

Tools



Objectives

After completing this section of the course you should be able:

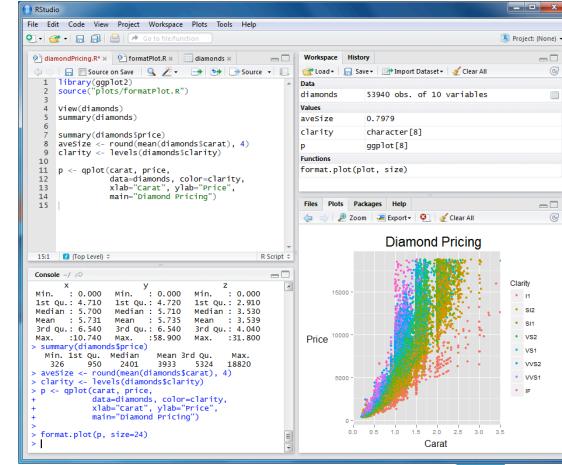
- » **Programming tools and open source libraries**
 - » List “Designer” tools for Data Visualization
 - » Know ISV offerings
 - » Differentiate Infographics and Data Visualization
- 

DataViz Libraries and Programming tools

R and RStudio

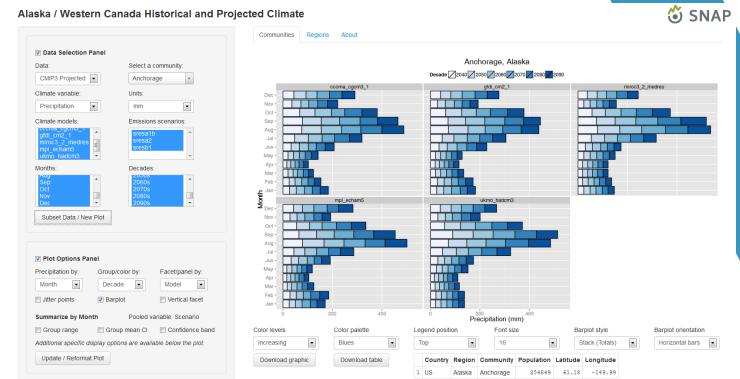
- » R is an interpreted computer language
- » Designed for Data Manipulation and Data Analysis
- » R Studio as IDE, Shiny
- » Extendable 800+ of packages

- » Ggplot2, ggigraph, dygraphs
- » googleVis
- » RColorBrewer
- » Flexdashboard
- » ...



A screenshot of the RStudio interface. The top menu bar includes File, Edit, Code, View, Project, Workspace, Plots, Tools, Help, and a tab for 'Get Started'. The left sidebar shows a project named 'diamondPricing.R' with files 'diamonds.R' and 'formatPlot.R'. The main workspace shows a scatter plot titled 'Diamond Pricing' with 'Carat' on the x-axis and 'Price' on the y-axis. The plot is color-coded by 'Clarity' levels. The bottom console window displays R code for data manipulation and plotting.

```
1 diamonds <- read.csv("diamonds.csv")
2 source("plots/formatPlot.r")
3 diamonds$carat <- round(diamonds$carat, 4)
4 diamonds$clarity <- levels(diamonds$clarity)
5 summary(diamonds)
6 nrow(diamonds)
7 summary(diamonds$price)
8 aveSize <- round(mean(diamonds$carat), 4)
9 clarity <- levels(diamonds$clarity)
10 p <- ggplot(carat, price)
11 p + geom_point(aes(color=clarity),
12 + xlab="Carat", ylab="Price",
13 + main="Diamond Pricing")
14 format.plot(p, size=24)
15 |
```

