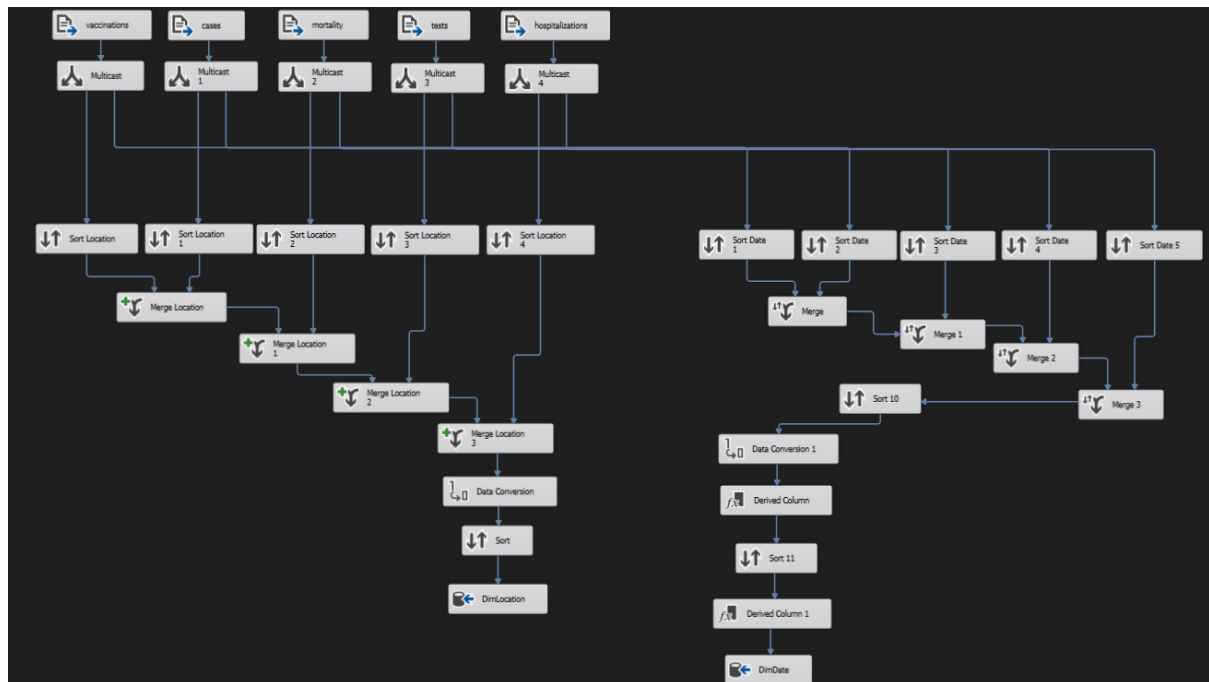


Dimension Tables Creating

Aim of the process

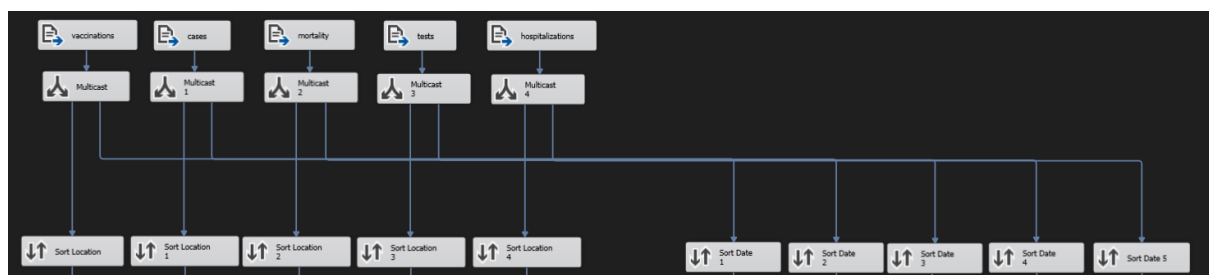
The goal is to create Dimension Tables (DimLocation and DimDate). The whole data flow at this stage is shown below:



Defining data sources, multicasting and sorting

Data are obtained from files preprocessed by python script, then split and sorted by location or date.

During sorting only the column needed for the respective dimension tables were retained:



Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns		
<input checked="" type="checkbox"/> Name	Pass Thr...	
<input type="checkbox"/> iso_code	<input checked="" type="checkbox"/>	
<input type="checkbox"/> continent	<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/> location	<input checked="" type="checkbox"/>	
<input type="checkbox"/> date	<input type="checkbox"/>	
<input type="checkbox"/> total_deaths	<input type="checkbox"/>	
<input type="checkbox"/> new_deaths	<input type="checkbox"/>	
<input type="checkbox"/> total_deat...	<input type="checkbox"/>	
<input type="checkbox"/> new_death...	<input type="checkbox"/>	

Input Column	Output Alias	Sort Type	Sort Order	Con
location	location	ascending	1	

☒ Remove rows with duplicate sort values

OK

Cancel

Help

Sort Transformation Editor

Specify the columns to sort, and set their sort type and their sort order. All nonselected columns are copied unchanged.

