

COVID-19 Spain

Storytelling interactiu sobre
dades del Coronavirus a Espanya

Postgrau:

'Presentació Efectiva de Dades'

Maig 2020 · Fundació UdG

Autor: Xavier Bolló

Tutor: Ferran Morales

Índex

Punt de partida	3
Pàgina web: covid19spain.org	3
Especificacions tècniques i oportunitats	3
Dades	3
Objectius	4
Finalitat	4
Públic objectiu	4
Mitjà de suport	4
Execució i resultat	5
Planificació	5
Recerca	5
Conceptes	5
Execució	8
Eines utilitzades	8
Processament de les dades	8
Estructura del contingut	9
Resultat final	11
Milliores pendents d'implementar	11
Conclusions i agraïments	12

Punt de partida

Pàgina web: covid19spain.org

Aquest projecte és una extensió del projecte digital covid19spain.org, una iniciativa d'Ignasi Alcalde —Director del curs de Postgrau— i en la que he col·laborat de manera activa fent suport a les tasques:

- Gestió de la creació del site.
- Disseny de la infraestructura de dades amb 2 objectius:
 1. Facilitar els fluxes de treball als responsables de la visualització de dades que s'adapti les eines que ja utilitzen: Google Sheets / Infogram
 2. Automatització en l'actualització de les dades visualitzades en temps real directament des de les fonts originals

Especificacions tècniques i oportunitats

Pàgina web creada mitjançant el gestor de continguts Wordpress.

Possibilitats àmplies per integrar-hi continguts tant estàtics com interactius oportunitat per practicar l'eficàcia de visualitzacions per a diferents tamany de pantalla.

Tanmateix la solució actual no exclou la possibilitat d'explorar i aprofitar el potencial de les noves tecnologies i *frameworks* web estàtics no basats en Wordpress — models tecnològics també conegut com JAMStack—i que actualment tenen un futur prometedori en UX/UI.

Els *frameworks* web estàtics tenen la particularitat de poder emular moltes funcions pròpies d'un servidor web —per exemple mostrar diferent contingut en funció d'un usuari— però que s'executen en el navegador client de l'usuari. Els avantatges més destacats són:

- Sense cost de hosting
- Millora substancial en la velocitat de càrrega
- 'Code Agnostic', això és, permeten incorporar-se a qualsevol tipus de tecnologia de gestió (php, python)
- Gran oferta gratuïta de publicació del contingut i redirecció de dominis (Netlify, Github, Vercel...)

Els principals *frameworks* compatibles que s'exploraran en aquest projecte són entre altres:

- Gatsby.js
- Vue.js
- Hugo
- Gridsome

Dades

Les dades utilitzades sobre el coronavirus a Espanya s'han obtingut a través del repositori públic de [Datadista](https://datadista.com).

Objectius

Finalitat

La finalitat d'aquest projecte és executar en valor els coneixements obtinguts al llarg del curs de postgrau de 'Presentació Efectiva de Dades' aplicats a enriquir la lectura que ofereixen les dades disponibles sobre el coronavirus 2020.

Tanmateix influirà el fet que personalment no tinc un *background* tècnic en quan a programació web però que alhora és una àrea on hi identifico un gran potencial per a la difusió visual d'informació de manera interactiva en diferents dispositius i mitjans.

Per aquest motiu he ponderat de manera rellevant que el projecte pogués tenir recorregut en aquesta direcció, no obstant sóc conscient que això implicarà importants limitacions creatives de concepte per la manca de expertesa en l'àrea.

tècniques, he convingut que encaixa millor un perfil de lector que ja té cert interès en la visualització de dades i que aprecia un enfocament diferent de dades.

Per afrontar la complexitat que pot derivar-se de visualitzacions no habituals ens acompanyarem amb el concepte *scrollytelling*, fent-ho així comprensible al lector general.

Mitjà de suport

Com s'ha comentat en el 'punt de partida' d'aquest document, es planteja un entorn web on exposar-hi el contingut de les visualitzacions i missatges explicatius.

A més del potencial interactiu que pugui oferir l'entorn digital, aquest també facilita la pròpia difusió del contingut a les xarxes socials com Instagram, Twitter, Facebook...

