

Data Visualization

Introduction

Andrea Marchetti

Obiettivi della lezione

Perchè visualizzare i dati

Il sistema visivo umano

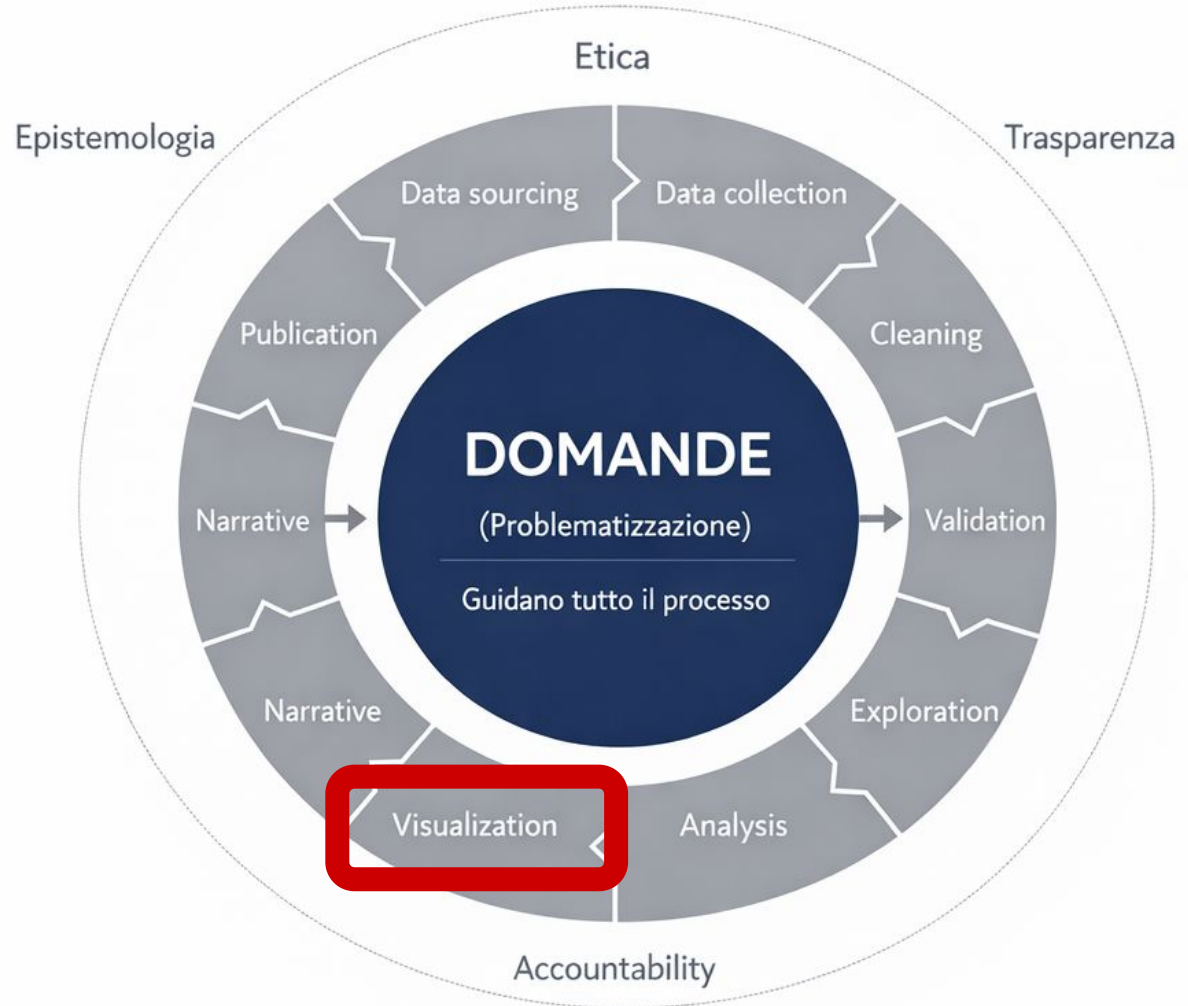
Come fare una buona visualizzazione

Come scegliere il grafico giusto

Errori di visualizzazione (encoding)

Fasi del workflow

1. Data sourcing
2. Data collection
3. Cleaning
4. Validation
5. Exploration
6. Analysis
7. Interpretation
- 8. Visualization**
9. Narrative integration
10. Publication



Perchè visualizzare i dati

I dati sono ovunque

I dati grezzi non sono comprensibili

- Bisogna passare dai dati alla conoscenza
- Accedere a fenomeni complessi

!!! Combattere il sovraccarico informativo !!!

Un immagine vale più di 1000 parole



[London 1940 after bombing](#)

Cos'è una visualizzazione

- una rappresentazione grafica di dati astratti
(numeri, testi, dati geografici, temporali, ecc.)
- uno strumento di analisi, comunicazione e comprensione
che sfrutta il **sistema visivo umano**

Obiettivi della data visualization

Registrare informazioni

- Fotografie, ECG, ...

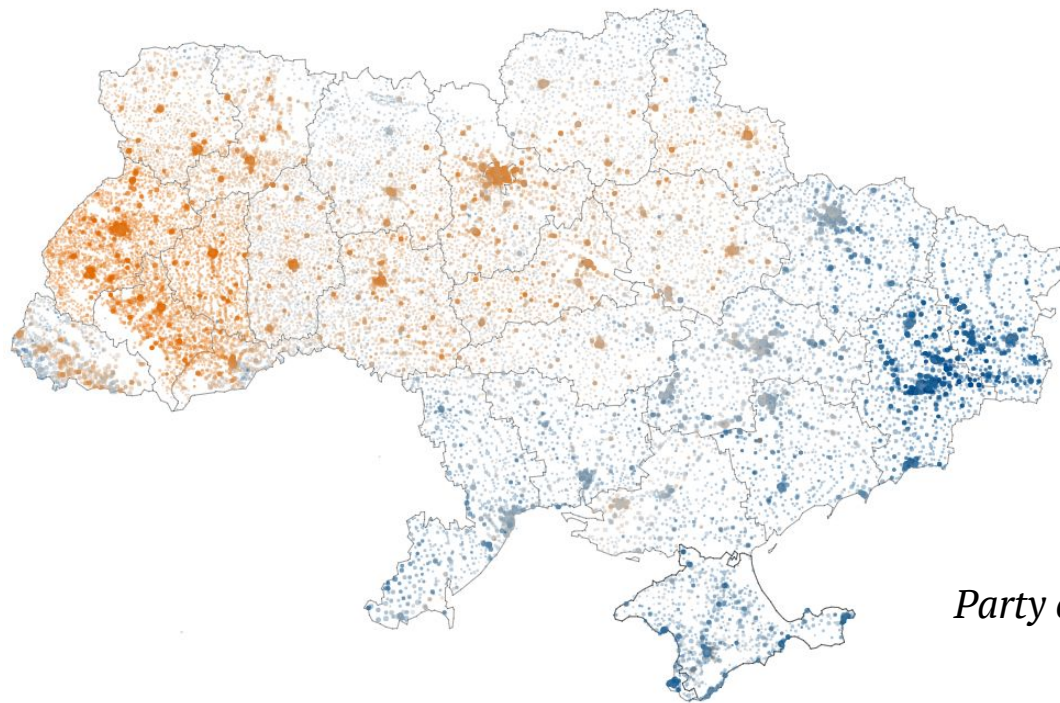
Esplorare i dati

- Verificare risposte o ipotesi
- Generare nuove domande
- Identificare pattern o outlier

Comunicare

- Condividere o mettere in evidenza gli insight emersi dai dati
- Convincere

Map of 2012 Ukrainian Election Results



Party of Regions (pro Russia)



Vs

Fatherland Party (pro Europe)



Il sistema visivo umano

I colori non ci aiutano, ma la nostra capacità di riconoscere i pattern, ci viene in soccorso.

