



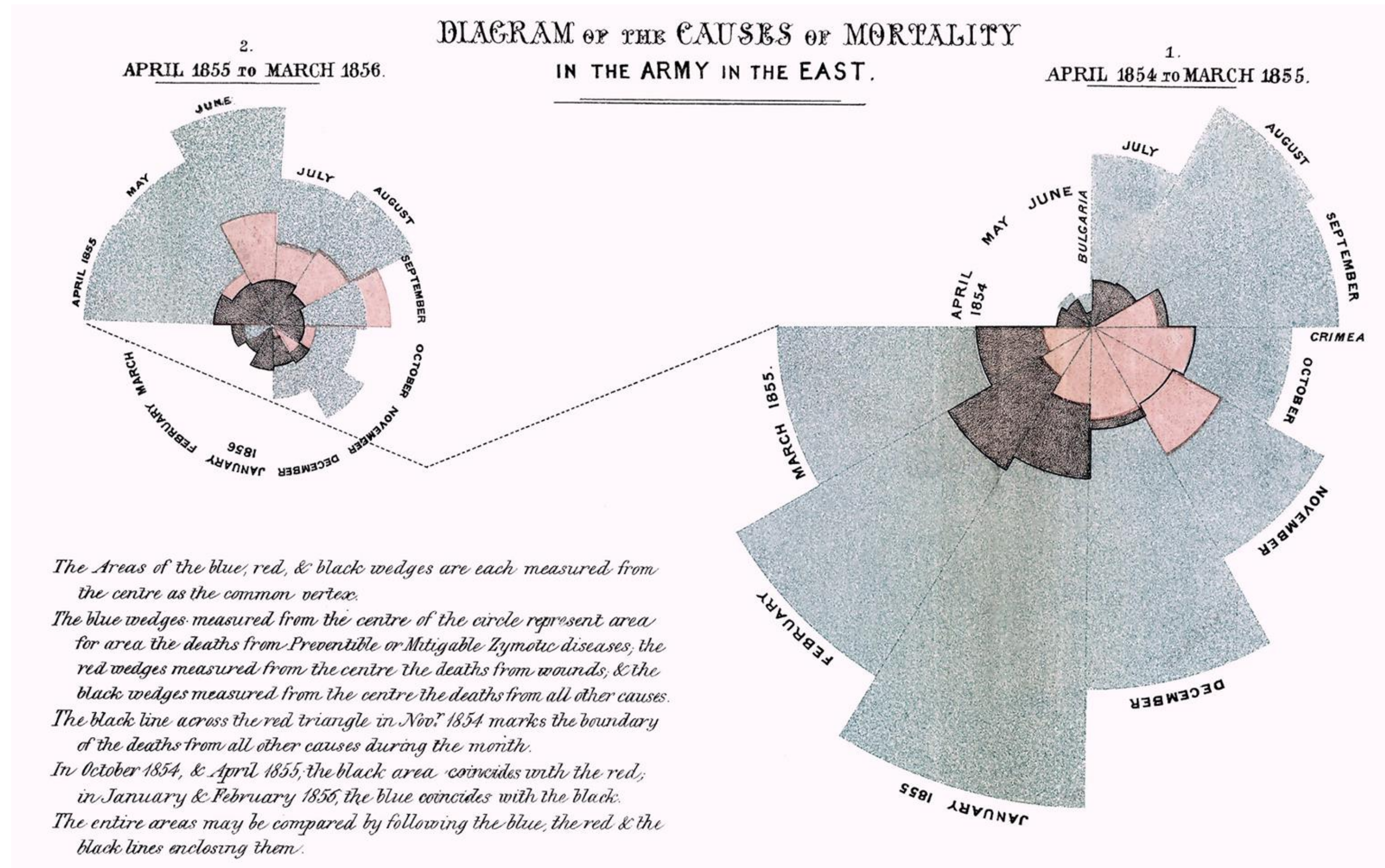
**WHAT IS DATA JOURNALISM**  
**(part one)**

# WHAT IS DATA JOURNALISM (part one)



## WHAT IS DATA JOURNALISM: INTRODUCTION

1858: Florence Nightingale writes a report on the condition of English soldiers in the Crimea, analysing the leading causes of death.







