Data Journalism

Introduction

InfoUma 2022 Andrea Marchetti

Overview

- Communication

- Newsroom composition and workflow

- How to transform numbers into stories

Communication

L'ultimo miglio - La Comunicazione

L'ultimo miglio del percorso di uno scienziato come quello di un giornalista investigativo è anche il passaggio decisivo, quello che rende una ricerca famosa o meno ed è
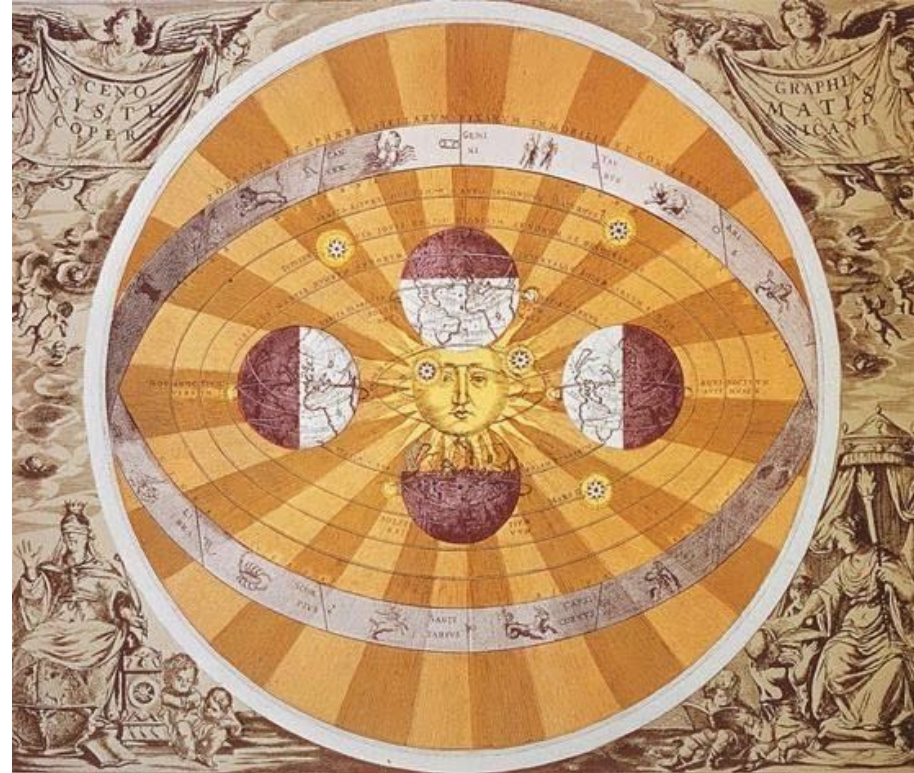**La Comunicazione dei Risultati**

Da come comunichi dipende il risultato

Copernico comunicò i suoi risultati con il libro *De Revolutionibus Orbium Caelestium* che solo gli scienziati potevano comprendere.

Galileo scrisse invece una commedia. *Il dialogo sui due massimi sistemi.* E lo capirono tutti molto bene !!!

Marco Paolini - ITIS Galileo [25 aprile 2012]

Scopo della comunicazione

La comunicazione efficace è quella che consente di cambiare il modo di pensare delle persone

La parabola del Buon Samaritano.

Dal vangelo secondo Luca

Chi è il mio prossimo?

...

Chi di questi tre ti sembra sia stato il prossimo di colui che è incappato nei briganti?

Qual'è la storia che vale la pena di raccontare?

Deve essere sicuramente qualcosa di semplice.

Se un grafico è troppo complicato devo **filtrare** i dati (meglio se lo faccio fare all'utente), eliminare le informazioni superflue magari a discapito della completezza.

L'intento dello scrittore è quello di comunicare qualcosa di utile che colpisca nel profondo le **emozioni** del lettore. L'importanza di conoscere l'audience

"*La vita non è quella che si è vissuta, ma quella che si ricorda e come la si ricorda per raccontarla*" Gabriel Garcia Marquez Vivere per raccontarla
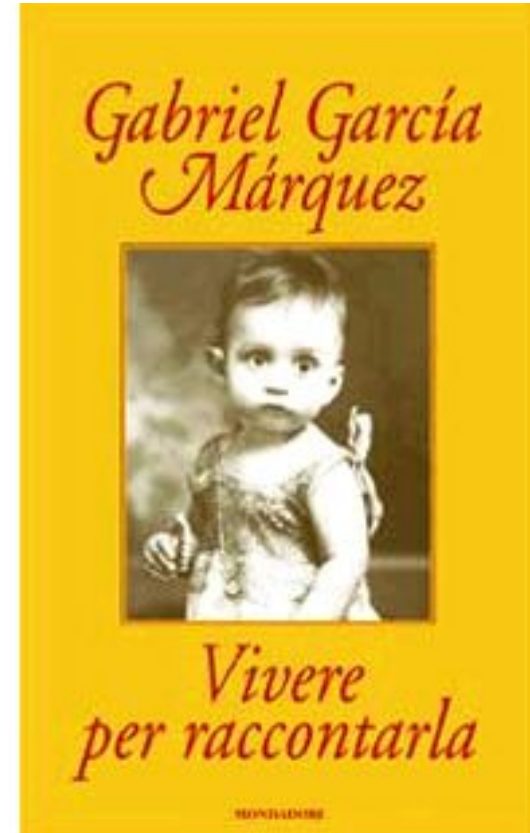
Proposta

