

Resources

Data Visualization and Data Storytelling workshop

Emmanuel Génard

Solution Architect, Data Analyst, Data Engineer, Data Scientist

IBM Global Solution Center

genard@fr.ibm.com



Objectives

- » This document contains a list of resources and references that are used throughout the course.
 - » **Data sets:** reference sites that contain free data sets classified by category
 - » **Data Visualization blogs and websites**
 - » **Tools:** links to DataViz products and technologies
 - » **Data Visualization examples:** links of all examples used throughout the course
- 

Data Sets

Data sets resources

- » **World Resources Institute** (multi domain data sets, contains data visualizations) -
<http://datasets.wri.org/dataset>
- » **Forbes** (multitude of Companies, personalities listings) - <https://www.forbes.com/lists>
- » **Statistical Computing** (selection of US related datasets) - <http://stat-computing.org/dataexpo/>
- » **Quora Wiki List of open Large Datasets** <https://www.quora.com/Where-can-I-find-large-datasets-open-to-the-public>
- » **US Data.gov** (200K Government related datasets) <https://www.data.gov/>
- » **Government data sources** (many several cities or Gov Agency publish open data)
 - ◊ **GovLoop** <http://data.govloop.com>
 - ◊ **City of Seattle** <http://data.seattle.gov>
 - ◊ **City of San Francisco** <http://data.sfgov.org>
 - ◊ **City of Chicago** <http://data.cityofchicago.org>
 - ◊ **Medicare data** <http://data.medicare.gov>
- » **1001 Datasets and Data repositories** (List of lists of lists...)
https://dreamtolearn.com/ryan/1001_datasets
- » **Kaggle** (the Home of Data Science & ML) <https://www.kaggle.com/datasets>
- » **DataCite** (find, access and reuse data) <https://search.datacite.org/>
- » **Re3data.org** (registry of research data repositories) <https://www.re3data.org/>

Websites and Blogs

The Functional Art <http://www.thefunctionalart.com/>

»Alberto Cairo's weblog

The Functional Art website homepage. The logo features a magnifying glass icon over a bar chart. Below the logo is the text "alberto cairo's weblog on visualization, infographics, and data journalism". The navigation menu includes links for Home, Contact the author, Speaking & Consulting, The Truthful Art book, The Functional Art book, Reviews, Tutorials & Resources, and Reading now.

Wednesday, November 15, 2017

Mapping HIV

I've been meaning to write this post for quite a long time. Months ago, I discovered [AIDSVu.org](#), an initiative by **Emory University's Rollins School of Public Health** in partnership with Gilead Sciences that tracks HIV in the U.S. The data comes from **CDC**.

As I liked what I saw, I wrote some quick notes but, as it often happens, completely forgot about them. Then, three weeks ago I met some of its creators at the **World Conference of Science Journalists**, where I gave a talk (here's [a summary](#) of what I said.) I promised that I'd recover the notes and publish them, as this is a visualization project that is worth your time.

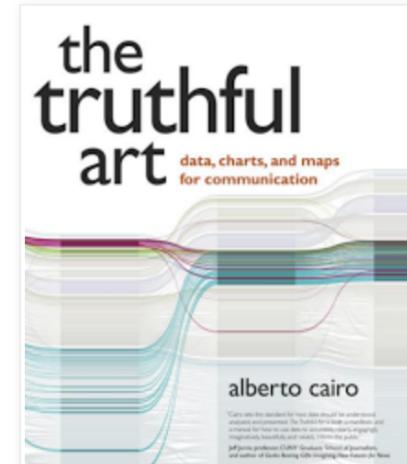
Once you enter the [AIDSVu.org](#), it'll locate you —if you don't block location services— and display statistics that pertain to your area. After that, you can jump to a **national map** with tons of filters:



Follow by Email

Submit

NEW BOOK AVAILABLE



ORDER IT NOW

The Functional Art book

Information is Beautiful

<http://www.informationisbeautiful.net/>

»David McCandless's weblog

≡ information is beautiful

About Blog Data Books Work With Us Training Jobs Contact [f](#) [t](#) [r](#)

