ČESKÝ STATISTICKÝ ÚŘAD

„SHINY“ APLIKACE PRO ÚMRTNOSTNÍ TABULKY V PROGRAMU R

David Morávek, Jiří Novák
ČSÚ

24. – 26. května 2023, Hradec Králové
52. konference České demografické společnosti

Obsah prezentace

- Cíl a motivace vývoje aplikace
- Metodika výpočtu úmrtnostních tabulek
- R package Shiny
- Ukázka aplikace CZSOLifeTables

Cíl a motivace

Cíl

- vytvoření webové aplikace v prostředí jazyka R pro data úmrtnostních tabulek:
  - pro jejich výpočet,
  - pro jejich export do struktury publikace,
  - pro jejich vizualizaci.

Motivace:

- zjednodušení procesu publikování dat,
- zpřístupnění aplikace dalším uživatelům (veřejnosti, VŠ, výzkumníkům, …),
- inovace a ukázat cestu ke zlepšení procesů na ČSÚ.

International Data Science Accelerator Programme

„Je skvělou příležitostí pro mezinárodní spolupráci v oblasti datové vědy přinášející odborné znalosti z různých zemí. Podporuje rozvoj, sdílení znalostí a vytváří propojení mezi různými částmi světa.“

Data Science Campus

Office for National Statistics

Metodika výpočtu úmrtnostních tabulek

Metodika výpočtu tabulek

- Od roku 2018 používá ČSÚ aktualizovanou metodiku.
- Tabulky se počítají pro muže a ženy od věku 0 do 105 let, pro ČR, regiony soudržnosti, kraje, okresy a správní obvody obcí s rozšířenou působností (SO ORP).
- Pro vyloučení nahodilých výkyvů v datech se tabulky počítají:
    - za dvouleté kalendářní období pro kraje (NUTS 3).
    - za pětileté kalendářní období pro okresy (LAU 1) a SO ORP.