Usare una storia per comunicare i risultati

Nessuno resiste ad una storia
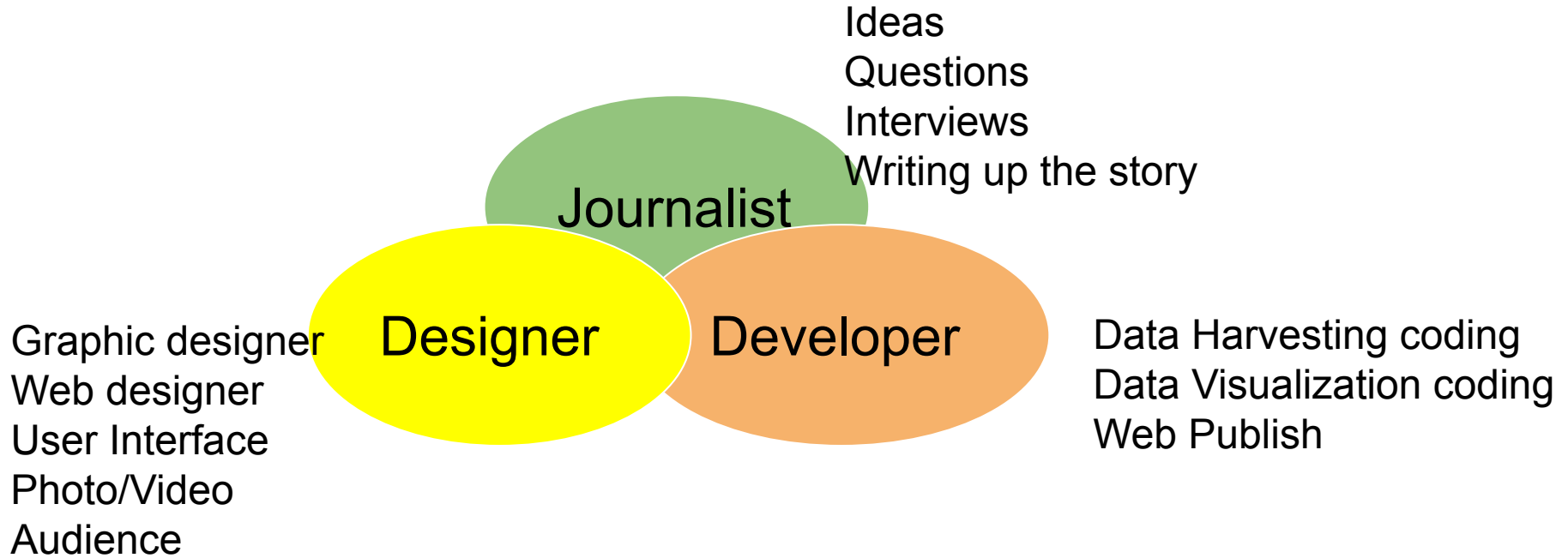
Data Journalism come vettore dei nostri risultati
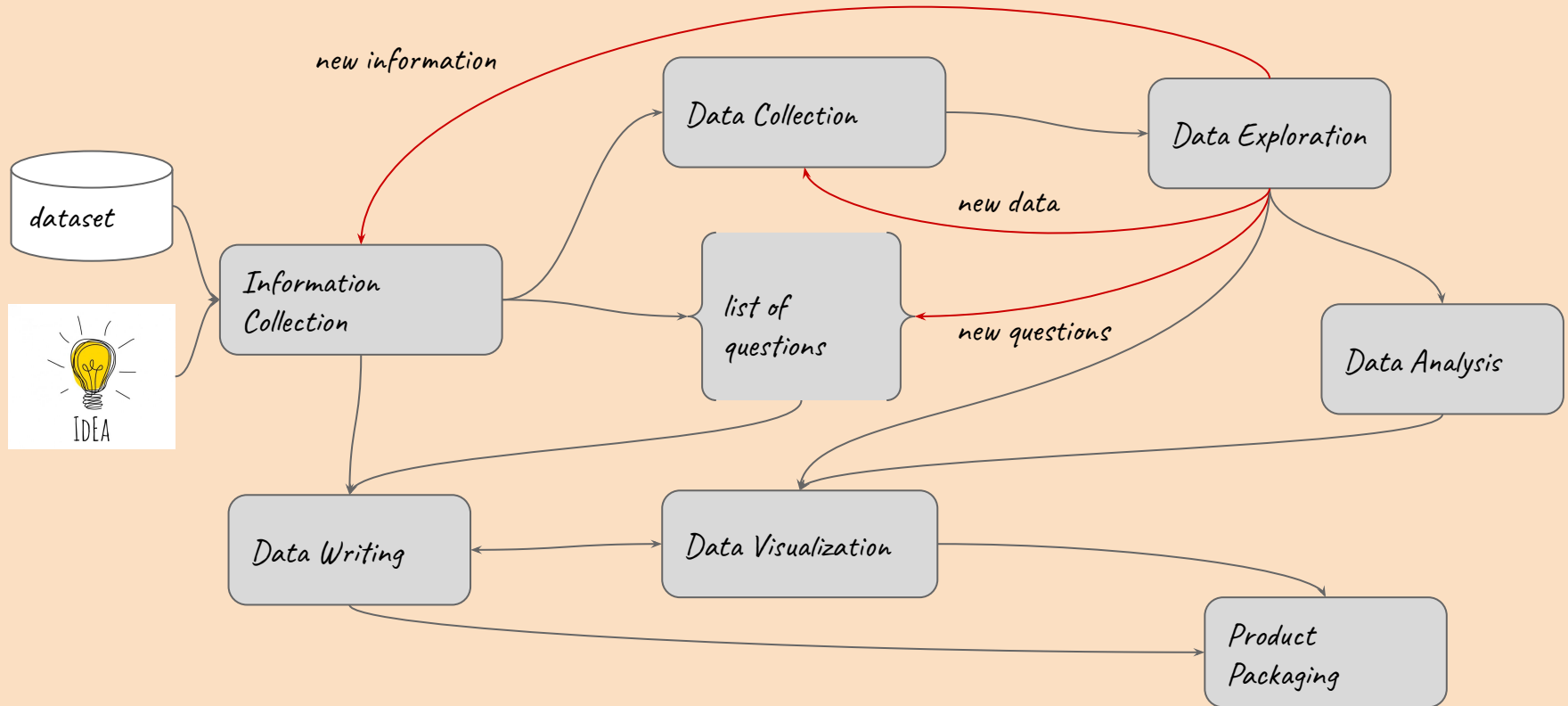
Newsroom composition and workflow

inside a data team

Small DJ Newsroom

Ideas
Questions
Interviews
Writing up the story

Journalist

Graphic designer
Web designer
User Interface
Photo/Video
Audience

Designer

Developer

Data Harvesting coding
Data Visualization coding
Web Publish

Data Journalism Workflow

The initial two possible Scenarios

Data journalism begins in one of two ways:

1. ***a question that needs data,*** (journalistic)

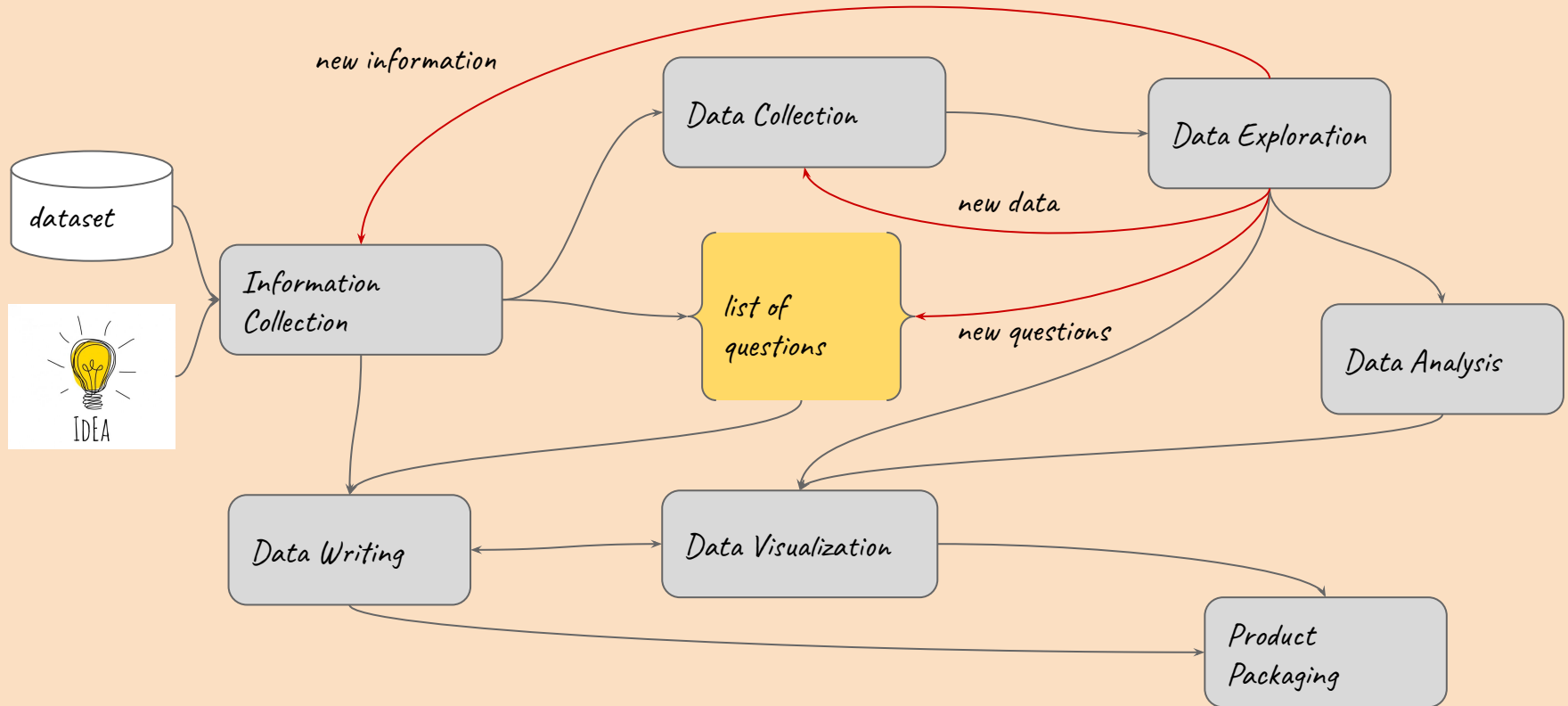2. ***an interesting dataset that needs questioning*** *(data scientist)*

An interesting dataset that needs a story

List of questions

List of questions

Define which questions the story should answer

- general questions
- specific questions

Each question should be **short** and **clear**.
The number of questions should be **limited**.
New questions could arise during the investigation.
The final list will be the chapter of our story

List of questions - Women's soccer

1. What is the situation in Italy with respect to women's soccer? In Europe and in the world?
2. How widespread is women's soccer in Italy?
3. Are there still discrimination against women's soccer?
4. How does the soccer government behave?
5. What are the difficulties a female soccer player encounters?

Interviews

The role of interviews

- Understand the domain of knowledge

- Discover interesting stories to tell

- Verify the final job

- Include the interviews inside the story

Data StoryTelling

Tipologie di Storie

Narrare i cambiamenti nel tempo

Dalla panoramica al dettaglio e viceversa

Esaltare i contrasti

Esplorare le intersezioni

Sezionare i fattori

Profilare i valori anomali (outlier)

* Kang, Martha (2015): Exploring the 7 Different Types of Data Stories. Mediashift [http://mediashift.org/2015/06/exploring-the-7-different-types-of-data-stories/].

Narrare i cambiamenti nel tempo

I dati sono utilizzati per narrare i cambiamenti nel tempo ed evidenziare le forze in gioco

