BellaDati

BellaDati Training - Data sets

Agenda

- **Data set**
  - basic usage, creating data set
  - managing indicators and attributes
  - ETL - import, data sources, transformation
  - data browsing
  - using mathematical formulas
  - appearance and translation

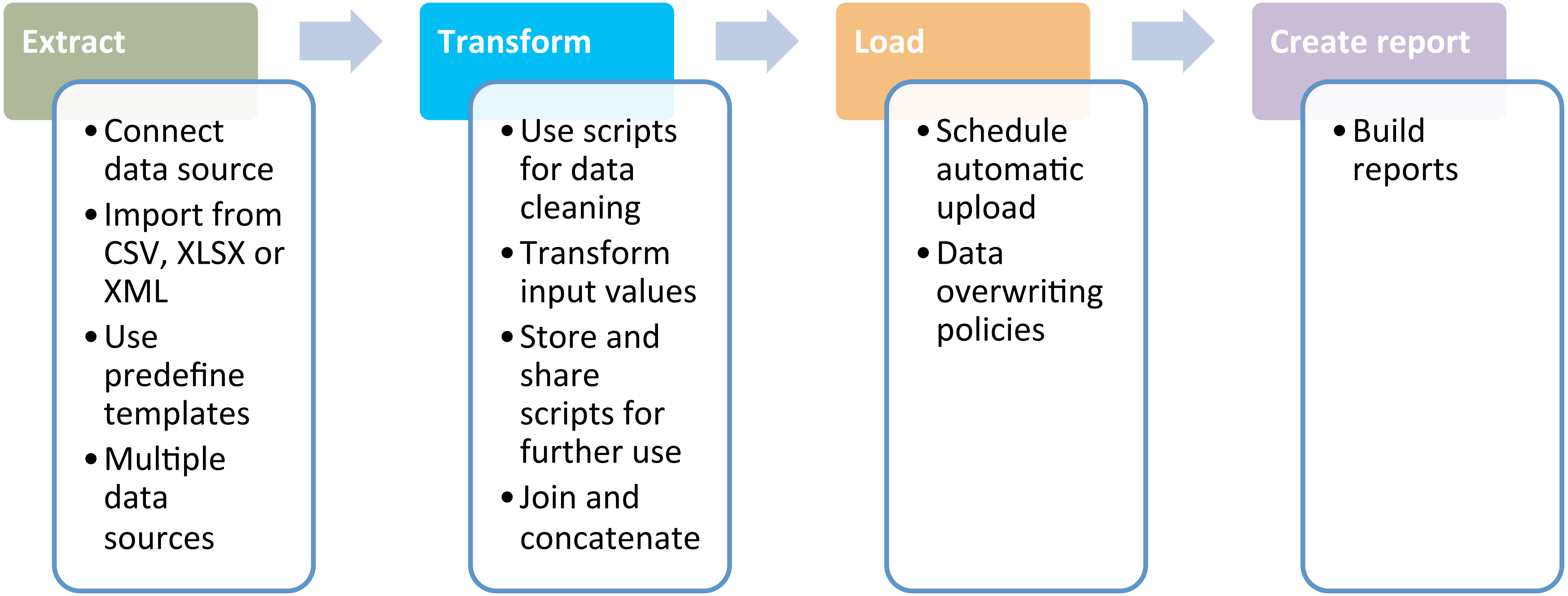
Content

- **Data set**
  - joining data sets
  - watching data changes
  - backup & restore
- **GEO mappers**
  - defining shapes and points

BellaDati

Data sets

Data set workflow



Extract

- Connect data source
- Import from CSV, XLSX or XML
- Use predefine templates
- Multiple data sources

Transform

- Use scripts for data cleaning
- Transform input values
- Store and share scripts for further use
- Join and concatenate

Load

- Schedule automatic upload
- Data overwriting policies

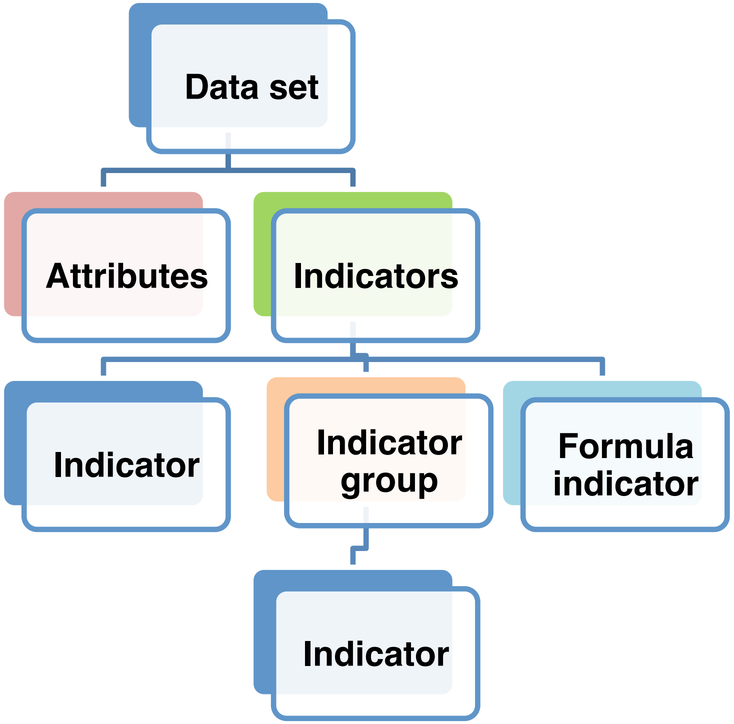
Create report

- Build reports

Data set functions

- Define the data model
- Extract-Transform-Load
  - connect to data sources
  - define import settings
  - transform and schedule
- Browse, manage and export data
- Watch data changes
- Join and concatenate datasets
- Share and setup data level permissions
- Backup & restore

