



**Barcelona  
Supercomputing  
Center**

Centro Nacional de Supercomputación



# Proyecto Final

---

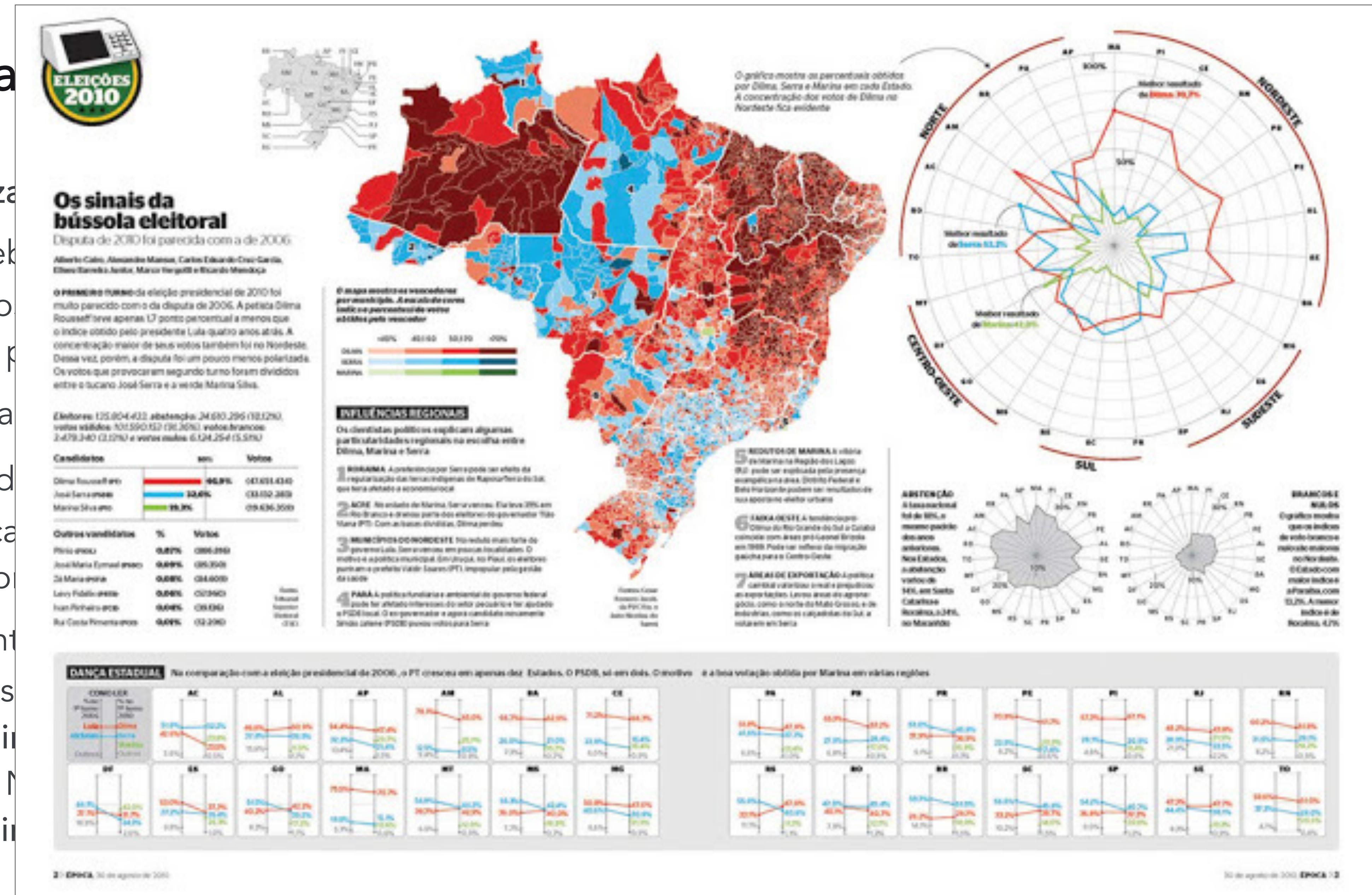
Guillermo Marin  
[guillermo.marin@uab.cat](mailto:guillermo.marin@uab.cat)

# Proyecto Final

- **Hacer una visualización** con datos de elección libre.
- La visualización debe servir a un propósito → Hipótesis
  - Puede guiarnos hacia la respuesta, o
  - Mostrarnos la pregunta visualmente para que el usuario busque las respuestas
- Compuesta de una o más gráficas
- Con todos los medios necesarios para ser comprensible de manera autónoma:
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elección libre
  - **Infografía:** Poster con una o varias gráficas estáticas + texto
  - **Visualización interactiva:** Dashboard o visualización multi-panel
  - **Scrollitelling:** Narración interactiva basada en browser de gráficas + texto
  - **Data storytelling:** Presentación basada en gráficas
  - ...

# Proyecto Final

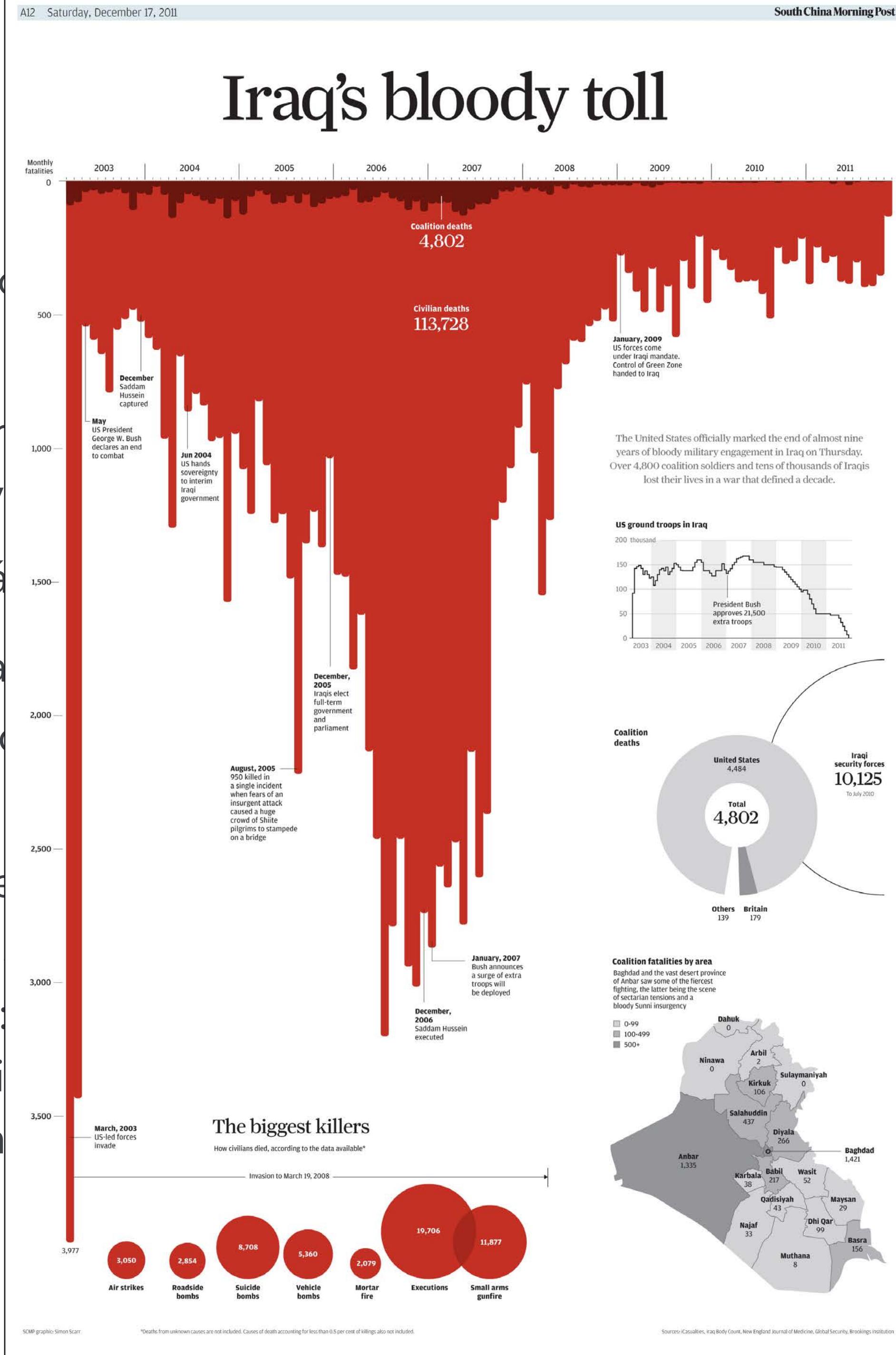
- Hacer una visualización
- La visualización debe:
  - Puede guiarnos
  - Mostrarnos la parte
- Compuesta de una
- Con todos los medios:
  - Leyendas, escala
  - Texto de soporte
- Soporte de presentación:
  - **Infografía:** Positivo
  - Visualización interactiva
  - Scrollitelling: Negativo
  - Data storytelling: Positivo
  - ...



The functional art: Radar graphs: Avoid them (99.9% of the time)

# Proyecto Final

- Hacer una visualización con datos
- La visualización debe servir a un propósito  
  - Puede guiarnos hacia la respuesta
  - Mostrarnos la pregunta que queremos responder
- Compuesta de una o más gráficas
- Con todos los medios necesarios  
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de datos  
  - **Infografía:** Poster con un diseño estético
  - **Visualización interactiva:** Web
  - **Scrollitelling:** Narración interactiva
  - **Data storytelling:** Presentación
  - ...



