Kyubit Self-Service BI

www.kyubit.com

Contents

1. Kyubit Self-Service BI Concepts

1.1. Self-Service Data Models

Kyubit BI software provides features and tools to quickly create data models (**Analytic Models**) in the Kyubit application from the data stored in **Excel/CSV files** or **SQL query results**, by the regular end-user without the involvement of BI professionals or third-party analytic software modules or special databases (OLAP). Such Self-Service **Analytic Models** can be used to create data analysis, reports, KPIs and dashboards by the same regular end-user with drag-and-drop and other user-friendly actions in the Kyubit application. In other words, Kyubit BI application provides all that it takes for a regular end-user to build data models for the analysis and data visualization. Data models (Analytic Models) consist of the Measures, Dimensions, Hierarchies and Details that are used while analyzing the prepared model in the analysis grid/chart view or preparing insights on the dashboard. While preparing data to create new Analytic Model, the source data needs to be organized in rows and columns. Later in the process, the user defines the Analytic Model structure by defining which column contains values for the model structures like **measures**, **dimensions** or **details** to complete the final Analytic Model processing.

1.2. Self-Service Data Analysis & Visualization

Once Self-Service data model (Analytic Models) is created in the Kyubit BI application, the same user, as well as other authorized users can quickly create data analysis, charts and reports based on the same 'Analytic Model' with a simple and comprehensive approach (**drag-and-drop**) that does not require special skills or training. Furthermore, created analysis/report could be used while creating dashboard charts, tables and KPIs, to **visualize prepared data insights by the regular end-user**. Kyubit BI includes features to quickly design the dashboard layout by drag-and-drop various charts, connecting with previously prepared queries or analyses, positioning and resizing dashboard elements, setting display options for the individual dashboard element and setting overall dashboard style.

2. Kyubit Self-Service BI Overview

Kyubit 'Analytic Model' is a Self-Service BI analytic feature that could be quickly utilized using your data from **Excel/CSV files** and **SQL query results**, without creating OLAP cubes. With analytic models, end-user can create pivot tables, analytic reports and dashboards, using measures, dimensions, slicers and many features similar to OLAP analysis.

In many situations, you have a set of data you wish to analyze, but you probably will not engage creating OLAP cubes, which almost always requires knowledge, time, tools, etc. With Kyubit Self-Service BI, end-user can quickly import and configure **Analytic Models**, which behaves almost like you have OLAP cubes ready for analysis. Set of values from Excel/CSV files or query results is transformed to analytic models and Self-Service BI tools are ready for all Kyubit users to use them in analysis and dashboards, while values from the same data sources could be scheduled to update regularly with new values based on our preference.
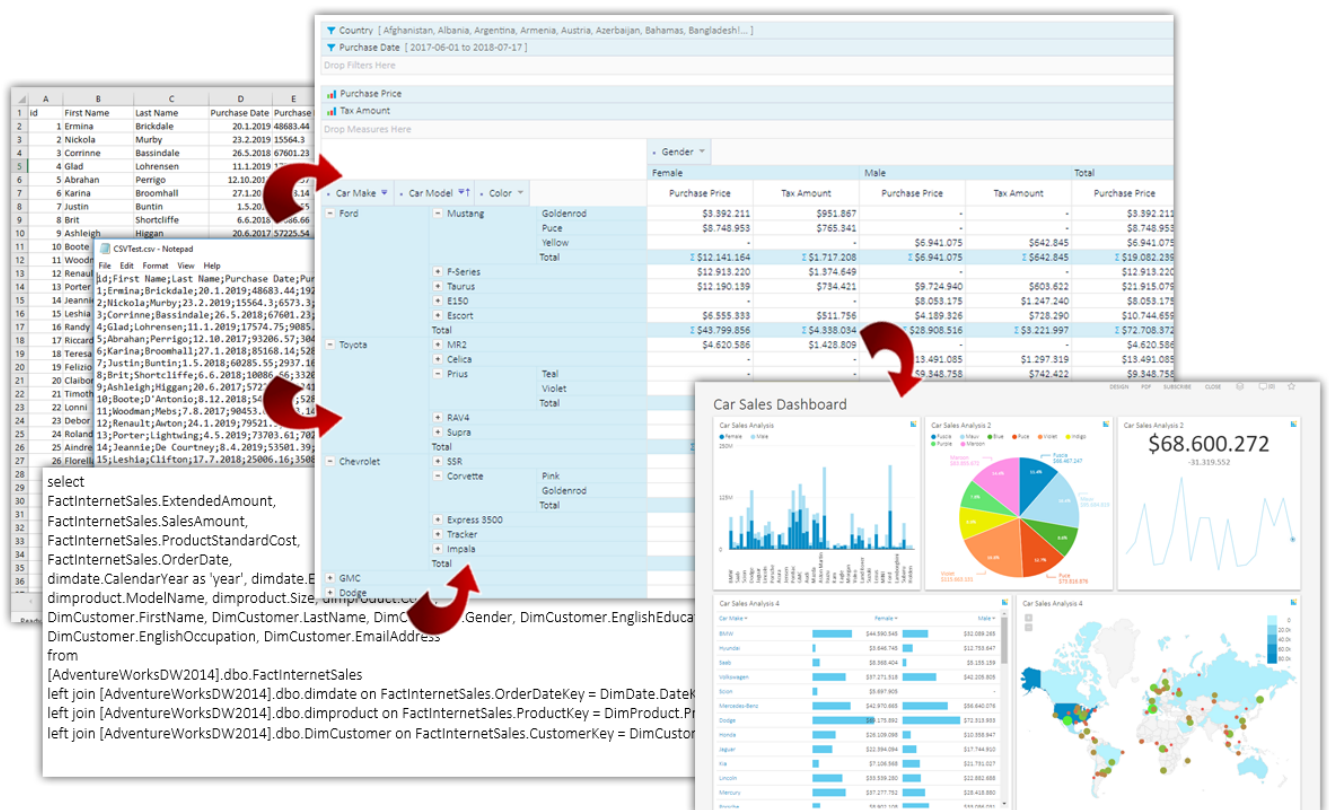
2.1.   End-user experience

After Kyubit Self-Service BI 'Analytic Model' is processed, authorized end-users can start a new analysis, that will look the same as if they are analyzing OLAP cube structures (very similar). End-users can create analysis, reports and create dashboards based on created analysis the same way they are doing with OLAP based analyses. Most features, like **drill-down**, **drill-through**, **expanding**, **slicing**, **ordering**, **isolating** are included in Analytic model analysis.

