# Advanced Data Visualizations with Stata: Part III

Asjad Naqvi

Austrian Institute of Economic Research (WIFO)
Vienna University of Economics and Business (WU)
International Institute for Applied Systems Analysis (IIASA)

Stata UK Conference
8 September 2022

## Background

- Based in Vienna, Austria
- Some 20 years of research
- Numerous projects using Stata since 2003
  - Data management, workflow, analysis

- Why do all the dataviz stuff?
  - Explosion of COVID-19 related dataviz + work-from-home + curiosity
  - Huge positive response from the online community

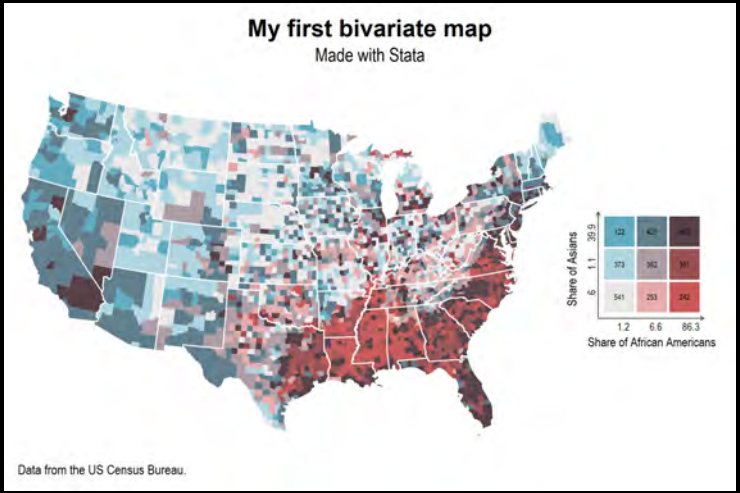
## Why we need more #StataViz?

- A lot of potential given the existing Stata/Mata structure
- Lots of exciting developments in terms of dataviz packages:
  - `palettes` (Jann), `spmap` (Pisati)
  - My own packages: bimap, joyplot, streamplot, clipgeo, delaunay, marimekko, arcplot, schemepack
  - Plus many more `#StataViz` packages by others are now out!

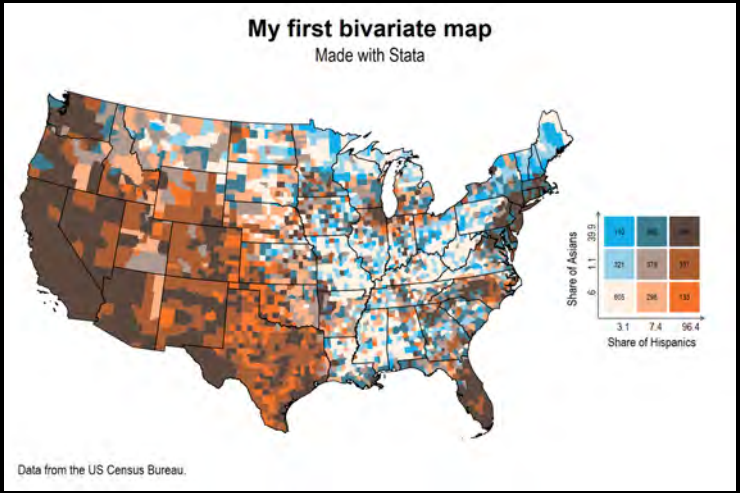
- What I will discuss today:
  - Maps
  - Polar plots
  - Plotting hierarchical data

- The ability to make maps in Stata has improved considerably, thanks to the `spmap` (Pisati) and the `palettes` (Jann) packages.
- A lot of additional work has been done in the past two years to further enhance mapping capabilities in Stata
- Some of this work is showcased below