Riesci a vedere il leopardo?



Identificare le forme



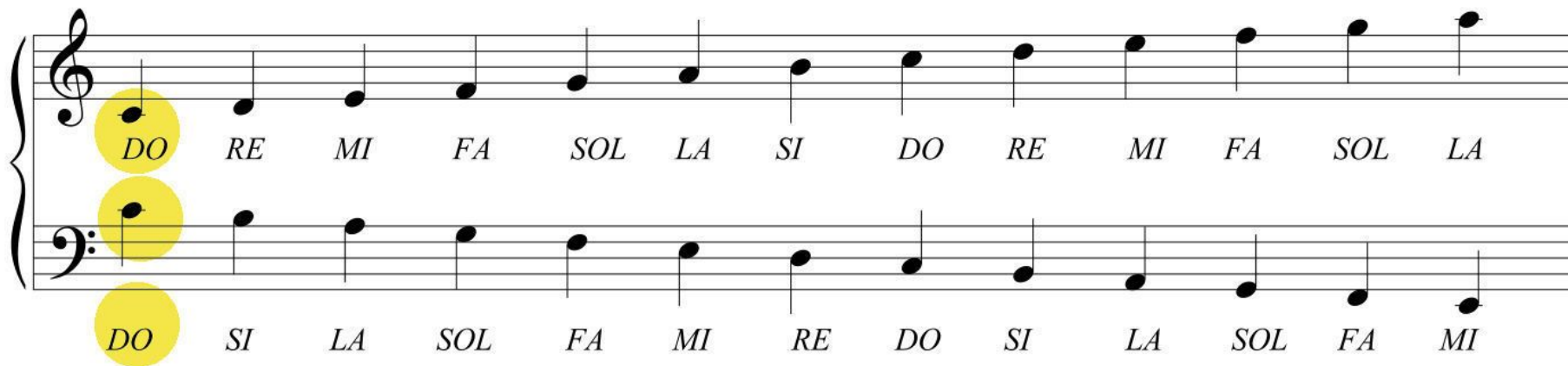
Patterns, Patterns Everywhere



Un pattern può essere definito formalmente come una regolarità osservabile nel mondo naturale e in quello artificiale, che si ripete in modo prevedibile.

Immagini Vs Testo

Il nostro **sistema visivo** è più veloce a decodificare **immagini** che **testo**



A musical score for a scale exercise on a grand staff. The treble clef staff contains the notes DO, RE, MI, FA, SOL, LA, SI, DO, RE, MI, FA, SOL, LA. The bass clef staff contains the notes DO, SI, LA, SOL, FA, MI, RE, DO, SI, LA, SOL, FA, MI. The first note in each staff, 'DO', is highlighted with a yellow circle.

DO RE MI FA SOL LA SI DO RE MI FA SOL LA

DO SI LA SOL FA MI RE DO SI LA SOL FA MI

Visualization System

Sfruttando l'ampia capacità del sistema visivo umano, è possibile:

- cogliere rapidamente grandi quantità di informazioni
- interpretare i dati più facilmente
- riconoscere pattern e tendenze in un insieme di dati

Come la data visualization sfrutta il sistema visivo umano



```
110100100101  
011100110101  
110110101010  
110101011100  
101010010101  
000111000111
```

Raw data



Data analysis



Analytical abstraction



Visual encoding



Visualization



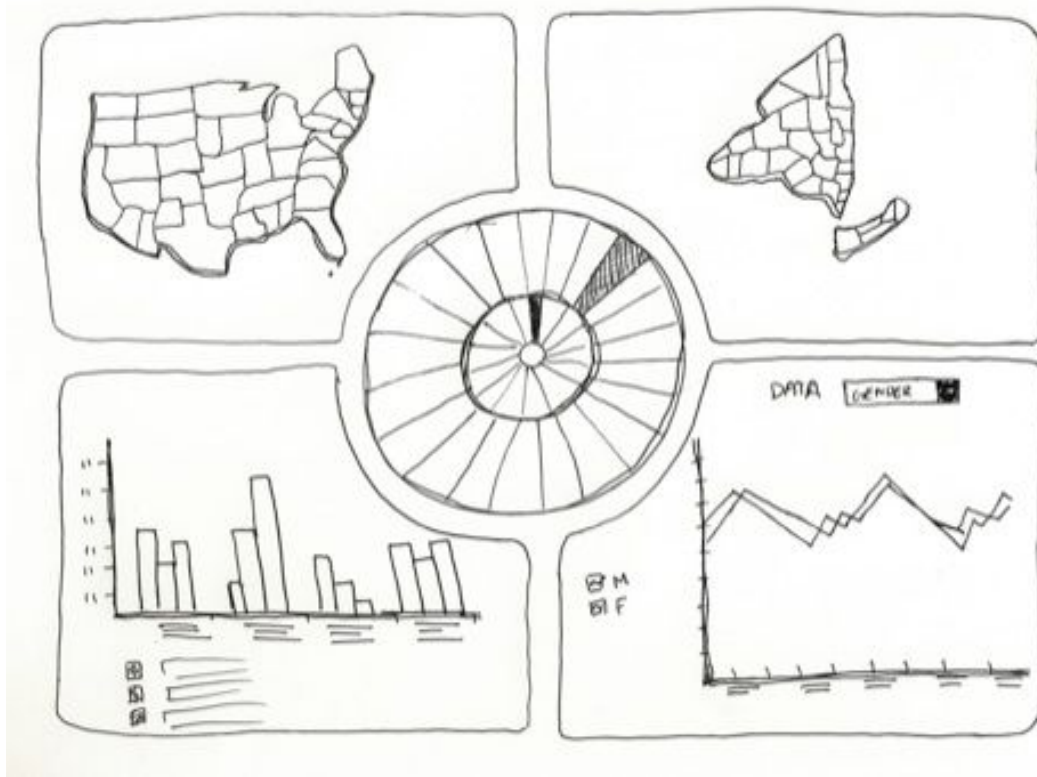
Visual decoding



Target user's insights

Come fare una buona visualizzazione

Fai uno shizzo della tua idea



False

‘Infographics is about designing cool pictures’