Públic objectiu

Inicialment plantejava el projecte com una alternativa visual d'impacte al públic general.

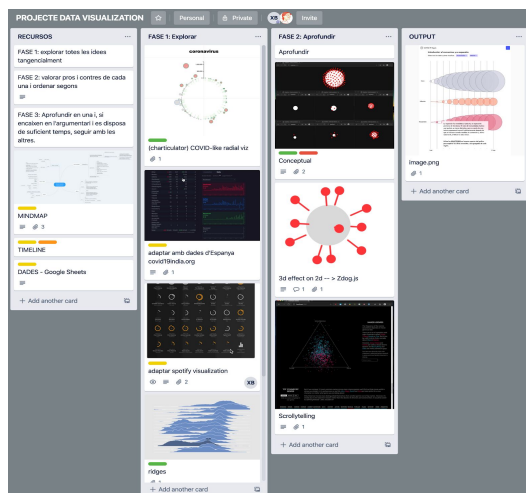
No obstant la il·lustració de dades sota la temàtica del coronavirus ha tingut ja molta difusió amb l'ús de visualitzacions estàndards, fet que a mi em suposaria més dificultats per obtenir resultats que fossin diferencials i rellevants.

Per aquest motiu i també per la pròpia voluntat de posar en pràctica diferents

Execució i resultat

Planificació

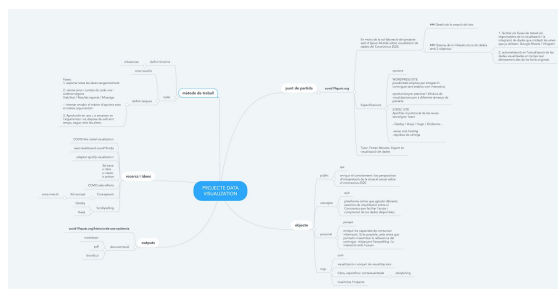
La planificació s'ha plantejat en 2 fases que consten una en **recerca** i l'altra en **aprofundir** en una de les opcions que es deriven de la recerca. El seguiment s'ha fet mitjançant el sistema Kanban que ofereix [Trello](#).



Taulell de Trello creat per al seguiment del projecte

També es va realitzar un mapa mental per estructurar els objectius i els resultats esperats. Aquest es pot consultar al següent enllaç:

<https://www.mindmeister.com/1496303717?t=SgFf4ZlouJ>



Recerca

En aquest projecte s'ha destinat especial esforç en la recerca per tal que a l'hora d'executar el projecte es pogués realitzar el recorregut desitjat sense incórrer en dificultats particularment tècniques que puguin afectar en detriment dels objectius del projecte.

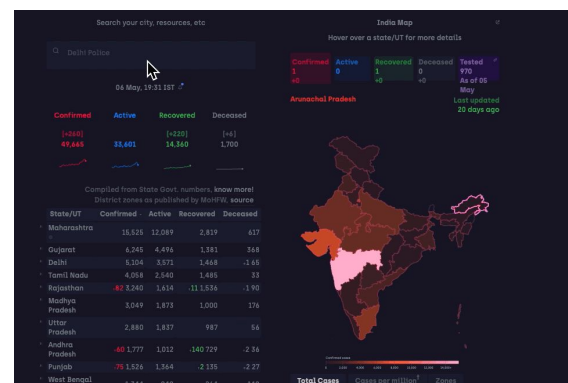
Conceptes

S'ha recercat sobre quins conceptes podrien respondre millor per visualitzar les dades principals sobre el coronavirus.

Alguns dels conceptes que s'ha tingut en compte són:

Dashboards:

Exploració de les dades mitjançant l'ús de taulells interactius



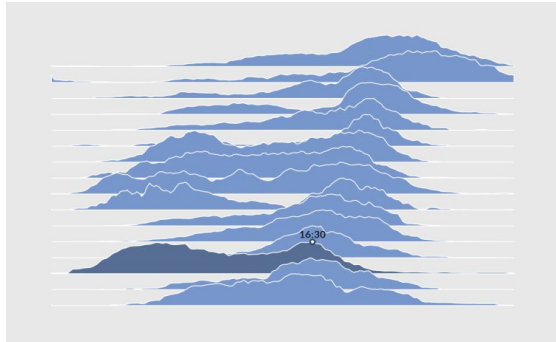
Dashboard interactiu - covid19india.org

Malgrat la riquesa de recursos disponibles per executar un dashboard, mancava dels conceptes de storytelling i no resultava fàcil obtenir un resultat diferencial.

