

# Tools!



# Plan for today:

Discussion of types of tools

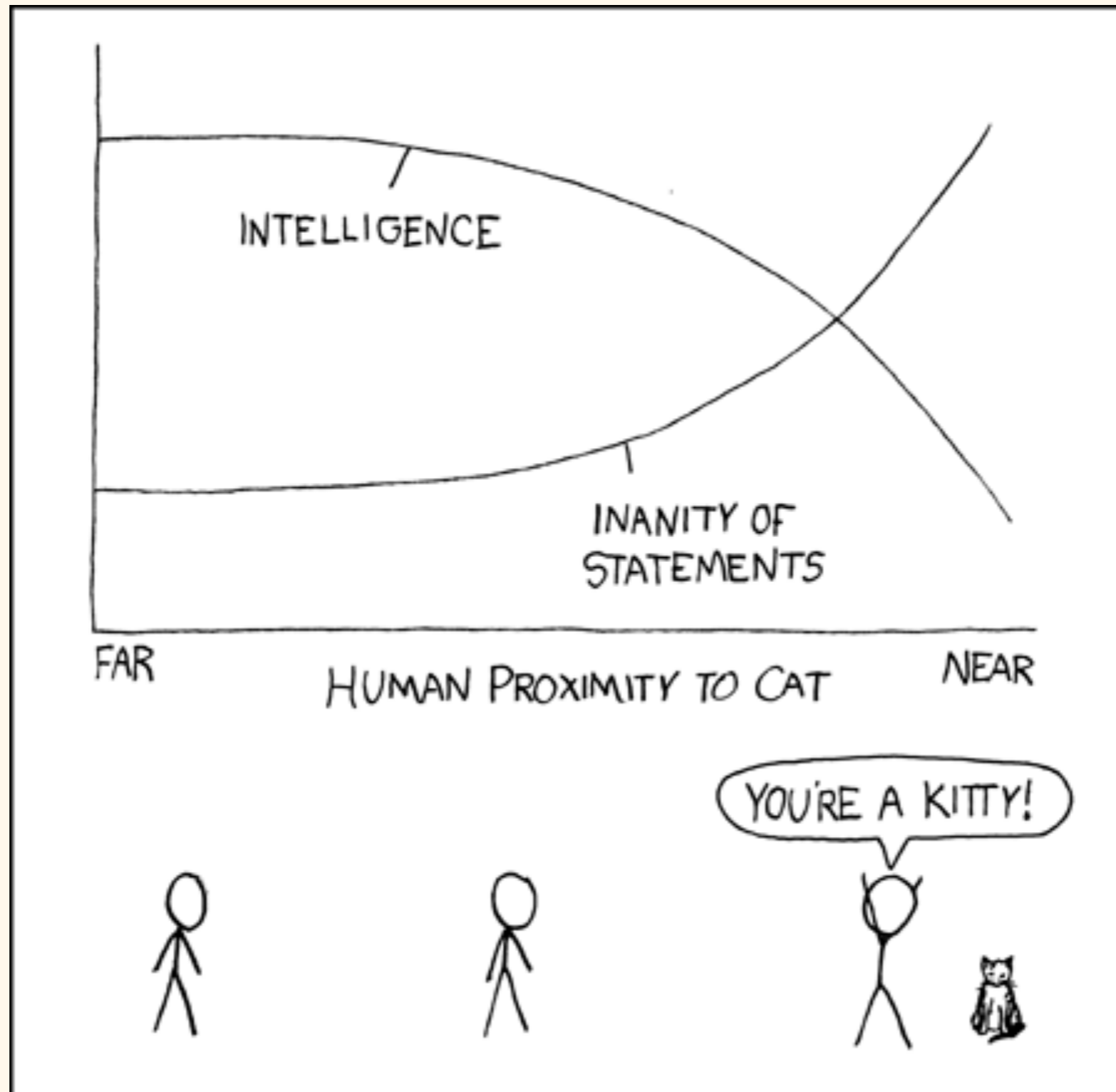
Example tools

Workflow considerations

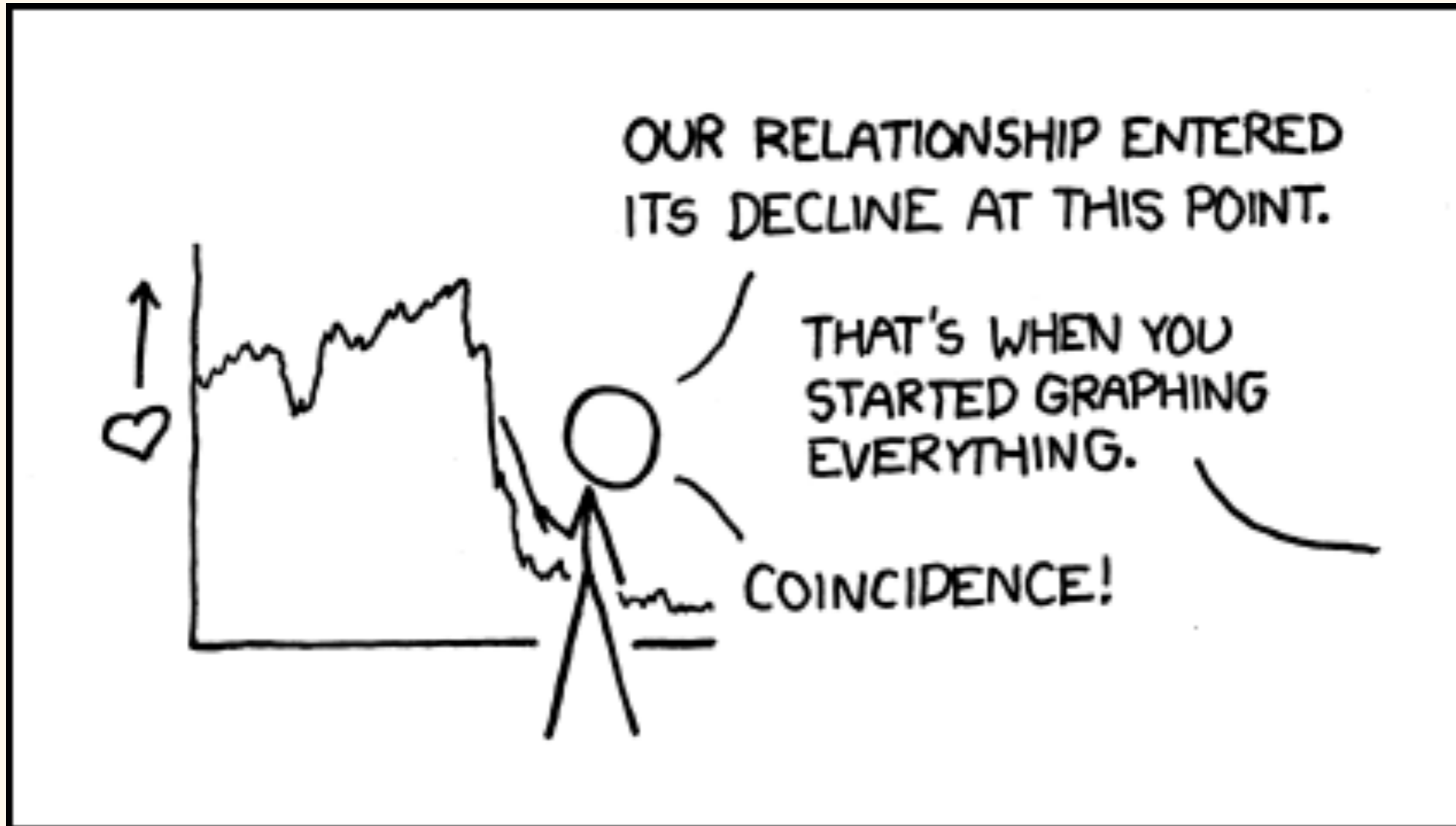
Other considerations

Avoid the default settings!

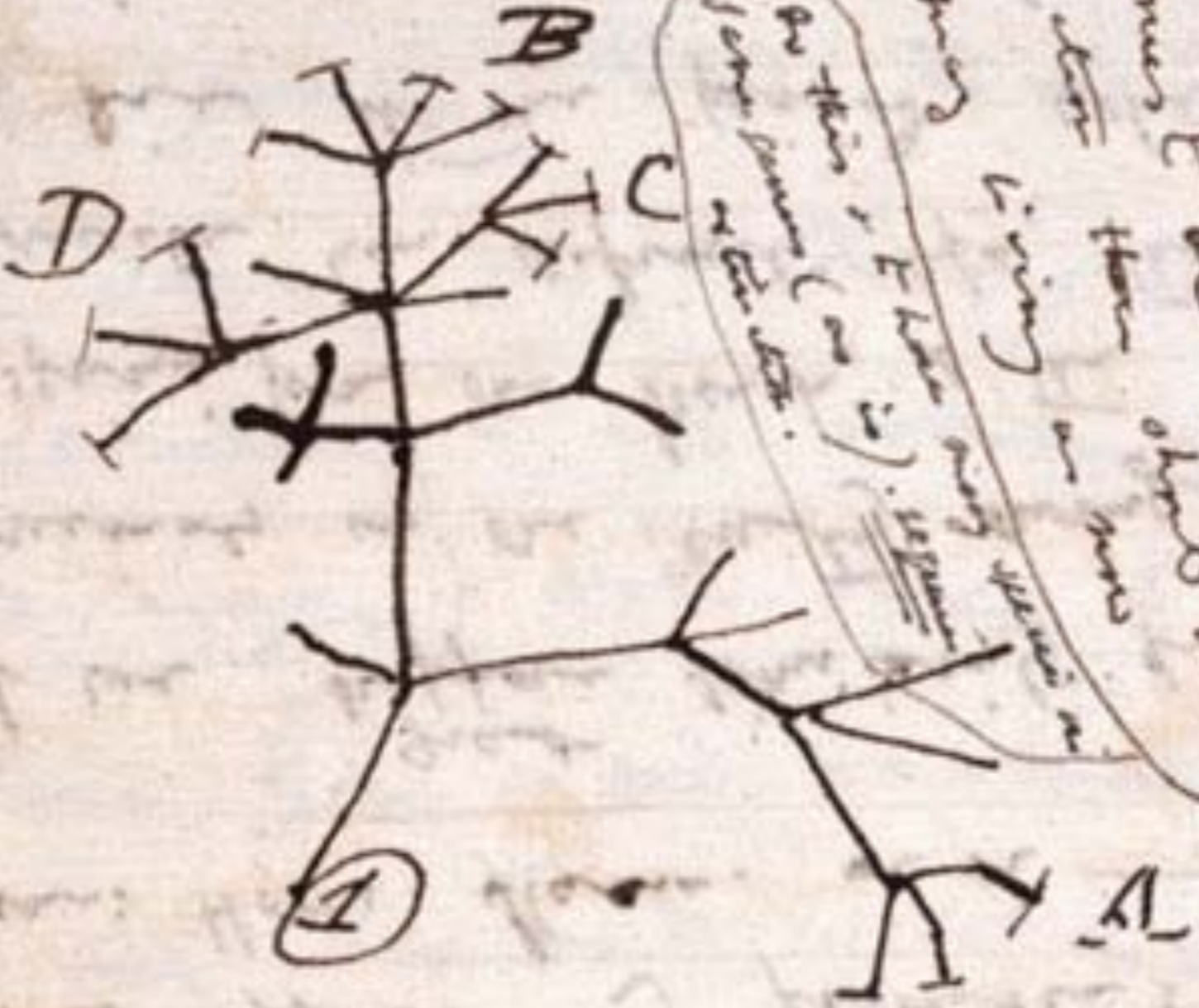
First: It is possible to do good visualization without a computer!



First: It is possible to do good visualization without a computer!



I think

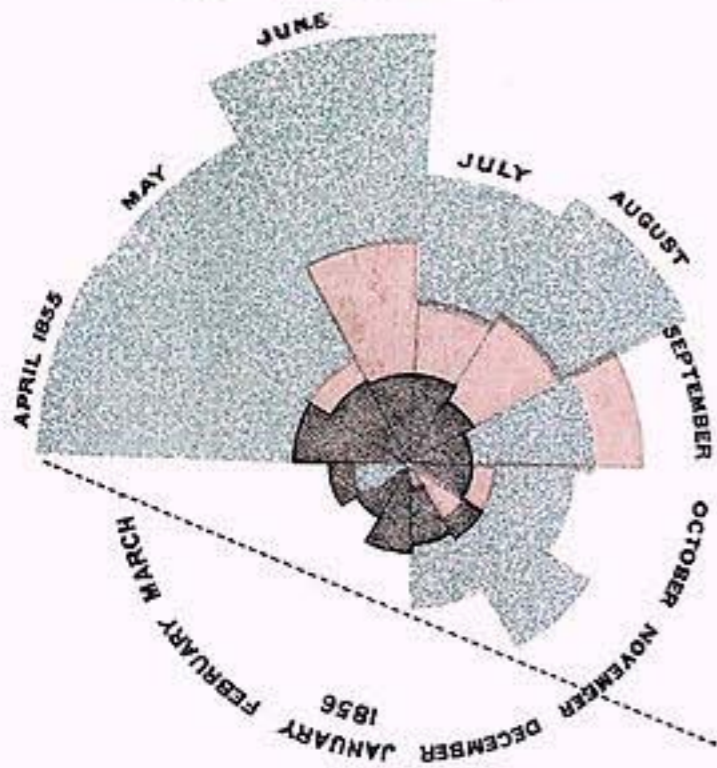


I think  
 some must be  
 some seen there  
 living in  
 some  
 I do not think it has any  
 or some species (as in) separate  
 or the same

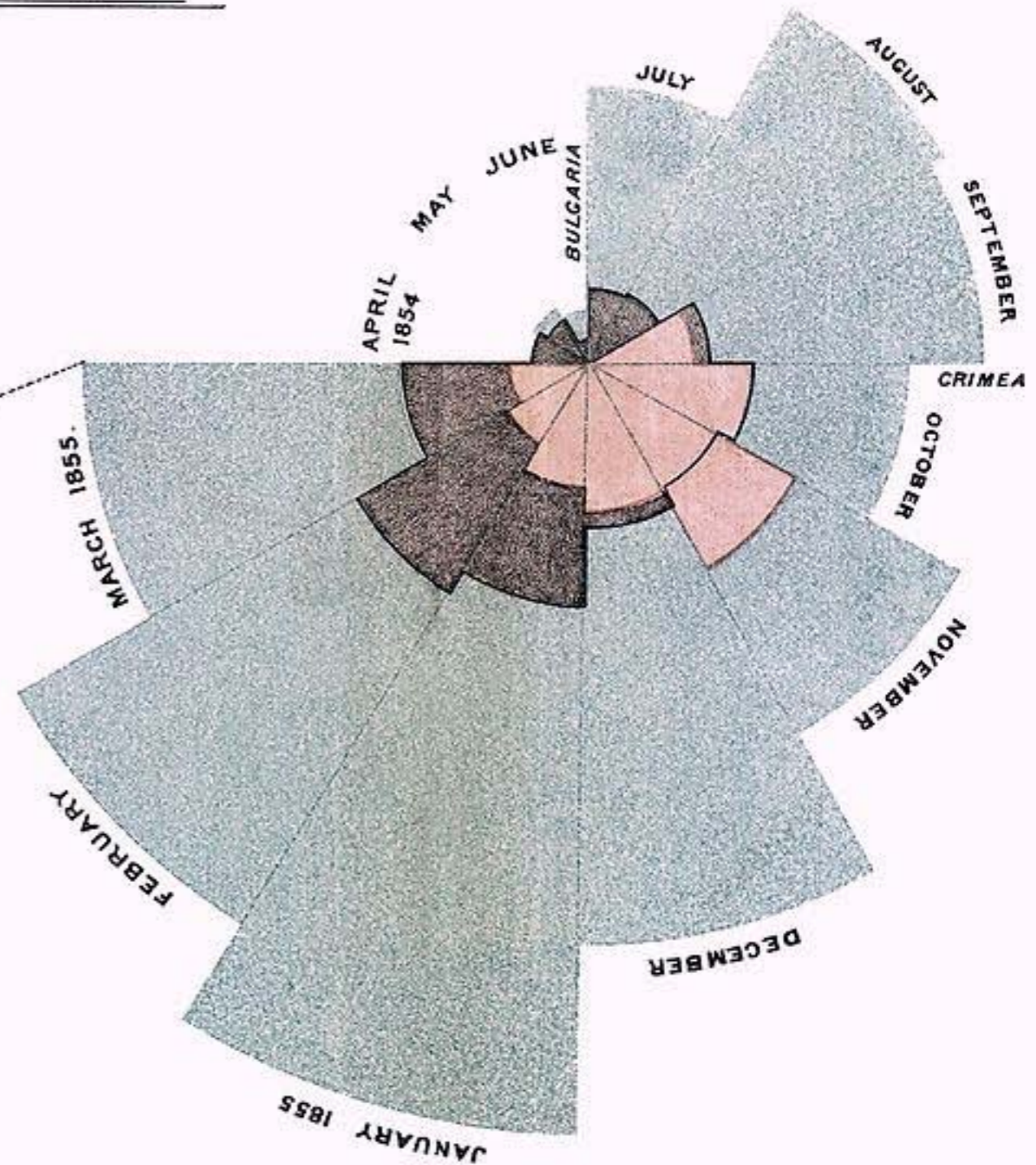
Ask me  
 should be  
 in some  
 living in

# DIAGRAM OF THE CAUSES OF MORTALITY IN THE ARMY IN THE EAST.

2.  
APRIL 1855 to MARCH 1856.



1.  
APRIL 1854 to MARCH 1855.



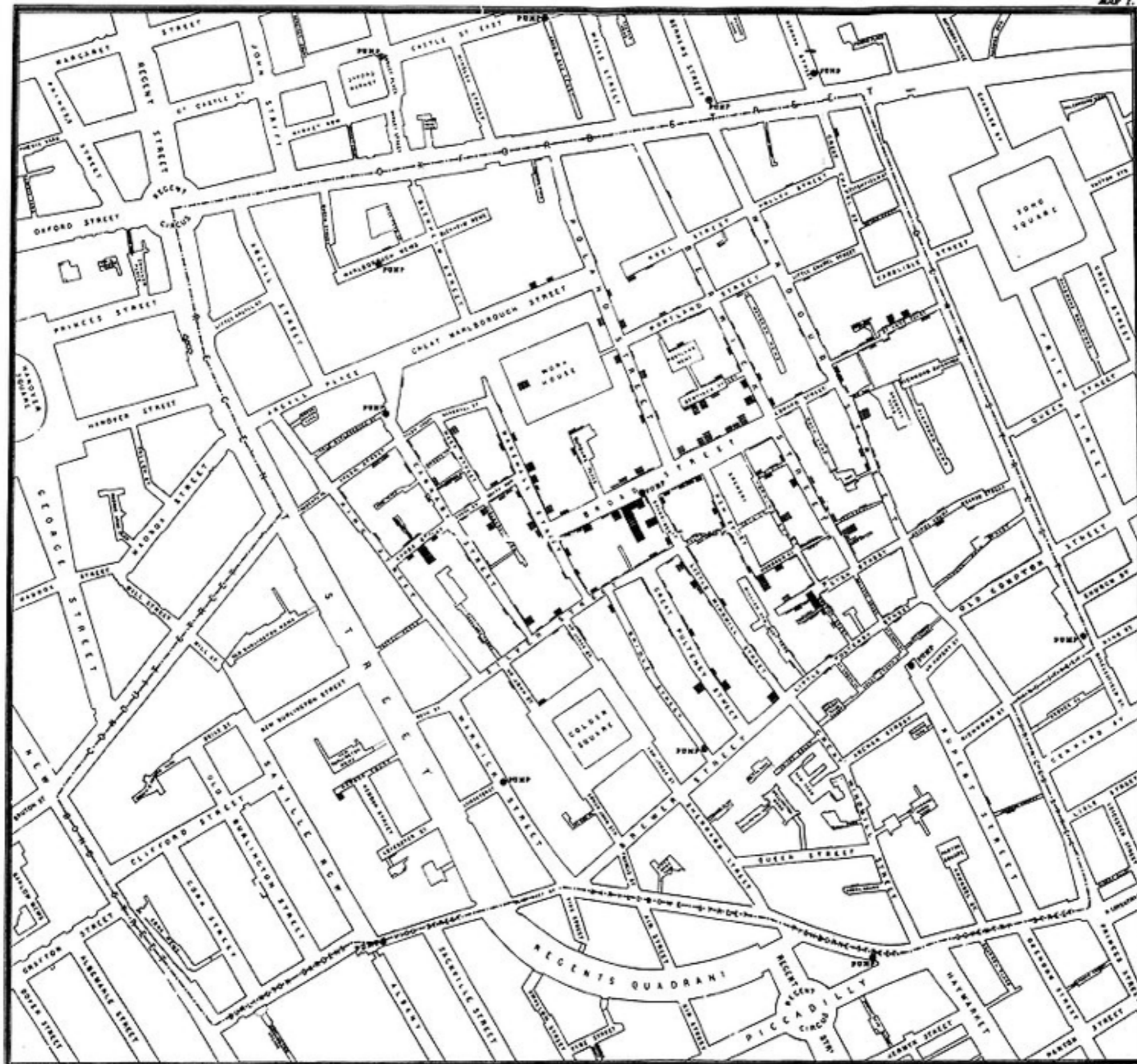
*The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.*

*The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases; the red wedges measured from the centre the deaths from wounds; & the black wedges measured from the centre the deaths from all other causes.*

*The black line across the red triangle in Nov<sup>r</sup> 1854 marks the boundary of the deaths from all other causes during the month.*

*In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.*

*The entire areas may be compared by following the blue, the red & the black lines enclosing them.*



C. J. Ogilby, Lith. Indagation by Lewis

SCALE 80 INCHES TO A MILE.

That said, computers make it a lot easier.

