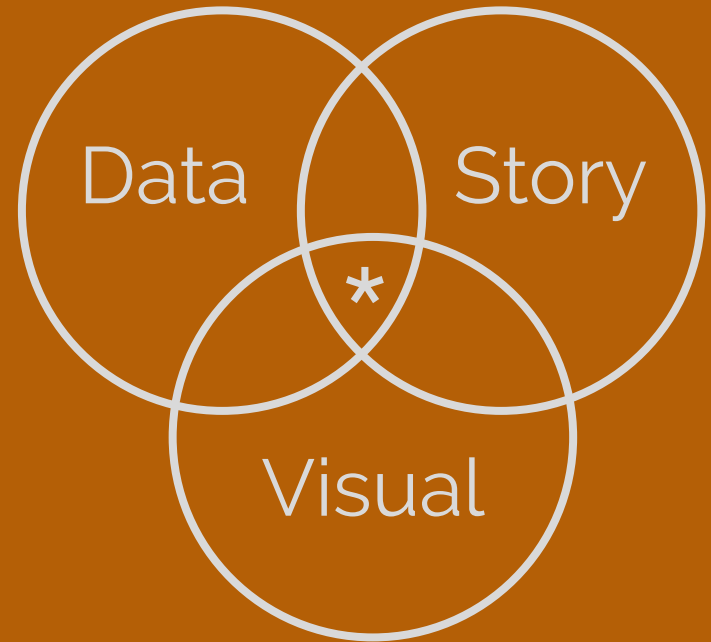


Crafting Visual Stories with Data

Amit Kapoor
@amitkaps

Tools & Resources



Tools Landscape



Abstract

Blackbox

Flexible

Limited

Difficult

Simple

Slow

Quick

Code

GUI

Expressive

Efficient

Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Charting

Collection of fixed charts that require data to be shaped in a particular way

[Excel](#)

[Mondrian](#)

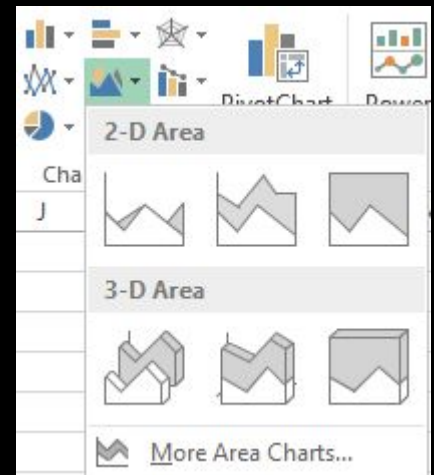
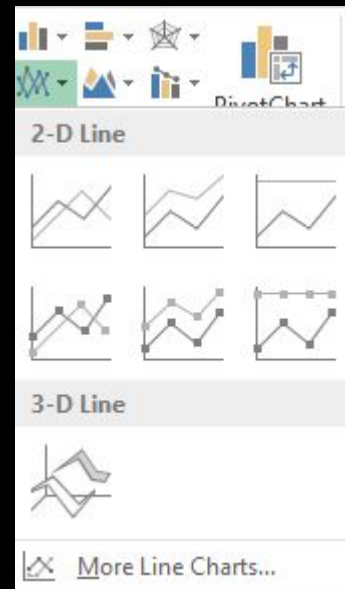
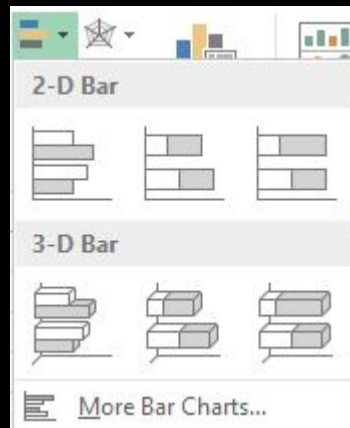
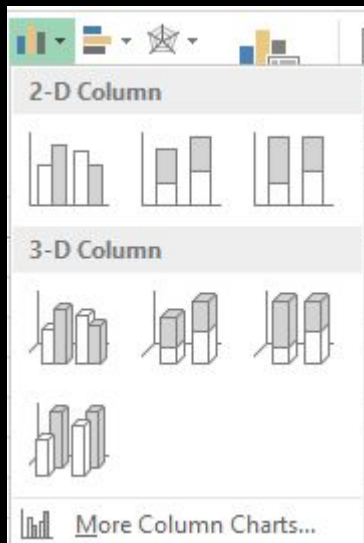
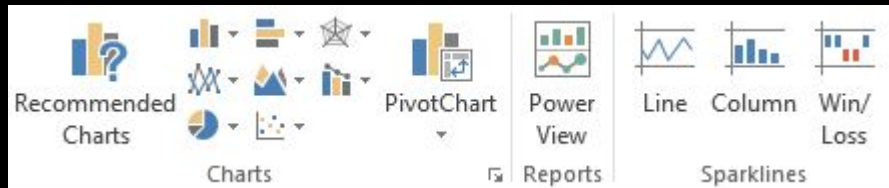
[Many Eyes](#)

[Google Charts](#)

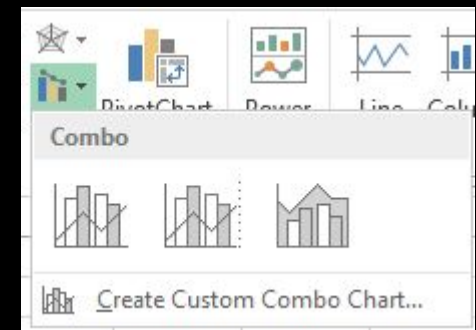
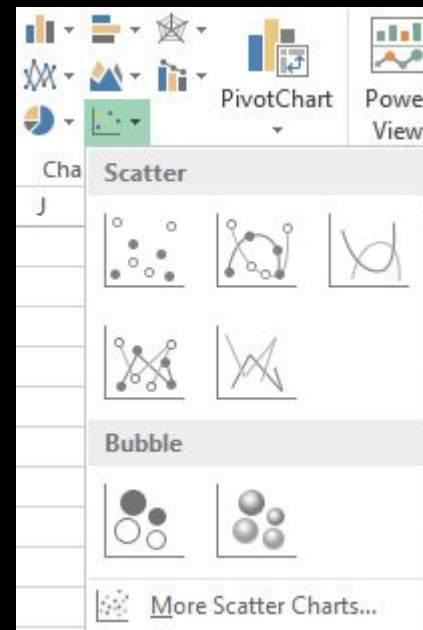
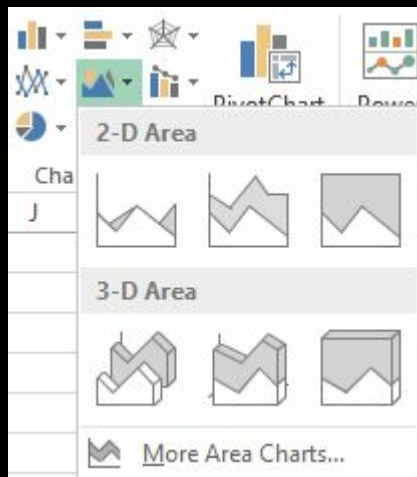
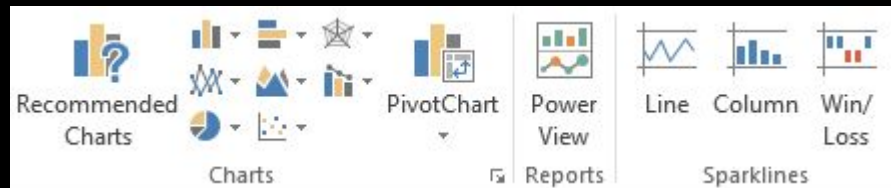
[HighCharts](#)

[Fusion Charts](#)

Excel - Chart Typology



Excel - Chart Typology



Excel - Trees, Streams

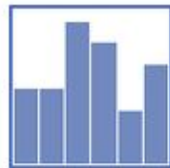
Microsoft Research Data Visualization Apps for Office

At Microsoft Research, we want to let users play with new types of visualizations. We're doing it with **Apps for Office**, a new feature for Office 2013 and Office 365. These let users easily add new functionality to Office.

Apps for Office require Office 2013 or Office 365. You can get a [free trial of Office 365](#). Our apps are in the [Microsoft Research Office Store](#). (Need help getting started? Check out this [handy guide](#).)

We'd Love Your Feedback!

Please ask questions and tell us about what you are doing at our [Online Forum](#).



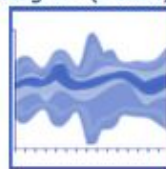
Histogram



Treemap



2D Histogram (Density Plot)



Streamgraph (Stacked Area Chart)

[Overview](#)

[Histogram](#)

[2D Histogram
\(Density Plot\)](#)

[Treemap](#)

[Streamgraph](#)

[Apps for Access](#)

People



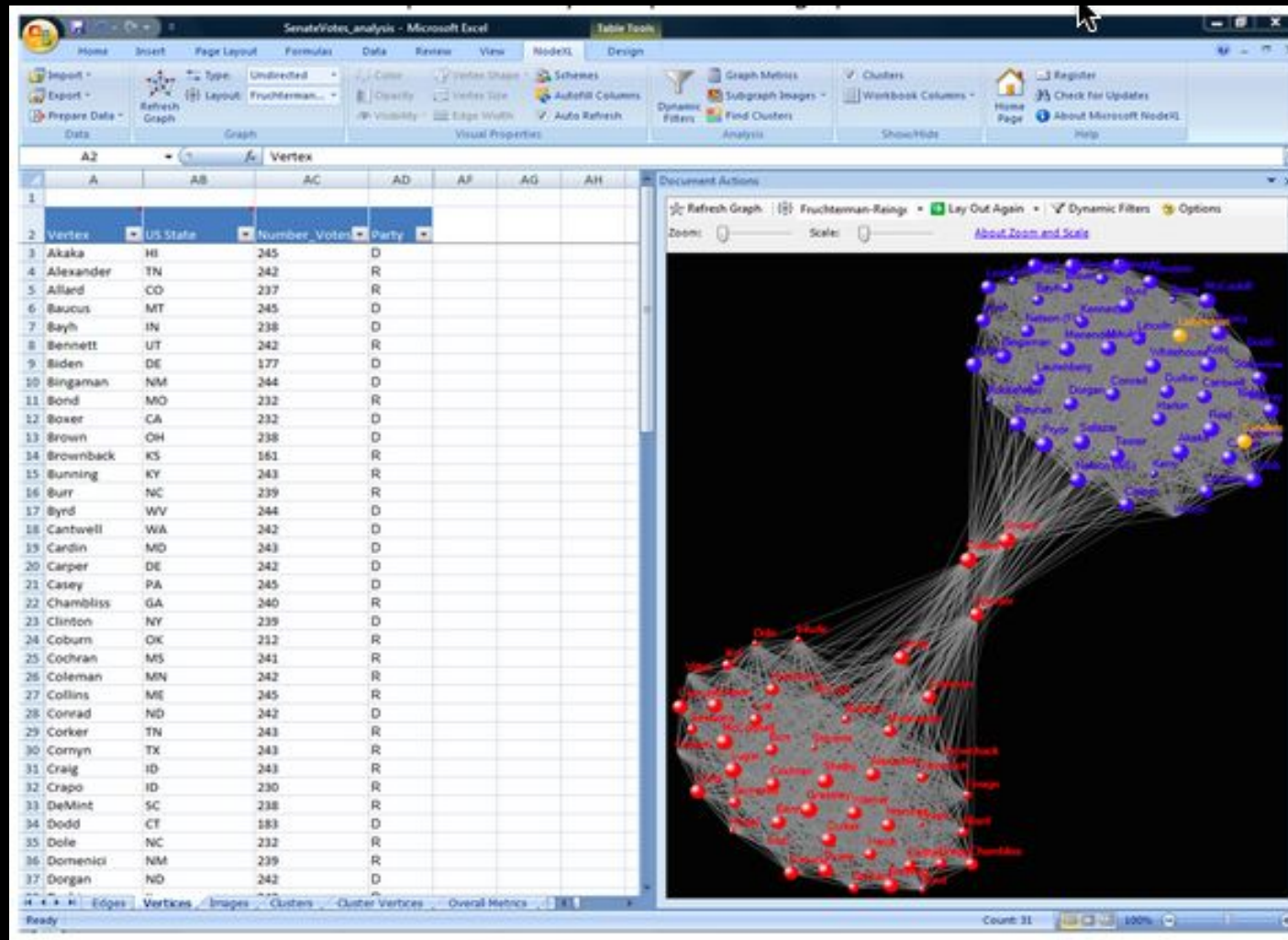
Danyel Fisher



Rick Gutierrez


Source: [Microsoft Research](#)