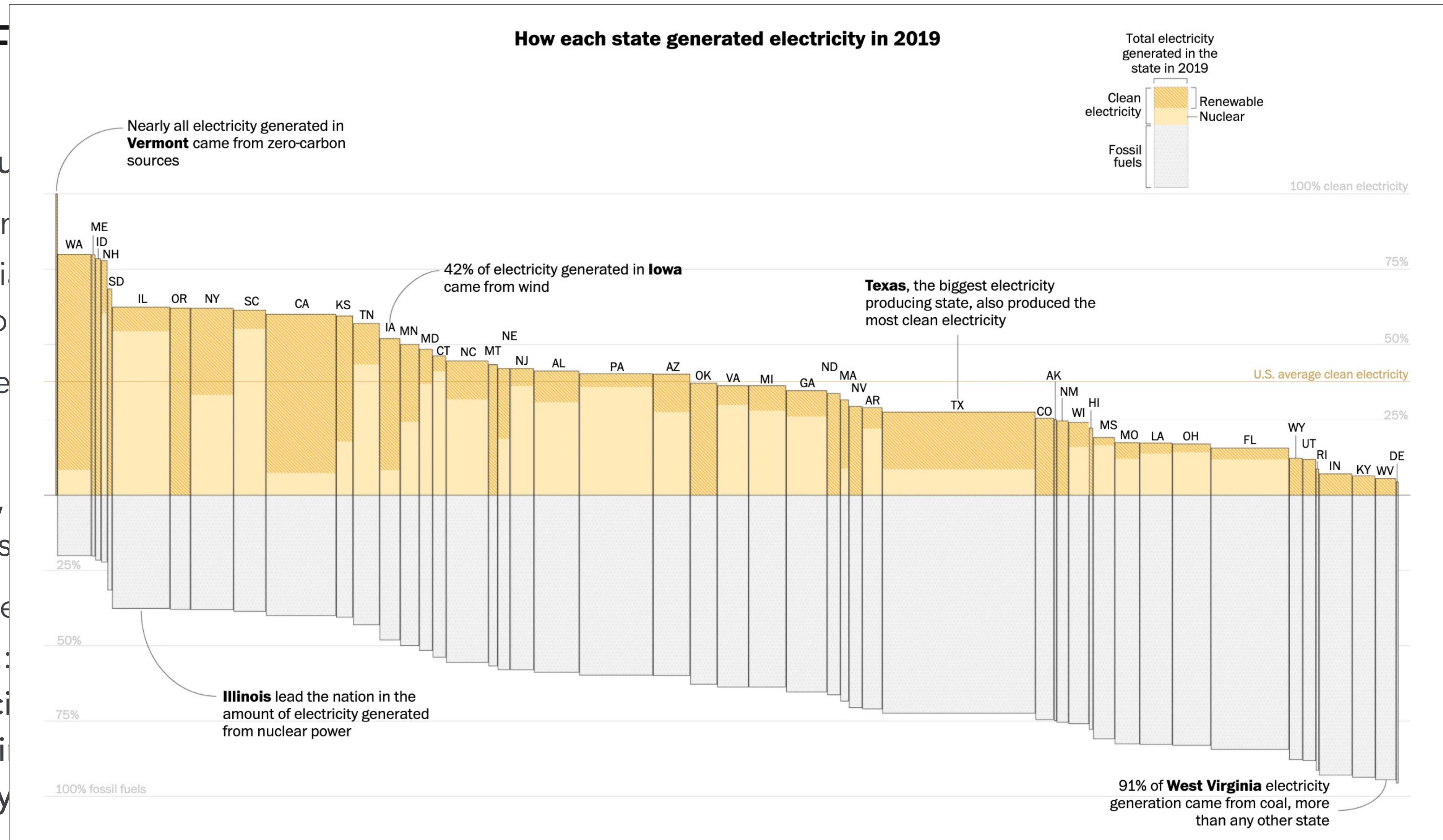
spuestas

oma:

exto

# Proyecto Final

- Hacer una visualización
- La visualización:
  - Puede guiar
  - Mostrarnos
- Compuesta de:
  - Con todos los
    - Leyendas,
    - Texto de subtitulos
- Soporte de presentación
  - **Infografía:**
  - Visualizaciones
  - Scrollitelli
  - Data story
  - ...

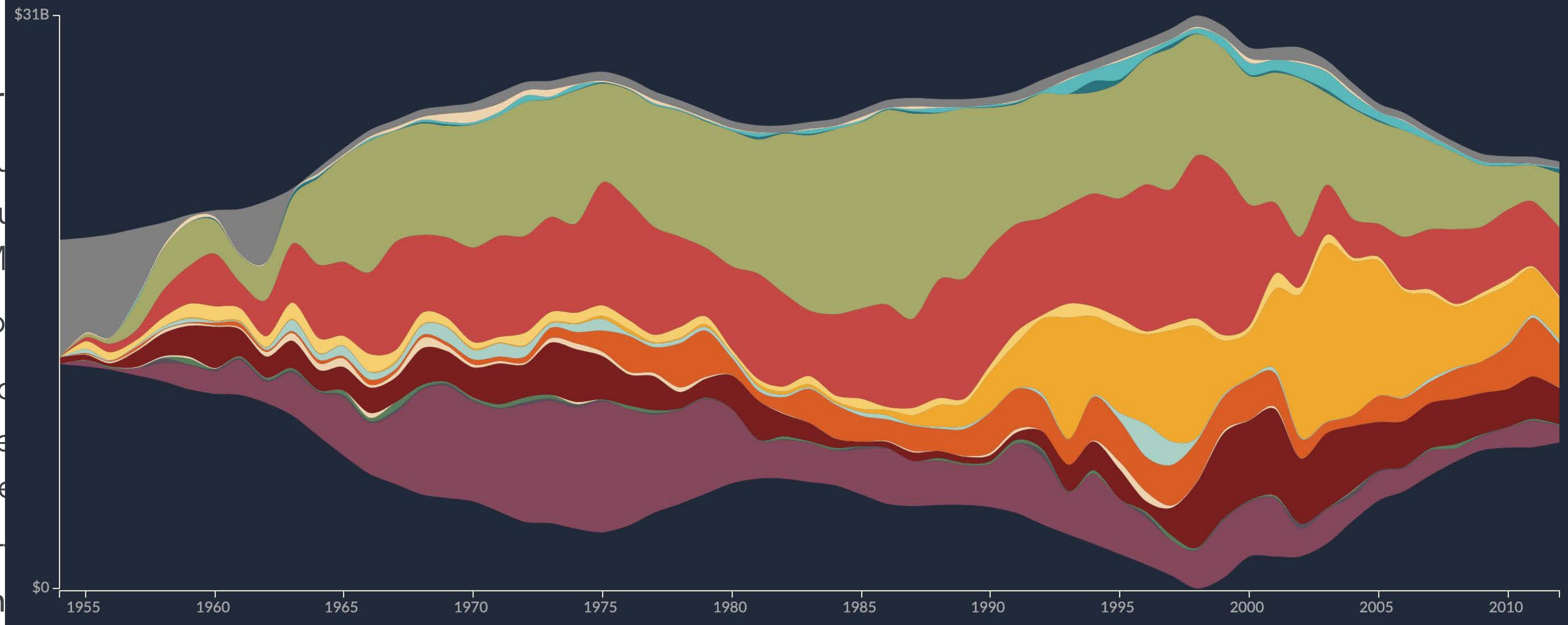


The Washington Post

# Proyecto

POPULARITY OF THE TOP MUSIC GENRES.  
DOUBLE CLICK TO SEE MORE

SCALE BY: PERCENTAGE  TOTAL SALES



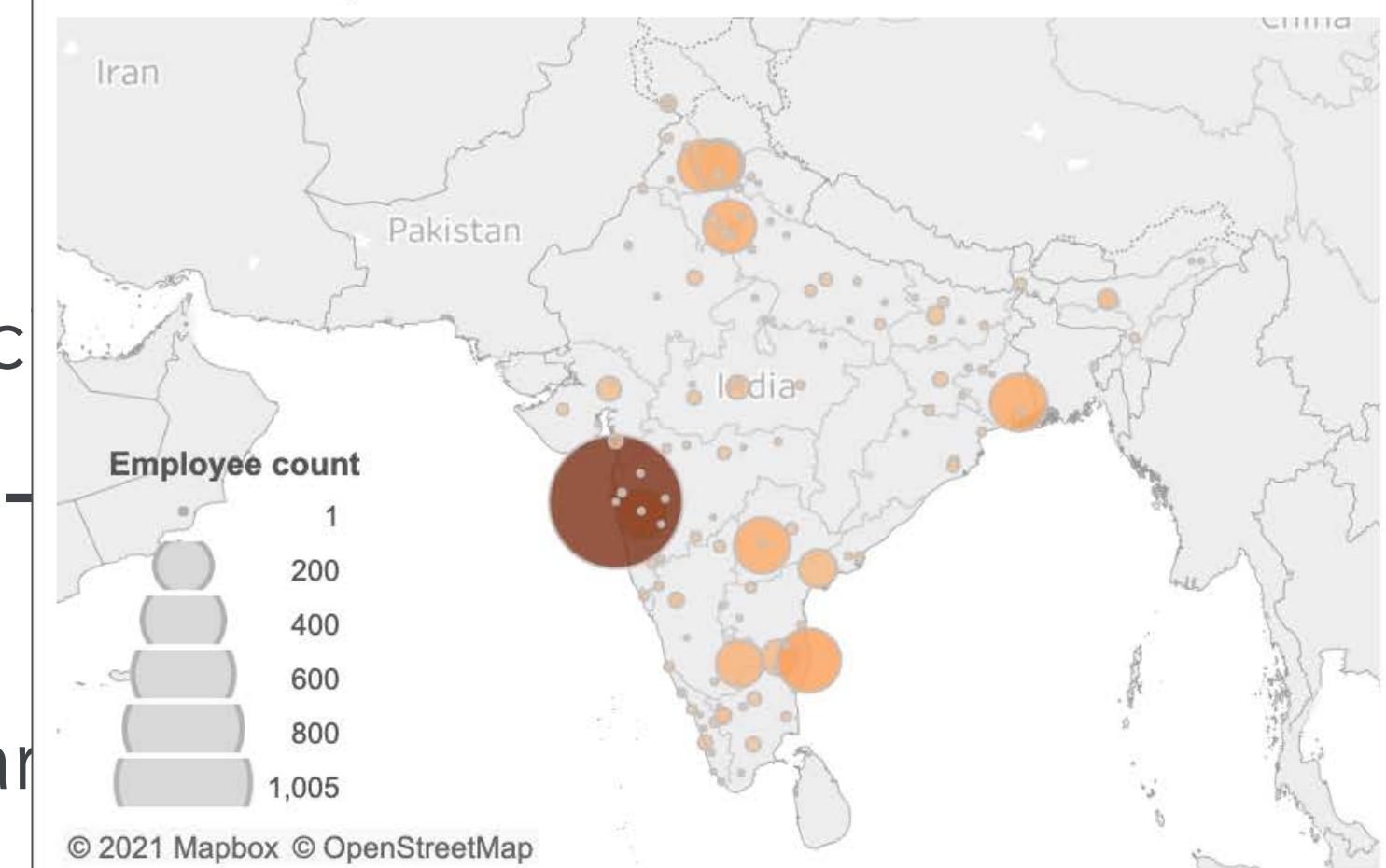
- Hacer dashboard interactivos
- La visualización interactiva
  - Puedes hacer scrollitelling
  - Me encanta la narración interactiva
- Compartir en diferentes plataformas
- Con todos los datos que necesitas
  - Leer los datos
  - Tenerlos en un repositorio
- Soporitir el desarrollo de las aplicaciones
- Integrar con otras tecnologías
  - Visualización interactiva: Dashboard o visualización multi-panel
  - Scrollitelling: Narración interactiva basada en browser de gráficas + texto
  - Data storytelling: Presentación basada en gráficas
  - ...