- Creating something attractive will attract readers.
- **After attracting readers, you also need to deliver something useful**
- Give a context to understand

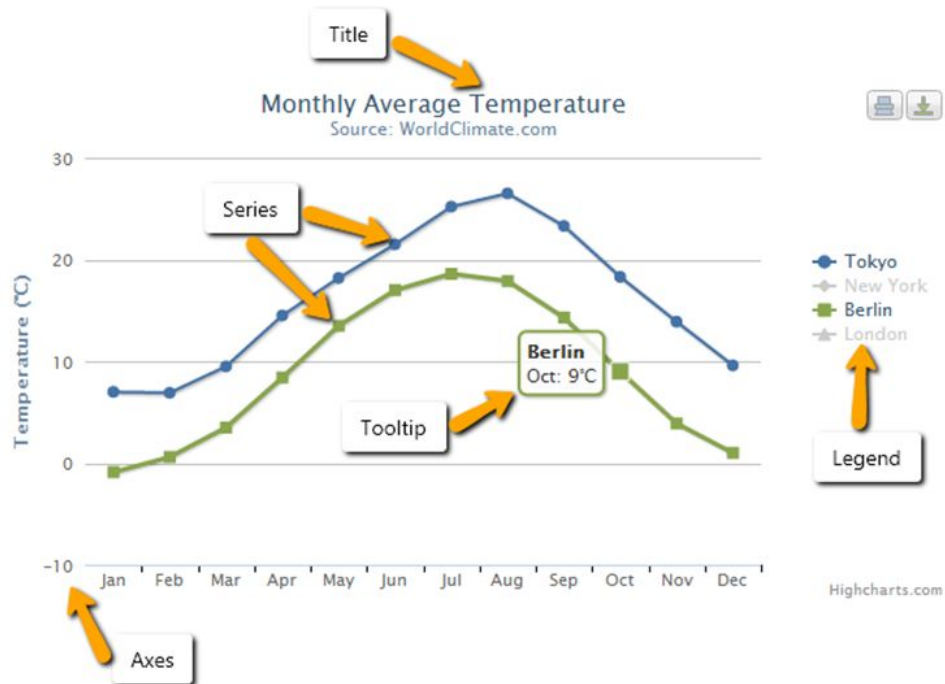
Un falso mito della data visualization

“L’infografica è solo una questione di immagini belle e accattivanti”

- Un design attraente aiuta ad attirare i lettori.
- Ma non basta catturare l’attenzione
- Bisogna anche comunicare qualcosa di utile

1 Fornire il contesto

E' veramente frustrante non capire

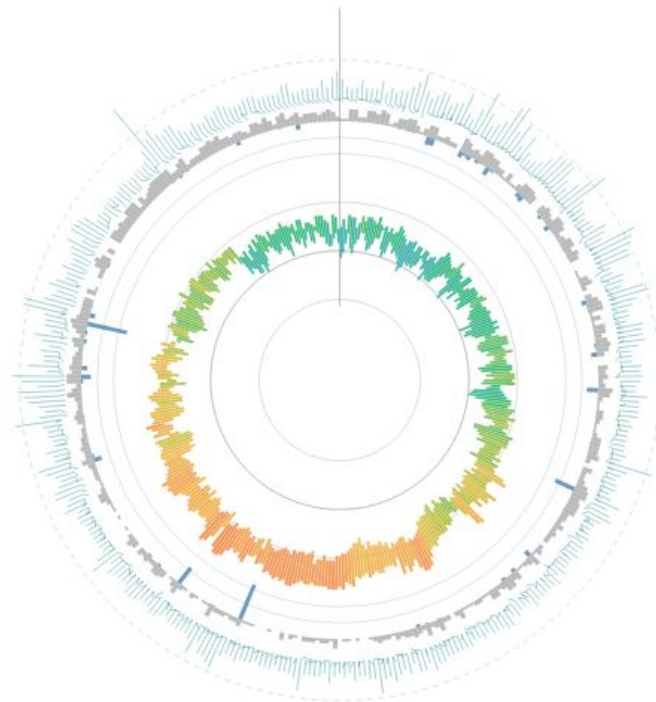


“Overview first, zoom and filter, then details-on-demand”

Prima la **visione d’insieme**: capire subito struttura e ampiezza del dataset

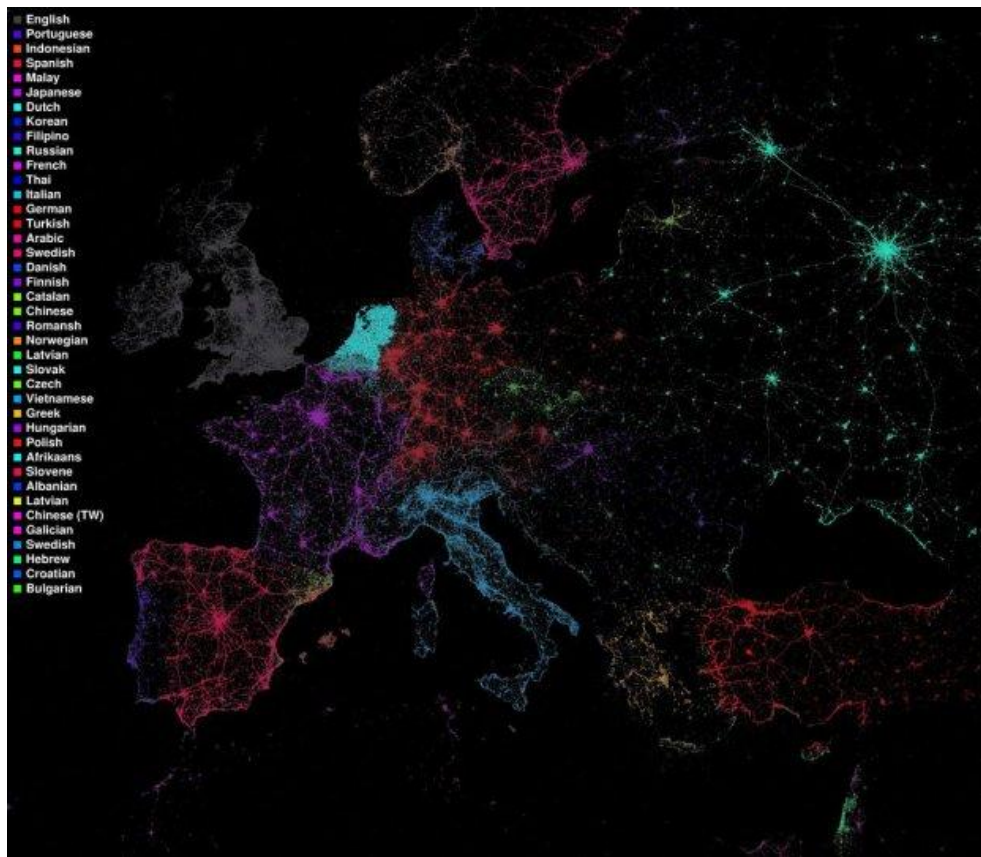
Poi **zoom e filtri**: concentrarsi su ciò che conta, riducendo il rumore

Infine i **dettagli** su richiesta: mostrare informazioni puntuali solo quando servono



The Weather Wheel

Le eccezioni sono le cose più utili



Language
communities of
Twitter – Eric
Fischer –
FlowingData:

Fondamenti della Data Visualization

pensare visivamente, progettando prima su carta

non solo attrarre, ma anche rendere i dati comprensibili

offrire una visione d'insieme e poi i dettagli

usare la forma visiva per far emergere insight dai dati

Come scegliere il grafico giusto

Cosa voglio comunicare

La prima domanda non è “quale grafico mi piace di più?”, ma “che cosa voglio comunicare?”.