2.2.   How it works

After you import your data from Excel/CSV files or SQL query Results and process 'Analytic Model', Kyubit creates special structures in Kyubit internal **"KyubitAnalyticModels" SQL database**, that are suitable for quick analytic SQL queries. While analyzing data Kyubit is creating SQL queries to bring analytic results from Kyubit Analytic Models database. In other words, Kyubit is using SQL technology, combined with **ColumnStore indexes** and some **smart caching** to bring data analysis. Only technology prerequisite is MS SQL Server, which is prerequisite for the whole product anyway.

2.3.   Pros

o   The main reason to use 'Analytic Model' is for a regular user to quickly add a set of data for analysis, dashboard usage, scheduled subscriptions and sharing with other users.

o   Excel/CSV data format should be friendly to all users while preparing data to be used

o   Great usage of Date filters (if data contains date values) that are much friendlier to be used than OLAP 'date' structures. Quickly select absolute or relative date filter values in the analysis, report or dashboard filters.

2.4.   Limitations

There are limitations to Kyubit Self-Service BI 'Analytic Model' usage, that should be known before using new Kyubit technology. Kyubit Analytic Model is not created in mind to replace more serious analytic engines, like OLAP technology, but to bring a simple solution for smaller data sets (below 10 million of rows) that should be analyzed quickly with very little knowledge of data analysis and structures.

▪   'Analytic Model' will perform great with hundreds of thousands of rows of data, while we would not recommend being used with more than 10 million of rows of data. This question greatly depends on the hardware on which SQL server is running, but more millions of rows of data should be used with in-any-case more robust and scalable OLAP technology.

▪   There are no limitations to the number of category members (rows) in grid analysis and reports, while analytic grid and report can contain a maximum of 128 series (columns) of values in analysis for each measure in the analysis.

▪   On category axis there could be multiple category levels expanding (drill-down) to explore data in more details, while series members cannot be expanded.

3. Step-by-step Self-Service BI

3.1. Create Analytic Model from Excel/CSV file

Data stored in Excel or CSV file delimited by the **semicolon** (;), can be quickly uploaded to Kyubit application and immediately is ready for the step of 'Data Definition'.

Numeric and **date values** in Excel/CSV file should respect current Kyubit BI server **regional settings** format for decimal separator and date format.

- Prepare Excel/CSV file based on the columns and rows that contain data for analysis.

Create a new Analytic model, add some name for the model, save it, and then upload an Excel/CSV file. Once uploaded, the file is automatically saved to this model.
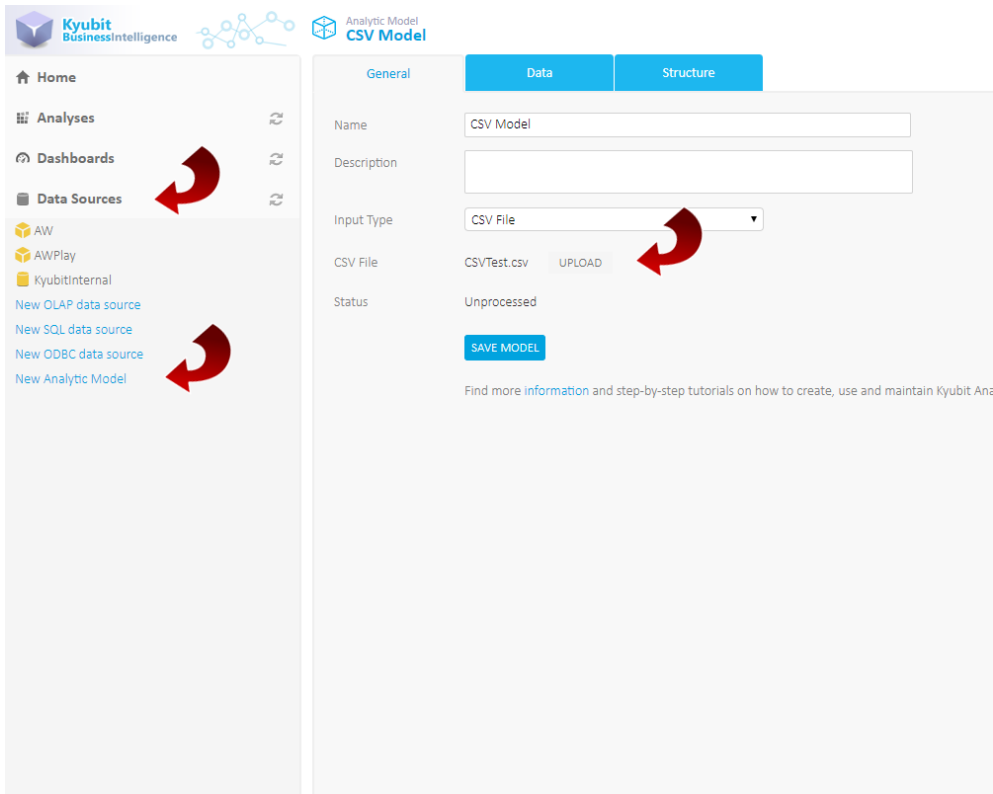
- After upload, select 'Data' tab to inspect the sample data from the Excel/CSV file. Creating analytic model is ready for the next step of 'Data Definition' explained below.

Measure ▼	Dimension ▼	Dimension ▼	Date ▼	Measure ▼	Measure ▼	Dimension ▼	Dimension ▼	Dimension ▼	D
id	First Name	Last Name	Purchase Date	Purchase Price	Tax Amount	Gender	Car Make	Car Model	
[Format Values]				[Format Values]	[Format Values]				
1	Ermina	Brickdale	20.1.2019	48683.44	1927.06	Female	Mazda	Tribute	
2	Nickola	Murby	23.2.2019	15564.3	6573.3	Male	Hyundai	Elantra	
3	Corrinne	Bassindale	26.5.2018	67601.23	5412.56	Female	Dodge	Charger	
4	Glad	Lohrensen	11.1.2019	17574.75	9085.54	Female	Lexus	ES	
5	Abrahan	Perrigo	12.10.2017	93206.57	3044.03	Male	Plymouth	Laser	
6	Karina	Broomhall	27.1.2018	85168.14	5286.2	Female	Lincoln	Mark VIII	
7	Justin	Buntin	1.5.2018	60285.55	2937.16	Male	Volkswagen	Scirocco	
8	Brit	Shortcliffe	6.6.2018	10086.66	3320.71	Male	Saab	9-7X	
9	Ashleigh	Higgan	20.6.2017	57225.54	1341.04	Female	Volkswagen	Fox	
10	Boote	D'Antonio	8.12.2018	54923.96	5288.26	Male	Ford	Thunderbird	
11	Woodman	Mebs	7.8.2017	90453.63	7883.14	Male	GMC	Yukon	
12	Renault	Awton	24.1.2019	79521.5	5063.29	Male	Suzuki	Grand Vitara	
13	Porter	Lightwing	4.5.2019	73703.61	7020.95	Male	Toyota	Sienna	
14	Jeannie	De Courtney	8.4.2019	53501.39	4835.19	Female	GMC	Yukon XL 1500	
15	Leshia	Clifton	17.7.2018	25006.16	3508.76	Female	Mazda	B-Series	
16	Randy	Halpeine	2.10.2018	33424.96	5442.23	Male	Nissan	Quest	
17	Riccardo	Djurisic	2.7.2018	57812.23	6477.99	Male	Mitsubishi	Mighty Max Macro	
18	Teresa	Redparth	22.3.2018	35617.85	9723.67	Female	Buick	LaCrosse	
19	Felizio	Silkston	26.3.2018	41550.7	2767.68	Male	Chevrolet	Tracker	
20	Claiborn	Finey	10.12.2018	99441.13	6519.52	Male	Subaru	Forester	
21	Timothea	Aloigi	18.12.2018	22584.67	6512.69	Female	Eagle	Talon	
22	Lonni	Poate	28.9.2018	27369.14	7138.44	Female	Bentley	Continental GT	
23	Debor	Roz	14.3.2019	17516.05	8179.03	Female	Saturn	VUE	
24	Roland	Bradshaw	18.6.2018	18095.04	9918.43	Male	Mazda	B-Series Plus	