Excel Add In - NodeXL

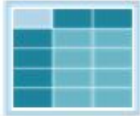


Source: [NodeXL](#)


Many Eyes

 **manyeyes** [Create](#) [Explore](#) [Articles](#) [Feedback](#)

Create a Visualization in Two Easy Steps



1. Upload or Select a Data Set




2. Choose a Visualization and Share

[Upload your own Data](#) [See guidelines \(opens in new tab\) for data format and style](#)

[Use Existing Data](#)


Visualizations You Can Create

Compare a Set of Values




Word Tree

A word tree is a visual search tool for unstructured text, such as a book, article, speech or poem. It lets you pick a word or phrase and shows you all the different contexts in which the word or phrase appears.



Word Cloud

A word cloud generator enables you to see how frequently words appear in a given text, or see the relationship between a column of words and a column of numbers.



Phrase Net

A phrase net diagrams the relationships between different words used in a text. It uses a simple form of pattern matching to provide multiple views of the concepts contained in a book, speech, or poem.

Track Rises and Falls Over Time



Bar Chart

A bar chart is a classic method for numerical comparisons. The Many Eyes bar chart can show one or more sets of variables



Block Histogram

A block histogram lets you see the distribution of numeric values in a data set. The x-axis is divided into "bins" that correspond to value ranges. Each item in the data set is drawn as a rectangular block, and the blocks are piled into the bins to show how many values in each range.



Bubble Chart

A bubble chart displays a set of numeric values as circles. It is especially useful for data sets with dozens to hundreds of values, or with values that differ by several orders of magnitude.



View in Context

The view in context chart is a type of line graph used to show the changes in data over a long set of intervals (often time, but not necessarily), as well as allowing the option of zooming in to see specific short intervals in great detail. View in Context charts are a good choice when there is a long series of data to display, and a combination of overview and detailed views is required.

See Parts of a Whole



Heat Map

Heat maps are a good way to compare three dimensional data. The first two dimensions form a grid, the third value is plotted within that grid as a color intensity. Heat maps don't provide a lot of precision, but are good for seeing broad trends and outliers in data.



Pie Chart

A pie chart is a familiar way of showing proportions. For example, the pie chart below shows the proportions of a budget for advertising: 50% television, 40% newspaper, and 10% yelling on a street corner. The percentages are encoded as "slices" of a pie, with the area corresponding to the percentage.



Tree Map

A tree map is a visualization of hierarchical structures. It is very effective in showing attributes of leaf nodes using size and color coding. Tree maps enable users to compare nodes and sub-trees even at varying depth in the tree, and help them spot patterns and exceptions.



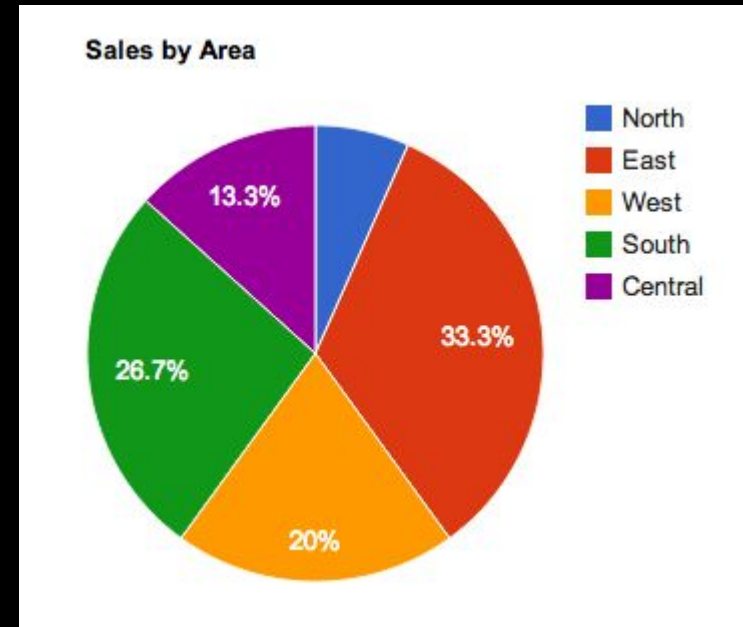
Scatter plot

A scatter plot is a classic statistical diagram that lets you visualize the relationship between numeric variables. For instance, if you have a table of data on cities, you could use a scatter plot to see if there is a relationship between population and crime levels.

Source: [IBM- Many Eyes](#)

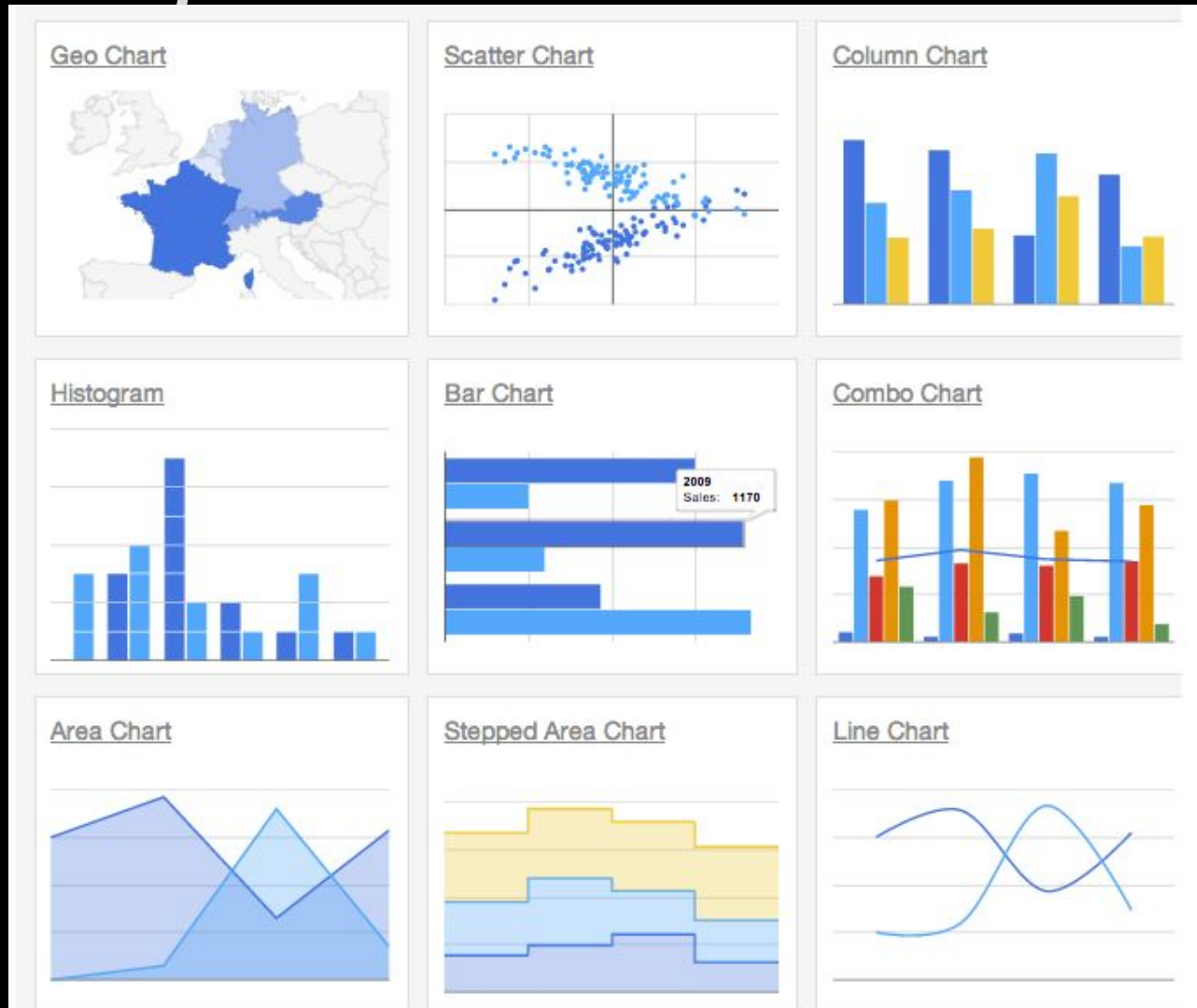
Google Charts - Pie

```
function drawVisualization() {  
    // Create and populate the data table.  
    var data =  
    google.visualization.arrayToDataTable([  
        ['Areas', 'Sales'],  
        ['North', 5],  
        ['East', 25],  
        ['West', 15],  
        ['South', 20],  
        ['Central', 10]  
    ]);  
  
    // Create and draw the visualization.  
    new  
    google.visualization.PieChart(document.ge  
tElementById('visualization')).  
        draw(data, {title:"Sales by  
Area"});  
}
```



Source: [Google Charts Playground](#)

Google Charts



Source: [Google Charts](https://www.google.com/charts/)

Google Charts

Pie Chart



Bubble Chart



Donut Chart



Org Chart



Treemap



Table

	Name	Salary	Full Time
1	Marie	\$24,700	✓
2	Albert	\$25,200	x
3	Enrico	\$25,700	✓
4	Lise	\$26,600	✓

