MRC Clinical Trials Unit

Smarter studies
Global impact
Better health

UCL

Manipulation of SVG graphs with Stata

Tim Morris, MRC Clinical Trials Unit at UCL

Robert Grant, BayesCamp

16 Nov 2017

Best feature in Stata 14?

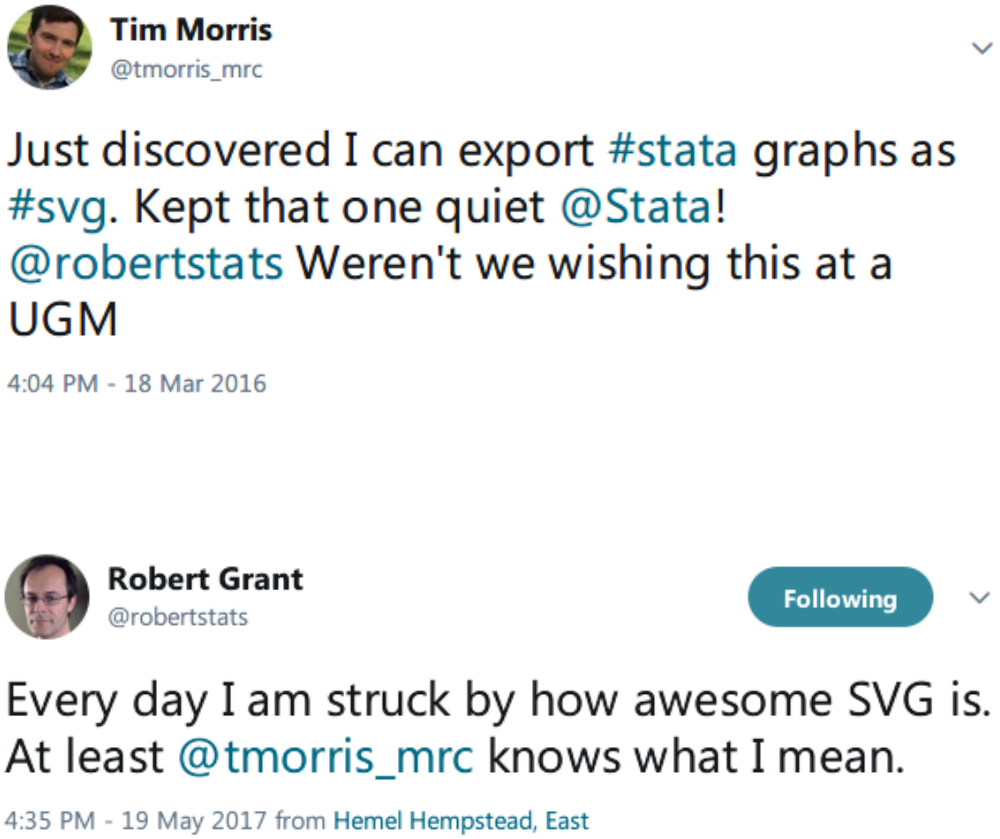
On upgrading Stata, one of the things I do first is to change bits with `graph set`. For example:

```
. graph set print logo off
```

With version 14, I submitted `graph set` and spotted something unfamiliar…

```
-> graph set svg

   svg                current
   setting            default          choices
   ------------------------------------------------
   baselineshift      off              on or off
   ignorefont         off              on or off
   nbsp               on               on or off
   clipstroke         on               on or off
   scalestrokewidth   off              on or off
   fontface           Open Sans        font name
   fontfacesans       Helvetica        font name
   fontfaceserif      Times            font name
   fontfacemono       Courier          font name
   fontfacesymbol     Symbol           font name
   ------------------------------------------------

   To change setting, type "graph set svg setting choice"
```

Best feature in Stata 14?



Tim Morris
@tmorris_mrc

Just discovered I can export #stata graphs as #svg. Kept that one quiet @Stata! @robertstats Weren't we wishing this at a UGM

4:04 PM - 18 Mar 2016

Robert Grant
@robertstats

Following

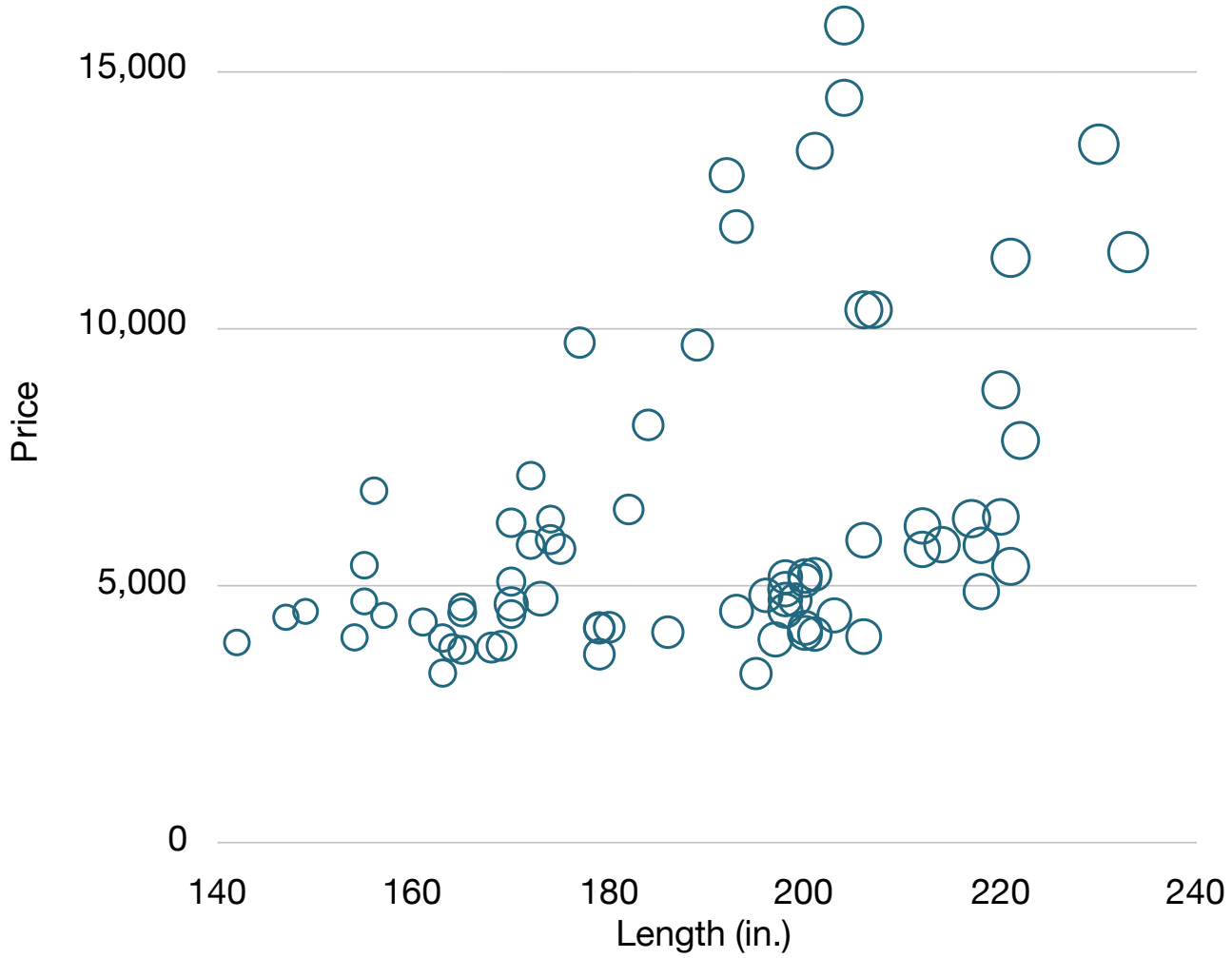
Every day I am struck by how awesome SVG is. At least @tmorris_mrc knows what I mean.

4:35 PM - 19 May 2017 from Hemel Hempstead, East

Creating an SVG file

```
. sysuse auto
. twoway scatter price length [pw=weight]
. graph export mygraph.svg, replace
```

Result (using mrc graph scheme)