'Ridges plot':

Utilitzar gràfics de densitat utilitzant l'eix Z per desglossar els països i així poder visualitzar l'efecte onada de l'expansió del

virus i la seva incidència. Per tal que resultés efectiva era especialment important prestar atenció al criteri d'ordenació dels elements en l'eix de coordenades Z



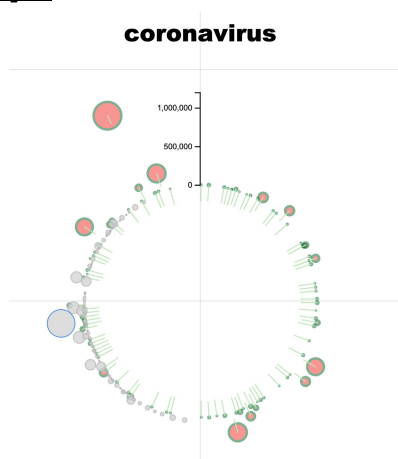
Exemple de ridge plot fet en R

Aquesta idea es va desestimar perquè oferia poc joc en la incorporació d'altres variables i era limitada en recorregut.

Associació visual sobre el concepte:

Representar les dades mitjançant elements visuals que remetien directament a la identificació de la temàtica

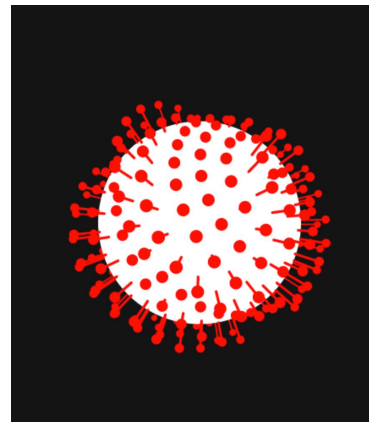
Prototip 1:



Prototip realitzat amb charticulator.com

Experimentar amb l'eina Charticulator va resultar molt interessant, no obstant es va desestimar per no ser prou versàtil en termes d'interactivitat

Prototip 2:



Prototip realitzat amb javascript d3

Paràmetres:

- Cercle blanc: població
- Punts vermells: infectats
- Distància del punt vermell al centre: mortalitat

Parametrització del prototip 2:

	Cap infectat	Pocs infectats	Molts infectats
Poca mortalitat			
Molta mortalitat			

Simulació d'escenaris per a la parametrització de les variables en el Prototip 2

Posada en productiu de la versió genèrica del prototip 2: <https://coviz.now.sh/>

Aquesta idea resultava alentadora en termes d'impacte alhora que la comprensió i explicació del mateix eren evidents *per se*. El motiu pel qual es va descartar va ser que en testejar-ne la reproductibilitat del codi es mostrava inestable en diferents suports tals com versions antigues del navegador

Google Chrome o dispositius mòbils, afectant així el potencial de difusió que busquem en el suport digital.

Es va treballar també un prototip paral·lel en aquesta mateixa línia utilitzant la llibreria [Zdog.js](#) especialitzada en creació d'objectes 3D amb renderitzat 2D, aquesta opció era més robusta en dispositius, però s'ha desestimat per incompatibilitat en els temps d'entrega del projecte.

Geodata:

Exploració de les dades en context geolocalitzat.



Simulació de les fases de desescalada segons municipis i el llindar de població de 5000 habitants feta amb Mapbox.