Il grafico va scelto in base

- alla relazione che vuoi mostrare nei dati
- alle domande a cui rispondere.

Cosa vuoi comunicare?

confronto

gerarchia

rete

evoluzione temporale

pattern geografico

dati spazio-temporali

testi

distribuzione

relazione tra variabili

composizione / parte del tutto

ranking

scostamento da un valore di riferimento, incertezza

flussi e movimenti

Grafici

Confronto

Bar chart · Column chart · Dot plot · Lollipop chart

Gerarchia

Treemap · Sunburst · Tree diagram

Reti

Network diagram · Node-link diagram · Arc diagram

Evoluzione temporale

Line chart · Area chart · Slope chart · Timeline, Horizon chart

Pattern geografico

Choropleth map · Symbol map · Dot map

Dati spazio-temporali

Small multiples maps · Animated map · Map + timeline

Testi

Word cloud

Distribuzione

Histogram · Box plot · Violin plot · Density plot

Relazione tra variabili

Scatterplot · Bubble chart · Heatmap

Composizione / parte del tutto

Stacked bar chart · 100% stacked bar chart · Pie chart · Treemap

Ranking

Sorted bar chart · Lollipop chart · Bump chart

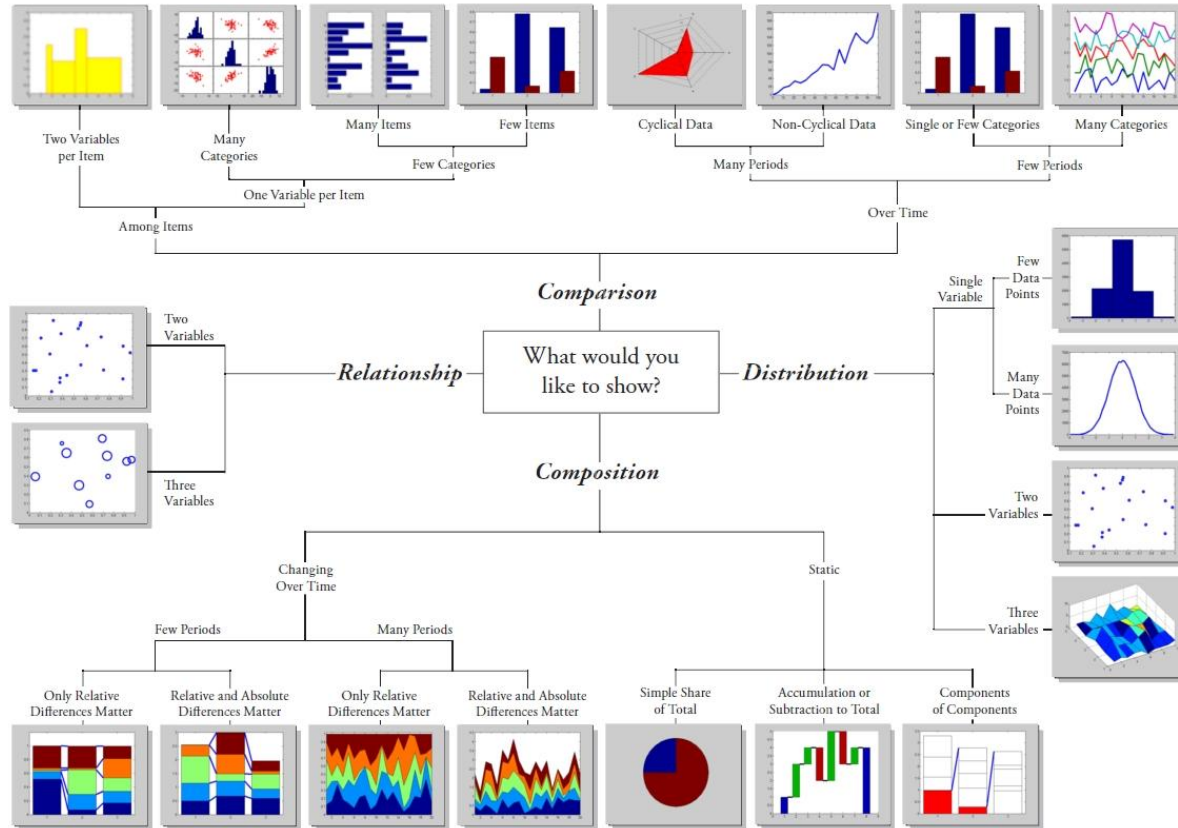
Scostamenti / incertezza

Diverging bar chart · Bullet chart · Error bars · Confidence band

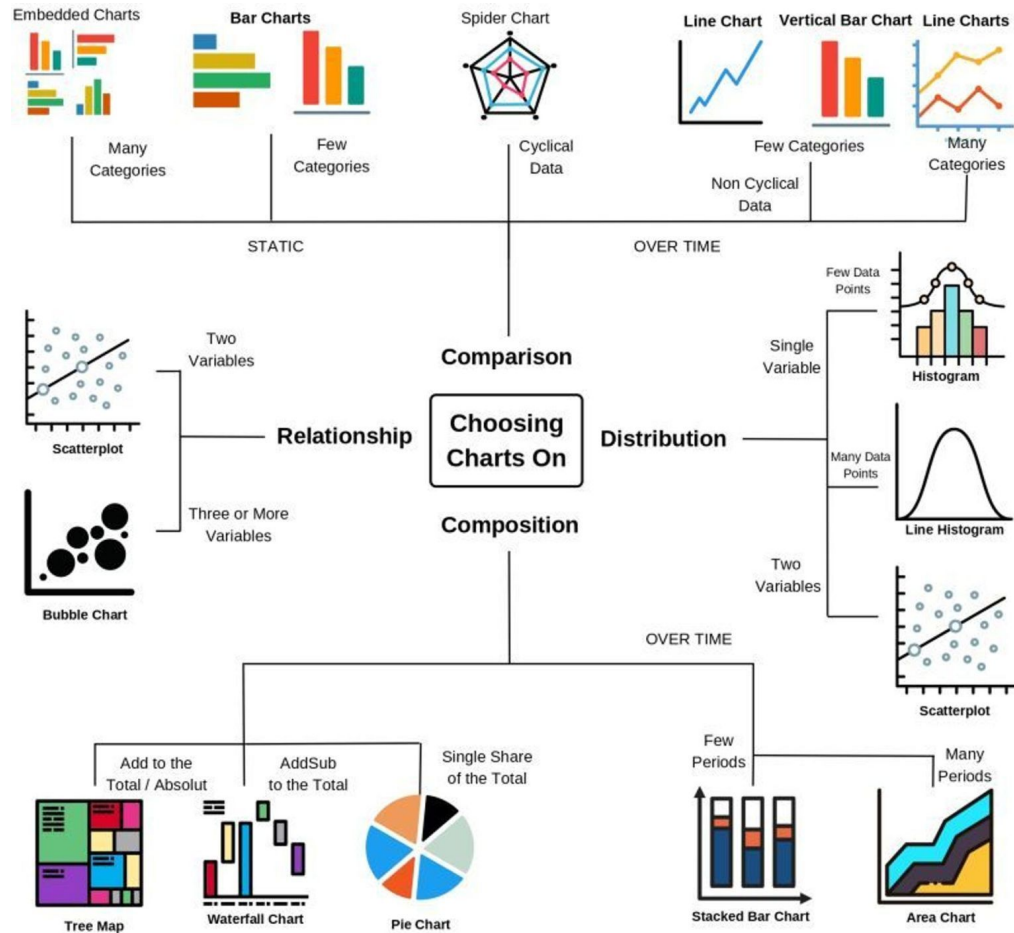
Flussi e movimenti

Sankey diagram · Alluvial diagram · Flow map · Chord diagram

Chart Suggestions—A Thought-Starter



Andrew Abela

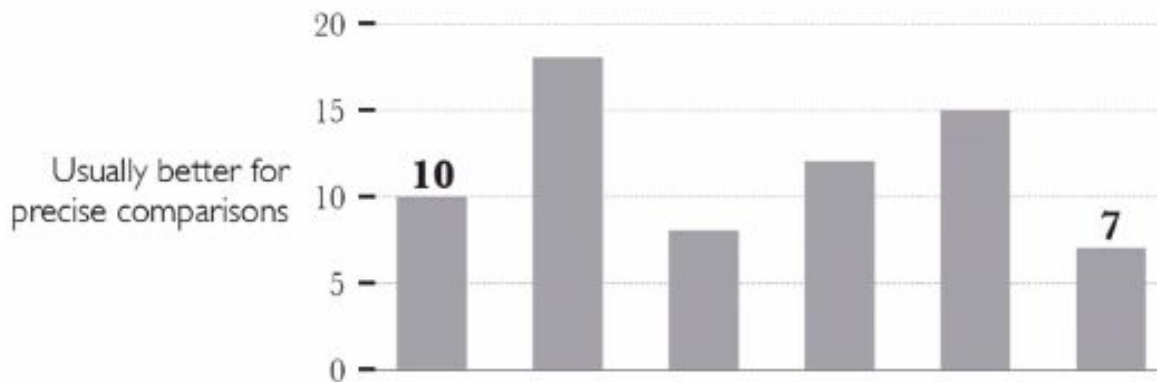


By Johnny Shollaj

Add me on:



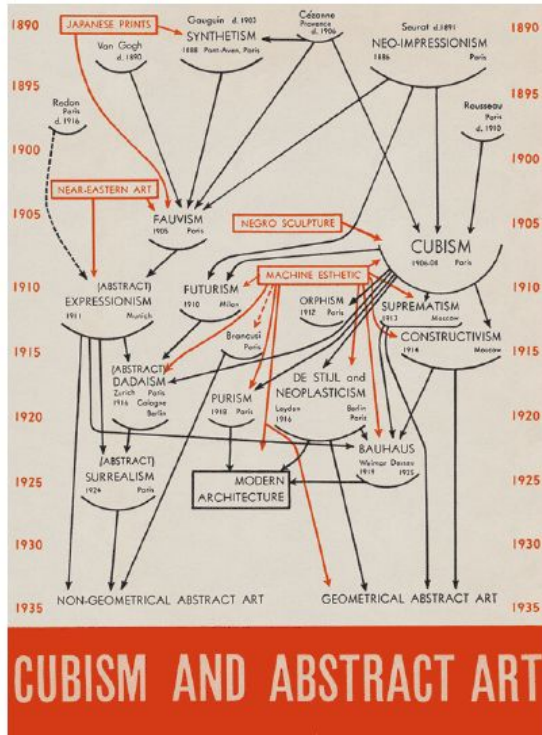
Comunicare confronti - altezze, dimensioni, colori



It may work for overviews of large datasets (choropleth maps, for instance)



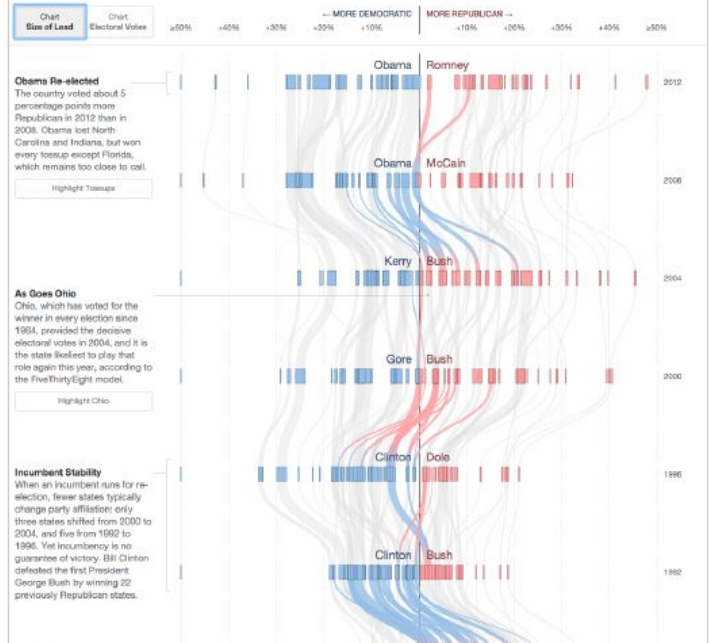
COMMUNICATE: TEMPORAL STRUCTURES



Cubism And Abstract Art (Alfred H. Barr 1936)