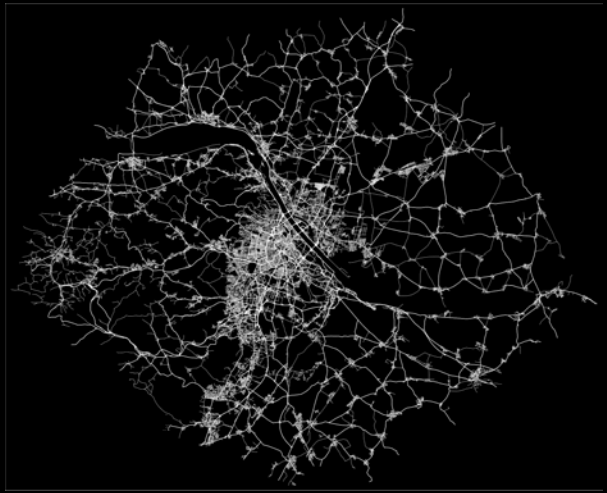
## Bi-variate maps (ssc install bimap, replace)



## Bi-variate maps (ssc install bimap, replace)



clipgeo (ssc install clipgeo, replace) - Line clipping



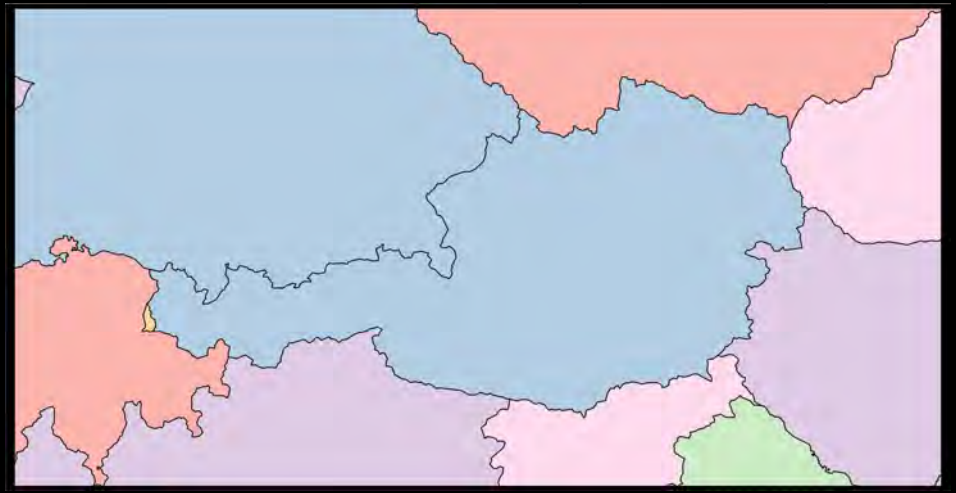
## clipgeo - Box clipping

## clipgeo v2 - Circle clipping

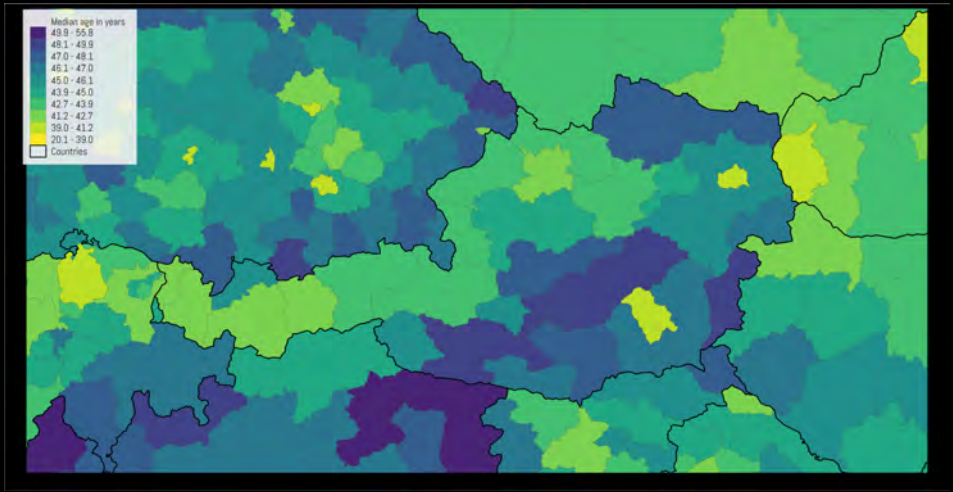
## clipgeo v2 - Polygon clipping and rotation