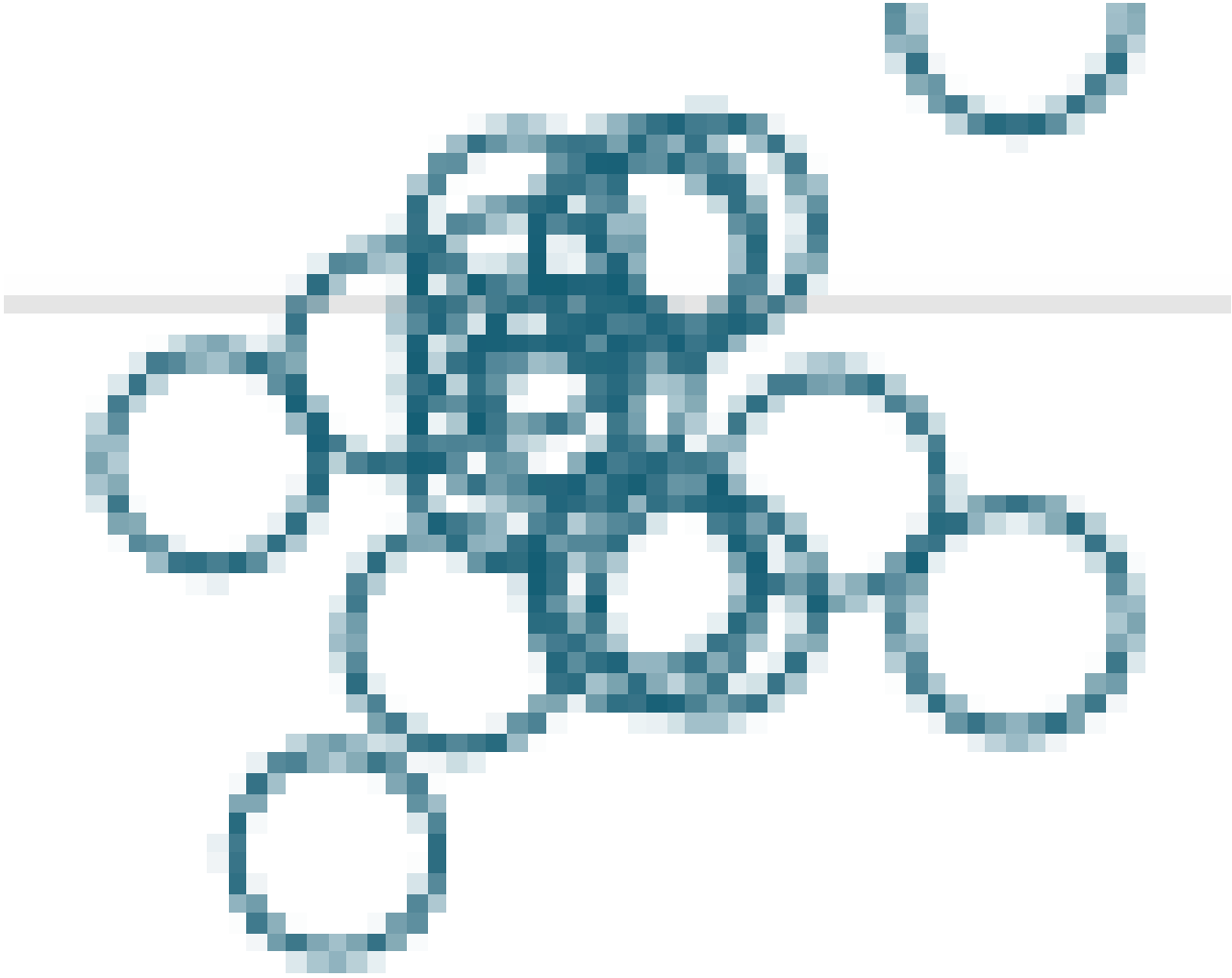


What is *SVG*?

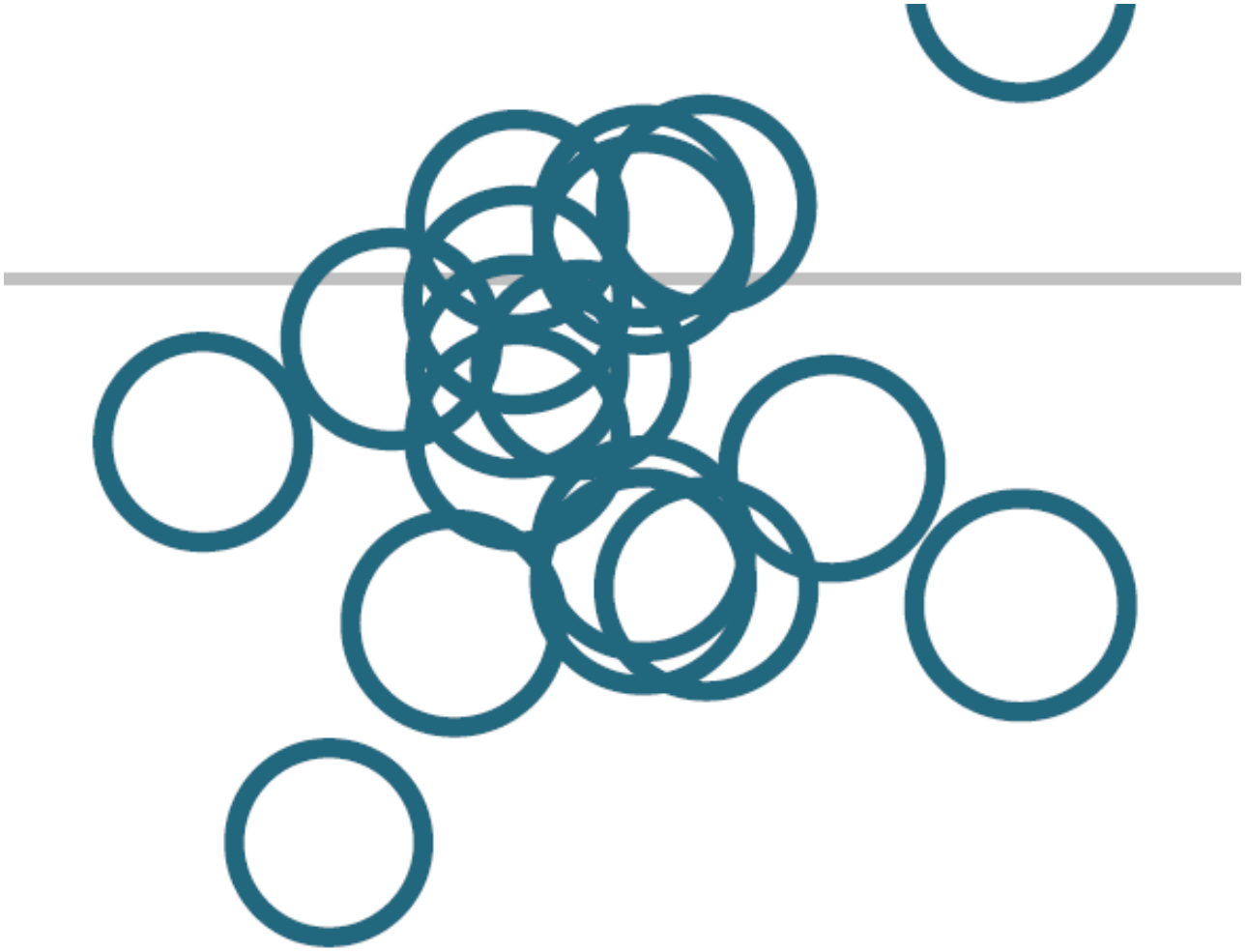
Scalable
Vector
Graphics

So
Very
Good

Raster graphics (23kb .png)



Vector graphics (10kb .svg)



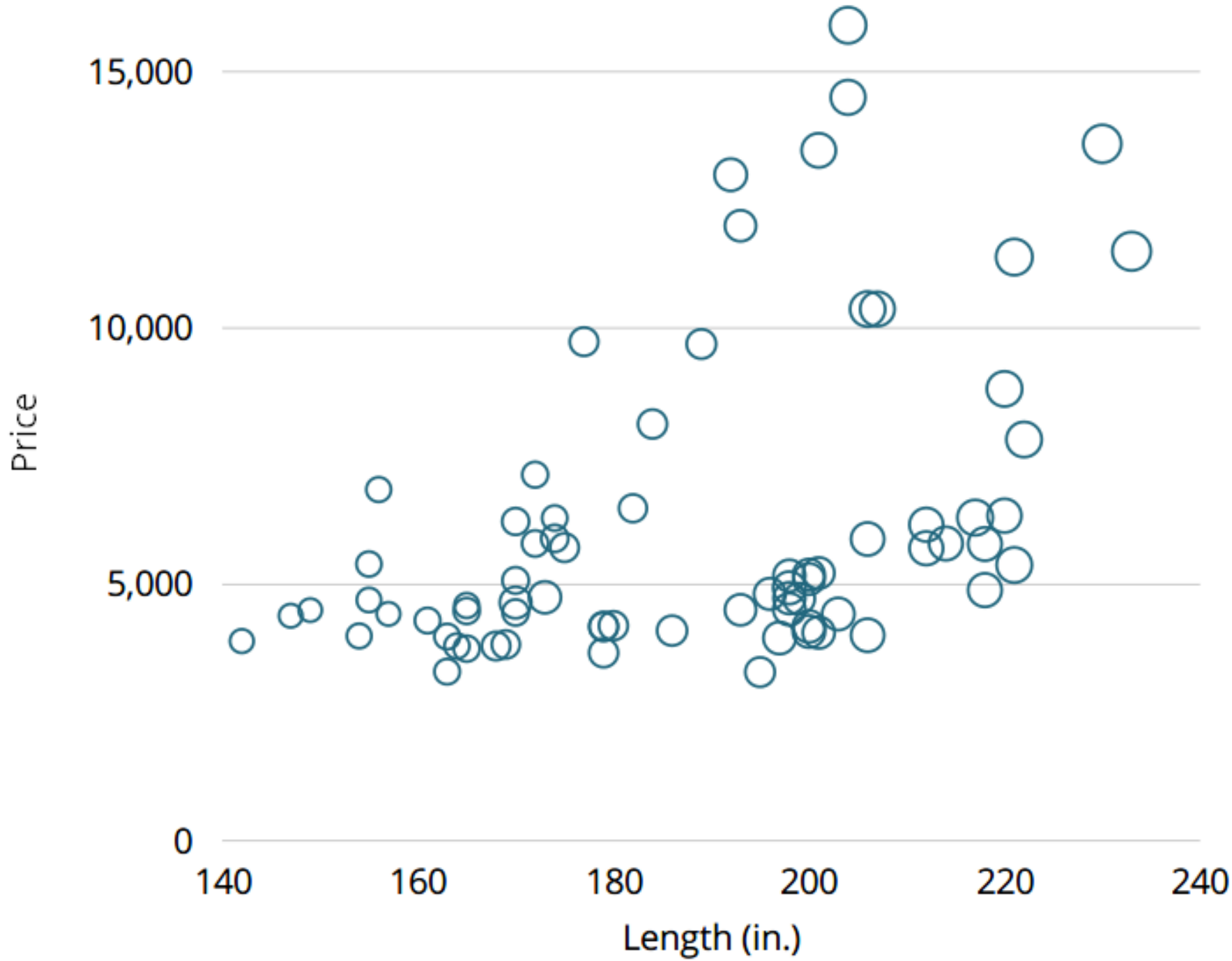
Read it

Read it

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- This is a Stata 15.0 generated SVG file (http://www.stata.com) -->