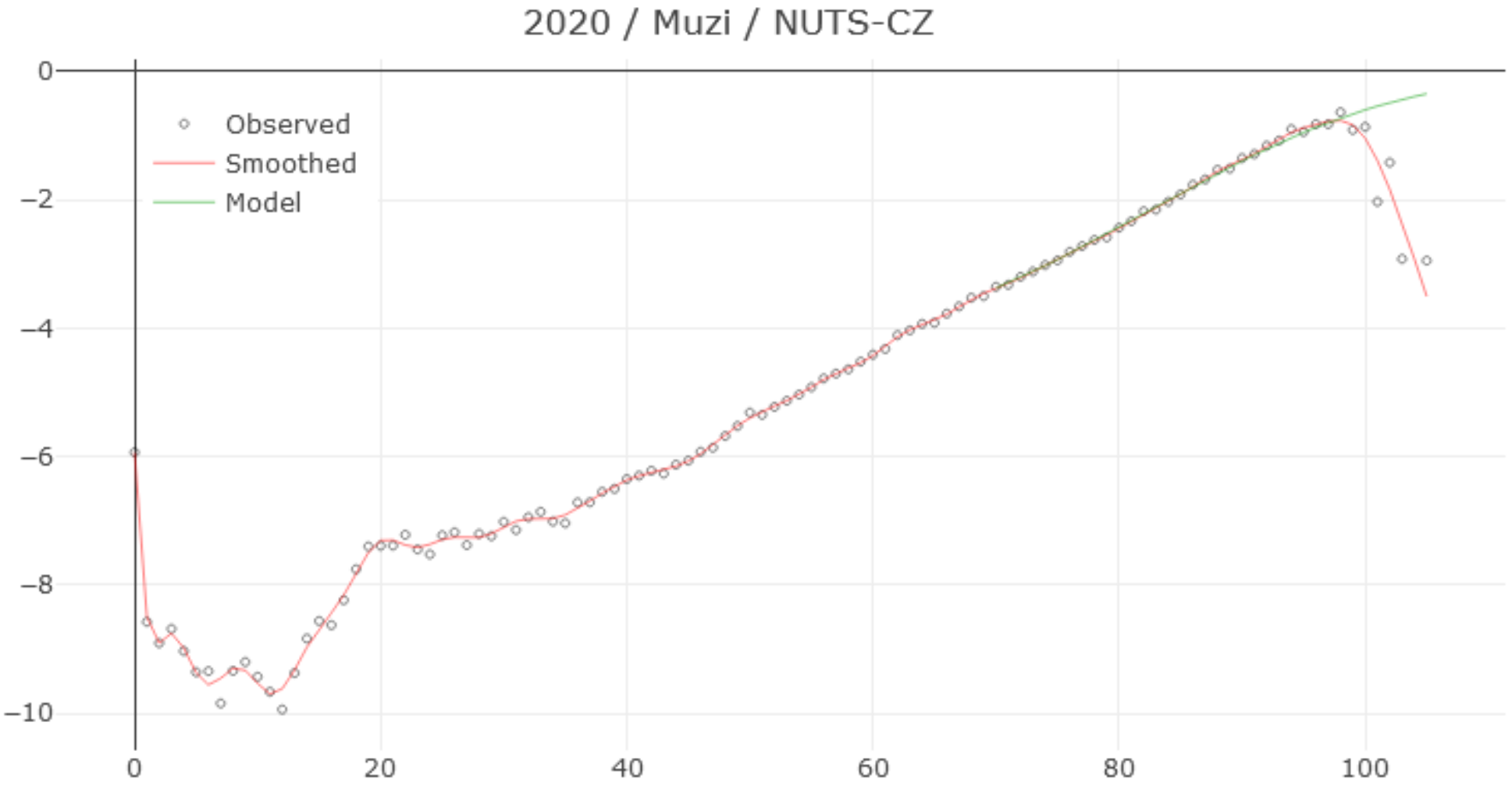
Vyrovnané míry úmrtnosti

- Pro vyloučení nahodilých výkyvů v datech ve věcích s nízkou úmrtností nebo nízkými počty žijících se používá metoda pro vyhlazení měr úmrtnosti.
- Míry úmrtnosti jsou vyhlazeny metodou zobecněných adaptivních modelů v kombinaci s tzv. P-spliny, někdy označováno jako metoda P-GAM (Eilers a Marx, 1996).

Modelové míry úmrtnosti

- S ohledem na nízké počty zemřelých ve vysokém věku a nižší spolehlivost údajů o středním stavu nejstarší populace jsou vyrovnané pozorované míry úmrtnosti v nejvyšších věcích nahrazeny modelovými hodnotami.
- Zvolený model je založen na logistické křivce, která zohledňuje zpomalení nárůstu úmrtnosti s věkem (Thatcher a kol., 1998).

Pozorované, vyhlazené a modelové míry úmrtnosti



Volba modelu

Model name	Model type
Makeham	exponential
Gompertz	exponential
Kannisto	logistic
Logistic2	logistic
Gamma-Gomp-Makeham	multiplicative frailty model
Gamma-Gomp	multiplicative frailty model