## clipgeo - Polygons

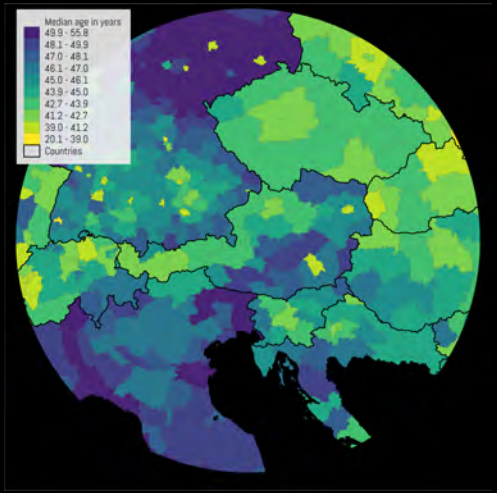
## clipgeo - Polygon box clipping



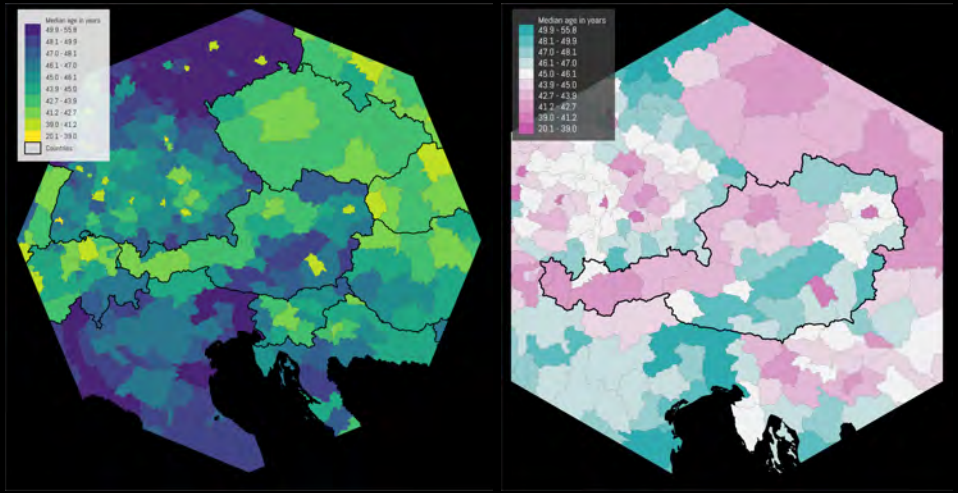
## clipgeo - Polygon box clipping + spmap + palettes