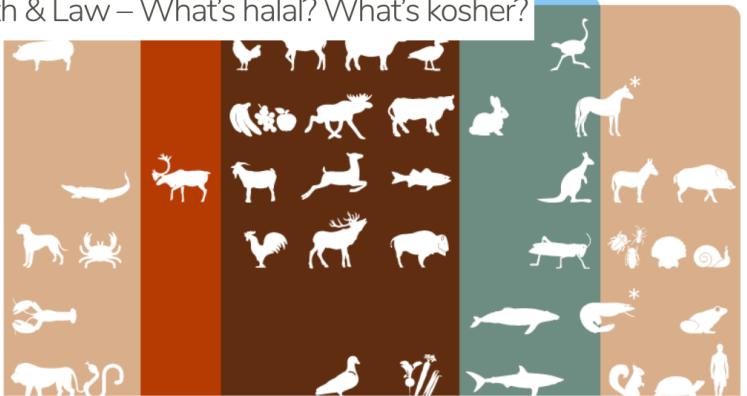


KANTAR
Information is Beautiful
Awards 2017

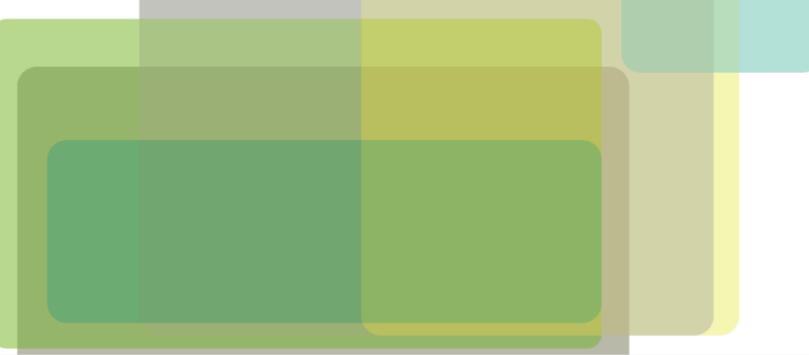
the best dataviz & infographics of the year
[explore the shortlist](#)

POPULAR: interactive belief food health ideas science web-tech

Tooth & Law – What's halal? What's kosher?



The UK's Brexit Options in the Eurozone Landscape – Visualized



TED

https://www.ted.com/talks/david_mccandless_the_beauty_of_data_visualization



Recent [SEE ALL →](#)

Global cycling and running heatmap

A few years back, cycling and running app Strava mapped the paths of its users. Now with a lot more ...

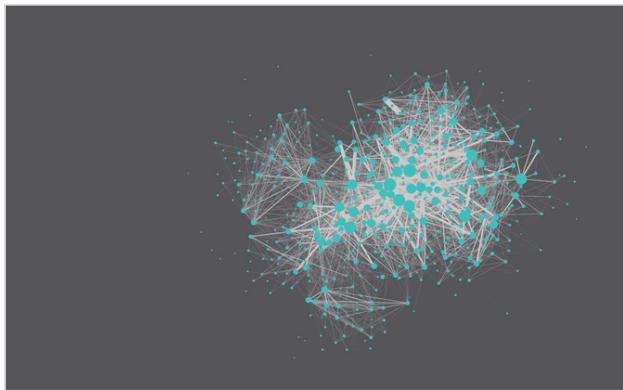


Scrollama.js, a JavaScript library for scrollytelling

Russell Goldenberg released Scrollama.js in an effort to make scrollytelling more straightforward to implement. Scrollytelling can be complicated to implement ...



MEMBERS ONLY



Switching Jobs

When people move to different jobs, here's where they go.



