

2024 UK Stata conference

Fancy graphics: Small multiples carpentry

Philippe Van Kerm

London School of Economics – September 13–14, 2024

“Small multiples” (trellis plots)?

Me:

“What are ‘small multiples’?”

ChatGPT:

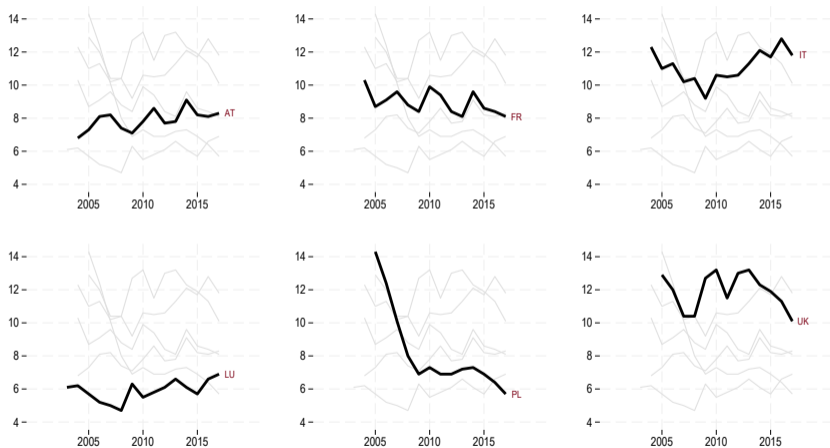
*Small multiples refer to a series of similar visualizations, such as charts or graphs, that use the same scale and axes, allowing them to be **easily compared side by side**. Each individual visualization **shows a different subset of data**, but because they are uniform in size and layout, they highlight variations and patterns across the multiple panels without the need to switch between different graph types or designs.*

Basic approaches

Some more flexible (fancier?) approach

Example 1

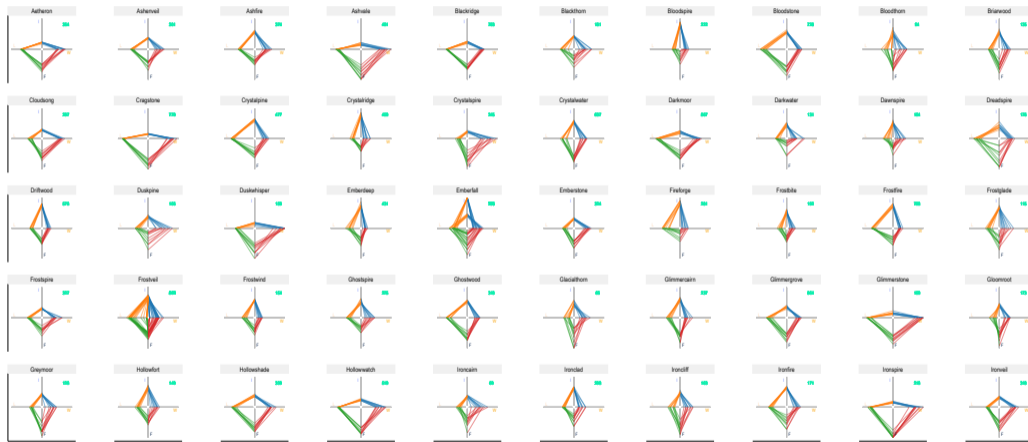
Proportion of households with low work intensity by country



Example 1: graph combine

```
separate lwi , gen(tmp) by(GEO)
gen pos = 2018
foreach c of local ctries {
    tw ///
    (line tmp? year , lcolor(gs14 ..) lpattern(solid ..) sort ) ///
    (line lwi year if GEO=="'c'" , lcolor(black) lwidth(thick)
        lpattern(solid)) ///
    (scatter lwi pos if GEO=="'c'" & year==2017, ms(i) mlab(GEO)
        mlabpos(0)) ///
    , xtitle("") yscale(noline) xscale(noline) xlabel(2005(5)2015)
        saving(g-'c' , replace) legend(off)
    loc gnames 'gnames' g-'c'.gph
}
graph combine 'gnames' , xcommon ycommon title("Proportion ...")
```

Example 2

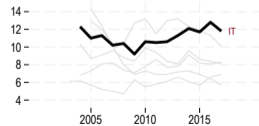
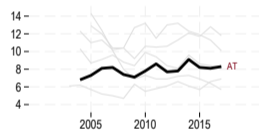
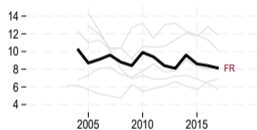
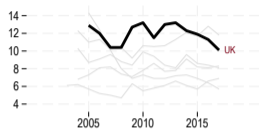