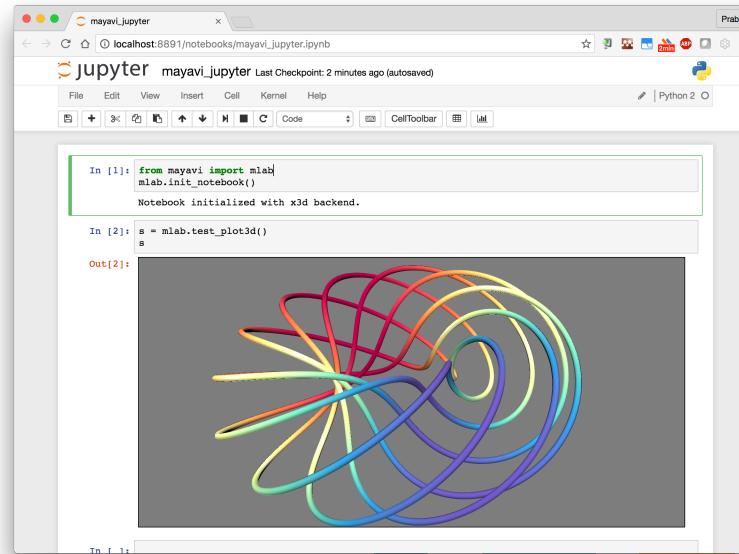


<http://cran.r-project.org/>

Python

- » Python is an interpreted computer language
- » Very active scientific computing community
- » Extendable thousands of libraries

- » Scipy, Matplotlib
- » Mayavi
- » Chaco
- » ...



The screenshot shows a Jupyter Notebook window titled "mayavi_jupyter" running on "localhost:8891/notebooks/mayavi_jupyter.ipynb". The notebook interface includes a toolbar with various icons, a menu bar with File, Edit, View, Insert, Cell, Kernel, and Help, and a code editor area. In the code editor, two cells are visible:

```
In [1]: from mayavi import mlab  
mlab.init_notebook()  
Notebook initialized with x3d backend.  
  
In [2]: s = mlab.test_plot3d()  
s
```

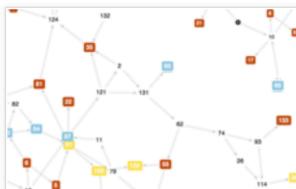
The output of the second cell, "Out[2]", displays a 3D plot consisting of several nested, multi-colored elliptical curves (red, orange, yellow, green, blue, purple) forming a complex, twisted shape.



- » **JavaScript** library for manipulating documents based data
- » Brings data life
- » Open web standards
- » Core of many Frameworks



<http://d3js.org>

**Arbor.js**

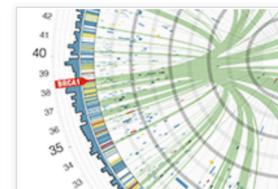
A library of force-directed layout algorithms plus abstractions for graph organization and refresh handling.

**CartoDB**

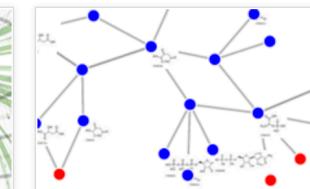
A web service for mapping, analyzing and building applications with data.

**Chroma.js**

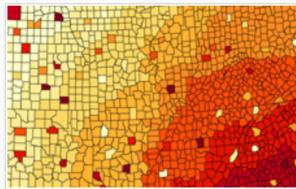
Interactive color space explorer that allows to preview a set of linear interpolated equidistant colors.

**Circos**

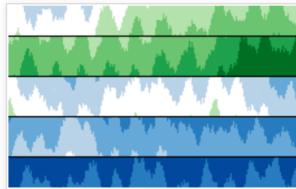
A software package for visualizing data in a circular layout.

**Cola.js**

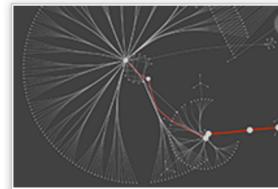
A library for arranging networks using constraint-based optimization techniques.

**ColorBrewer**

A web tool for selecting colors for maps.