There are two main categories of tool:

1. Data analysis software that also does visualization;
2. Dedicated visualization packages.

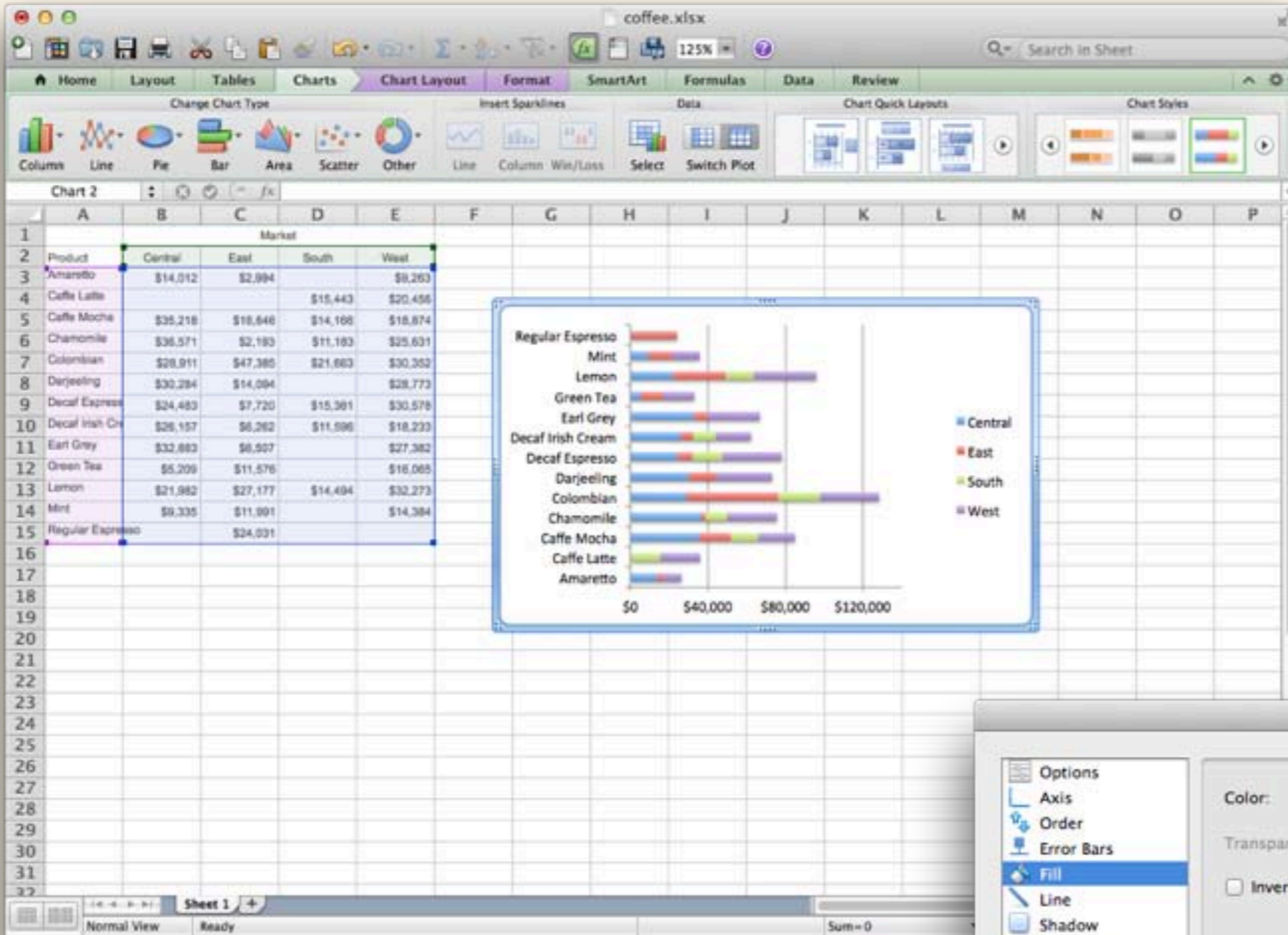


# 1. Data analysis software that also does visualization;

The image displays three overlapping software windows used for data analysis and visualization:

- Microsoft Excel (Sales Data - Microsoft Excel):** Shows a spreadsheet with sales data. It includes summary tables for 'Total Sales by Store', 'Total Sales by Region', and 'Total Sales by Category'. A 3D pie chart titled 'TOTAL SALES BY CATEGORY' is also visible.
- SPSS Data Editor (IEF - SPSS Data Editor):** Displays a data table with columns for 'country', 'ppp2000', and various statistical variables. The data includes entries for Albania, Algeria, and Angola.
- GraphPad Prism (Project1:Exponential decay):** Shows a graph titled 'Dissociation of  $\alpha_2$  receptors'. The graph plots 'Minutes' on the x-axis (0 to 60) and receptor count on the y-axis (0 to 12000). Two data series are shown: 'Control' (blue diamonds) and 'Treated' (red squares). Both series show an exponential decay over time. A 'Viewer Window' is highlighted, showing a histogram of 'ppp2000' values.

	Control	Treated
Y0	9253 to 10732	8362 to 10825
PLATEAU	254.3 to 1720	825.3 to 1482
K	0.06569 to 0.1128	0.2341 to 0.3743
Half Life	6.142 to 10.55	1.852 to 2.960
Span	8211 to 9800	7259 to 9621



Format Data Series

Solid Gradient Picture or Texture Pattern

Options

Axis

Order

Error Bars

Fill

Line

Shadow

Glow & Soft Edges

3-D Format

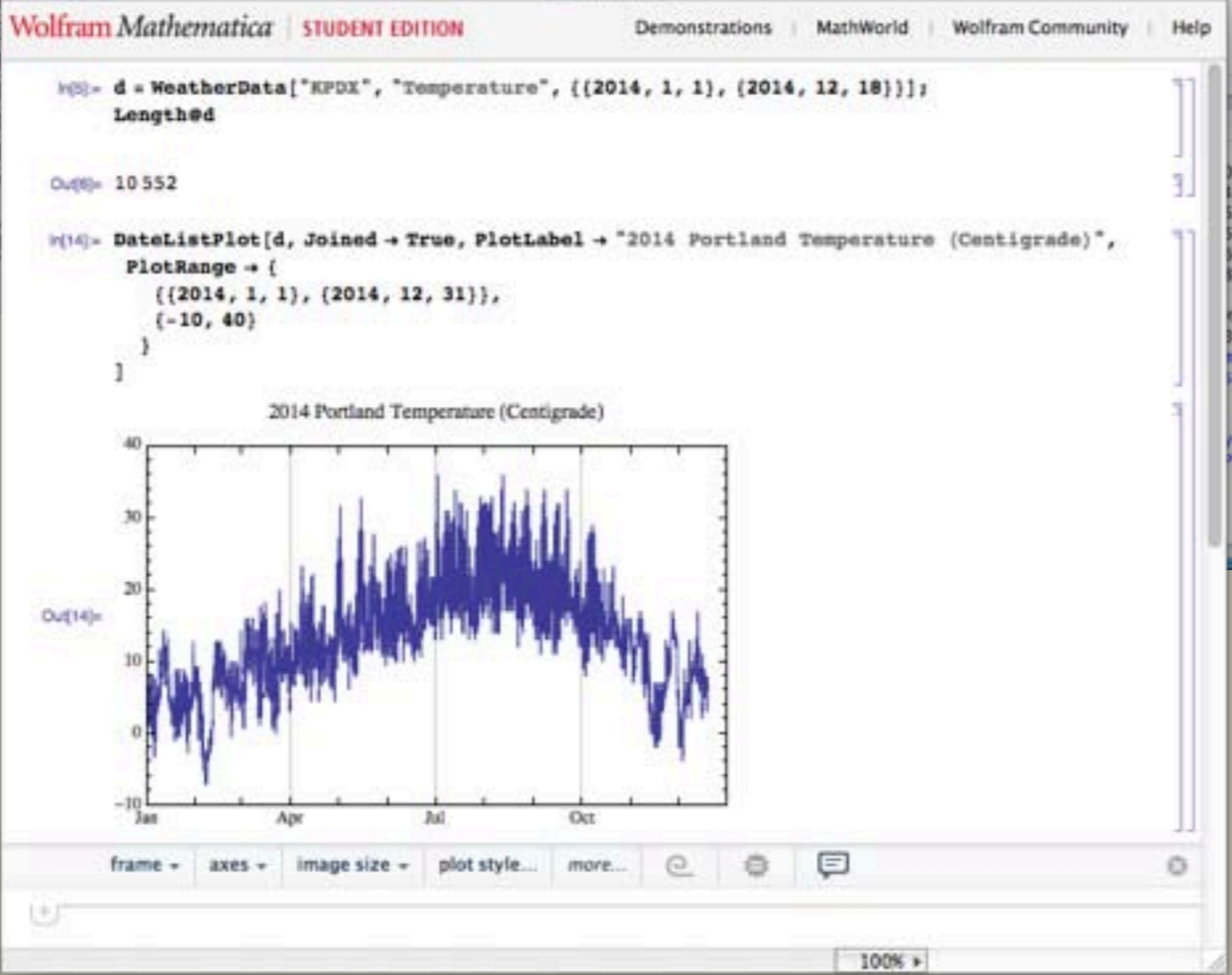
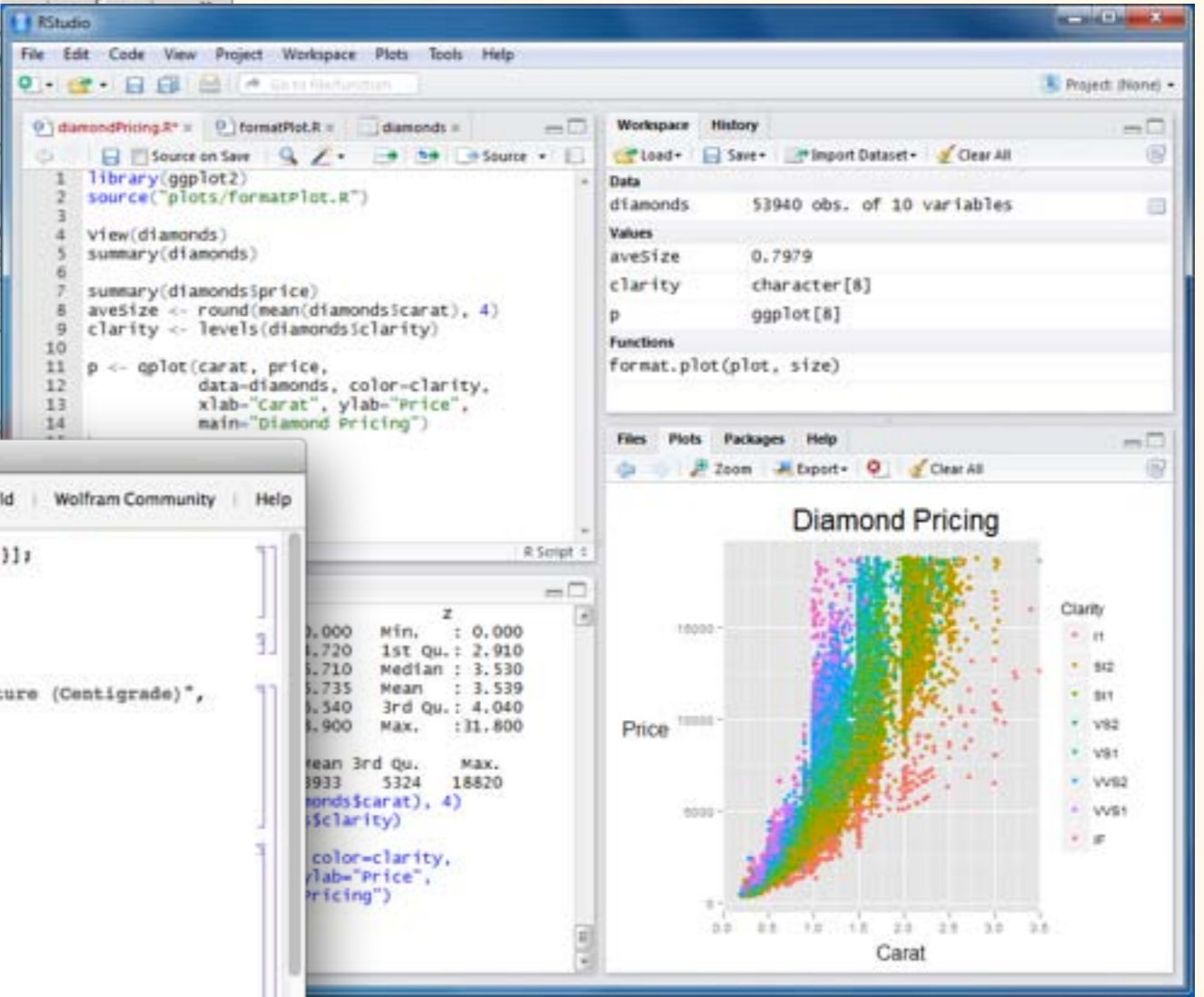
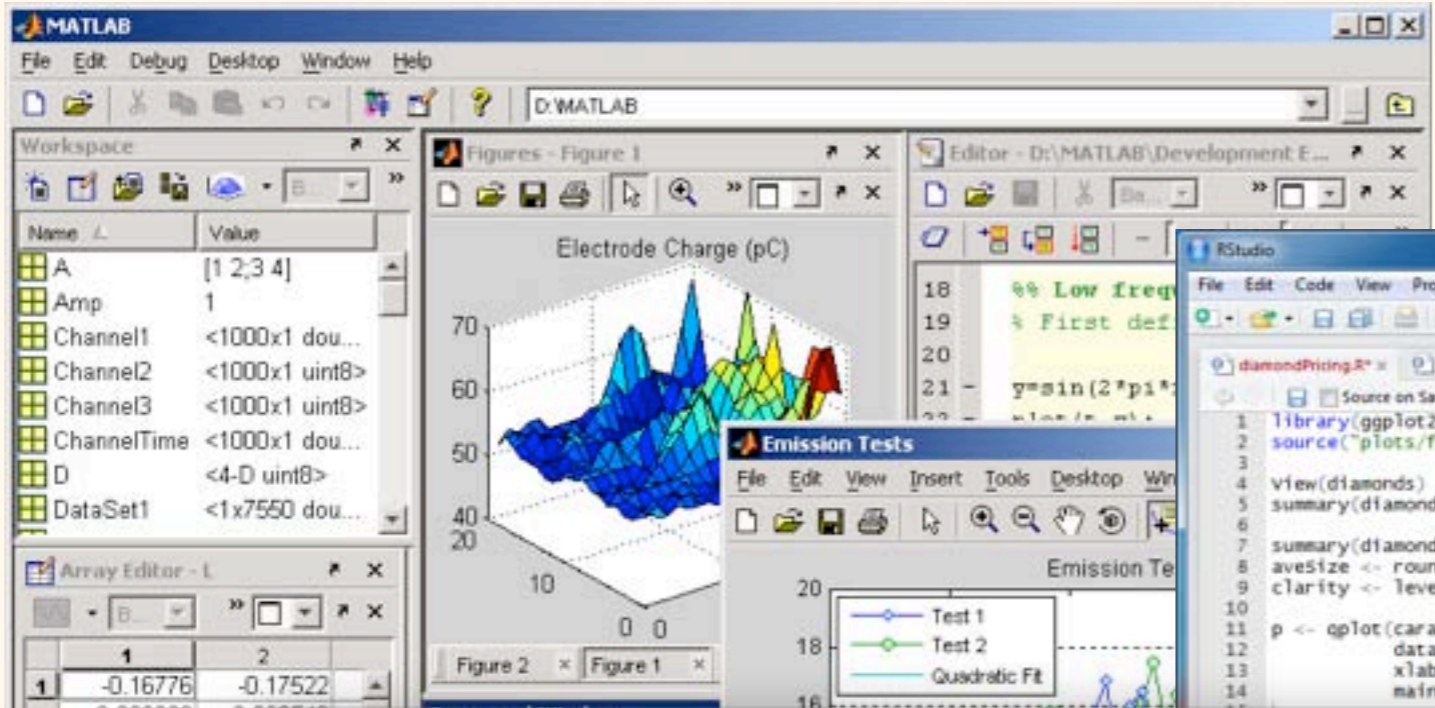
Color: Automatic

Transparency: 0 100

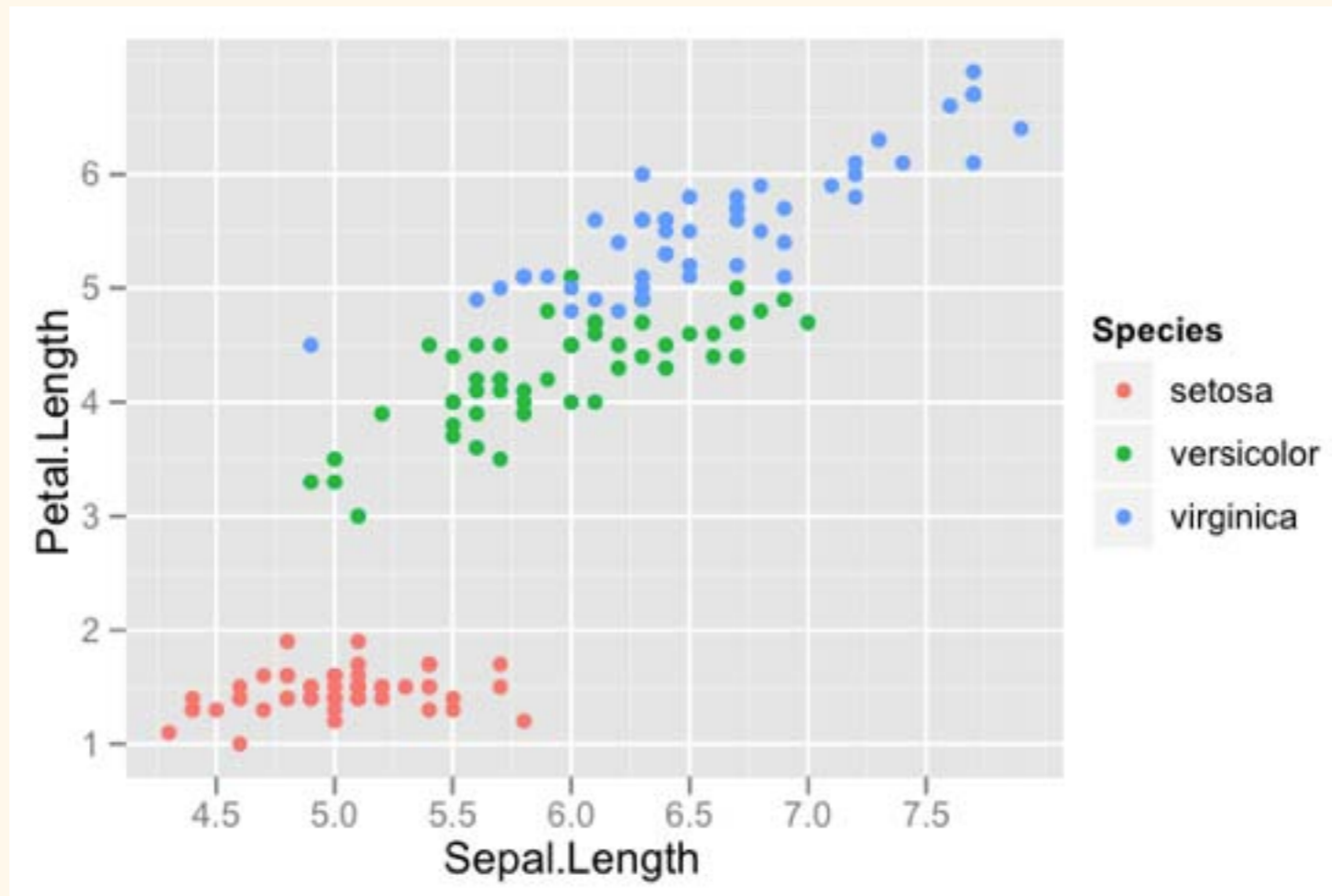
Invert if negative

Cancel OK

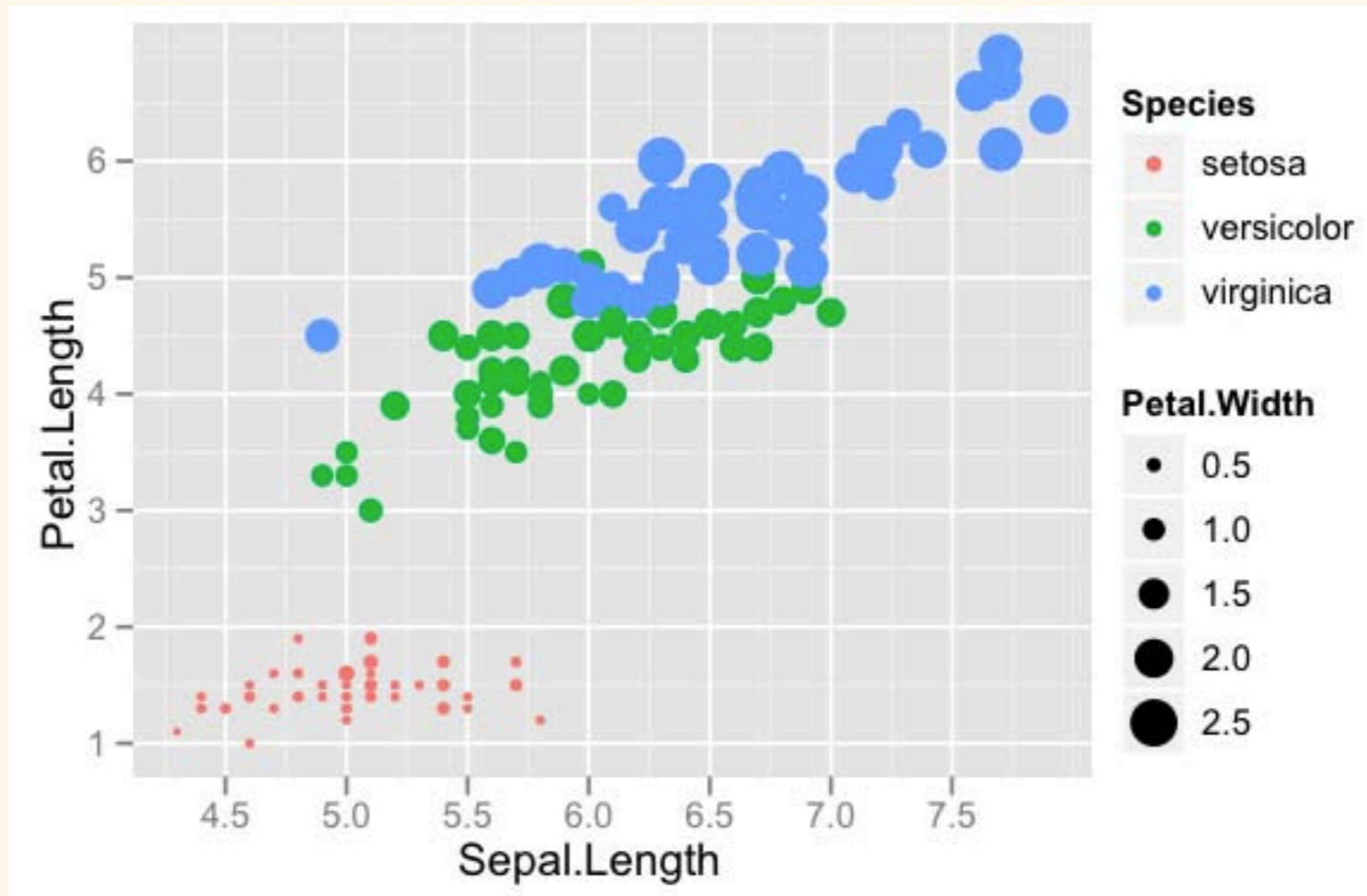
# 1. Data analysis software that also does visualization;