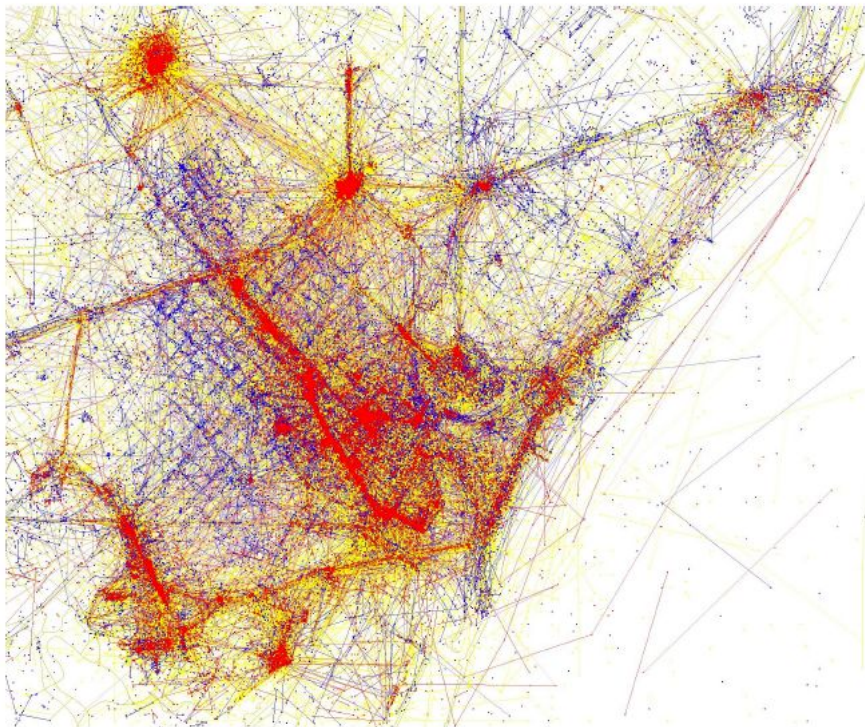
Over the Decades, How States Have Shifted

Recent elections have placed a heavy emphasis on "swing states" — Ohio, Florida and the other competitive states. Yet in the past, many more states shifted between the Democratic and Republican parties. A look at how the states stacked up in the 2012 election and how they have shifted over past elections.

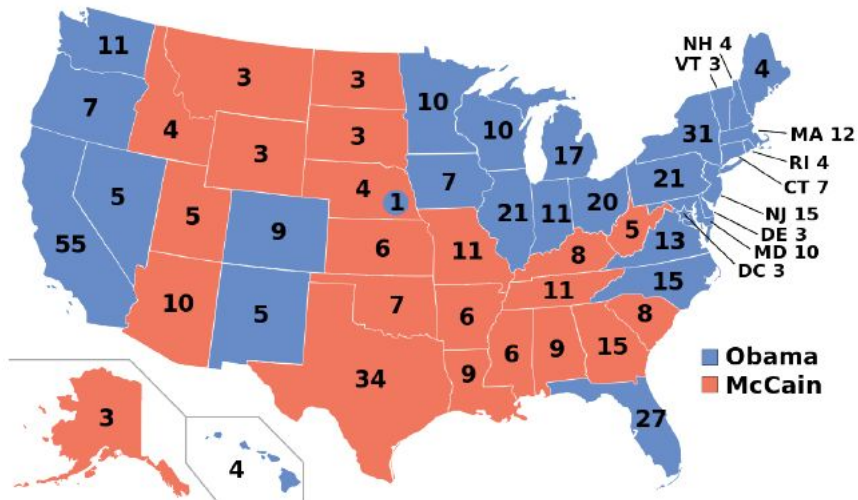


<http://www.nytimes.com/interactive/2012/10/15/us/politics/swing-history.html>

COMMUNICATE: MAPS



<https://www.flickr.com/photos/walkingsf/sets/72157624209158632/>



"ElectoralCollege2008" by Gage - Own work. Licensed under Public Domain via Wikimedia Commons - <http://commons.wikimedia.org/wiki/File:ElectoralCollege2008.svg#mediaviewer/File:ElectoralCollege2008.svg>

Risorse online per scegliere il grafico giusto

[Data to Viz](#)

Una guida decisionale che parte dal tipo di dati e porta verso i grafici più adatti; segnala anche i principali errori da evitare

[Datawrapper](#) – A friendly guide to choosing a chart type

Una panoramica molto chiara dei principali grafici e di quando usarli; utile soprattutto per chi inizia.

[FT Visual Vocabulary](#)

Organizza i grafici per funzione comunicativa ed è nato come strumento per migliorare la chart literacy in redazione.

[UNHCR Dataviz Platform](#)

Permette di esplorare i tipi di grafico principali e offre per ciascuno consigli pratici, con “dos and donts”.

[The Data Visualisation Catalogue](#)

Un repertorio molto utile dei principali tipi di grafico, con descrizioni ed esempi