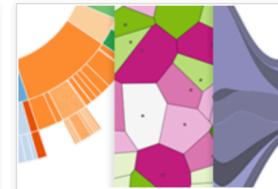
**Cubism.js**

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

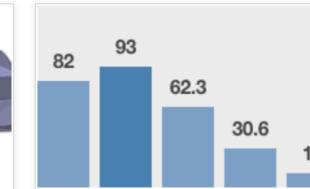
**Cytoscape**

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





visualising
data

HOME BLOG RESOURCES REFERENCES TRAINING SERVICES BOOK ABOUT



24 OCT ADOBE SNEAK PEAK: PROJECT LINCOLN >>

DATA HANDLING

CHARTING

PROGRAMMING

MULTIVARIATE

MAPPING

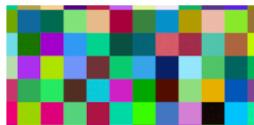
WEB-BASED

SPECIALIST

COLOUR

10 Best Practices for Creating Effective Dashboards

+ a b l e a u



0 TO 255



ABBYY



ABLE2EXTRACT



ADIOMA



ADOBE AFTER EFFECTS



ADOBE ANIMATE



ADOBE COLOR

Intuitive Workflow for Data Blending and Advanced Analytics



ADOBE EDGE



ADOBE ILLUSTRATOR



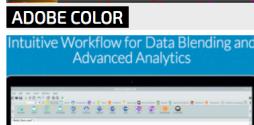
AESOP STORY ENGINE



AFFINITY DESIGNER



AI2HTML



ALTERYX



ANIMAPS



ANYCHART



APPS FOR EXCEL



ARBOR.JS



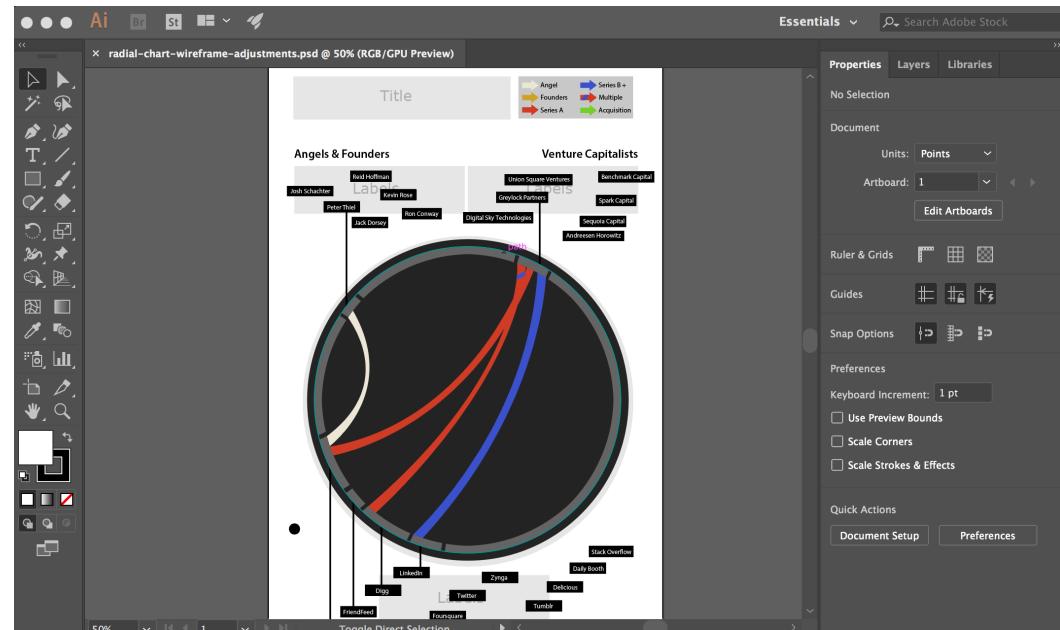
ARCGIS

<http://www.visualisingdata.com/resources>

Designer Tools

Adobe Illustrator & Photoshop

- » Designer kind of Experience
 - » Perfect for Visualization Posters and Infographics
 - » **No analytics** capabilities
 - » **Illustrator** has basic Chart tools embedded
 - » Accepts imports of Excel graphs