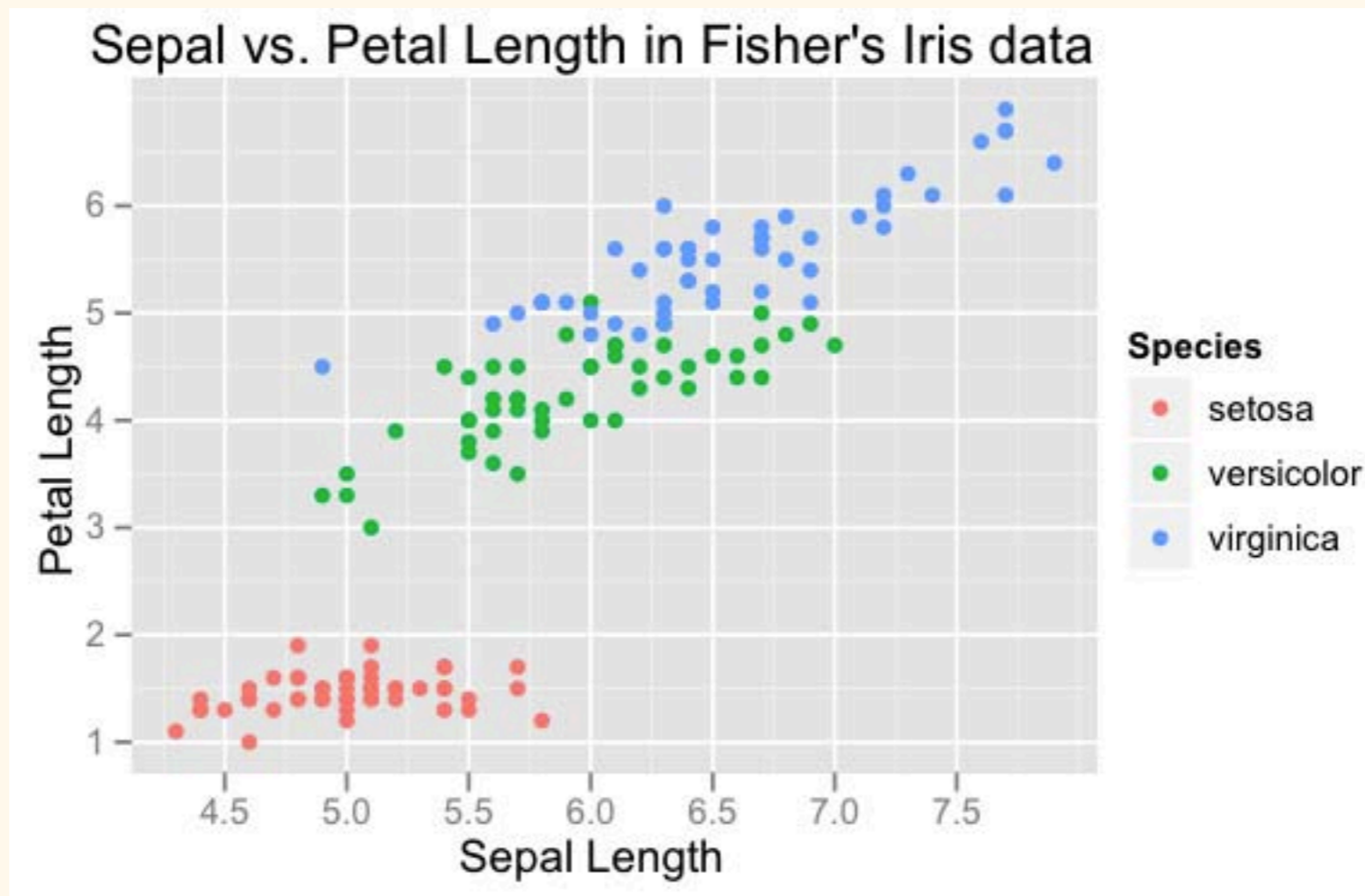
```
qplot(Sepal.Length, Petal.Length, data = iris, color = Species)
```



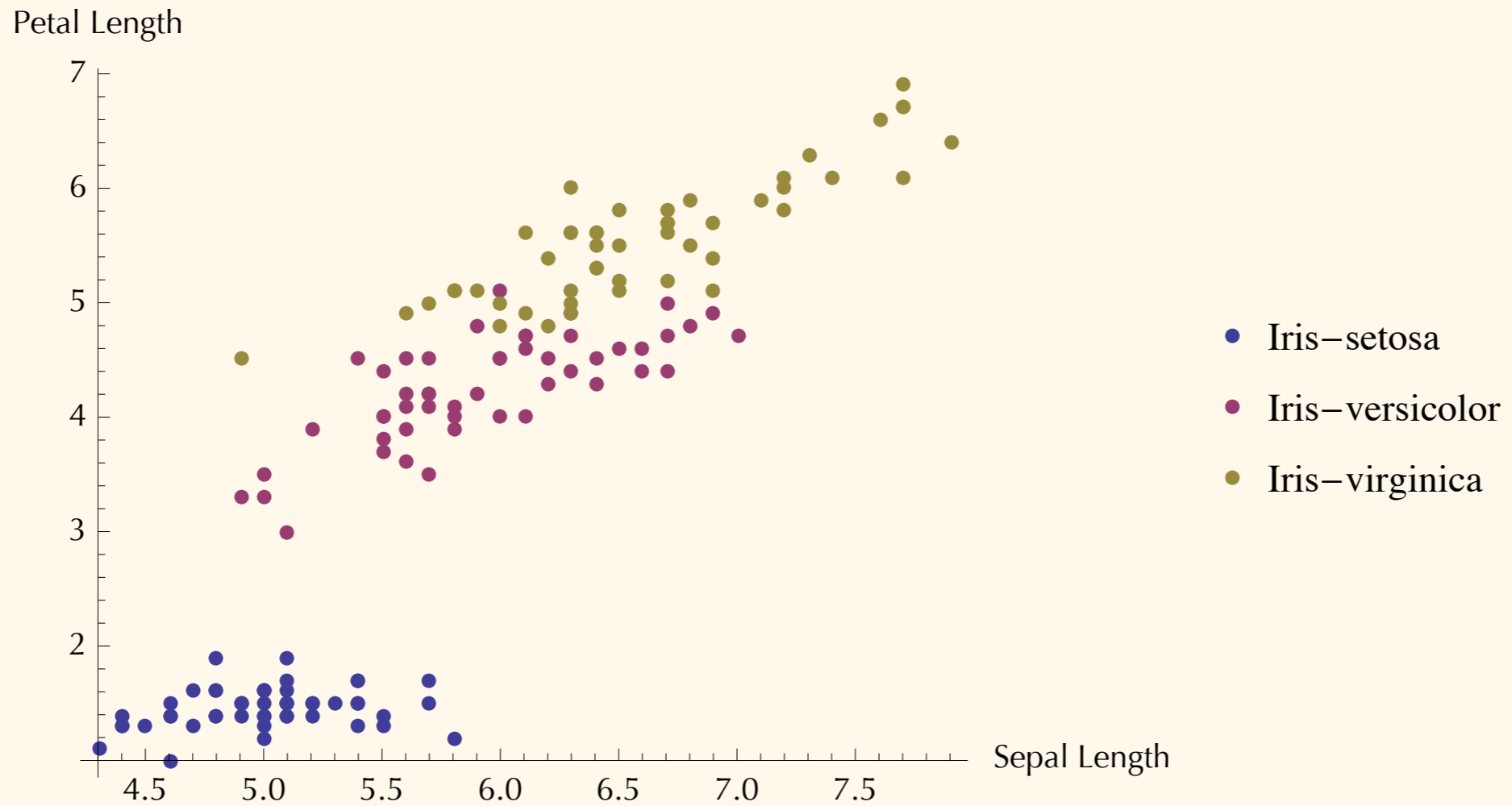
```
qplot(Sepal.Length, Petal.Length, data = iris, color = Species, size = Petal.Width)
```



```
ggplot(Sepal.Length, Petal.Length, data = iris, color = Species,  
       xlab = "Sepal Length", ylab = "Petal Length",  
       main = "Sepal vs. Petal Length in Fisher's Iris data")
```



```
ListPlot[
  irises[[All, All, {1, 3}]],
  PlotRange -> All, PlotStyle -> PointSize[Medium],
  BaseStyle -> {FontFamily -> "Optima"},
  PlotLegends -> z[[All, 1, 5]],
  AxesLabel -> {"Sepal Length", "Petal Length"},
  AspectRatio -> 2 / 2.5
]
```



# Graphical systems (Excel, etc.)

Pro:

“Easier”

Quick path to results, esp. for simple plots

Might be what you're already using

One-stop-shopping (analysis & vis. together)

Easier collaboration, probably



# Graphical systems (Excel, etc.)

## Cons:

Less control

Hideous/boring defaults

Lots of repetitive clicking around

Capabilities are (often) limited

Hard to exactly reproduce “what you did last time”

“Programmatic” systems (R, etc.)

Pro:

“Easier”

More control

Easier reproducibility

R is free!

Fewer limits

Default settings for graphs are less ugly

“Programmatic” systems (R, etc.)

Con:

Very steep learning curve

Getting around limitations can be a lot of work

Can be a huge time-sink (especially at first!)

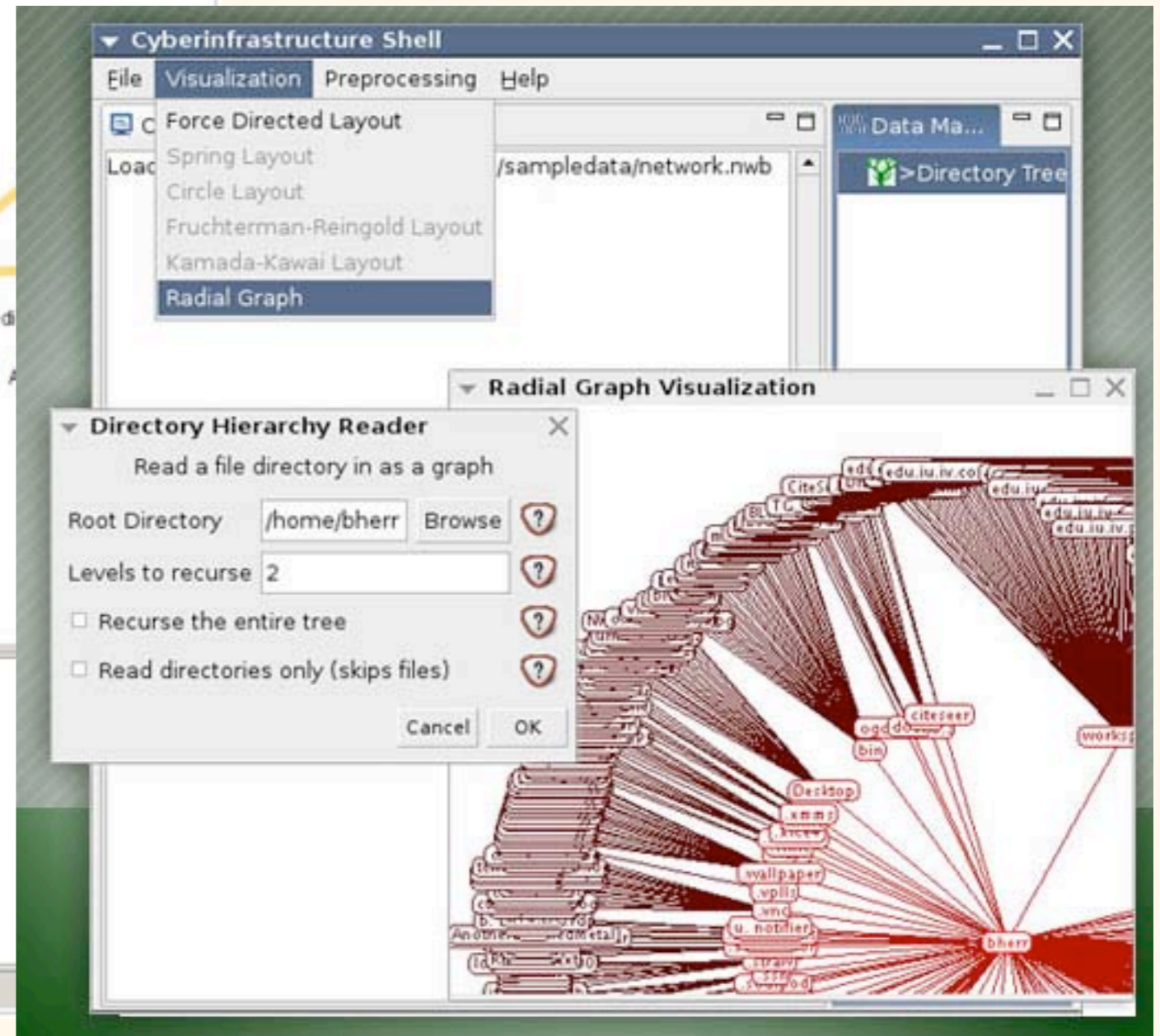
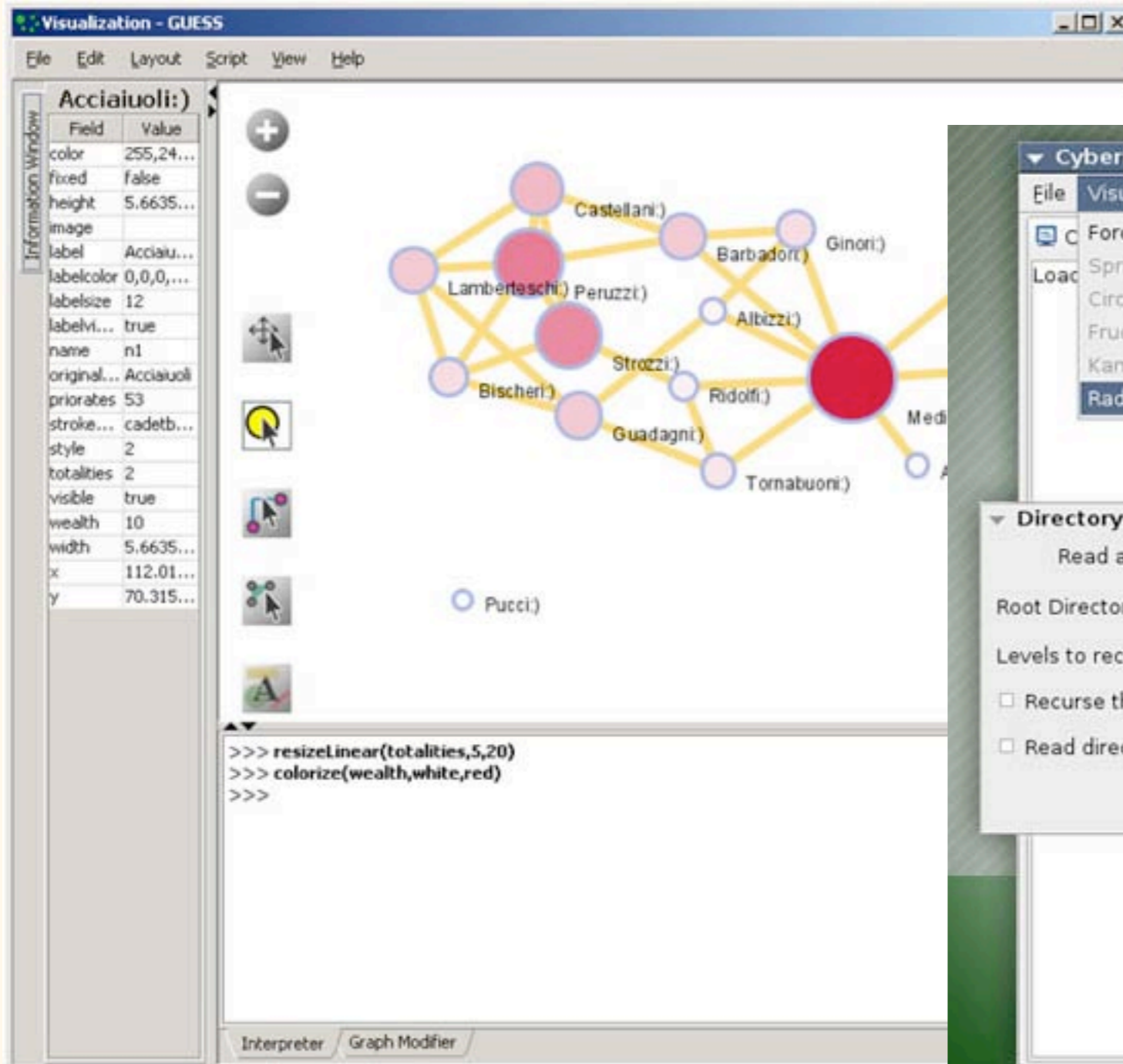
Neutral:

May make collaboration easier or harder, depending.

1. Data analysis software that also does visualization;
2. Dedicated visualization packages.

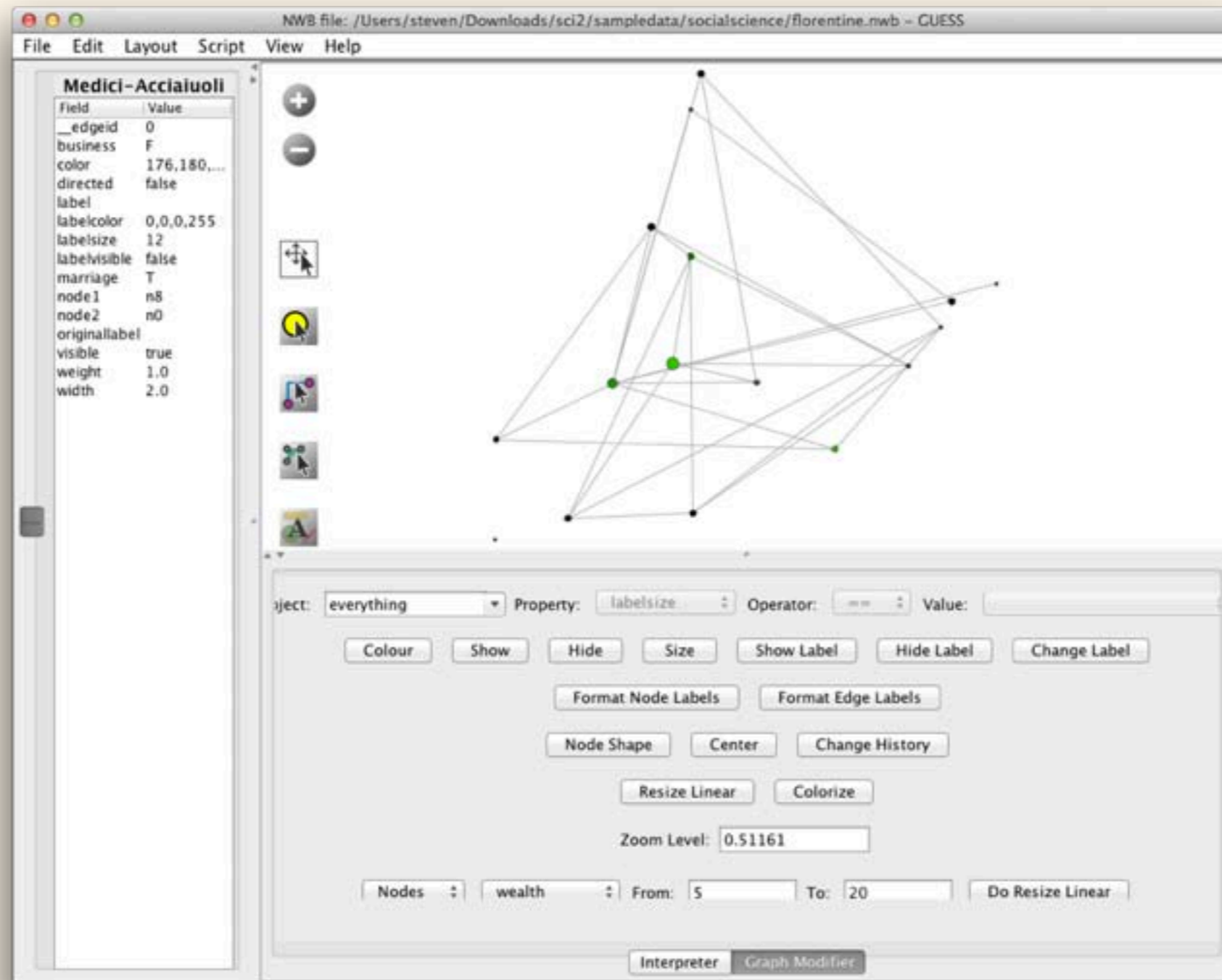
## 2. Dedicated visualization packages.

Sci2: network diagrams, simple maps, etc.



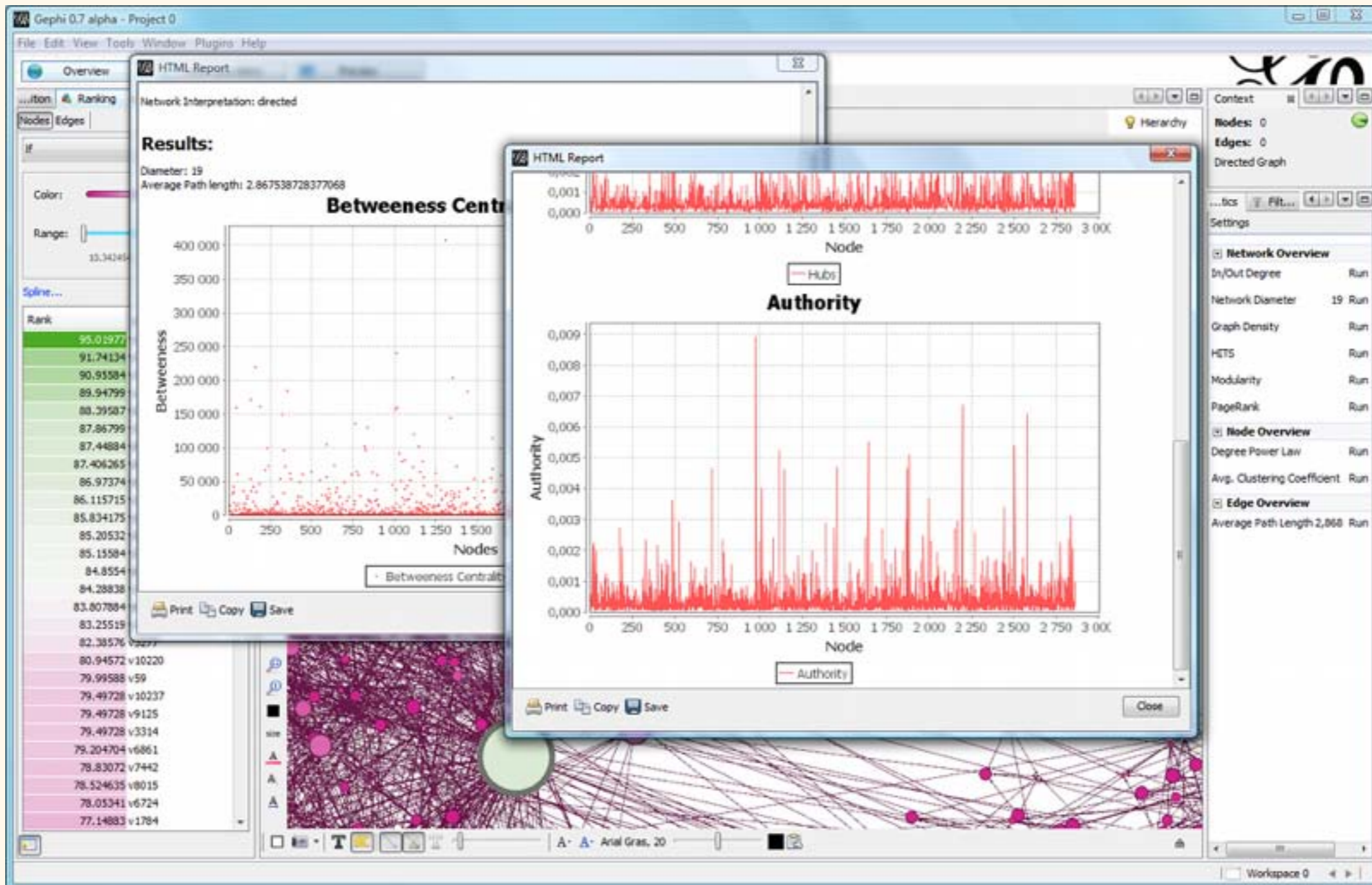
## 2. Dedicated visualization packages.

Sci2: network diagrams, simple maps, etc.