<svg version="1.1" width="5.00in" height="4.00in" viewBox="0 0 3600 2880" xmlns="http://www.w3.org/2000/svg
    <desc>Stata Graph - Graph</desc>
    <rect x="0" y="0" width="3600" height="2880" style="fill:#FFFFFF;stroke:none"/>
    <line x1="685.46" y1="2453.29" x2="3473.32" y2="2453.29" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="1721.93" x2="3473.32" y2="1721.93" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="990.56" x2="3473.32" y2="990.56" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="259.20" x2="3473.32" y2="259.20" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <circle cx="1967.85" cy="1853.78" r="42.64" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1605.38" cy="1758.71" r="45.67" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1465.99" cy="1897.65" r="40.50" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2246.63" cy="1748.93" r="45.00" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2971.46" cy="1308.49" r="50.40" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2859.86" cy="1606.72" r="47.81" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1521.79" cy="1802.03" r="37.24" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2358.11" cy="1694.36" r="45.11" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
```

Edit it

Let’s take a look at the file we created: auto.svg

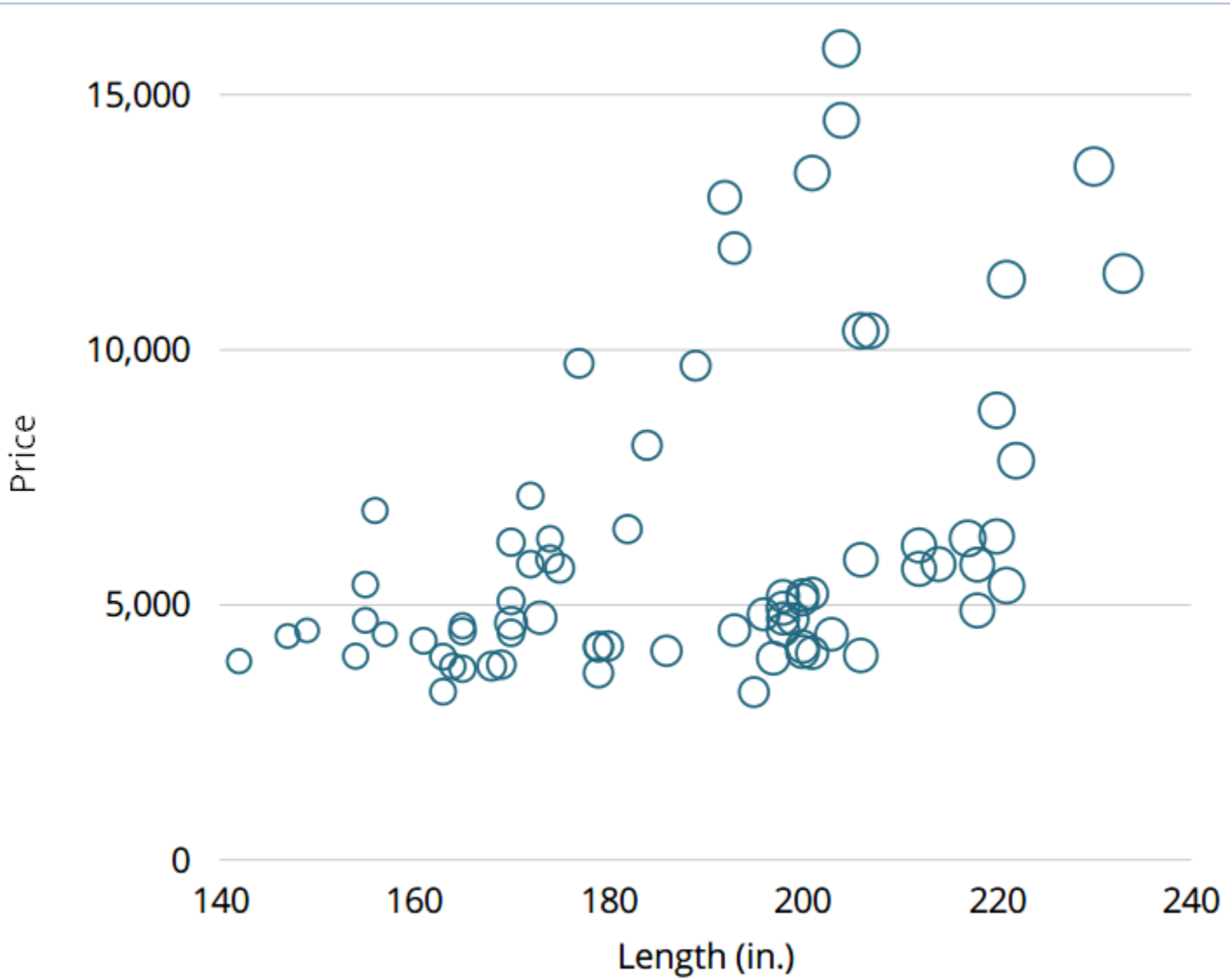

15,000
10,000
5,000
0
Price
140
160
180
200
220
240
Length (in.)

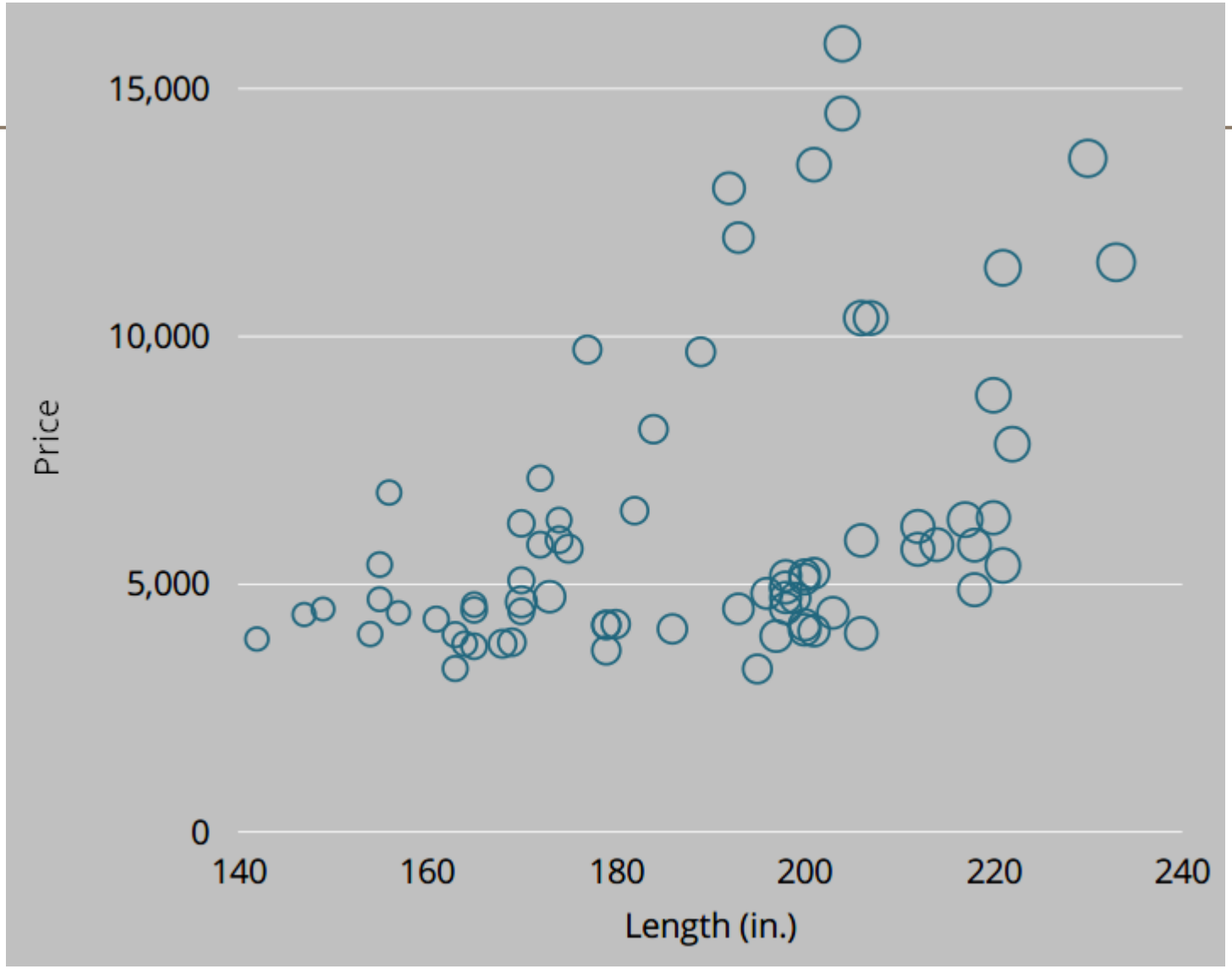
Edit it (from)

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- This is a Stata 15.0 generated SVG file (http://www.stata.com) -->

<svg version="1.1" width="5.00in" height="4.00in" viewBox="0 0 3600 2880" xmlns="http://www.w3.org/2000/svg
    <desc>Stata Graph - Graph</desc>
    <rect x="0" y="0" width="3600" height="2880" style="fill:#FFFFFF;stroke:none"/>
    <line x1="685.46" y1="2453.29" x2="3473.32" y2="2453.29" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="1721.93" x2="3473.32" y2="1721.93" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="990.56" x2="3473.32" y2="990.56" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <line x1="685.46" y1="259.20" x2="3473.32" y2="259.20" style="stroke:#C0C0C0;stroke-width:4.32"/>
    <circle cx="1967.85" cy="1853.78" r="42.64" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1605.38" cy="1758.71" r="45.67" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1465.99" cy="1897.65" r="40.50" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2246.63" cy="1748.93" r="45.00" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2971.46" cy="1308.49" r="50.40" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2859.86" cy="1606.72" r="47.81" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1521.79" cy="1802.03" r="37.24" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2358.11" cy="1694.36" r="45.11" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
```

