

# DATOS BALSTĪTS PLĀNOŠANAS RĪKS

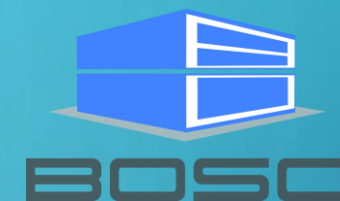
RAITIS BĒRZIŅŠ, SIA BOSC, RAITISBE@GMAIL.COM



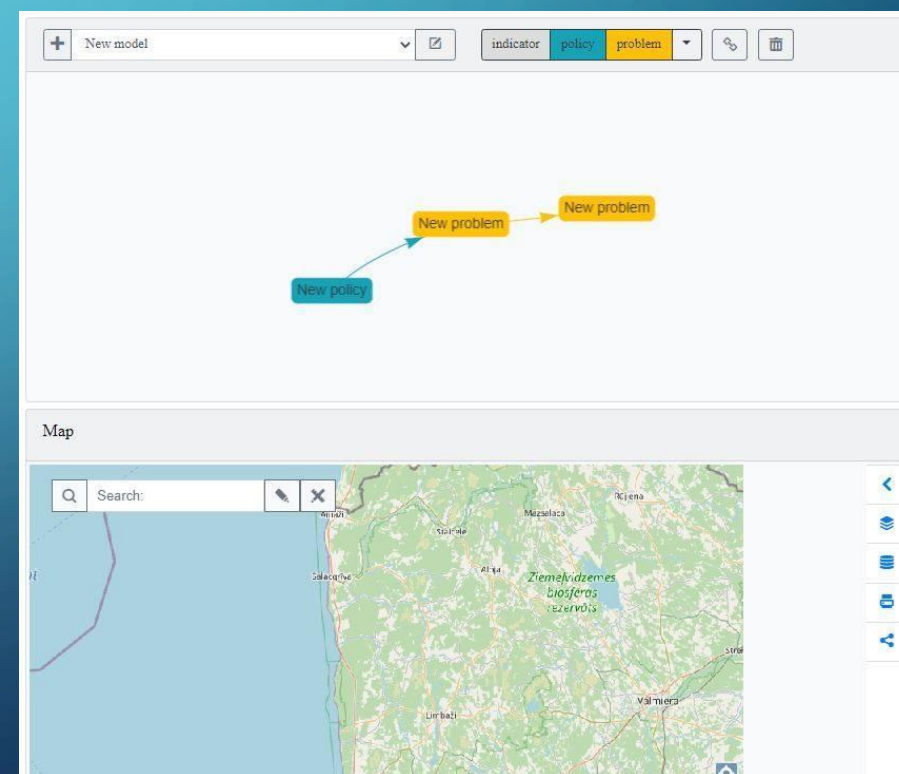
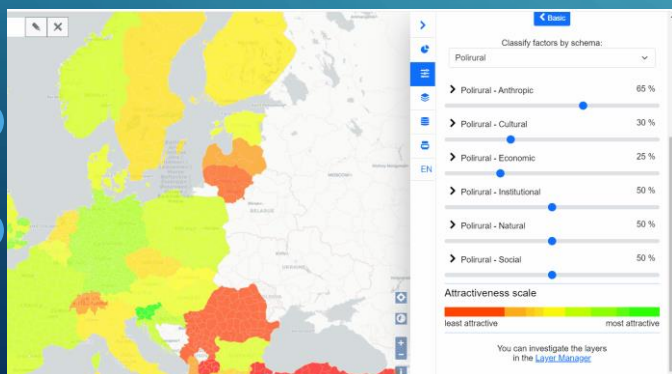
It received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 818496