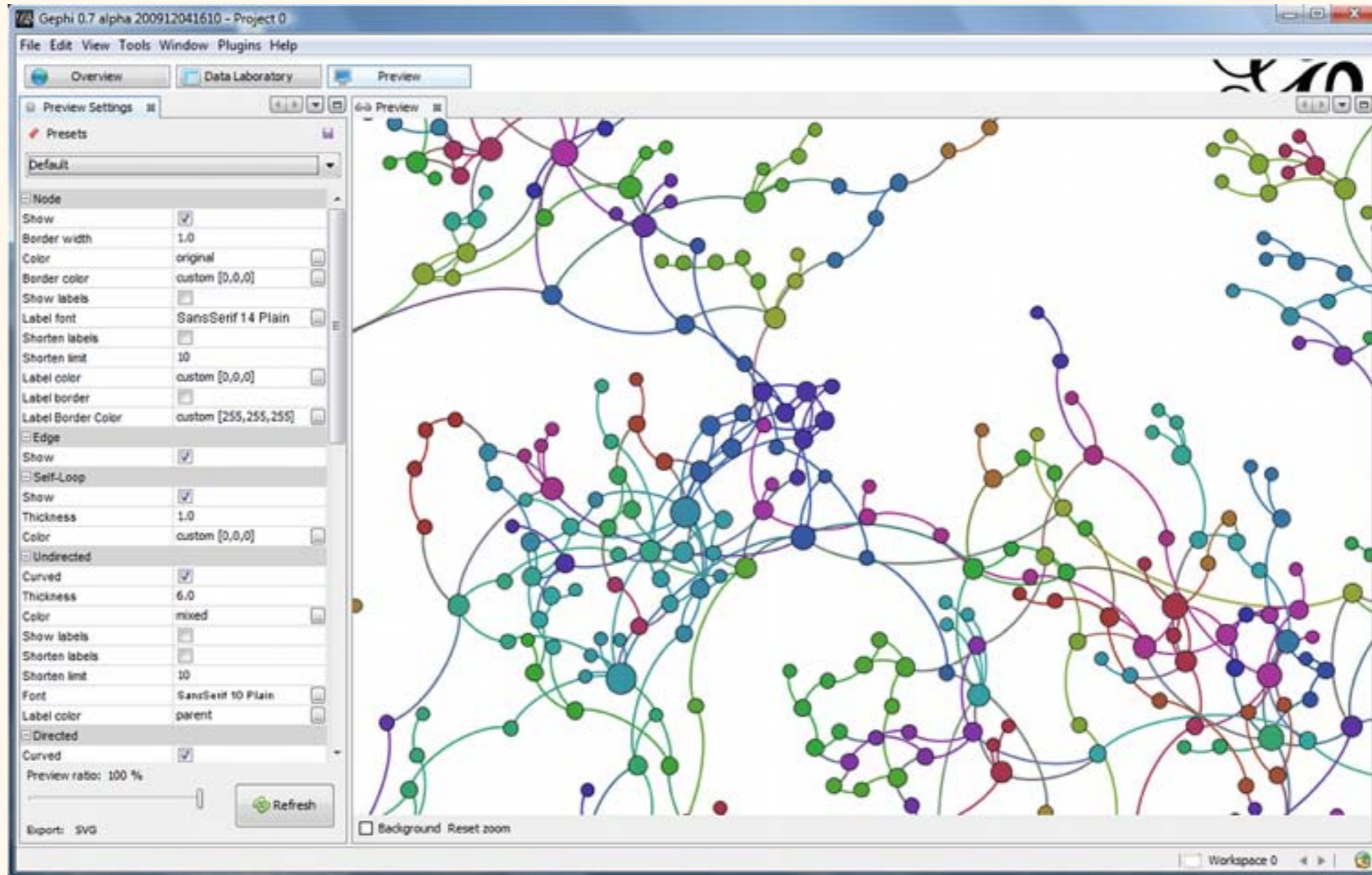
## 2. Dedicated visualization packages.

# Gephi: network diagrams



## 2. Dedicated visualization packages.

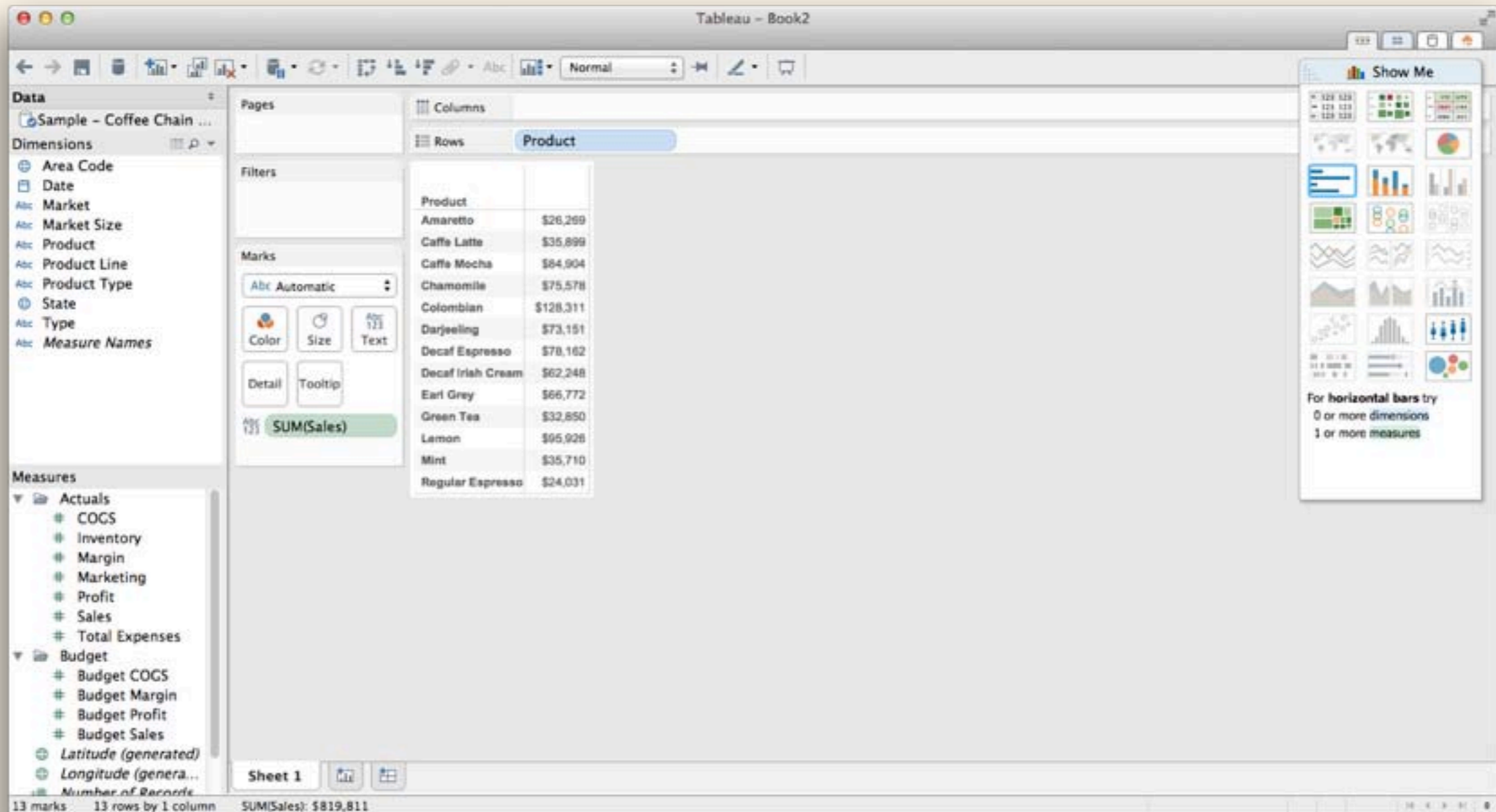
# Gephi: network diagrams





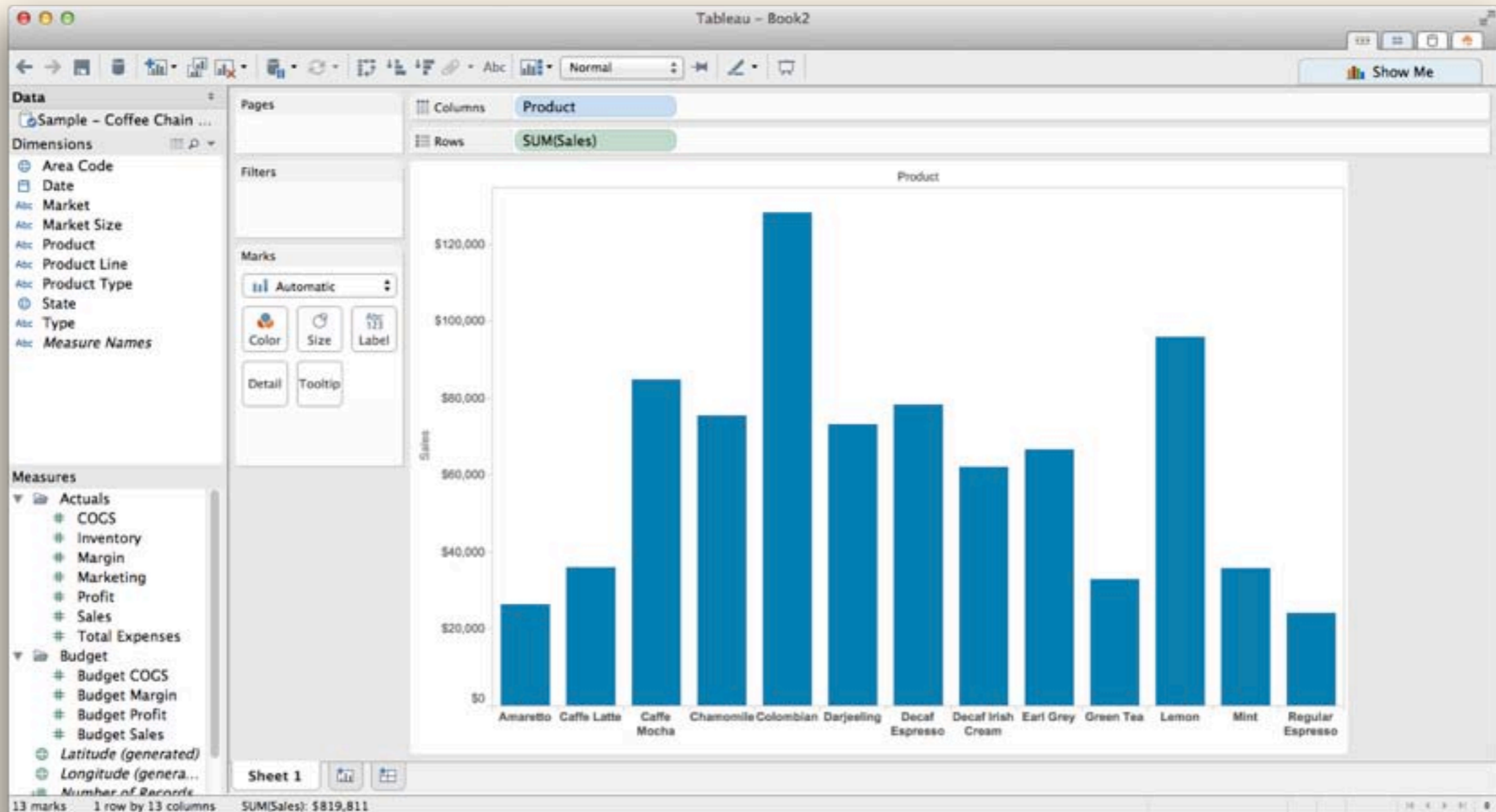
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



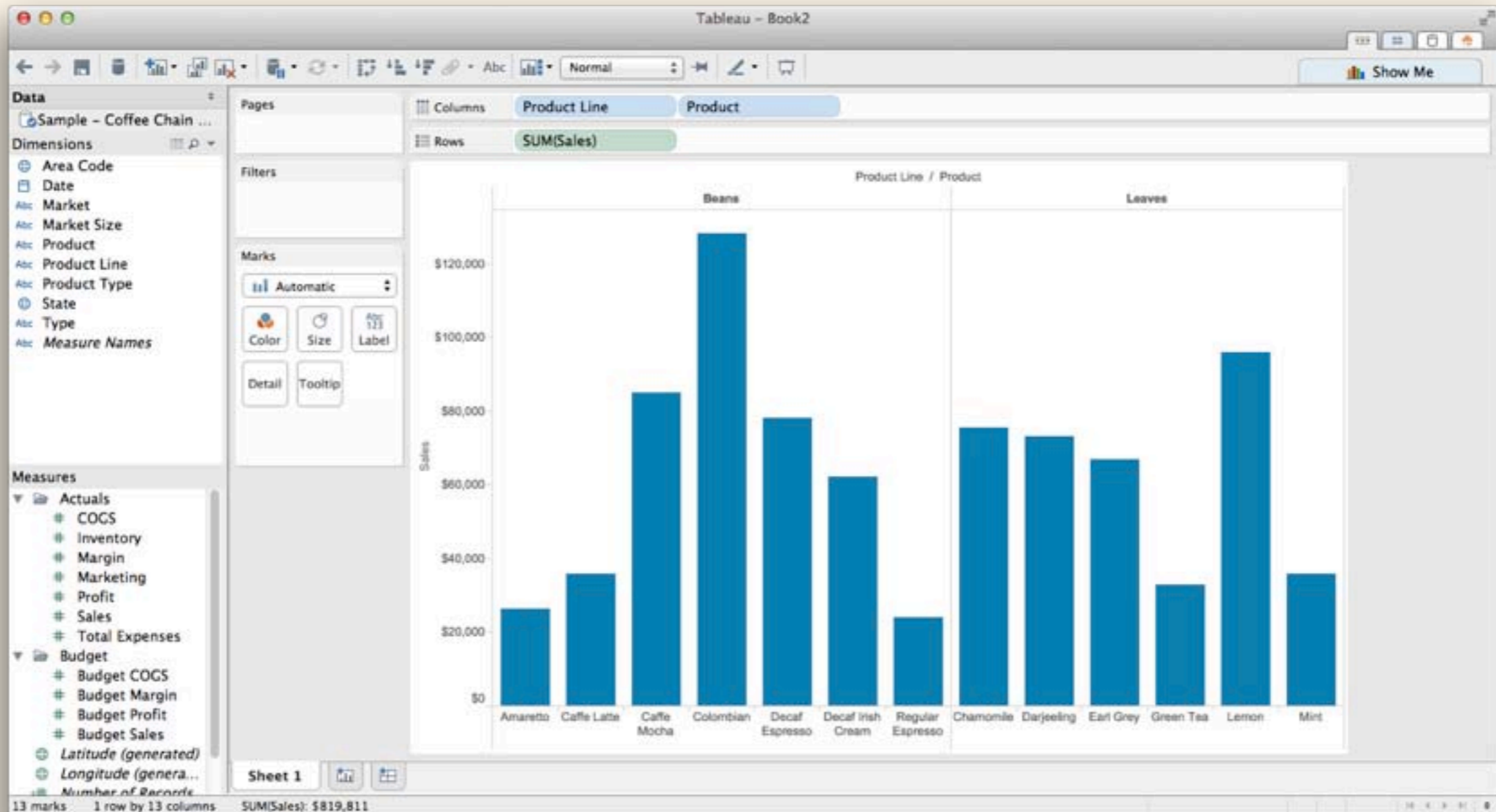
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



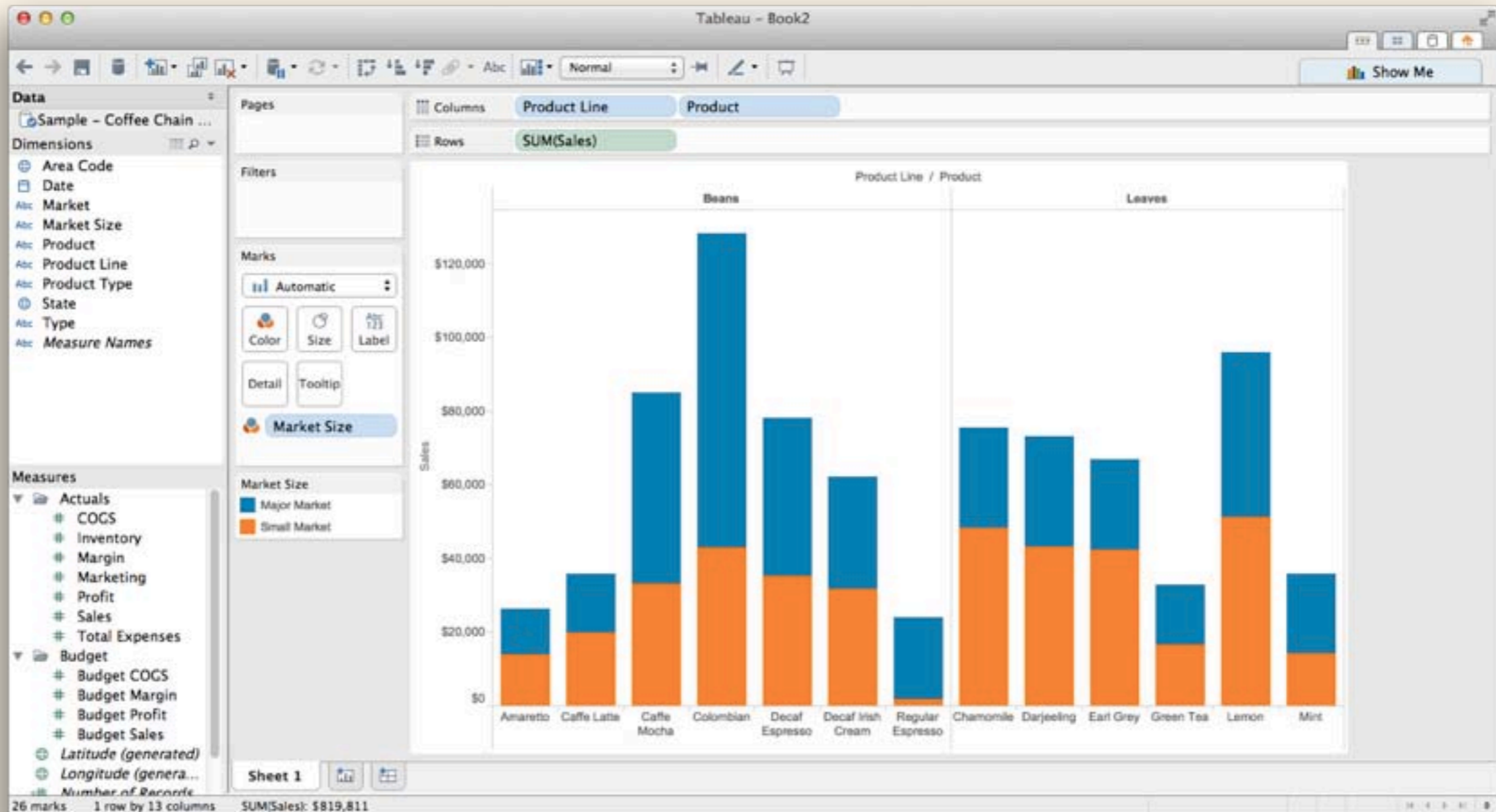
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



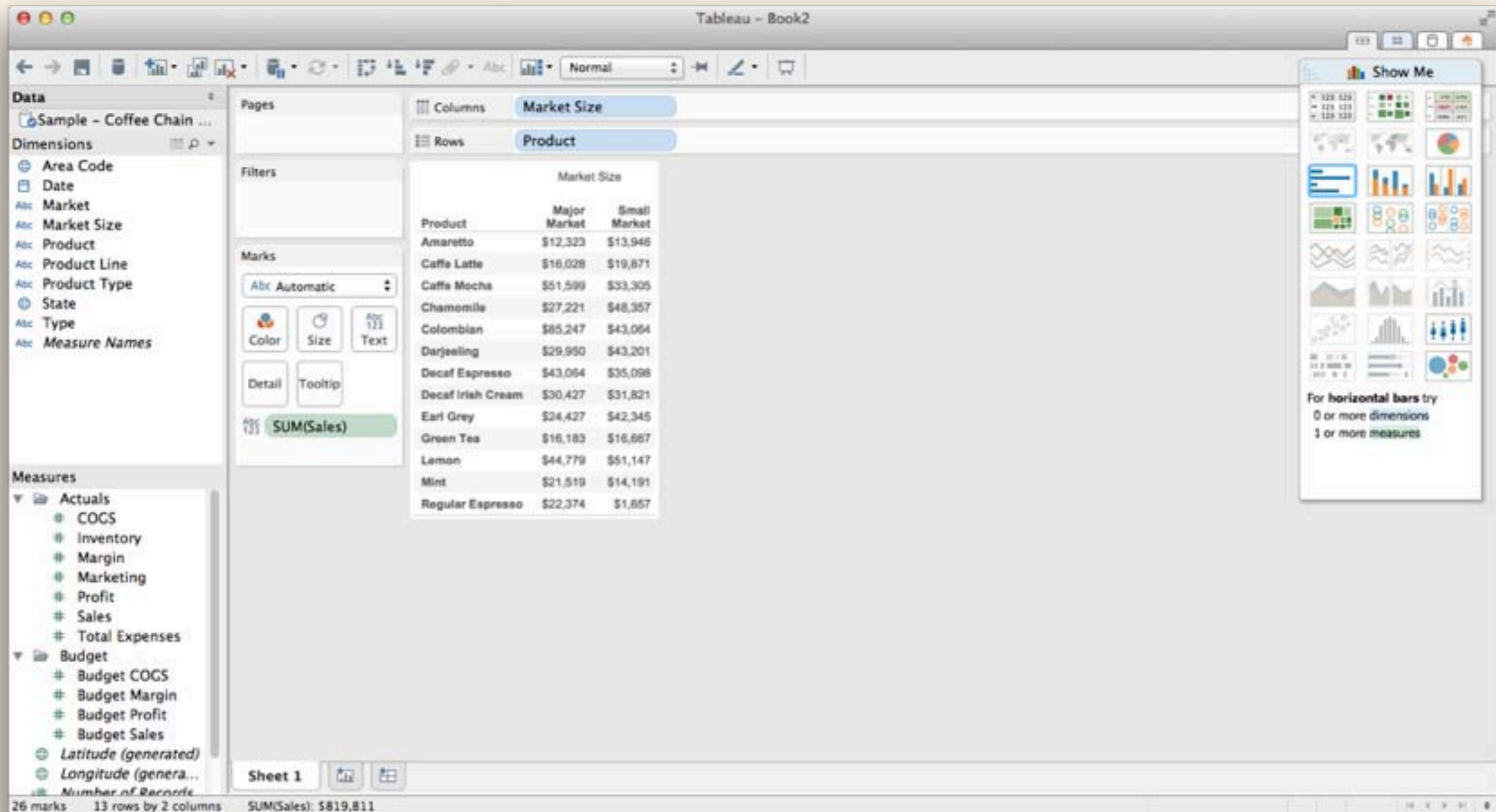
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