Edit it (to)

```
<?xml version="1.0" encoding="UTF-8" standalone="no"?>
<!-- This is a Stata 15.0 generated SVG file (http://www.stata.com) -->

<svg version="1.1" width="5.00in" height="4.00in" viewBox="0 0 3600 2880" xmlns="http://www.w3.org/2000/svg
    <desc>Stata Graph - Graph</desc>
    <rect x="0" y="0" width="3600" height="2880" style="fill:#C0C0C0;stroke:none"/>
    <line x1="685.46" y1="2453.29" x2="3473.32" y2="2453.29" style="stroke:#FFFFFF;stroke-width:4.32"/>
    <line x1="685.46" y1="1721.93" x2="3473.32" y2="1721.93" style="stroke:#FFFFFF;stroke-width:4.32"/>
    <line x1="685.46" y1="990.56" x2="3473.32" y2="990.56" style="stroke:#FFFFFF;stroke-width:4.32"/>
    <line x1="685.46" y1="259.20" x2="3473.32" y2="259.20" style="stroke:#FFFFFF;stroke-width:4.32"/>
    <circle cx="1967.85" cy="1853.78" r="42.64" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1605.38" cy="1758.71" r="45.67" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1465.99" cy="1897.65" r="40.50" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2246.63" cy="1748.93" r="45.00" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2971.46" cy="1308.49" r="50.40" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2859.86" cy="1606.72" r="47.81" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="1521.79" cy="1802.03" r="37.24" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
    <circle cx="2358.11" cy="1694.36" r="45.11" style="fill:none;stroke:#21677E;stroke-width:8.64"/>
```


15,000
10,000
5,000
0
Price
140
160
180
200
220
240
Length (in.)


15,000
10,000
5,000
0
Price
140
160
180
200
220
240
Length (in.)

Edit it

You don’t have to edit the text as I showed

You can edit using software like Inkscape or Illustrator

Manipulating SVG: principle

- Produce an SVG file with Stata (14+)
- Open it up and edit to make it even better
- Receive praise and accolades from peers

What can I manipulate?

A lot, actually:

- Add transparency to graph elements (ooh)
- Embed images behind your graph (aah)
- Create shapes Stata doesn’t allow (phwoar)
- Introduce interactive elements

Why use Stata at all?

Stata creates great SVG files. They are well organised (consistent order) and small.

See: **SVG from stats software: the good, the bad and the ugly**

goo.gl/q6no1L


SVG permits ‘opacity=#’

‘Opacity’

opaque

adjective

not able to be seen through; not transparent.

translucent

adjective

(of a substance) allowing light, but not detailed shapes, to pass through; semi-transparent.

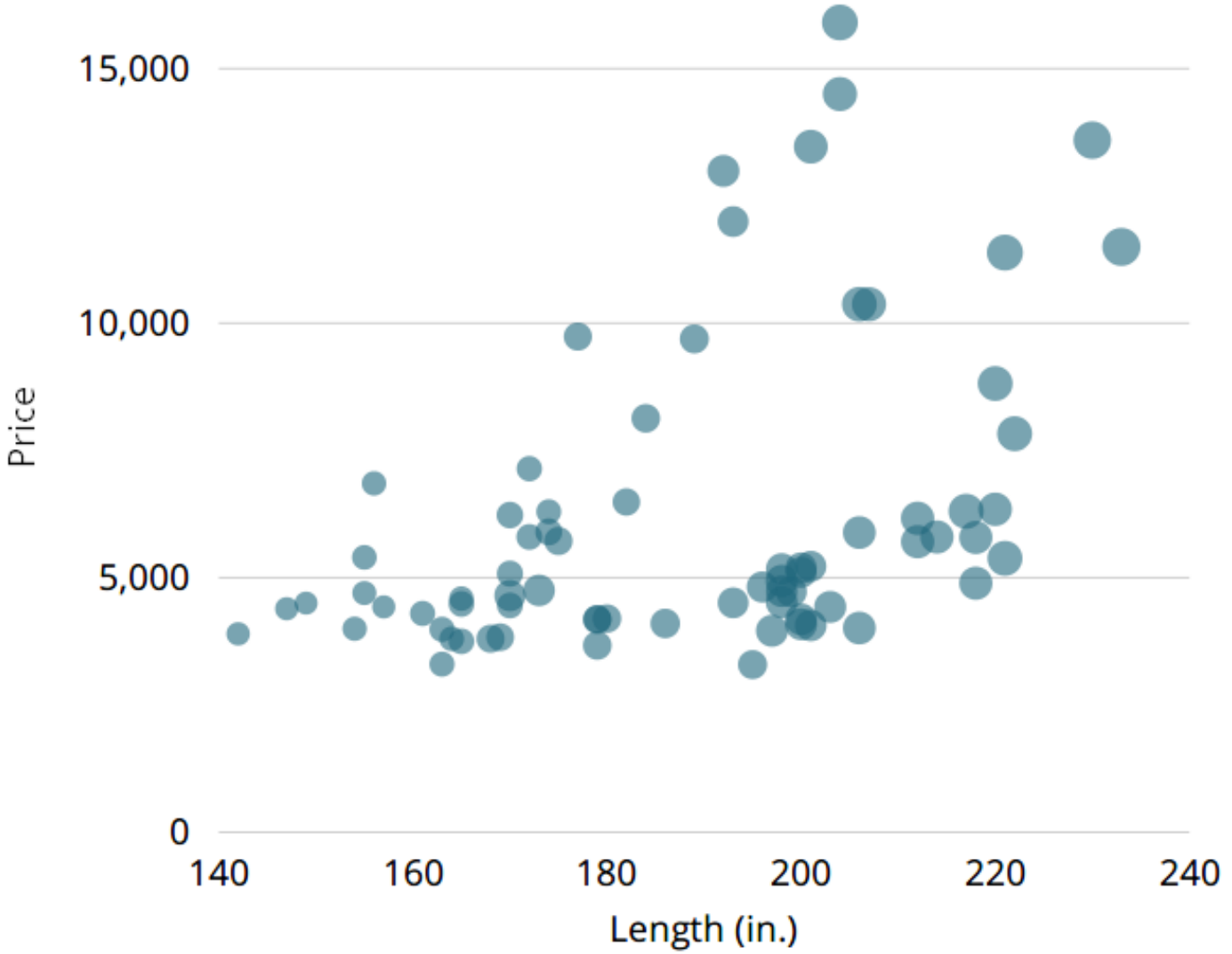
Translucency

- Add to the `style` of objects
- E.g `stroke-opacity` or `fill-opacity`
- All done

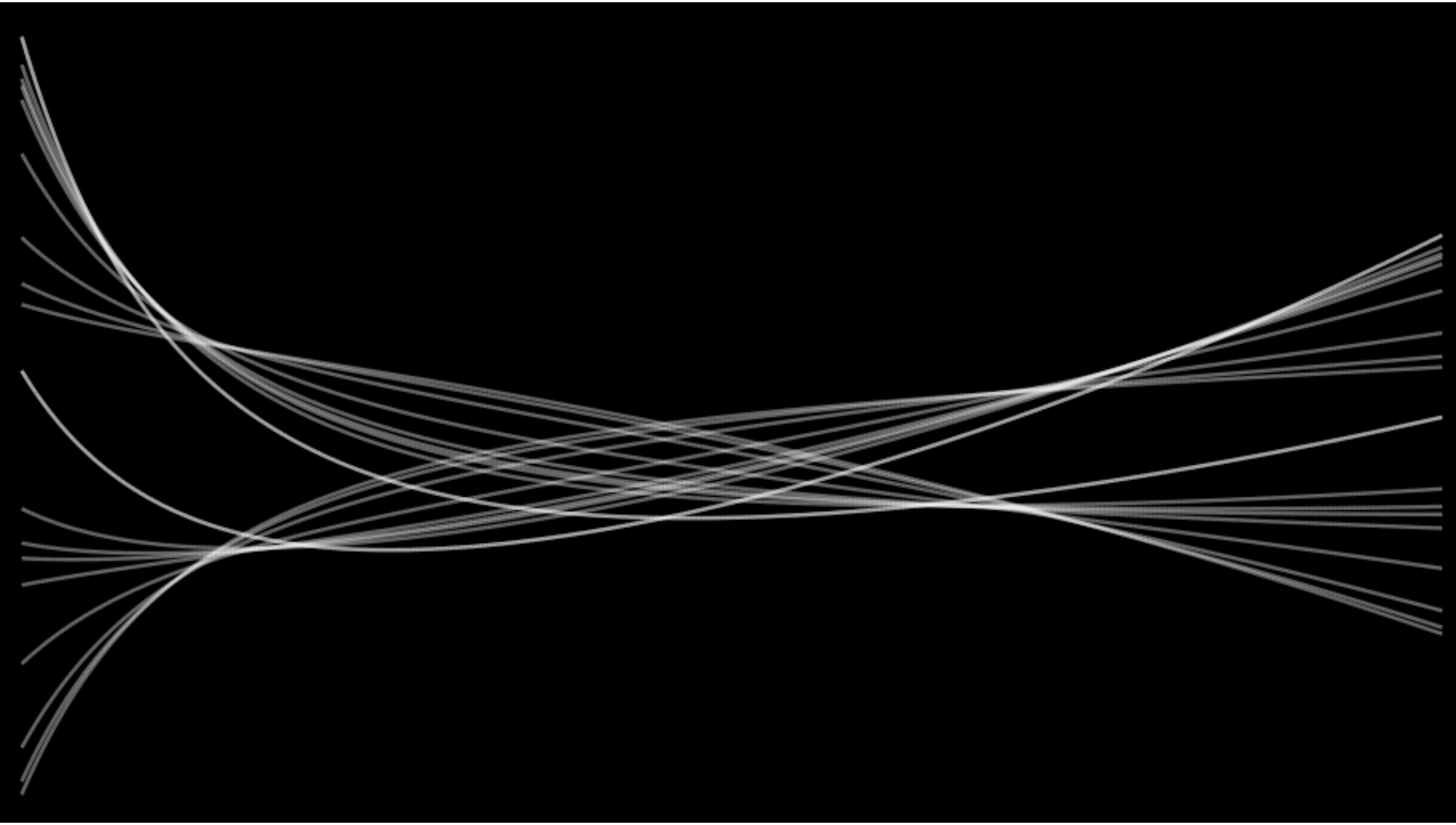
Translucency

- As of Stata 15, you can do this within Stata.