Other tools with designers UX

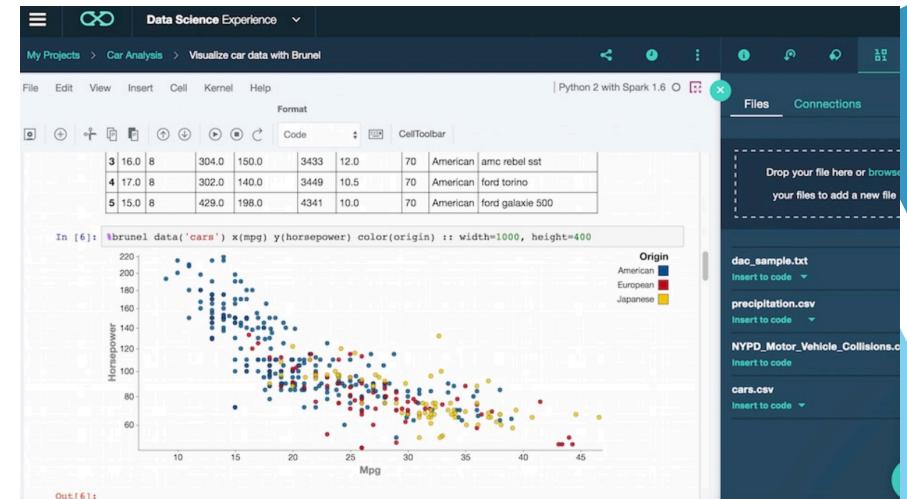
- » **Infogram:** Best for creating interactive charts and infographics
 - » **ChartBlocks:** Best for creating simple charts in minutes
 - » **Raw:** Best for visualizing complex data in vector format
 - » **Datamatic:** Best for quickly creating a wide variety of charts
 - » **Silk:** best for creating interactive stories with datasets
 - » **Datahero:** Best for rapidly visualizing data from various sources
 - » **Quadigram:** Best for combining texts, images and interactive visualizations to bring your data to life
- 

Independant Software Vendors

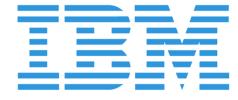


IBM Data Science Experience

- » Single place to learn, create and collaborate
- » Community-based, tutorial rich
- » Cloud-based or local
- » Any programming language (R, Python, Scala, Java)
- » Rstudio and RShiny embedded
- » Fully Spark enabled
- » ...



<http://datascience.ibm.com>



IBM Watson Analytics

- » Self-service Analytics Platform
- » Predictive analytics embedded
- » No programming
- » NLP capabilities
- » Social Media Analysis
- » ...

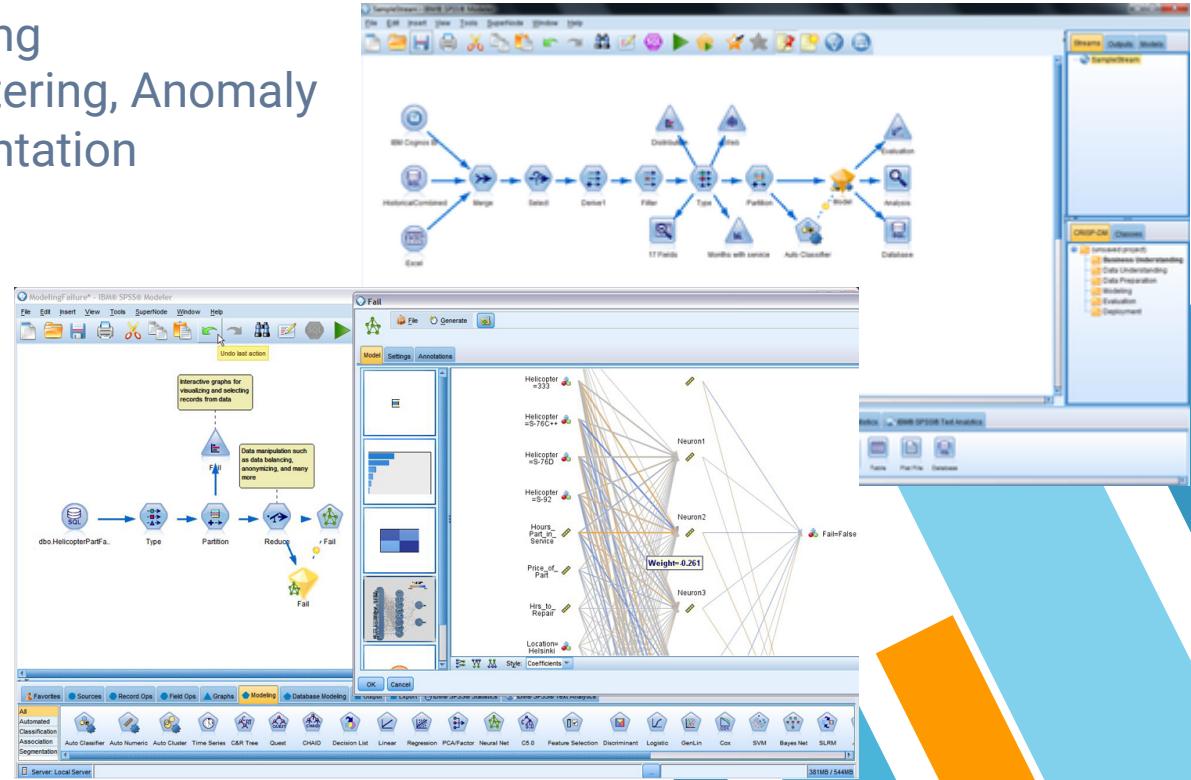
**“Data scientist-like
capabilities meets
iPhone-like usability”**