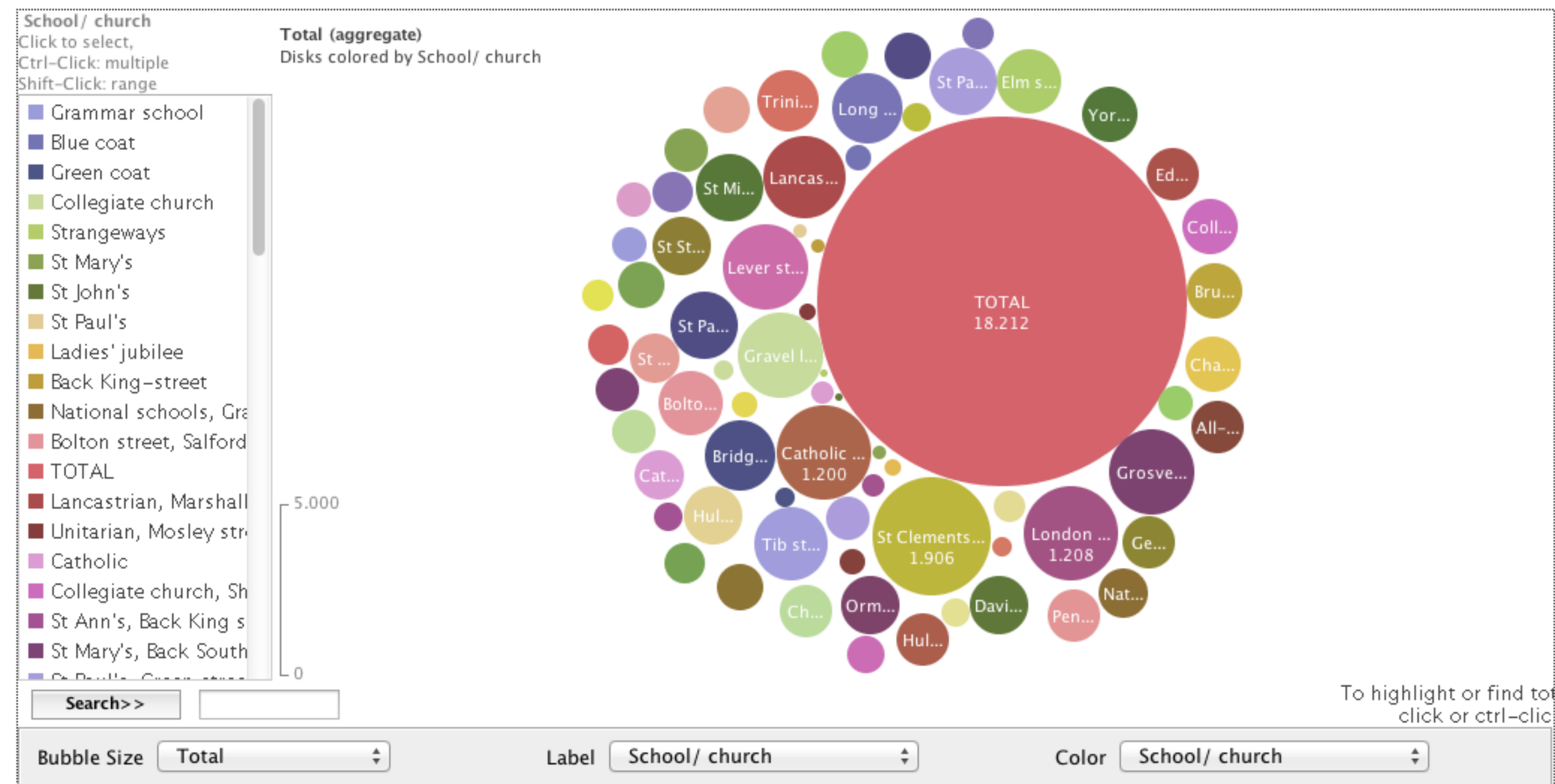


# WHAT IS DATA JOURNALISM: INTRODUCTION

## Visualizations : Schools in Manchester, 1821

Uploaded by: [smfrogers](#)  
Description:

Created at: Sep 25 2011



2000s: Simon Rogers  
(Guardian Datablog)  
reproduces the table  
published in 1821 using  
simple Data Visualisation  
tools

## WHAT IS DATA JOURNALISM: INTRODUCTION

Data Journalism is not new: you could say that it has **existed as long as data has.**

The difference is that in the past the data were published in very costly books, now computers with spreadsheets and specific tools do half the work by themselves.

