioning	
		Coll11 parking	$\sim$

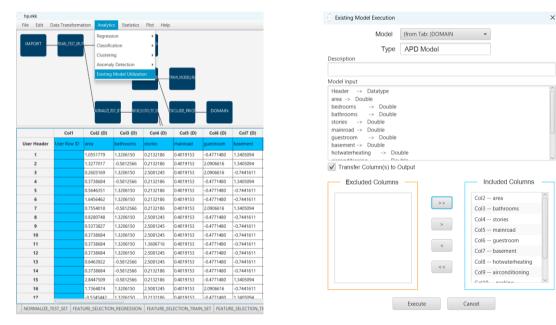
le Edit D	ata Transforma	tion Analytic	cs Statistics	Plot Help														
MPORT	TRAIN, TEST, SPUT	T NORMUZE TR	90.25		RAIN, MODEL, RC	DOMAIN	1											
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col9 (D)		Col1	Col2	Col3	Col4	Col5	Col6	Col7	2
er Header		price	area	bathrooms	stories	mainroad	guestroom	basement	hotwaterheati	User Header	User Row ID							
4		12215000.000	1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	1	-							_
2		9870000.0000	1.3277017	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	3								
3		9800000.0000	0.2603169	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.2438028	4								
4		9100000.0000	0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	5								_
5		000 8855000.0000 000	0.5546251	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	7								-
		000 8575000.0000 000	0.3040331							8								
6		000	1.6456462	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	9								_
7		8120000.0000 000	0.7554018	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	10								_
8		8080940.0000 000	0.8280748	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	11								-
9		7840000.0000	0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	13								
		7700000.0000	0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	14								
10		7560000.0000		1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	-0.2438028	15								_
10		000																

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\_PRICE\_TEST\_SET" tab by right-clicking on the input spreadsheet and then choosing "Import from SpreadSheet".