IBM SPSS Statistics, Modeler

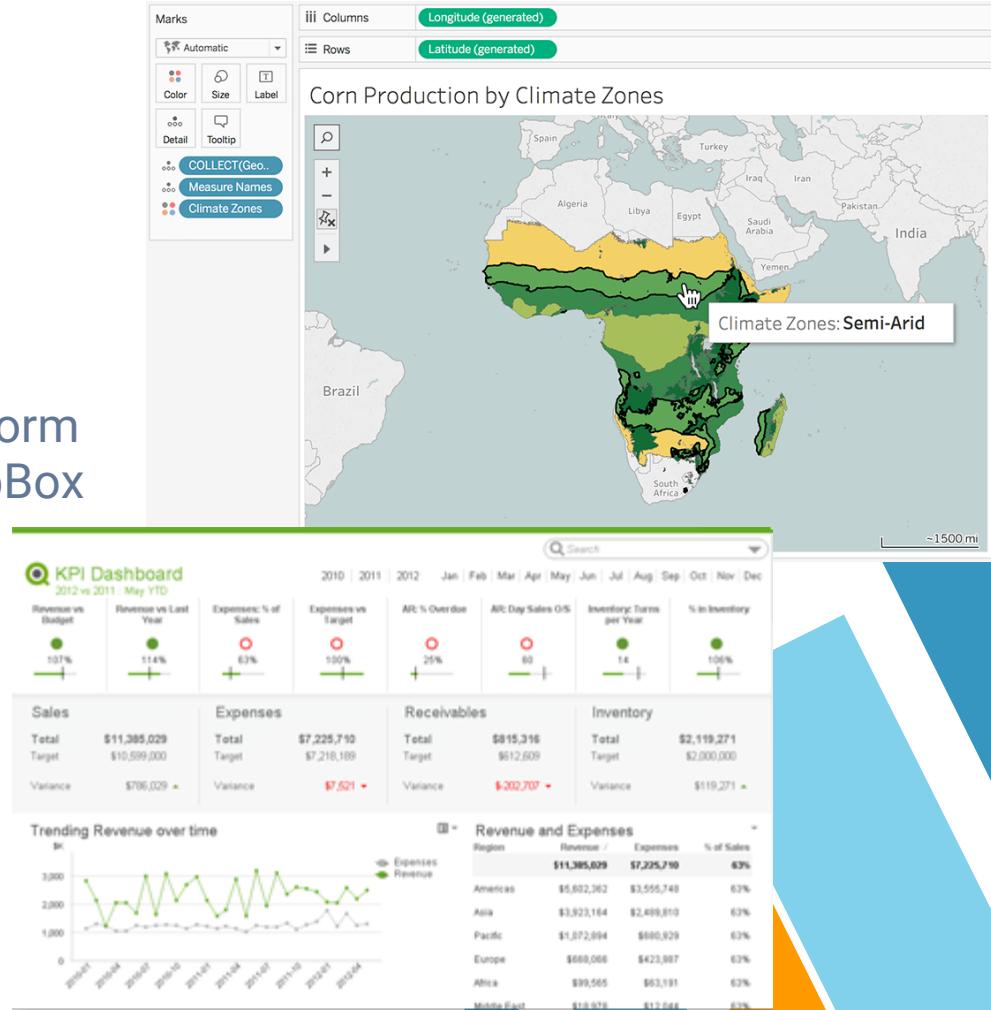
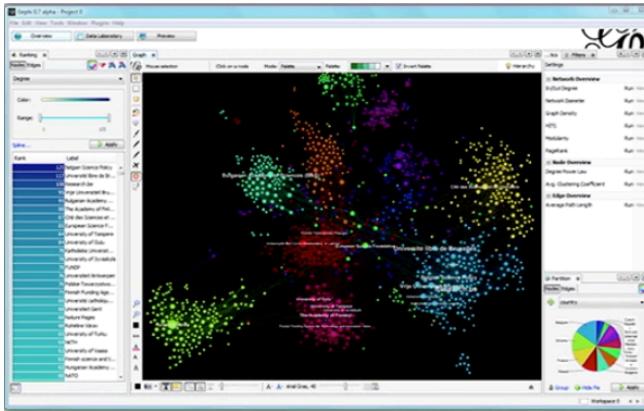