Data set definition



- **Is virtual database which holds all data**
- **Consists of**
  - Attributes
  - Indicators
- **Each report is build on single data set, but you can:**
  - join data sets together
  - cross reference data set
  - concatenate data sets (custom join)

BellaDati

Data set – Attribute

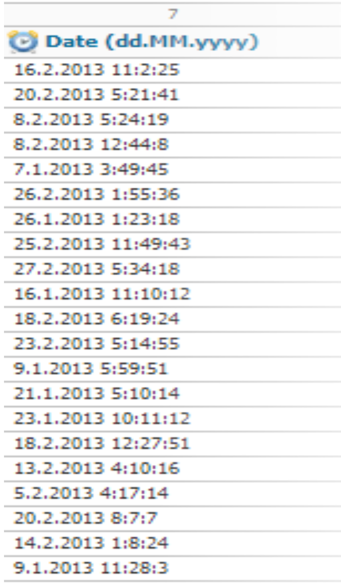
- **Attributes are describing indicators**
- **Provide categories for organizing items**
- **Usually in a form of general terms such as:**
  - country, department, product, employee, hire date
- **Attributes contains members (values):**
  - Korea, Sales Department, Product A, John Smith, 2011-01-01
- **Attribute types:**

- Text
- Date
- Time
- Datetime
- Geo Point
- Geo JSON
- Long text

Date/Time, Datetime attribute

- Used in reports to show date/time axis in charts/tables
- Not mandatory, multiple date/time attributes allowed eg. CreatedDateTime, CompletedDateTime etc.
- Current date/time can be appended during import
- Various date time formats, from years to seconds
- Can be used in scripts to derive values

  e.g. daysBetween(),

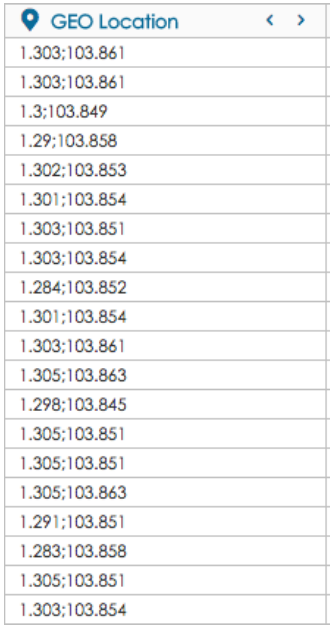


7
Date (dd.MM.yyyy)
16.2.2013 11:2:25
20.2.2013 5:21:41
8.2.2013 5:24:19
8.2.2013 12:44:8
7.1.2013 3:49:45
26.2.2013 1:55:36
26.1.2013 1:23:18
25.2.2013 11:49:43
27.2.2013 5:34:18
16.1.2013 11:10:12
18.2.2013 6:19:24
23.2.2013 5:14:55
9.1.2013 5:59:51
21.1.2013 5:10:14
23.1.2013 10:11:12
18.2.2013 12:27:51
13.2.2013 4:10:16
5.2.2013 4:17:14
20.2.2013 8:7:7
14.2.2013 1:8:24
9.1.2013 11:28:3

GEO point attribute

- Used in reports to display values in GEO maps
- Format must be LAT;LONG (g.e. 1.303;103.861)

- Create/Import place definition (point)
  - Shapes: GeoJSON with associated drill-down members
    - g.e. {"properties":{"names":[”REG1"],"name":”Region1"},"type":"Feature","geometry":{"type":"Polygon","coordinates”:[[[9.921906,54.983104],[9.93958,54.596642], [9.921906,54.983104]]]}}
  - Points: Latitude & Longitude with associated place
    - g.e. Chicago,"-87.636368","41.866212","United States"
  - Places definition needs to be associated with the data set



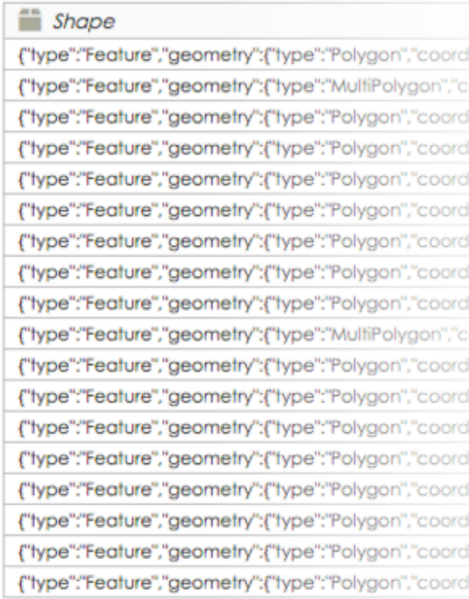
GEO Location ‹ ›
1.303;103.861
1.303;103.861
1.3;103.849
1.29;103.858
1.302;103.853
1.301;103.854
1.303;103.851
1.303;103.854
1.284;103.852
1.301;103.854
1.303;103.861
1.305;103.863
1.298;103.845
1.305;103.851
1.305;103.851
1.305;103.863
1.291;103.851
1.283;103.858
1.305;103.851
1.303;103.854