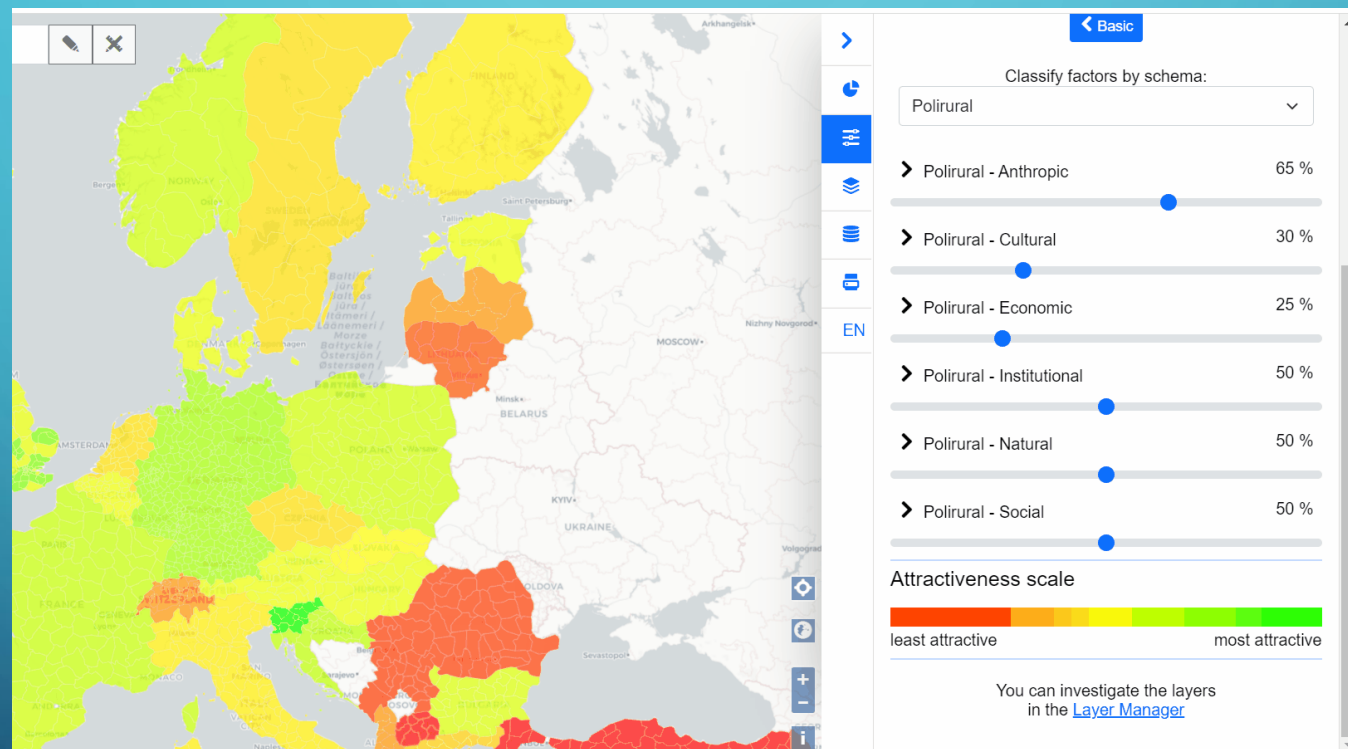
# "BOSC" LOMA POLIRURAL PROJEKTĀ



- Tehniskais atbalsts
- Programmatūras izstrāde
  - Pirmais prototips "Attractiveness Explorer"
  - Modeļu veidotājs un semantiskā pārlūka integrācija
  - Statistiskās analīzes un prognožu rīks