Webová interaktivní aplikace: R Shiny

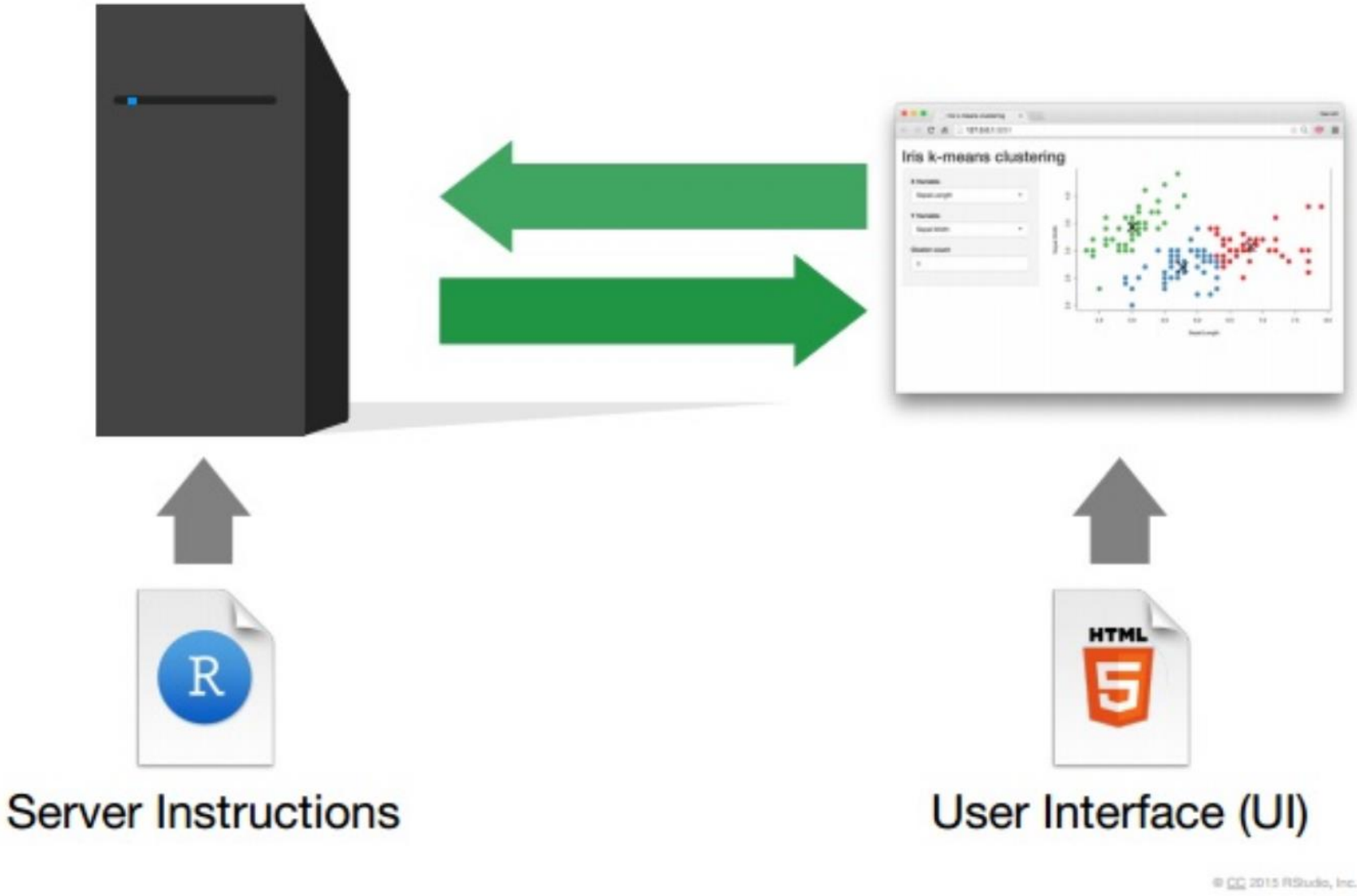
Shiny aplikace

- Shiny je balíček v jazyce R
- Pro vytváření interaktivních webových aplikací

- Shiny kombinuje výkon R s interaktivitou moderního webu.

- První verze na podzim 2012
- Author Joe Cheng z Posit (dříve Rstudio)

Shiny aplikace — princip



© CC 2015 RStudio, Inc.