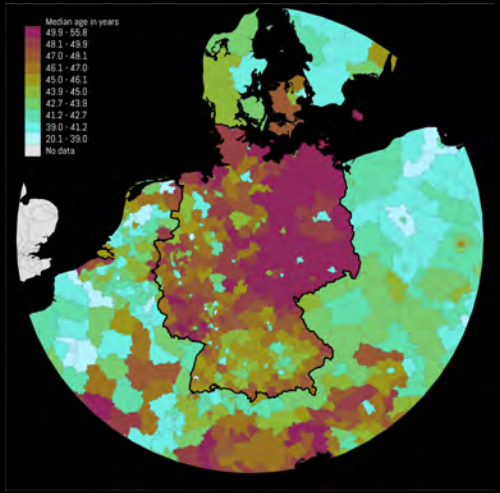
# clipgeo v2 - Polygon circle clipping

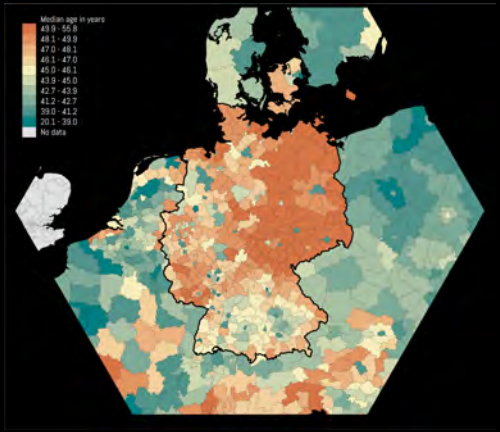


## clipgeo v2 - Polygon clipping variations



## clipgeo v2 - Polygon clipping variations



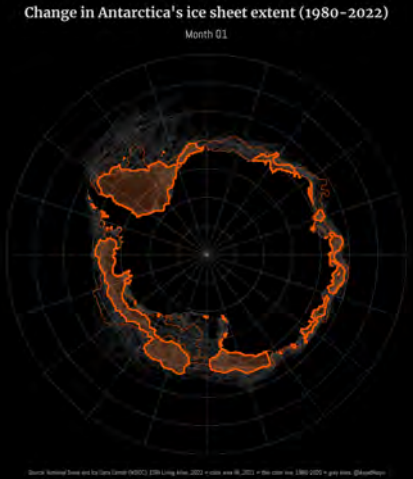
## clipgeo v2

The clipgeo v2 makes several enhancements to the package:

- Much faster plotting
- We can now clip on any regular polygon (polygons with equal sides)
- Circular clipping
- Polygon rotation

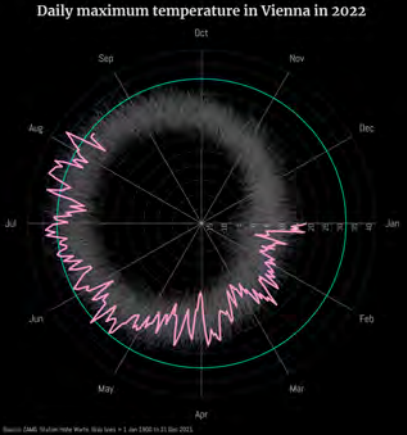
→ Demo

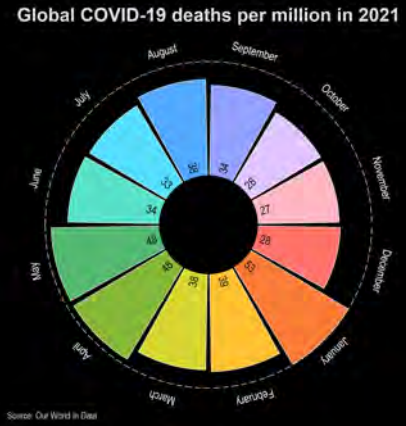
# Polar plots


Change in Antarctica's ice sheet extent (1980–2022)
Month 01

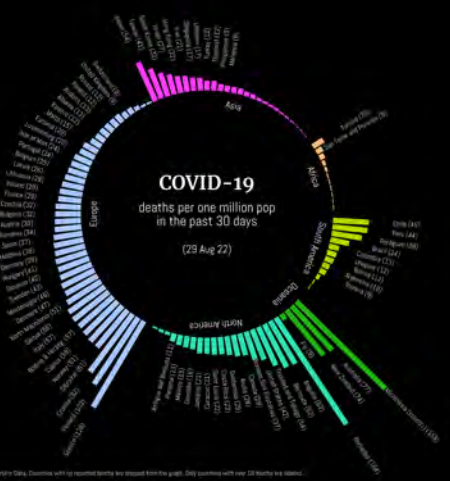
Change in Antarctica's ice sheet extent (1980–2022)
Month 02

# Polar plots


Daily maximum temperature in Vienna in 2022
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep

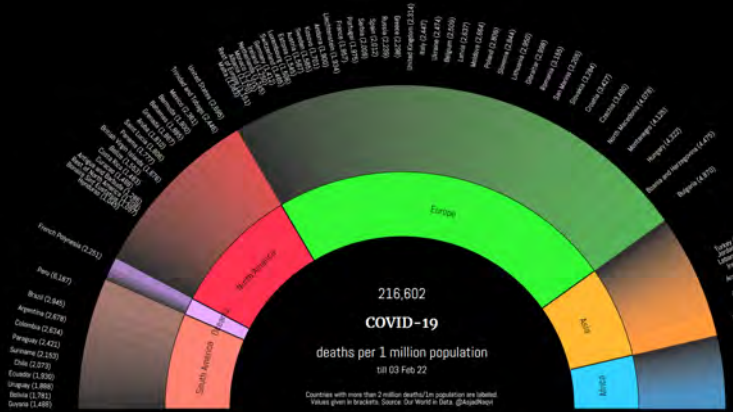

Global COVID-19 deaths per million in 2021
August
September
October
November
December
January
February
March
April
May
June
July
Source: Our World in Data

# Polar plots

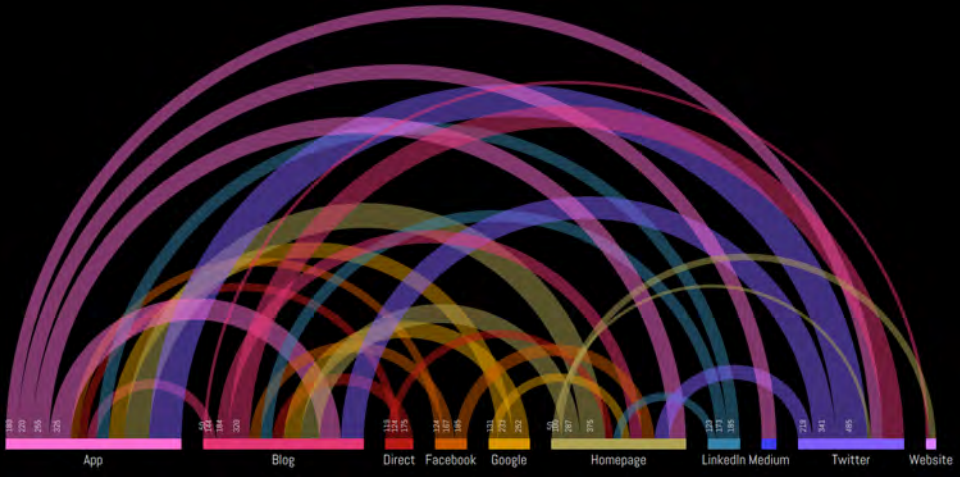

COVID-19
deaths per one million pop
in the past 30 days
(29 Aug 22)
Asia
Africa
South America
Oceania
North America
Europe