BellaDati

GEO JSON attribute

- Used in reports to display shapes in GEO maps
- Format must be in valid Geo JSON format (geojson.org)

- Create/Import place definition (point or shapes)
  - Shapes: GeoJSON with associated drill-down members
    - g.e. {"properties":{"names":["REG1"],"name":"Region1"},"type":"Feature","geometry":{"type":"Polygon","coordinates":[[[9.921906,54.983104],[9.93958,54.596642], [9.921906,54.983104]]]}}
  - Places definition needs to be associated with the data set



Shape
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"MultiPolygon","c
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"MultiPolygon","c
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord
{"type":"Feature","geometry":{"type":"Polygon","coord

Long text attribute

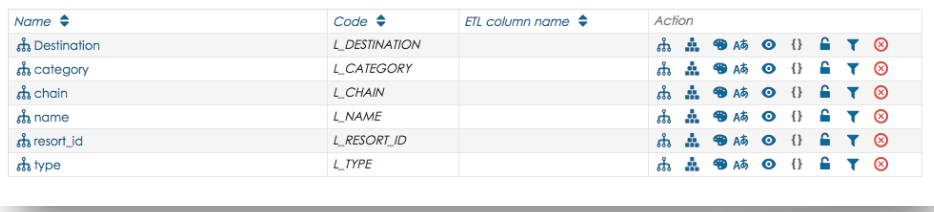
- **Values with length longer than 220 characters**
- **Cannot be used in aggregations and visualisations**
- **Suitable for following use cases:**
  - View source data with values longer than 220 characters
  - Display values longer than 220 characters in KPI labels and tables (firstValue() or lastValue() functions need to be used)

Data set – indicator

- Everything you can measure = quantitative data
- Represents **raw not aggregated numerical value** (fact):
  - employee wage, transaction cost, temperature, …
- This value is described by **attributes**, e.g.:
  - wage of employee living in Seoul, Female, 34 years old, …
- Data set indicators are available in **all reports**
  - but also in cross reference function
- Can be defined by mathematical formula (count, unique, etc.)

Working with attributes

- **Add, modify, remove**
  - all changes are propagated to the reports, where they are applied!
- **Each attribute has got unique code, starting with L_**
  - code is used in formulas and scripts
- **Appearance and localization support**
- **Permissions and data level access settings**
- **Hide system attributes – not displayed in reports**
  - i.e. primary or foreign keys used in joining
- **Transform existing values**
- **Predefined drill-down paths**
- **Subsets**

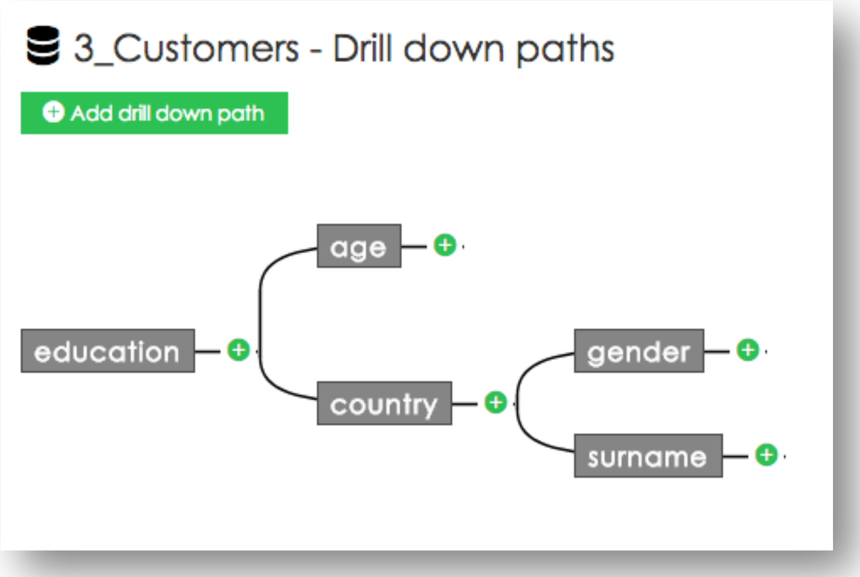


Name	Code	ETL column name	Action
Destination	L_DESTINATION		
category	L_CATEGORY		
chain	L_CHAIN		
name	L_NAME		
resort_id	L_RESORT_ID		
type	L_TYPE		