Timeline



Gauge

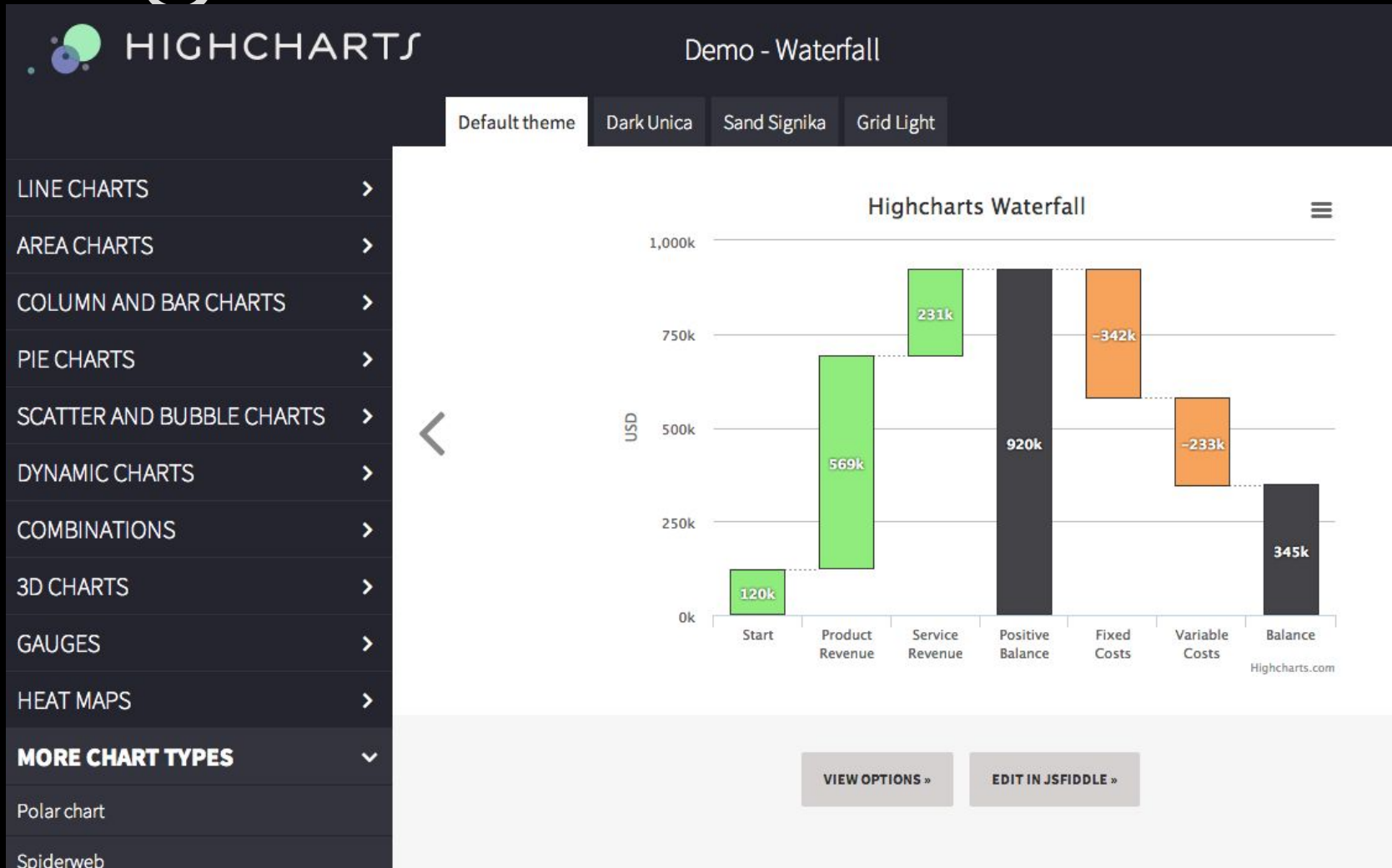


Candlestick Chart



Source: [Google Charts](https://www.google.com/charts/)

HighCharts



Source: [HighCharts](https://www.highcharts.com/)

Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Visual analysis
languages allowing
flexibility to design
many variants

[Tableau](#)
[PowerBI](#)

[raw](#)

[Polestar](#)
[Gephi](#)

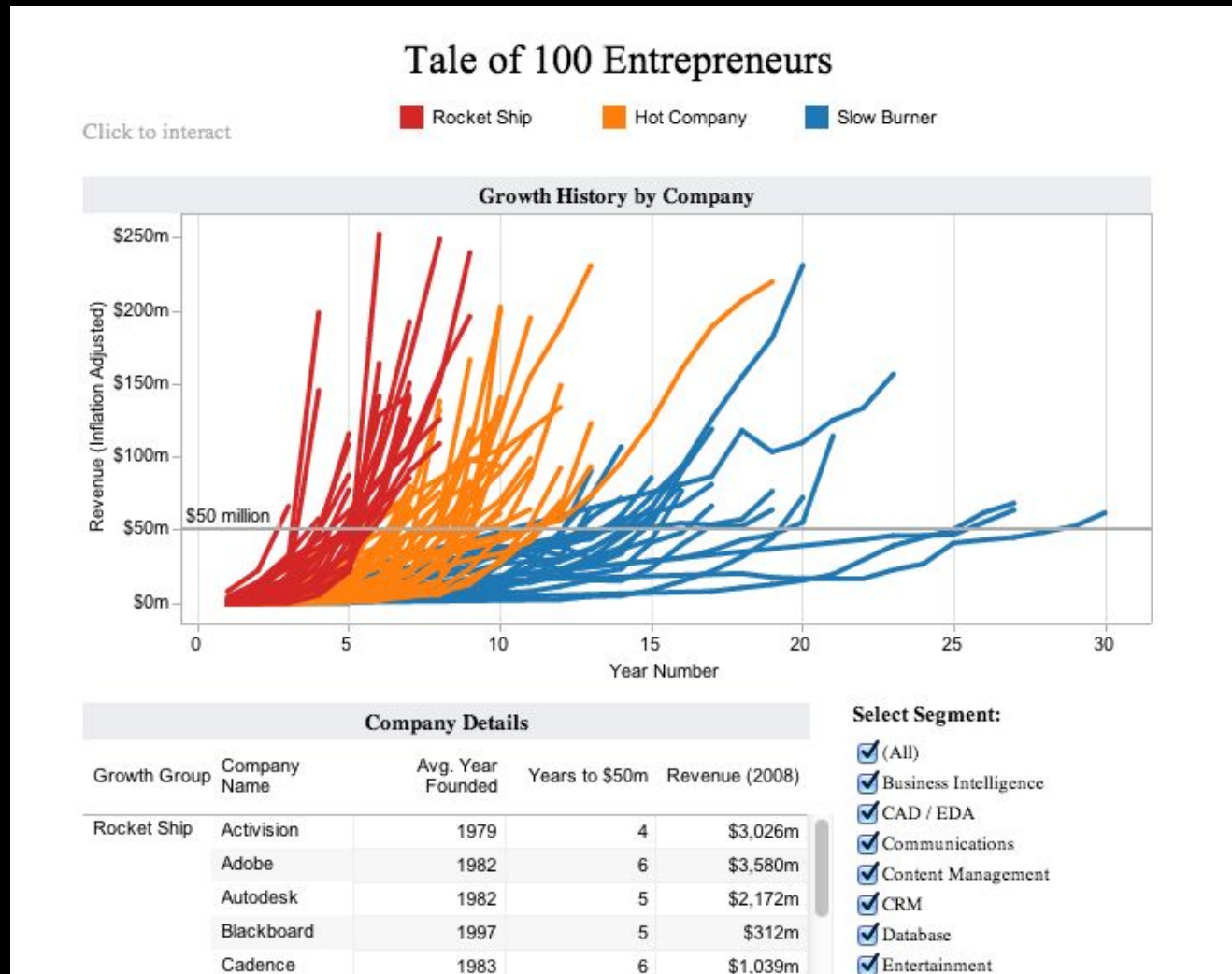
Collection of fixed
charts that require
data to be shaped
in a particular way

[Excel](#)
[Mondrian](#)

[Many Eyes](#)

[Google Charts](#)
[HighCharts](#)
[Fusion Charts](#)

Tableau - VizQL



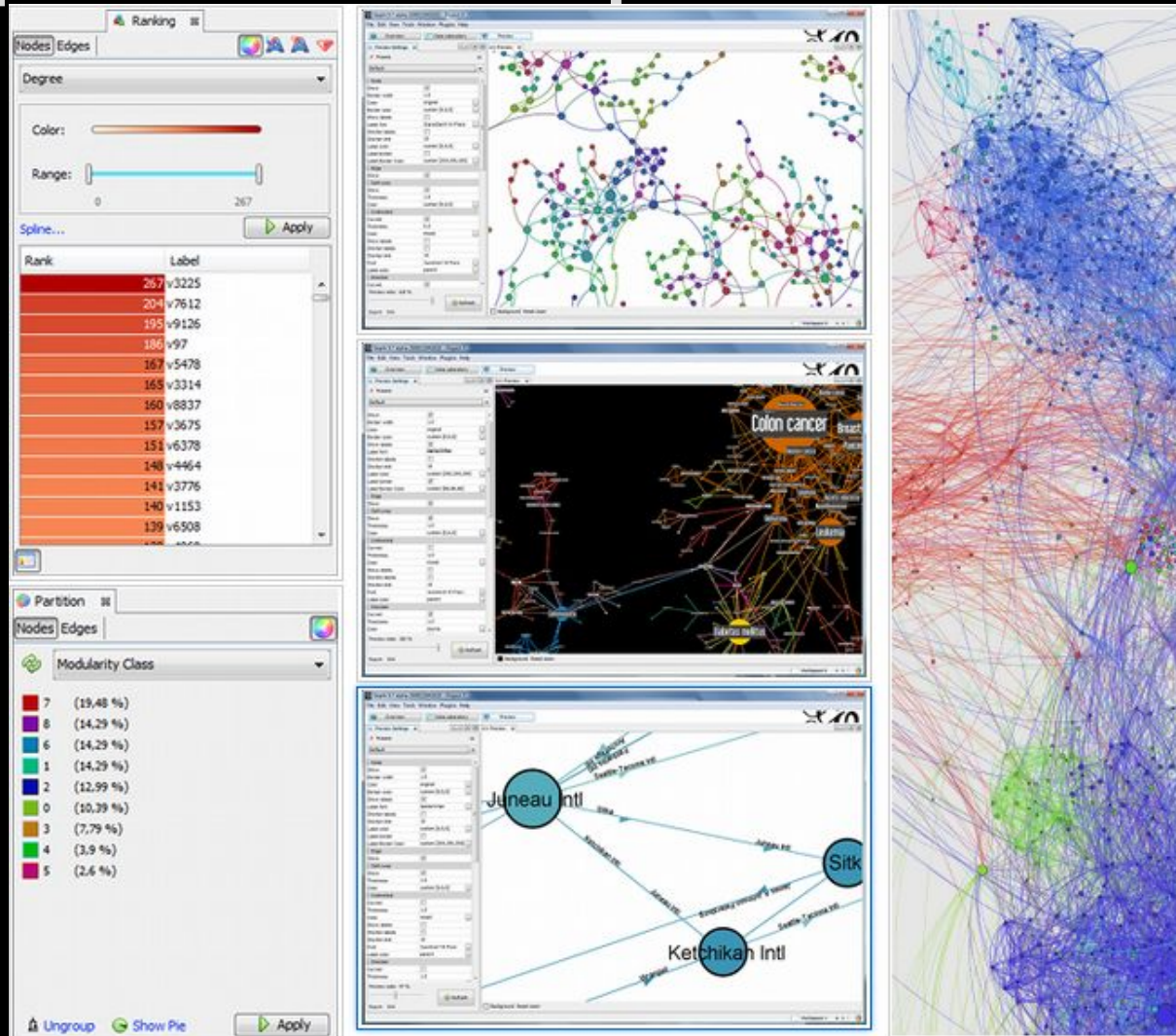
Source: [Tableau Software Public](#)

