Geocoding

Encode region names into geographical coordinates, or reverse-geocode latitude and longitude pairs into regions.

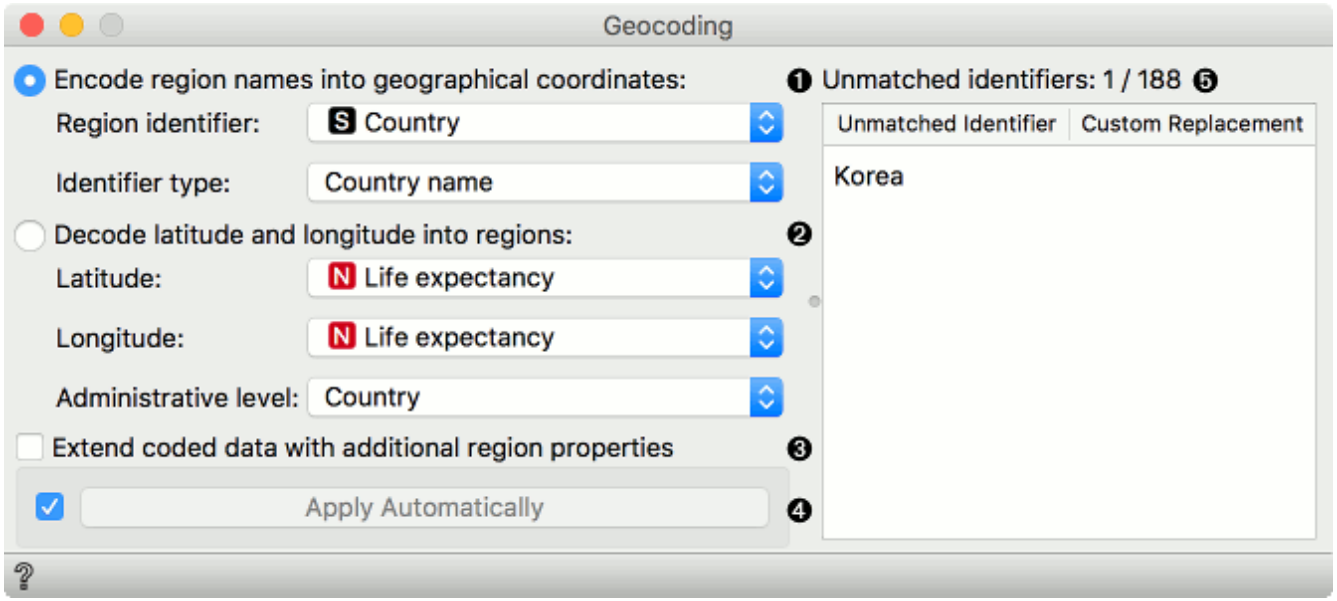
Inputs

- Data: An input data set.

Outputs

- Coded Data: Data set with new meta attributes.

Geocoding widget extracts latitude/longitude pairs from region names or synthesizes latitude/longitude to return region name. If the region is large, say a country, encoder with return the latitude and longitude of geometric centre.



1. Use region names to extract the corresponding latitude/longitude pairs:
    - Region identifier: attribute with the information on region names. Can be discrete or string.
    - Identifier type: define how the data is coded. Supports ISO codes and some major cities and countries.

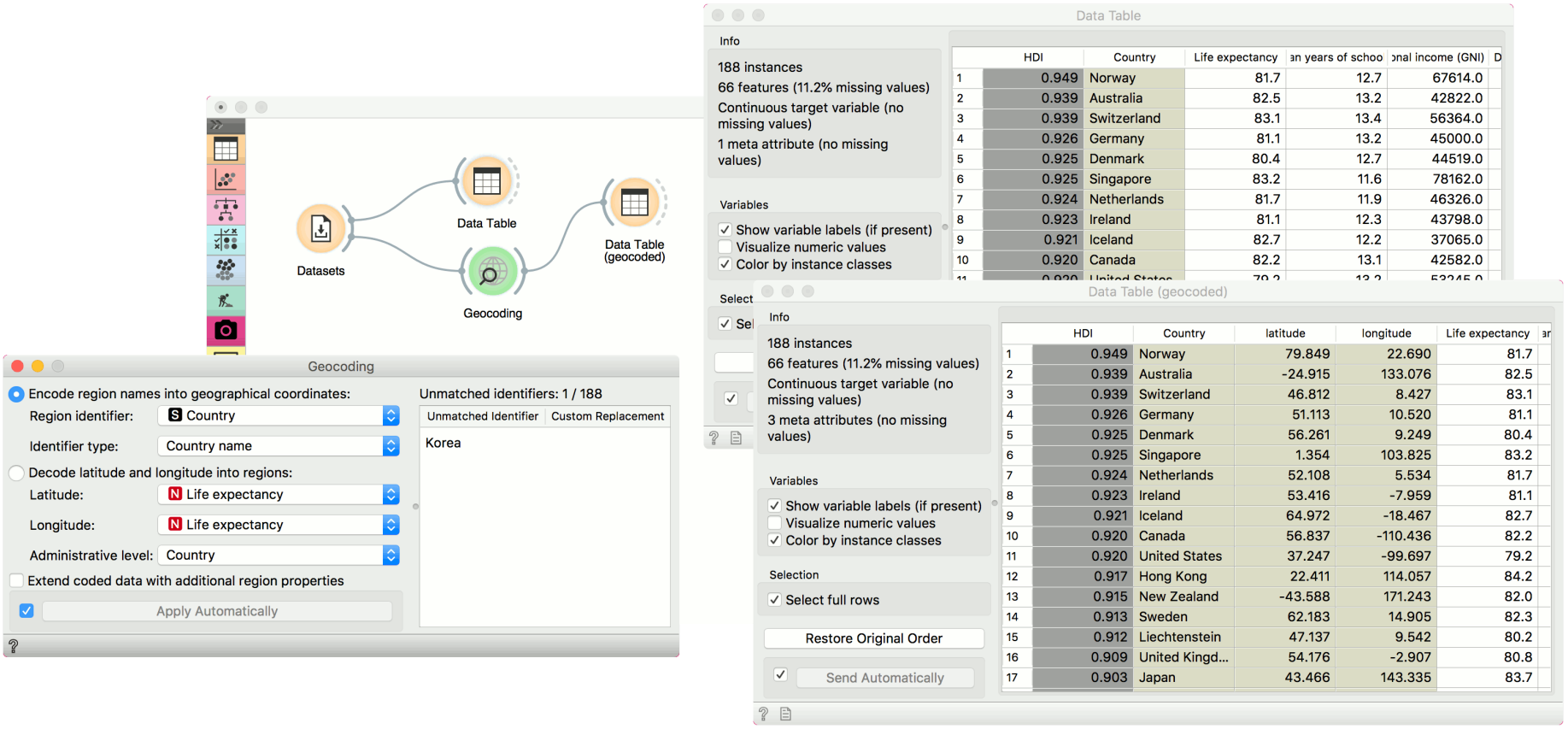
2. Use latitude and longitude pairs to retrieve region names:
    - Latitude attribute.
    - Longitude attribute.
    - Administrative level of the region you wish to extract.
3. Extend coded data adds additional information on the region of interest. For countries, for example, one would get economy, type, continent, etc.
4. If *Apply Automatically* is ticked, the changes will be communicated automatically. Alternatively, press *Apply*.
5. Unmatched identifiers editor. Match regions names that couldn't be matched automatically with their corresponding name.

