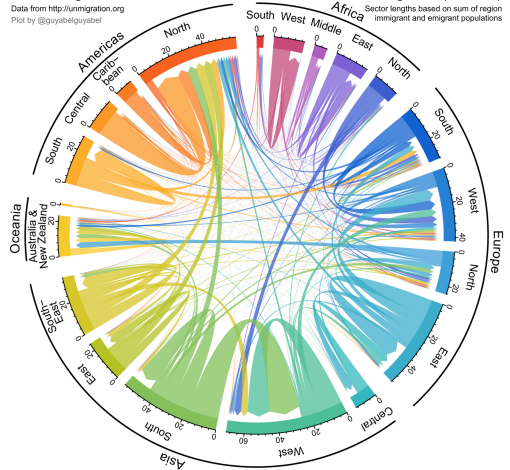# Use of R in Demography

Dilek Yıldız[1,2]

[1]Vienna Institute of Demography/Austrian Academy of Sciences

[2]International Institute for Applied Systems Analysis

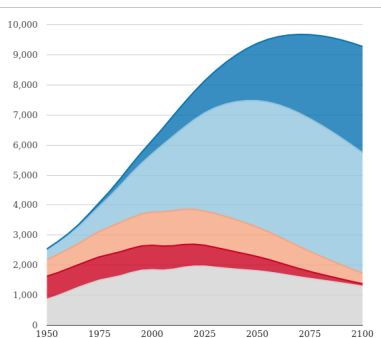R Day on Demography and Population - 14 July 2021

Abel and Cohen, Scientific Data, 2019
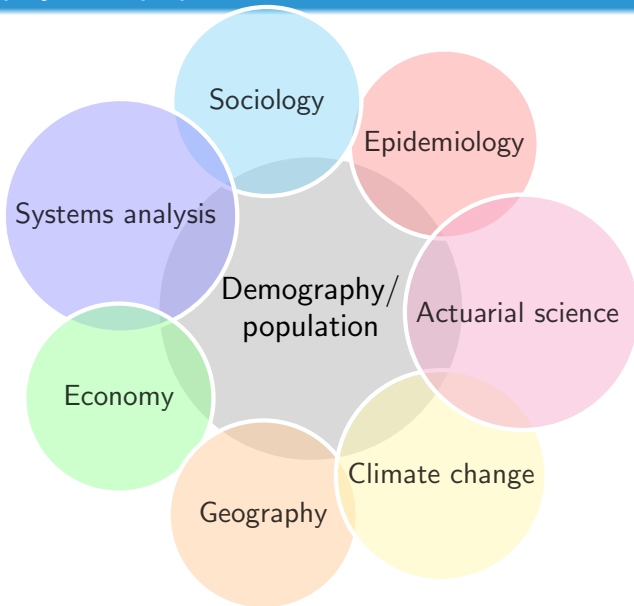
Multidimensional demography

The study of populations stratified by age, sex, and one or several attributes such as

- Region of residence
- Marital status
- Number of children
- Labour force participation
- Religion
- Educational attainment
- Health status

WIC Data Explorer Version 2.0.
2020, Medium scenario

Demography and population studies
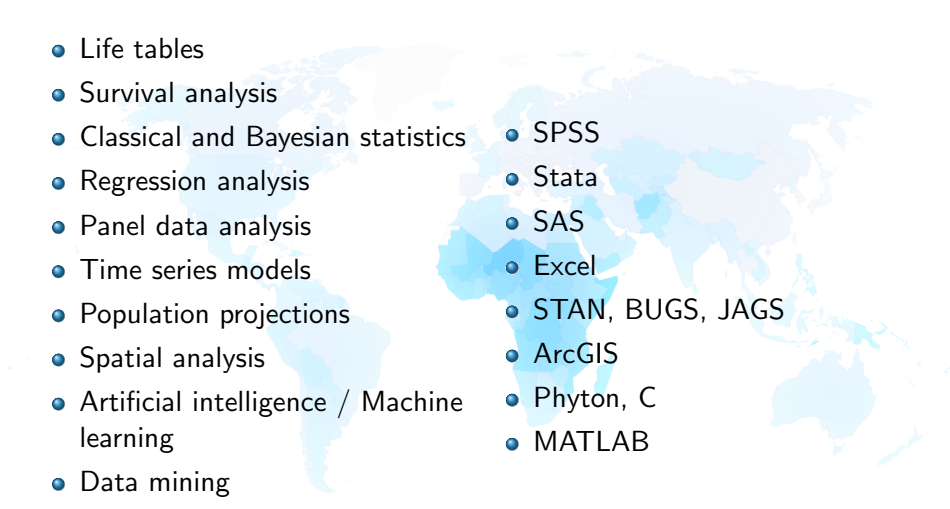
Data sources