Power BI



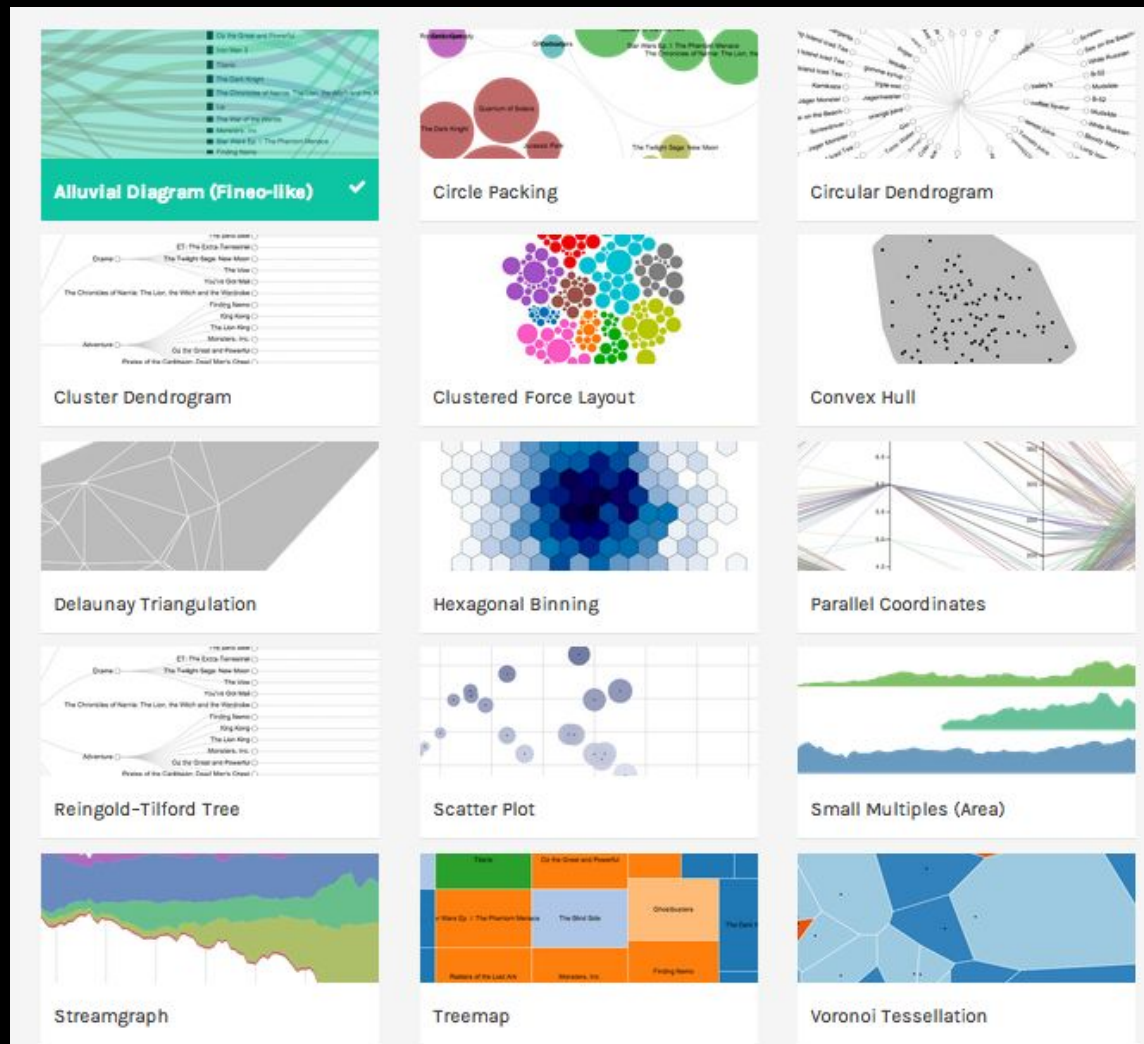
Source: [PowerBI](#)

Gephi - Graph Viz



Source: [Gephi](https://gephi.org/)

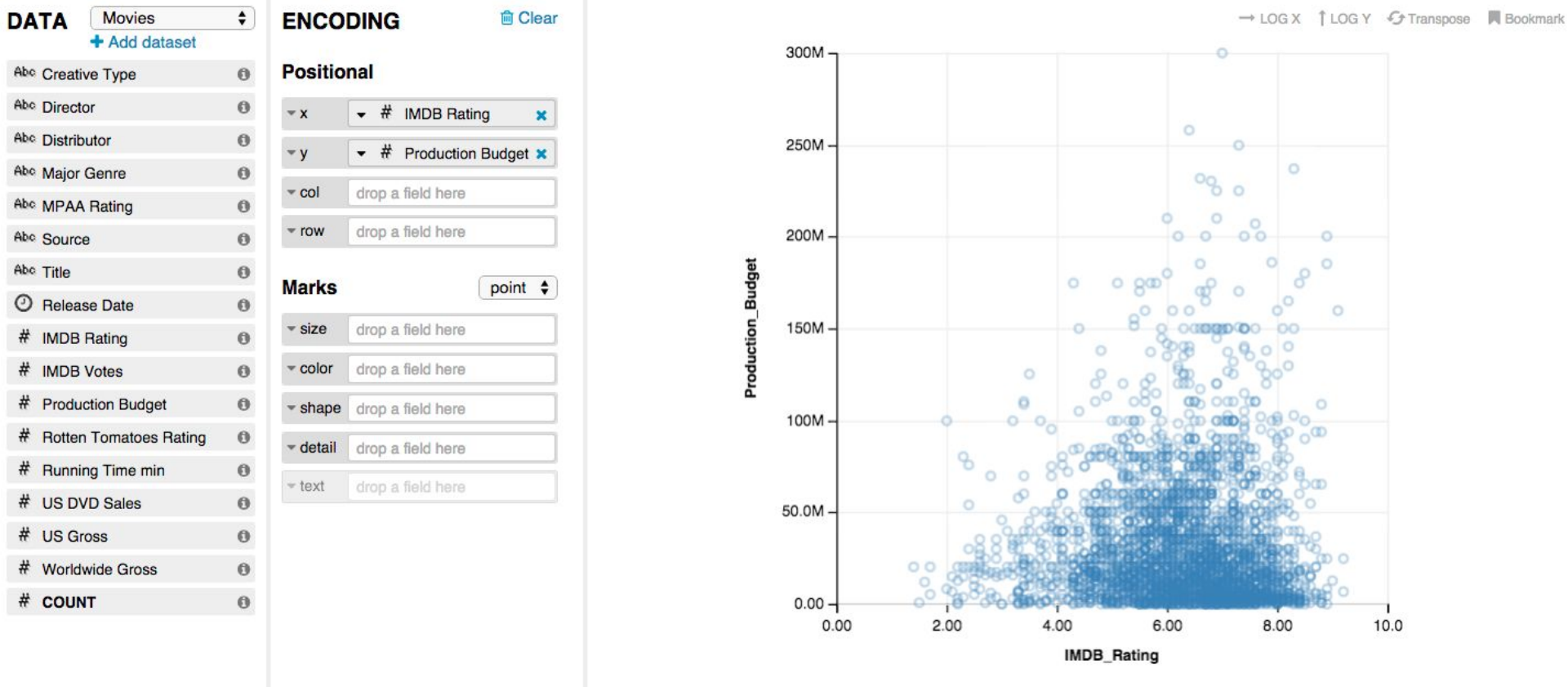
Raw



Source: [Raw](#)

Polestar

Polestar



Source: [Polestar](#)

Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Grammar

Visual

Charting

Collection of graphical primitives for composing data driven graphics

Visual analysis languages allowing flexibility to design many variants

Collection of fixed charts that require data to be shaped in a particular way

[R-ggplot2](#)
[SPSS](#)

[Tableau](#)
[PowerBI](#)

[Excel](#)
[Mondrian](#)

[plot.ly](#)

[raw](#)

[Many Eyes](#)

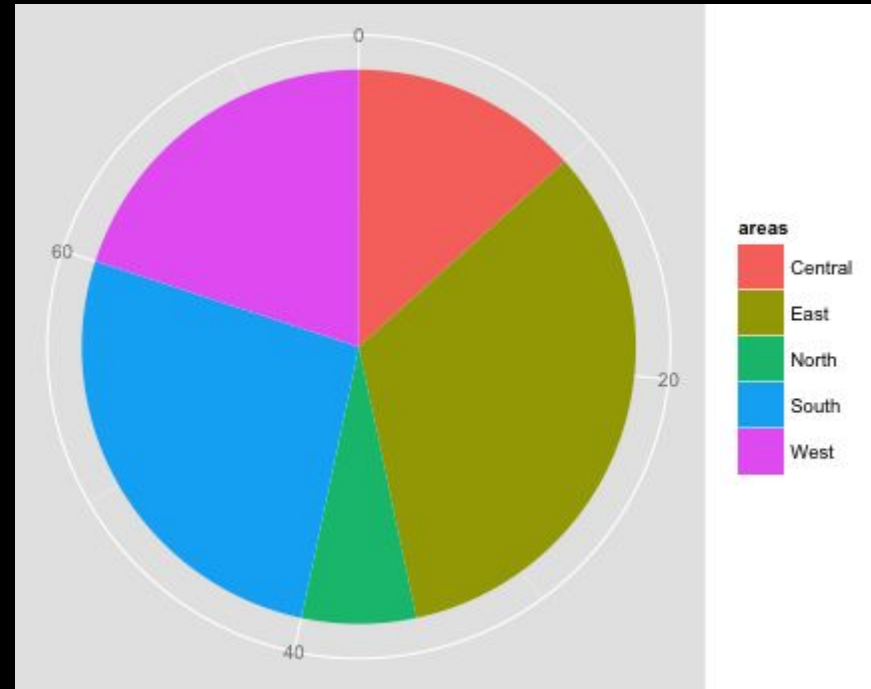
[d3.js](#)
[Bokeh](#)
[vega](#)

[Polestar](#)
[Gephi](#)

[Google Charts](#)
[HighCharts](#)
[Fusion Charts](#)

R - ggplot2

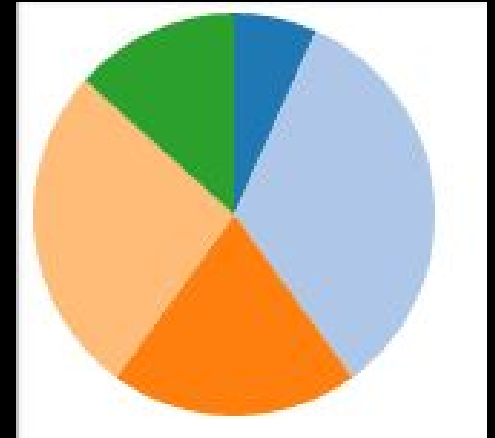
```
areas <- c("North", "East", "West",  
           "South", "Central")  
sales <- c(5, 25, 15, 20, 10)  
humble <- data.frame(areas, sales)  
  
ggplot(humble, aes(x = "",  
                   fill = areas,  
                   weight = sales)) +  
  geom_bar(width = 1) +  
  coord_polar("y")
```



Source: [ggplot2](#)