Example 2: `twoway ... , by(...)`

```
tw ///
  (pcspike minpos_work_y minpos_work_x  maxpos_work_y
    maxpos_work_x  , lcol(gs12)) ///
  (pcspike minpos_income_y minpos_income_x  maxpos_income_y
    maxpos_income_x  , lcol(gs12) ) ///
  ...
  (scatter ypos xpos , mlabel(nstud) ms(i) mlabpos(0)) ///
  (scatteri '=-0.1*'ray'' '=0.9*'ray'' "W" , ms(i) mlabpos(0)) ///
  ...
  , legend(off)  aspectratio(1) yscale(off) xscale(off) ylabel(none
    ) xlabel(none) ///
  by(cidbis , cols(10) legend(off)  noedge noiyaxe noixaxe
    noiylabel noixlabel note("") plotregion(style(none))
    graphregion(style(none)))
```

Example 3

Proportion of households with low work intensity by country

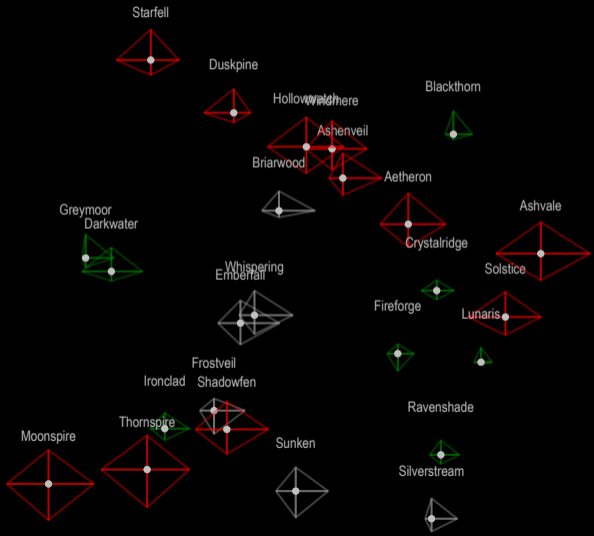


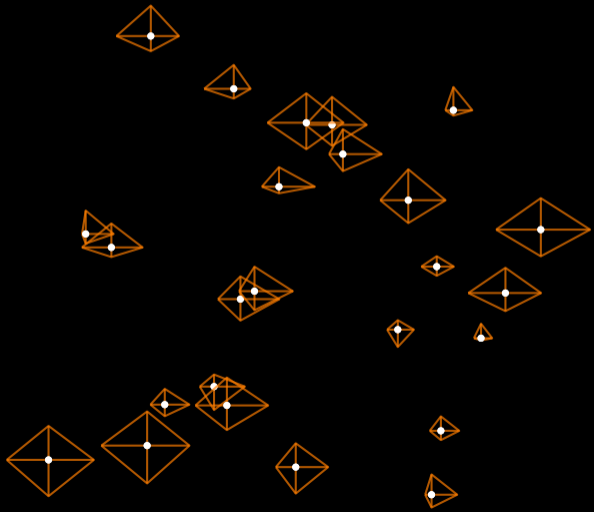
Example 3: graph combine – again

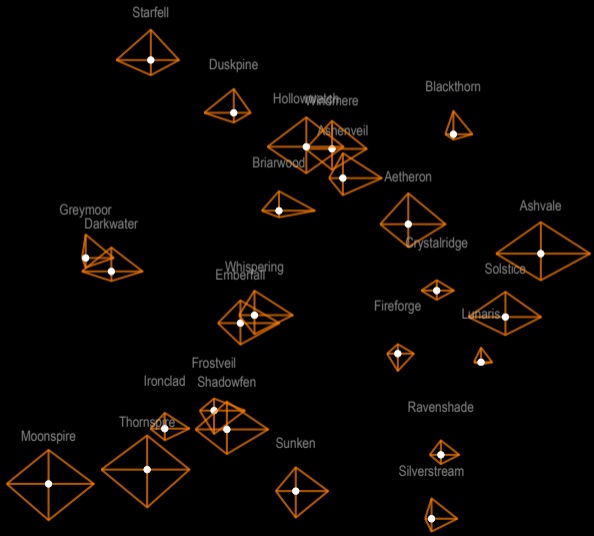
```
graph combine g-UK.gph g-LU.gph g-PL.gph g-FR.gph g-AT.gph g-
  IT.gph , xcommon ycommon ///
  title("Proportion of households with low work intensity by
    country") ///
  holes(4 7 8) rows(3)
```