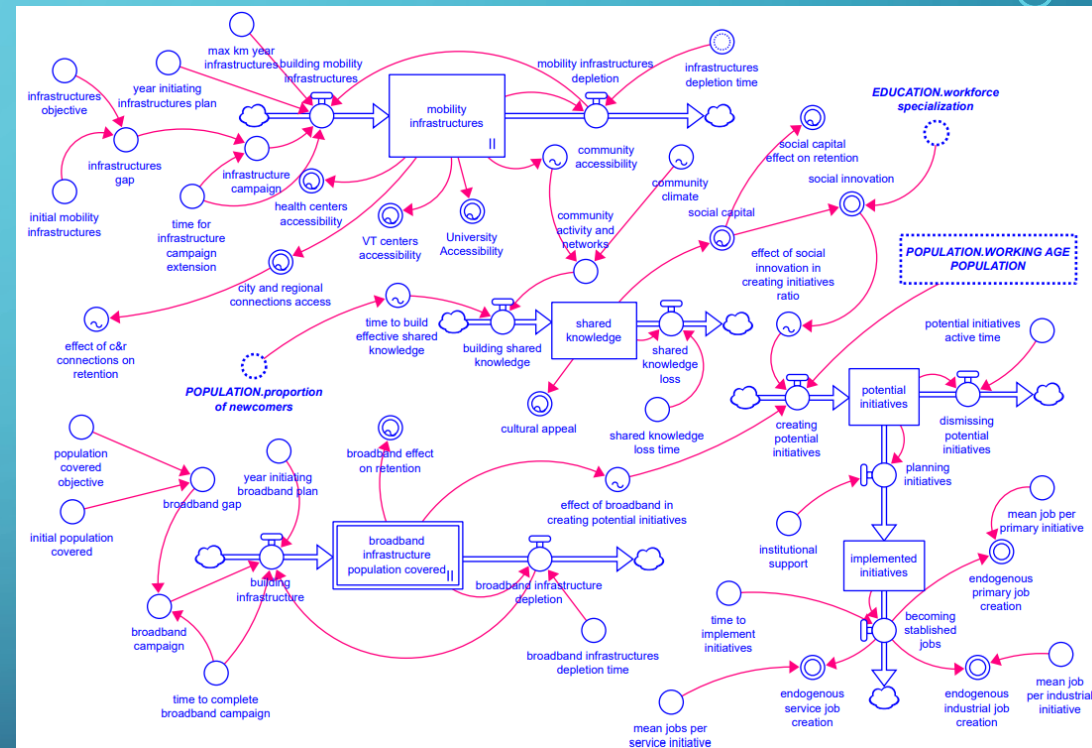
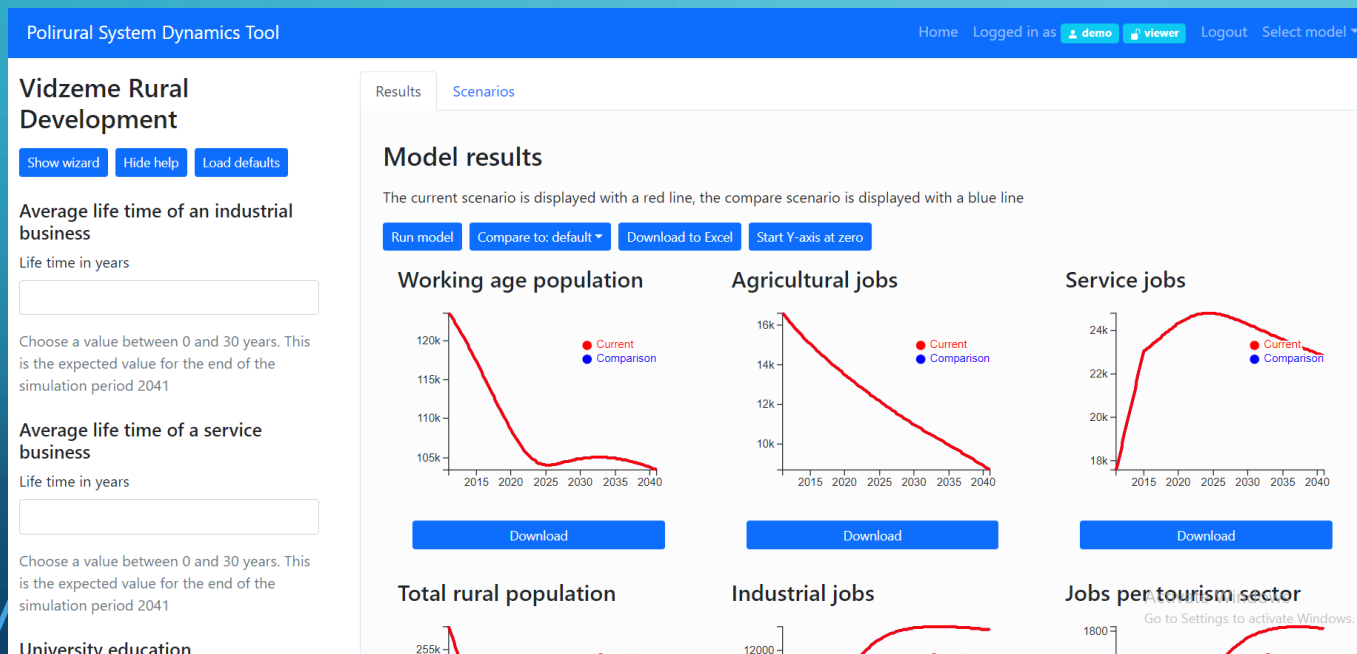
# LAUKU PIEVILCĪBAS PĀRLŪKS