Altre risorse

informationisbeautiful.net

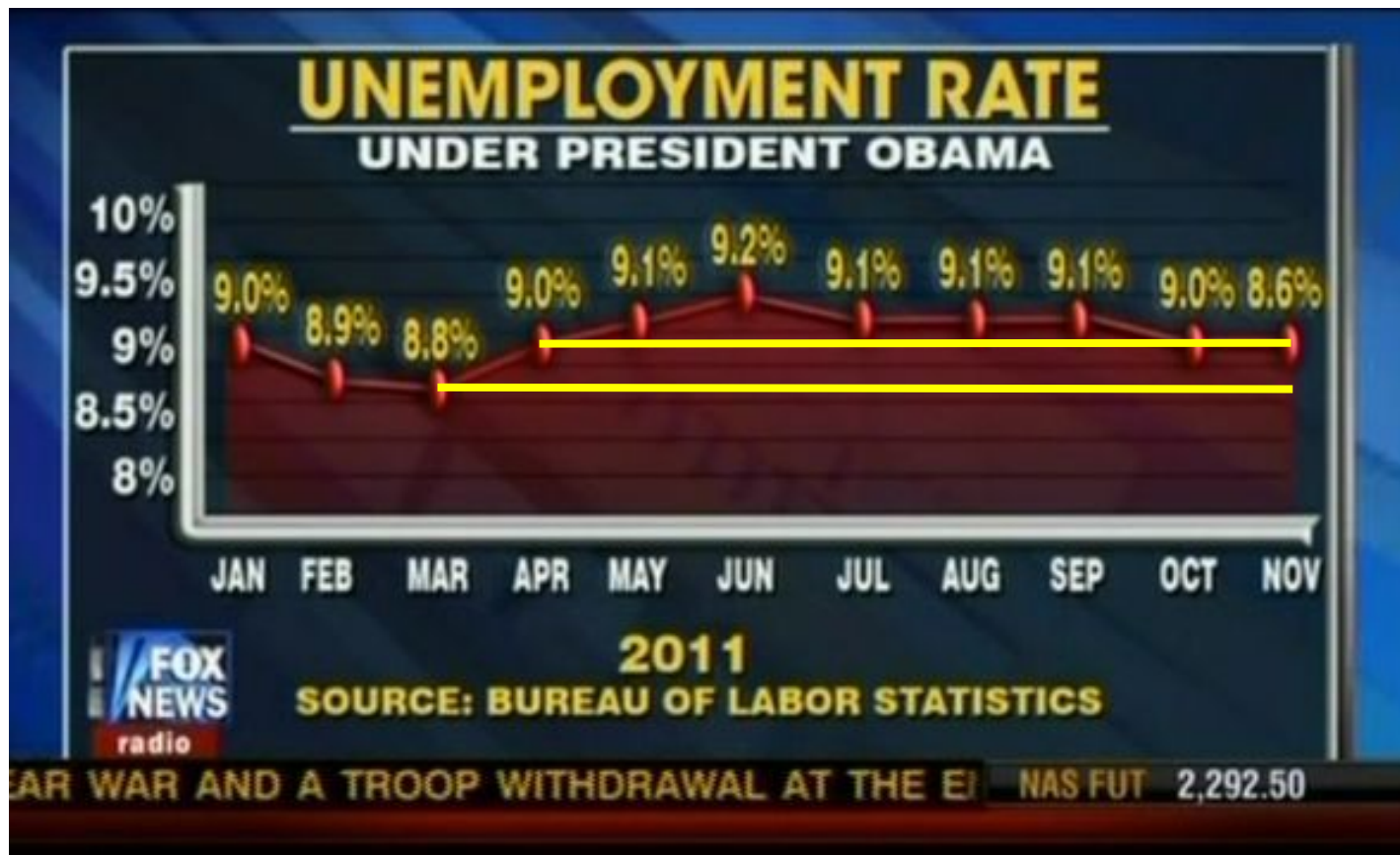


flowingdata.com



Errori di encoding

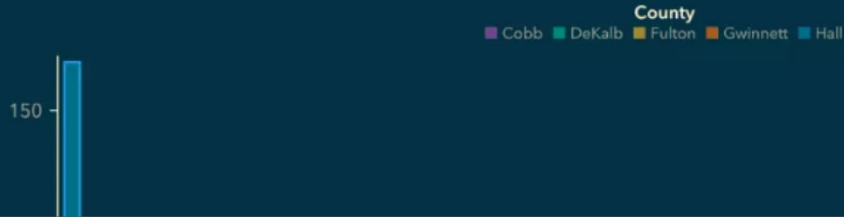
Fox News – grafico sulla disoccupazione del 2011



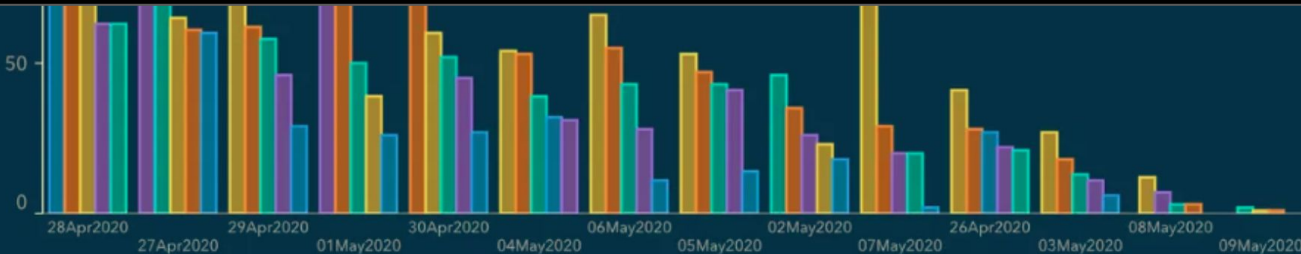
Georgia Department of Public Health - 2020

Top 5 Counties with the Greatest Number of Confirmed COVID-19 Cases

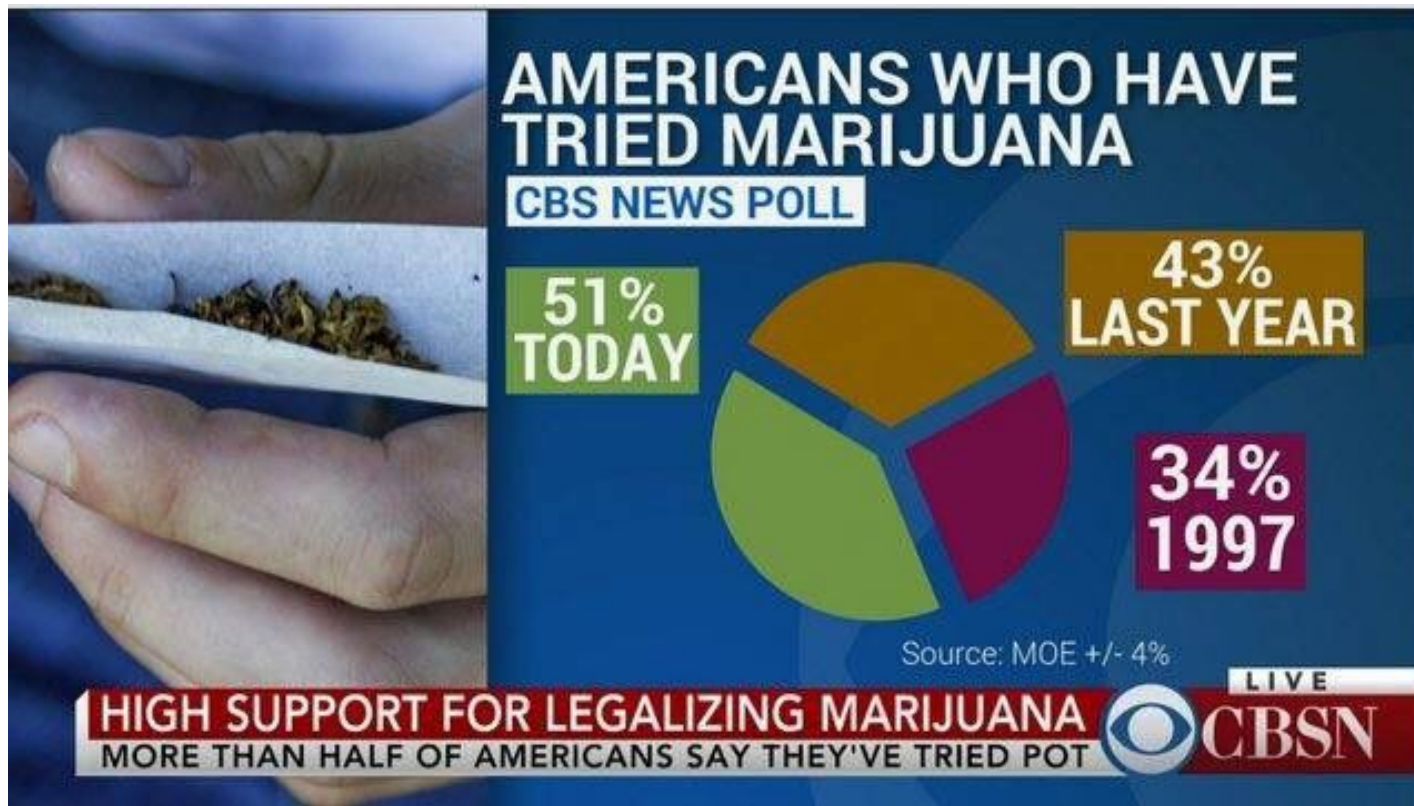
The chart below represents the most impacted counties over the past 15 days and the number of cases over time. The table below also represents the number of deaths and hospitalizations in each of those impacted counties.