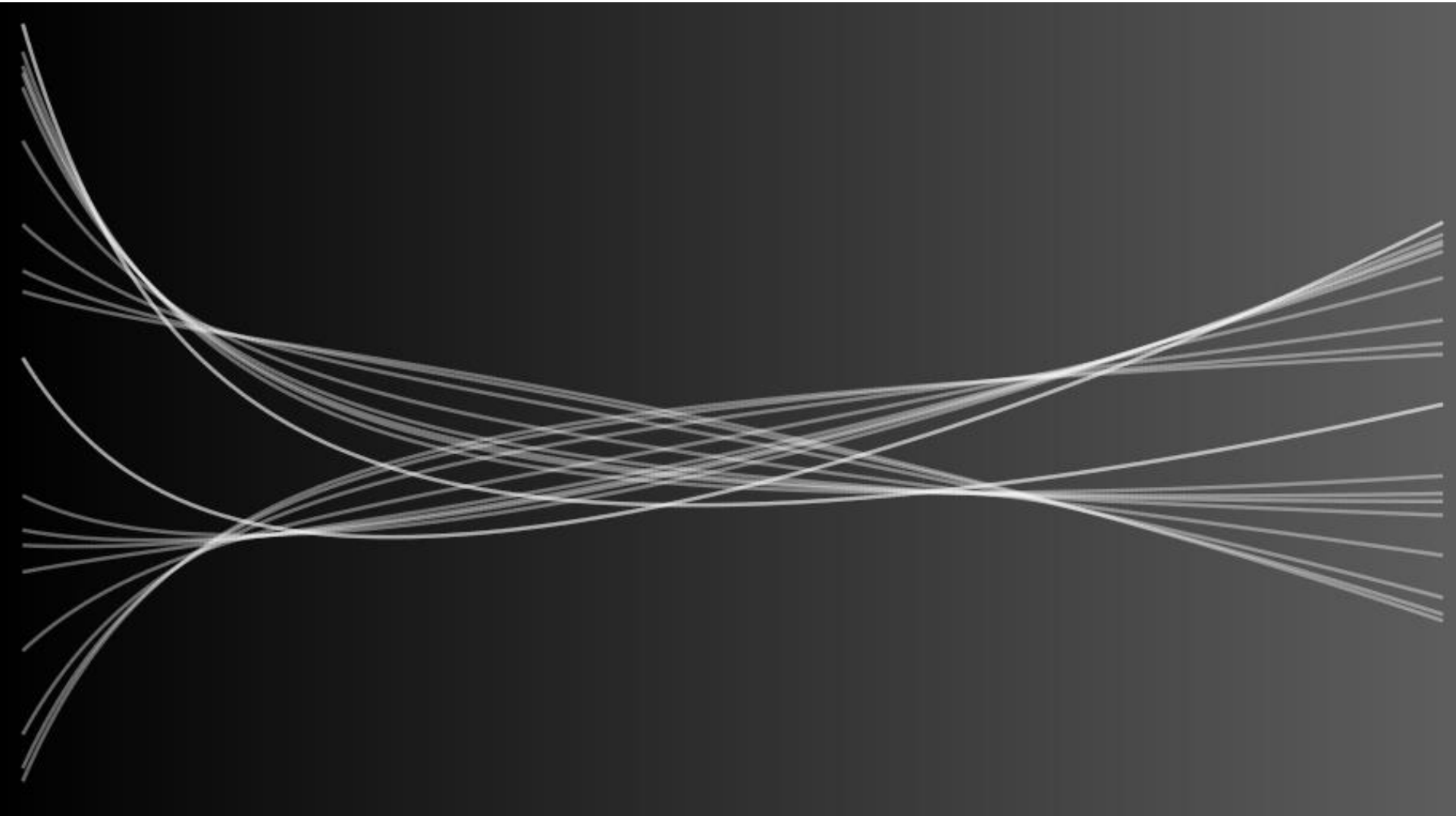
Translucency in action (1)



Translucency in action (2)



Add gradient to background



Add gradient to background

```
<svg version="1.1" width="7.00in" height="4.00in" viewBox="0 0 5040 2880" xmlns="http://www.w
    <desc>Stata Graph - Graph</desc>
  <defs>
    <linearGradient id="MyGradient">
      <stop offset="0%" stop-color="#000000"/>
      <stop offset="100%" stop-color="#606060"/>
    </linearGradient>
  </defs>
  <rect x="0" y="0" width="5040" height="2880" style="fill:none;stroke:none"/>
    <rect x="0.00" y="0.00" width="5039.84" height="2880.00" style="fill:url(#MyGradient)"/>
```

Embed an image/images

- Hexagons are cool right now.
- Putting a picture in a hexagon makes the picture cooler by association.

Embed an image/images

- Take an SVG file
- Include a line referring to image:

  <image href="tim.jpg" … />
- Include more such lines if you want
- The image can be JPG, PNG, another SVG… probably GIF and more

tim.jpg

tim.jpg in an SVG hexagon

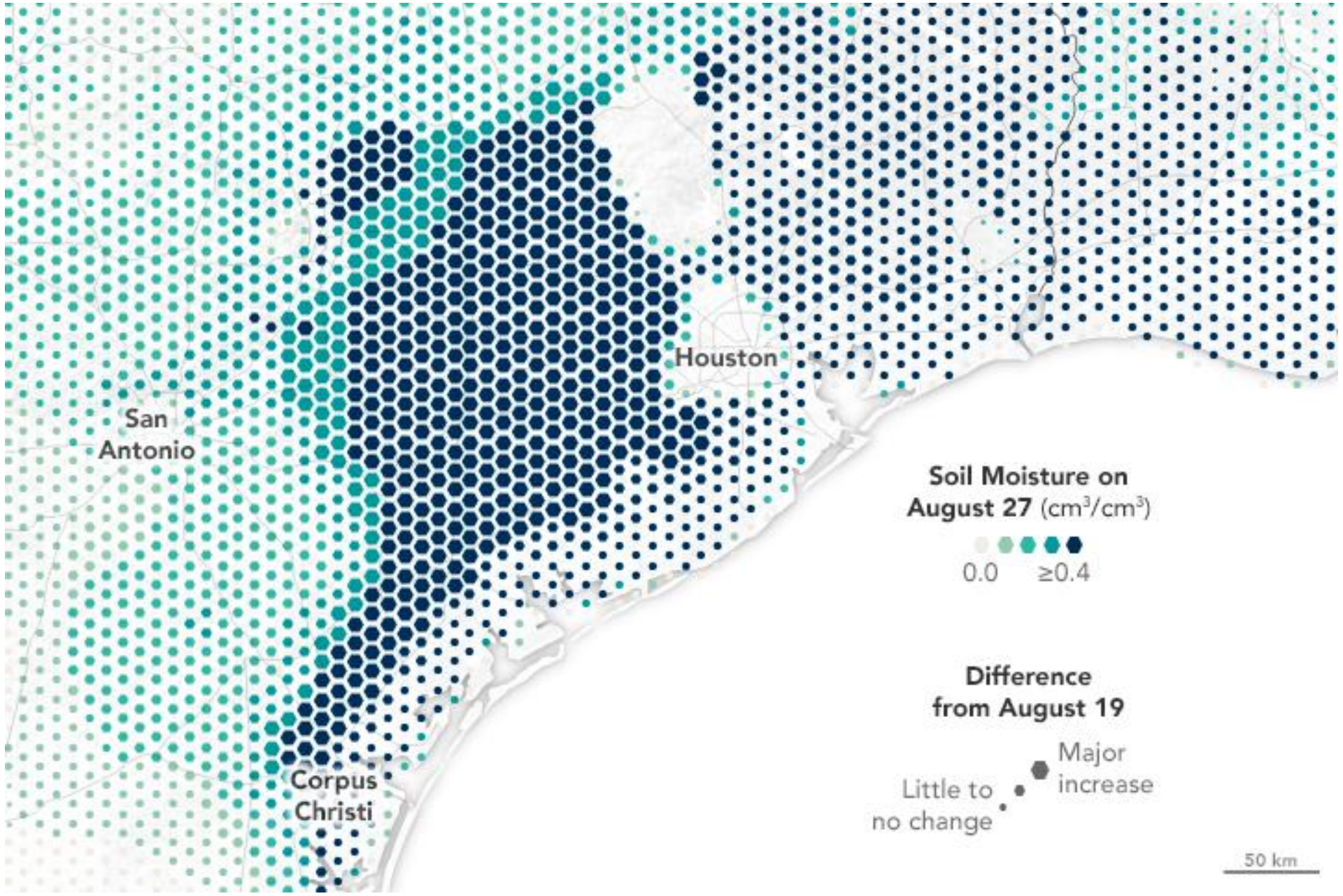
tim.jpg in an SVG hexagon

```
<svg width="100%" height="100%" viewBox="0 0 100 100" xmlns="http://www.w3.org/2000/svg">
  <rect x="0" y="10" width="30" height="30" style="fill:#808080;stroke:none"/>
  <image href="tim.jpg" x="5" y="15" height="20" width="20"/>
  <polygon points="15,12.5 25.8,18.75 25.8,31.25 15,37.5 4.2,31.25 4.2,18.75" fill="none"
</svg>
```

in same directory
as the hexagon
.svg file.