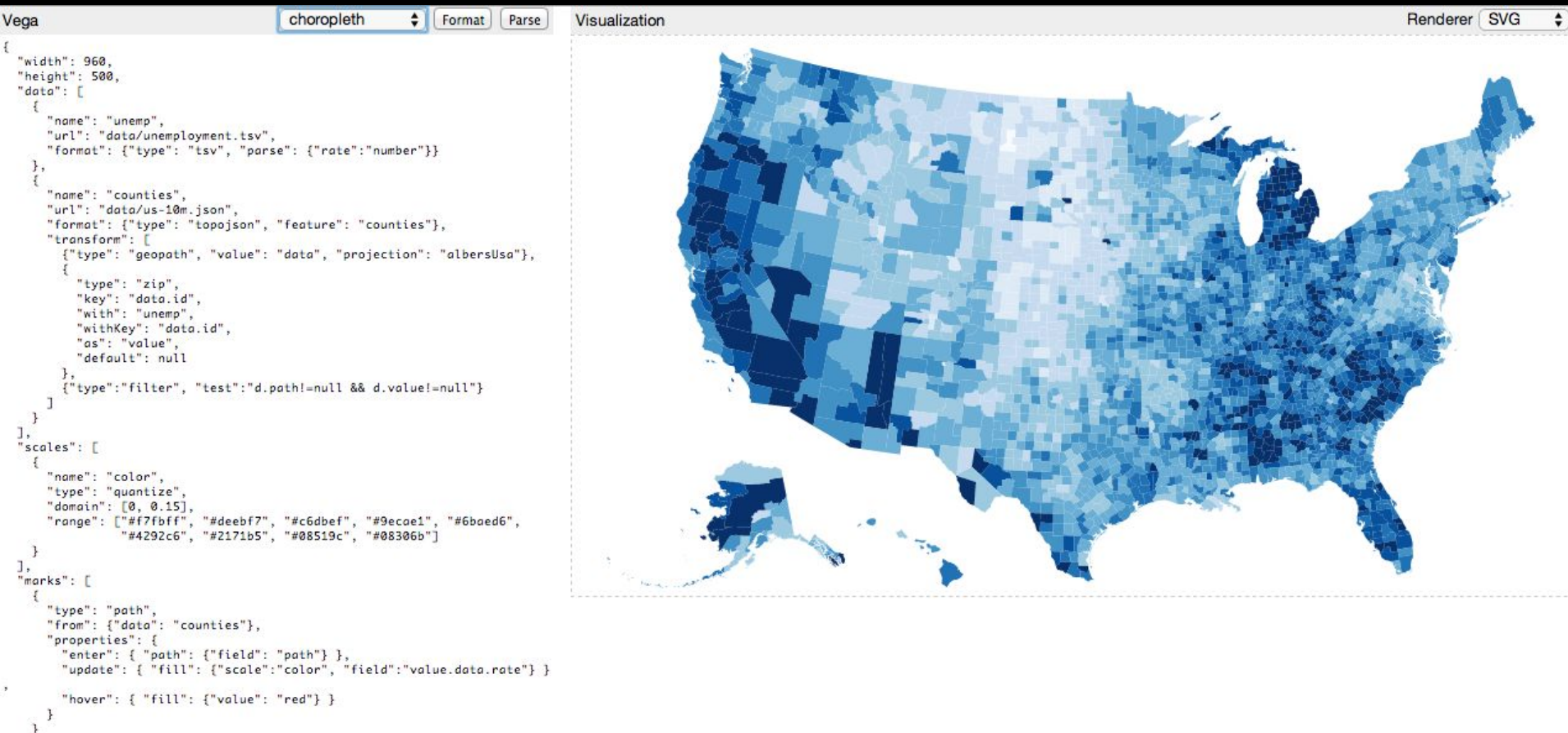
d3.js

```
var data = [5, 25, 15, 20, 10];  
var color = d3.scale.category20();  
var svg = d3.select("body").append("svg")  
    .attr("width", 150)  
    .attr("height", 150);  
  
var g = svg.append("g")  
    .attr("transform", "translate(75,  
75)");  
  
var arcs = g.selectAll("path")  
    .data(d3.layout.pie().sort(null)(data))  
    .enter().append("path")  
    .style("fill", function(d,i) {  
        return color(i); })  
    .attr("d",d3.svg.arc()  
        .innerRadius(0)  
        .outerRadius(70));
```



Source: [Data-Driven Documents](#)

Vega



Source: [Vega Visualization](#)

Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Canvas

Paint directly on a pixel grid. Design & manage every element of chart

[Processing](#)
[Nodebox](#)

[sketchpad](#)

[Raphael.js](#)
[Paper.js](#)
[Processing.js](#)

Grammar

Collection of graphical primitives for composing data driven graphics

[R-ggplot2](#)
[SPSS](#)

[plot.ly](#)

[d3.js](#)
[Bokeh](#)
[vega](#)

Visual

Visual analysis languages allowing flexibility to design many variants

[Tableau](#)
[PowerBI](#)

[raw](#)

[Polestar](#)
[Gephi](#)

Charting

Collection of fixed charts that require data to be shaped in a particular way

[Excel](#)
[Mondrian](#)

[Many Eyes](#)

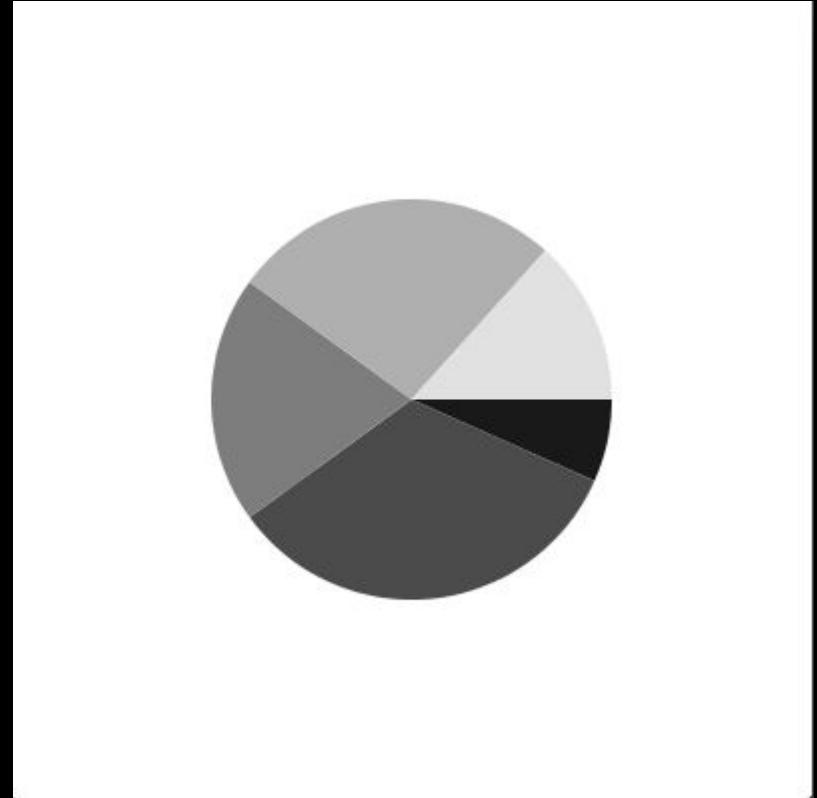
[Google Charts](#)
[HighCharts](#)
[Fusion Charts](#)

Processing - Pie Chart

```
size(400, 400);
background(255);
smooth();
noStroke();

int diameter = 200;
float[] data = {5, 25, 15, 20, 10};
int[] sat = {25, 75, 125, 175, 225};
float lastAngle = 0, sum = 0;
for (int i = 0; i < data.length; i++)
    sum += data[i];

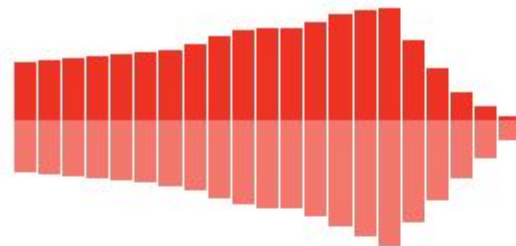
for (int i = 0; i < data.length; i++) {
    fill(sat[i]);
    float angle = data[i] / sum * 2 * PI;
    arc(width / 2, height / 2, diameter,
    diameter,
    lastAngle, lastAngle + angle);
    lastAngle += angle;
}
```



Source: [Processing](#)

Processing - Example

France
Germany
Italy
 **Japan**
United Kingdom
USA
South Korea
China

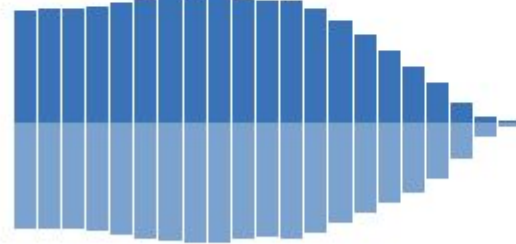


2050 Population: 102 million

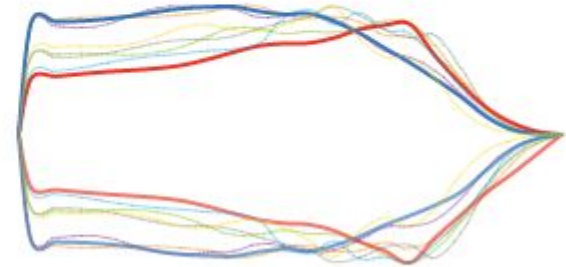
An Aging Population

This visualization depicts shifts in population by age group from eight industrialized nations. On the left, select two countries for comparison, or see all eight as a composite below. Use the slider at the bottom of the screen to cycle through population growth from 1950 to 2050.

France
Germany
Italy
Japan
United Kingdom
 **USA**
South Korea
China



2050 Population: 404 million



▶  2050


Source: [Fathom](#)