# Polar plots

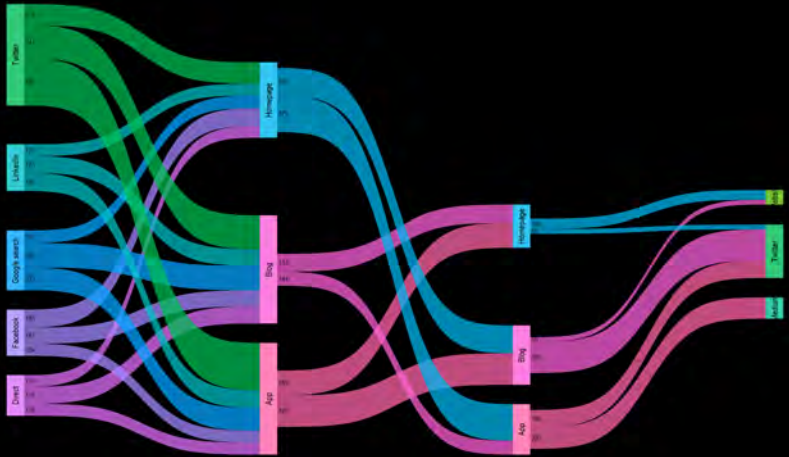
# Half sunburst plot



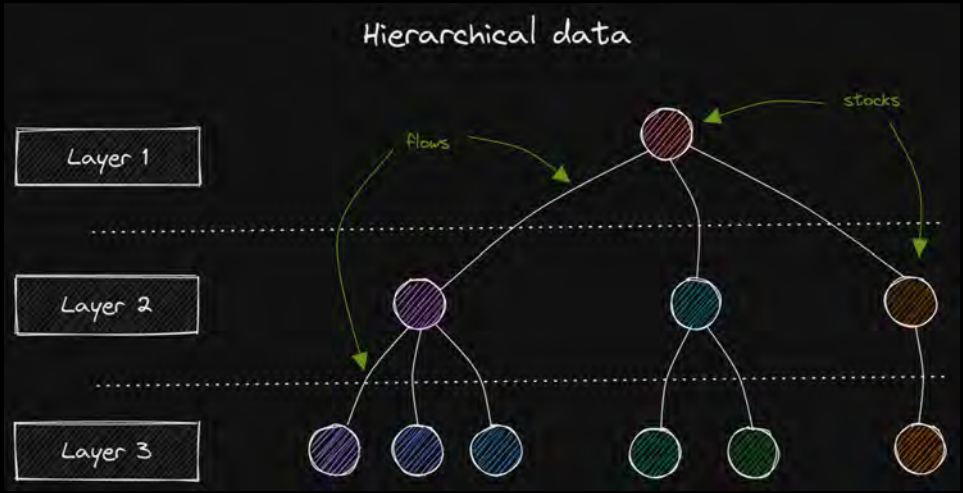
## Arc plots (now available as the arcplot package)



# Hierarchical flows: Sankey



## Data hierarchies



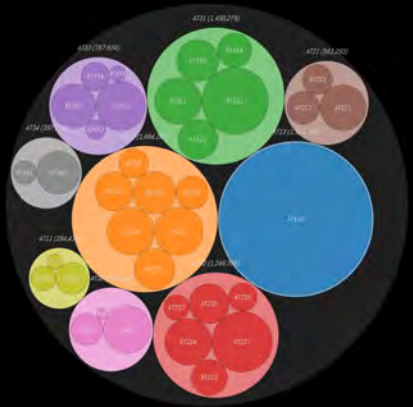
## Hierarchical stocks: `circlepack`

- Plot nested lists as circles within circles.
- Circle packing requires a packing algorithm that maximize the area of a starting circle, such that no circles overlap.
- The circlepack package uses the "circlify" algorithm
- Like squarify, "circlify" is an iterative algorithm, that starts with the largest circle and iteratively adds other circles, until the best fit is achieved.