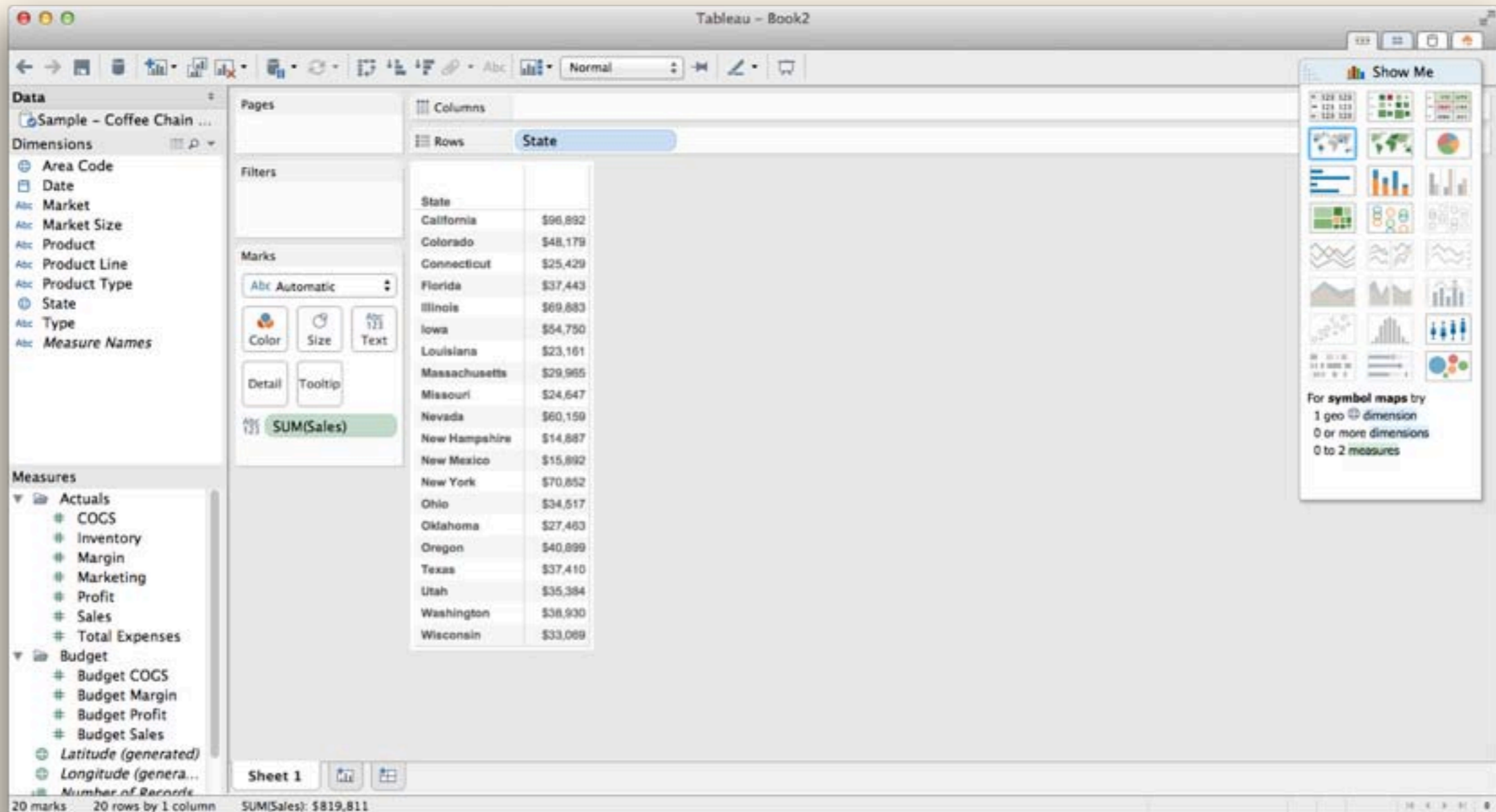
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



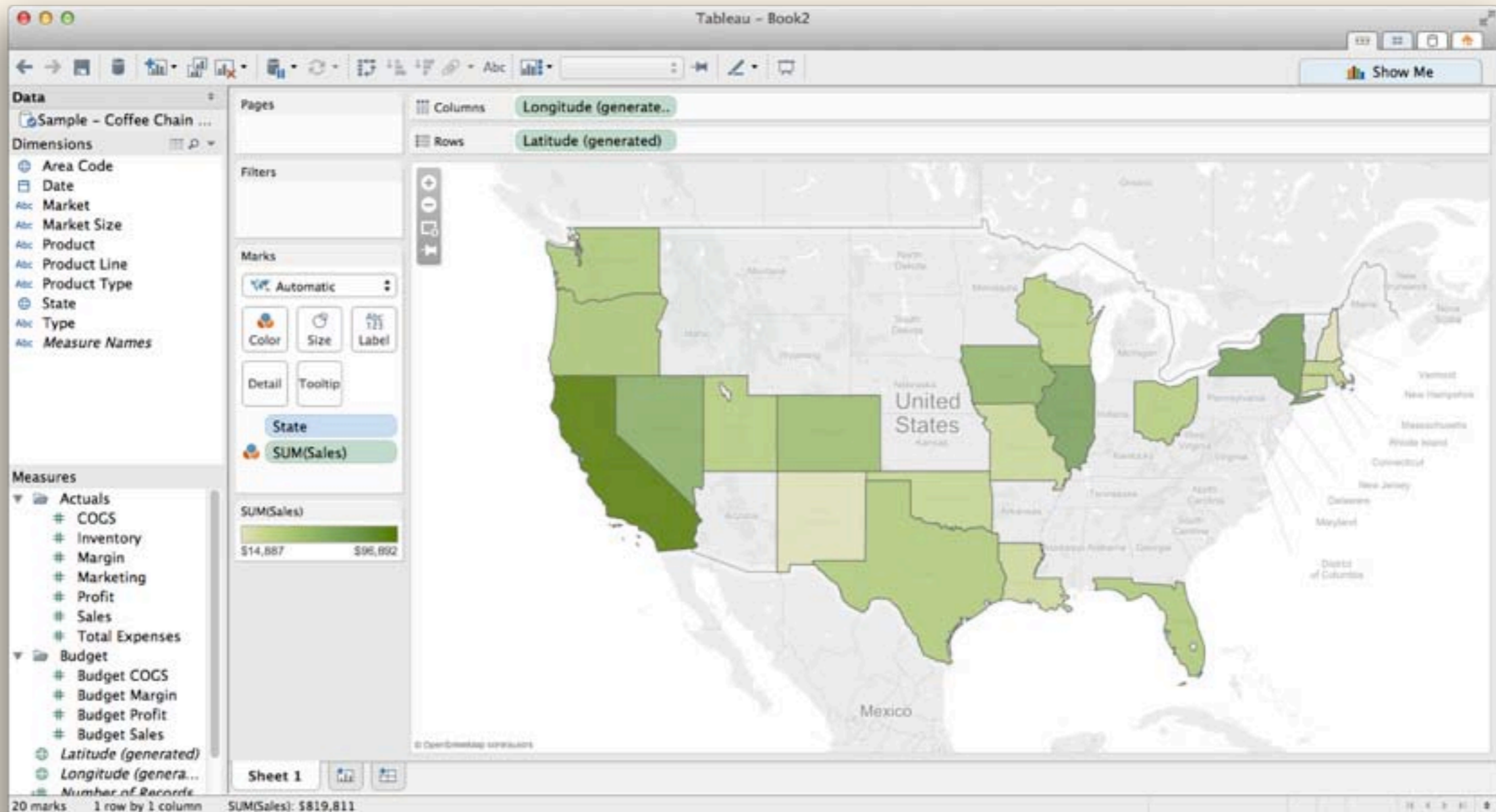
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



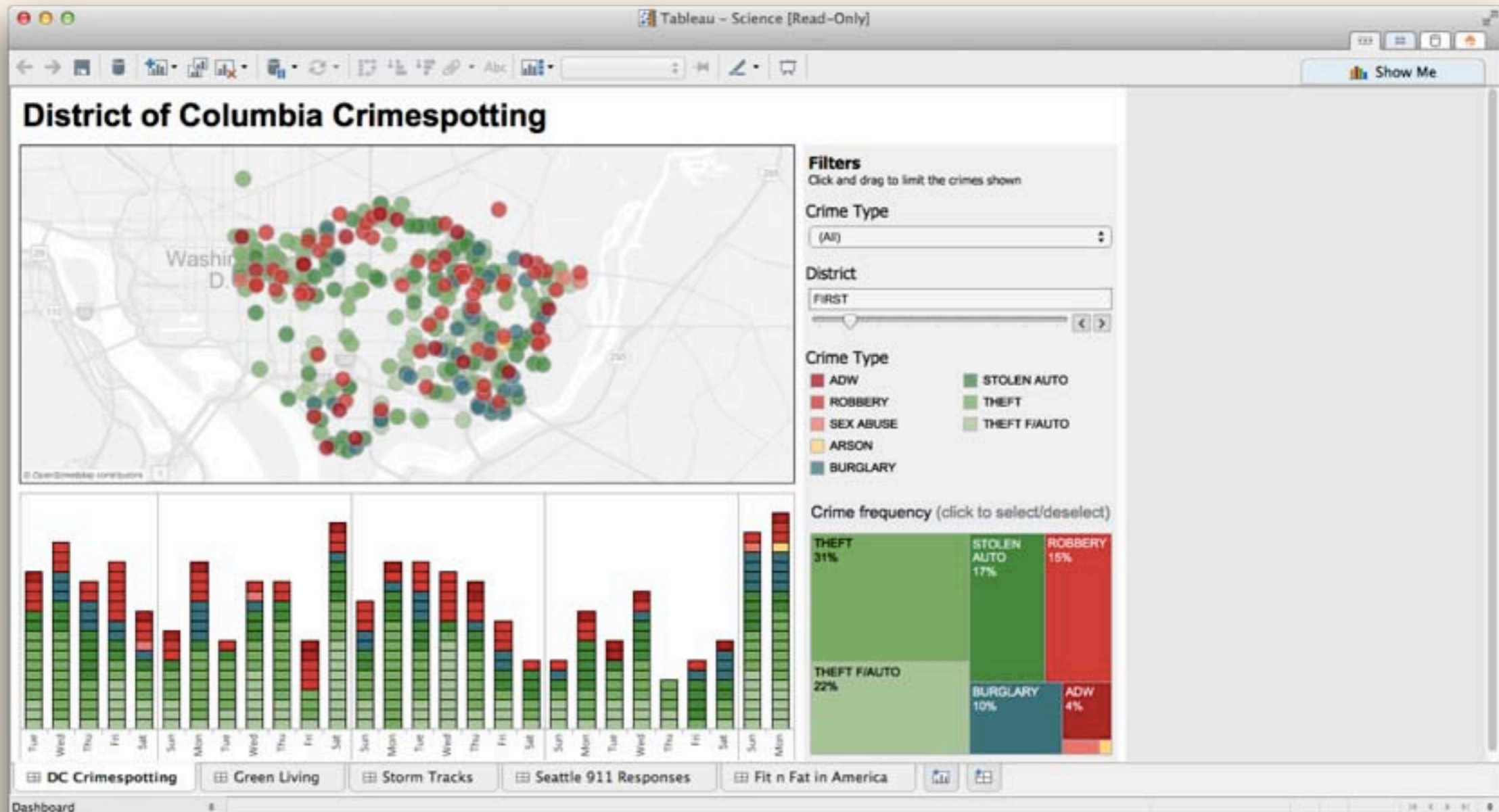
## 2. Dedicated visualization packages.

### Tableau: the kitchen sink



## 2. Dedicated visualization packages.

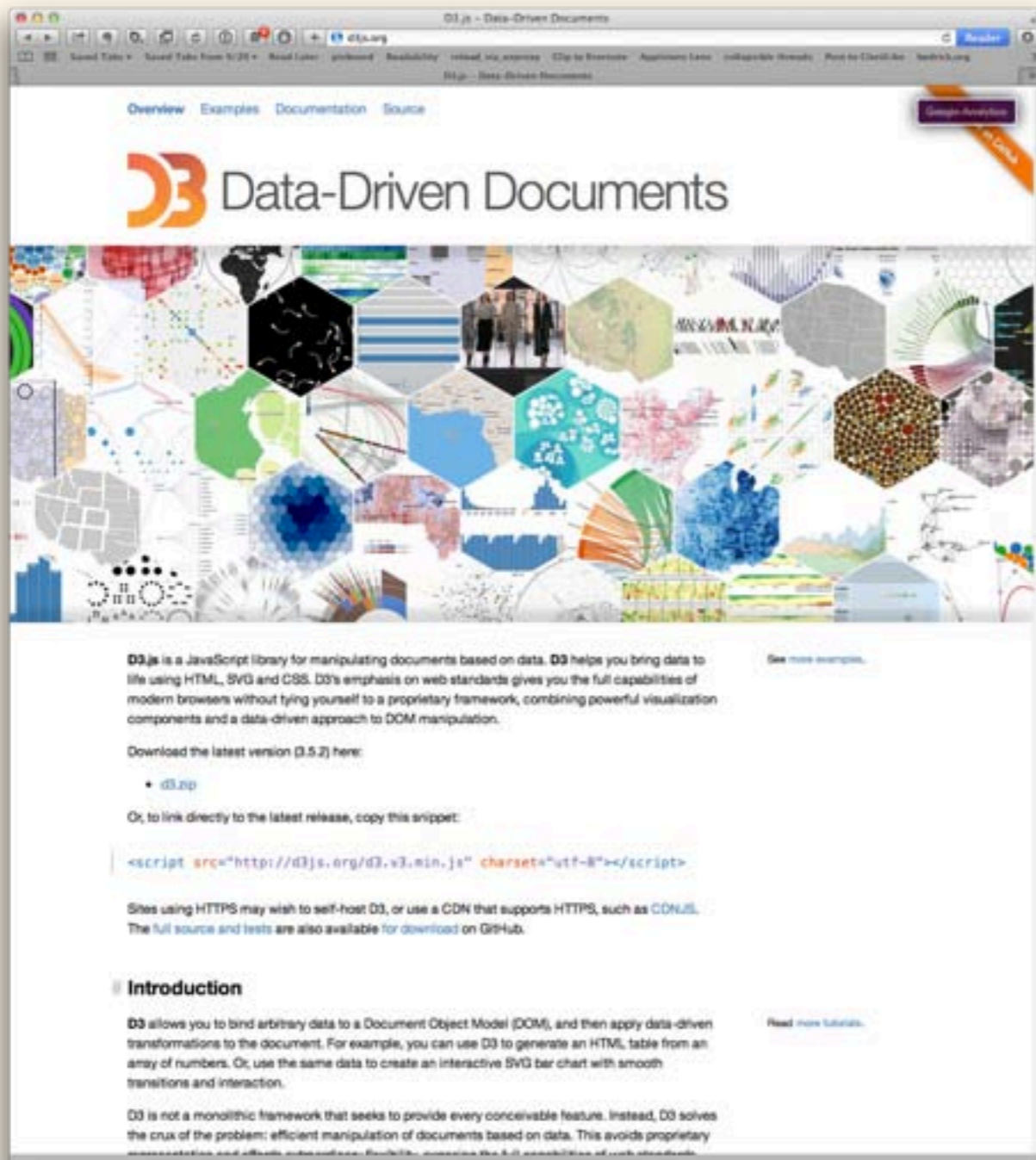
### Tableau: the kitchen sink



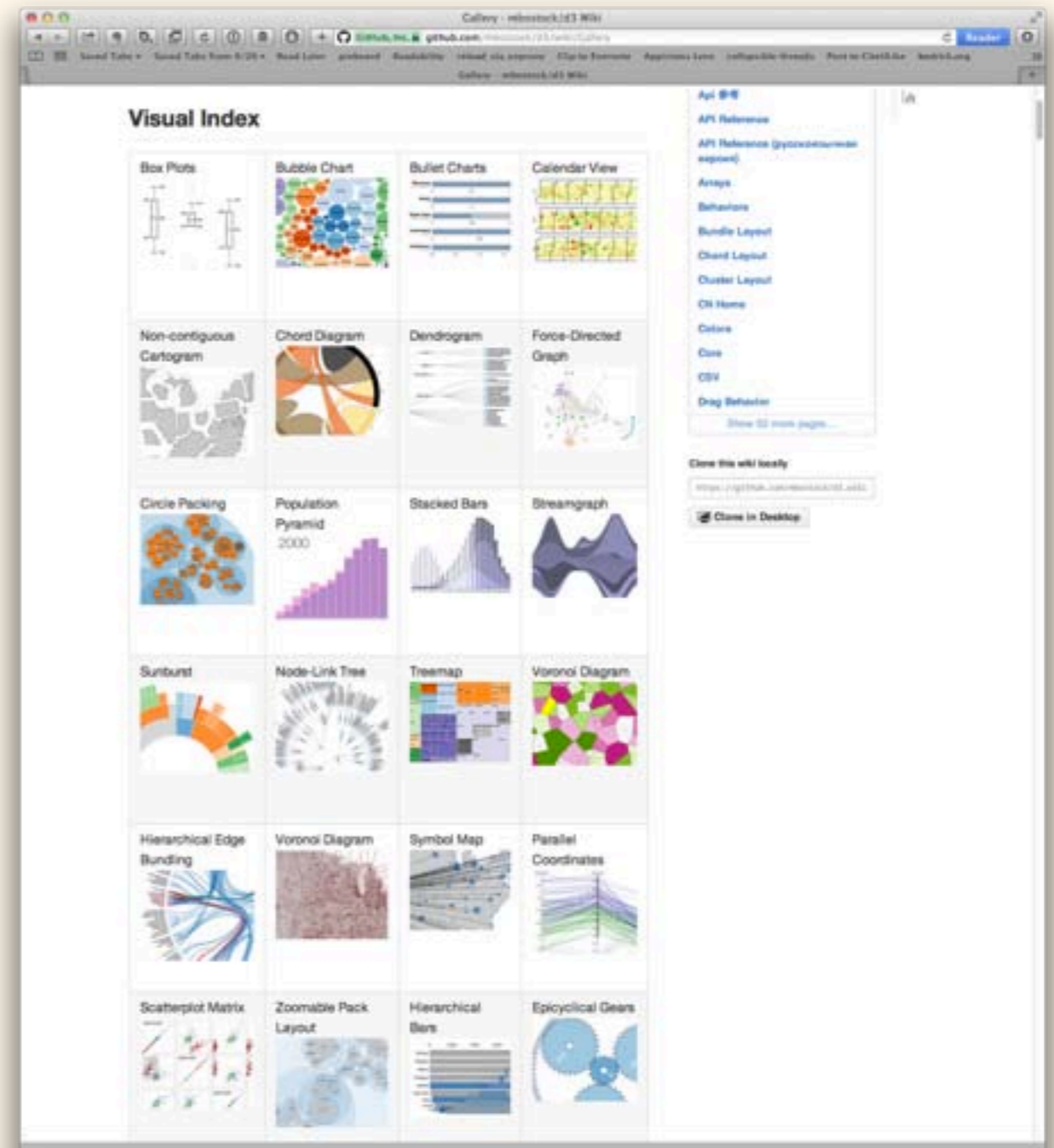


## 2. Dedicated visualization packages.

d3: Great for interactive graphics...



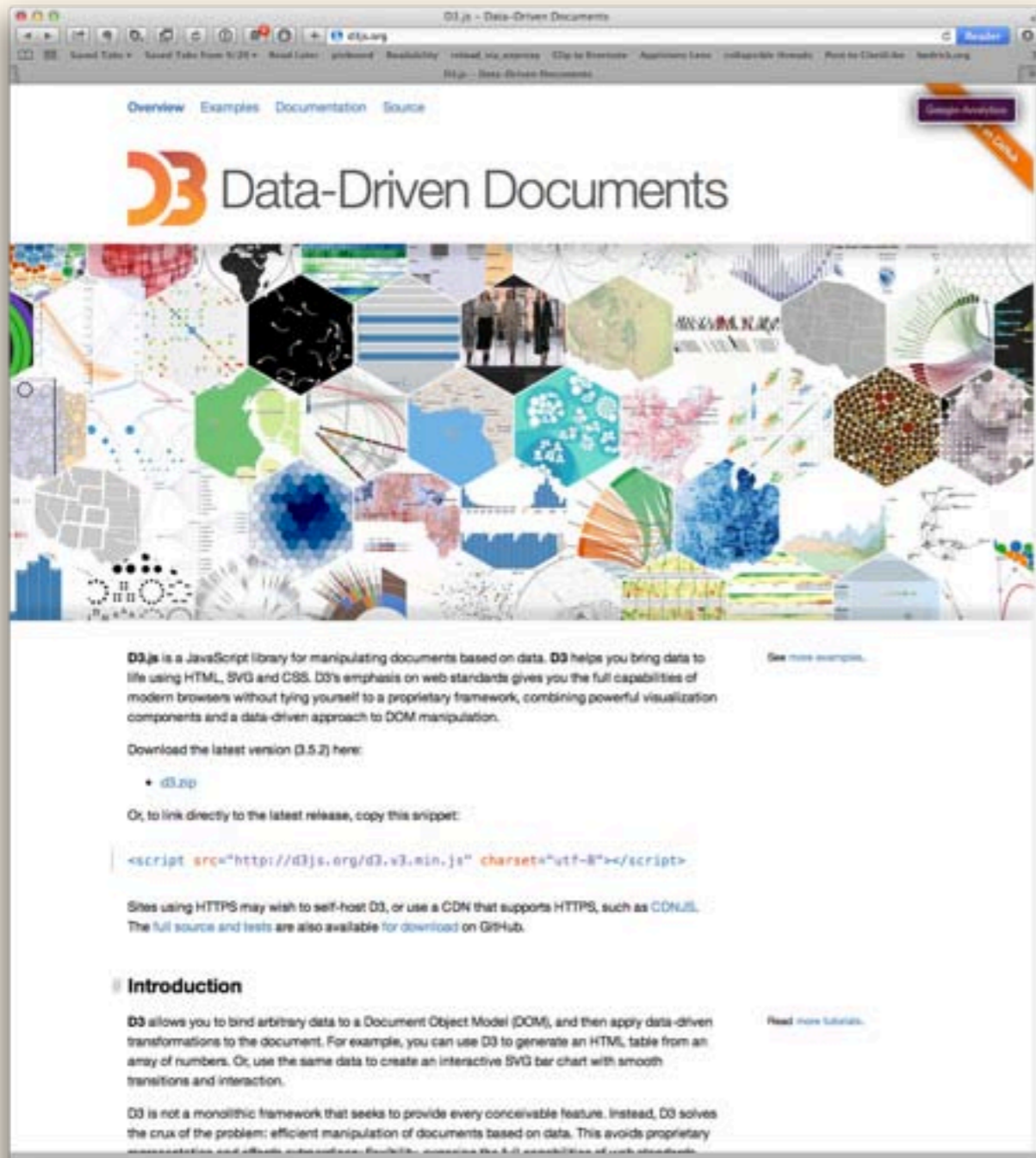
The screenshot shows the homepage of the d3.js library. At the top, there is a navigation bar with links for "Overview", "Examples", "Documentation", and "Source". Below this is the d3.js logo and the text "Data-Driven Documents". A large, colorful collage of various data visualizations, including maps, charts, and diagrams, is displayed in the center. Below the collage, there is a section titled "Introduction" which describes the library's purpose and provides a code snippet for including the library. The text reads: "d3.js is a JavaScript library for manipulating documents based on data. d3 helps you bring data to life using HTML, SVG and CSS. D3's emphasis on web standards gives you the full capabilities of modern browsers without tying yourself to a proprietary framework, combining powerful visualization components and a data-driven approach to DOM manipulation." It also provides a link to download the latest version (3.5.2) and a code snippet: `<script src="http://d3js.org/d3.v3.min.js" charset="utf-8"></script>`. There are also links for "See more examples" and "Read more tutorials".