3.2.   Create Analytic Model from the Query results

If the data for the analytic model is based on existing data from the relational databases, create the **SQL query** that will be used to retrieve data for the new model. Any valid SQL query could be used to run against the **SQL server** or **ODBC data sources** registered in the Kyubit application.

- Under 'Data Sources' select New Analytic Model, change input type to 'Query', select the data source for query and set query text. This query will be used to get data for this analytic model.

- When the query is defined, click on the 'Run Query' button or 'Data' tab to get the query sample results for the inspection. At the same time, the analytic model creation is ready for the 'Data Definition' step explained below.

Analytic Model
Adventure Works Analytic Model

General	Data	Structure

Measure ▼	Measure ▼	Measure ▼	Date ▼	Dimension ▼	Dimension ▼	Dimension ▼	Dimension ▼	Dimension ▼
ExtendedAmount	SalesAmount	ProductStandardCost	OrderDate	year	month	day	ModelName	Size
[Format Values]	[Format Values]	[Format Values]						
3578,2700	3578,2700	2171,2942	29.12.2010 0:00:00	2010	December	29	Road-150	6:
3399,9900	3399,9900	1912,1544	29.12.2010 0:00:00	2010	December	29	Mountain-100	4
3399,9900	3399,9900	1912,1544	29.12.2010 0:00:00	2010	December	29	Mountain-100	4
699,0982	699,0982	413,1463	29.12.2010 0:00:00	2010	December	29	Road-650	6:
3399,9900	3399,9900	1912,1544	29.12.2010 0:00:00	2010	December	29	Mountain-100	4
3578,2700	3578,2700	2171,2942	30.12.2010 0:00:00	2010	December	30	Road-150	4
3578,2700	3578,2700	2171,2942	30.12.2010 0:00:00	2010	December	30	Road-150	6:
3374,9900	3374,9900	1898,0944	30.12.2010 0:00:00	2010	December	30	Mountain-100	4
3399,9900	3399,9900	1912,1544	30.12.2010 0:00:00	2010	December	30	Mountain-100	3
3578,2700	3578,2700	2171,2942	31.12.2010 0:00:00	2010	December	31	Road-150	4
3578,2700	3578,2700	2171,2942	31.12.2010 0:00:00	2010	December	31	Road-150	4
699,0982	699,0982	413,1463	31.12.2010 0:00:00	2010	December	31	Road-650	5:
3578,2700	3578,2700	2171,2942	31.12.2010 0:00:00	2010	December	31	Road-150	5:
3578,2700	3578,2700	2171,2942	31.12.2010 0:00:00	2010	December	31	Road-150	5
3578,2700	3578,2700	2171,2942	1.1.2011	2011	January	1	Road-150	5

3.3. Data Definition

When the source for the new analytic model is defined (CSV file or query), the next step is to define the **role** of **each column** in the provided data set on the 'Data' tab. Possible definition for the columns are 'Measure', 'Dimension', 'Date' and 'Details'.

Measure

The measure is a value from our data set that will be used for the aggregations while analyzing the analytic model. Typically, using Sum, Avg, Count and other aggregations. The measure has to be of the **numeric values** in the Kyubit Analytic Model.

Dimension

The dimension values are used to slice and analyze measure values, used in the required context. Analytic results are based on measures and dimension members on the categories and series.

Date

The Date column could be used as a filter in the analysis. Because analyzing business data mostly include filtering based on a certain period, this is an extremely useful element for the analysis. The Date cannot be used on analysis categories and series.

Details

Some information from data set are not good analytic material to be used on analysis categories and series, but should only be available when asked for details of aggregated data. For example, information such as 'address' and 'phone number', we are hardly going to use as aggregating data, but still, we like to use as 'Details' on the aggregated results. Setting such columns as the 'Details' ensures a more compact analytic model, faster processing and smaller size on the disk.

Example data definitions...

Example usage of defined structures later in analysis.

Column Caption and Description

While defining data for the analytic model, optionally click on the column name and set its caption (if should be different from the source) and column Description.

Format Values

Measure values should be provided as a pure numeric value. To configure measure to be presented as a formatted numeric value (Currency for example), click on the "Format values" below column name and choose one of the formatting options or write your own.

Key Sorting Column

Assign another column that contains values (numbers) that will be used for sorting of the dimension level. For example, if the dimension level contains months (January, February, March,…), sorting these values alphabetically does not make sense. Here comes the 'Key column' that contains values that will be used for sorting. For example, if we have column 'MonthNumbers' that contains values (1,2,3, etc.) corresponding to the order of the month, we will apply the 'MonthNumber' column as a 'Key Column' for the dimension level 'Months'.

When the Analytic Model is processed, in the analysis click on the 'Level Sorting' > Member Key to apply sorting with values from the 'Sorting Key Column'.

Ranges

If the values for a particular dimension are of the numeric type, it would be convenient to organize them in ranges. Typically, we would like to analyze data based not on the particular number, but the range of values that present certain scope we can give a new for. To organize numeric values of the particular dimension into ranges, click on the dimension column and the **Range** button. Defined any number of ranges by providing the range name, min, and max values. After the data processing, the dimension will be presented with the defined ranges as its members.

3.4 Organize Structures Additionally

When each column definition is defined, we can immediately proceed to the processing of the Analytic Model, but on the third tab 'Structure' there are additional options to **organize analytic model** structures. Instead to leave all dimension levels in 'Default dimension', we can create new dimensions and organize levels appropriately to be **more comprehensive** for end-users.