Numero di persone **uccise** negli anni (**1970-2017**) da attacchi terroristici
fonte GDT
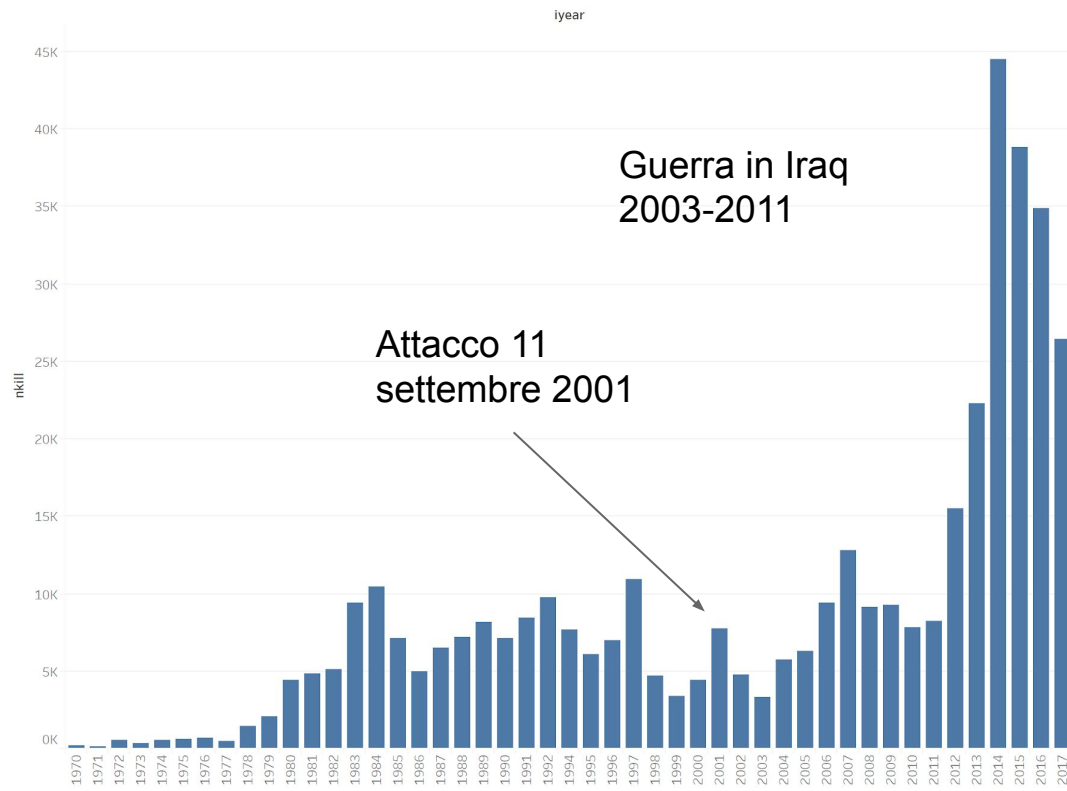
Overview Vs Detail

Dal grande al piccolo

i dati guidano il lettore da una situazione generale verso un caso particolare. Il lettore può muoversi dal generale al particolare e viceversa (ad esempio attraverso una mappa) e accedere ai dati per ogni area specifica. Si possono avere anche dei filtri interattivi per l'accesso ad alcuni dati

Dal piccolo al grande

il lettore inizialmente si concentra su un particolare dell'intera informazione disponibile (ad esempio su di un paese specifico oppure su di una particolare città) e dopo può espandere la sua vista ad una prospettiva più generale

Shneiderman's Mantra

How to understand and visualise large quantities of data?
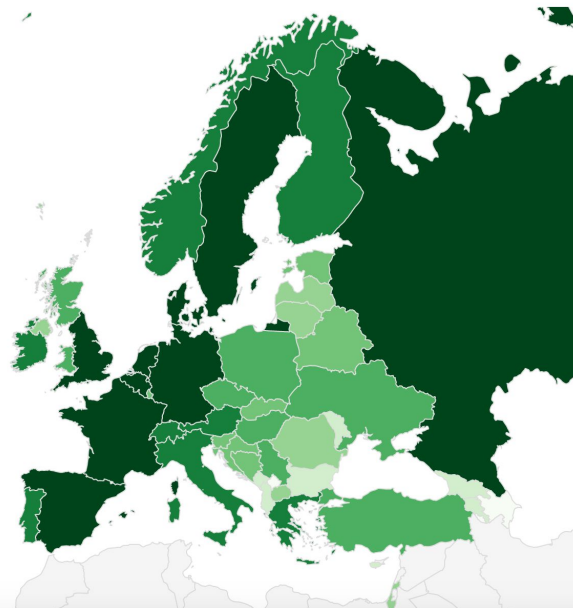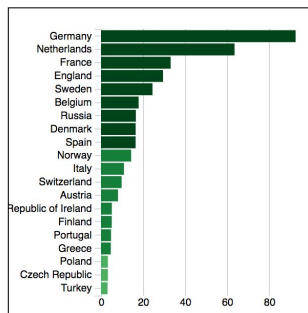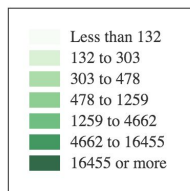
"Overview first, zoom and filter, then
-on-demand"
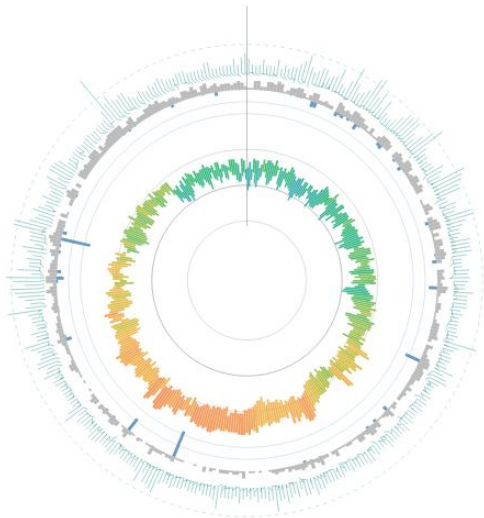
The Weather
Wheel
by WebVisLab