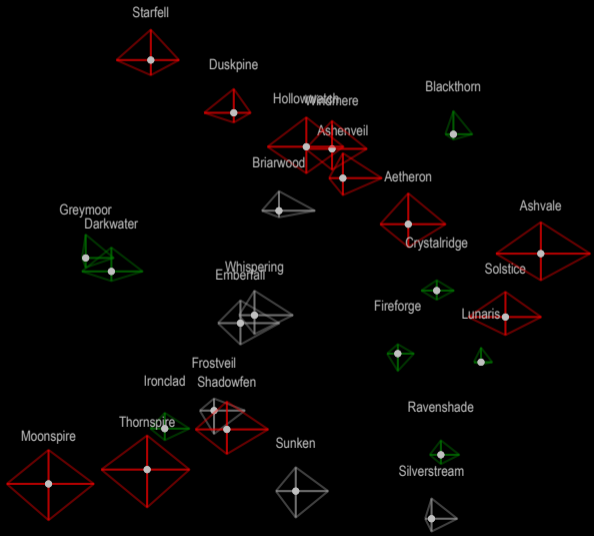
Basic approaches

Some more flexible (fancier?) approach









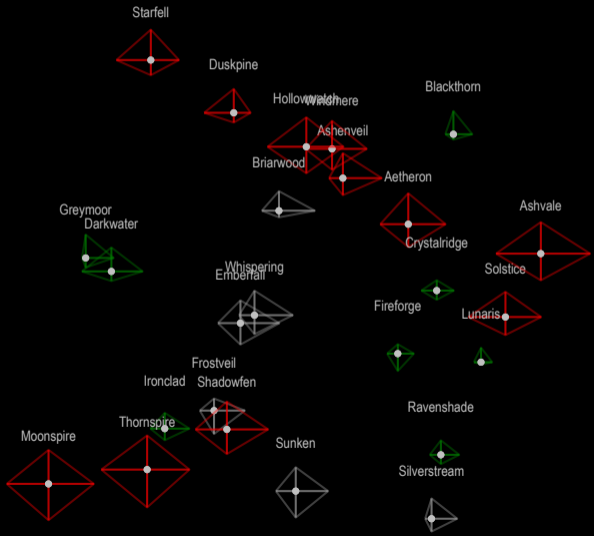
A little helper for the fancier approach: `sketchtoplot`


1. use Stata's immediate graph command syntax to *sketch* one 'panel' in a (-1,-1)-(1,1) box
2. populate, position and scale multiples *sketches* from the data
3. pass the string to `graph twoway` for producing the graph

```
local mysketch ///
  (scatteri y{0} x{0} , mcol(gs15%50) )
sketchtoplot 'mysketch' , xvarshift(_CX2) yvarshift(_CY2) global(
  plot)
graph tw $plot , $gropts
```


A little helper for the fancier approach: sketchtoplot

```
local mysketch ///  
  (pci ///  
    y{0} x{0}   y{yp_v1} x{xp_v1} ///  
    y{0} x{0}   y{yp_v2} x{xp_v2} ///  
    y{0} x{0}   y{yp_v3} x{xp_v3} ///  
    y{0} x{0}   y{yp_v4} x{xp_v4} , color(z{color}%70)) ///  
  (scatteri y{yp_v1} x{xp_v1} y{yp_v2} x{xp_v2} y{yp_v3} x{xp_v3} y{yp_v4} x{xp_v4} y{yp_v1} x{xp_v1} , recast(line) lcolor(z{color}%40)) ///  
  (scatteri y{.9} x{0} "z{name}", ms(i) mlabpos(12) mlabc(gs12))///  
  (scatteri y{0} x{0}, msym(o) mc(gs12))  
sketchtoplot 'mysketch' , global(plot) xvarshift(_CX2) yvarshift(_CY2) scale(5000)  
graph tw $plot , $gropts
```



 `sketchtoplot` – An immediate graph command parser and small multiple generator

```
sketchtoplot string [if][in][ , scale(real) xvarshift(varname)  
yvarshift(varname) global(string) ... ]
```

`sketchtoplot` parses *string* for *x{}*, *y{}* and *z{}* content. If content is a variable name, it populates it with data values. *x{}* and *y{}* content are then shifted by the optional 'shift' arguments. This is repeated for all observations selected by the *if/in* clauses and collated in the macro `global()` or `local()`.

Closing

- Small multiples (trellis) plots are excellent and visualisation tools – at least I find
- Simple widgets like those presented here may help (a lot) in producing custom graphs (esp. ‘data-driven’ parameters)
 - `sketchtoplot` has more potential uses than shown here
- Ben Jann’s `geoplot` provides functionalities similar to what I described for overlaying plots located ‘on shapefiles’
- Also, don’t miss Asjad Naqvi’s series of articles on Medium for great examples and tutorials <https://asjadnaqvi.medium.com/>
- Yes, Stata graphics capabilities are great (and flexible)!