Shiny aplikace — implementace

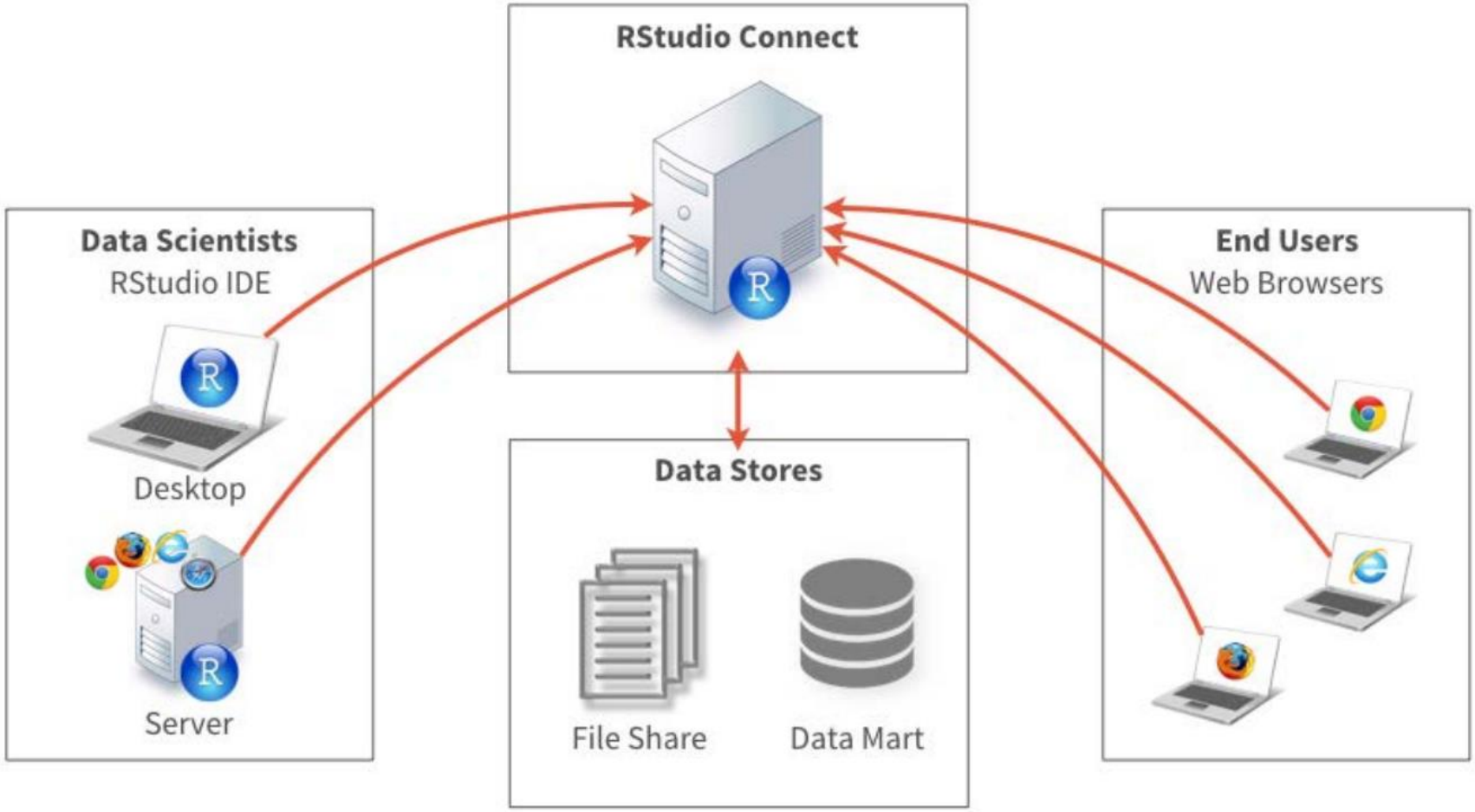
Lokální řešení

- Lokální spouštění scriptu přes R
- Vytvoření balíčku obsahujícího aplikaci
  - Lokální spouštění přes R
  - Spouštění přes R-front-end, tzv. přes ikonu

Cloudové řešení

- Využití cloudových služeb — https://www.shinyapps.io/
- Využití R serveru — https://www.rstudio.com/products/workbench/

Shiny aplikace — R server



Ukázka aplikace

CZSOLifeTables

CZSO Life Tables

Parameter selection

Population

Model

Change language

en

Home | Explore | Customize | Map | Development

Welcome to CZSOLifeTables!