	hp.ekk																	-	0	×
Kart Austriku         Kart Kart Kart Kart Kart Kart Kart Kart	File Edit Da	ata Transforma	tion Analyti	cs Statistics	Plot Help															
InterfaceNormalNo	IMPORT	TRAIN_TEST_SPLI					DOMAIN													
Name         Name <th< th=""><th></th><th>Col1</th><th>Col2 (D)</th><th>Col3 (D)</th><th>Col4 (D)</th><th>Col5 (D)</th><th>Col6 (D)</th><th>Col7 (D)</th><th>Col8 (D)</th><th>Col</th><th></th><th>Col1</th><th>Col2</th><th>Col3</th><th>Col4</th><th>Col5</th><th>Col6</th><th>Col7</th><th>c</th><th>810</th></th<>		Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Col		Col1	Col2	Col3	Col4	Col5	Col6	Col7	c	810
1       1051779       1320510       0.421216       0.401913       0.401913       0.400913       0.400914       0.430028       1         2       132707       0.508156       0.21216       0.401913       2.000516       1.405094       0.243028       1         3       0.263160       1.326150       2.38126       0.401913       2.000516       1.405094       0.2438028       1.4         4       0.373864       0.32150       0.41218       0.401913       0.47140       1.40094       0.2438028       1.4         5       0.564515       1.320510       0.213216       0.401913       0.471410       0.418028       1.4         6       1.645642       1.326510       2.13216       0.401913       0.471410       0.4248028       1.4         7       0.755018       0.3273826       1.3205150       0.401153       0.471410       0.4248028       1.4         9       0.3373827       1.3205150       2.081246       0.401153       0.471410       0.4248028       1.4         10       0.373864       1.3205150       2.081246       0.401153       0.471410       0.4248028       1.4         12       0.373864       1.3205150       2.081246       0.471416	User Header	User Row ID	area	bathrooms	stories	mainroad	guestroom	basement	hotwaterheat	i aircc	User Header	User Row ID								
2         1327071         -0.81256         0.213216         0.401953         2.006616         1.405944         -0.248028         1.41	1		1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	g 1.41	1									
3         0.269160         1.326150         2.981240         0.401933         0.201660         0.741611         0.248028         1.4           4         0.373864         0.320560         0.213180         0.001933         0.47140         1.340920         0.248028         1.4         4         6         1         1         1         1         1         1         0.248028         1         6         1         1         1         2         2         1         1         1         1         2         2         1         1         2         2         1         1         2         2         1         1         2 <th2< th=""> <th1< th="">         1         <!--</td--><td>2</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>2</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></th1<></th2<>	2										2									
4       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.373804       0.320150       0.071400       0.340024       0.4         6       0.564535       0.320150       0.213286       0.401913       0.471400       0.241611       0.248028       1.4         7       0.7554018       0.320550       0.213286       0.401913       0.471400       0.248028       1.4         9       0.3273827       1.3205150       25081245       0.401913       0.471400       0.248028       1.4         9       0.373864       1.3205150       25081245       0.401913       0.471400       0.248028       1.4         10       0.373864       1.3205150       25081245       0.401913       0.471410       0.248028       1.4         12       0.373864       1.3205150       25081245       0.401913       0.471410       0.248028       1.4         13       0.645922       0.312186       0.401913       0.471410       0.248028       1.4         14       0.373864       1.3205150       25081246       0.471410       0.448028       1.4         14       0.373864       0.3											3									
S       0.6546351       1.320550       0.2132180       0.4019133       0.471440       0.244002       4.14         6       1.6646251       1.320150       0.2132180       0.4019133       0.471440       0.244002       1.41         7       0.754018       -3320550       2.312160       0.4019133       0.471401       0.2430228       1.41         8       0.820748       1.320510       2.081246       0.4019133       0.471401       0.243028       1.41         9       0.573827       1.320510       2.081246       0.01133       0.471400       0.741611       0.243028       1.41         10       0.373864       1.320510       2.081246       0.01133       0.471400       0.741611       0.243028       0.71         11       0.373864       1.320510       2.081246       0.401133       0.471400       0.741611       0.243028       1.41         12       0.373864       1.3205150       2.081246       0.401133       0.471400       0.741611       0.243028       1.41         13       0.6649922       0.931256       0.30133       0.471400       0.741611       0.243028       1.41         14       0.4649922       0.931256       0.31216       0.4019133	4		0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	-0.7	4									
6       1.454462       1.306150       0.2132180       0.409153       0.474400       0.248108       0.414         7       0.7594108       0.531256       0.2132160       0.409153       2.0407140       0.430940       0.248028       1.41         8       0.8280748       1.3206150       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         9       0.537827       1.3206150       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         10       0.373864       1.3206150       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         12       0.373864       1.3206150       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         13       0.645992       0.581256       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         14       0.3738644       0.320150       2.5081245       0.409153       0.477140       0.744161       0.248028       1.41         15       2.8447509       0.581256       2.5081246       0.409153       0.471410       0.438028       1.41         16       2.844	5		0.5646351	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	5									
7         0.759/018         0.832960         0.2132180         0.401913         2.080616         1.340594         0.240808         1.41           8         0.8229744         1.320550         2.081245         0.401913         0.471410         0.241611         0.248028         1.41           9         0.5373827         1.320550         2.081245         0.401913         0.471410         0.241611         0.248028         1.41           10         0.373864         1.320550         2.081245         0.401913         0.471410         0.248028         1.41           10         0.373864         1.320550         2.081245         0.401913         0.471410         0.248028         1.41           12         0.373864         1.320550         2.081245         0.471480         0.741611         0.248028         1.41           13         0.645922         0.581256         2.081246         0.471410         0.248028         1.41           14         0.373864         0.381256         0.471410         0.748082         1.41           14         0.381266         0.411913         0.471410         0.248028         0.41           15         2.847509         0.81256         0.281266         0.471410 <t< td=""><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>6</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></t<>	6										6									
a       0.820748       13206150       25081246       0.4019153       0.471440       0.741461       0.4240028       14         9       0.573827       13206150       25081246       0.4019153       0.471440       0.741611       0.248028       14         10       0.373804       13206150       25081246       0.4019153       0.471440       0.741611       0.248028       14         11       0.373804       13206150       25081246       0.4019153       0.471400       0.741611       0.248028       14         12       0.373804       13206150       25081246       0.4019153       0.471400       0.741611       0.248028       14         13       0.645922       0.581256       25081246       0.4019153       0.471400       0.741611       0.248028       141         14       0.378804       0.581256       0.281256       0.4019153       0.471400       0.248028       141         15       2.447599       0.581256       0.281266       0.471400       0.248028       141         16       2.448799       0.581256       0.281266       0.471400       0.248028       141         16       2.44799       0.581256       0.281266       0.471410       0.248	7		0.7554018	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	1.41	7									