<https://hub.polirural.eu/rural-attractiveness-clustering>

# SISTĒMDINAMIKA

Kā pierādīt izmaiņu ātrumu (koeficientus)?



# DATU AVOTI

- Eurostat Data Explorer
- Centrālā statistikas pārvalde
- RAIM

<https://ec.europa.eu/eurostat/data/database>

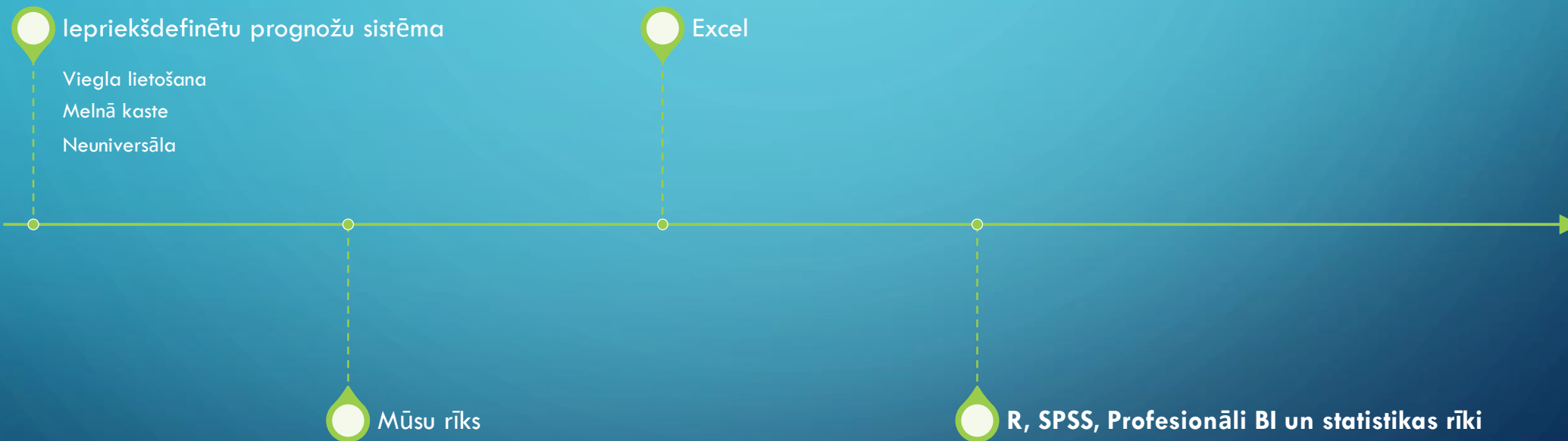


# IDEJA

Rīks, kas ļauj viegli vizualizēt statistiskos datus CSV formātā, aprēķināt korelācijas starp mainīgajiem, kā arī izveidot un darbināt lineārās regresijas prognožu modeļus.



# SAREŽĢĪTĪBA



# DARBA PLŪSMA

Augšupielāde

Visualizēšana

Korelēšana

Modeļu  
ģenerēšana

Prognozes

- Karte
- Laika sērijas
- Aprakstošā statistika

- Lineārā regresija
- Multiplā lineārā regresija

- Laika sērijas (līkņu diagrammas)



# DEMO

- <https://dih.bosc.lv/statistical-predictions/>

# CSV DATU FAILU IELĀDE

Upload data

Upload data in CSV format. A template can be downloaded [here](#).  
Example dataset for life satisfaction of people aged over 16: [life\\_satisfaction\\_over\\_16.csv](#).  
Example dataset for energy consumption in kilograms of oil per capita per year: [energy\\_consumption.csv](#).