# Proyecto Final

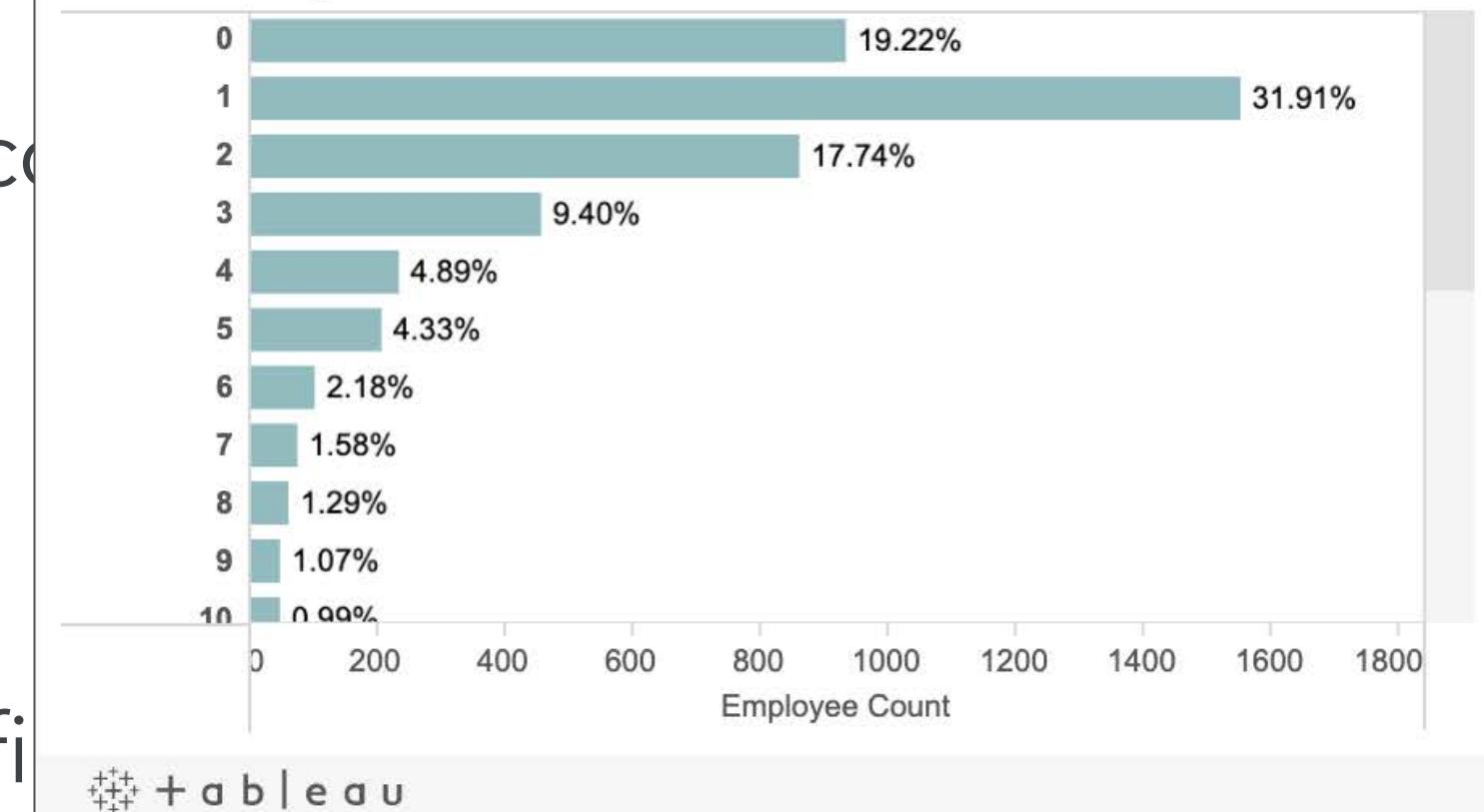
- Hacer una visualización con datos de elección libre
- La visualización debe servir a un propósito -
  - Puede guiarnos hacia la respuesta, o
  - Mostrarnos la pregunta visualmente para que sea más fácil de responder.
- Compuesta de una o más gráficas
- Con todos los medios necesarios para ser comprensible
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elección libre
  - **Infografía:** Poster con una o varias gráficas
  - **Visualización interactiva:** **Dashboard** o visualización multi-panel
  - **Scrollitelling:** Narración interactiva basada en browser de gráficas + texto
  - **Data storytelling:** Presentación basada en gráficas
  - ...

## HR Dashboard

Where are they located?



Tenure Length?



++ + a b | e a u

Hire Date

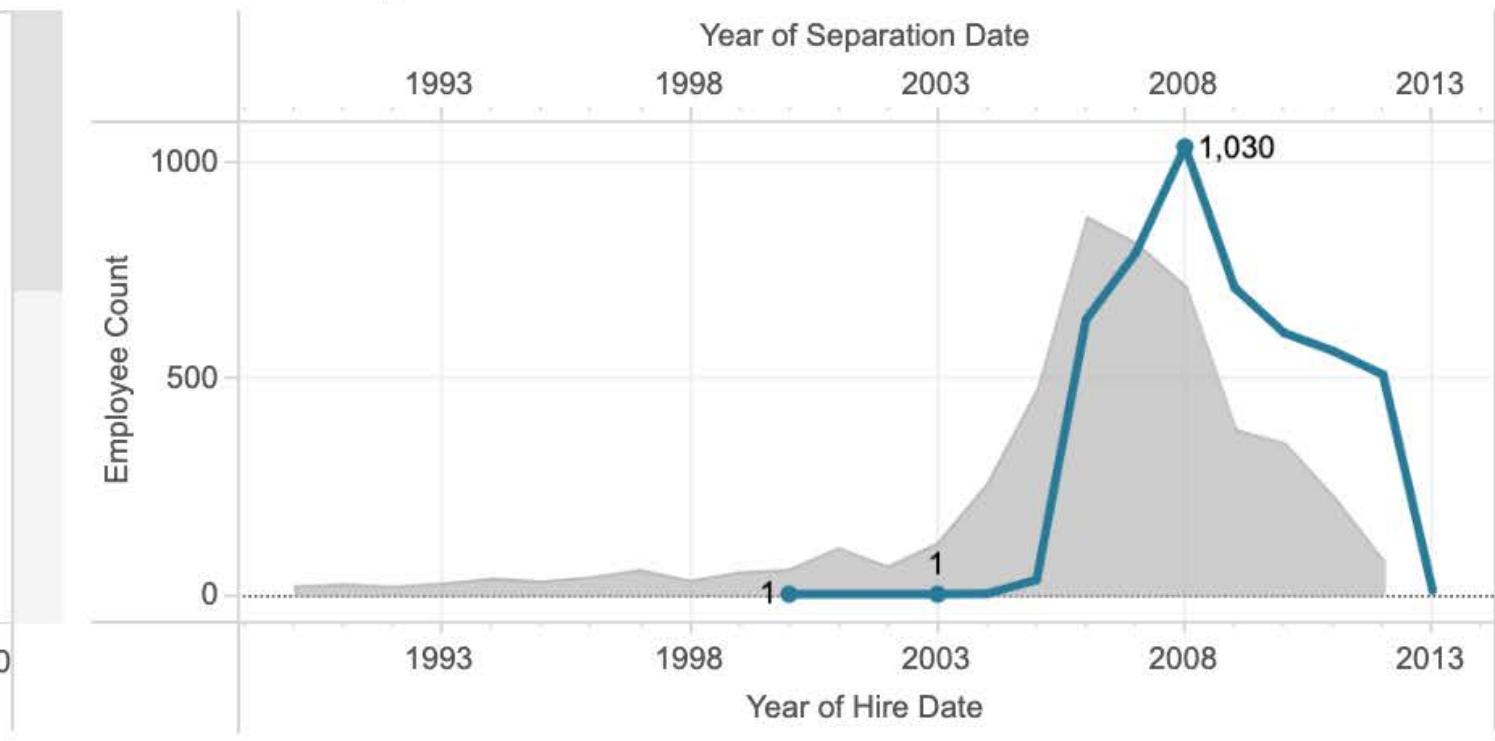
1/1/1990

10/15/2012

Location

(All)