9         0.537827         13206150         2508124         0.4019133         0.471400         0.241610         0.248028         1           10         0.373864         13206150         2508124         0.4019153         0.471400         0.241610         0.248028         1           11         0.373864         13206150         2508124         0.4019153         0.471400         0.241611         0.248028         1           12         0.373864         13206150         2508124         0.4019153         0.471400         0.248028         1           13         0.645922         -0.581266         25081246         0.4019153         -0.471400         0.248028         1           14         0.373864         -0.381266         0.213216         0.4019153         -0.471400         0.4248028         1           15         2.8447509         -0.391256         0.213216         0.4019153         0.471400         0.4248028         1           16         2.8447509         -0.391256         0.213216         0.4019153         0.471400         0.4248028         1           16         2.8447509         -0.391266         0.4019153         0.471400         0.4248028         1           16         2.84475	8		0.8280748	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	8									
10         0.373884         1320550         258874         0.4019153         0.471400         0.741601         0.428028         0.41           10         0.373864         1320550         150576         0.4019153         0.471400         0.741611         0.248028         0.41           12         0.373864         1320550         25081245         0.4019153         0.471400         0.741611         0.248028         14           13         0.645922         0.531256         25081245         0.4019153         0.741611         0.248028         14         12         13         0.645922         0.531266         0.213216         0.4019153         0.471401         0.748082         14         14         0.41161         0.428028         14           14         0.373864         0.531256         0.213216         0.4019153         0.471401         0.248028         14         14         0.41161         0.428028         14           15         2.8447509         0.581256         0.213216         0.4019153         0.471410         0.248028         14         15         16         0.411         1.428028         14           16         1.300150         0.581256         0.4019153         0.471410         0.438028	9		0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	9									
11       0.373864       1.3206150       1.360716       0.4019153       0.471480       0.741611       0.248008       1.41         12       0.373864       1.3206150       2.508126       0.4019153       0.471480       0.741611       0.2480028       1.41         14       0.373864       0.381256       2.508126       0.4019153       0.471400       0.741611       0.2480028       1.41         15       2.6447509       0.508126       0.4019153       0.4071400       0.741611       0.2480028       0.41         16       0.373864       0.381256       0.213216       0.4019153       0.4071400       0.2480028       0.41         16       0.736874       0.3208150       2.081746       0.4019153       0.4071400       0.2480028       0.41         17       0.548474       1.200150       0.21316       0.4019153       0.4071400       0.2480028       0.47         16       0.736874       1.20116       0.013166       0.474161       0.2480028       0.47         17       0.548474       1.20116       0.01316       0.477480       0.474480       0.2480028       0.47         16       0.5081746       0.01316       0.471491       0.474480       0.474808       0	10		0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	-0.7(	10									
12       0.3738084       1.3205105       2.5081245       0.4019153       0.471480       0.744161       0.2438028       1.41         13       0.6569922       0.581256       2.5081245       0.4019153       0.471480       0.744161       0.2438028       1.41         14       0.373864       -0.581256       0.213218       0.4019153       0.471480       0.744161       0.2438028       1.41         15       2.8447509       -0.581256       0.213218       0.4019153       0.471480       0.744161       0.2438028       1.41         16       1.754874       1.3205150       2.681245       0.4019153       2.005616       0.744161       0.2438028       1.41         16       1.754874       1.3205150       2.681246       0.4019153       2.005616       0.744161       0.2438028       1.41         16       1.754874       1.3205150       2.581246       0.4019153       2.005616       0.744161       0.2438028       1.41         17       0.5544424       1.320156       0.4019153       0.471480       1.400944       0.4338028       1.41         17       0.5544424       1.320156       0.012116       0.471480       0.474480       1.41       0.414       0.414       0.414	11		0.3738684	1.3206150	1.3606716	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41									1	
13         0.6469922         -0.891566         2.081740         0.0471915         -0.747140         -0.741801         -0.741802         1.4           14         0.373864         -0.931556         0.213216         0.4071953         -0.477140         -0.741611         -0.2438028         1.41           15         2.847509         -0.931566         0.213216         0.4071953         -0.477140         0.741611         -0.2438028         0.47           16         1.3954674         1.305050         2.581246         0.4071953         -0.0741611         -0.2438028         0.47           16         1.3954674         1.305050         2.581246         0.005165         -0.7441611         -0.2438028         0.47           17         0.5554424         1.305050         0.307186         0.4771400         1.407044         -0.2438028         1.47           17         0.5554244         1.305050         0.307186         0.4771401         1.407044         0.4238028         1.47	12		0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41									-	
14         0.3738084         -0.581256         0.2132180         0.4019153         -0.471400         -0.248028         1.41           15         2.847509         -0.581256         0.2132180         0.4019153         -0.471400         1.3405094         -0.248028         1.41           16         1.7364874         1.3206150         25081245         0.4019153         -0.471400         1.3405094         -0.248028         1.41           17         0.5544744         1.3206150         25081245         0.4019153         -0.4714101         -0.248028         1.41           17         0.5544744         1.300150         0.019153         -0.471401         1.340704         -0.248028         1.41           17         0.5544744         1.300150         0.19153         0.471491         1.407044         0.3483028         1.41	13		0.6463922	-0.5812566	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41										
15         2.8447599         -0.812656         0.2132166         0.401913         -0.47140         1.240394         -0.243028         0.71           16         1.7564874         1.320556         2.5081245         0.4019133         2.906565         0.7441611         -0.243028         0.71           17         0.5554454         1.301546         0.4019135         0.4714461         -0.243028         1.41           17         0.5554454         1.301546         0.101316         0.41074345         0.4744611         -0.243028         1.41	14		0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41										
16 1.7368/4 1.2305/50 2308745 0.409793 229006/6 4.0.44161 0.2489028 1.41 17 0.655452 123065/6 0.31328.6 A.041915 3.027906 1.41,01 0.2489028 1.41 17 0.655452 1.23065/6 0.31328.6 A.041915 3.027906 1.41,01 0.2489028 1.41	15		2.8447509	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	-0.70									-	
17 -0.5345442 1.3206150 0.2132186 0.4019153 -0.4771480 1.3405094 -0.2438028 1.1.41	16		1.7364874	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.2438028	1.41									-	
	17		-0 5345442	1 3206150	0 2132186	0.4019153	-0 4771480	1 3405094	-0.2438028	141~	17	20								