Evidenziare i contrasti

L'obiettivo consiste nel sottolineare le **differenze** tra i dati

In Italia non abbiamo una grande percentuale di donne che praticano il calcio ma questo non è vero nel resto di europa
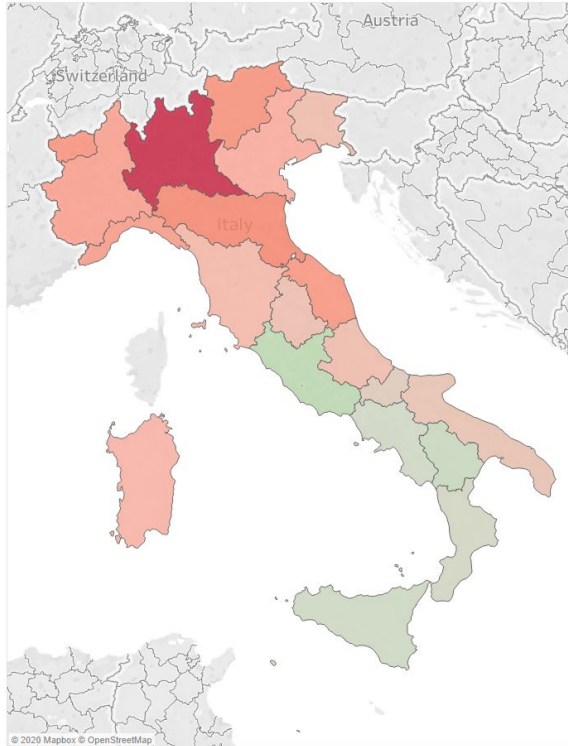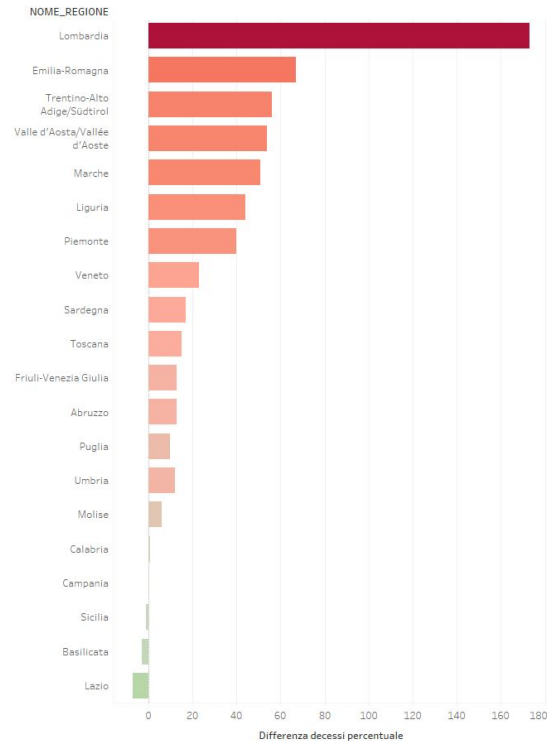
Differenza nel numero dei decessi primavera 2020

Esplorare le intersezioni

L'obiettivo consiste nel tentativo di spiegare il caso di due linee divergenti di dati che si intersecano e una sorpassa l'altra

Sezionare i fattori

Il data journalist cerca di scoprire la relazione tra i diversi fattori di una storia

Profilare i valori anomali

La storia si concentra sui valori anomali e cerca di spiegarne le cause. Di solito la ricerca di tali valori anomali comporta l'esplorazione dei dati

Language communities of Twitter – Eric Fischer – FlowingData

October 27, 2011

Language comunities on twitter

When looking at a visualisation, try to read it and extract meaning from it.

- Most countries show that the main language used in their country is their own.

- It shows expected information, which is not insightful.

Rather than the general trends, the exceptions are more useful.
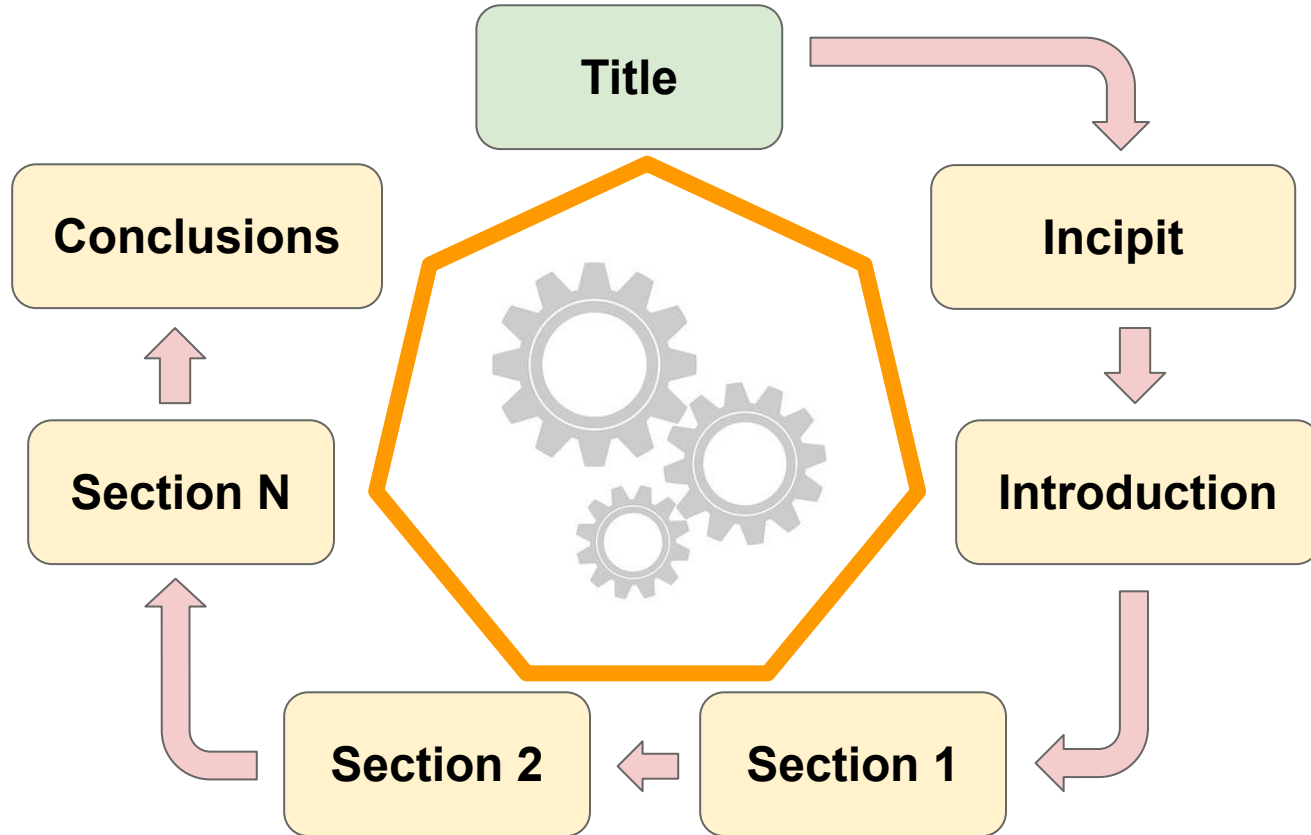
How to build a story

StoryBoard

- The StoryBoard helps to write the story
- The List of questions will be our storyboard
- Try to guess the visualizations useful for the story
- The story it's the centre
- Enter real stories not just **facts**

Story Board

Story Board - Women's soccer

L'altra metà del calcio

21 milioni di italiani interessati, oltre 7 milioni incollati agli schermi per seguire le gesta della nazionale di Milena Bartolini all'ultimo mondiale. In molti hanno imparato a riconoscere e apprezzare le giocate del capitano Sara Gama o i gol dell'attaccante Barbara Bonansea.[..]