Become a member.
*Learn to visualize your data.
From beginner to advanced.*

WHAT YOU GET

Categories

Visualization

Seeing data

Statistics

Analyzing data

Maps

Seeing geographic data

Software

Working with data

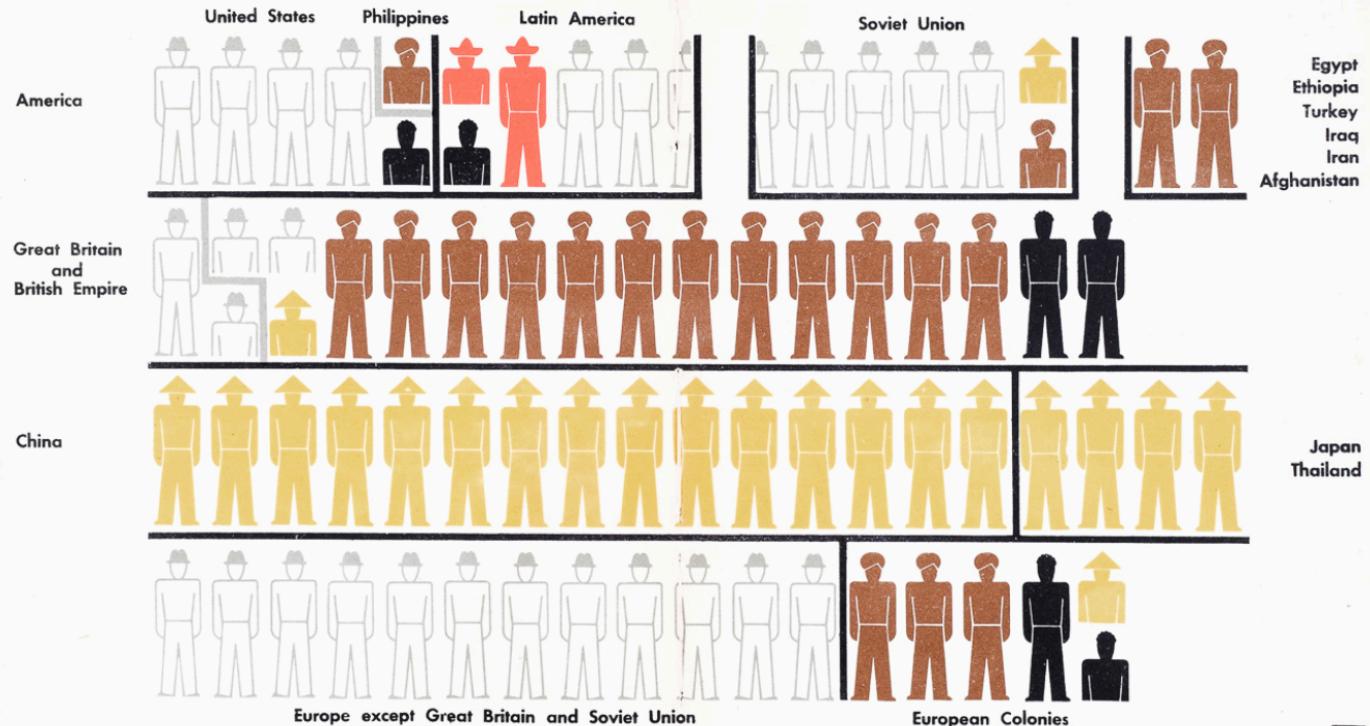
Infographics

Explaining data

Sources

Getting data

United States and Great Britain in the World



Each complete figure represents 30 million population

Here we are—the peoples of the world—black, white, yellow and brown. The man and a half in Latin America are Red Indians ; there are so few left in the United States that it would be difficult to carve a man up small enough to show them. The brown men are mostly Indians. It is surprising to find that there are more Negroes in the British Empire than in the United States.



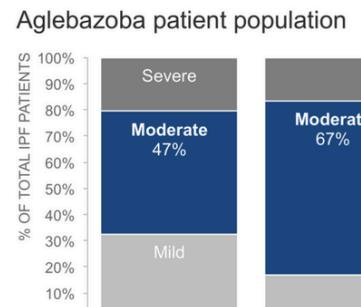
<http://www.storytellingwithdata.com/gallery/>

storytelling WITH data®

BLOG BOOK PUBLIC WORKSHOPS CUSTOM WORKSHOPS GALLERY ABOUT CONTACT

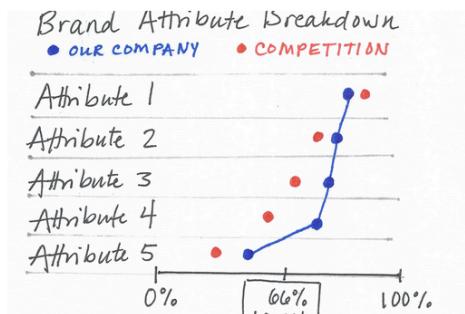
Browse the gallery to see featured blog posts (scroll down and right).