New Dimension

Create a new dimension based on some topic (for example, 'Customer' or 'Product') and assign appropriate analytic levels to this dimension. This way you group analytic levels to certain topic and makes analysis more comprehensive.

New Hierarchy

When analytic levels are related to each other in parent-child relation, it is practical to organize them inside 'Hierarchies'. For example, Year-Month-Day or Continent-Country-City. This way makes analysis easy to drill-down data from a higher view to more detailed values for end-user.

Default structure.
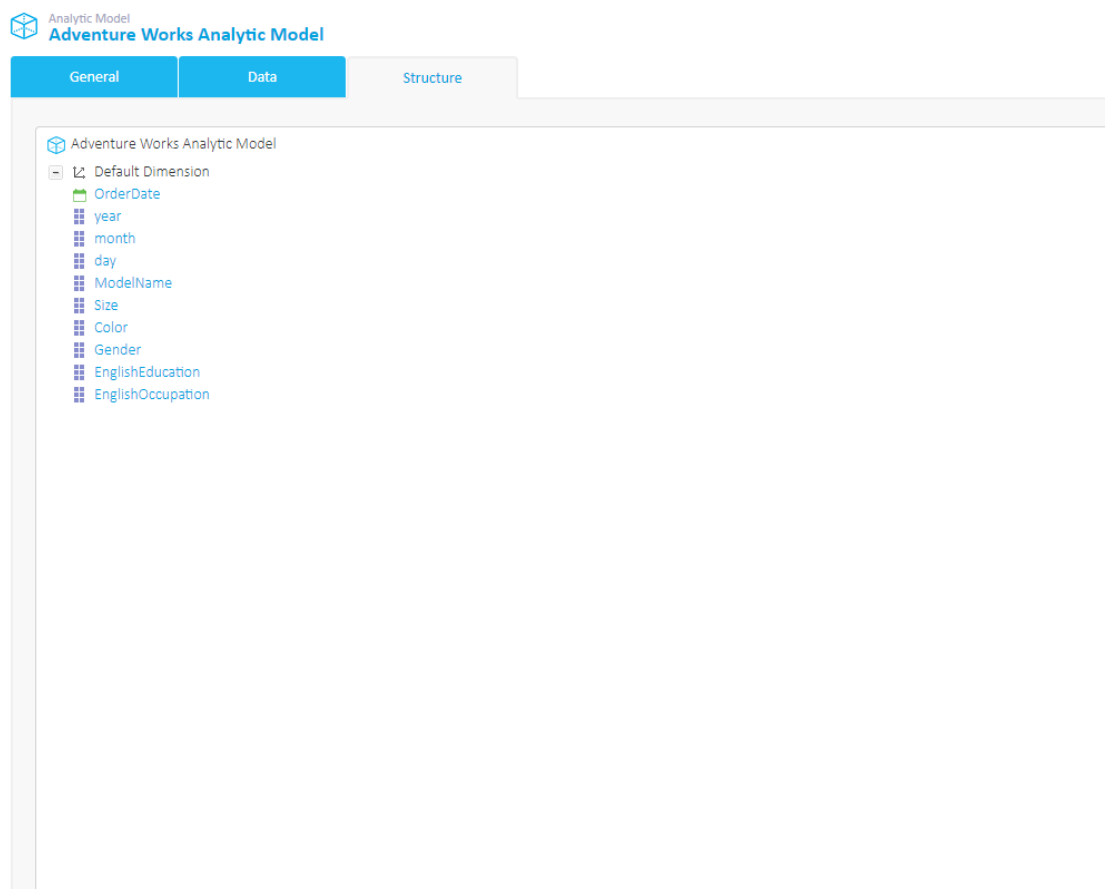
Organized structure.

3.5. Processing of Analytic Model
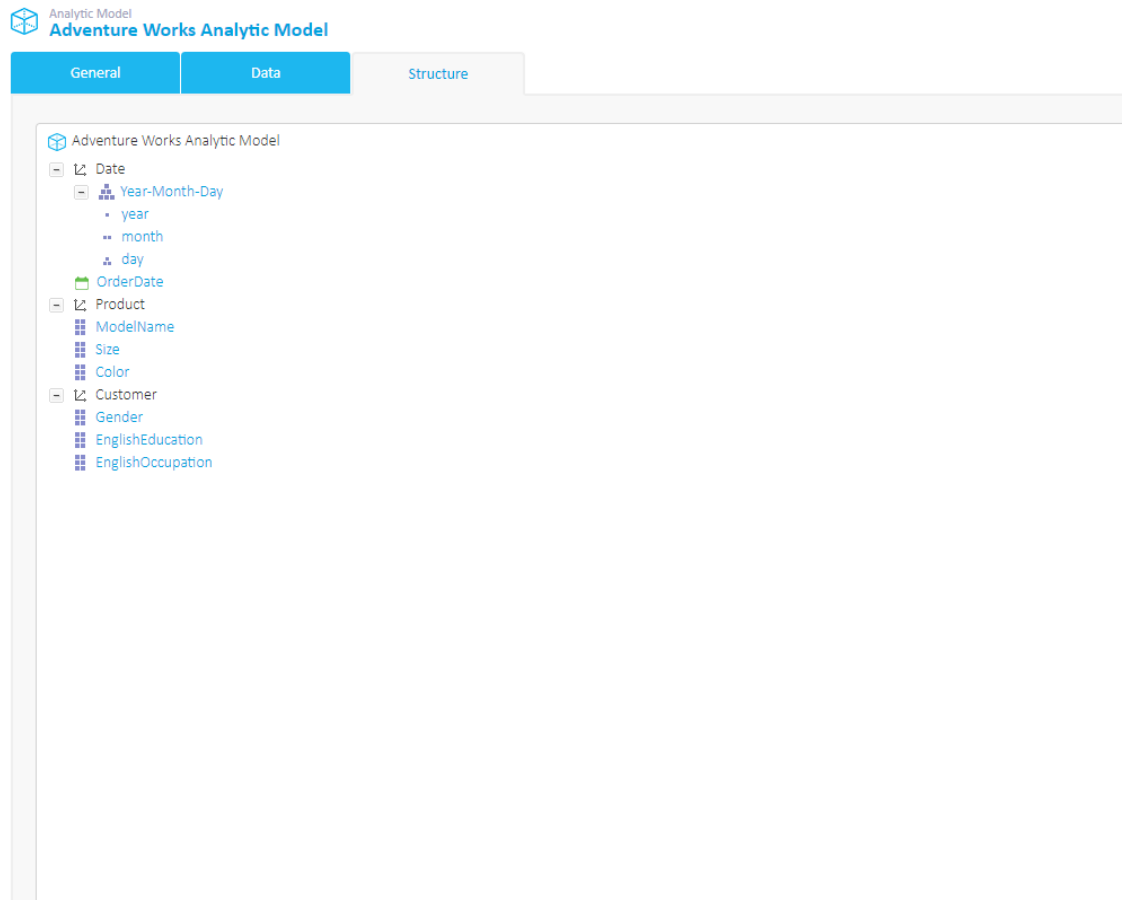
While preparing the analytic model, the user can save and open same analytic model many times, which is in 'Unprocessed' status. When all data and structures are prepared, click the **Process Analytic Model** button on the 'Structure' tab to actually start processing of analytic model data and make it ready for analysis and visualizations. The process could take from **few seconds to several minutes** (or more) depending on a number of rows and columns defined for the analytic model. All columns not required to be 'Dimension' set to 'Details' structure type, which will speed processing and save space in the 'models' database.