Today, Data Journalism is a **form of journalism** that:

- Collects, cleans and analyses **data**
- Uses **new technologies**
- Translates data into **information, visualisations, analyses...** in a few words: **TELLING STORIES (in the most understandable way possible)**

## Starting point...



**Data** are the raw material for a data journalist. As in traditional journalism, it's also important to start out with a good **research question**, identify the **story** that you want to tell and understand the **final objective** of the project.



## Think like a journalist...

The data journalist, usually, thinks about data as a journalist would and not as an analyst.

Why are some numbers interesting?

What's in there that's new? What might happen if it's crossed with other data?

Answering these questions is the most important thing.



**WHICH DATA?**

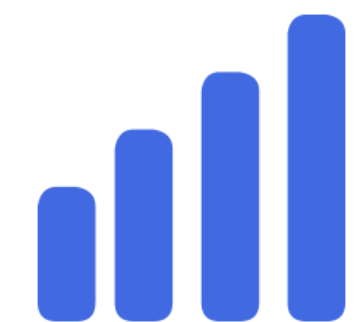


## Data useful for our research:

- Associated with **economic or social** issues
- Taken from **reliable** sources
- **Reusable** (published by sources with licences for reuse)
- **Can be aggregated**, when we can build them ourselves by putting multiple sources together

## The tools you need

- Spreadsheet (Excel or Libre Office Calc) which allows you to **analyse** small or large scale tables
- Software and applications that allow you to **visualise** the data we have aggregated using maps or graphs