Conclusioni

Se l'interesse cresce

Introduzione

Un gioco da dilettanti

In minoranza

A great visualization

The 4 features that define a great visualisation

1. **Functional** The shape of the graphic is adapted to the questions the visualisation should help answer.
2. **Beautiful/Attractive/Aesthetic level** If it is not attractive, readers won't stop to read and interact with it. Attention-Grabbing is 50% of the work.
3. **Insightful** Revealing hidden facts that are unexpected or surprising or extremely important.
4. **Enlightening** The information that the visualisation reveals changes the perception of the reader.

Try to objectively guess the best graphic form

1. **Think about the AUDIENCE and the publication that you are working for**
    a. teenagers or serious people
2. **Think of the questions your graphic should help answer**
    a. ~~I love maps so I will use maps forever~~
3. **Can you understand the graphic without reading every single number?** *The goal of visualising data is to provide rapid access to trends and patterns in the data without having to read the data.*

Sketch out your idea
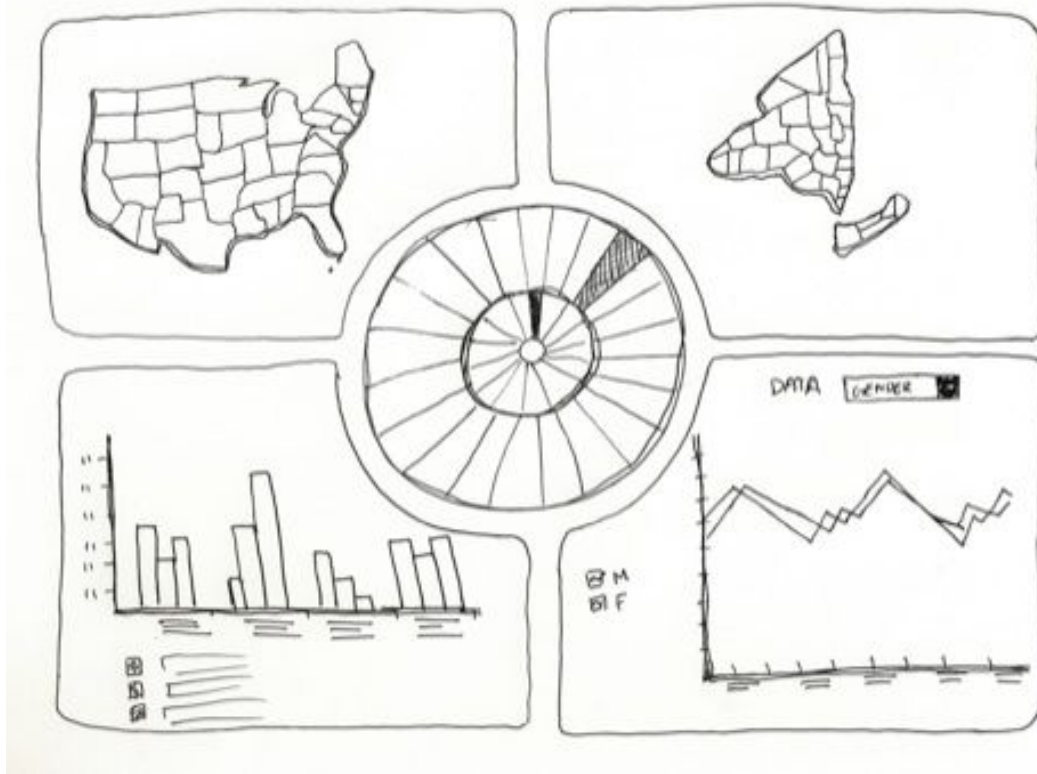
Sketch out your idea

Making visualisations depends on thinking visually, not on a computer, so:

- Start with a **sketch of an idea** for visualising a story (don't start with a tool such as Tableau, Excel, Google Fusion, …)
- If the visualisation does not work on paper, you are easily able to start a new one.

Something of useful

- Creating something attractive will attract readers.

- **After attracting readers, you also need to deliver something useful**

- Give a context to understand

Analyse graphic published

After making the graphic attractive, you need to make it functional

- What are you supposed to do with the graphic?
- What questions are supposed to be answered?