28-Apr, 27-Apr, 29-Apr, 1-May, 30-Apr, 04-May, 06-May, 02-May, 07-May, 26-Apr, 03-May, 08-May, 09-May



CBS News - 2017



Top Donating Countries - 2009



BANKING ON THE WORLD BANK

Developing countries around the world depend on assistance from the World Bank to help support their often fragile economies. We have highlighted the regional recipients and respective outlays, as well as the nine largest donor countries contributing to the Fund's efforts to fight poverty around the globe.

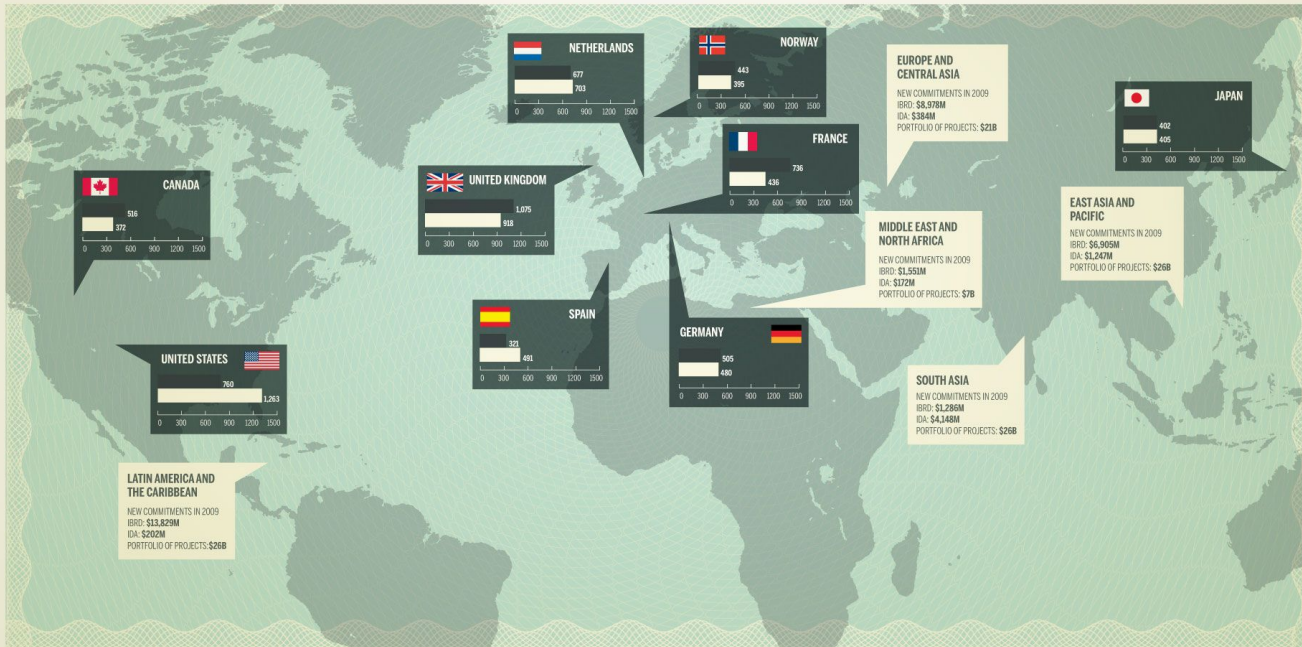
IBRD: THE INTERNATIONAL BANK FOR RECONSTRUCTION AND DEVELOPMENT
IDA: THE INTERNATIONAL DEVELOPMENT ASSOCIATION

TOP DONATING COUNTRY

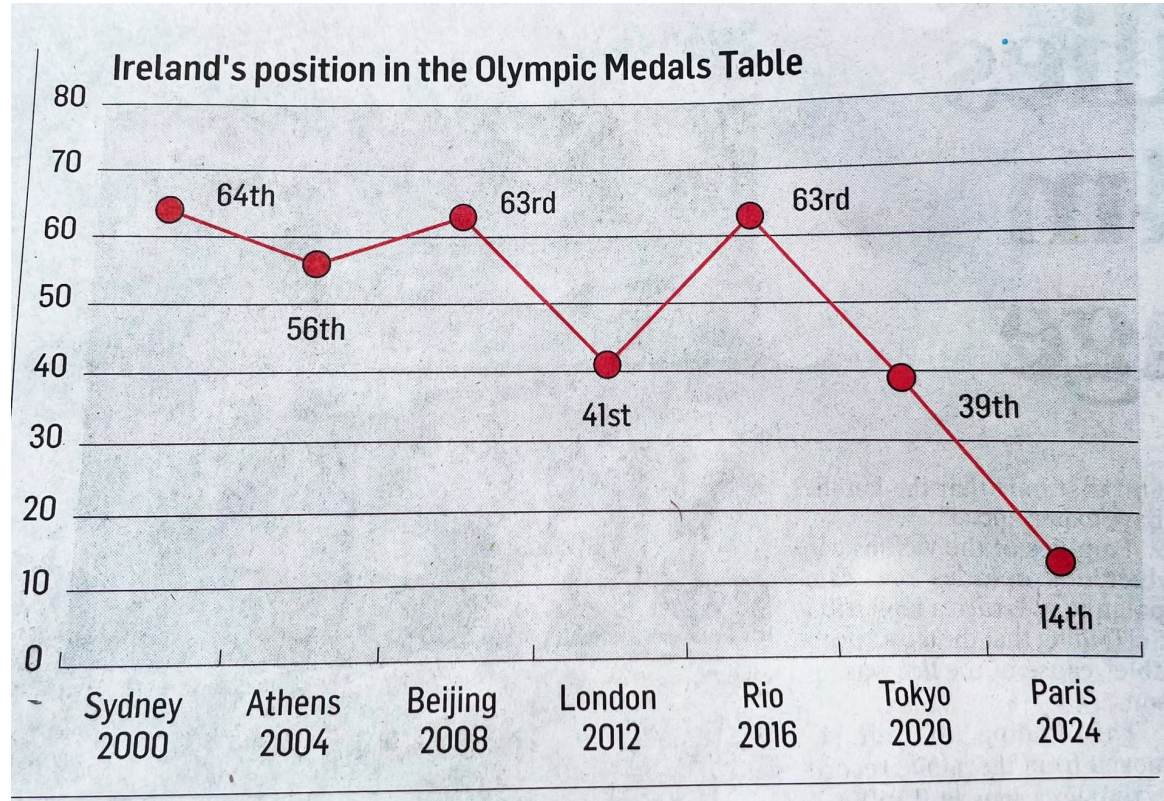


AID RECEIVED BY REGION

SOURCE: WORLD BANK.ORG



The Irish Times - 2024



Esercizio - Olympic Games

1. Estrai i numeri da [All-time Olympic Games medal table - Wikipedia](#)
2. Pensa alla domanda che vuoi rispondere
3. Riprogetta il grafico
4. Fai più di una visualizzazione