makeovers



a tale about opportunity

[Read More →](#)



novel vs. the boring old bar chart



visualizing change via slopegraph

VizWiz

Data Viz Done Right



MY PORTFOLIO

MAKEOVER MONDAY

TABLEAU TIP TUESDAY

WORKOUT WEDNESDAY

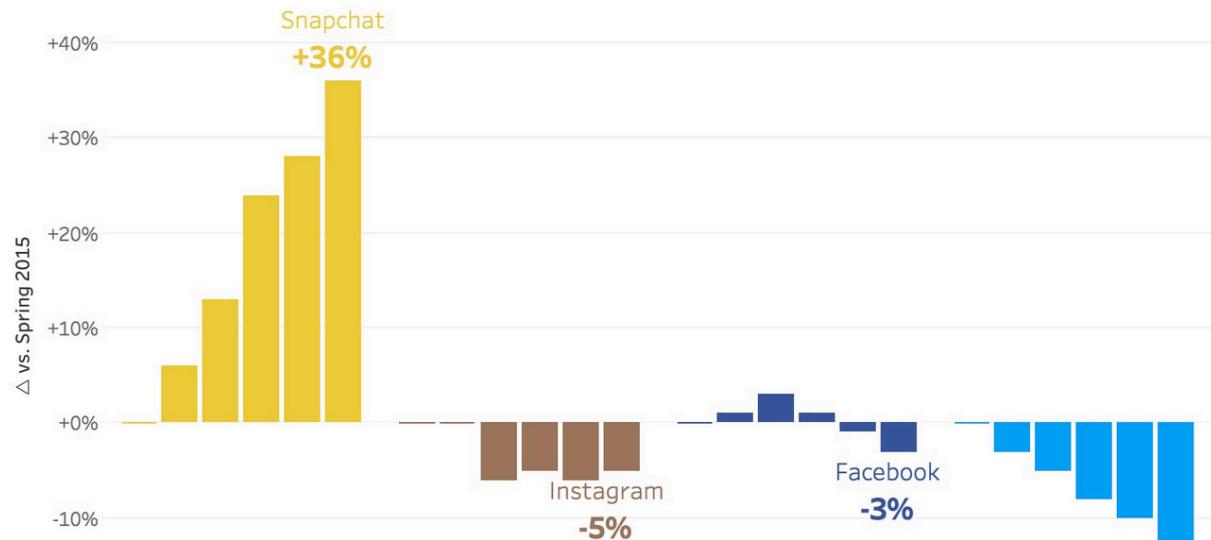
DATA

Makeover Monday: Snapchat is tops with American teens

Business Insider, facebook, instagram, Makeover Monday, snapchat, social media, teens, trends, twitter 4 comments

Snapchat is tops with American teens

Since the spring of 2015, Snapchat has seen a 37% increase in teens preferring their app, whereas Instagram, Facebook and Twitter have all seen declines.



Tools



Tools

- » **R and Rstudio** popular tools for data scientists and analysts with scripting and limited programming experience. (++ numerous packages and guides available to users)
 - » **Python** general purpose programming language focusing on readability and simplicity. Easy learning curve. Large collection of Libraries
 - » **IBM DataScience Experience** fully integrated workbench supporting several languages (Scala, R, Python and Java)
 - » **IBM SPSS Statistics and SPSS Modeler** full range of data mining capabilities
 - » **IBM Watson Analytics** self-service analysis empowered by Watson Cognitive capabilities
 - » **Microsoft Excel** needless to present
 - » **Goolge Docs** if you don't have a Microsoft Office licence.
 - » **Tableau Public** intuitive drag and drop experience for data analysis and data visualization
 - » **RapidMiner Studio** covers entire life-cycle of prediction modeling, starting from data preparation to model building and deployment.
 - » **QlikView**
 - » **D3** javascript library for building interactive data visualizations within web applications. Highly customizable and foundation of many DataViz libraries.
 - » **Data Wrangler** interactive tool for data cleaning and transformation
- 

DataViz Examples



Many DataViz examples are available from the website and blogs already listed.

- » **Mike Bostock's Blocks** by the founder of D3.js
<https://bl.ocks.org/mbostock>
- » **D3.js** homepage to the Data-Driven Document library <http://d3js.org>

- » **Tableau Public** data visualization community for Tableau Developers
<https://public.tableau.com>
- » **Visme** a blog with great links to Data Visualizations
<http://blog.visme.co/best-information-graphics-2016/>

... and so many others ...
Google is your friend