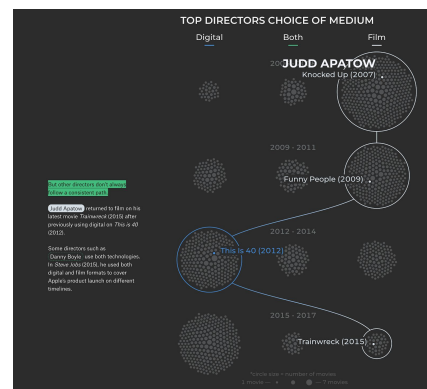


[Interessant](#) aplicació creada per a la visualització animada i interactiva de la incidència del COVID-19 per països

El concepte de geodata m'interessava molt personalment ja que opino que ofereix moltes possibilitats. No obstant he optat en no prioritzar-ho enguany, ja que no hi tinc cap mena d'experiència i que qualsevol obstacle en l'execució podria resultar en no poder culminar-ne l'execució sense opció a aprofitar la feina feta de cares a altres alternatives.

Scrollytelling:

Concepte web amb el qual es genera una relació directa d'interacció entre el contingut i l'usuari mitjançant les variables de desplaçament per la pàgina ('scrolling') i curs d'esdeveniments de la història que tracem amb dades.



[The Pudding](#) és un mitjà digital de visual storytelling amb execucions del concepte scrollytelling molt interessants

Aquest és el concepte pel qual he optat construir el projecte. A la xarxa hi ha una àmplia riquesa de recursos pel que qualsevol obstacle podria tenir fàcil solució o reenfoc. També ofereix versatilitat per adaptar el missatge i la història proactivament amb les limitacions que pugui anar trobant.

Execució

Eines utilitzades

- **Framework web:** Vue.js
- **Llibreria de gràfics:** d3.js
- **Scrollytelling:** Scrollama.js
- **Treball de dades:** R i Google Sheets
- **IDE de programació:** VSCode
- **Hosting estàtic:** Vercel (aka Zeit)

Per construir un entorn flexible on integrar el contingut de l'*scrollytelling* no vaig ser capaç d'identificar cap eina per a Wordpress prou flexible amb l'ambició del projecte, fet que hauria m'hauria permès dedicar més temps a visualitzacions creatives.

La llibreria **Scrollama.js** era una de les eines més complertes per coordinar interacció i continguts en la línia que esperava, però era necessari identificar on integrar-ne les funcionalitats.

Inicialment vaig començar l'execució amb el generador de pàgines estàtiques **Gatsby.js** ja que l'havia explorat tangencialment altres vegades. No obstant Gatsby encaixa més en projectes tipus 'blog' o multi-pàgina, l'estructura de contingut resulta molt versàtil però alhora molt fragmentada pel que resultava confús i poc còmode estar treballant en diversos fitxers l'estructura d'una sola pàgina, pel que vaig abandonar-ne el desenvolupament.

Finalment després de testejar diferents alternatives vaig decidir optar per **Vue.js**, potent i amb una estructura interna més unificada que s'adaptava millor als meus objectius.

Processament de les dades

Part de les dades les havia ja pre-processat a **Google Sheets** per al projecte covid19spain.org, no obstant vaig processar-les totes en R per tenir-ho integrat en el mateix **IDE** amb el que treballava i alhora alguns procediments son bastant més àgils.

Per exemple vaig requerir sovint la parametrització de diverses columnes com una sola i en la que el llenguatge R va estalviar-me molt treball manual:

1. Exemple de taula a convertir

fecha	variable1	variable2
1 abr	1	6
2 abr	2	5

Exemple fitxer *dades_csv*

2. Exemple de codi R:

```
data <- pivot_wider(dades_csv,
                    id_cols='Semana',
                    names_from='Fecha',
                    values_from = 'Valor',
                    values_fill = 0 )
```

3. Taula convertida

fecha	variables	valor
1 abr	variable1	1
2 abr	variable2	2
1 abr	variable1	6
2 abr	variable2	5

Estructura del contingut

Estructura del discurs basic de l'storytelling:

- Introducció
- Cos
- Desenllaç

Tenint en compte aquesta estructura definirem el següent recorregut en el discurs:

A. Introducció

1. Què és el coronavirus

B. Cos

2. Anatomia de les dades
3. Incidència demogràfica a Espanya
4. Interpretació i particularitats

C. Desenllaç

5. Tenència, què podem esperar?

Dades:

Per traçar la història centrem-nos en les principals xifres disponibles:

- Casos de contagi
- Número de morts
- Número de recuperats

Aquestes les hem aconseguit categoritzades per:

- Comunitat Autònoma
- Demografia: sexe / edat

Construcció de l'argument:

A. INTRODUCCIÓ

Per presentar el coronavirus no ho hem complicat massa, es dona per fet que és un tema àmpliament conegut arreu. Per evitar cap mena ambigüitat en la interpretació oferim directament la definició que en dona la OMS.