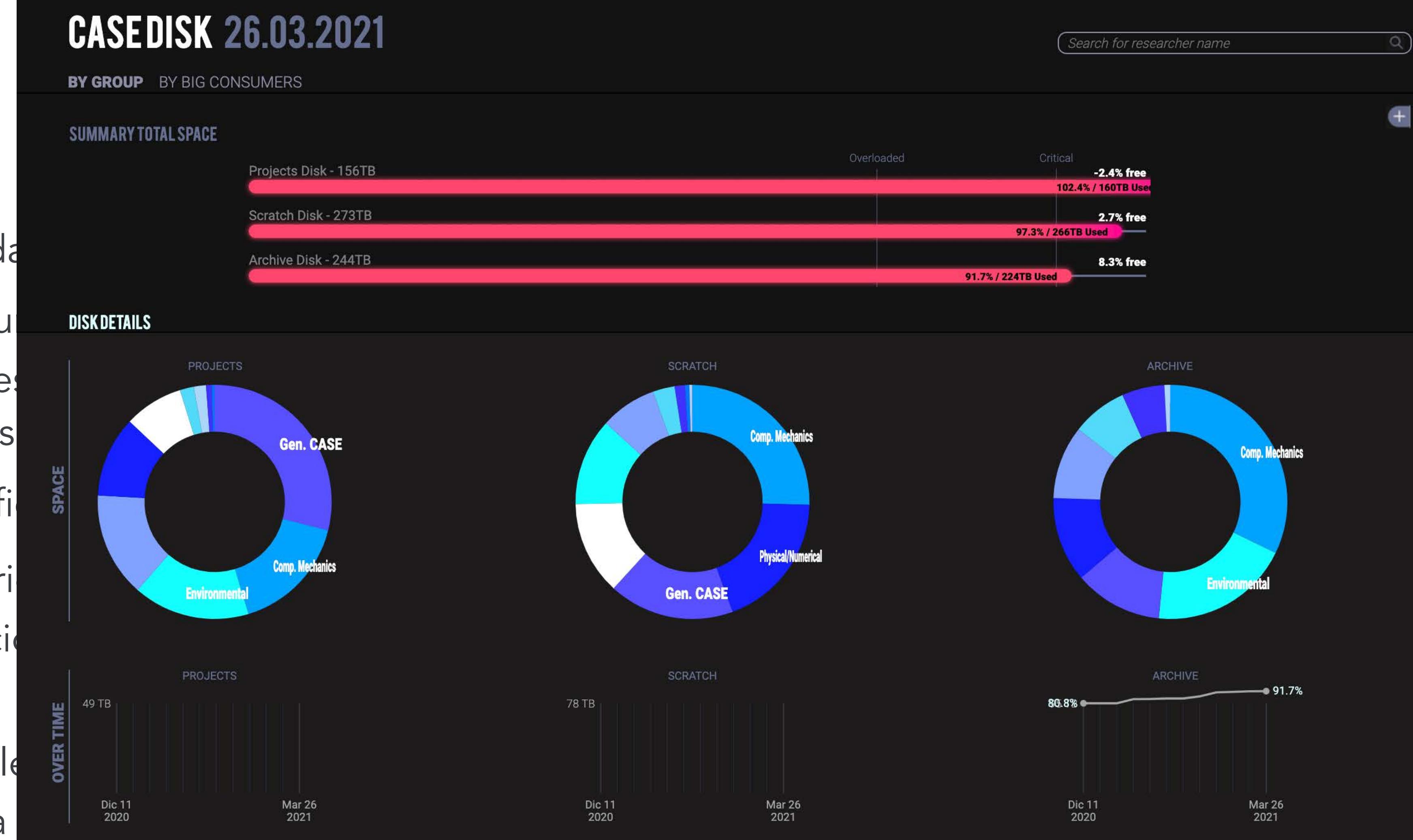
When did they leave?



Viz Author: [Core Product Marketing](#)

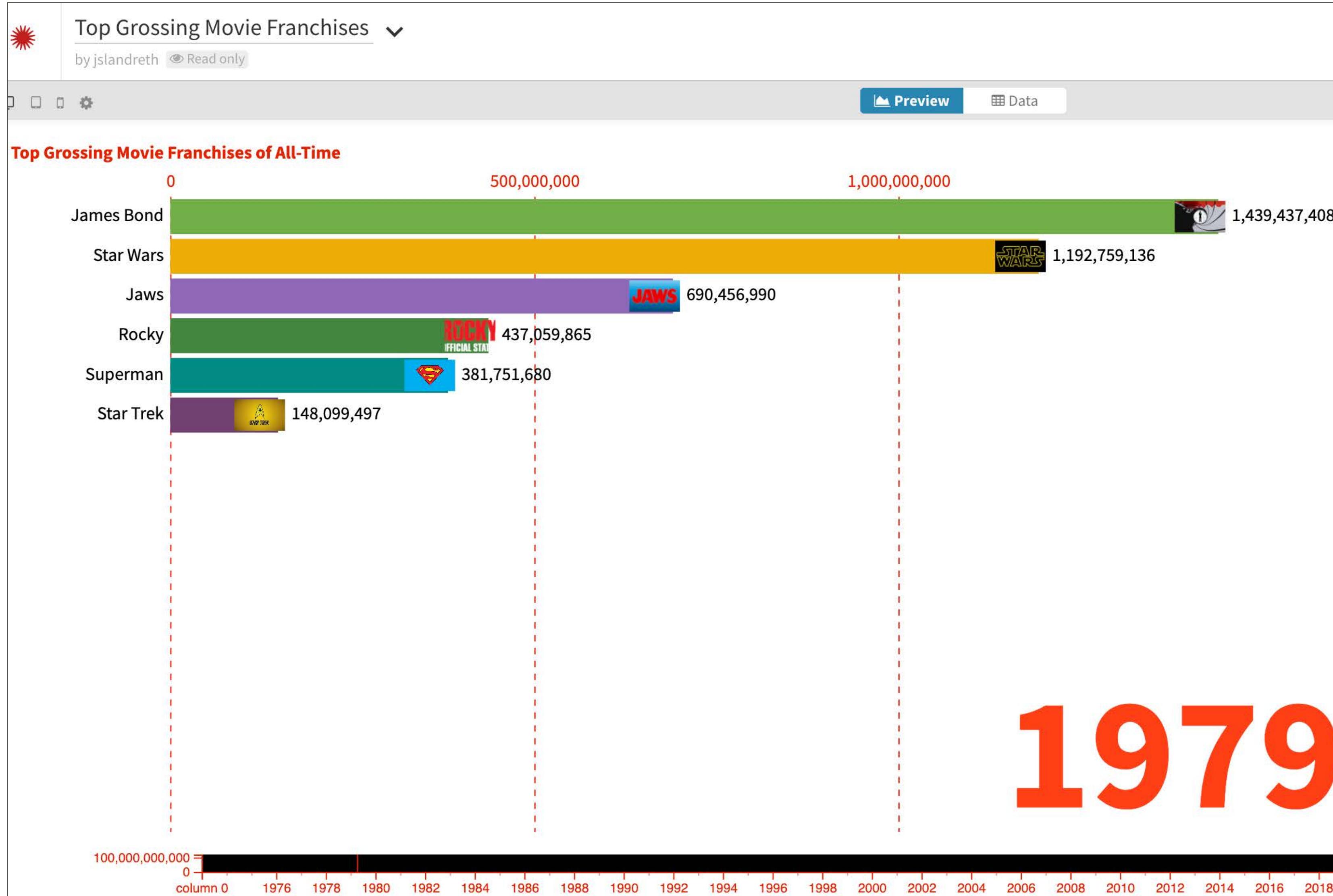
# Proyecto Final

- Hacer una visualización con datos
- La visualización debe servir a un público objetivo
  - Puede guiarnos hacia la respuesta
  - Mostrarnos la pregunta visualizada
- Compuesta de una o más gráficas
- Con todos los medios necesarios
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elementos
  - Infografía: Poster con una visión general
  - Visualización interactiva: **Dashboard** o visualización multi-panel
  - Scrollitelling: Narración interactiva basada en browser de gráficas + texto
  - Data storytelling: Presentación basada en gráficas
  - ...