Nodebox - Humble Pie

humble_pie.ndbx

root

Viewer Data ☒ Handles ☐ Points ☐ Point Numbers ☐ Origin ☐ Bounds

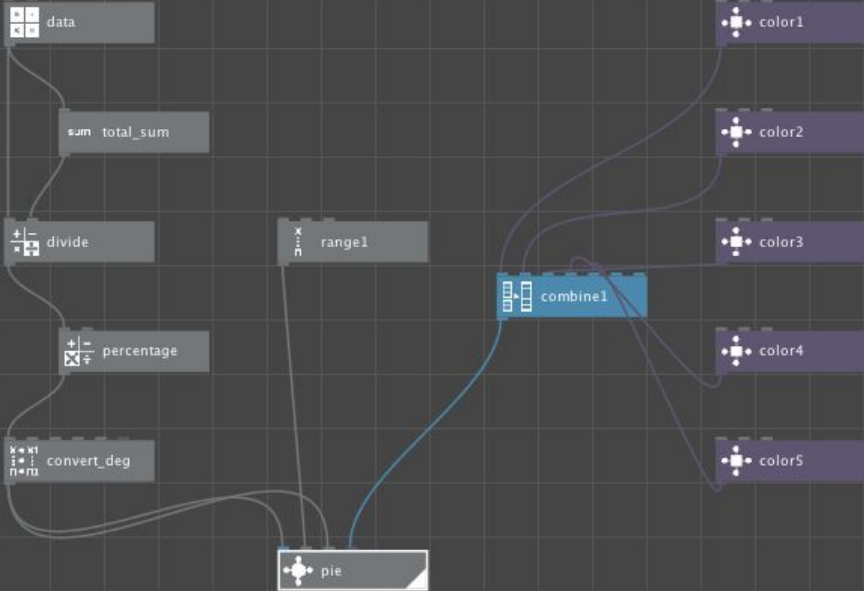


pie

	Metadata
Size	<connected>
Degrees	<connected>
Fill	<connected>

Network

+ New Node



data

sum total_sum

divide

percentage

convert_deg

range1

combine1

color1

color2

color3

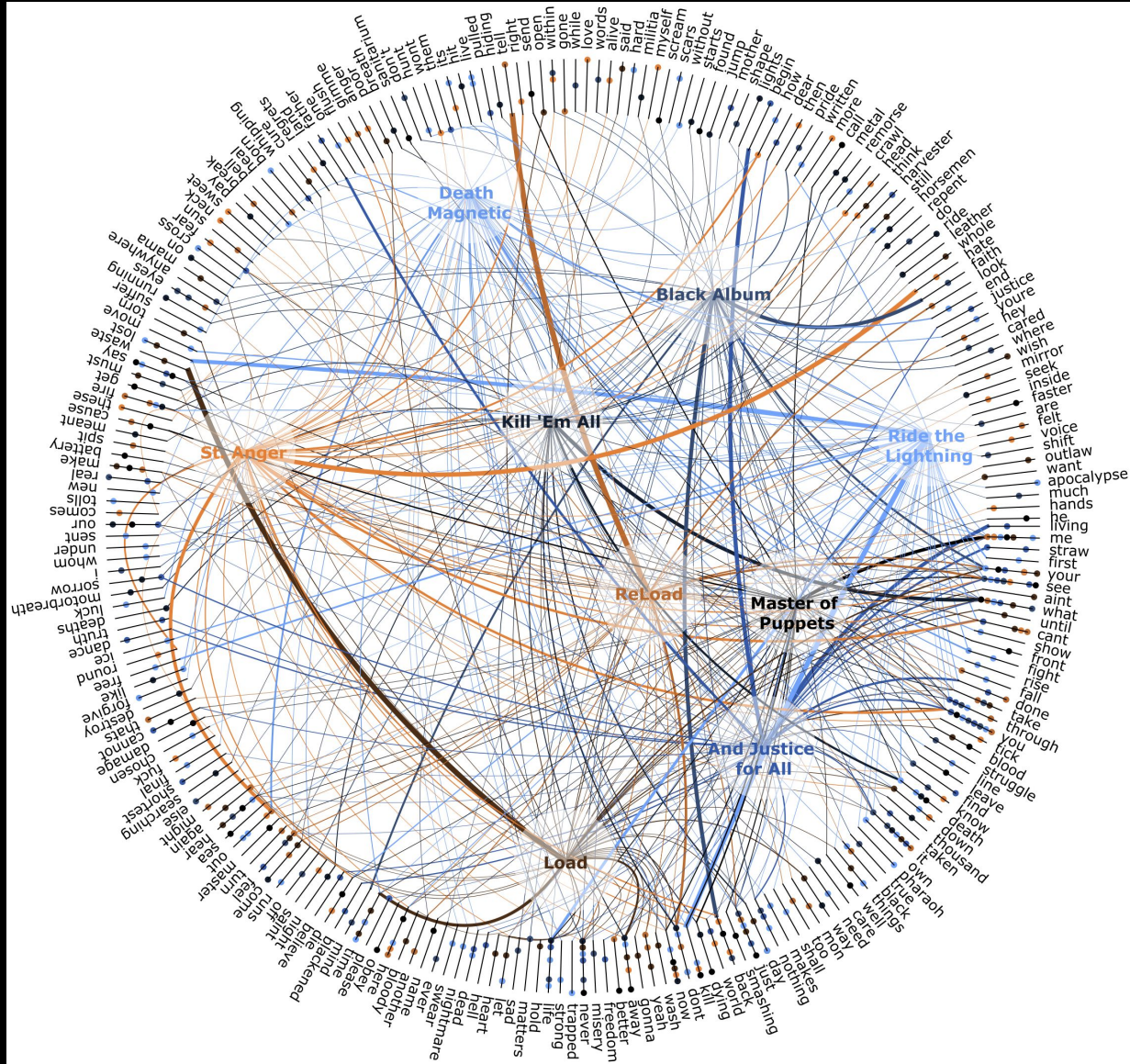
color4

color5

pie

1.00 Play Rewind

Nodebox - Example



Tools Landscape

Abstract, Flexible, Difficult
Slow, Code, Expressive

Blackbox, Limited, Simple
Quick, GUI, Efficient



Canvas

Paint directly on a pixel grid. Design & manage every element of chart

[Processing](#)
[Nodebox](#)

[sketchpad](#)

[Raphael.js](#)
[Paper.js](#)
[Processing.js](#)

Grammar

Collection of graphical primitives for composing data driven graphics

[R-ggplot2](#)
[SPSS](#)

[plot.ly](#)

[d3.js](#)
[Bokeh](#)
[vega](#)

Visual

Visual analysis languages allowing flexibility to design many variants

[Tableau](#)
[PowerBI](#)

[raw](#)

[Polestar](#)
[Gephi](#)

Charting

Collection of fixed charts that require data to be shaped in a particular way

[Excel](#)
[Mondrian](#)

[Many Eyes](#)

[Google Charts](#)
[HighCharts](#)
[Fusion Charts](#)

Tools - Chart (1/2)



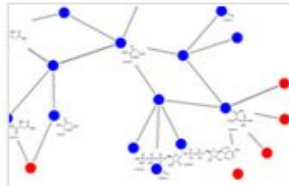
Arbor.js

A library of force-directed layout algorithms plus abstractions for graph organization and refresh handling.



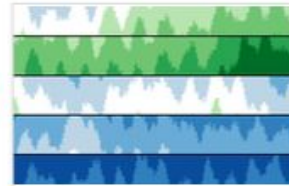
Circos

A software package for visualizing data in a circular layout.



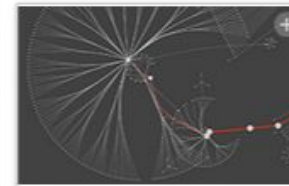
Cola.js

A library for arranging networks using constraint-based optimization techniques.



Cubism.js

A library for creating interactive time series and horizon graphs based on D3.js



Cytoscape

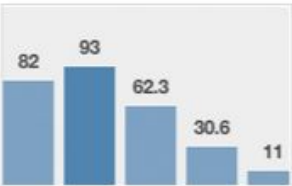
An application for visualizing complex networks and integrating these with any type of attribute data.

<http://www.cytoscape.org>



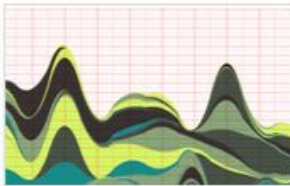
D3.js

An small, flexible and efficient library to create and manipulate interactive documents based on data.



Dance.js

A simple data-driven visualization framework based on Data.js and Underscore.js



Degrafa

A powerful declarative graphics framework for rich user interfaces, data visualizations and mapping.



Envision.js

A library for creating fast, dynamic and interactive time series visualizations.



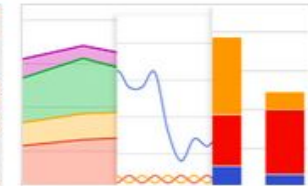
Flare

A set of software tools for creating rich interactive data visualizations in ActionScript.



Gephi

A visualization and exploration platform for networks with dynamic and hierarchical graphs.



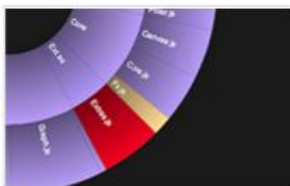
Google Chart Tools

A collection of simple to use, customizable and free to use interactive charts and data tools.



Google Fusion Tables

A web application that makes it easy to host, manage, collaborate on, visualize, and publish data tables.



JavaScript InfoVis Toolkit

A JavaScript library that provides tools for creating interactive data visualizations for the web.



Many Eyes

A web application to build, share and discuss graphic representation of user uploaded data.



NVD3.js

A collection of re-usable charts and chart components for d3.js.



NodeBox

A desktop application that lets you create generative, static, animated or interactive visuals.

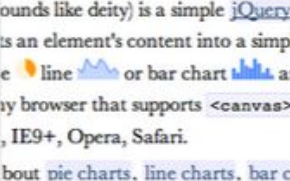

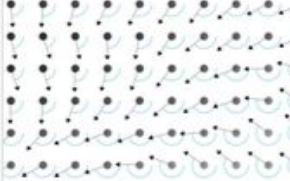





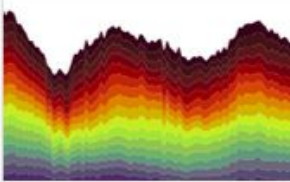

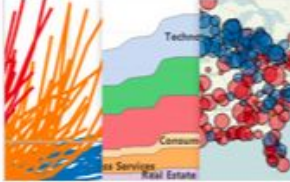
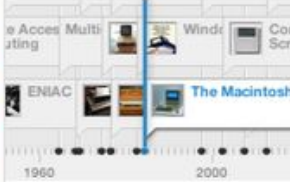
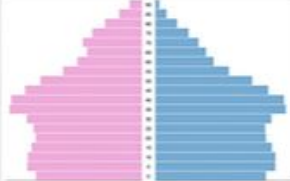




Paper.js

A vector graphics scripting framework in a well designed, consistent and clean programming interface.

Source: datavisulization.ch

Tools - Chart (2/2)

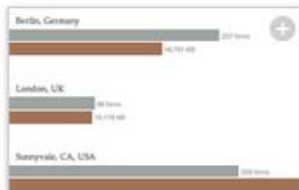
 <p>ounds like deity) is a simple jQuery plugin that converts an element's content into a simple line or bar chart. It works in any browser that supports <code><canvas></code>, IE9+, Opera, Safari.</p> <p>about pie charts, line charts, bar charts.</p> <p>Peity</p> <p>A simple jQuery plugin that converts an element's content into a simple mini pie, line or bar chart.</p>	 <p>Prefuse</p> <p>A set of software tools for creating rich interactive data visualizations in Java.</p>	 <p>Processing</p> <p>An open source programming language and environment to create images, animations, and interactions.</p>	 <p>Processing.js</p> <p>The sister project of Processing that makes projects work using web standards and without any plug-ins.</p>	 <p>Protovis</p> <p>A library that composes custom views of data with simple marks such as bars and dots.</p> <p>http://protovis.org</p>	 <p>R</p> <p>A software environment for statistical computing and graphical techniques.</p>
 <p>Raphaël</p> <p>A small library that simplifies working with vector graphics on the web.</p>	 <p>Raw</p> <p>An application to create custom vector-based visualizations on top of D3.js.</p>	 <p>Rickshaw</p> <p>A library for creating interactive time series graphs based on D3.js.</p>	 <p>Sigma.js</p> <p>An open-source lightweight library to display interactively static and dynamic graphs.</p>	 <p>Tableau Public</p> <p>A desktop application to build and post interactive graphs, dashboards, maps and tables to the web.</p>	 <p>Timeline.js</p> <p>A tool to create timelines with data and media from different sources like Google Docs, Twitter, Flickr or Vimeo.</p>
 <p>Vega</p> <p>A visualization grammar, a declarative format for creating, saving and sharing visualization designs.</p>	 <p>Visage</p> <p>An application to create beautiful, branded reports.</p>	 <p>ZingCharts</p> <p>A library for building HTML5 charts with an API set designed to render charts across browsers and devices.</p>			

Tools - Data



D3.js

A small, flexible and efficient library to create and manipulate interactive documents based on data.