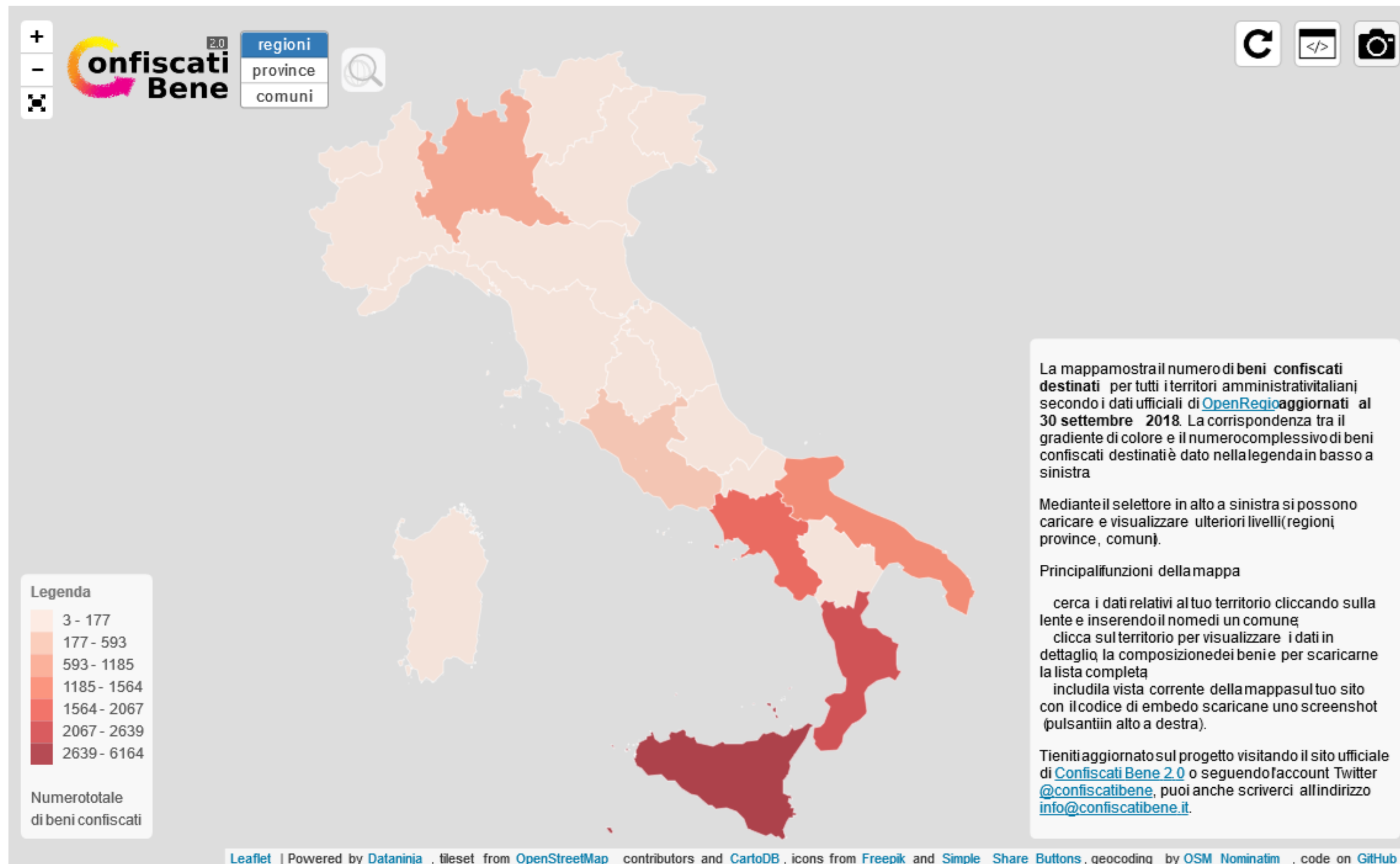
# SOME EXAMPLES

## Analyse a social phenomenon across territories

- The aim is to compare different geographical areas in terms of a phenomenon
- We require and need to make uniform data for all territories that we wish to compare
- They need to be compared starting from a common element (in relation to the population, for example)



# Example: investigation into confiscated Mafia assets



<https://www.confiscatibene.it>

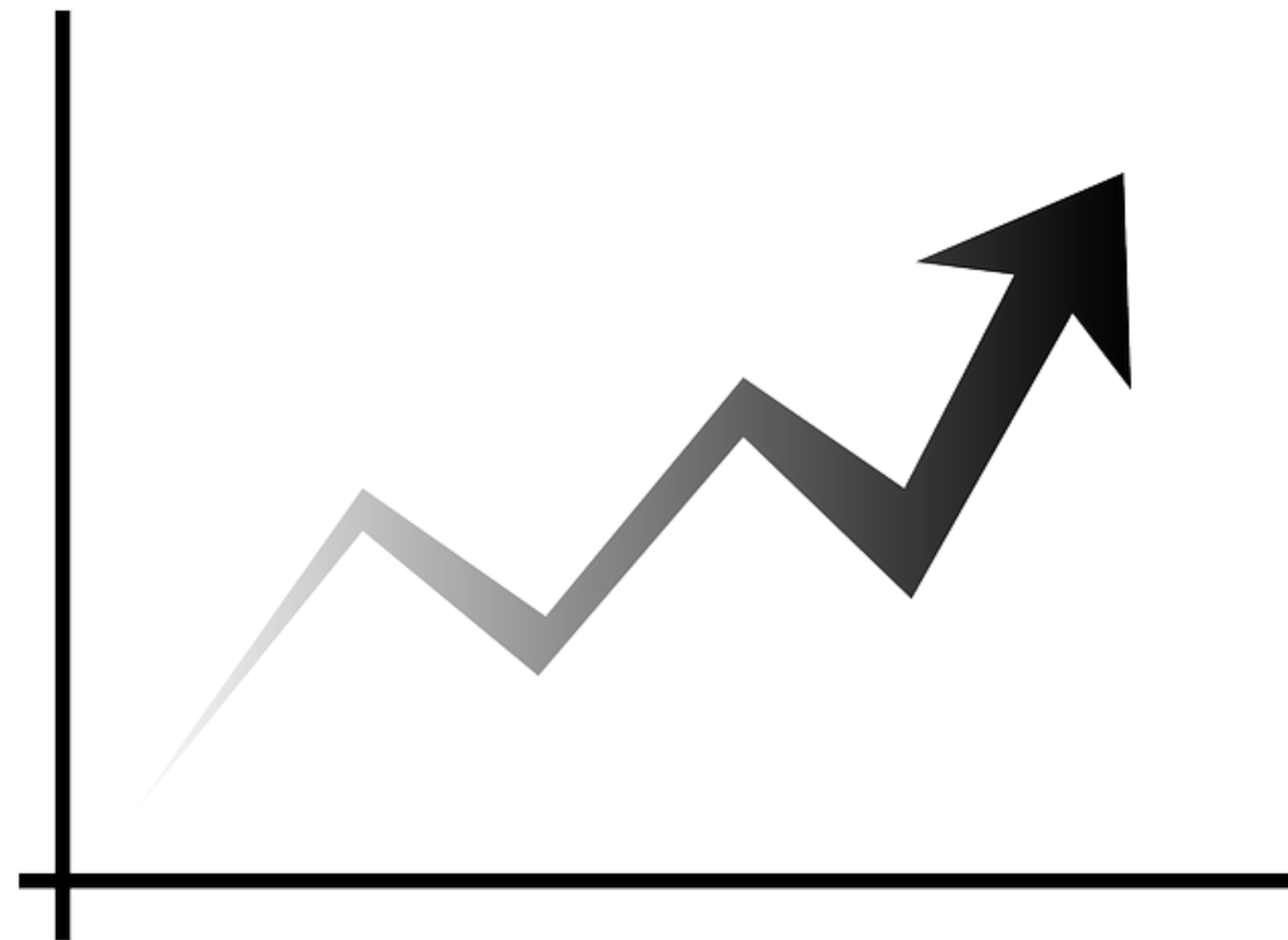
## The results of the analysis allow you to:

- Understand whether a phenomenon has a **greater incidence** in one territory rather than another
- Evaluate – given these differences – whether it is necessary to look for and analyse **new data** to understand whether there is correlation with other phenomena.



## The development of a phenomenon over time

- Given a certain phenomenon, the aim is to understand how it recurs over time



## Some examples

- An accident occurs on a section of road: is it the first time this happened or is that section often affected by similar events?
- A flood causes damage: was it more intense than similar episodes or is it a territorial/town-planning problem?



## The example: the impact of vaccines in the 20th century

- What was the trend of infectious diseases in the United States?
- What impact have vaccines had on individual diseases?
- Are there other significant territorial differences in proportion to the number of infected people?
- Are there some interesting outliers?



## SOME EXAMPLES: DATA JOURNALISM INVESTIGATIONS



## Battling Infectious Diseases in the 20th Century: The Impact of Vaccines

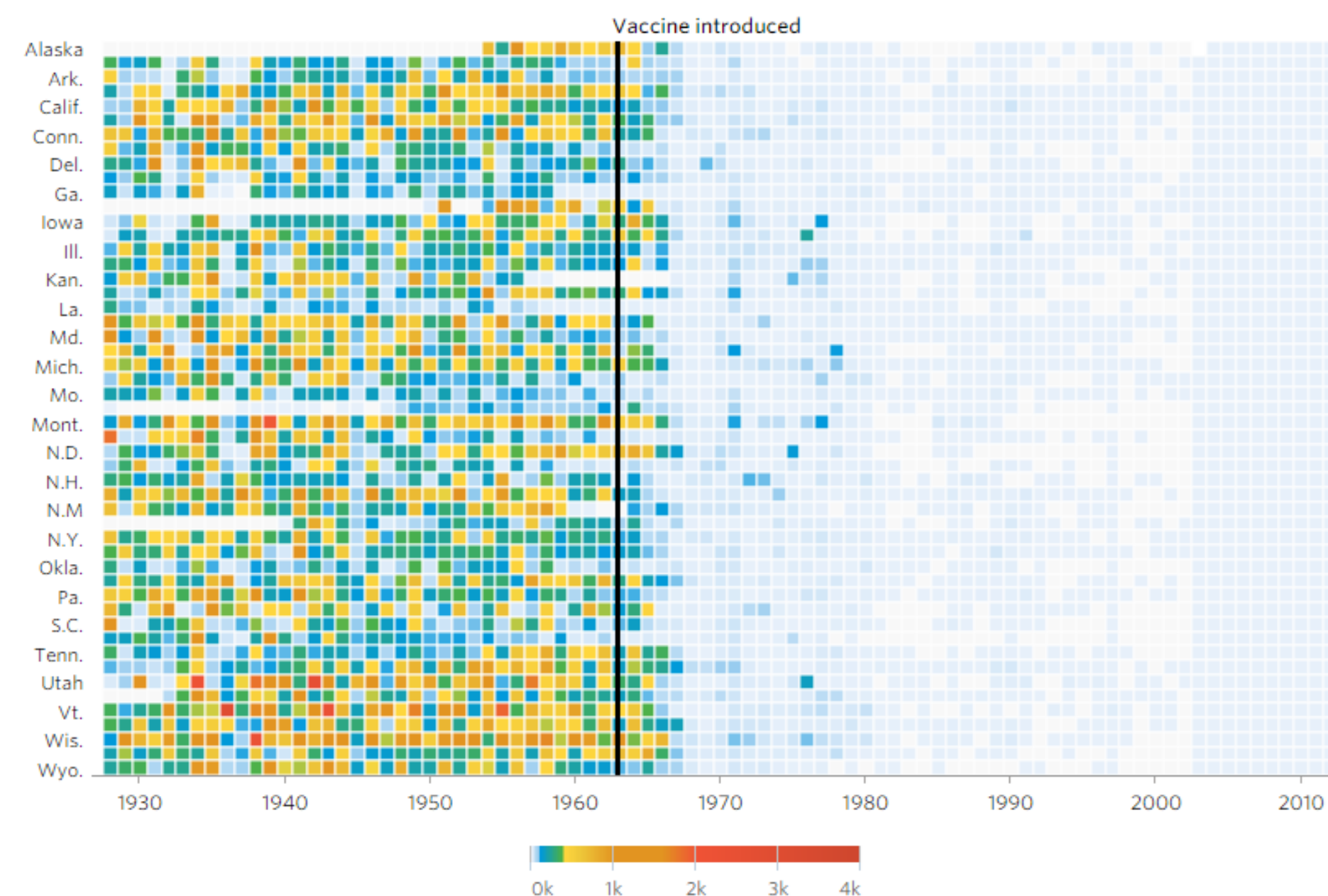
By [Tynan DeBold](#) and [Dov Friedman](#)