# Animación. Ranking de franquicias de cine por recaudación

<https://public.flourish.studio/visualisation/3268762/>



# Proyecto Final

## Wealth shown to scale

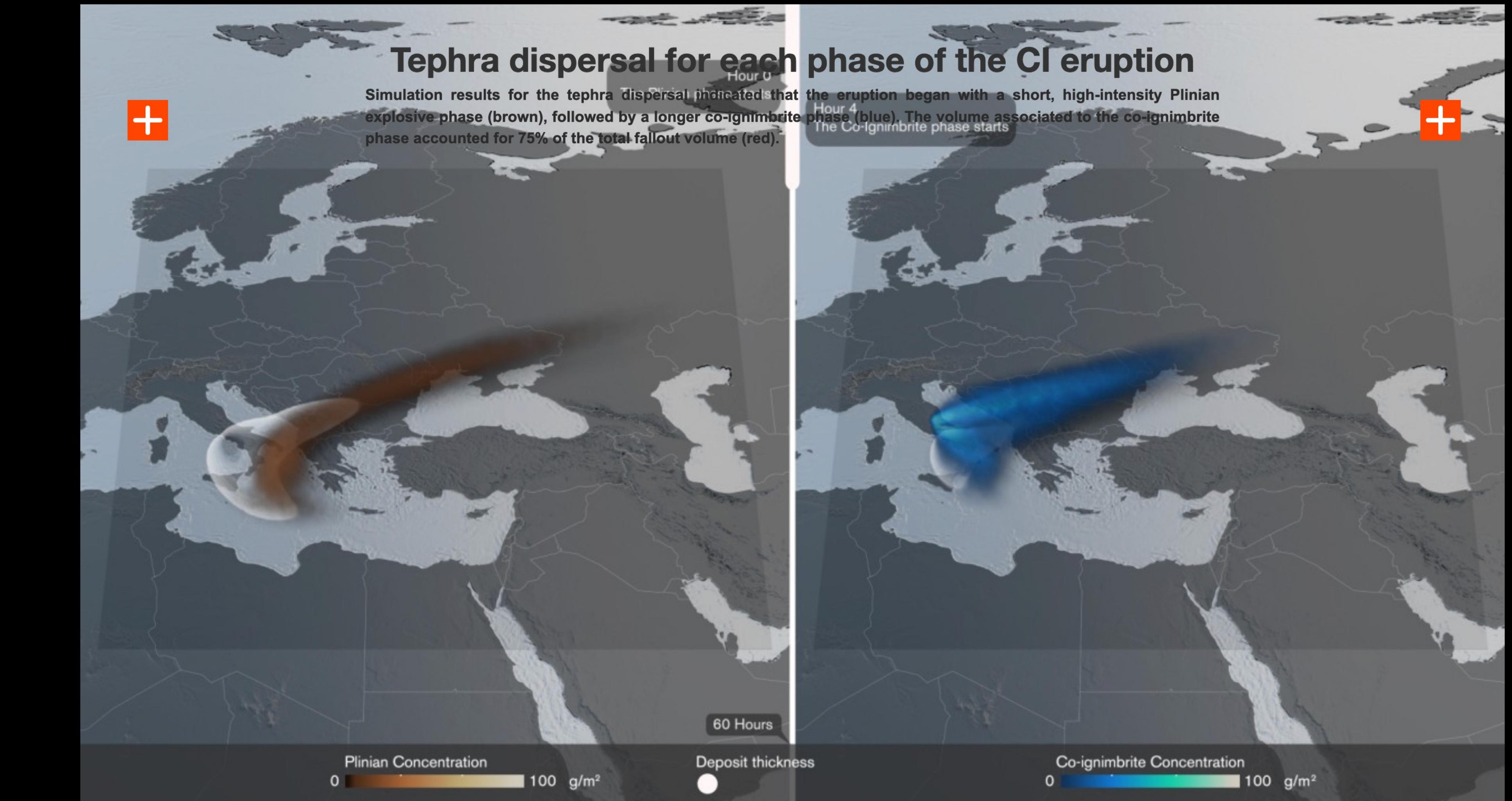
- Hacer una visualización con datos de elección
- La visualización debe servir a un propósito →
  - Puede guiarnos hacia la respuesta, o
  - Mostrarnos la pregunta visualmente para responderla.
- Compuesta de una o más gráficas
- Con todos los medios necesarios para ser comprendida
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elección libre
  - **Infografía:** Poster con una o varias gráficas estáticas + texto
  - **Visualización interactiva:** Dashboard o visualización multi-panel
  - **Scrollitelling:** Narración interactiva basada en browser de gráficas + texto
  - **Data storytelling:** Presentación basada en gráficas
  - ...



Wealth shown to scale - Matt Korostoff <https://mkorostoff.github.io/1-pixel-wealth/>

# Proyecto Final

- Hacer una visualización
- La visualización:
  - Puede guiar
  - Mostrarnos
- Compuesta de:
  - Con todos los medios
  - Leyendas, etc.
  - Texto de soporte
- Soporte de presentación
  - Infografía:
  - Visualización
  - **Scrollitelling**: Narración interactiva basada en browser de gráficas + texto
  - Data storytelling: Presentación basada en gráficas
  - ...



[https://www.bsc.es/viz/campanian\\_ignimbrite/](https://www.bsc.es/viz/campanian_ignimbrite/)

BSC-CNS

# Proyecto Final



- Hacer una visualización de datos
- La visualización debe servir:
  - Puede guiarnos hacia la respuesta
  - Mostrarnos la pregunta
- Compuesta de una o más visualizaciones
- Con todos los medios necesarios:
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elección libre
  - **Infografía:** Poster con una o varias gráficas estáticas + texto
  - **Visualización interactiva:** Dashboard o visualización multi-panel
  - **Scrollitelling:** Narración interactiva basada en browser de gráficas + texto
  - **Data storytelling:** Presentación basada en gráficas
  - ...

US Gun killings in 2018 – Periscopic <https://guns.periscopic.com/>

# Objetivos

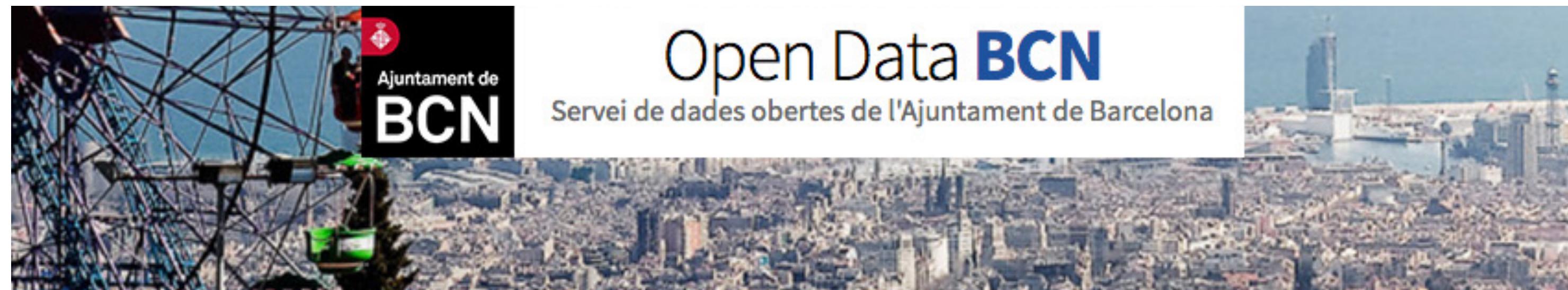
- Hacer una visualización con datos de elección libre [Una o varias gráficas]
- Formular hipótesis y usar la visualización para responderla
- La visualización puede ser un póster, una web interactiva, un dashboard, una animación, una presentación ...
- La entrega consistirá en la visualización, un reporte del proceso, y la defensa del proyecto.
- Aplicar conceptos vistos en clase:
  - Data treatment: scrapping, infraestructura (BBDD), Data Massage, ...
  - Análisis: Estadística, cruzar datasets relacionados, Dimensionality Reduction, ML, ...
  - Análisis visual exploratorio.
  - Gráficas avanzadas (multipanels/facets, SPLOM, mapas, etc.)
  - Interactividad, animación, diseño, ...

# Datos

Fuentes de datos recomendadas:

- Open Data BCN: Servei de dades obertes de l'Ajuntament de Barcelona  
<https://opendata-ajuntament.barcelona.cat/es>
- Our World in Data: Banc de dades amb activitats humanes  
<https://ourworldindata.org/>
- Kaggle: Banc de dades obert per usuaris <https://www.kaggle.com/datasets>
- Datos propios: Smartwatch, tracking de Smartphone, historial de Spotify, ...

# Datos



Open Data BCN: Servei de dades obertes de l'Ajuntament de Barcelona

<https://opendata-ajuntament.barcelona.cat/es>

- Tablas formateadas
- Padrón de habitantes (demografía), cambios de domicilio, altas y bajas..
- Fallecimientos, fecundidad, esperanza de vida
- Niveles de renta, servicios, hospitales, servicios sociales, centros de día...
- Territorio: Catastro, patrimonio Arq., precio vivienda...
- Ruido, calidad del aire, temperatura.
- Accidentes e incidentes gestionados por Guardia Urbana
- ...
- Casi todos tienen localización geográfica, muchos son históricos

## Temas

**Administración (89)**

**Ciudad y servicios (155)**

**Economía y empresa (17)**

**Población (111)**

**Territorio (117)**

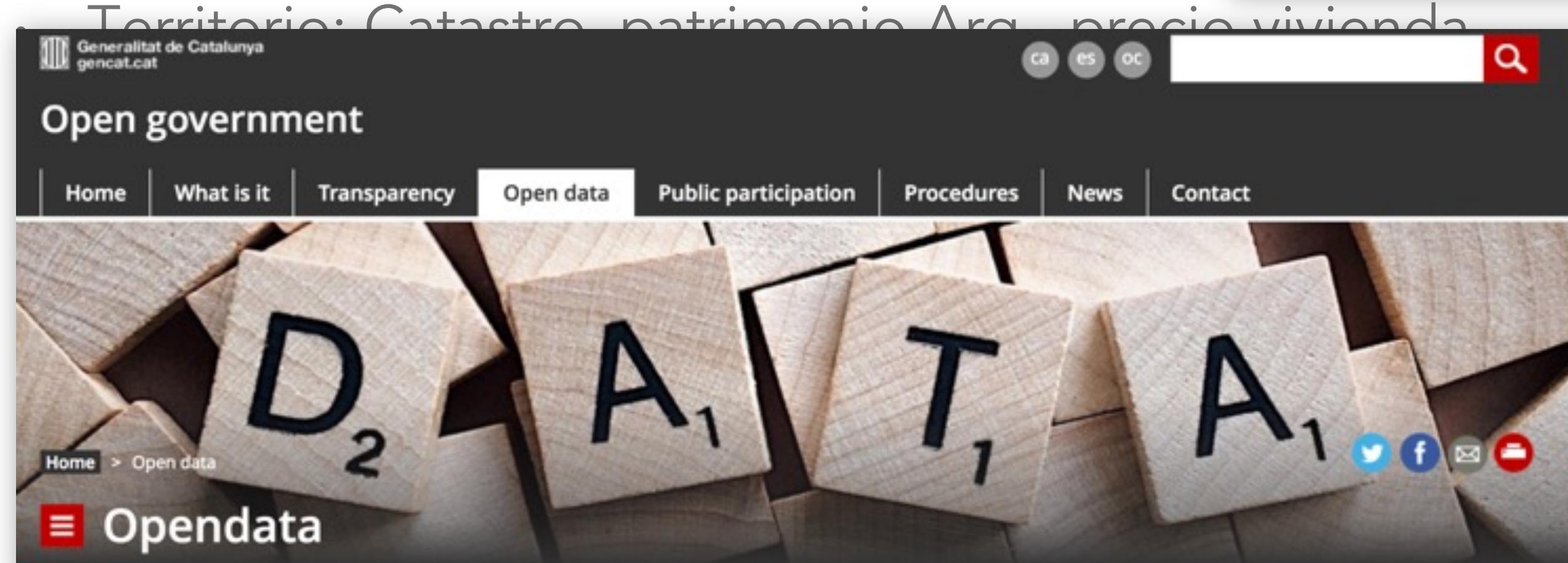


The screenshot shows the Open Data BCN website. At the top, there's a banner with a photo of a cable car over a cityscape, the text 'Ajuntament de BCN', and 'Open Data BCN'. Below the banner, it says 'Servei de dades obertes de l'Ajuntament'. A blue header bar has a magnifying glass icon and the word 'Temas'. On the right, there are links for 'Contacte', 'CA', and a search bar. The main content area has a dark header 'gencat' with 'Mossos d'Esquadra'. Below it is a navigation bar with 'Inici', 'Els Moss...', 'Temes', 'Serveis i...', 'Consells...', 'Actualitat', and 'Cita prè...'. Underneath is a breadcrumb trail: 'Inici > Els M > Indic > Dades > Catàleg'. To the right are social media icons for Twitter, Facebook, Telegram, WhatsApp, and Print. The main section is titled 'Catàleg' with a subtitle 'Cercador' and a search input field. It features four categories with icons: 'Dades organitzatives' (organization chart), 'Activitat institucional' (building), 'Activitat policial' (police car), and 'Dades delinqüencials' (downward arrow).