An interactive open-source software tool for calculating life tables for the Czech Republic and its cohesion regions, regions, districts, and administrative districts with extended powers.

Calculate, Explore, Customize, and Visualize life table data!

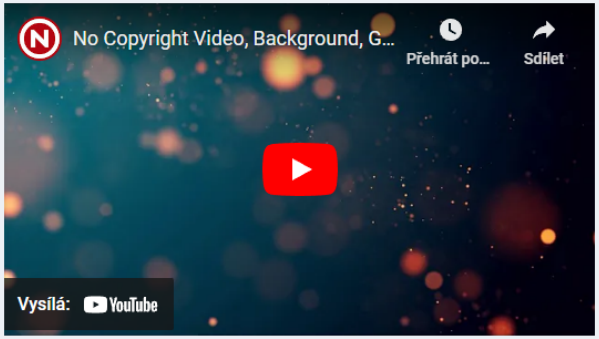
From the left panel menu, you will be able to set the parameters for the population for which you will calculate the life table, and parameters for mortality modelling, and graph settings.

Explore – View and visualize data

Customize – Prepare data for export and publishing

Map – Visualize data territorially

For a quick look, you can watch a tutorial video below.



Credits

The application was developed as part of the International Data Science Accelerator program in 2022 under the title

"Developing a web application from scratch for life table data using R package Shiny".
The main goal was to simplify the process of processing and publishing life table data at the Czech Statistical Office.

Authors:

Jiří Novák, David Morávek, Alberto Sanchez Rodelgo

CZSO Life Tables

Parameter selection

Population

Model

Model name: Kannisto

Alpha: 0,1

Max age: 105

Est. range: 70 to 90

Fit from: 70

Min merge age: 75

Column to calculate: ex

Change language: en

CZSO

Home | Explore | Customize | Map | Development

Life Table

Select region: CZ010

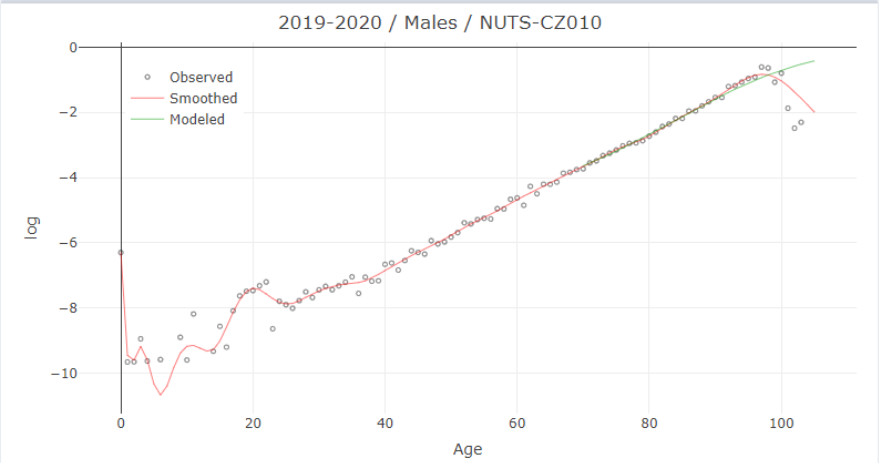
Specify x axis range: to

Specify y axis range: to

Show 7 entries Search:

	Year	Sex	Region	Age	Dx	Px	mx	qx	lx	dx	Lx
1	2019-2020	1	CZ010	0	28.00	15,215.00	0.00	0.00	100,000.00	183.72	99,834.65
2	2019-2020	1	CZ010	1	1.23	15,596.00	0.00	0.00	99,816.28	7.86	99,812.35
3	2019-2020	1	CZ010	2	1.05	15,517.00	0.00	0.00	99,808.42	6.76	99,805.04
4	2019-2020	1	CZ010	3	1.58	15,327.00	0.00	0.00	99,801.66	10.31	99,796.51
5	2019-2020	1	CZ010	4	1.03	15,129.00	0.00	0.00	99,791.35	6.82	99,787.94
6	2019-2020	1	CZ010	5	0.48	14,772.00	0.00	0.00	99,784.53	3.27	99,782.89
7	2019-2020	1	CZ010	6	0.34	14,469.00	0.00	0.00	99,781.26	2.32	99,780.10