Check the predictions' reliability by browsing: "Analytics"  $\rightarrow$  "Existing Model Utilization". Then select as Model "(from Tab:) DOMAIN".



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

IMPORT	TRAIN_TEST_SPLF		IN ST BARE	2000/1904.55	RAIN_MODELLAD	DOMAIN												
	Col1	Col2 (D)	Col3 (D)	Col4 (D)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (D)	Cot		Col1	Col2 (D)	Col3 (D)	Col4 (S)	Col5 (D)	Col6 (D)	Col7 (D)	Col8 (
lser Header	User Row ID	area	bathrooms	stories	mainroad	questroom	basement	hotwaterheati	-	User Header	User Row ID	Domain	APD	Prediction	area	bathrooms	stories	mainroad
	Oper Now 10					-		ng	9	1		2.6593096	3.9052670	reliable	1.0551779	1.3206150	0.2132186	0.40191
1		1.0551779	1.3206150	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	1.41	2		1.5827052	3.9052670	reliable	1.3277017	-0.5812566	0.2132186	0.40191
2		1.3277017	-0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	1.41	3		2.7287725	3.9052670	reliable	0.2603169	1.3206150	2.5081245	0.401915
3		0.2603169	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.2438028	1.41	4		1.4748561	3.9052670	reliable	0.3738684	-0.5812566	0.2132186	0.401915
4		0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	-0.7(	5		0.0590468	3.9052670	reliable	0.5646351	1.3206150	0.2132186	0.40191
5		0.5646351	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	6		1.6613661	3.9052670	reliable	1.6456462	1.3206150	0.2132186	0.40191
6		1.6456462	1.3206150	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	7		0.9333940	3.9052670	reliable	0.7554018	-0.5812566	0.2132186	0.40191
7		0.7554018	·0.5812566	0.2132186	0.4019153	2.0906616	1.3405094	-0.2438028	1.41	8		0.6813095	3.9052670	reliable	0.8280748	1,3206150	2.5081245	0.40191
8		0.8280748	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	9		1.1477765	3.9052670	reliable	0.5373827	1.3206150	2.5081245	0.40191
9		0.5373827	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	10		1.1866370	3.9052670	reliable	0.3738684	1.3206150	2.5081245	0.40191
10		0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	-0.7(	11		1.1474529	3.9052670	reliable	0.3738684	1.3206150	1.3606716	0.40191
11		0.3738684	1.3206150	1.3605716	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	12		0E-7	3.9052670	reliable	0.3738684	1.3206150	2.5081245	0.40191
12		0.3738684	1.3206150	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	13		2.5206654	3.9052670	reliable	0.6463922	-0.5812566	2.5081245	0.40191
13		0.6463922	-0.5812566	2.5081245	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	14		1.1486412	3.9052670	reliable	0.3738684	-0.5812566	0.2132186	0.40191
14		0.3738684	-0.5812566	0.2132186	0.4019153	-0.4771480	-0.7441611	-0.2438028	1.41	15		2.0736358	3.9052670	reliable	2.8447509	-0.5812566	0.2132186	0.40191
		2.8447509	+0.5812566	0.2132186	0.4019153	-0.4771480	1.3405094	-0.2438028	-0.7(	16		2.7449893	3.9052670	reliable	1.7364874	1.3206150	2.5081245	0.40191
15 16		1.7364874	1.3206150	2.5081245	0.4019153	2.0906616	-0.7441611	-0.2438028	1.41	10							10.000	

# Final Isalos Workflow

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