BellaDati

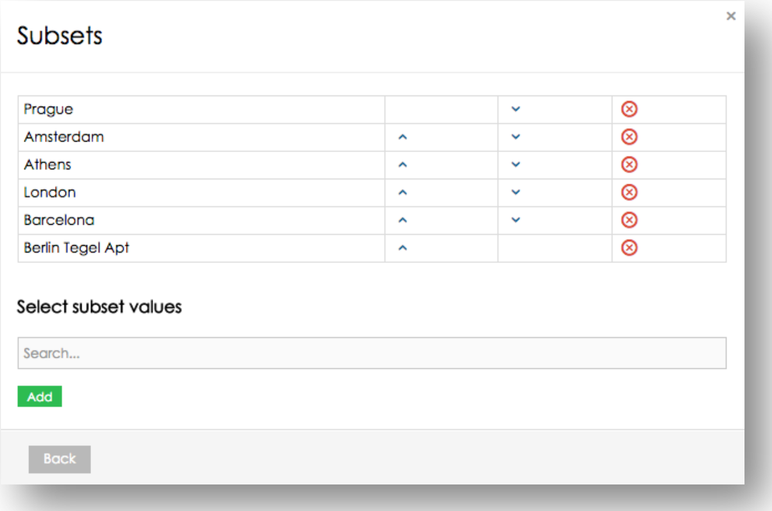
Working with drill-downs

- **Allows to define relation among attributes**
  - Tree structure
  - Multiple branches are allowed
- **Leveraged in reports to move up/down in detail**
- **Suggested by + sign next to each member**


3_Customers - Drill down paths

Working with subsets

- Attribute subset is a virtual copy of attribute
- Select and use only desired members
- Define custom order of members



Subsets

×

Prague		⌄	⊗
Amsterdam	⌃	⌄	⊗
Athens	⌃	⌄	⊗
London	⌃	⌄	⊗
Barcelona	⌃	⌄	⊗
Berlin Tegel Apt	⌃		⊗

Select subset values

Search...

Add

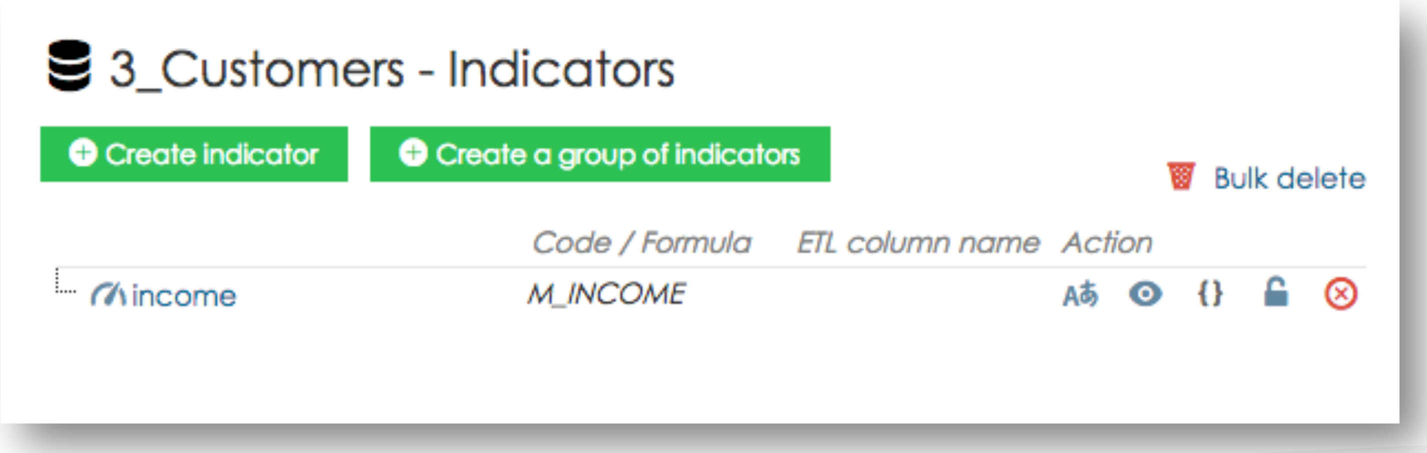
Back

Working with indicators

- **Add, modify, remove**
  - all changes **are propagated** to the reports, where they are used!
- **Each indicator has got unique code, starting with M_**
  - code is used in formulas and scripts
- **Formula indicators (starting with F_)**
- **Import columns are mapped on indicator with codes only**
- **Appearance and localization support**
- **Hide system indicators – not displayed in reports**

Working with indicators

- **Transform existing values**
- **Define the indicator groups**
  - abstract structure only – doesn't have a code
  - possible aggregation operations in report
  - in-report definition possible