Analytic Model Status

There are 3 analytic model status, **Unprocessed**, **Processing** and **Processed**. While the analytic model is processing, it cannot be used by any user. If processing fails for any reason, it reverts to 'Unprocessed' status.

Log

After processing of the analytic model, **details of processing** could be inspected by clicking on the 'Log' button in 'General' tab. If the processing of the analytic model has failed, this is a good place to start troubleshooting for the possible cause of the problem.
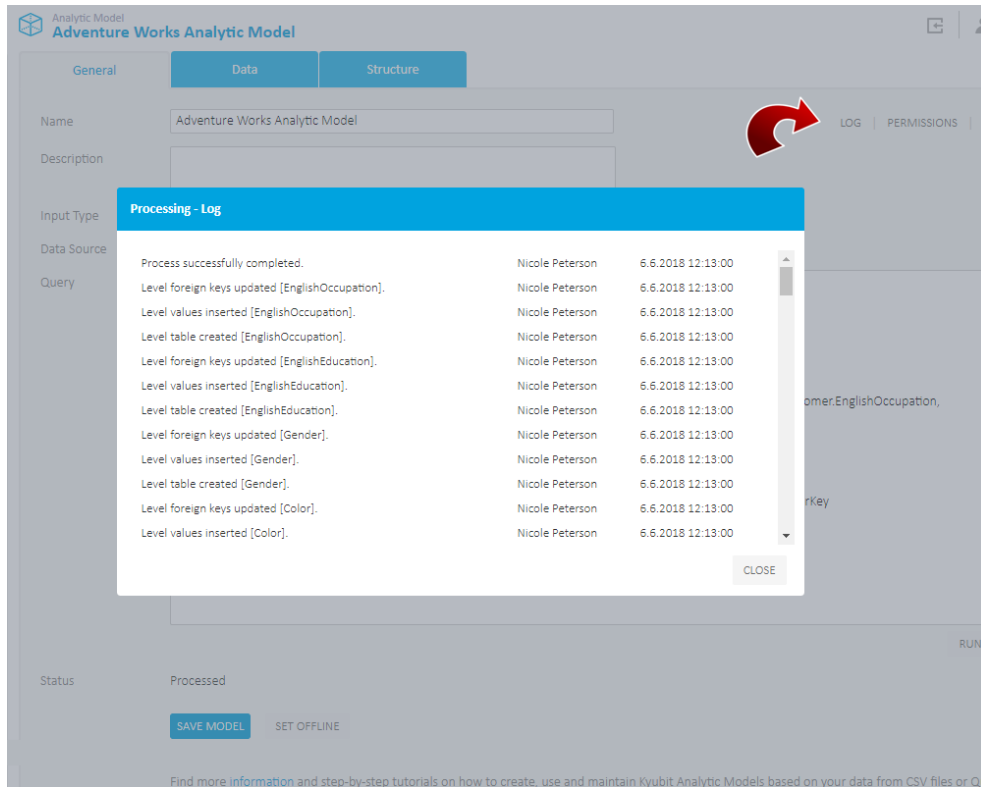
Schedule model updates