→ Demo

# Hierarchical stocks: `circlepack`

# Hierarchical stocks: `circlepack`



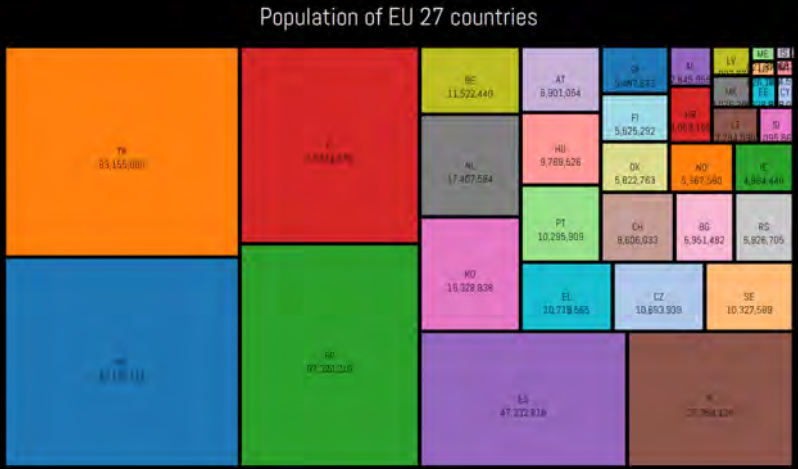
# Hierarchical stocks: `circlepack`

## Hierarchical stocks: `treemap`

- Treemaps plot nested lists as rectangles within rectangles.
- Treemaps requires a tiling algorithm.
- The treemap package uses the "squarify" algorithm, the industry standard (OEC, OWID, D3).
- "squarify" optimizes placement of tiles to approach a desired aspect ratio.
- Tiles are sorted by largest-to-smallest size, and iteratively placed by either "slicing" or "dicing" a starting rectangle, until all tiles are close to the desired aspect ratio.