The screenshot shows the datos.gob.es website. The logo 'datos.gob.es' is in large purple letters, with 'reutiliza la información pública' in smaller purple text below it. To the right, it says 'obertes de na.cat/es'.

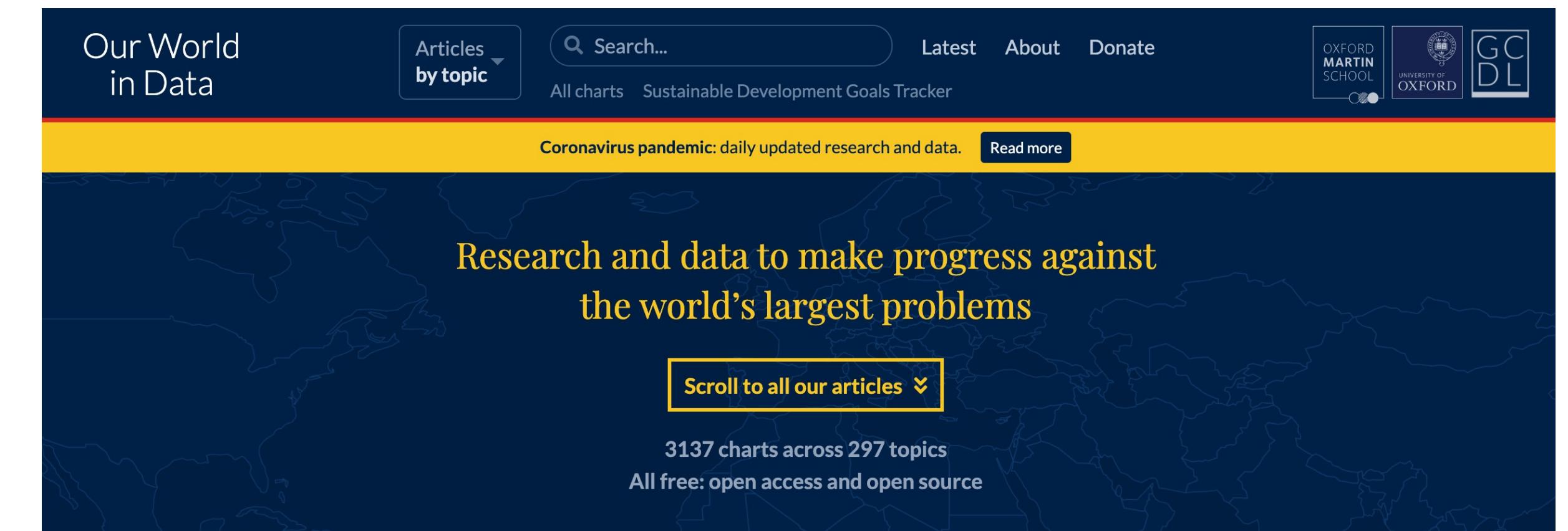
- Padrón de habitantes (demografía), cambio
- Fallecimientos, fecundidad, esperanza de v
- Niveles de renta, servicios, hospitales, servi



The screenshot shows the Open government website. The header includes the logo 'Generalitat de Catalunya gencat.cat', language options 'ca', 'es', 'oc', and a search bar. The main menu has 'Home', 'What is it', 'Transparency', 'Open data' (which is highlighted in white), 'Public participation', 'Procedures', 'News', and 'Contact'. Below the menu is a large image of wooden blocks spelling out 'DATOS'. The footer has a 'Home' link, 'Open data', and social media icons for Twitter, Facebook, Email, and Print. The page title is 'Territorio · Catastro · patrimonio · Árgo · precio vivienda'.

# Datos

- [www.datacatalog.worldbank.org](http://www.datacatalog.worldbank.org)
- [www.undatacatalog.org](http://www.undatacatalog.org)
- ...
- Complementarios a:  
Our World in Data <https://ourworldindata.org/>  
Artículos abiertos sobre datos socio económicos:
  - Mas enfocada en conocimiento que en datos solamente
  - Sirve para inspiración y porque todas las fuentes de datos están citadas y son utilizables
  - Datasets actualizados y completos de COVID-19
  - Datos sociales, económicos, salud, conflicto...



# Datos

Kaggle: Banc de dades obert per usuaris <https://www.kaggle.com/datasets>

- Comunidad de usuarios, orientada a data sc.
- Proponen desafíos y tareas, sobre todo de ML, pero también de análisis y viz
- Repositorio enorme y abierto de datos de cualquier tipo:
  - Catálogo de Netflix, uso de marihuana, nombres de perros...

- **Twitter?**

- **Datos vuestros:**

- Smart phone, watch
- Historial de Netflix, Spotify,
- ...

The screenshot shows the Kaggle platform interface. On the left is a sidebar with navigation links: Home, Compete (which is highlighted), Data, Notebooks, Communities, Courses, and More. Below this is a 'Recently Viewed' section with links to the Titanic competition, a machine learning tutorial, a neural network example, a project example, and an EDA notebook. The main content area features a large banner for the 'Titanic - Machine Learning from Disaster' competition, which is described as a 'Getting Started Prediction Competition'. It shows a black and white photo of the Titanic ship at night. Below the banner, there's a summary: 'Kaggle · 18,119 teams · Ongoing'. A navigation bar below the banner includes Overview, Data, Notebooks, Discussion, Leaderboard, Rules, Team, My Submissions, and a prominent 'Submit Predictions' button. The 'Overview' tab is selected. In the bottom right corner of the main content area, there's a welcome message: 'Ahoy, welcome to Kaggle! You're in the right place.' followed by detailed instructions for new users about the competition rules and how to get started.

## VISUALITZACIÓ DE DADES

### GRAU EN ENGINYERIA DE DADES

#### CURS 2021-22

#### PORTFOLI

#### VIOLÈNCIA D'ARMES ALS ESTATS UNITS

Álvaro Rodríguez, Roberto Manresa, Javier Ortega, Víctor García, Ona Folch

Bellaterra, Maig 2022

### 3.3.7. Nombre d'habitants i la violència d'armes

Finalment, per la setena i última hipòtesi hem de veure si existeix correlació entre el nombre d'habitants en un estat (població) i la quantitat d'actes violents que s'han produït, utilitzant l'últim dataset explicat de *popstates*. Hem fet un diagrama doble de barres que ens mostri el nombre d'habitants per cada estat (població estimada) al costat esquerre i el nombre d'incidents per milió d'habitants per l'altre costat, fet amb Tableau. Es pot veure el resultat a la Figura 38.