The **Analytic model** could be **updated automatically** with 'Scheduled Jobs'. Go to Schedule -> Jobs and create new Job 'Process Analytic Model' and time preference for updates to occur. If Analytic model is based on the query, the same query will be run against the defined data source to bring fresh data into Analytic model. If Analytic model is based on the Excel/CSV file, schedule job can be created only if Excel/CSV file is uploaded from shared folder and path begin with "\\..." (For example, \\SomeMachine\FolderWithData).
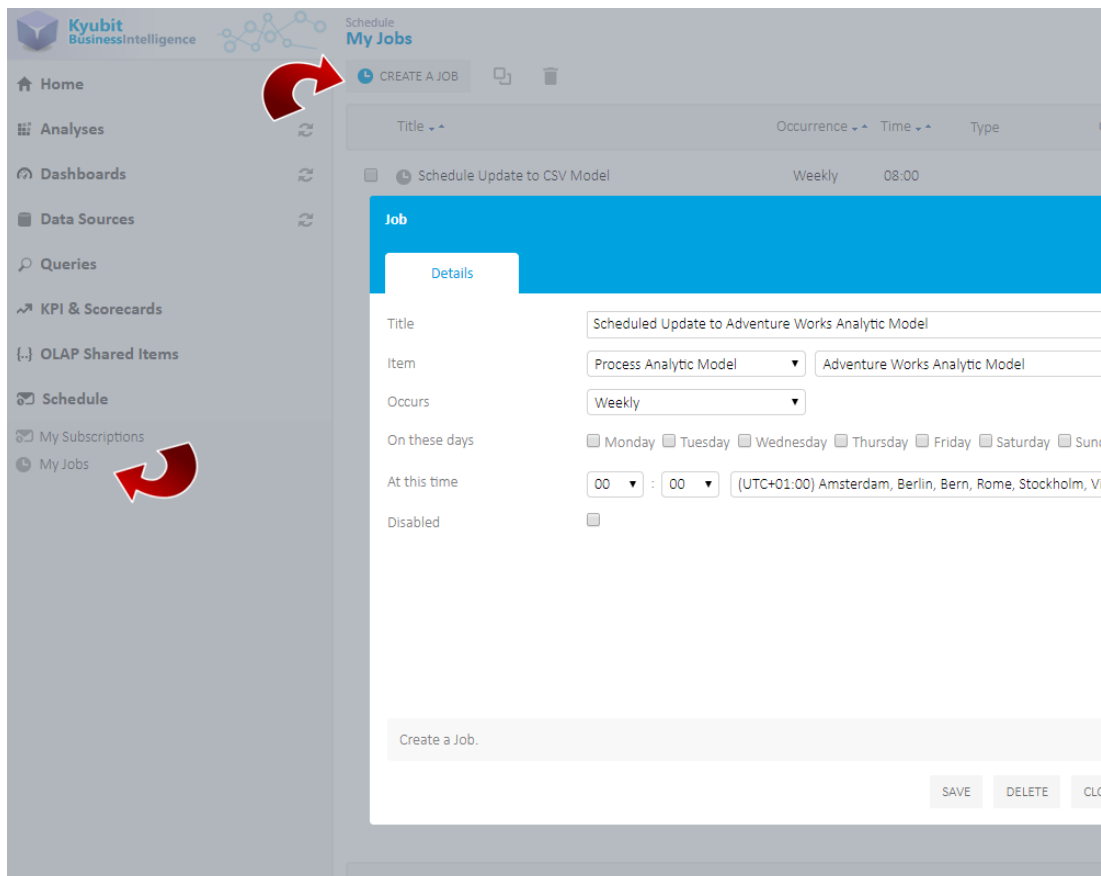
Processing of Analytic model could be started manually or with 'Scheduled Jobs'.
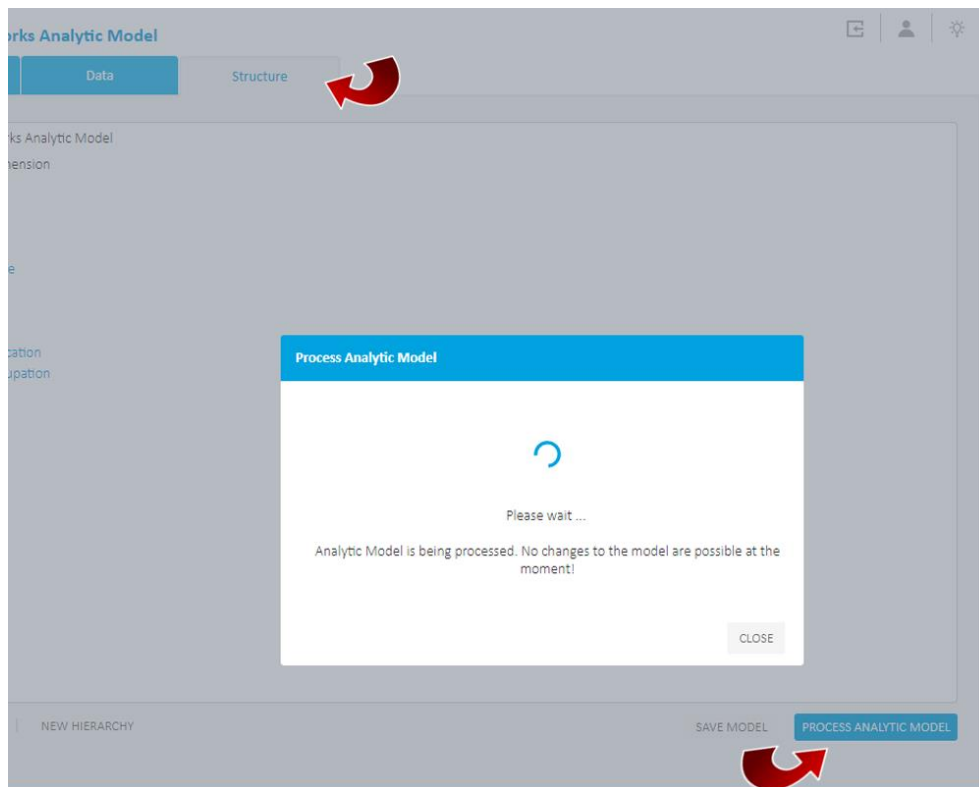
Find useful / troubleshooting information in the processing 'Log' form.

Schedule Analytic model update from the defined model data sources.

3.6. Usage in Analysis and Dashboards

After the Analytic model is processed, it is ready for the **analysis** and **visualizations**. Create the analysis and reports in the grid/chart view by adding measures and dimensions on the categories/series axis or slicer axis. Use ordering, aggregate, isolating and other analytic actions to prepare the analytic report of your interest. Once the analysis is prepared it could be added to the **Dashboard** and be visualized by many charts and visualization widgets, while at the same time data could be further analyzed by the end-user working with the dashboard (Drill-down, drill-through, drill-by, slicing and other useful analytic actions).

Analysis with Analytic models.

Reports prepared with the Analytic models.

Car Sales Analysis

▼ Filters:
(1) Purchase Date: 2017-06-01 to 2018-07-17

|			Gender					
			Female		Male		Total	
▪ Car Make	▪ Car Model	▪ Color	Purchase Price	Tax Amount	Purchase Price	Tax Amount	Purchase Price	Tax Amount
Ford	Mustang	Goldenrod	$3.392.211	$951.867	-	-	$3.392.211	$951.867
		Puce	$8.748.953	$765.341	-	-	$8.748.953	$765.341
		Yellow	-	-	$6.941.075	$642.845	$6.941.075	$642.845
		Total	Σ $12.141.164	Σ $1.717.208	Σ $6.941.075	Σ $642.845	Σ $19.082.239	Σ $2.360.053
	E-Series		-	-	$19.273.783	$1.627.772	$19.273.783	$1.627.772
	F-Series		$12.913.220	$1.374.649	-	-	$12.913.220	$1.374.649
	Taurus		$12.190.139	$734.421	$9.724.940	$603.622	$21.915.079	$1.338.043
	E150		-	-	$8.053.175	$1.247.240	$8.053.175	$1.247.240
	Total		Σ $37.244.523	Σ $3.826.278	Σ $43.992.973	Σ $4.121.479	Σ $81.237.496	Σ $7.947.757
Chevrolet	Corvette	Puce	-	-	$1.311.782	$945.136	$1.311.782	$945.136
		Pink	-	-	$7.472.839	$759.567	$7.472.839	$759.567
		Goldenrod	$4.791.645	$551.251	-	-	$4.791.645	$551.251
		Purple	$9.657.927	$377.213	-	-	$9.657.927	$377.213
		Total	Σ $14.449.572	Σ $928.464	Σ $8.784.621	Σ $1.704.703	Σ $23.234.193	Σ $2.633.167
	S10		-	-	$12.438.309	$1.851.364	$12.438.309	$1.851.364
	Suburban 2500		-	-	$10.084.200	$1.678.699	$10.084.200	$1.678.699
	SSR		$493.538	$637.013	$8.011.041	$901.958	$8.504.579	$1.538.971
	Blazer		$7.941.808	$805.417	$400.034	$716.637	$8.341.842	$1.522.054
	Total		Σ $22.884.918	Σ $2.370.894	Σ $39.718.205	Σ $6.883.361	Σ $62.603.123	Σ $9.224.255
Toyota	MR2		$7.765.757	$1.684.817	-	-	$7.765.757	$1.684.817
	Camry Solara		-	-	$12.478.011	$1.426.067	$12.478.011	$1.426.067
	RAV4		$207.608	$181.779	$9.191.991	$1.120.120	$9.399.599	$1.301.899
	Celica		-	-	$13.491.085	$1.297.319	$13.491.085	$1.297.319
	Prius	Teal	-	-	$9.348.758	$742.422	$9.348.758	$742.422
		Violet	$9.522.962	$540.952	-	-	$9.522.962	$540.952
		Total	Σ $9.522.962	Σ $540.952	Σ $9.348.758	Σ $742.422	Σ $18.871.720	Σ $1.283.374
	Total		Σ $17.496.327	Σ $2.407.548	Σ $44.509.845	Σ $4.585.928	Σ $62.006.172	Σ $6.993.476
GMC			$95.670.271	$8.691.333	$60.939.884	$6.419.696	$156.610.155	$15.111.029
Dodge			$69.175.892	$7.924.466	$72.313.933	$6.065.491	$141.489.825	$13.989.957
Mercedes-Benz			$42.970.665	$6.422.972	$56.640.076	$6.248.937	$99.610.741	$12.671.909
Audi			$42.950.274	$6.055.726	$55.508.426	$5.898.201	$98.458.700	$11.953.927
Mitsubishi			$52.698.890	$5.372.810	$43.073.410	$4.619.740	$95.772.300	$9.992.550
Mazda			$70.642.937	$5.835.948	$12.420.874	$3.461.070	$83.063.811	$9.297.018
Pontiac			$89.173.942	$5.318.738	$51.571.712	$3.560.301	$140.745.654	$8.879.039
Total			Σ $676.365.123	Σ $67.400.799	Σ $667.131.162	Σ $77.663.010	Σ $1.343.496.285	Σ $145.063.809