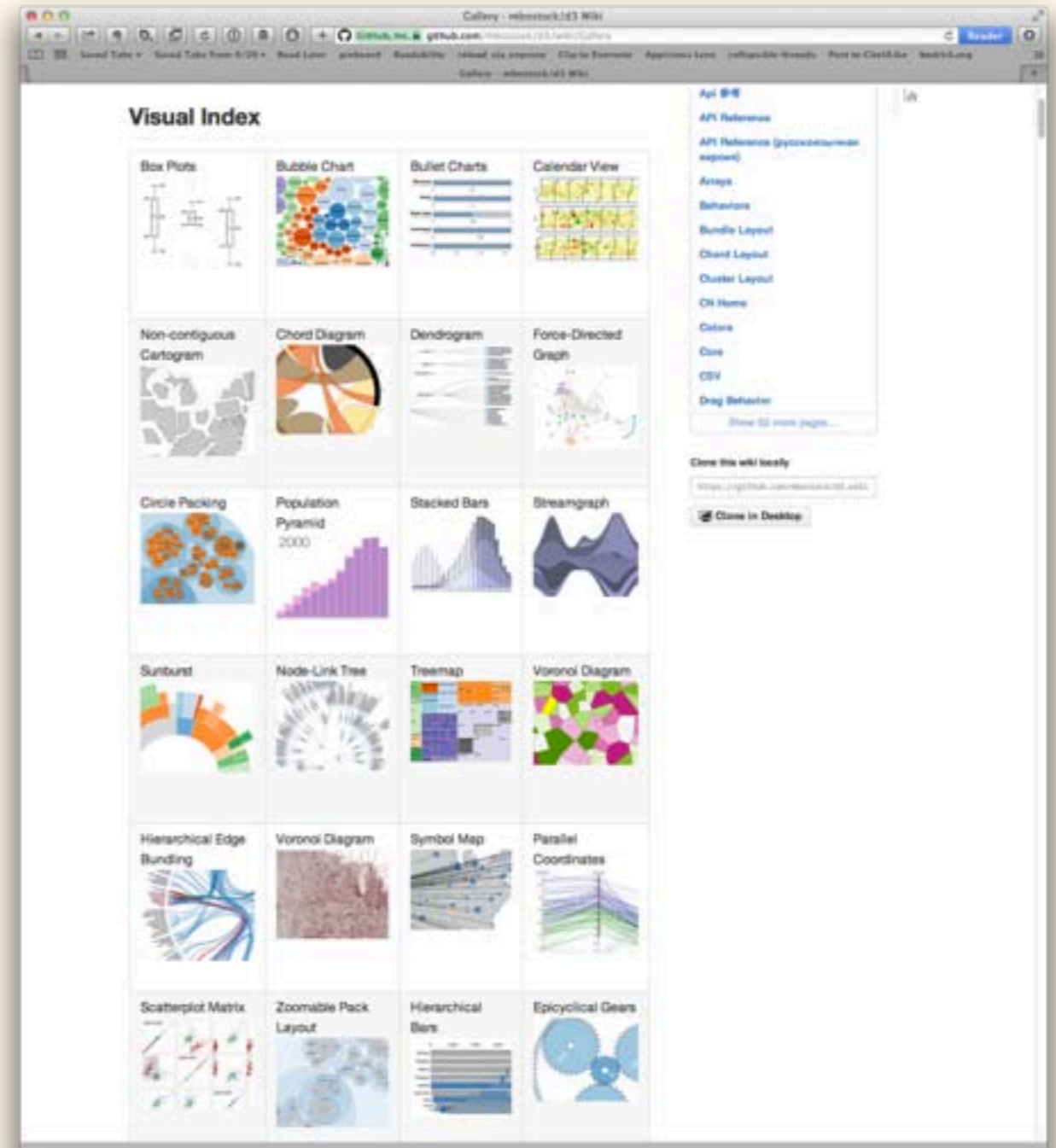
The screenshot shows the "Visual Index" page of the d3.js library. It features a grid of 20 small thumbnail images, each representing a different visualization type. The thumbnails are arranged in a 5x4 grid. The types of visualizations include: Box Plots, Bubble Chart, Bullet Charts, Calendar View, Non-contiguous Cartogram, Chord Diagram, Dendrogram, Force-Directed Graph, Circle Packing, Population Pyramid 2000, Stacked Bars, Streamgraph, Sunburst, Node-Link Tree, Treemap, Voronoi Diagram, Hierarchical Edge Bundling, Voronoi Diagram, Symbol Map, Parallel Coordinates, Scatterplot Matrix, Zoomable Pack Layout, Hierarchical Bars, and Epicyclical Gears. On the right side of the page, there is a sidebar with a search bar and a list of navigation links, including "API Reference", "API Reference (previous version)", "Arrays", "Behaviors", "Bundle Layout", "Chord Layout", "Cluster Layout", "D3 Home", "Colors", "Cores", "CSV", "Drag Behavior", and "Show 52 more pages...". There is also a "Clone this wiki locally" section with a "Clone in Desktop" button.

## 2. Dedicated visualization packages.

d3: ... but very programming-intensive.



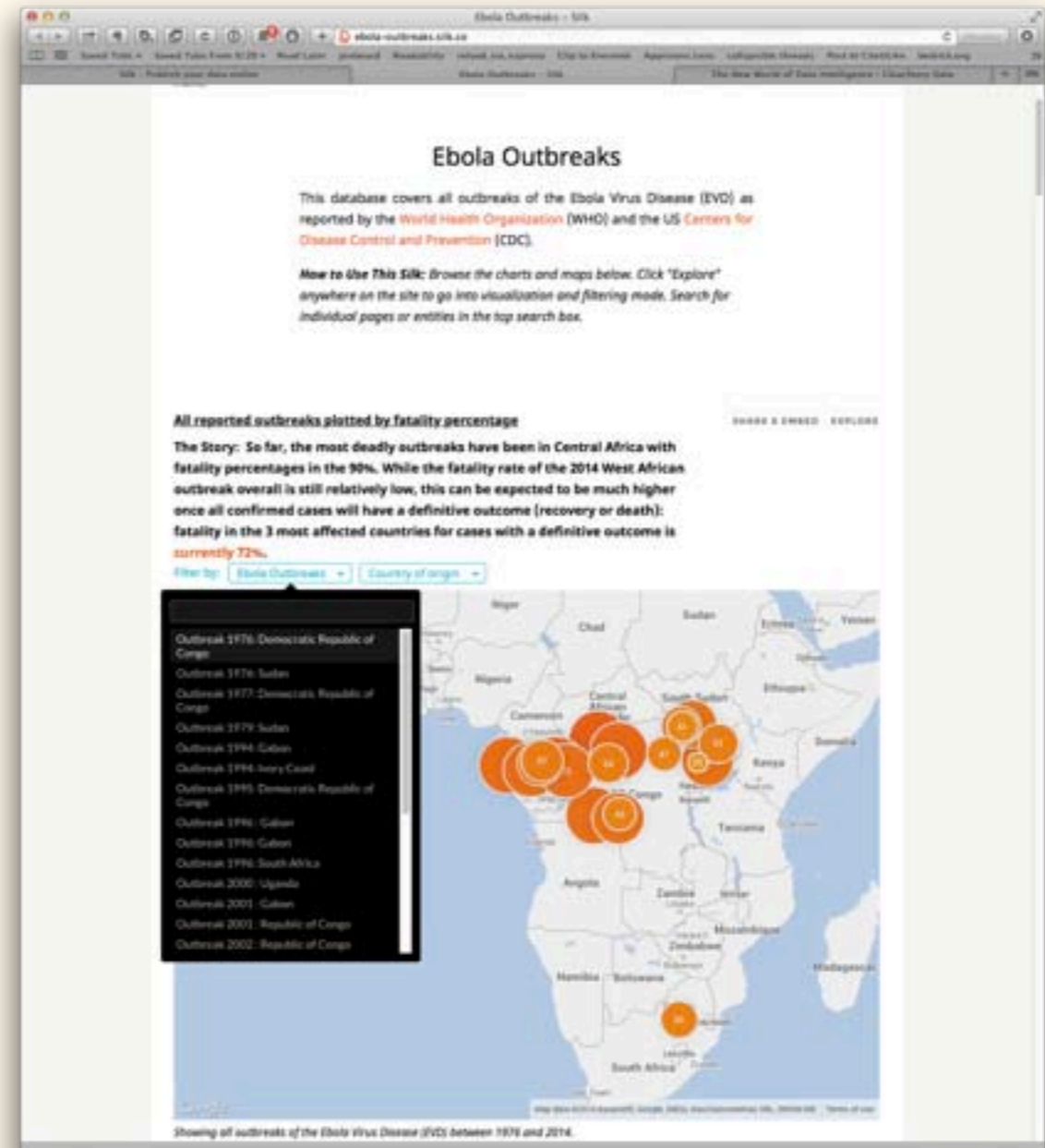
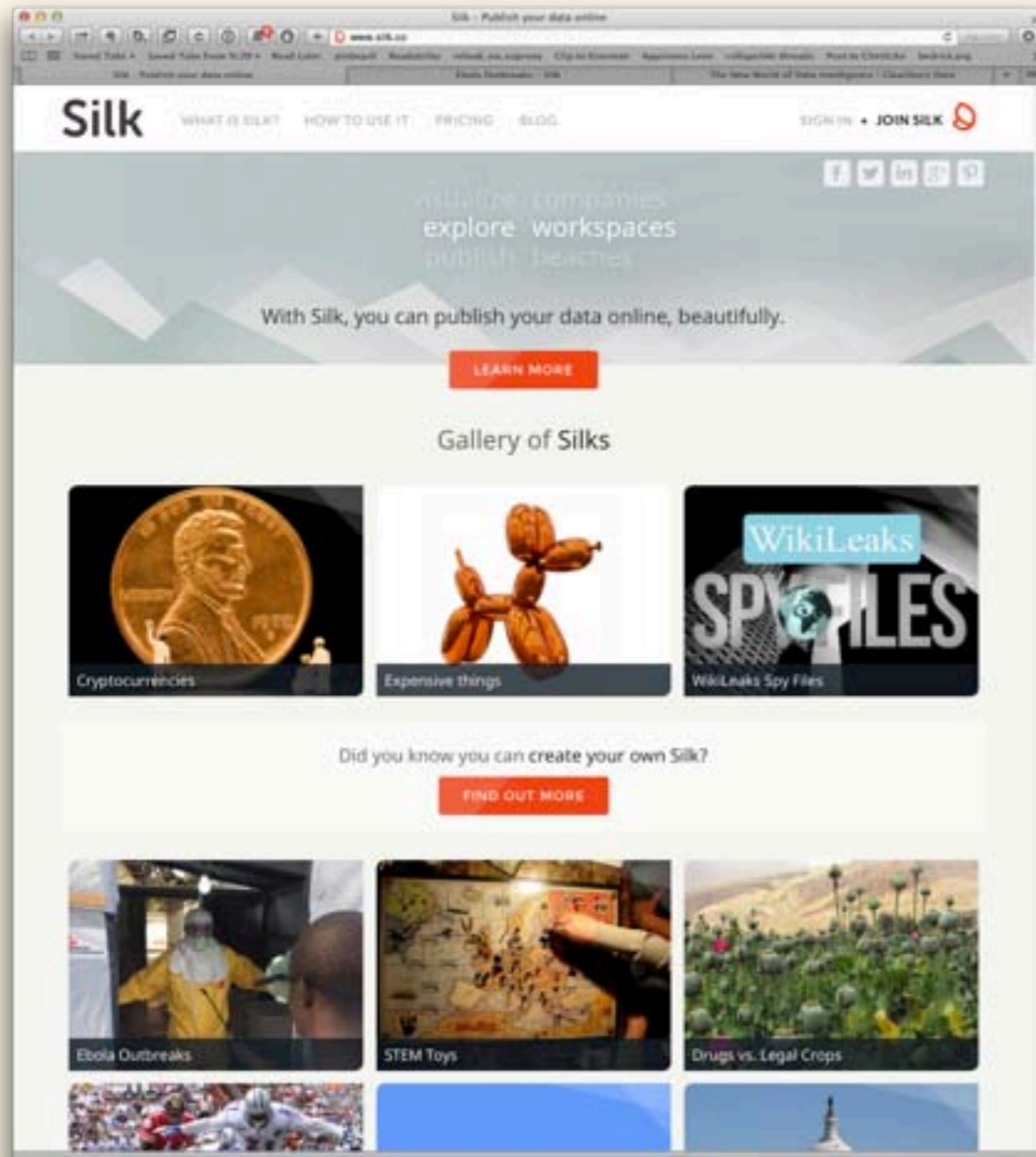
The screenshot shows the homepage of the d3.js library. At the top, there are navigation links for "Overview", "Examples", "Documentation", and "Source". The main heading is "Data-Driven Documents" with a large orange "d3" logo. Below the heading is a collage of various data visualizations, including bar charts, pie charts, and network graphs. A purple "Download" button is visible in the top right corner. The main text describes d3.js as a JavaScript library for manipulating documents based on data, emphasizing its use of HTML, SVG, and CSS. It provides instructions on how to download the latest version (3.5.2) and includes a code snippet for linking to the library via a script tag. The page also features an "Introduction" section and a "Read more tutorials" link.



The screenshot shows the "Visual Index" page of the d3.js library. The page is titled "Visual Index" and features a grid of 20 small thumbnail images, each representing a different visualization type. The thumbnails are arranged in a 5x4 grid. The visualization types include: Box Plots, Bubble Chart, Bullet Charts, Calendar View, Non-contiguous Cartogram, Chord Diagram, Dendrogram, Force-Directed Graph, Circle Packing, Population Pyramid 2000, Stacked Bars, Streamgraph, Sunburst, Node-Link Tree, Treemap, Voronoi Diagram, Hierarchical Edge Bundling, Voronoi Diagram, Symbol Map, Parallel Coordinates, Scatterplot Matrix, Zoomable Pack Layout, Hierarchical Bars, and Epicyclical Gears. On the right side of the page, there is a sidebar with a search bar and a list of navigation links, including "API Reference", "API Reference (previous version)", "Arrays", "Behaviors", "Bundle Layout", "Chord Layout", "Cluster Layout", "D3 Home", "Colors", "Cores", "CSV", "Drag Behavior", and "Show 52 more pages...". There is also a "Clone this wiki locally" section with a "Clone in Desktop" button.

## 2. Dedicated visualization packages.

“Cloud” solutions (Silk, etc.):



## 2. Dedicated visualization packages.

Notes:

Sci2 and Gephi are free!

Tableau is *not*!

However, it *is* free for students!

Important consideration: getting data out.

# Notes on workflow:

1. Where will your graphic end up?
2. External tools
3. IP & Security
4. Change the defaults!

Notes on workflow:

Where will your figure end up?

In print?

In a journal?

On a poster?

On a screen?

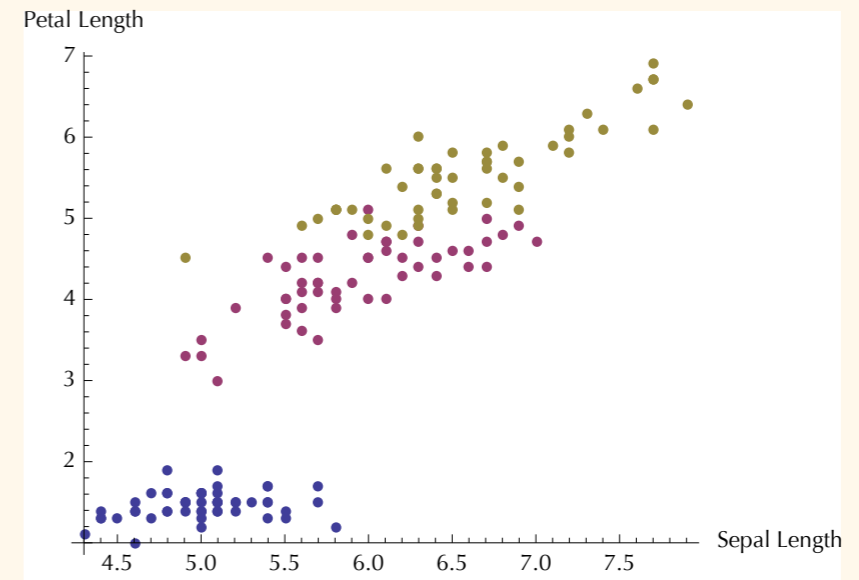
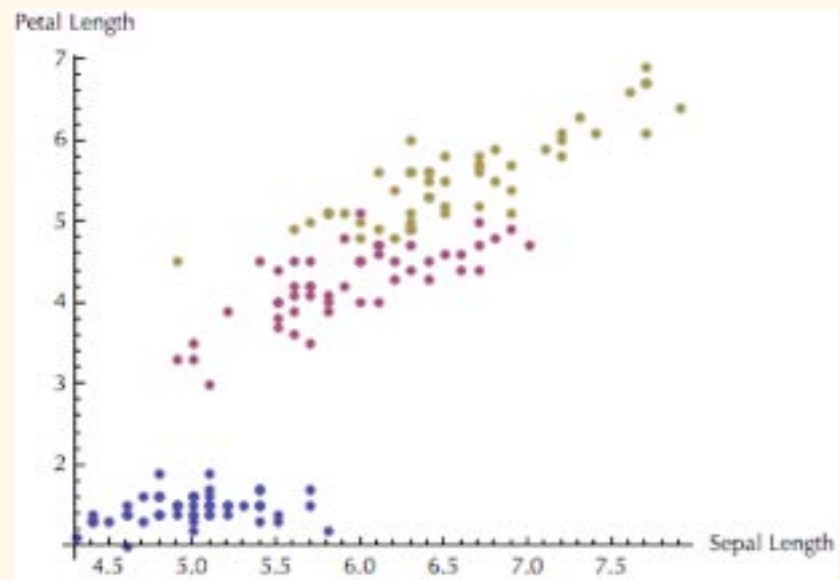
Online?

In a presentation?

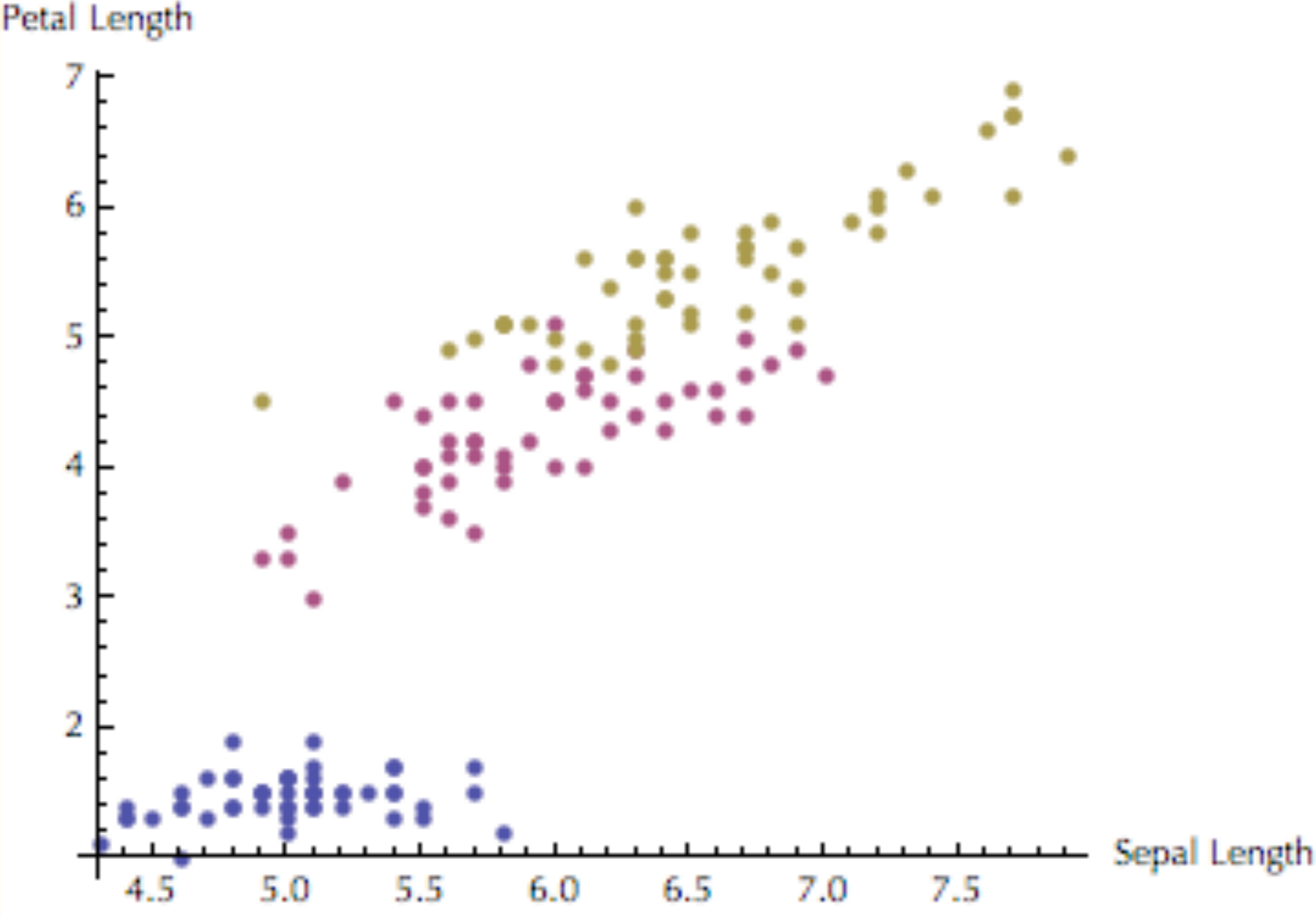
Will it be static...

... or interactive?