Traditional data sources

- Censuses
- Population registers
- Demographic surveys (DHS, GGS)
- Vital registrations (birth and death records)
- International databases (UN data, Eurostat)

New and innovative data sources

- Satellite images (night time lights)
- Mobile phone call records
- Social media (Twitter, Facebook)
- Webpages (LinkedIn)

1.13 6.3

Methods and software

- Life tables
- Survival analysis
- Classical and Bayesian statistics
- Regression analysis
- Panel data analysis
- Time series models
- Population projections
- Spatial analysis
- Artificial intelligence / Machine learning
- Data mining
- Microsimulation

- SPSS
- Stata
- SAS
- Excel
- STAN, BUGS, JAGS
- ArcGIS
- Phyton, C
- MATLAB

Why R?

- R can be used in every step of demographic research: importing data, descriptive statistics, demographic and statistical analysis, visualization, writing up the results and publishing.
- More comprehensive than most of the other software.
- It is possible to use some of the other software such as (JAGS) via R.

Importing data → Analysis → Visualization → Presentation

Importing data

- Tablo, readr: .csv
- Excel, readxl: selecting a specific sheet or part of data
- Reading data from a website, httr, jasonlite, rvest, rtweet, etc.
- Other, haven, foreign: SPSS, Stata, SAS Minitab

Demographic databases I

wpp
United Nations World Population Prospects 1950-2100, population estimates, demographic rates, projections: wppExplorer, wpp2019, wpp2017, etc.

bayesPop
Uni. Washington & UN, Bayesian estimations and projections of WPP: bayesTFR, bayesLife, bayesDem

eurostat
Statistical Office of European Union database, data on many other areas in addition to demography

ipumsr
Read data from IPUMS extract into R along with metadata, variable labels, value labels etc. Spatial data and hierarchical data structures are also supported.

Demographic databases II

rdhs	Useful to import Demographic and Health Survey (DHS) data with API
HMD-HFDplus	mortality data for 40+ countries and fertility data for about 30 countries
US Census data	22 R packages available on CRAN including acs

rdhs — Useful to import Demographic and Health Survey (DHS) data with API

dhs.rates — birth and death rates from DHS surveys
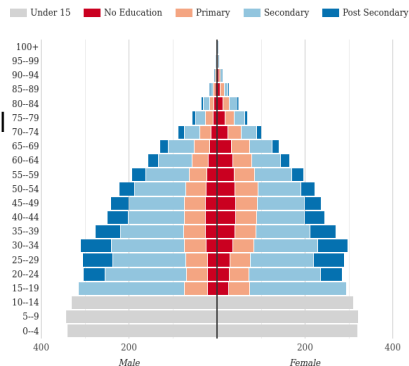
HMD-HFDplus — mortality data for 40+ countries and fertility data for about 30 countries

US Census data — 22 R packages available on CRAN including acs

wcde

- To download data from Wittgenstein Centre Human Capital Data Explorer

- Developed by Guy Abel, Claudia Reiter, Dilek Yildiz and Samir KC

- Available in CRAN

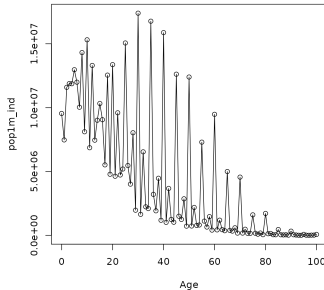- Please let us know about your comments and recommendations

WIC Data Explorer Version 2.0., medium scenario

Data cleaning and preparation

- Long and wide formats gather() and spread() functions
- Separating and uniting variables separate() and unite() functions
- Changing type of variables as.numeric(), as.character() etc. functions
- Selecting, filtering, ordering and summarizing variables select(), filter(), arrange(), group_by(), summarise() functions
- Analyzing survey data with survey package
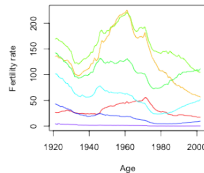
DemoTools