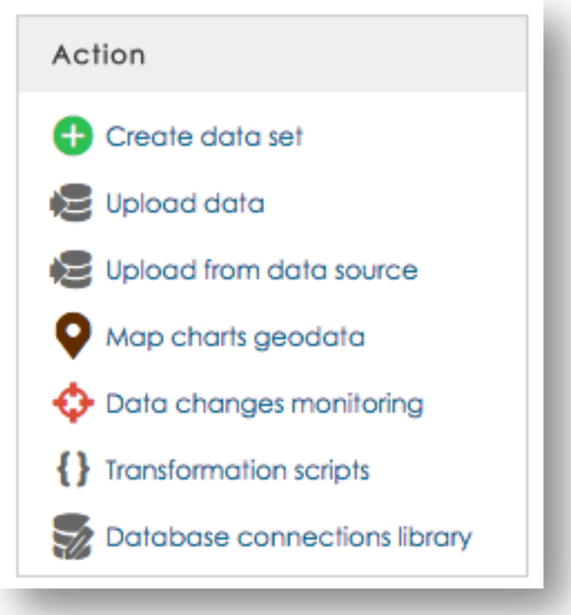


3_Customers - Indicators

	Code / Formula	ETL column name	Action
income	M_INCOME		

Creating data set

- **Data manager** role is required
- Data set name must be unique
- Each data set has unique **system code**
- Can be created:
  - manually
  - by import
  - by connecting to data source

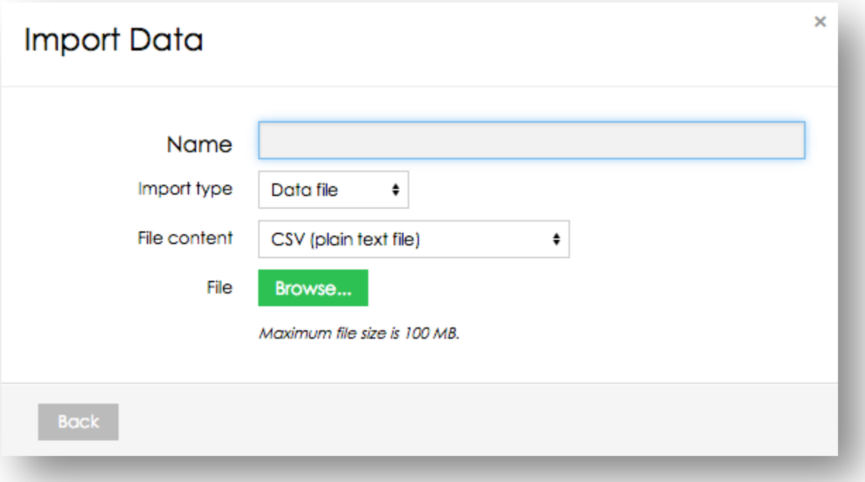


Action

- Create data set
- Upload data
- Upload from data source
- Map charts geodata
- Data changes monitoring
- Transformation scripts
- Database connections library

Data upload - settings

- **Copy & paste or file import**
- **Supported formats:**
  - CSV, XLS (Excel < 2003), XLSX (Excel > 2007), XML
  - ZIP archive
  - various encodings support
  - maximum file size – by default 2MB (configurable)



Import Data

Name:

Import type: Data file

File content: CSV (plain text file)

File:

Maximum file size is 100 MB.

Import settings

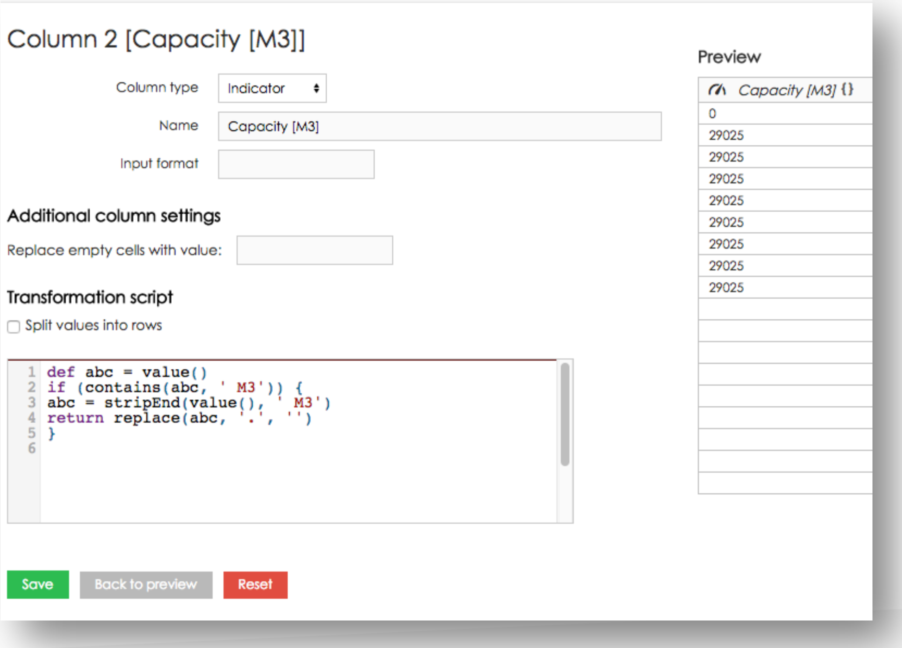
these settings are same also for all data sources

- **Provides mapping on indicator and attributes**
- **Possible column types:**
  - attribute or indicator
  - date/time
  - translation
  - GEO point
  - don’t import
- **Basic operations**
  - file setting: header row, excluded rows count, encoding
  - use import templates
  - reset to default