B. COS

Fig 1- Procurem transmetre la naturalesa exponencial del contagi del coronavirus mitjançant una seqüència temporal que permeti al lector identificar ràpidament l'explosivitat relativa de les xifres.

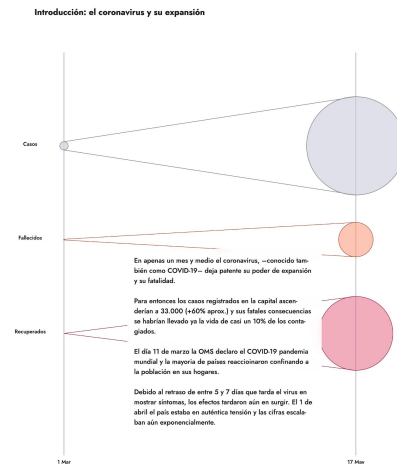


Fig 1

Fig 2- Ara que el lector ja coneix el context de la història procurem retenir-lo oferint-li la informació més detallada i fent-lo protagonista deixant-lo consultar dinàmicament les dades posant en valor la seva pròpia curiositat.

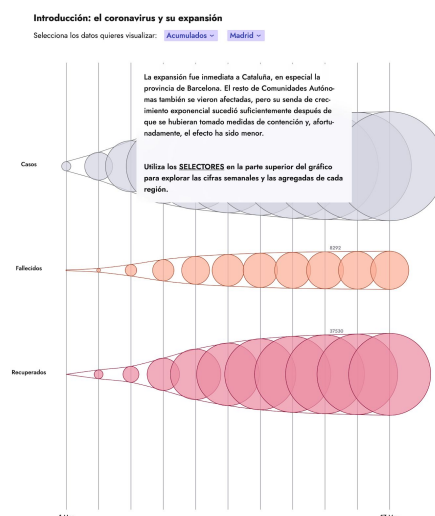


Fig 2

*Nota: la darrera versió publicada ja assigna correctament el color vermell a 'fallecidos'.

Fig 3- També procurem il·lustrar amb senzillesa la naturalesa de cada variable, per tal de consolidar la comprensió de les dades en l'ideari del lector per posteriorment poder incorporar-hi major complexitat.