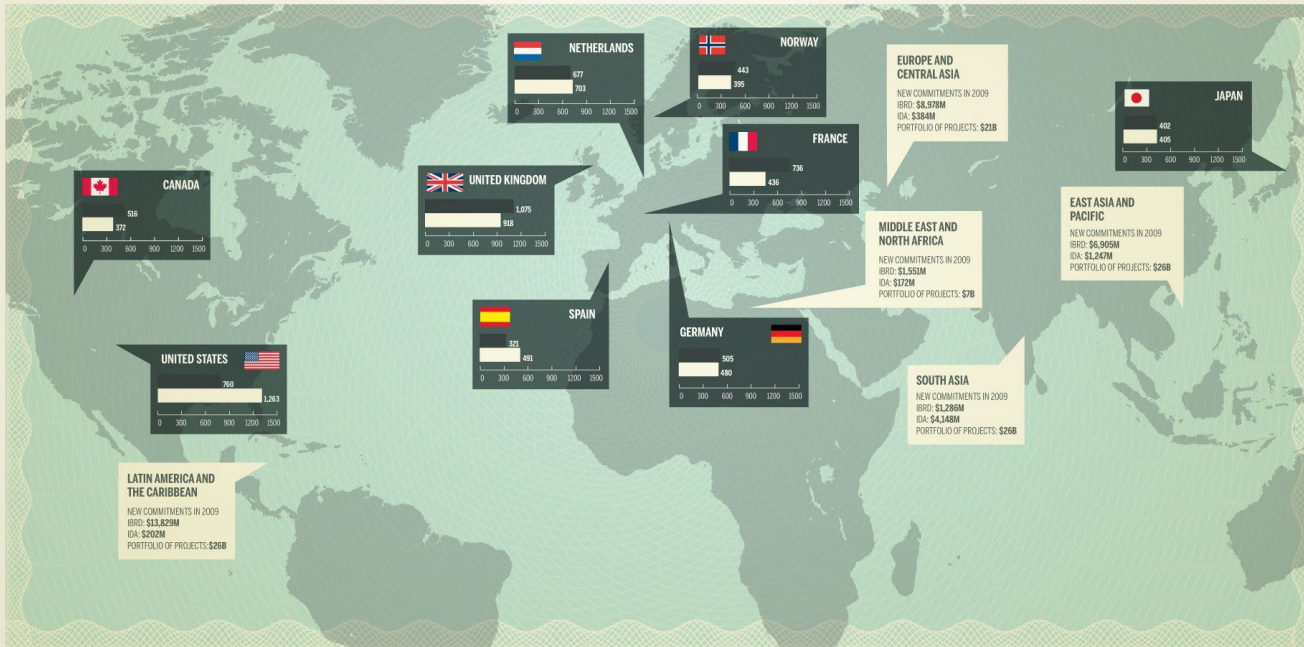Example - Banking on the World Bank – extract from GOOD Magazine - **The Donating countries**

Top Donating Countries

Comments

User question: **What is the ranking of countries by aid given? I want to see who gave more money?**

If you want to know the answer you need to memorize the numbers and put in a rank.

Other question: **Who has been the most generous?**

- This overview is incomplete, a variable is missing to make the comparison between the countries complete.
- You need to provide the amount given per capita per country.

The eventual goal of the graphic seems to be to locate the different countries. Is it important?

Rule

A designer needs to anticipate the readers' questions and facilitate what their needs are.

Tips

- Declare the **source** of data
- Numbers without context are just numbers, **words give context**
- **Explain the charts**, wrap a story around the numbers, ensure the numbers are clearly understood

The source

- Where did the data come from? **The source**

- **Transparency** about the source is critical.

  - How do you know if the information is correct? We need to use data from reliable sources. We need to declare which modifications have been carried out in order to have reproducible results

- Do not blindly trust the accuracy of supplied data

Give the context

It's really frustrating don't understand
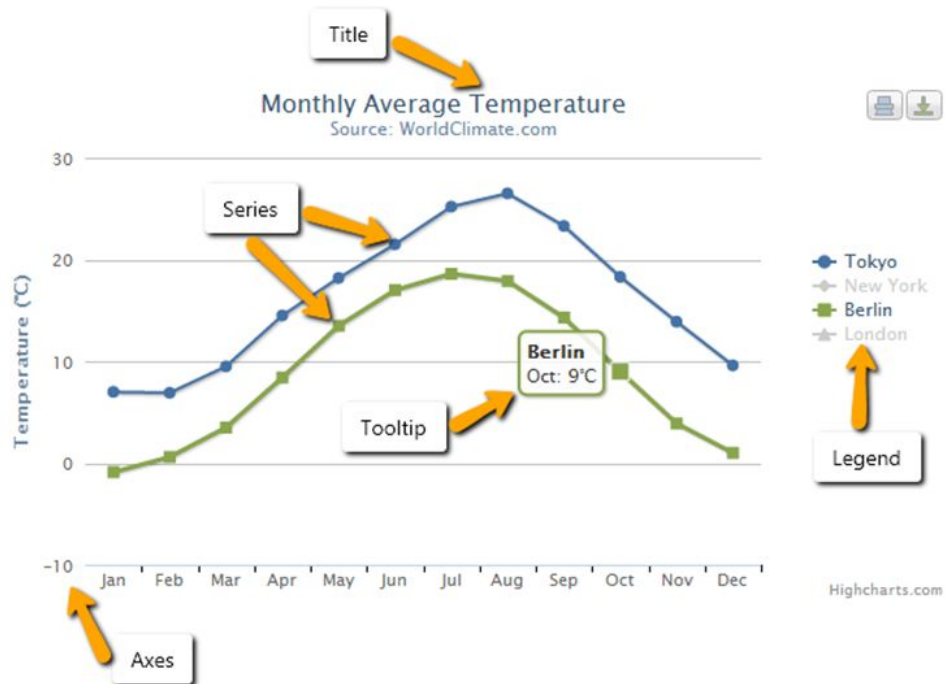
Summary of main principles of DV

DV creation doesn't require software

- Think visually
- Sketch first

DV can't be only attractive

- make readable with the context
- give overview then details

Provides insight into the data

- It's not just transforming data into visual forms
- The visual form has to reveal hidden facts