Data.js

A data representation framework providing a uniform interface to domain data.

<http://substance.io/michael/data.js>



DataWrangler

An interactive web application for data cleaning and transformation.



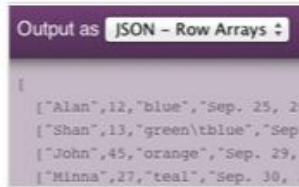
Google Fusion Tables

A web application that makes it easy to host, manage, collaborate on, visualize, and publish data tables.



Miso

A toolkit to expedite the creation of interactive storytelling and data visualisation content.



Mr. Data Converter

A simple console that converts Excel data into web-friendly formats, including HTML, JSON and XML.



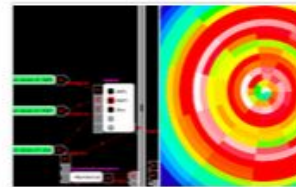
Mr. Nester

A simple console for learning and experimenting with d3.js data nesting.



OpenRefine

A tool for working with data, cleaning it up, reformatting it or extending it with web services.



Quadrigram

A visual programming language aimed to gather, process and visualize information.



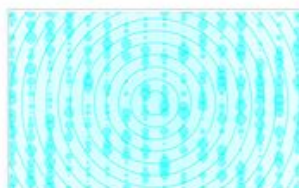
R

A software environment for statistical computing and graphical techniques.



Recline.js

A simple but powerful library for building data applications in pure JavaScript and HTML.



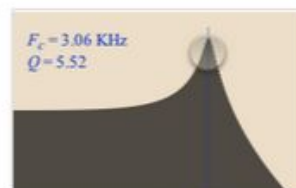
SVG Crowbar

A bookmarklet that extracts SVG nodes from an HTML document into a SVG file.

Apellido y Nombre	
ABDALA de MATARAZZO, Norma Amanda	Fronte Civic
ALBRIEU, Oscar Edmundo Nicolas	Fronte para
ALONSO, Maria Luz	Fronte para
ARENA, Celia Isabel	Fronte para
ARREGUI, Andres Roberto	Fronte para
AVOSCAN, Herman Horacio	Fronte para
BALCEDO, Maria Ester	Fronte para
BARRANCO, David Federico	Fronte para

Tabula

A tool to extract CSV formatted data from text tables in PDF documents.



Tangle

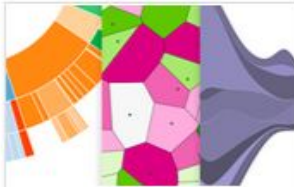
A library that allow to interactively explore, play, and see the document update immediately.

Tools - Maps



CartoDB

A web service for mapping, analyzing and building applications with data.



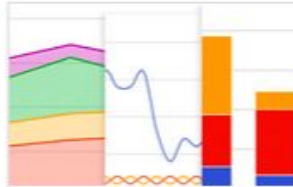
D3.js

An small, flexible and efficient library to create and manipulate interactive documents based on data.



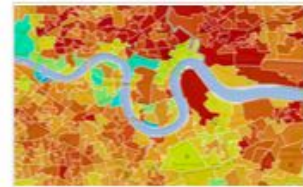
GeoCommons

A public community and set of tools to access, visualize and analyze data with compelling map visualizations.



Google Chart Tools

A collection of simple to use, customizable and free to use interactive charts and data tools.



Google Fusion Tables

A web application that makes it easy to host, manage, collaborate on, visualize, and publish data tables.



Kartograph

A simple and lightweight framework for creating beautiful, interactive vector maps.



Leaflet

A lightweight JavaScript library for making tile-based interactive maps for desktop and mobile browsers.



Many Eyes

A web application to build, share and discuss graphic representation of user uploaded data.



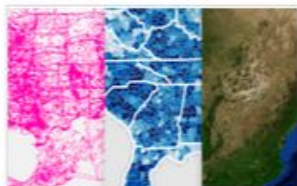
MapBox

A web platform for hosting custom designed map tiles and a set of open source tools to produce them.



Modest Maps

A display and interaction library for tile-based maps in Flash, JavaScript and Python.



Polymaps

A library for making dynamic, interactive maps with image- and vector-based tiles.

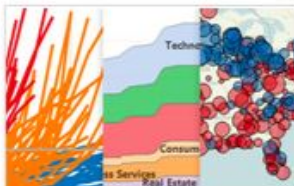


Tableau Public

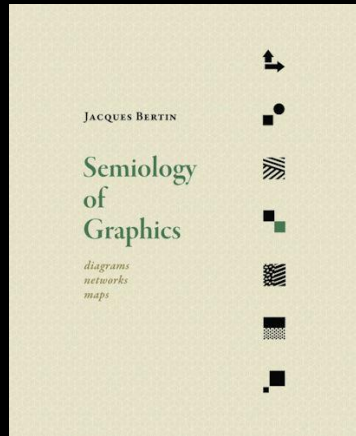
A desktop application to build and post interactive graphs, dashboards, maps and tables to the web.



Unfolding

A library to create interactive maps and geovisualizations in Processing and Java.