Hexagonal binning

- A ‘hexbin’ plot is a sort of bivariate histogram
- Hexagons form a regular tessellating grid (and look better than triangles/rectangles)
- Number of points covered by each hexagon is counted
- Hexagon coloured/sized by count


Houston
San Antonio
Corpus Christi
Soil Moisture on August 27 (cm³/cm³)
0.0
≥0.4
Difference from August 19
Major increase
Little to no change
50 km

Hexbin using Stata

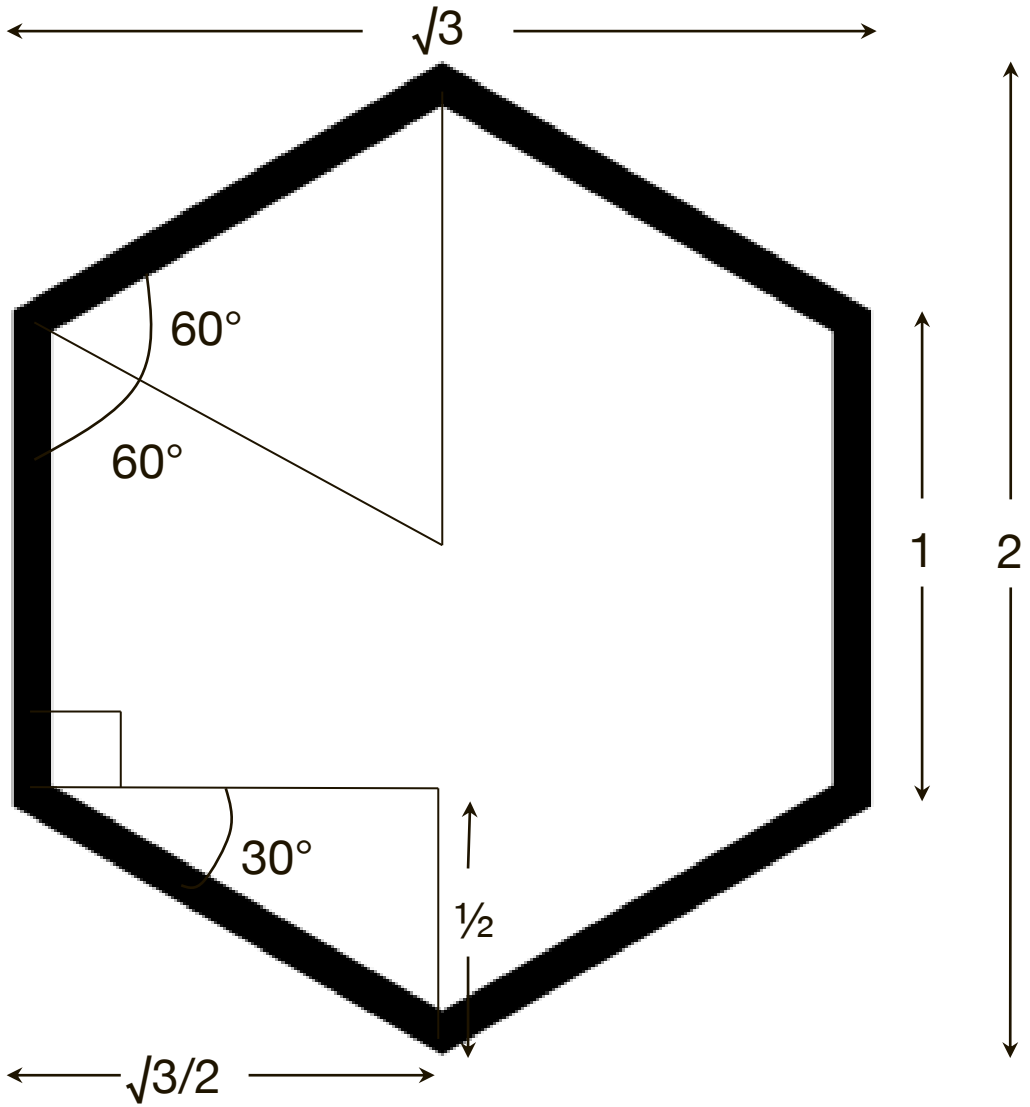
Here, we are producing a plot that Stata does not

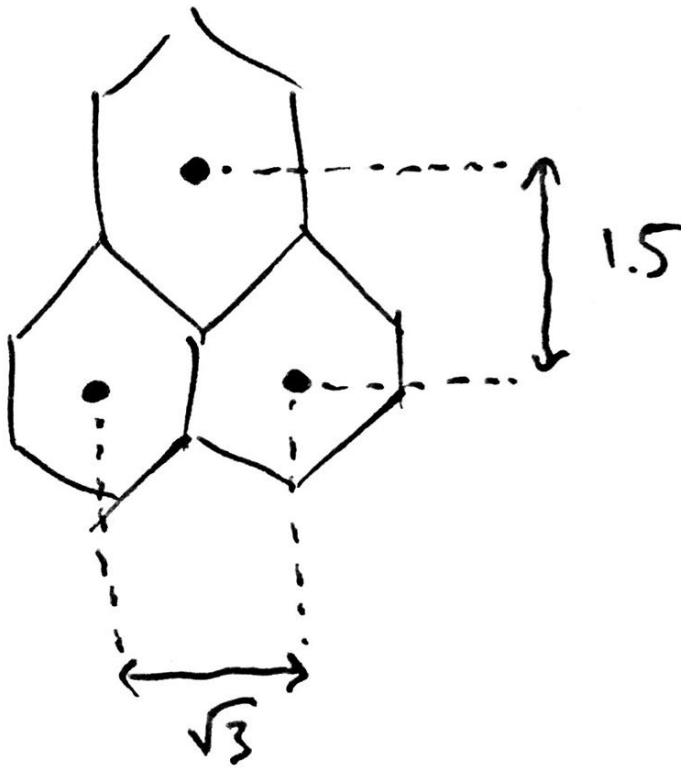
- Set up a grid of hexagons in Stata and count the number of points in each hexagon's territory
- Output a scatterplot with points on the grid
- Define a hexagon in SVG and replace each point output by Stata

xkcd.com/1050/

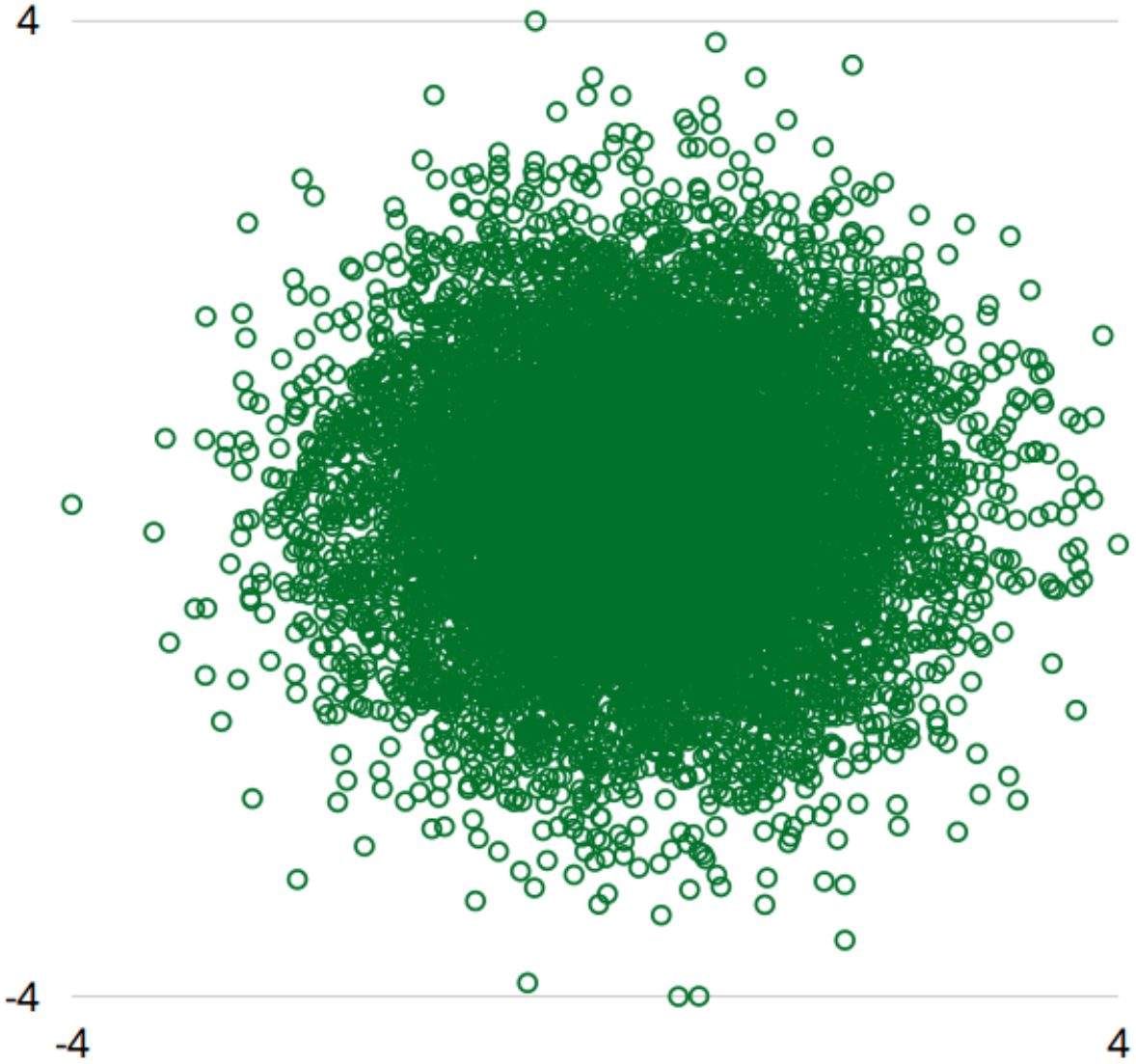
To do hexagon counting, we will use some high-school algebra