Gender
■ Female ■ Male

PURCHASE PRICE
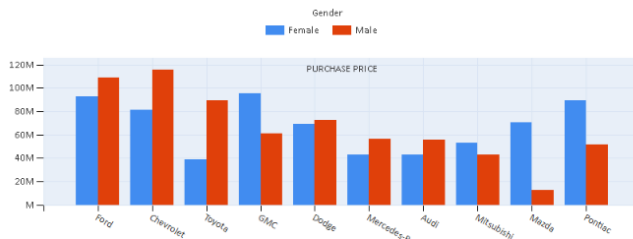
Dashboards based on the Analytic models.
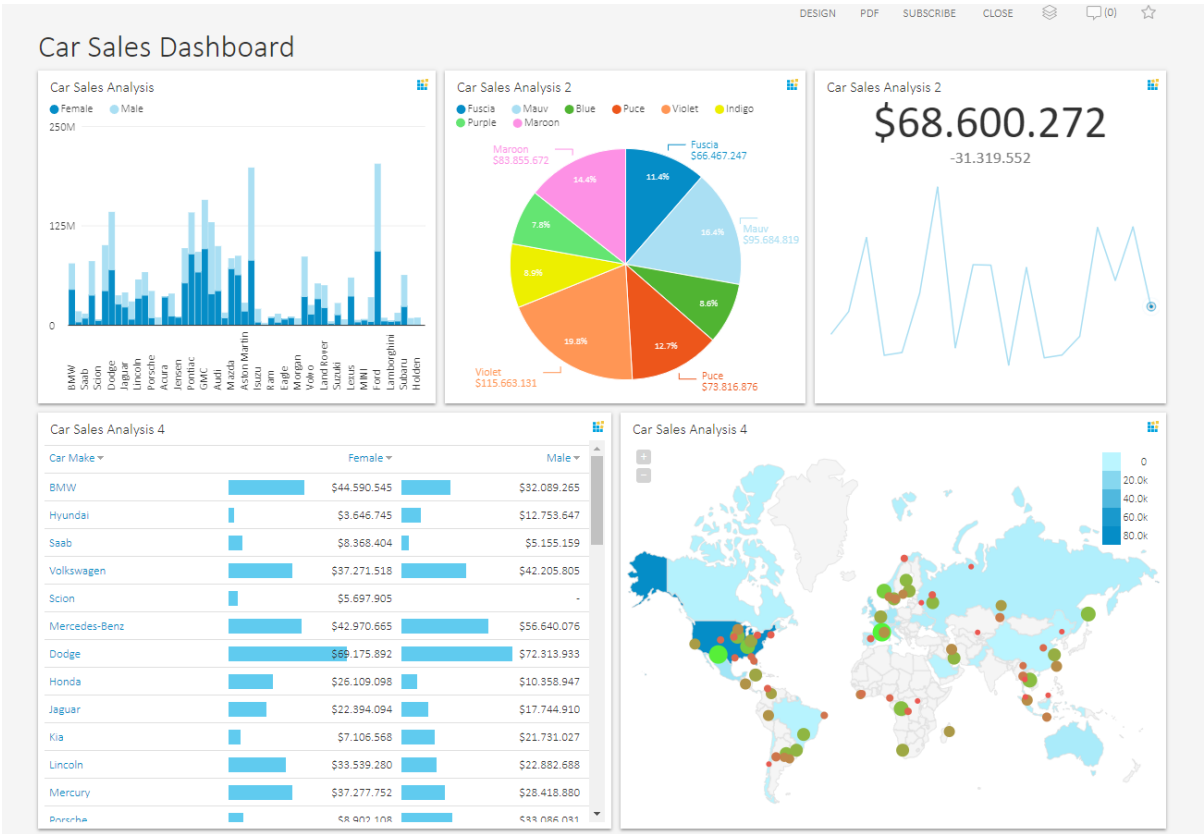
3.7. Date filtering