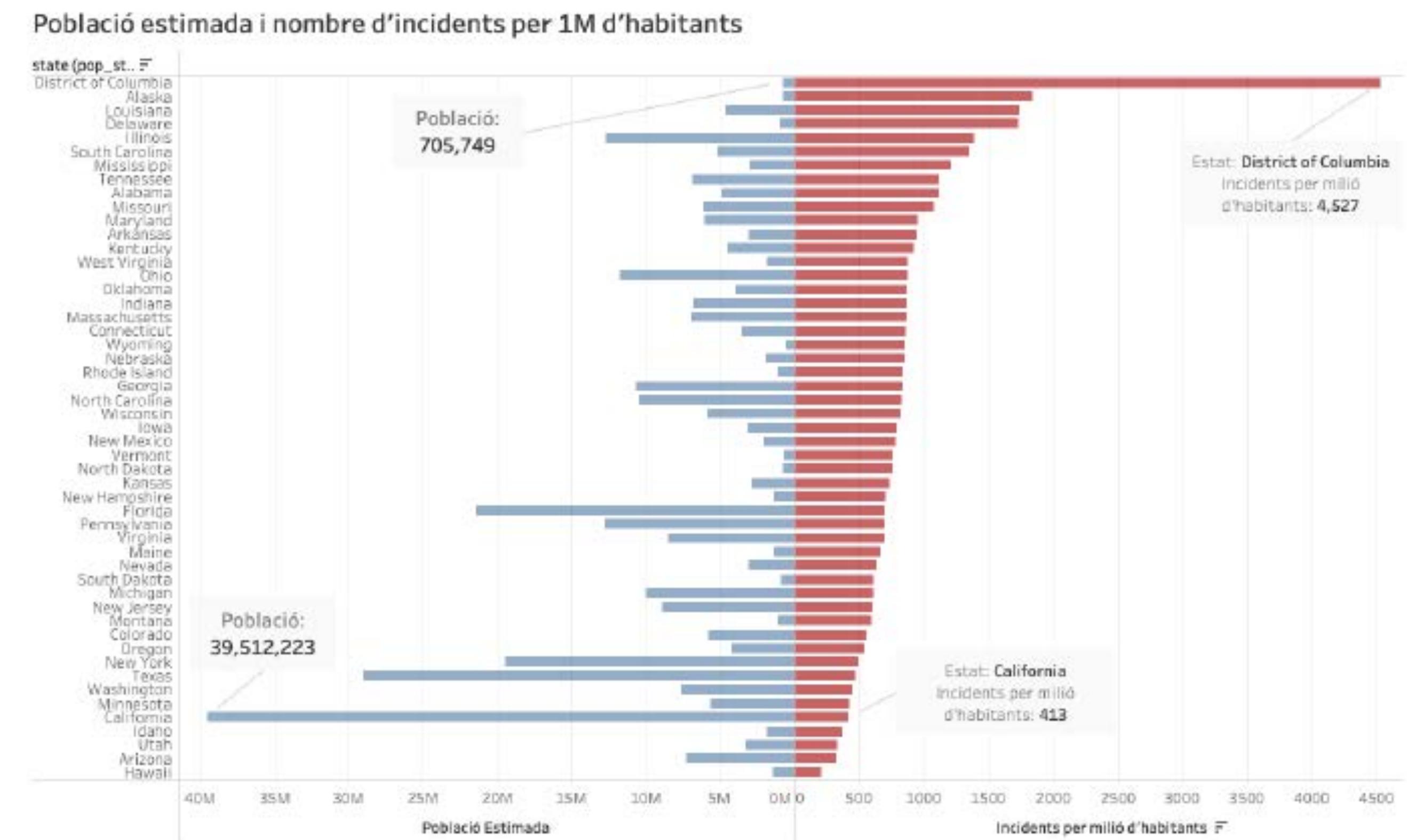
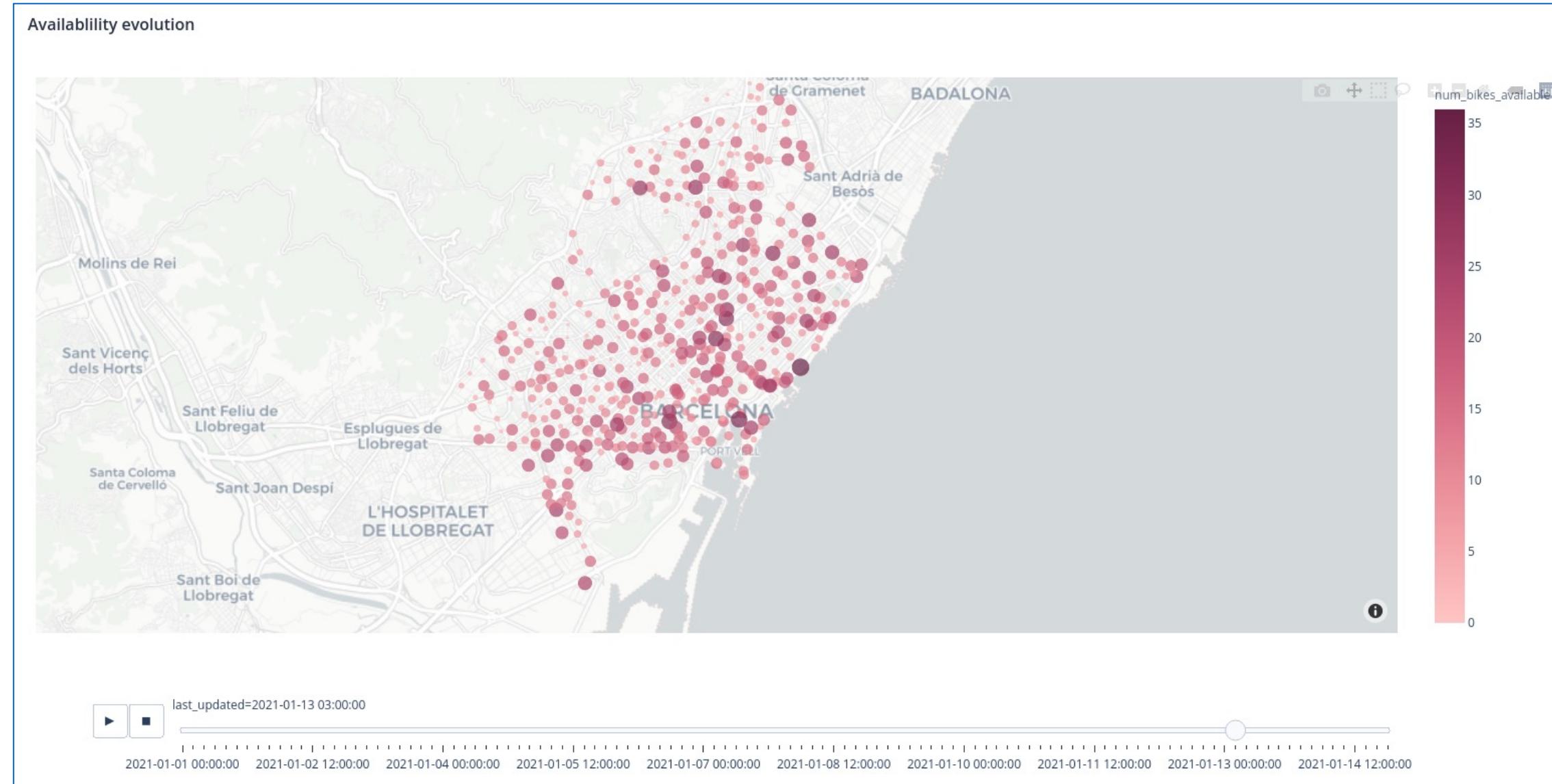
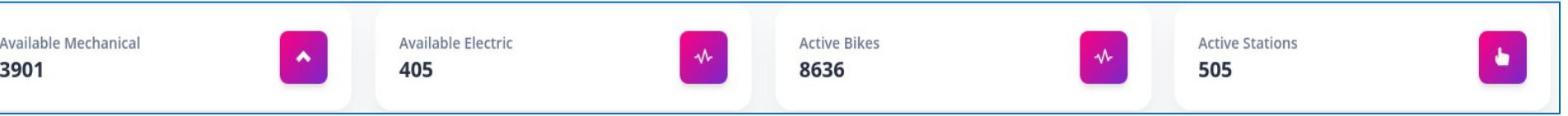