Showing 1 to 7 of 106 entries Previous 1 2 3 4 5 … 16 Next



Export the displayed table

Export all tables in one csv

Export all tables separately in zip

Export the displayed plot

Export all plots in pdf

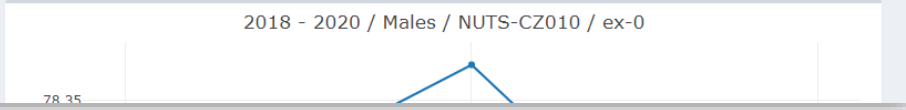
Export all plots in zip

Aggregated Life Table

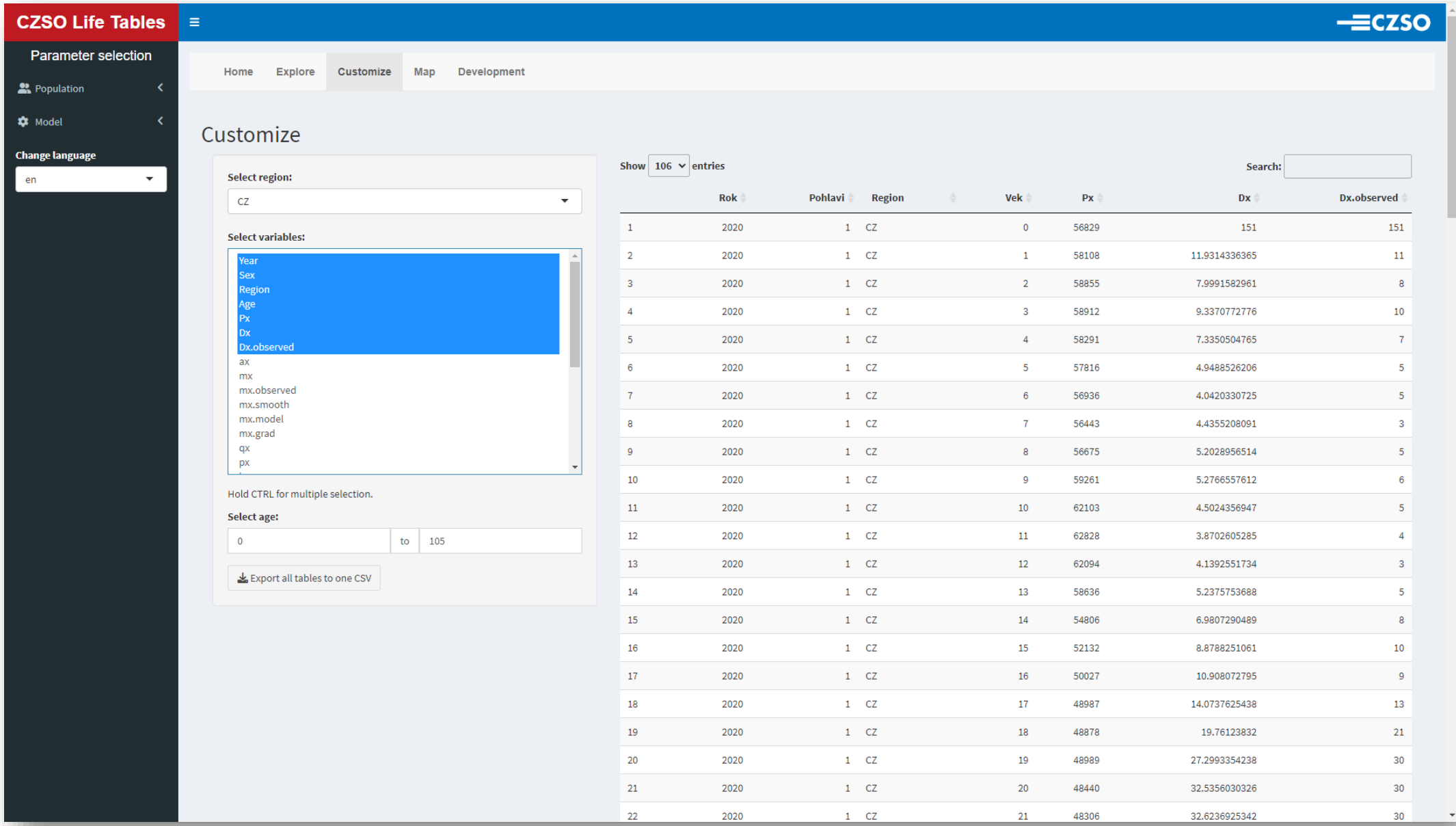
Select column: All

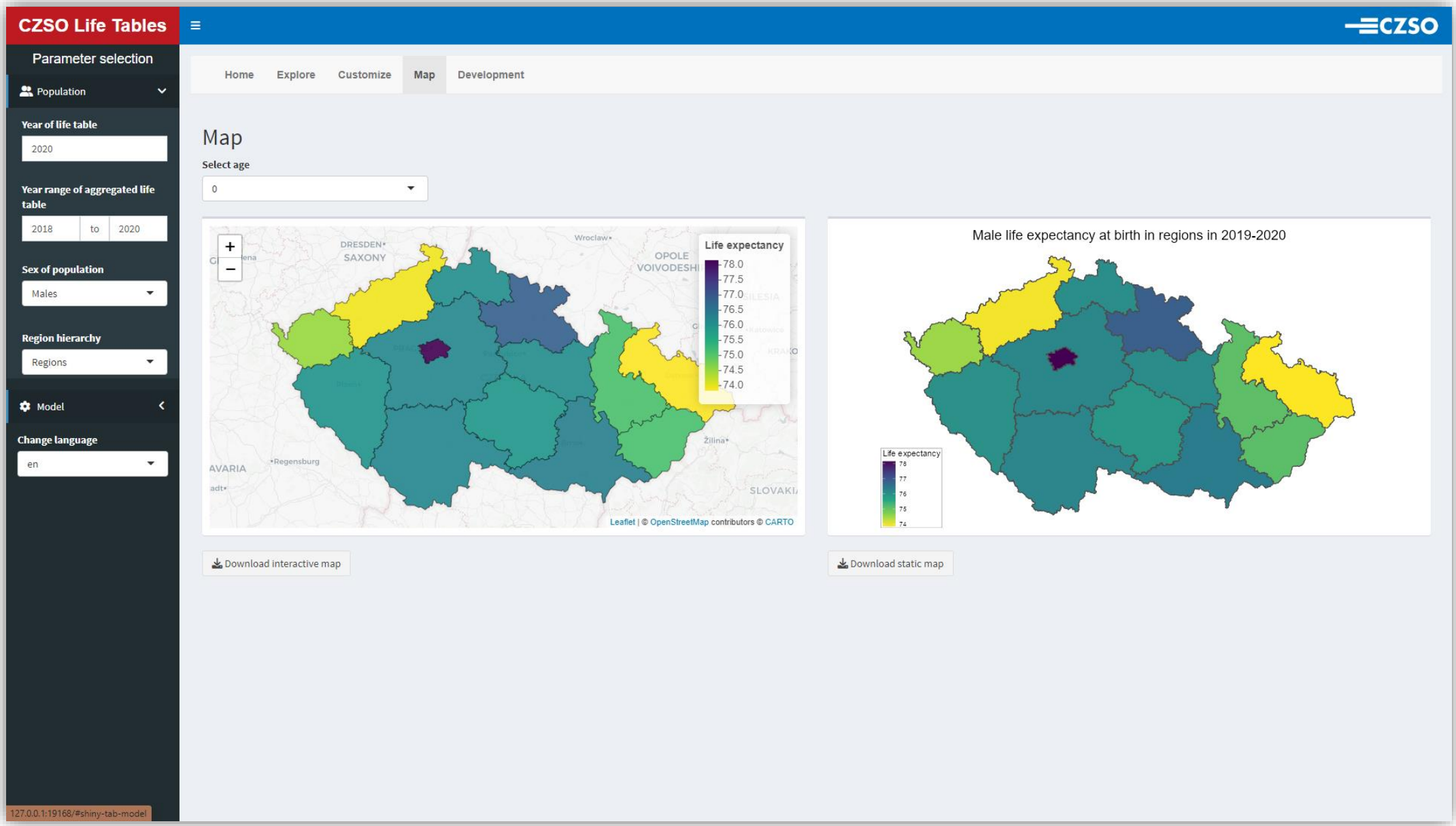
Show 10 entries Search:

Year	Region	ex.0	ex.1	ex.2	ex.3	ex.4	ex.5	ex.6	ex.7	ex.8	ex.



ČSÚ


CZSO Life Tables
Parameter selection
Population
Model
Change language
en
Home
Explore
Customize
Map
Development
CZSO
Customize
Select region:
CZ
Select variables:
Year
Sex
Region
Age
Px
Dx
Dx.observed
ax
mx
mx.observed
mx.smooth
mx.model
mx.grad
qx
px
Hold CTRL for multiple selection.
Select age:
0
to
105
Export all tables to one CSV
Show 106 entries
Search:
Rok Pohlavi Region Vek Px Dx Dx.observed
1 2020 1 CZ 0 56829 151 151
2 2020 1 CZ 1 58108 11.9314336365 11
3 2020 1 CZ 2 58855 7.9991582961 8
4 2020 1 CZ 3 58912 9.3370772776 10
5 2020 1 CZ 4 58291 7.3350504765 7
6 2020 1 CZ 5 57816 4.9488526206 5
7 2020 1 CZ 6 56936 4.0420330725 5
8 2020 1 CZ 7 56443 4.4355208091 3
9 2020 1 CZ 8 56675 5.2028956514 5
10 2020 1 CZ 9 59261 5.2766557612 6
11 2020 1 CZ 10 62103 4.5024356947 5
12 2020 1 CZ 11 62828 3.8702605285 4
13 2020 1 CZ 12 62094 4.1392551734 3
14 2020 1 CZ 13 58636 5.2375753688 5
15 2020 1 CZ 14 54806 6.9807290489 8
16 2020 1 CZ 15 52132 8.8788251061 10
17 2020 1 CZ 16 50027 10.908072795 9
18 2020 1 CZ 17 48987 14.0737625438 13
19 2020 1 CZ 18 48878 19.76123832 21
20 2020 1 CZ 19 48989 27.2993354238 30
21 2020 1 CZ 20 48440 32.5356030326 30
22 2020 1 CZ 21 48306 32.6236925342 30


CZSO Life Tables
Parameter selection
Population
Year of life table
2020
Year range of aggregated life table
2018
to
2020
Sex of population
Males
Region hierarchy
Regions
Model
Change language
en
Home
Explore
Customize
Map
Development
Map
Select age
0
Life expectancy
78.0
77.5
77.0
76.5
76.0
75.5
75.0
74.5
74.0
Leaflet | © OpenStreetMap contributors © CARTO
Male life expectancy at birth in regions in 2019-2020
Life expectancy
Download interactive map
Download static map
127.0.0.1:19168/#shiny-tab-model

ČESKÝ STATISTICKÝ ÚŘAD

czso.shinyapps.io/demo/

Děkujeme za pozornost.

jiri.novak@czso.cz
david.moravek@czso.cz