Aburto, Kashnitsky, Pascariu, Tim Riffe, 2021

- Tim Riffe, Sean Fennell, José Manuel Aburto, Ilya Kashnitsky, Marius Pascariu, Jorge Cimentada, Monica Alexander, Patrick Gerland
- Funded by UNPD and Gates Foundation
- Includes many functions that are used in demographic data evaluation, transformation and adjustment
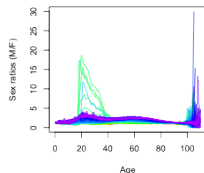- Such as checking age heaping, smoothing age groups, and Roger-Castro fitting for migration
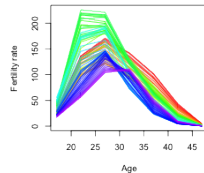
demography

- Rob J Hyndman, Heather Booth, Leonie Tickle, John Maindonald
- Life table calculations, Lee-Carter models, death and birth rates,
- Life expectancy, net migration estimation, visualization,
- Population projections,
- Read data from Human Mortality Database

demogR - popbio

demogR

- James Holland Jones
- Similar to other demography packages, latest update in 2018
- Evolutionary demography, ecology and biology
- Analysis of matrix population models, eigenvalues
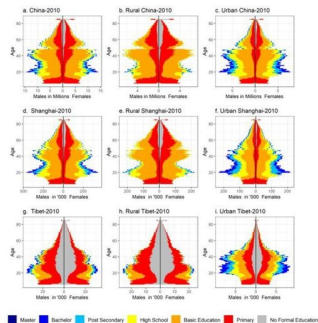- Allows using model life tables

popbio

- Chris Stubben, Brook Milligan, Patrick Nantel
- Matrix population models
- Caswell (2001) and Morris & Doak (2002) Matlab code translation
- Capture-recapture models

MSDem

- Multi-State (Dimensional) Demography
- Developed at IIASA by Marcus Wurzer, Samir KC ve Markus Speringer
- Population projections by age, sex, region, urban/rural and education breakdown

Age
Gender-
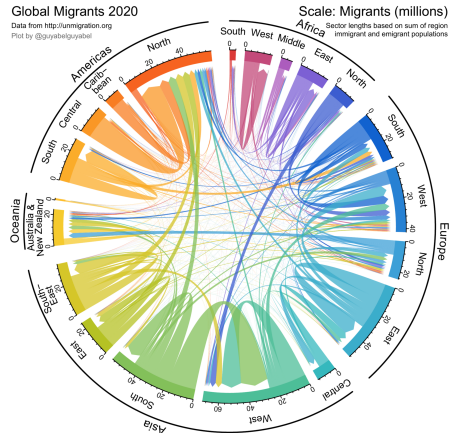**S**ubnational-
Rural/Urban
Education

KC, S., 2019

Visualization

- ggplot2
- geofacet
- circlize
- shiny
- Rmarkdown

Data Explorer   Graphic Explorer   About

Wittgenstein Centre
Human Capital Data Explorer

Explore, select and download data on past reconstructions and future projections of the global population by age, sex and education, published in Lutz, Goujon, KC, Stonawski, and Stilianakis (Eds.) (2018) ⧉.

Wittgenstein Centre
FOR DEMOGRAPHY AND
GLOBAL HUMAN CAPITAL

Selection   Data   Assumptions

1. Indicators

Indicator Type

Population & Human Capital Stocks▾

Indicator

Population Size (000's)   ▾

2. Geography

Country

Type or click for multiple choices

Region

Type or click for multiple choices

☐ Include countries of selected regions

3. Breakdown

Sex

Both

Age

All

☐ Include all age groups

4. Time Horizon

Scenario

Medium (SSP2)

Year

2015

☐ Include all times

View Data ⧉

⬇ Download

ÖAW   AUSTRIAN ACADEMY OF SCIENCES   International Institute for Applied Systems Analysis   IIASA www.iiasa.ac.at   ADRI   Asian Demographic Research Institute

Dilek Yıldız
Wittgenstein Centre (IIASA, VID/ÖAW, Uni.Vienna),
Vienna Institute of Demography/Austrian Academy of Sciences,
International Institute for Applied Systems Analysis

E-mail: yildiz@iiasa.ac.at
Twitter: dil_ yildiz