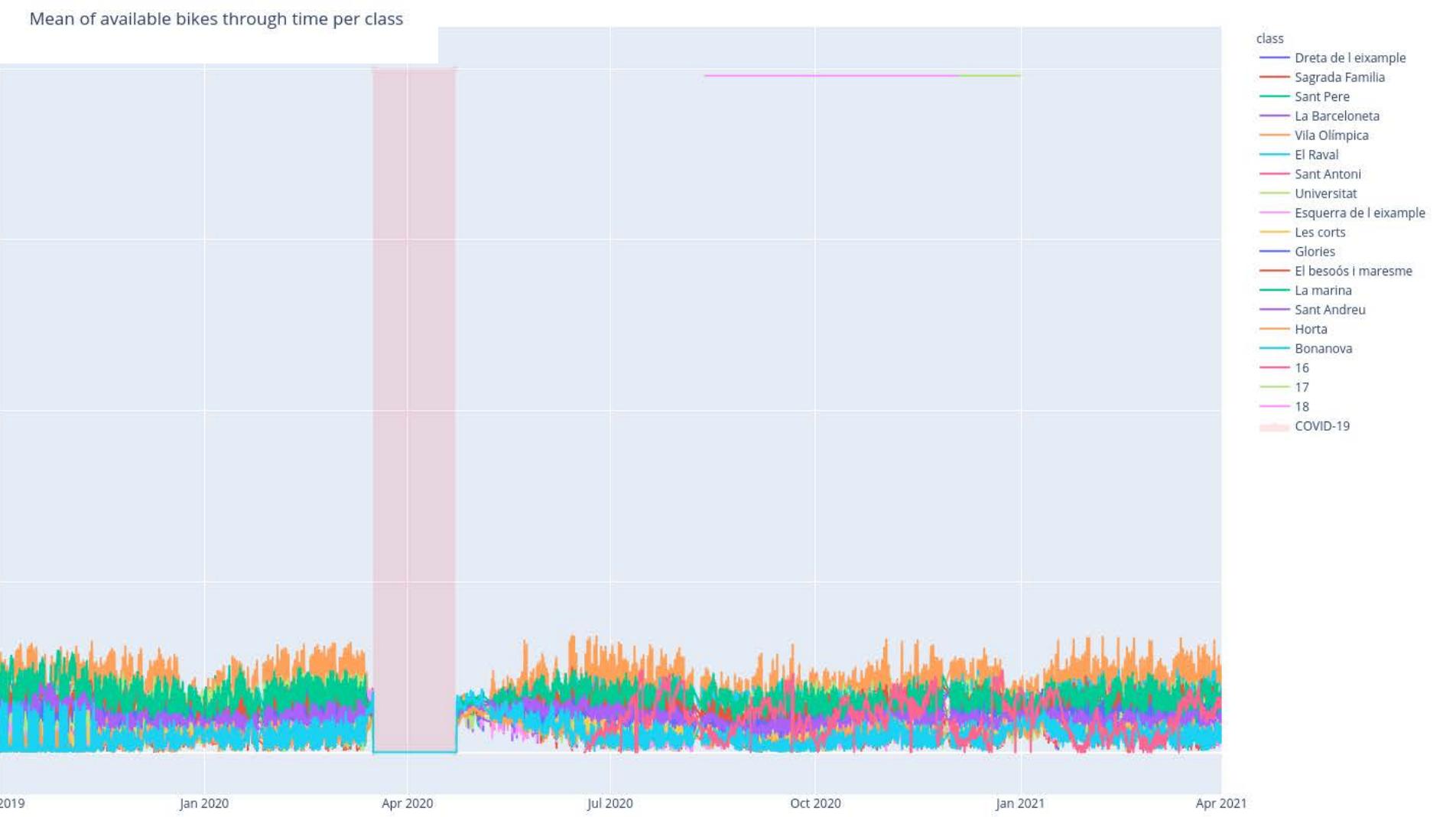
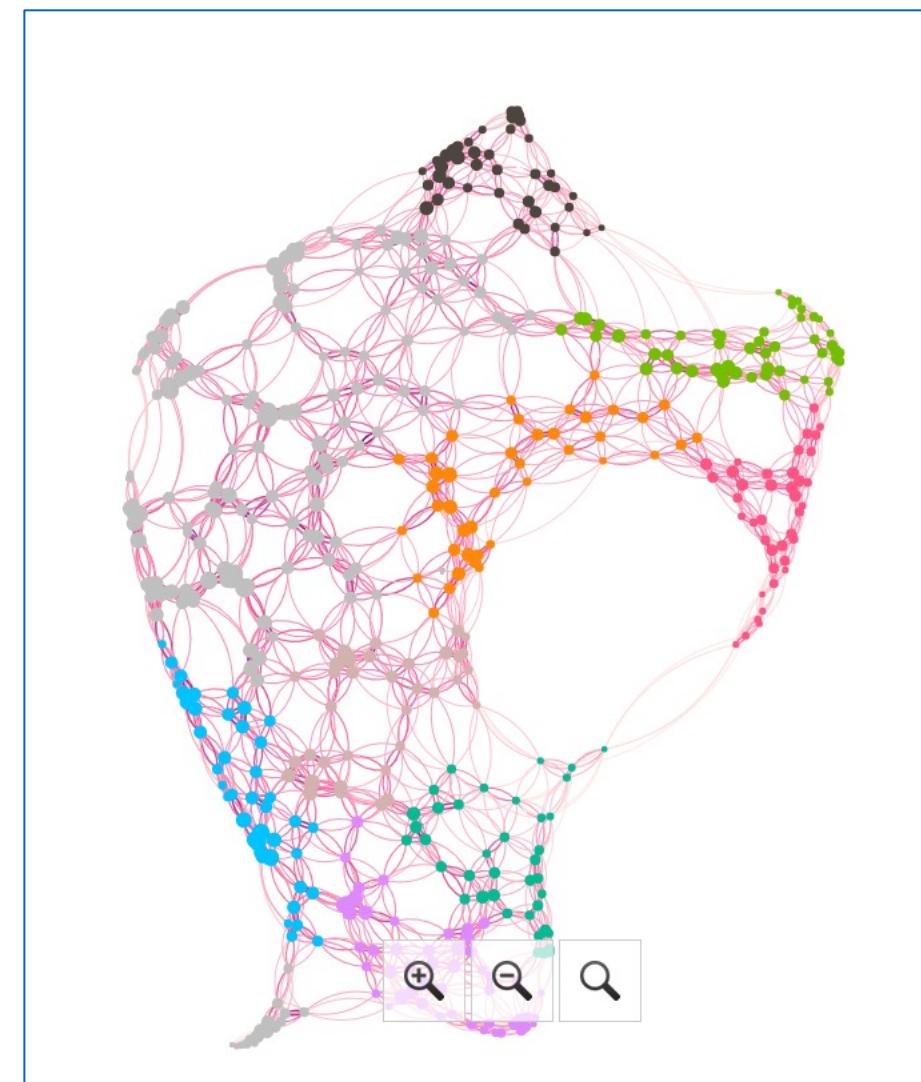
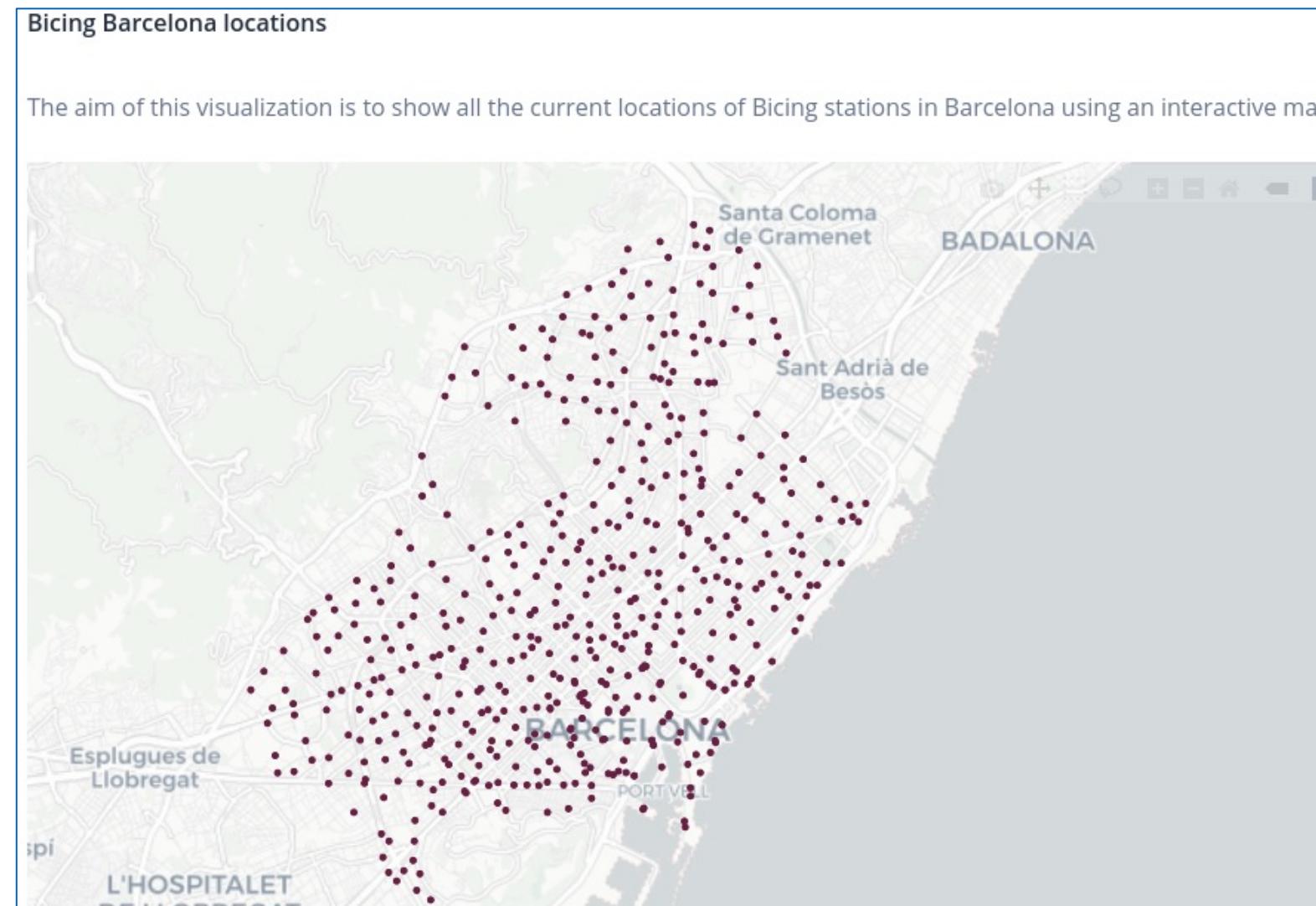
Figura 38. Diagrama de barres de la població estimada per estat i nombre d'incidents per 1M d'habitants<sup>24</sup>

# WeBike



## Top 10 most used stations

ADDRESS	TOTAL SLOTS	AVAILABLE SLOTS	USAGE
Av. Diagonal, 668	31	0	100.0%
Av. J. V. Foix, 63	34	0	100.0%
Av. Diagonal, 664	32	1	96.88%
Gran Via De Les Corts Catalanes, 375-385	44	3	93.18%
C/ Rocafort, 72-74	38	3	92.11%
Pg. Vall D'Hebrón, 168	50	6	88.0%
Gran Via De Les Corts Catalanes, 1157 (D)	52	7	86.54%
Rambla De Guipúzcoa, 43/Fluvia	42	6	85.71%
Av. Paral·Lel, 146 Bis	39	6	84.62%
Gran Via Corts Catalanes, 760	44	13	70.45%



# Procedimiento

- Brainstorming, inspeccionar datasets y elegir el dataset principal de trabajo. Repartir tareas
- Exploración visual y estadística. Buscar patrones, outliers y correlaciones -> Formular hipótesis
- Análisis avanzado, buscar fuentes complementarias de datos si es necesario
- Estructura de la historia que se quiere explicar:
  - Definir gráficas y formato (web, poster, reporte, etc.)
- Diseño final

- Se puede hacer en R o Tableau.
- Otras herramientas permitidas, pero sin soporte técnico.

SESSIÓ	Dia	Hora	Aula
1	19 Abril	Divendres, 13:00 a 15:00	Q5/1012
2	10 Maig	Divendres, 13:00 a 15:00	Q5/1012
3	17 Maig	Divendres, 13:00 a 15:00	Q5/1012
4	24 Maig	Divendres, 13:00 a 15:00	Q5/1012
5 (Examen Individual)	28 Maig	Dimarts, 15:00 a 17:00	Q5/1012
6 (Presentació Projecte)	31 Maig	Divendres, 13:00 a 15:00	Sala de Graus

# Hoy

- Explorar datos y definir tema
- Examinar data sets y plantear hipótesis posibles.
- Definir tareas a realizar para la siguiente sesión y asignar a cada miembro del grupo.
- **Subir al Campus Virtual un acta de la sesión**

# !Manos a la obra!

- **Hacer una visualización** con datos de elección libre.
- La visualización debe servir a un propósito → Hipótesis
  - Puede guiarnos hacia la respuesta, o
  - Mostrarnos la pregunta visualmente para que el usuario busque las respuestas
- Compuesta de una o más gráficas
- Con todos los medios necesarios para ser comprensible de manera autónoma:
  - Leyendas, escalas, anotaciones
  - Texto de soporte. Etc.
- Soporte de presentación de elección libre
  - **Infografía:** Poster con una o varias gráficas estáticas + texto
  - **Visualización interactiva:** Dashboard o visualización multi-panel
  - **Scrollitelling:** Narración interactiva basada en browser de gráficas + texto
  - **Data storytelling:** Presentación basada en gráficas
  - ...