Advantage of Analytic Model over OLAP is clearly date filtering, which offers date picker dialog with options to set fixed or relative date periods that would be used to slice data in analysis, report or dashboard usage.
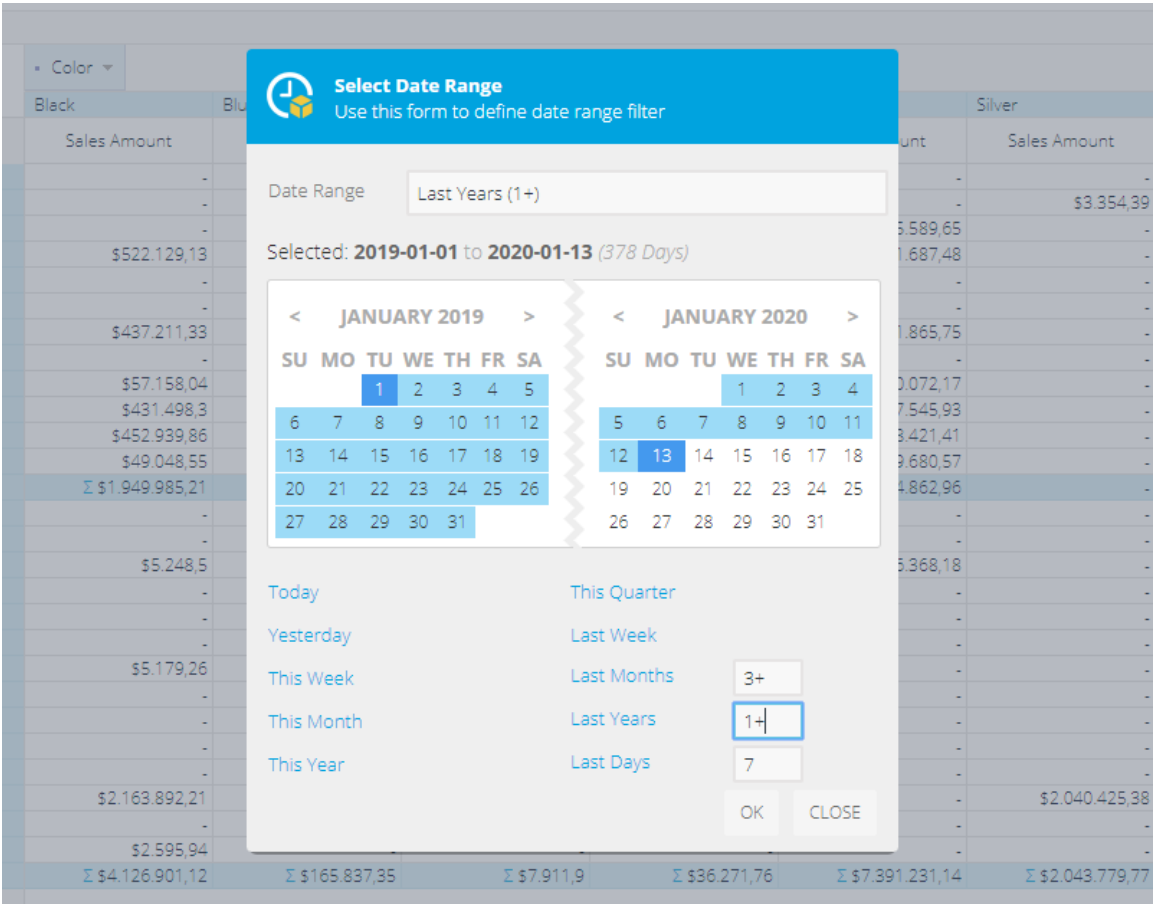
To set relative date period, enter a number, which would return Last x months, years or days to the current date. For example, for the last 3 years without the current year, set 'Last Year' to '3'. To also include the current year, set 'Last Year' to '3+'.

3.8.   'Calculated Measures' in Analytic Model

The prepared analytic model contains model measures that could be used by end-user to create 'Calculated Measures' based on the required expression and formatting options.

To create new 'Calculated Measure', in grid analysis view click the 'Options' and select 'Create User Calculated Measure'. A form to edit 'Calculated Measure' expression will be displayed. Any expression that contains analytic model expressions could be used to create a new 'Calculated Measure'.
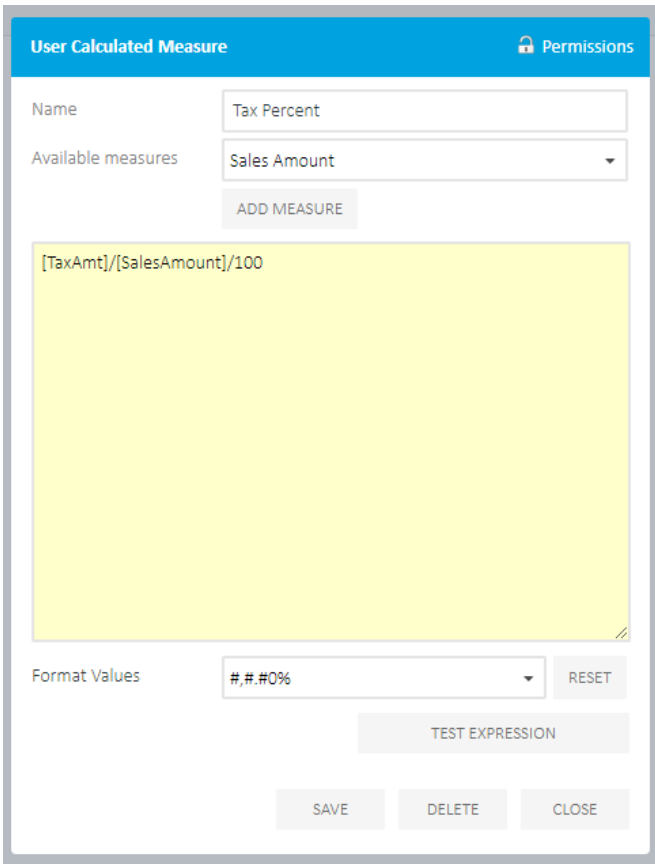
'Calculated Measures' expression for the values that cannot be null (division by zero issue) have to be included in the **NULLIF([Measure],0)** statement. This will ensure that expression will never be evaluated resulting in 'division by zero' exception.

While constructing the calculated measure, a user could use arithmetic operators and mathematical functions to prepare the final expression.

+ (Add)   - (Subtract)   * (Multiply)

/ (Divide)   % (Modulo)

ABS	DEGREES	RAND
ACOS	EXP	ROUND
ASIN	FLOOR	SIGN
ATAN	LOG	SIN

'Test Expression' feature could warn that used expression is not valid. After saving the calculated measure, it is immediately visible along with the model measures in the Analytic Models structure tree.

To edit existing 'Calculated Measure' right-click on the measure in grid analysis or open in 'OLAP Share Items' section. Each created 'Calculated Measure' is not visible to other users, unless appropriate permissions are set for other users and groups to see or manage the same calculated measure.

Kyubit
BusinessIntelligence

- **Home**
- **Analyses** ⟳
- **Dashboards** ⟳
- **Data Sources** ⟳
- **Queries**
- **KPI & Scorecards**
- **{..} OLAP Shared Items**
- {..} User Named Sets
- User Drillthrough Columns
- User Calculated Measures ⬅

Calculated measures defined in Kyubit BI
User Calculated Measure

🗑

Name	Data Source
☐ 🔢 Mix	Adventure Works Analytic Model
☐ 🔢 Tax Percent	Adventure Works Analytic Model