Import settings

these settings are same also for all data sources

- **Preview data before proceeding with import**
- **Add new columns**
  - create new indicators and attributes
- **Column operations**
  - map on attribute or indicator
  - transformation scripts
  - merge



Column 2 [Capacity [M3]]

Column type: Indicator

Name: Capacity [M3]

Input format:

Additional column settings

Replace empty cells with value:

Transformation script

□ Split values into rows

```
def abc = value()
if (contains(abc, ' M3')) {
abc = stripEnd(value(), ' M3')
return replace(abc, '.', '')
}

```

Preview

Capacity [M3] {}
0
29025
29025
29025
29025
29025
29025
29025
29025

Transformations scripts

- Can be used for data manipulation, data cleaning, creating new columns, split/merge columns
- Example:

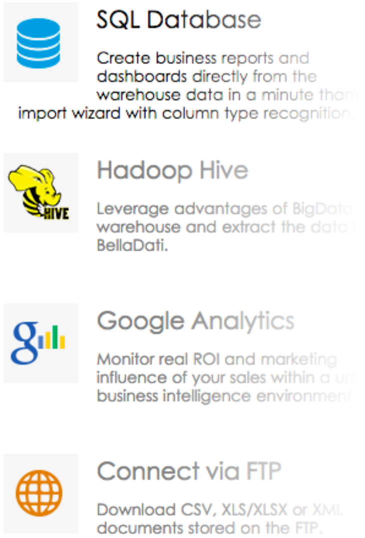
```
age = yearsBetween(date(value(4)), date(actualDate()))
if (age<26) {
    group = 'Junior‘
} else if ((age>=26) && (age <59)) {
    group = 'Adult‘
} else if (age >=60) {
    group = 'Senior'
}
return group;
```

Browsing imported data

- Available for all users with the at least with read access
- Add, modify and delete rows (for RW access only)
- Filter and order
- Appearance settings and localization
- Transformation
- Export – CSV, ZIP (encoded in UTF8)

Automatic import

- **Always from a data source**
- **Supported data sources:**
  - SQL databases
  - Google Analytics, Google Spreadsheets
  - URL, FTP (CSV, XML, XLS, XLSX, ZIP)
  - Facebook, Twitter, YouTube
  - SalesForce, Intuit
  - Existing data set
- **Single data set supports multiple data sources**
- **Automatic changes propagation**



SQL Database

Create business reports and dashboards directly from the warehouse data in a minute than import wizard with column type recognition.

Hadoop Hive

Leverage advantages of BigData warehouse and extract the data BellaDati.

Google Analytics

Monitor real ROI and marketing influence of your sales within a un business intelligence environment.

Connect via FTP

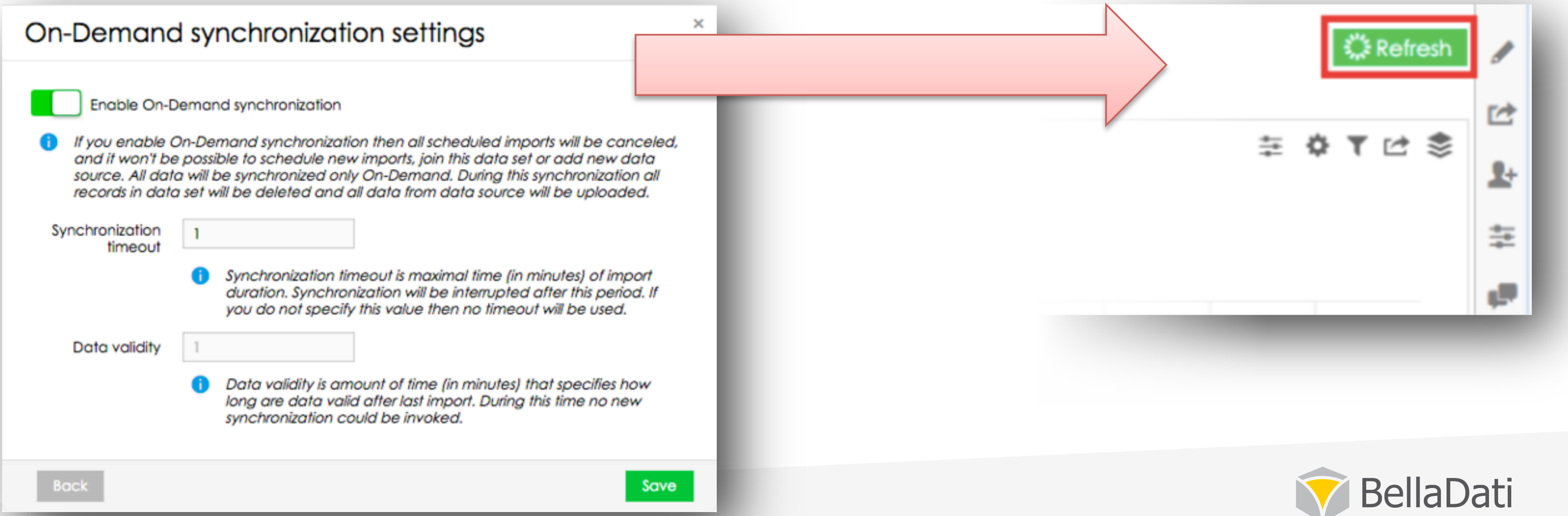
Download CSV, XLS/XLSX or XML documents stored on the FTP.