Drag and drop files

or

Choose files...

List variables

Current variables: 0

Time series chart

To map

Clear all data

Visualize time series

Location:

☐ Merge charts

Darbības notiek klienta pusē

# LAIKA SĒRIJAS

## Visualize time series ?

Location: ▼

☐ Merge charts

☒ Life satisfaction

Frequency: year

☒ Energy consumption

Frequency: year

## Visualize on map ?

To visualize data on a map we must first choose the layer on which to draw and specify attribute (country code, ZIP, field number, etc.) on the geographical objects by which to find matching records in the CSV. Values must match with the column marked as 'location' when uploading the CSV data.

Select a variable (phenomenon) which to visualize and time dimension (year) because it makes sense to compare values between regions only in the same year.

Layer:

Latvian municipalities ▼

Location attribute in map layer: ▼

## Visualize on map

Layer:

Latvian municipalities ▼

Location property:



Variable:

Life expectancy ▼

Filter by time:

2013 ▼

Minimum 4.8

Maximum 8

Descriptive statistics

Visualize



> Layermanager

Filter:

Baselayers ☒

☒ OpenStreetMap

Map content

☐ Nuts3 regions

☐ Nuts2 regions

☐ Nuts1 regions

☐ Nuts0 regions (countries)

☐ Vidzeme statistical region municipalities

☒ Latvian municipalities

# KORELĀCIJAS

Correlate



Regression

Predict

## Correlate

Variable list

Correlation matrix

Variable 1	Variable 2	Correlation	N pairs
Energy consumption Shift 0 years 	Life satisfaction	0.66	73
Life satisfaction Shift 0 years 	Energy consumption	0.66	73

# MODELA IZVEIDE

## Create predictive models

In case of a significant correlation we can choose one of the variables as dependant, the other as independent variable.

Scatter plots are used to plot data points on a horizontal and a vertical axis in the attempt to show how much one variable is affected by another. Each row in the data table is represented by a marker whose position depends on its values in the columns set on the X and Y axes.

Linear regression analysis estimates the coefficients of the linear equation, involving one or more independent (choose multiple-linear) variables that best predict the value of the dependent variable. Linear regression fits a straight line or surface that minimizes the discrepancies between predicted and actual output values.

The calculated regression coefficients (statistical model) can be saved to make predictions for the dependent variable in case we know or can make an educated guesses for the future values of independent variable.

Linear ▾ Regression

Variable: ▾

☒ Energy consumption

Shift 0 years

☒ Life satisfaction

Shift 0 years

## Input X and Predict Y

By selecting one of the previously created linear regression models we can predict the future values of the dependent statistical variable.

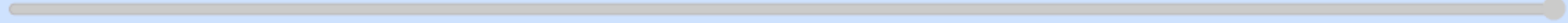
Please fill the date range to generate input fields to be filled for the independent variable. You can fill them manually or use a sketch area to draw a line describing the function.

Stored statistical models: ▾



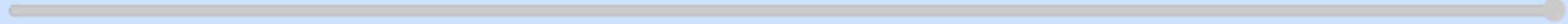
# MODEĻA IZPILDE

Shift 0 years



☒ Life satisfaction

Shift 0 years



## Input X and Predict Y



By selecting one of the previously created linear regression models we can predict the future values of the dependent statistical variable.

Please fill the date range to generate input fields to be filled for the independent variable. You can fill them manually or use a sketch area to draw a line describing the function.

Stored statistical model: ▼

From 2022

Till 2032

Variable

Location (optional): ▲

# REZULTĀTI

- Rīks pieejams <https://dih.bosc.lv/statistical-predictions/>
- Atvērtais kods <https://github.com/hslayers/statistical-predictions>
- Rīks tika veiksmīgi izmantots, lai analizētu augstas pievienotās vērtības industriju attīstību.