# Vector vs. Raster images:

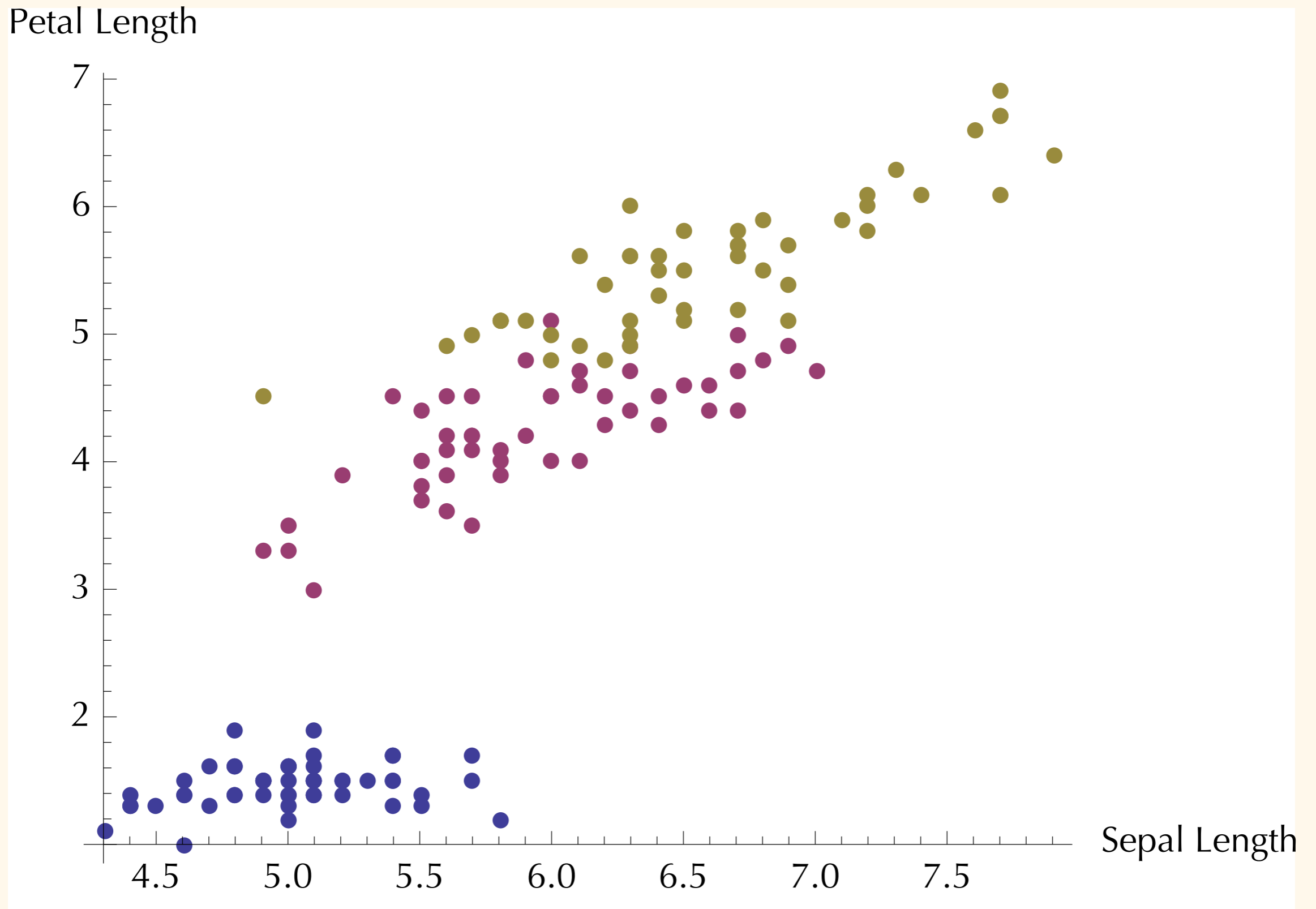


# Vector vs. Raster images:



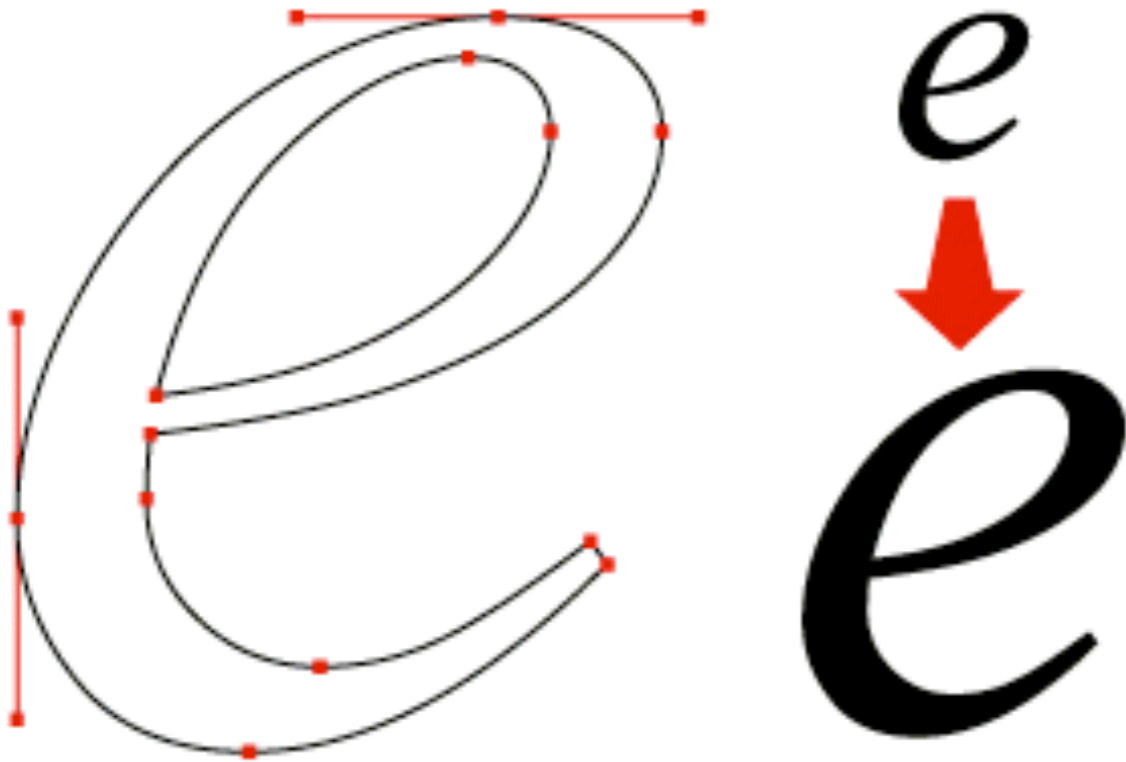


# Vector vs. Raster images:

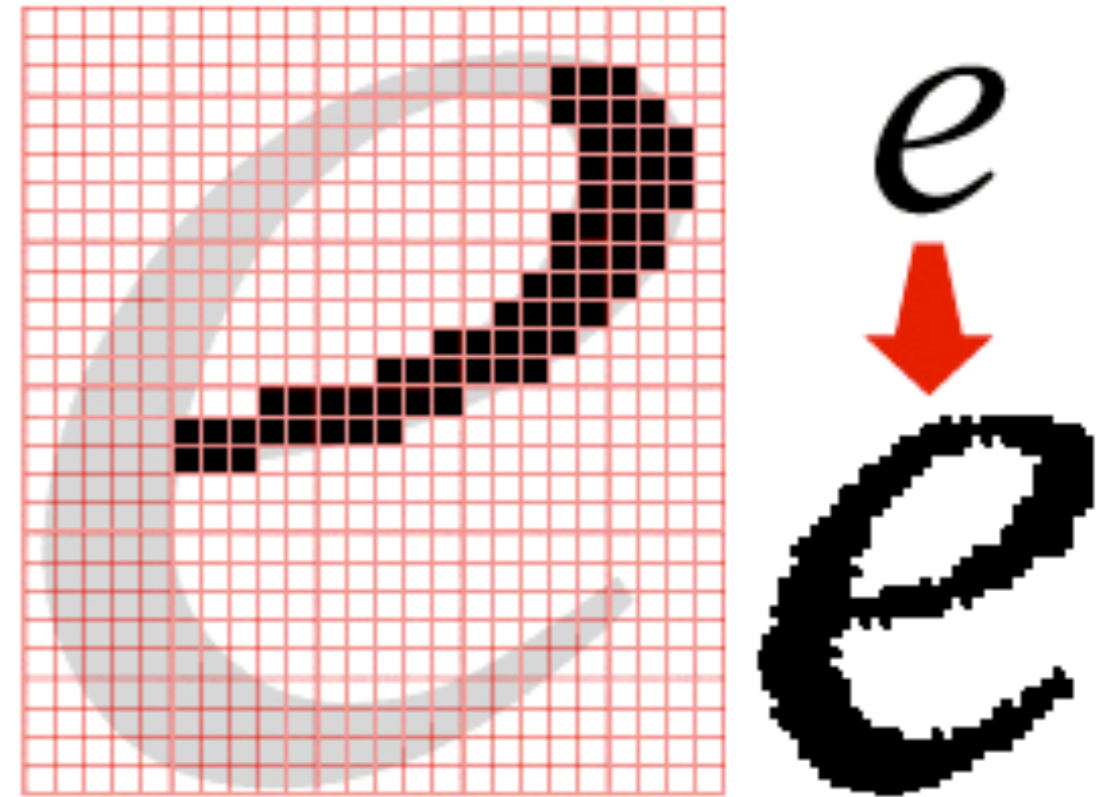


# Vector vs. Raster images:

VECTOR GRAPHICS



BITMAPMED (RASTER) GRAPHICS



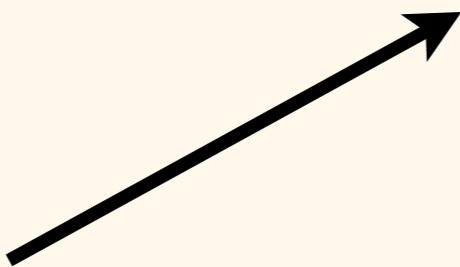
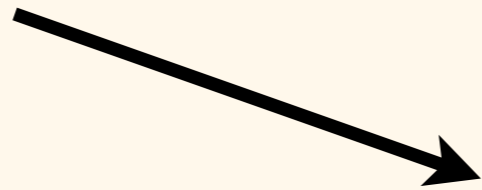
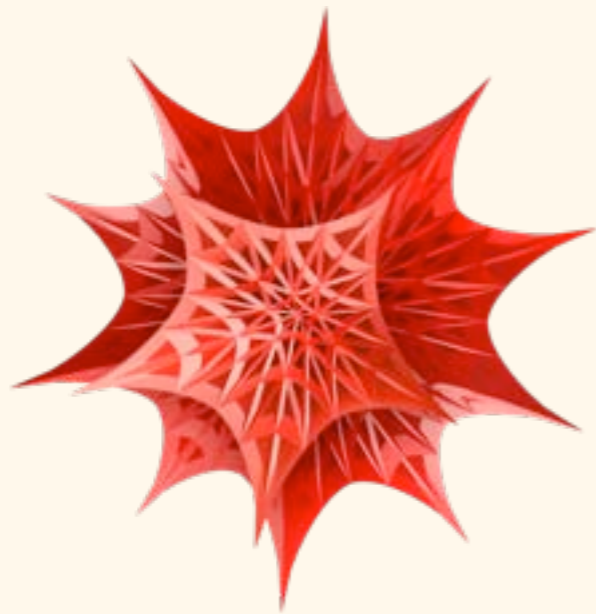
The upshot:

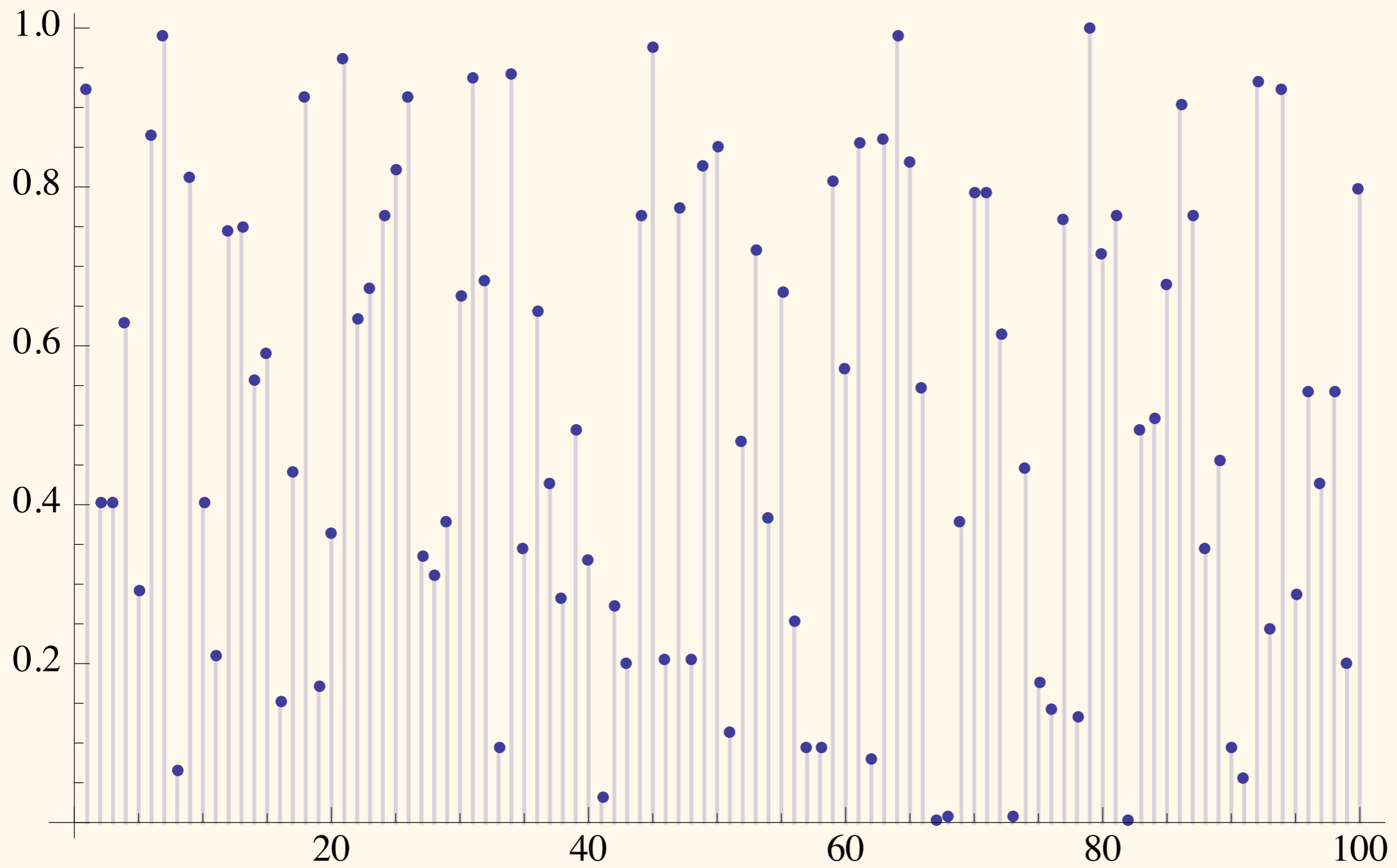
When possible, generate vector images...

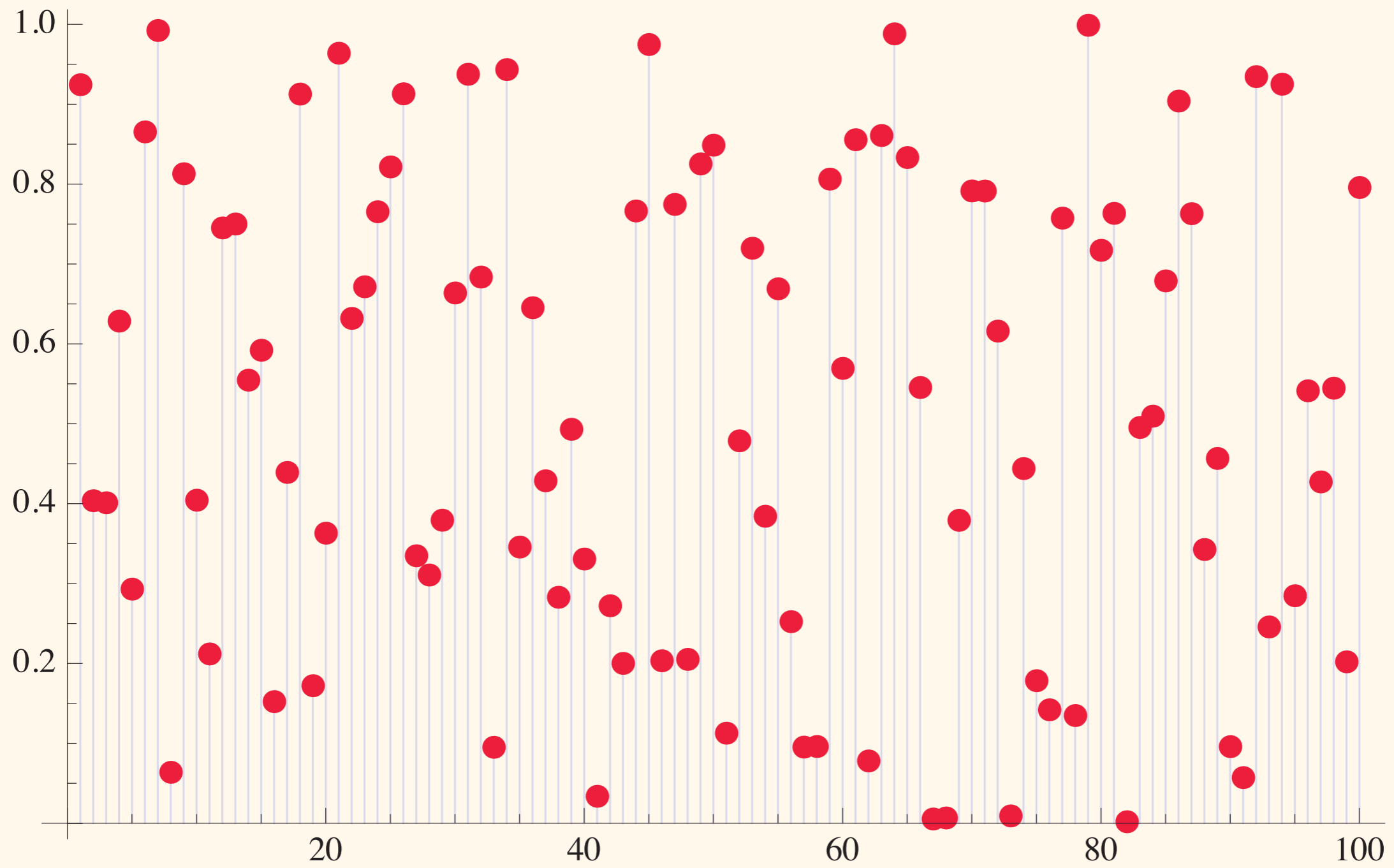
... especially for any print application!

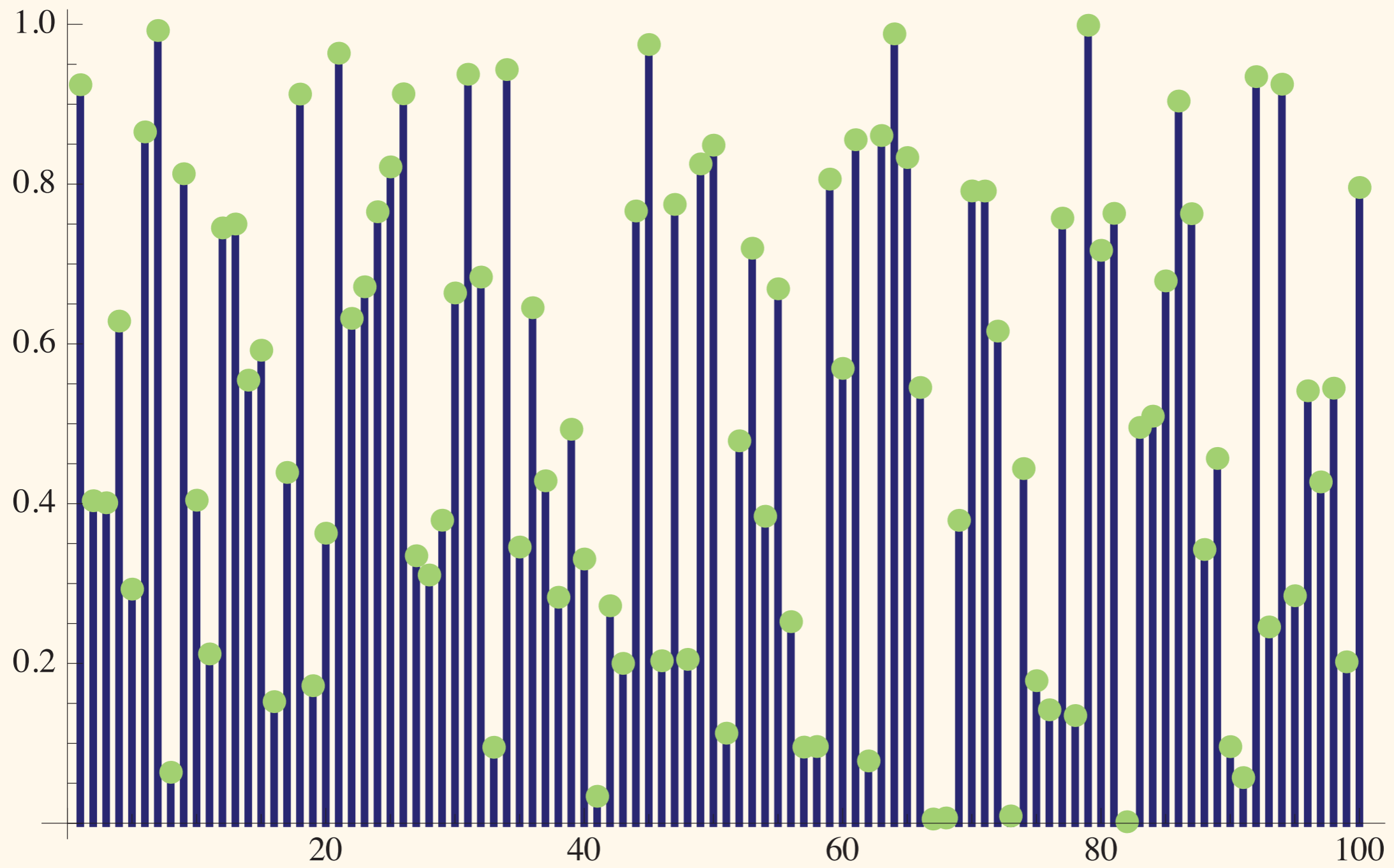
Most visualization tools (including Excel!) can do this.

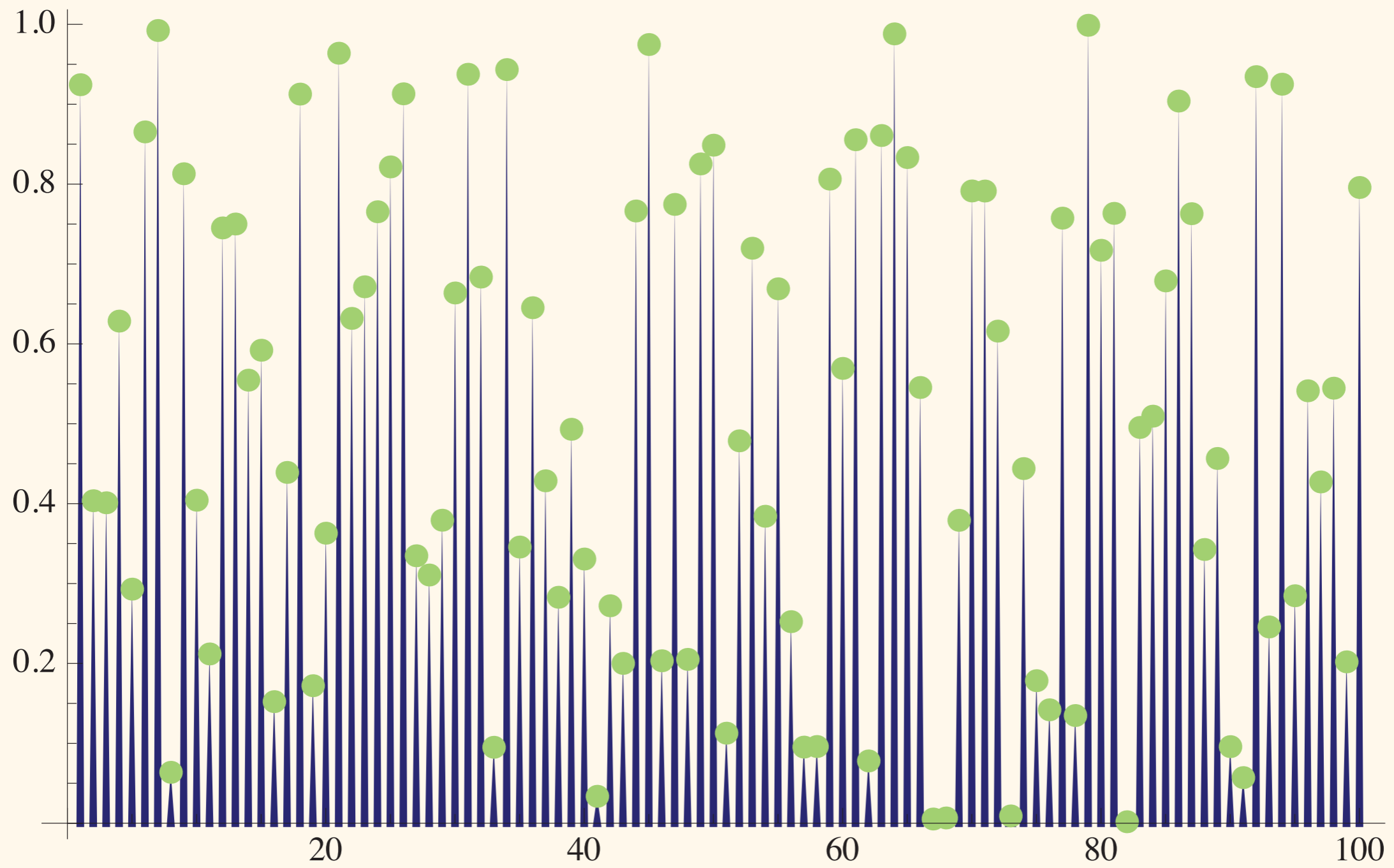
Often, multiple tools can work together.