Available Input Columns		
<input checked="" type="checkbox"/> Name	Pass Thr...	
<input type="checkbox"/> country	<input type="checkbox"/>	
<input checked="" type="checkbox"/> date	<input checked="" type="checkbox"/>	
<input type="checkbox"/> iso_code	<input type="checkbox"/>	
<input type="checkbox"/> people_fully_vaccinated_per_hundred	<input type="checkbox"/>	
<input type="checkbox"/> people_vaccinated_per_hundred	<input type="checkbox"/>	
<input type="checkbox"/> total_vaccinations_per_hundred	<input type="checkbox"/>	
<input type="checkbox"/> sputnikv_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> sinovac_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> sinopharm_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> pfizer_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> astrazeneca_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> novavax_total_vaccinations	<input type="checkbox"/>	
<input type="checkbox"/> moderna_total_vaccinations	<input type="checkbox"/>	

Input Column	Output Alias	Sort Type	Sort Order	Con
date	date	ascending	1	

☒ Remove rows with duplicate sort values

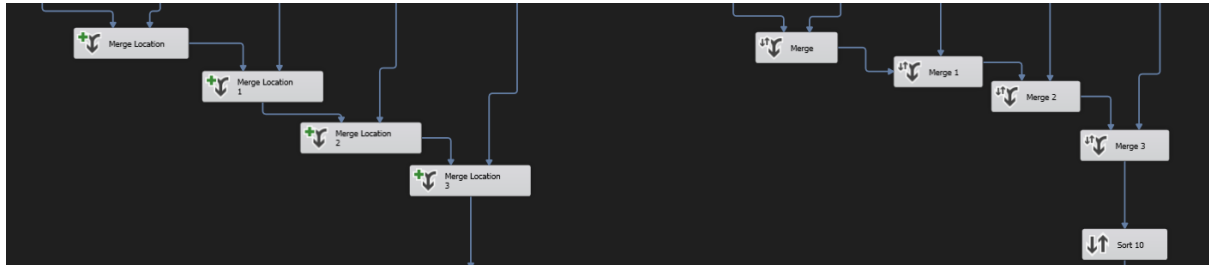
OK

Cancel

Help

Merge, Merge Join

Data are merged using Merge Join with 'Full outer join' join type for the location data, and using Merge for the date data:



Merge Join Transformation Editor

Configure the properties used to join two sources of sorted data. Select the join type and then specify the columns to be used as the join key. Join keys must be used in the order specified by the sort-key position of the column.

Join type: Full outer join Swap Inputs

Sort Location

<input type="checkbox"/>	Name	Ord...	Join...
<input type="checkbox"/>	country	1	<input checked="" type="checkbox"/>
<input type="checkbox"/>	iso_code	0	<input type="checkbox"/>

Sort Location 1

<input checked="" type="checkbox"/>	Name	Ord...	Join...
<input checked="" type="checkbox"/>	iso_code	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	continent	0	<input type="checkbox"/>
<input checked="" type="checkbox"/>	location	1	<input checked="" type="checkbox"/>

Input	Input Column	Output Alias
Sort Location 1	location	location
Sort Location 1	continent	continent
Sort Location 1	iso_code	iso_code

OK Cancel Help

Configure the properties used to merge two sorted inputs into one output by creating mappings between columns.

Output Column Name	Merge Input 1	Merge Input 2
date (Sort key: 1)	date (Sort key: 1)	date (Sort key: 1)

OK

Cancel

Help

Data Conversion, Derived Column, Final Destination

Data are converted to the appropriate data types, sorted and then exported into appropriate database tables (created previously using SQL scripts) with the surrogate keys. For date data, new columns containing day, month, quarter, and month name are also created:



1 Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input C...

- ☒ Name
- ☒ location
- ☒ continent
- ☒ iso_code

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
location	Copy of location	string [DT_STR]	50			1252 (ANSI - Latin I)
continent	Copy of continent	string [DT_STR]	30			1252 (ANSI - Latin I)
iso_code	Copy of iso_code	string [DT_STR]	10			1252 (ANSI - Latin I)

< >

Configure Error Output... OK Cancel Help

Data Conversion Transformation Editor

Configure the properties used to convert the data type of an input column to a different data type. Depending on the data type to which the column is converted, set the length, precision, scale, and code page of the column.

Available Input...

☒ Name
☒ date

Input Column	Output Alias	Data Type	Length	Precision	Scale	Code Page
date	date_converted	database date [DT_DB...				

Configure Error Output...

OK

Cancel

Help

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters
Columns

Mathematical Functions

String Functions

Date/Time Functions

NULL Functions

Type Casts

Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length
year	<add as new column>	YEAR(date_converted)	four-byte signed inte...	
month	<add as new column>	MONTH(date_converted)	four-byte signed inte...	
day	<add as new column>	DAY(date_converted)	four-byte signed inte...	
quarter	<add as new column>	DATEPART("qq",date_converted)	four-byte signed inte...	

Configure Error Output...

OK

Cancel

Help

Derived Column Transformation Editor

Specify the expressions used to create new column values, and indicate whether the values update existing columns or populate new columns.

Variables and Parameters

Columns

Mathematical Functions

String Functions

Date/Time Functions

NULL Functions

Type Casts

Operators

Description:

Derived Column Name	Derived Column	Expression	Data Type	Length
month_name	<add as new column>	month == 1 ? "January" : month == 2 ? "Febru...	Unicode string [DT_...	9

Configure Error Output...

OK

Cancel

Help