Automatic import

- **Immediate import**
- **Scheduled execution**
  - available repeating intervals: from minutes to years
  - automatic execution for joined data sets on source data set change
  - multiple executions
- **Overwriting policies**
  - by data/time match, members match
  - delete all data, delete by filter
- **Modifying settings – data source specific**
- **Import settings – same as manual import**

On-demand data synchronization

- Import can triggered by user manually directly from the report
- Imports cannot be scheduled
- Complete data set is overwritten with the new data



On-Demand synchronization settings

Enable On-Demand synchronization

If you enable On-Demand synchronization then all scheduled imports will be canceled, and it won't be possible to schedule new imports, join this data set or add new data source. All data will be synchronized only On-Demand. During this synchronization all records in data set will be deleted and all data from data source will be uploaded.

Synchronization timeout: 1

Synchronization timeout is maximal time (in minutes) of import duration. Synchronization will be interrupted after this period. If you do not specify this value then no timeout will be used.

Data validity: 1

Data validity is amount of time (in minutes) that specifies how long are data valid after last import. During this time no new synchronization could be invoked.

Import results

- **You can display the detailed results for each import:**
  - green - successfully finished import without errors
  - orange - finished import containing errors
  - red - aborted or unfinished import
  - gray - deleted import
  - blue - scheduled future import
- **If errors occurred during import, you can find the reasons in the detail popup**
  - column name containing errors is highlighted errors count and location is displayed

Data set details

- List of All, Successful, failed and upcoming reports
  - Failed reports with error description
- Schedule import
- Reports list
- Localizable name
- Description (WYSIWYG editor)

Data source details

- Change connection parameters
- See imports overview
- Check data source availability
- Send email notification after import is completed
  - Select email triggers:
    - Completed
    - Completed with errors
    - Error
    - You don’t have enough space in your data warehouse
    - Wrong data source configuration

Joining data sets

- **Principle is similar to SQL joins**
- **One primary data set – multiple join points**
- **Joining conditions (attributes match):**
  - left outer join: related record in second data set is optional
  - inner join: related record in second data set is mandatory
  - cross join with condition
- **All joins must be built (recalculated) before they can be used**
  - you can see the progress bar
- **Changes in source data sets are automatically triggered by join**
- **Disabling of build is possible (can be used when playing with large joins)**

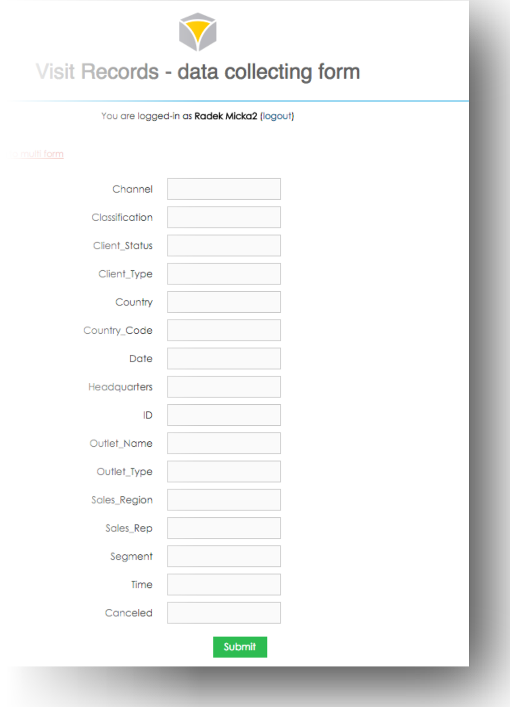
BellaDati

Connections library

- List of SQL data source connections
- Create connection
- Edit connection
- Delete connection
- See the list of data sets connected to each data source
- This list is used when connecting to SQL data source

Data collecting form

- Create and publish form connected to BellaDati Data sets
- Distribute form URL to all users responsible for collecting data
- Multiple forms for one data set



Visit Records - data collecting form

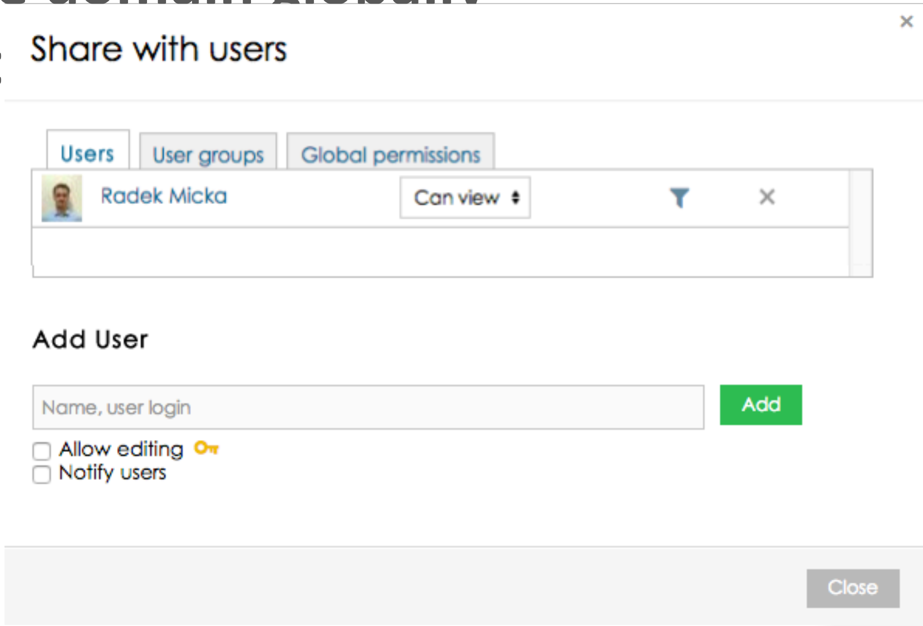
You are logged-in as Radek Micka2 (logout)