Data Journalists Bibliography

Web Sites

DataJournalism.com

International Journalists' Network

Google News Initiative, Google News Lab

Journalism Courses

Microsoft Data Journalism

Books

2012 Data Journalism Handbook - various authors

2013 The Functional Art - An Introduction to Information Graphics and Visualization - Alberto Cairo

2015 Storytelling with Data - Cole Nussbaumer

2016 The Truthful Art Data Charts and Maps for Communication - Alberto Cairo
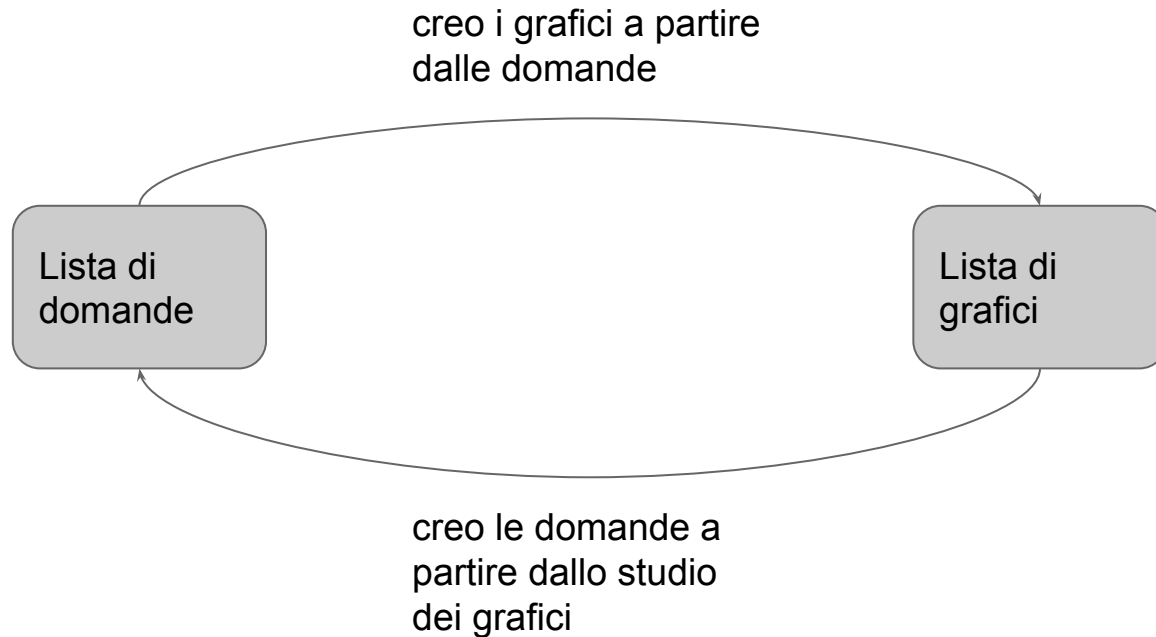
2018 Data Journalism Handbook 2

La Sprezzatura

… e, per dir forse una nova parola, usar in ogni cosa una certa sprezzatura, che nasconda l'arte e dimostri ciò che si fa e dice venir fatto senza fatica e quasi senza pensarvi…

Castiglione Baldassarre 1528

Comunicare il nostro **lavoro** tramite una **storia**

creo i grafici a partire
dalle domande

Lista di
domande

Lista di
grafici

creo le domande a
partire dallo studio
dei grafici

what data journalism is

In times of uncertainty, let the facts be your guide.

Subscribe for $1 a week.

What Data Journalism is

- **Telling stories** using numbers
- Using **visual technologies** to aid communication of the numbers
- **Technologies** are not important, the focus is on the **story**
- Journalism is **reporting facts** in a way that **people can understand**

Why Data Journalism is so popular?

- Many governments around the world have published thousands of **open data.**

  - on May 2013 Obama started the Open Data Policy: Transparency + Participation + Collaboration
    - **data.gov**
    - **data.gov.uk**
    - **dati.gov.it**

- Many **tools** for data cleansing, data exploration, data analisys, data visualization are available

Open Datasets

User generated data: [Wikipedia](#), [DBPedia](#), [Wikidata](#)

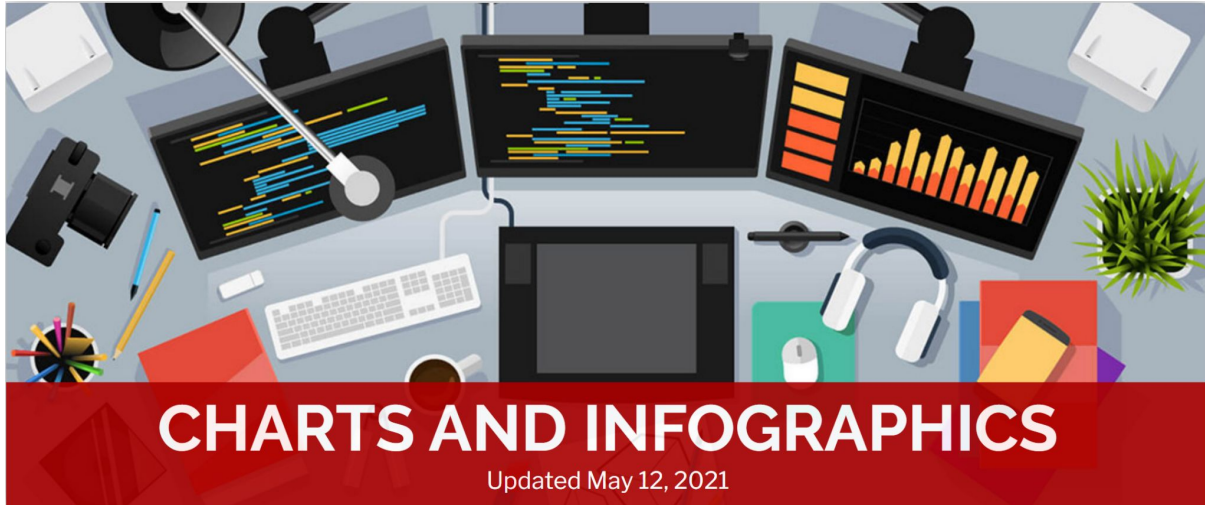Open Government Data: [data.gov](#), [data.gov.uk](#), [data.gov.it](#), …

Statistics Institutes: [istat](#), [eurostat](#), [worldbank](#), [oecd](#), …

Open Data aggregators: [awesome dataset](#), [cooldataset](#), …

Open Data search engines: [Google Data Search](#), [Google trends](#)

DJ Tools - [Journalist Toolbook](#)

CHARTS AND INFOGRAPHICS
Updated May 12, 2021

JOURNALIST'S TOOLBOX

Created by **Mike Reilley** · Designed by **Billy O'Keefe**

The Sigma Awards 2021

[The Sigma Awards 2021 Projects Short List](#)

140 projects

[Download the ShortList](#)

International Journalism Festival

- The biggest annual media event in Europe
- Perugia, Italy
  - *from Wednesday 01 to Sunday 05 April 2020* CANCELED
  - *from Wednesday 14 to Sunday 18 April 2021* CANCELED
  - *from Wednesday 06 to Sunday 10 April 2022*
- All sessions are free entry, No registration
  - *Access to all sessions is on a first-come-first-served basis*
- The **Sigma Awards** premiation

#ijf22

INTERNATIONAL JOURNALISM FESTIVAL

PERUGIA, ITALY | **6-10 APRIL 2022**

XVI EDITION | FREE ENTRY