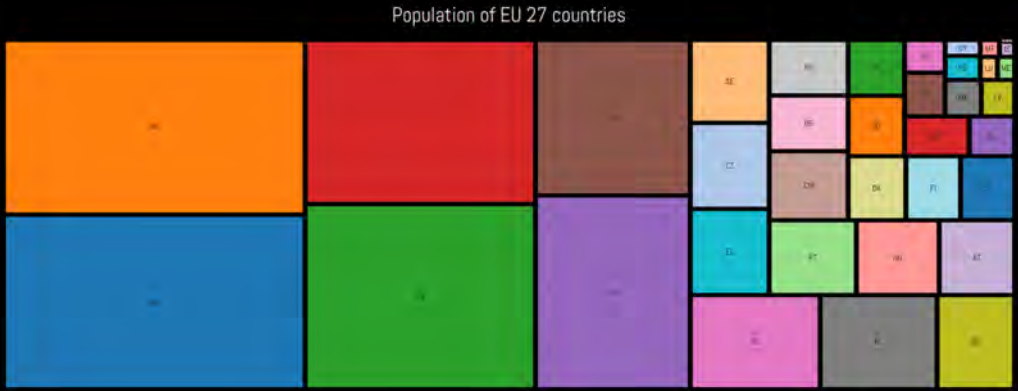
→ Demo

# Hierarchical stocks: `treemap`



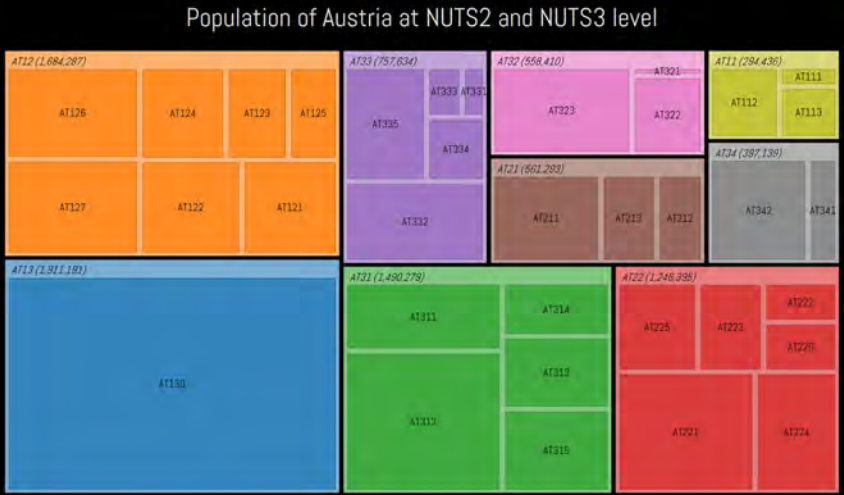
# Hierarchical stocks: `treemap`

# Hierarchical stocks: `treemap`

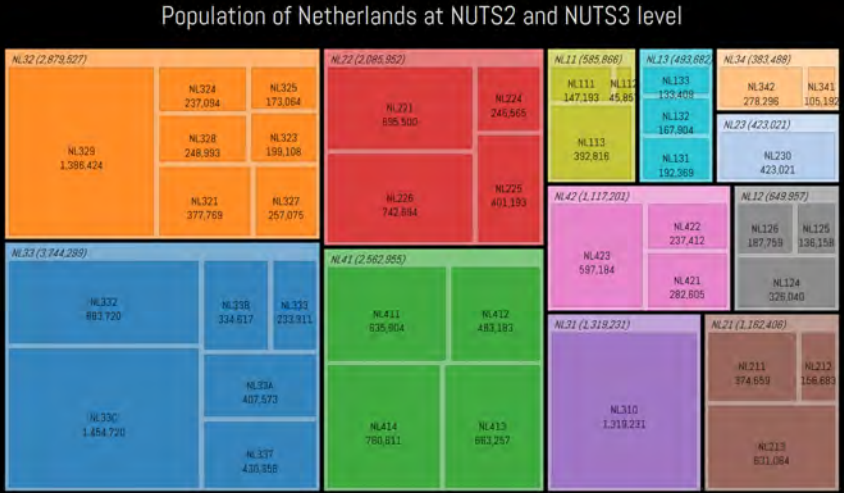


# Hierarchical stocks: `treemap`

# Hierarchical stocks: `treemap`



# Hierarchical stocks: `treemap`



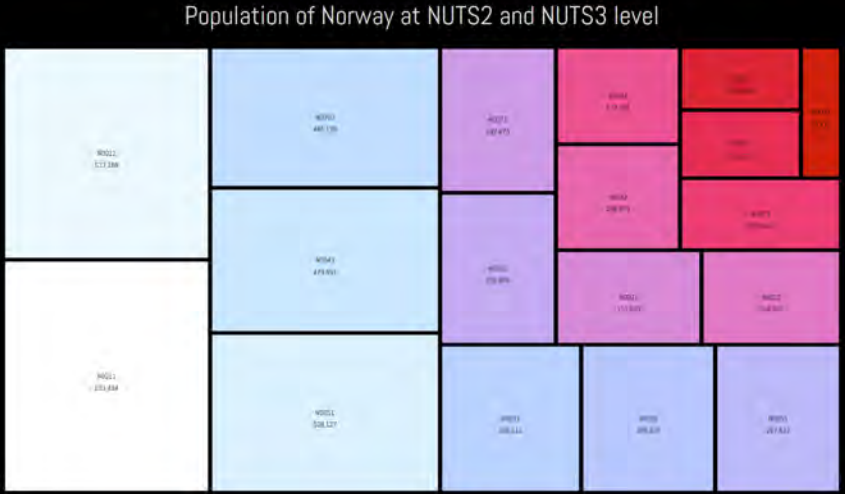
# Hierarchical stocks: `treemap`

Population of Netherlands at NUTS2 and NUTS3 level

Population of Netherlands at NUTS1-NUTS3 level

# Hierarchical stocks: `treemap`

# Hierarchical stocks: `treemap`



## What can be improved

Some routines are very time and memory expensive:

- Exporting data back to Stata from Mata
- Plotting by colors
- Making shapes, for example, circles should be easier

Some optimization requests

- Allow point, line, area properties to be modified more easily
  - Allow angles, sizes, colors, clock positions, etc. to be read from variables (like mlabels)
  - Line weights (like marker weights)
  - Color scaling (starting and ending color)
- Allow adding custom markers (or increase marker pool)
- Allow colored text in graphs
- Ability to read images (pixel data)
- Ability to add images in figures

## More info

More Stata viz:

- The Stata Guide on Medium
- The Stata Gallery on Medium (anyone can contribute!)
- #30DayMapChallenge 2021
- The Stata Gallery

Connect with me:


- @ asjadnaqvi@gmail.com
- github.com/asjadnaqvi
- @AsjadNaqvi
- AsjadNaqvi