Example

We will use *HDI* data from the **Datasets** widget. Open the widget, find *HDI* data, select it and press *Send*. First, let us observe the data in a **Data Table**. We have a meta attribute names *Country*, which contains country names. Now we would like to plot this on a map, but **Geo Map** widget requires latitude and longitude pairs. **Geocoding** will help us extract this information from country names.

Connect **Geocoding** to **Datasets**. Region identifier in our case is the attribute *Country* and the identifier type is *Country name*. If our data contained major European cities, we would have to select this from the dropdown. On the right there is the *Unmatched identifier*editor, which shows those instances, for which **Geocoding** couldn't find corresponding latitude/longitude pairs. We can help the widget by providing a custom replacement. Click on the field and start typing *Korea*. The widget will suggest two countries, Democratic Republic of Korea and South Korea. Select the one you wish to use here.

Finally, we can observe the data in the second **Data Table**. We can see our data now has two additional attributes, one for the latitude and one for the longitude of the region of interest. Now, you can plot the data on the map!


Datasets
Data Table
Geocoding
Data Table (geocoded)
Geocoding
Encode region names into geographical coordinates:
Region identifier: Country
Identifier type: Country name
Decode latitude and longitude into regions:
Latitude: Life expectancy
Longitude: Life expectancy
Administrative level: Country
Extend coded data with additional region properties
Apply Automatically
Unmatched identifiers: 1 / 188
Unmatched Identifier
Custom Replacement
Korea
Data Table
Info
188 instances
66 features (11.2% missing values)
Continuous target variable (no missing values)
1 meta attribute (no missing values)
Variables
Show variable labels (if present)
Visualize numeric values
Color by instance classes
HDI Country Life expectancy
1 0.949 Norway 81.7
2 0.939 Australia 82.5
3 0.939 Switzerland 83.1
4 0.926 Germany 81.1
5 0.925 Denmark 80.4
6 0.925 Singapore 83.2
7 0.924 Netherlands 81.7
8 0.923 Ireland 81.1
9 0.921 Iceland 82.7
10 0.920 Canada 82.2
Data Table (geocoded)
Info
188 instances
66 features (11.2% missing values)
Continuous target variable (no missing values)
3 meta attributes (no missing values)
Variables
Show variable labels (if present)
Visualize numeric values
Color by instance classes
Selection
Select full rows
Restore Original Order
Send Automatically
HDI Country latitude longitude Life expectancy
1 0.949 Norway 79.849 22.690 81.7
2 0.939 Australia -24.915 133.076 82.5
3 0.939 Switzerland 46.812 8.427 83.1
4 0.926 Germany 51.113 10.520 81.1
5 0.925 Denmark 56.261 9.249 80.4
6 0.925 Singapore 1.354 103.825 83.2
7 0.924 Netherlands 52.108 5.534 81.7
8 0.923 Ireland 53.416 -7.959 81.1
9 0.921 Iceland 64.972 -18.467 82.7
10 0.920 Canada 56.837 -110.436 82.2
11 0.920 United States 37.247 -99.697 79.2
12 0.917 Hong Kong 22.411 114.057 84.2
13 0.915 New Zealand -43.588 171.243 82.0
14 0.913 Sweden 62.183 14.905 82.3
15 0.912 Liechtenstein 47.137 9.542 80.2
16 0.909 United Kingd... 54.176 -2.907 80.8
17 0.903 Japan 43.466 143.335 83.7