- » Data-mining workbench
- » Predictive modeling
- » Association, Clustering, Anomaly detection, Segmentation algorithms
- » No programming
- » Text mining
- » ...



And...

- » Tableau Software
- » QlikView
- » MatLab
- » Microsoft Excel
- » Gephi: Open Graph Viz Platform
- » Google Maps, CartoDB, MapBox
- » ...



Infographics vs Data Visualization

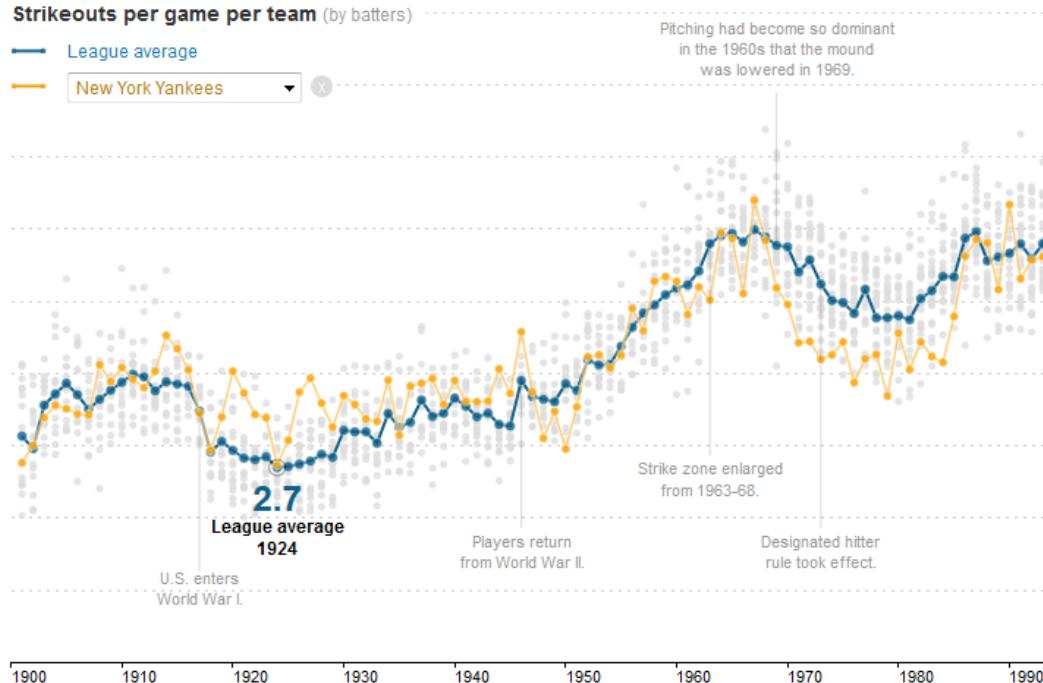
Data Visualization vs. Infographics

Data Visualization

- » In Data Visualization programs create the graphics
 - ◊ Needs data as inputs
 - ◊ Algorithms creates the graphics (computer analysis)
 - ◊ Visualization
- » Focus on Exploration and Understanding of the data
- » Less attention to the “flashiness”
- » Interactive
- » Capable of handling Large volume of Data

Strikeouts on the Rise