to multi form

Channel	
Classification	
Client_Status	
Client_Type	
Country	
Country_Code	
Date	
Headquarters	
ID	
Outlet_Name	
Outlet_Type	
Sales_Region	
Sales_Rep	
Segment	
Time	
Canceled	

Access control & sharing

- **Grant access to selected users or user groups**
  - Read-only access: Only reports or dashboards can be created.
  - Full access: All functions except sharing and data set deleting are
- **Grant access to all users in the domain globally**
- **Restrict access to data subset**
  - Leverage filters
  - Based on attributes/indicators



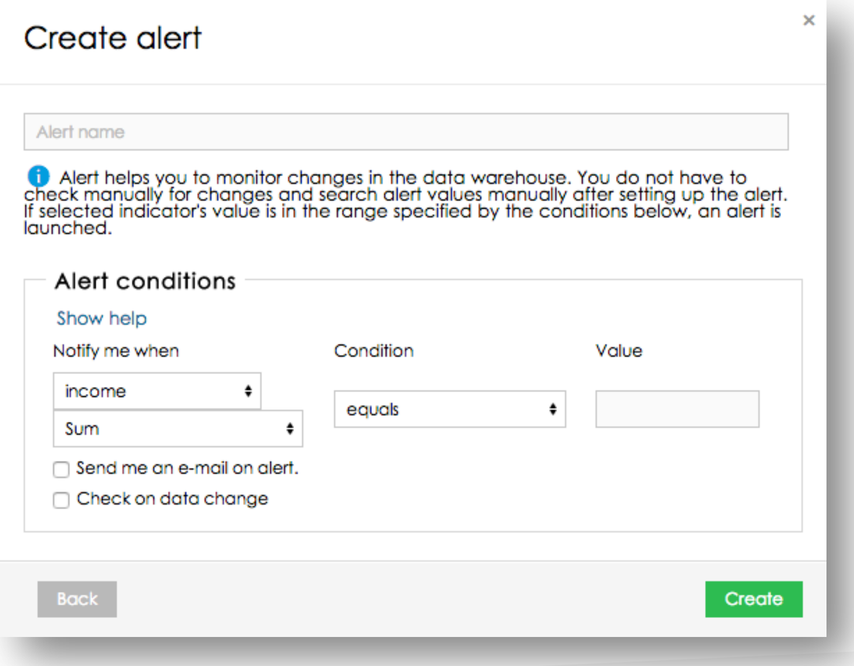
Radek Micka Can view

□ Allow editing
□ Notify users

BellaDati

Watching data changes

- Monitors actual data and checks the alarm conditions
- Triggers automatic notifications
- Multiple alarms for one data set
- Email notifications support
- Mobile device notification
- Known also as alarm



Create alert

Alert name

Alert helps you to monitor changes in the data warehouse. You do not have to check manually for changes and search alert values manually after setting up the alert. If selected indicator's value is in the range specified by the conditions below, an alert is launched.

Alert conditions

Show help

Notify me when: income, Sum

Condition: equals

Value:

□ Send me an e-mail on alert.

□ Check on data change

BellaDati

Concurrent data access

- **One import at same time**
- **Structure is locked for changes while:**
  - data transformation is running
  - joins rebuild is running
- **Reports are not affected – they are using the read-only mode**

Backup & restore

- **The following structures are included:**
  - Data sets
  - Data source settings
  - Alarms
  - Reports
- **Backup of multiple data sources support**
- **Configurable restore**
  - selecting structures to be restored
  - users and user groups replacement

GEO mappers

- **Serves to properly position data on the GEO Map**
- **Point based vs. region (shape) definition**
  - Import points database
  - Import GEO JSON
- **Leverage existing geo-data**
- **Create own definition**
- **Definition is associated with the domain**

Withdrawal amount (mil. $)
Gelang Patah
Nusajaya
Singapore
Sungai Rengit

Visits
1–182.875
182.875–364.75
364.75–546.625
546.625–728.5
728.5–910.375
910.375–1,092.25
1,092.25–1,274.125
1,274.125–1,456
Stockholm
Moscow
Paris
France
Barcelona
Roma
Napoli
Malta
Algiers
Gibraltar
Türkmenistan

Documentation

- Visit the documentation at

http://support.belladati.com

BellaDati

Q&A