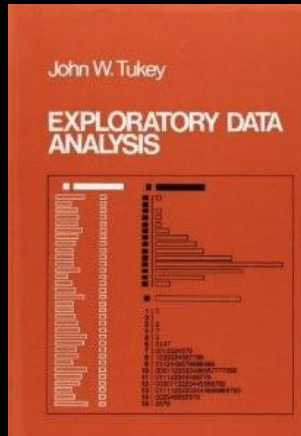
Foundational



Semiology of
Graphics

Jacques
Bertin

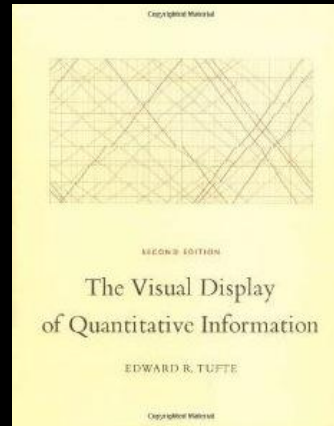
Theory



Exploratory
Data Analysis

John
Tukey

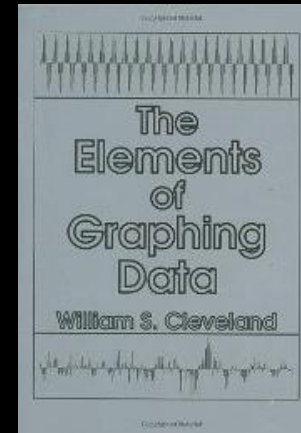
EDA



The Visual
Display of
Quantitative
Information

Edward
Tufte

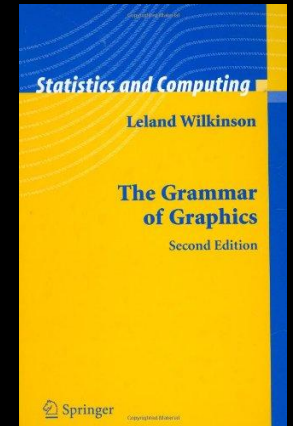
Excellence



The Elements
of Graphing
Data

William
Cleveland

Methods

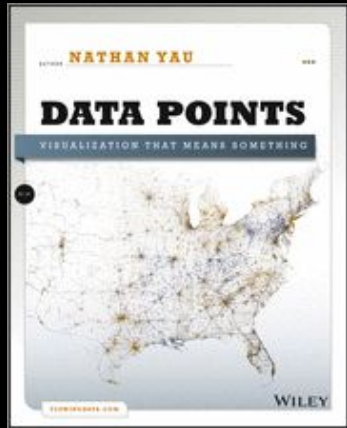


The Grammar
of Graphics

Leland
Wilkinson

Grammar

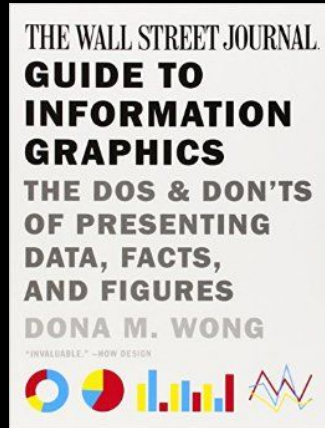
Starter Books



Data Points:
Visualisation that
means something

Nathan
Yau

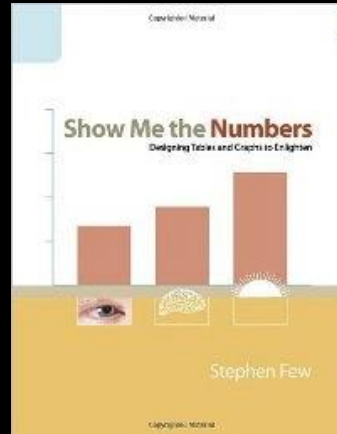
Viz Overview



Guide to
Information
Graphics

Dona
Wong

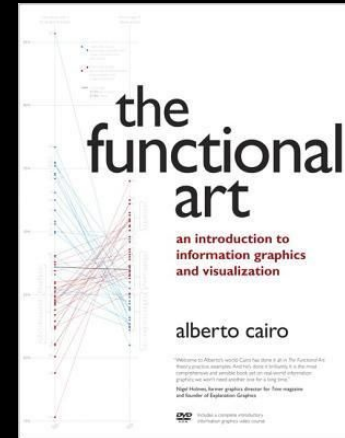
Do's & Don'ts



Show Me
the
Numbers

Stephen
Few

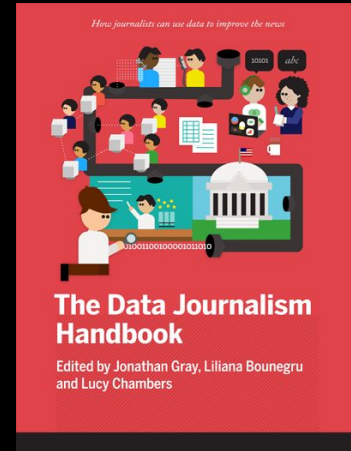
Basic Graphs



The
Functional
Art

Alberto
Cairo

Infographics

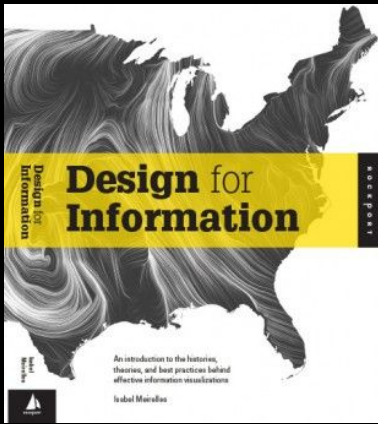


Data
Journalism
Handbook

Open
Knowledge

Journalism

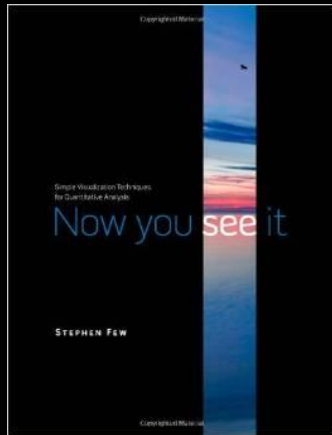
Advanced



Design for
Information

Isabel
Meirelles

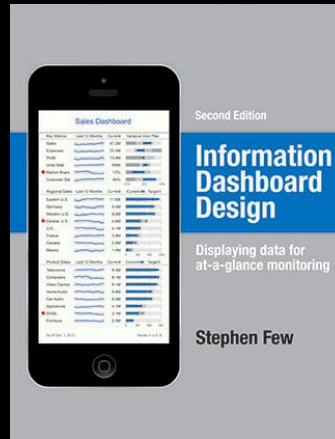
History + Best
Practices



Now you
see it

Stephen
Few

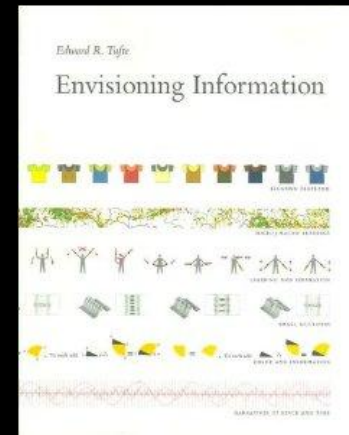
Exploratory



Information
Dashboard
Design

Stephen
Few

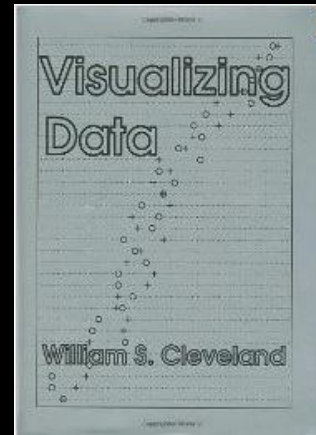
Dashboard



Envisioning
Information

Edward
Tufte

Methods

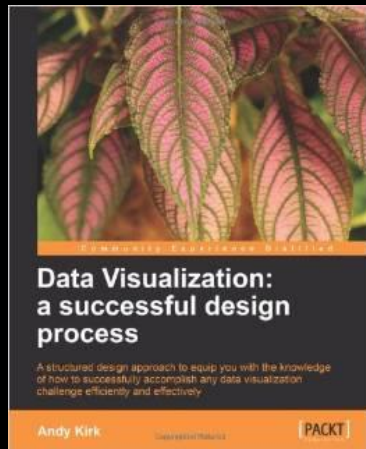


Visualising
Data

William
Cleveland

Methods

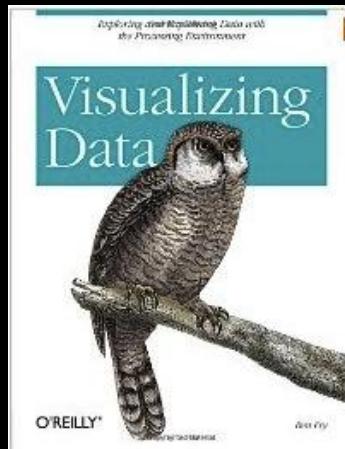
Approach & Tools



Data
Visualization: A
Successful
Design Process

Andy
Kirk

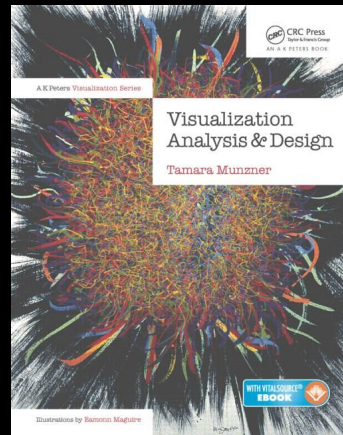
Approach
(Business)



Visualizing
Data

Ben
Fry

Approach
(Processing)



Visualization
Analysis &
Design

Tamara
Munzner

Approach
(Academic)



Interactive
Data
Visualization

Scott
Murray

D3.js

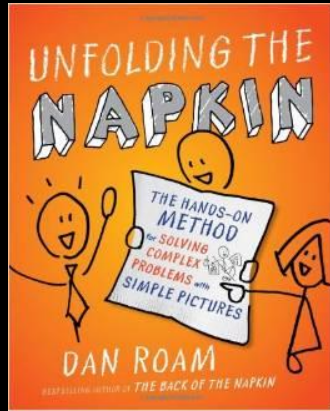


ggplot2:
Elegant
Graphics for
Data Analysis

Hadley
Wickam

R: ggplot2

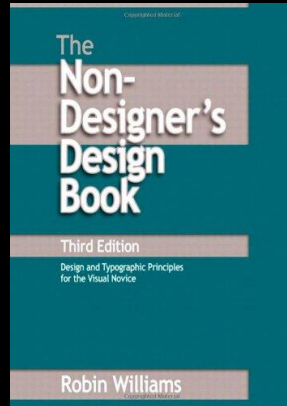
Think, Perception, Design



Unfolding the
Napkin / The
Back of the
Napkin

Dan
Roam

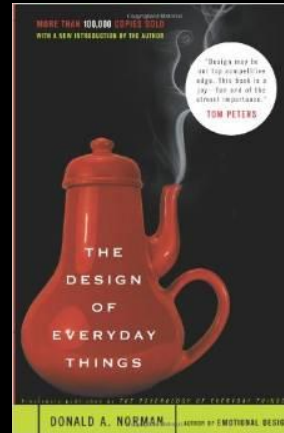
Thinking



The Non-
Designer's
Design Book

Robin
Williams

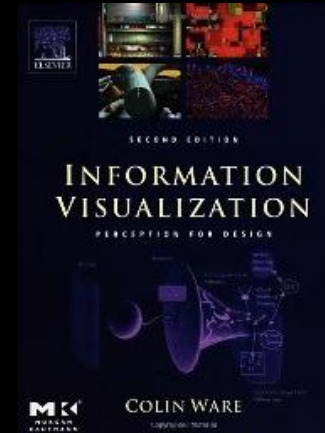
Graphic
Design



The Design of
Everyday
Things

Donald
Norman

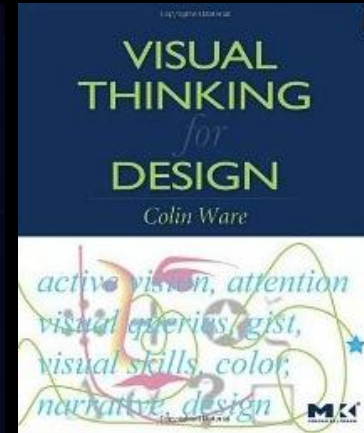
Human Object
Interaction



Information
Visualization:
Perception for
Design

Colin
Ware

Perception

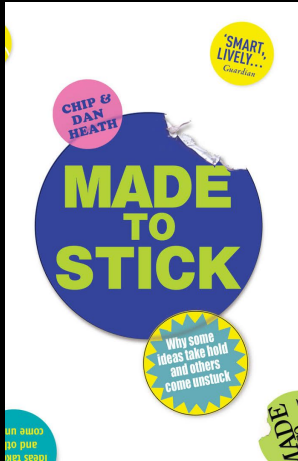


Visual
Thinking for
Design

Colin
Ware

Visual
Cognition

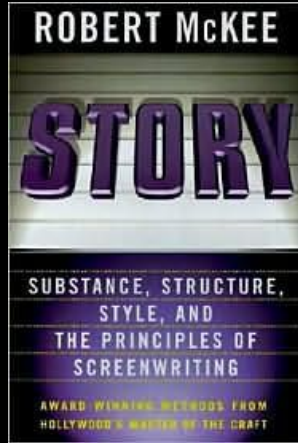
Storytelling



Made
to Stick

Chip and Dan
Heath

Stories
(Oral)



Story:
Principles of
Screenwriting

Robert
McKee

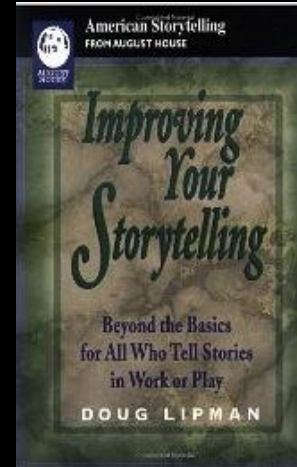
Stories
(Movies)



Making
Comics

Scott
McCloud

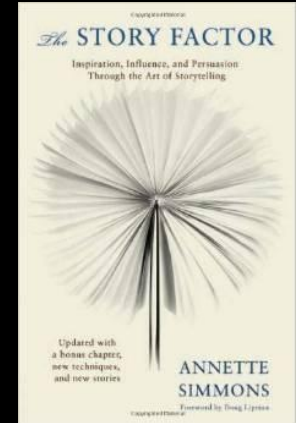
Stories
(Comics)



Improving
your
Storytelling

Doug Lipman

Oral
Storytelling



The Story
Factor

Annette
Simmons

Business
Storytelling

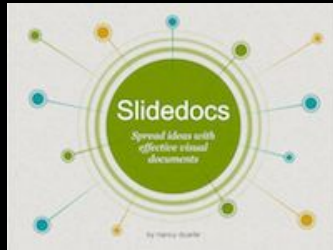
Communication



Resonate

Nancy
Duarte

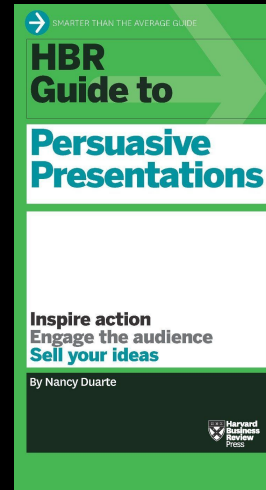
Stories



Slidedocs

Nancy
Duarte

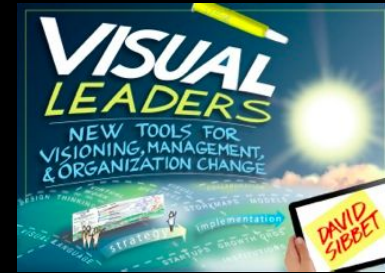
Slides +
Documents



HBR Guide to
Persuasive
Presentation

Nancy
Duarte

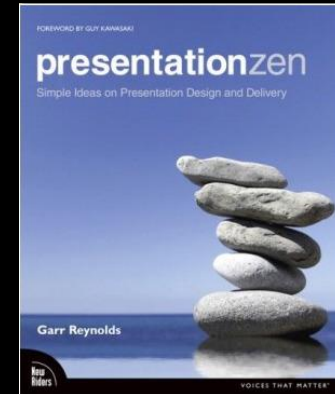
Presentation



Visual
Leaders

David
Sibbet

Visual
Comm.



Presentation
Zen

Gary Reynolds

Presentation

Blogs and Links

- [Guardian Data Blog](#)
- [New York Times](#)
- [Excel Charts](#)
- [Visualising Data](#)
- [Fell in Love with Data](#)
- [Flowing Data](#)
- [datavisualization.ch](#)
- [Eager eyes](#)
- [Junk Charts](#)
- [Information Aesthetics](#)
- [HBR - Persuading with Data](#)

[A collection of blogs by Flowing Data](#)

Courses

- Jeffrey Heer - [CSE512 Data Visualization](#)
- Tamara Munzer - [CS533-09 Information Visualization](#)
- Udacity - [Intro to Data Science](#)
- Udacity - [Exploratory Data Analysis](#)

Amit Kapoor
@amitkaps

amitkaps.com