Published Feb. 11, 2015 at 3:45 p.m. ET

The number of infected people, measured over 70-some years and across all 50 states and the District of Columbia, generally declined after vaccines were introduced.

The heat maps below show number of cases per 100,000 people.

### Measles



**Interactive heat maps** showing the number of infected people, measured over a period of 80 years (from 1930 to 2010) in 50 American states. How have infections diminished following the introduction of vaccines?

**WALL STREET JOURNAL**  
(Data Journalism Award 2015,

[Best Data Visualization](#))

<http://graphics.wsj.com/infectious-diseases-and-vaccines/>

## Civic and participatory journalism

Many Data Journalism investigations, especially those related to government transparency, spring from an organised system of **information exchange**, where researchers, sociologists, experts or lay citizens who are passionate or interested about the topic, work alongside journalists.

## SOME EXAMPLES: DATA JOURNALISM INVESTIGATIONS

lanacion.com  
VozData

Últimas noticias | Edición impresa | Blogs | LN Data | Servicios | Guía LA NACION

lanacion.com | VozData | Gastos Del Senado

Acerca | Preguntas frecuentes | Aviso legal | Contacto

Login

### Gastos del Senado



LA NACION te invita a colaborar en la construcción de una base de datos sobre documentos públicos de gastos del Senado entre 2010 y 2012. En este primer proyecto, VozData te permite participar del proceso de revisión y clasificación de más de 6500 rendiciones del Senado para transformar estos datos en información que se irá publicando en tiempo real en este sitio. Ver notas sobre gastos del Senado

Liberá un documento

¡Ya revisamos 2265 de 6657 documentos!

Gastos del Senado procesados: \$ 22.727.632

Compartir

144 277 13

Compartir Tweet +1 Enviar por e-mail

Ranking de adjudicatarios ver mas

1	Honorable Senado de la N...	\$ 12.416.751
2	Aerolíneas Argentinas S.A.	\$ 1.322.595
3	Derudder hnos. S.R.L.	\$ 835.504
4	Empresa General Urquiza ...	\$ 709.206

Ranking de rubros ver mas

1	Pasajes aéreos, terrestre...	\$ 6.952.863
2	Caja chica	\$ 4.791.156
3	Viáticos para agentes	\$ 2.937.438
4	Viáticos y gastos	\$ 1.955.022

Ranking de usuarios Por cantidad de documentos revisados

	Jeepika	511
	jaelariadna	428
	alkymia	411

The **VozData** initiative in Argentina, a platform that combines participation in civic monitoring with data journalism investigations.

### LA NACION

<http://blogs.lanacion.com.ar/projects/data/vozdata-ii-civic-participation-for-investigative-reporting-and-educational-platform/>





**WHAT IS DATA JOURNALISM**  
**(part one)**