Hexbin

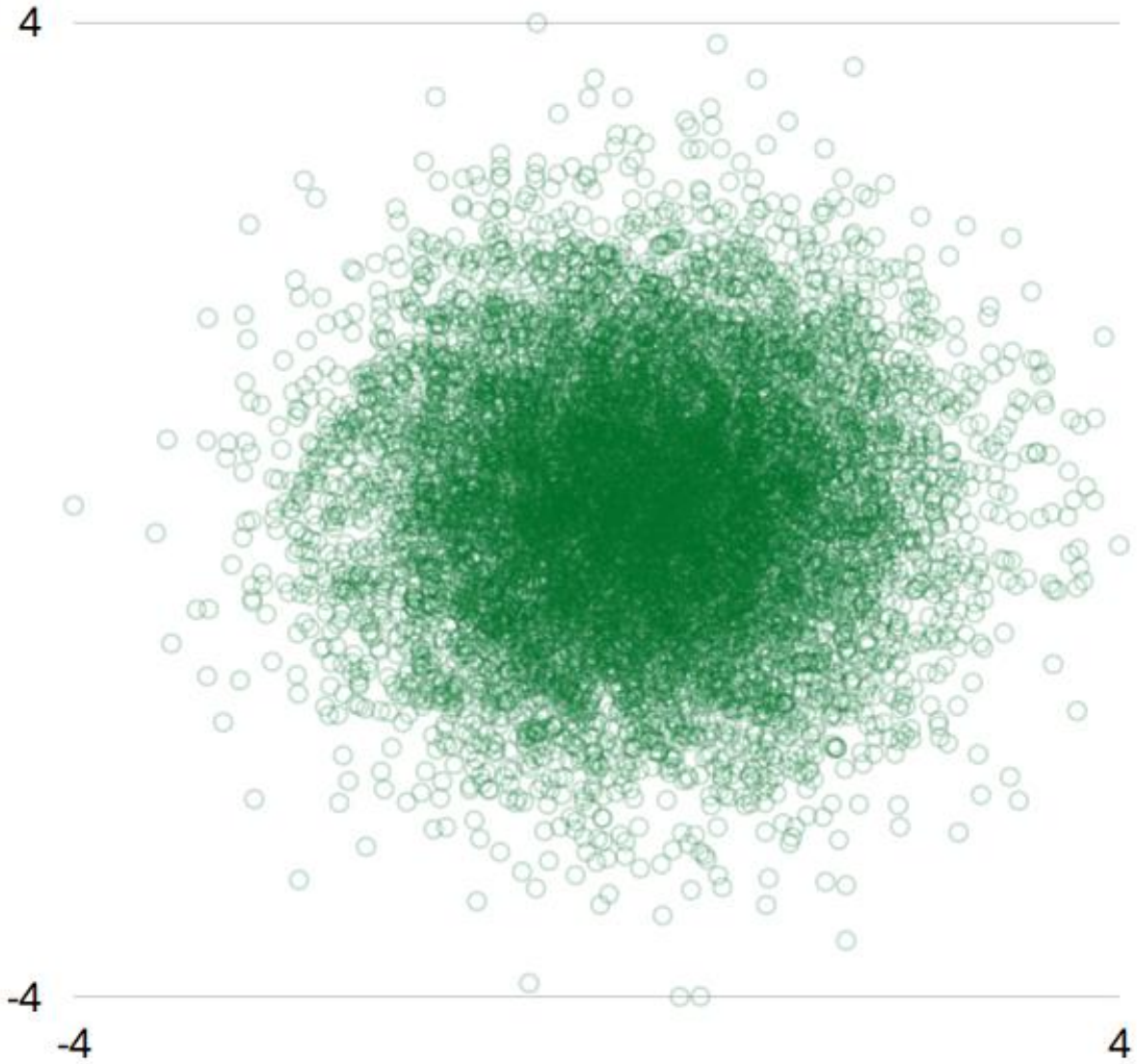




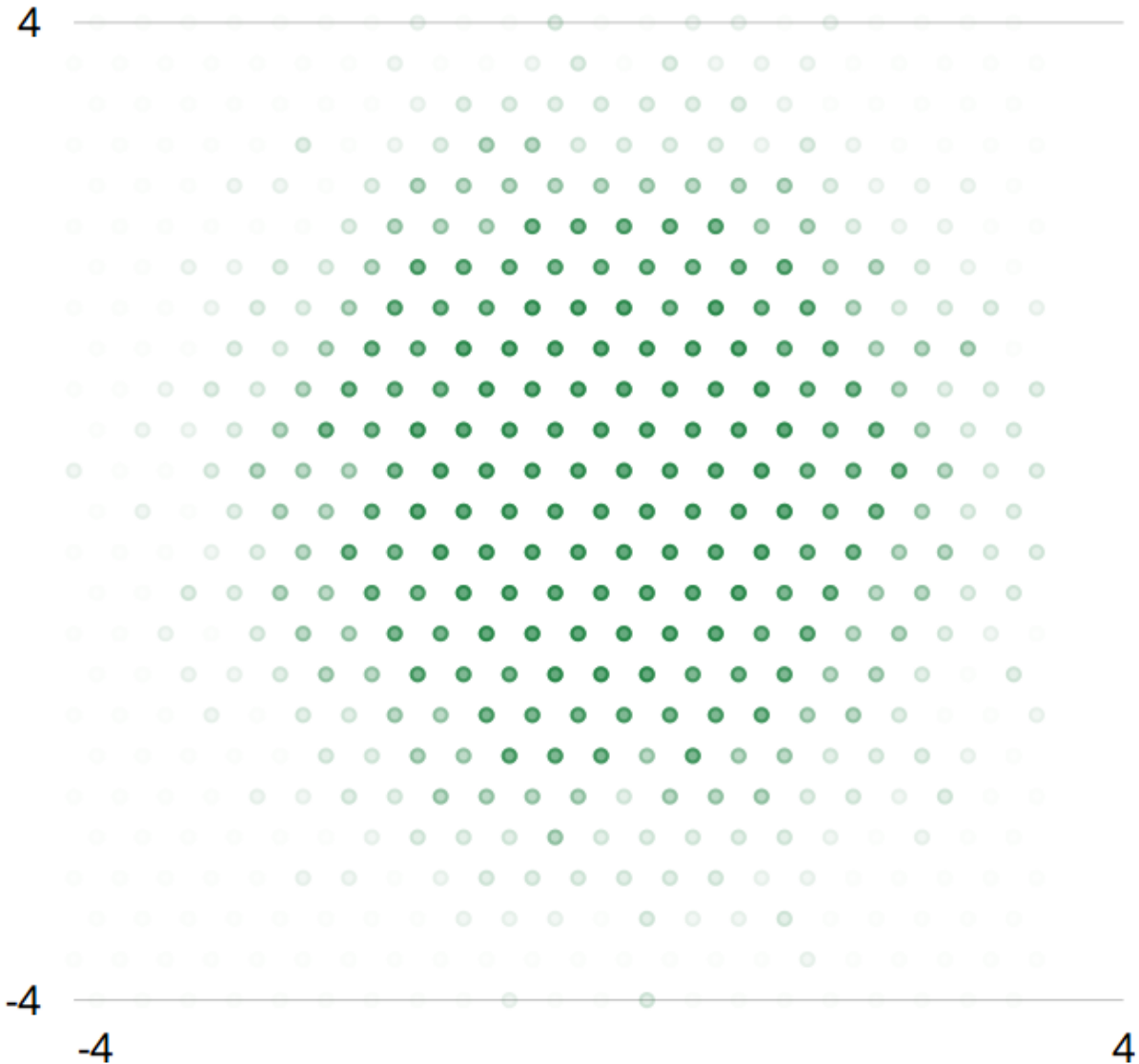
Bivariate normal



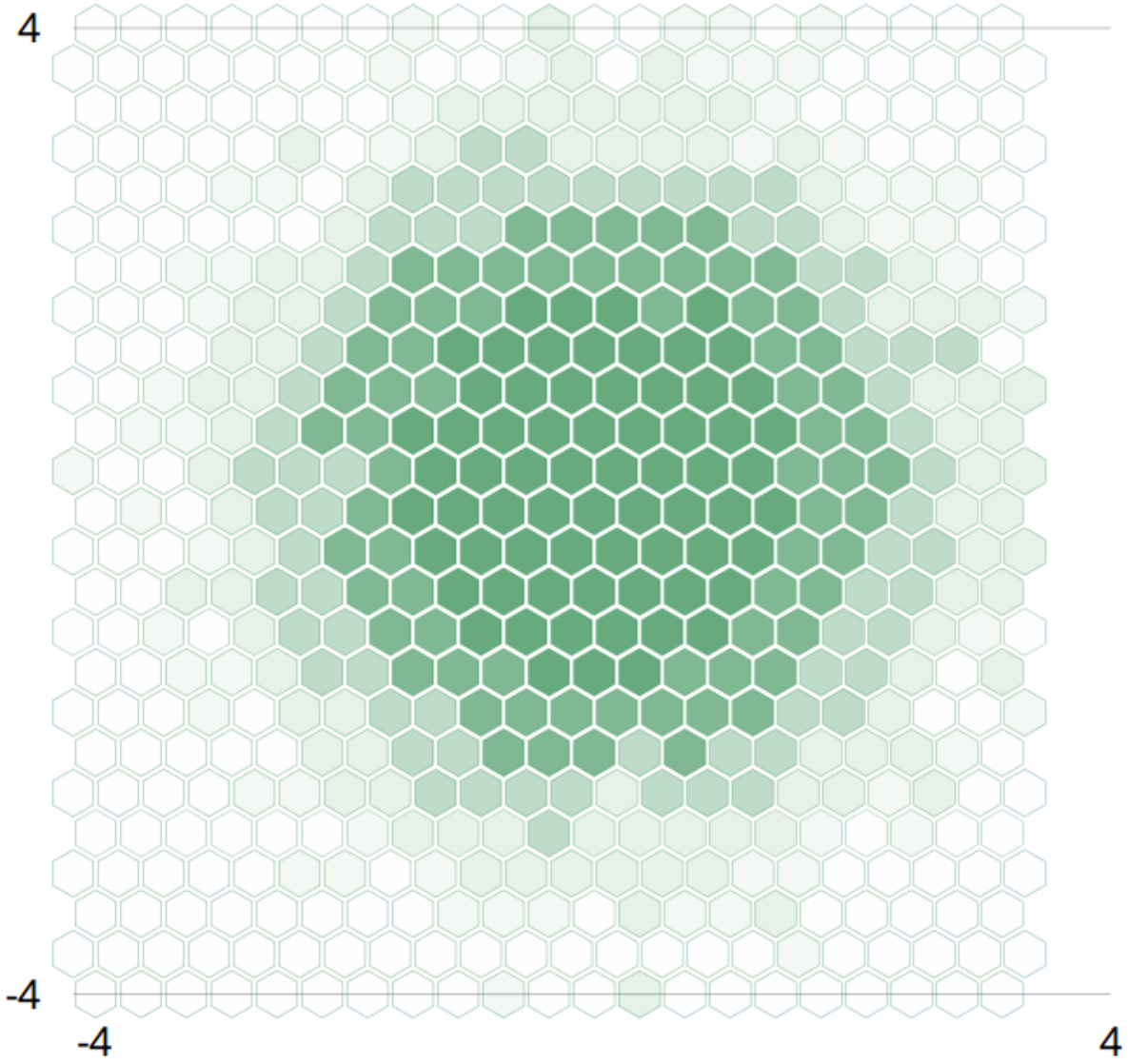
Better with translucency…



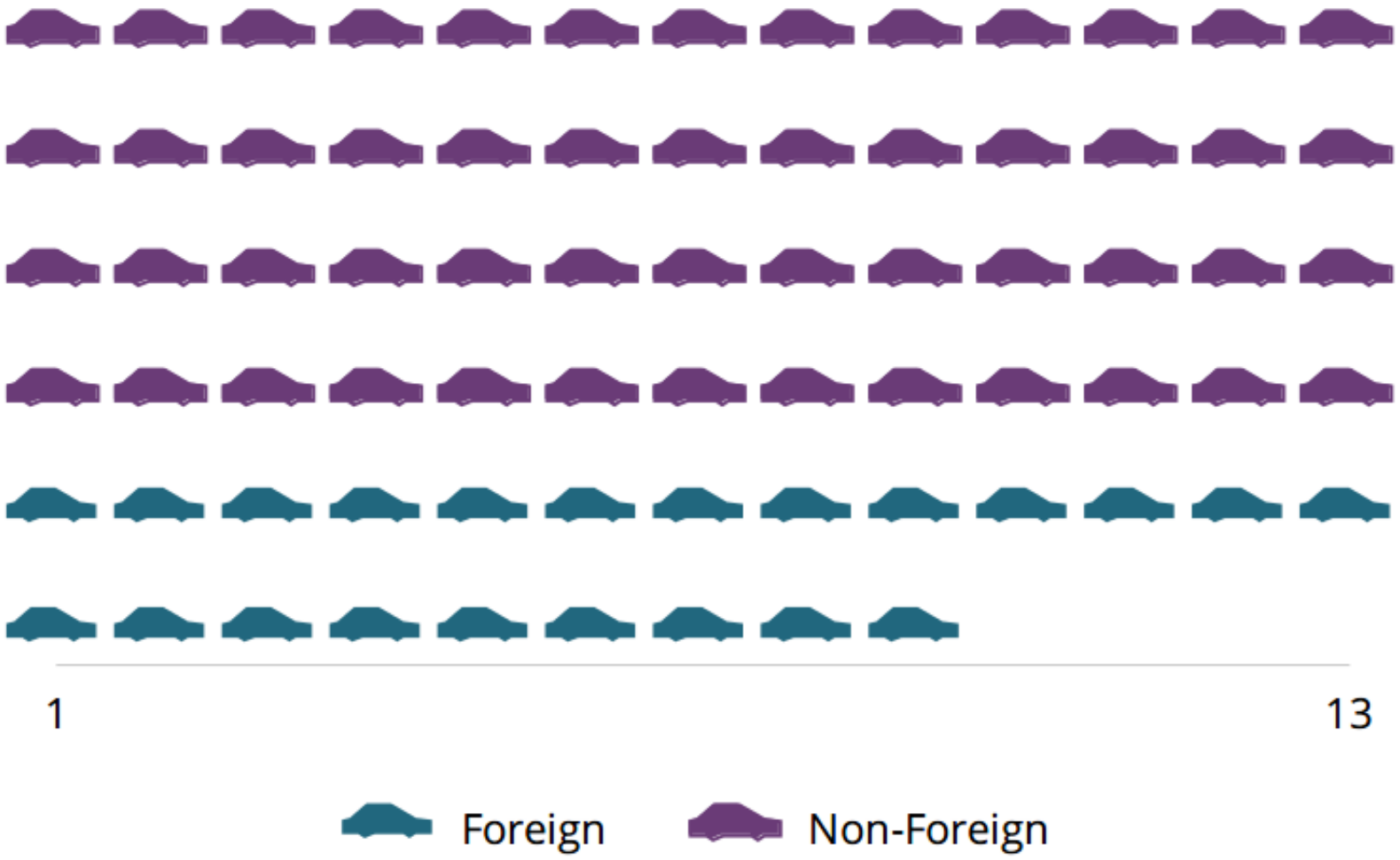
Stata’s grid of hex centres



Hexagons from SVG



Waffle charts (similar concept)



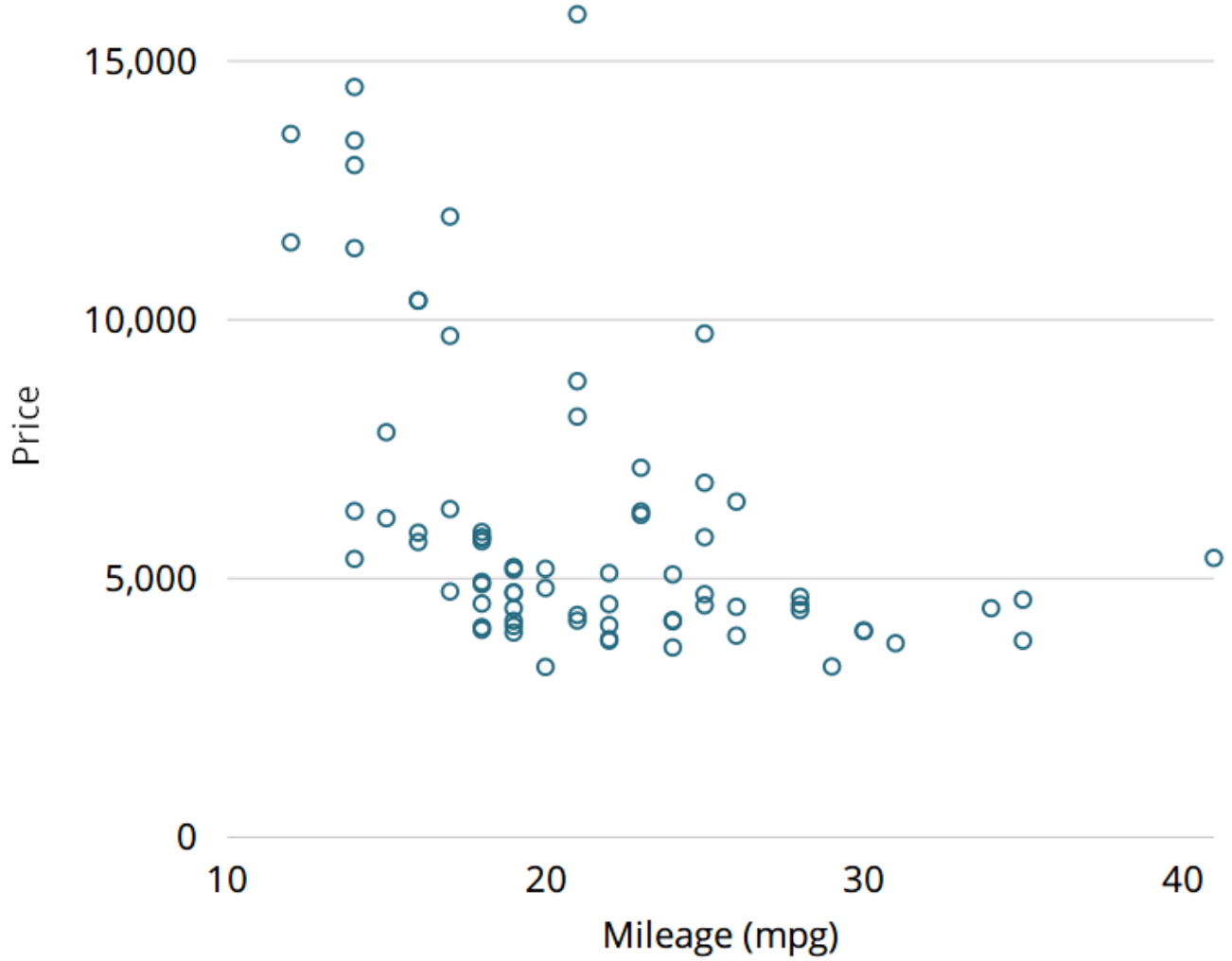
Maps

We now have the tools to produce hexbin maps

- Create hexbin using latitude and longitude
- Find (or draw) a map and embed as an image behind your hexbin
- Ensure they line up. This is hard:
  - Start with a scatterplot with a few key points and check that you can make these line up with map
  - Produce hexbin with same axis ranges, xsize, ysize, aspect etc.

Interactivity

- Add basic HTML before & after the SVG
- Now you have a web page
- Connect objects to controls like sliders, buttons, etc. with JavaScript



The make of this car is:

Olds 98

Much more is possible!

See:

- Nadieh Bremer's magic https://www.visualcinnamon.com/2016/04/svg-beyond-mere-shapes.html
- Sarah Drasner's animations

What are your ideas?

Thanks for listening