There were more strikeouts in 2012 than at any other time in major league history.



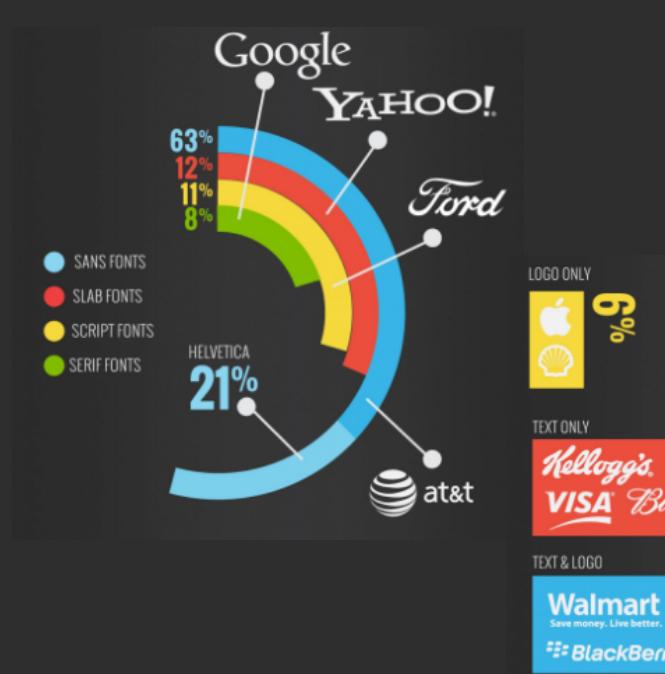
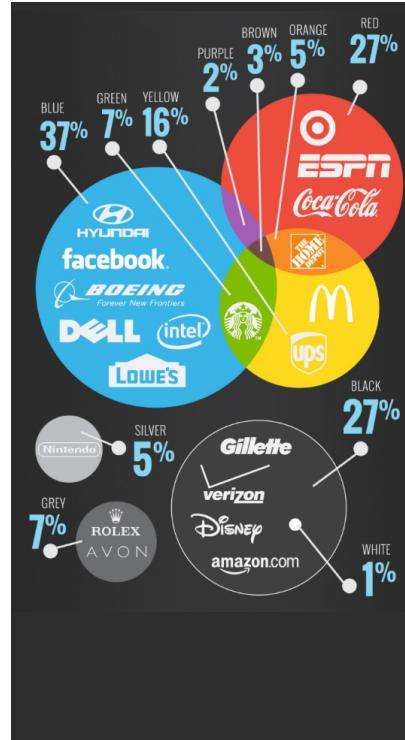
Data Visualization vs. Infographics

Infographics

- » **Infographics** have graphics that are created before hand
 - ◊ Not rendered programmatically
 - ◊ It is pre-generated
 - ◊ Specific graphics tools like Illustrator & PhotoShop
- » Greater interest on the editorial process
- » Design to entertain
- » Less data, data pre-processed
- » Focus on the **conclusions**

“Information presented for a particular audience, with context and with purpose to persuade”

<http://visual.ly/fonts-colors-drive-world's-top-brands>





THANKS!

Any questions?

You can find me at

- » [@manuGenard](https://twitter.com/manuGenard)
 - » genard@fr.ibm.com
- 