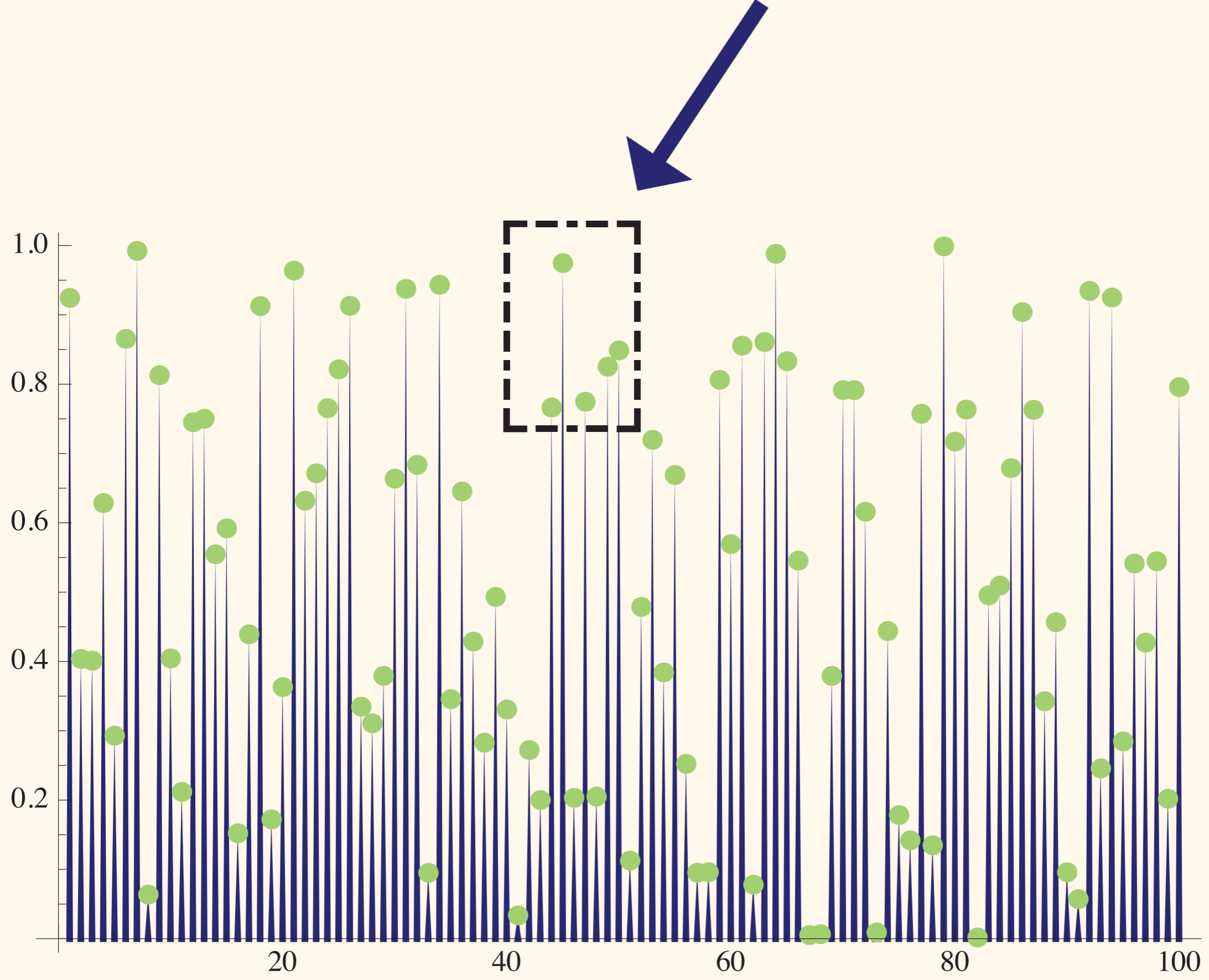












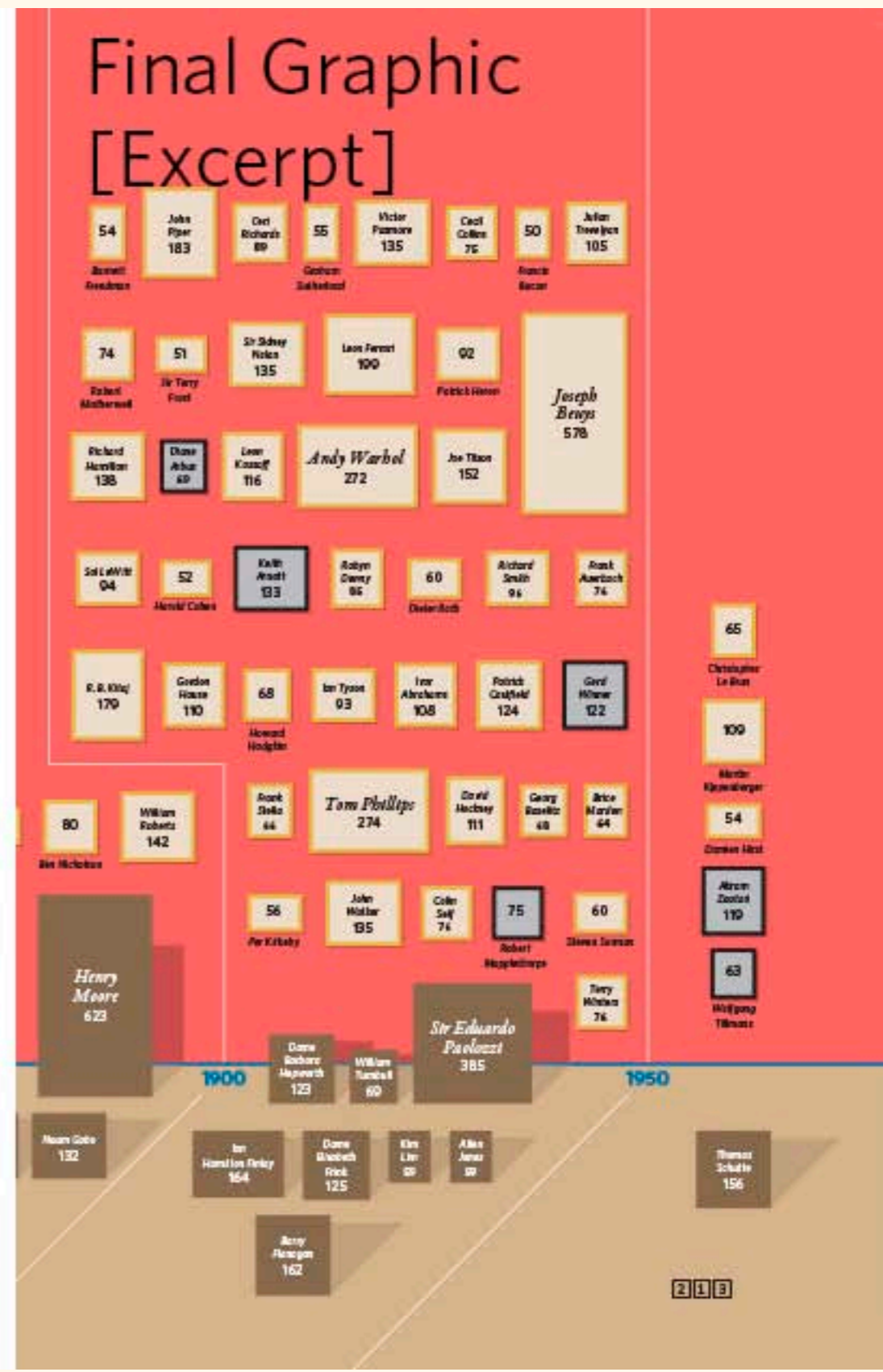


# A less stupid example:

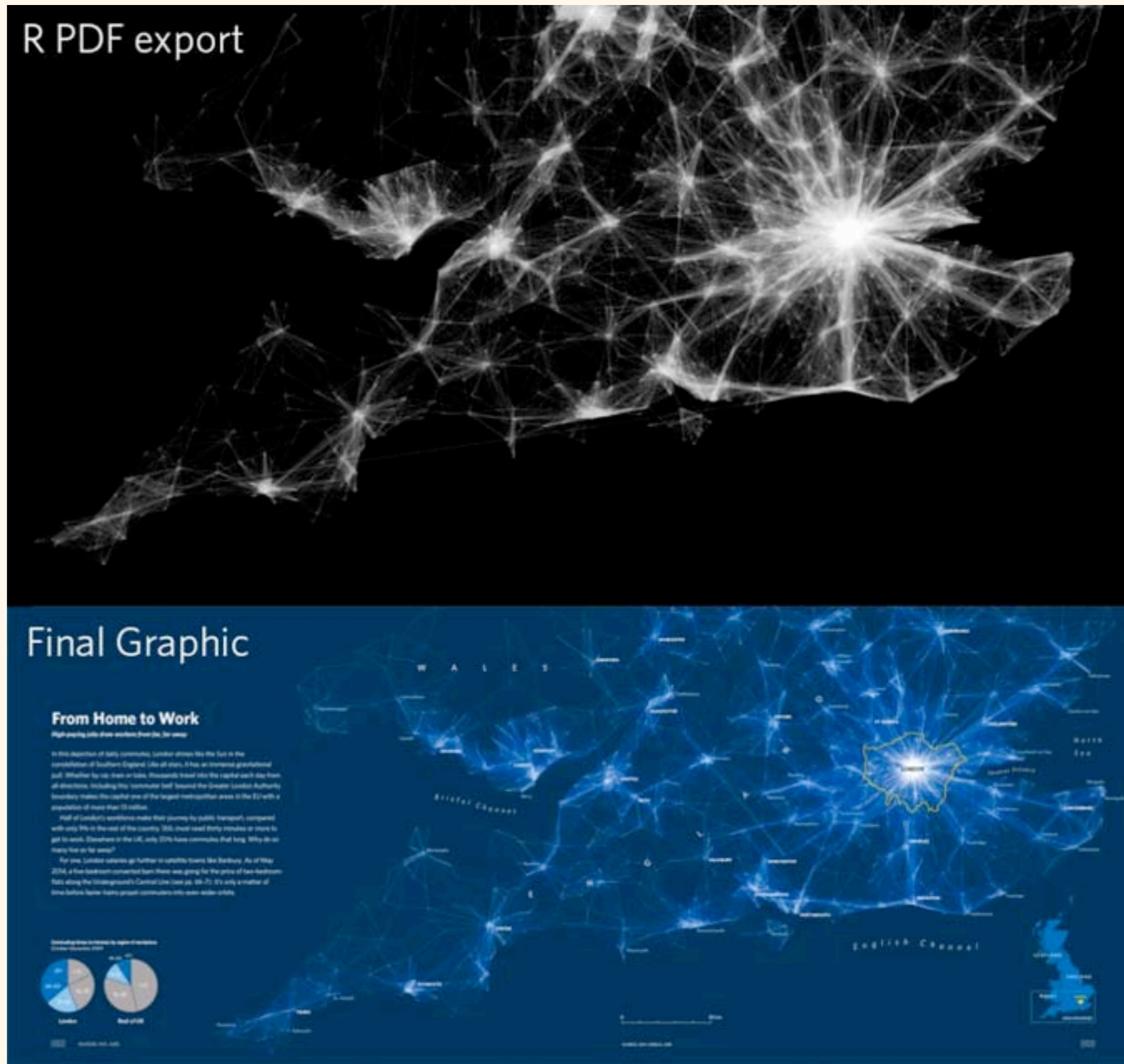
## R PDF export [Excerpt]



## Final Graphic [Excerpt]

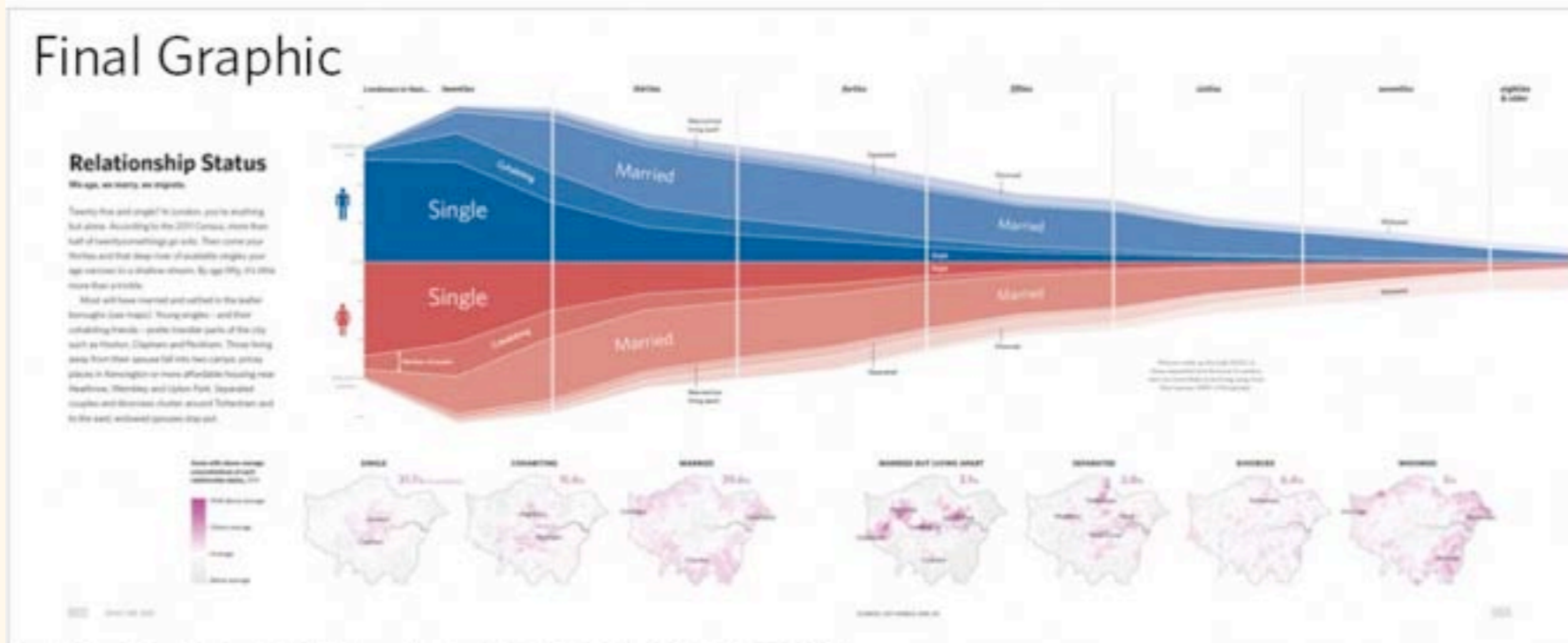
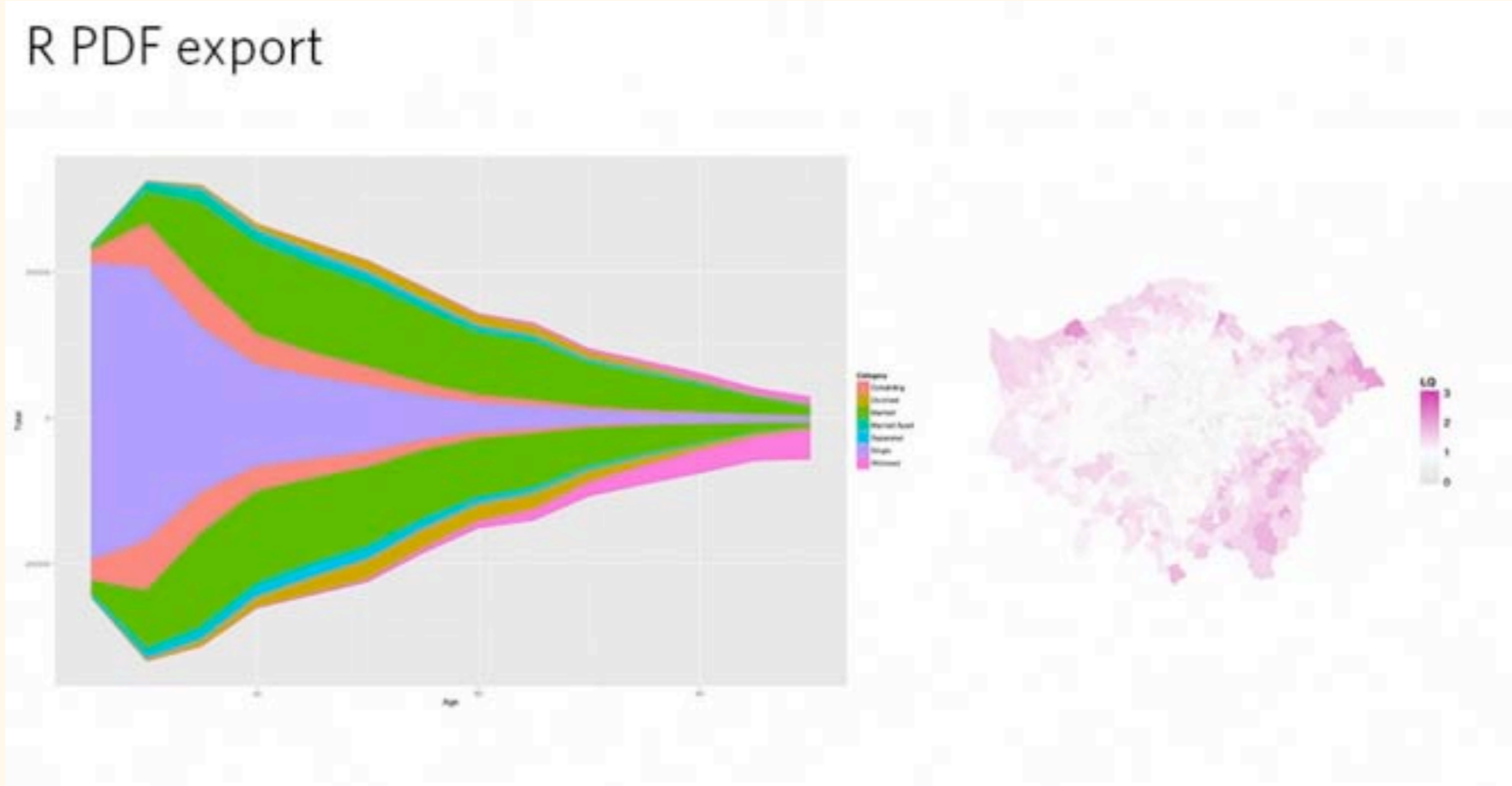


# A less stupid example:



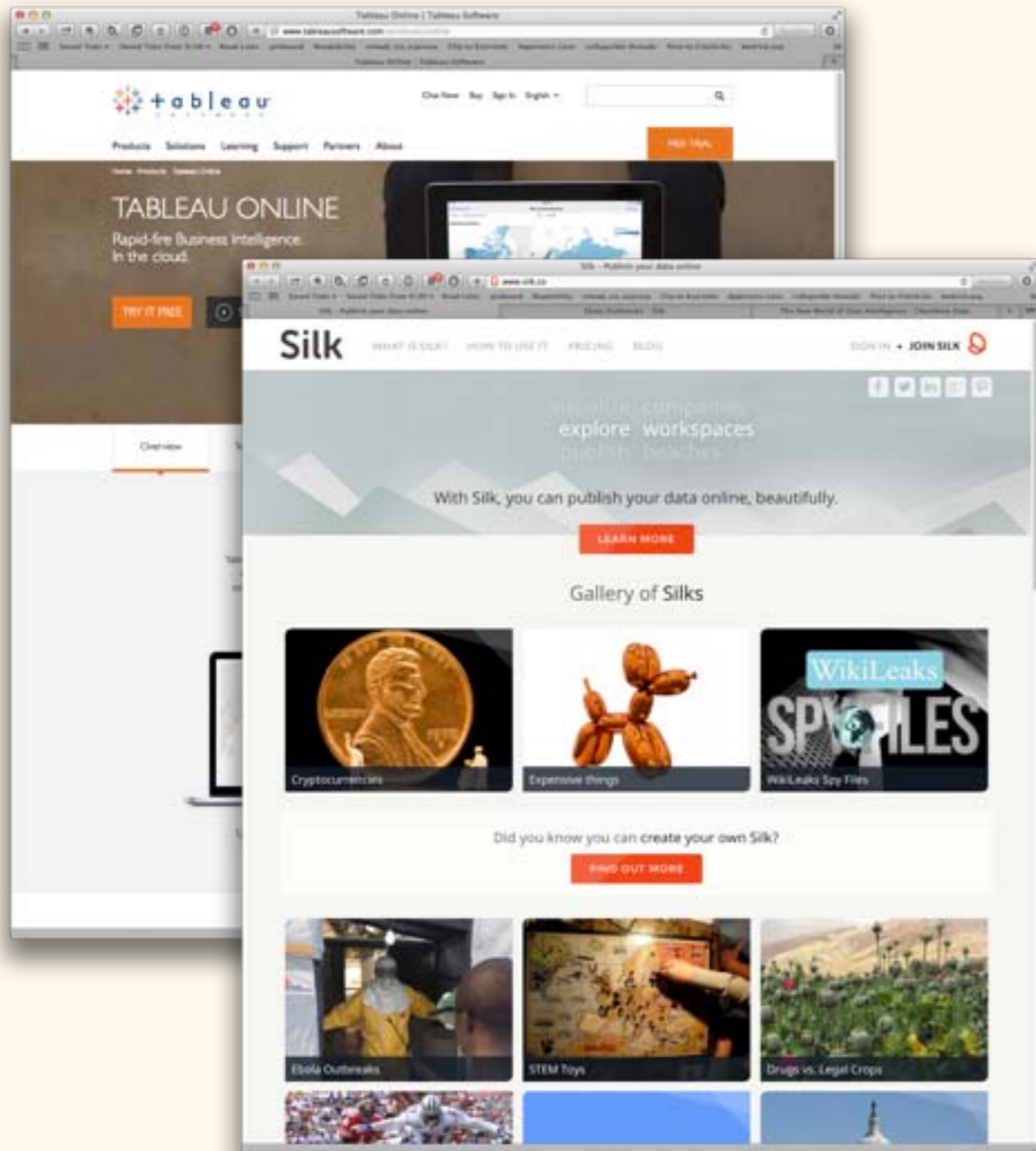
Excerpted from *London: The Information Capital* by James Cheshire and Oliver Uberti (Particular Books, 30 October 2014)

# A less stupid example:



Excerpted from London: The Information Capital by James Cheshire and Oliver Uberti (Particular Books, 30 October 2014)

# IP & Security:



Think before you upload!

Terms of use: "What am I giving away?"

Unpublished data?

IRB & HIPAA?

*When in doubt, ask your neighborhood librarian!*

Change the defaults:

Your tool's default settings are *not* the final word.

Small changes can have a large cumulative effect!

Example: Jackie's competitive eating chart!