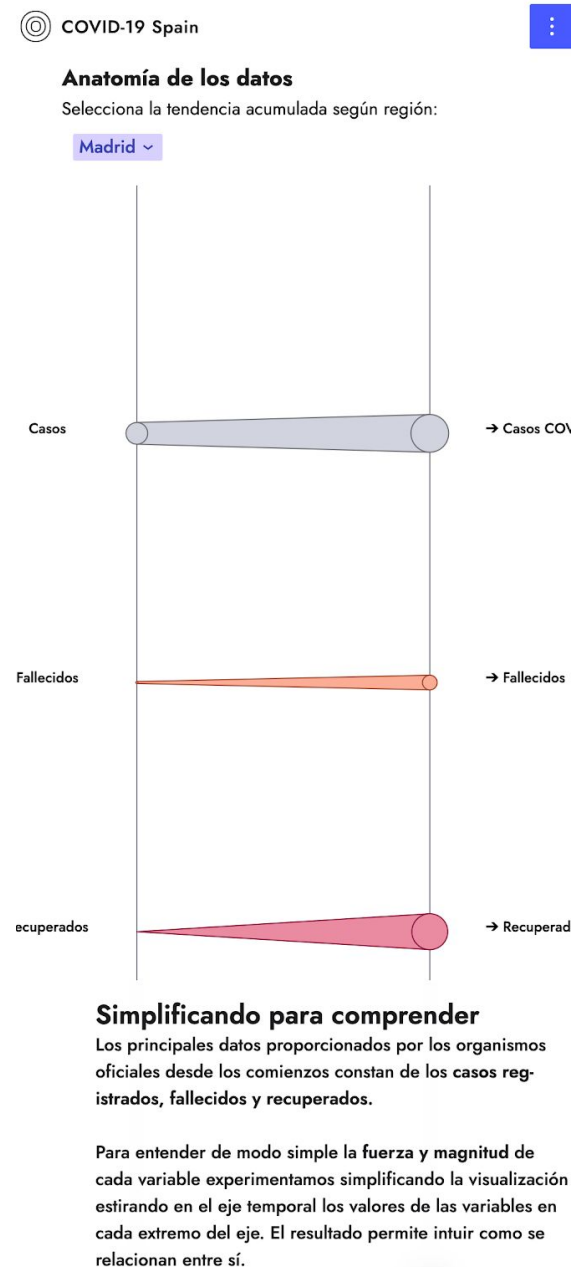


Fig 3

Fig 4- Per tal que l'usuari pugui percebre major dimensionalitat de les dades l'introduïm en una visualització de part sobre el tot sobre la incidència demogràfica.

Procurem alleujar-ne el pes visual fent només visibles les xifres passant el cursor per sobre cada gràfic.

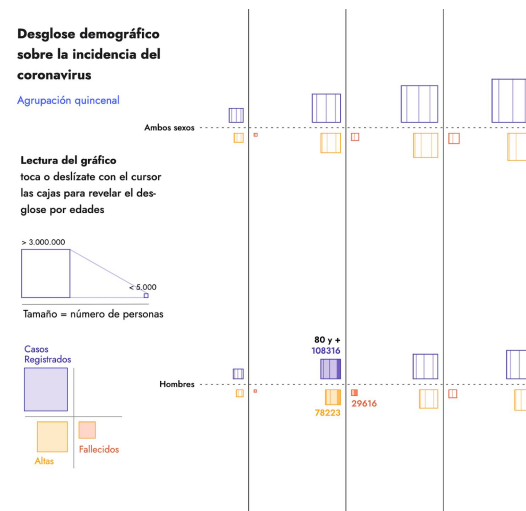


Fig 4

Fig 5- Recuperem un nivell de comunicació visual més senzill per incloure un curiositat que enriqueix al lector facilitant-li la comprensió de com interpretar algunes dades i les decisions institucionals que en deriven.

La variable de *recuperats* ha entrat en desús les darreres setmanes, ja que aquesta en desestimar-se cap relació amb el concepte d' 'immunitat de grup' s'interpreta que no aporta informació sobre l'evolució de la situació. Aquest factor és interessant en la interpretació pel que l'hem incorporat en el punt (4) per enriqueix els matissos de la història i de la naturalesa de les interpretant-ho amb motiu pel qual Regne Unit ha rectificat la polèmica posició inicial particularment laxa en quan a restriccions a la mobilitat de persones.

La responsabilidad del superviviente

■ Casos ■ Recuperados ■ Fallecidos

La noticia sobre el señor japonés que tras haberse recuperado se infectó de nuevo plantea un nuevo marco de análisis en el que reducir la tasa de contagios se convierte en una tarea mucho más compleja con y cualquier resultado dependerá en gran manera de la capacidad de gestión colectiva y —desgraciadamente— del número de fallecidos.

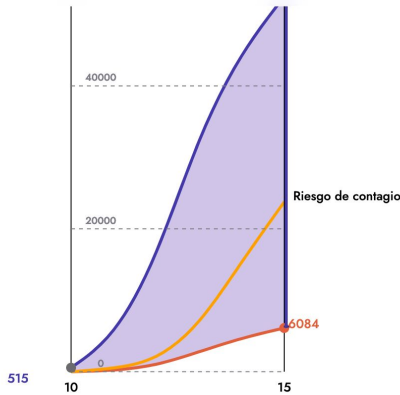


Fig 5

C. DESENLLAÇ

Fig 6 - Concluim l'exercici posant èmfasi en la tendència que podem esperar. Aquesta afortunadament es mostra positiva, deixant una bona sensació al lector en el pas per la nostra història.

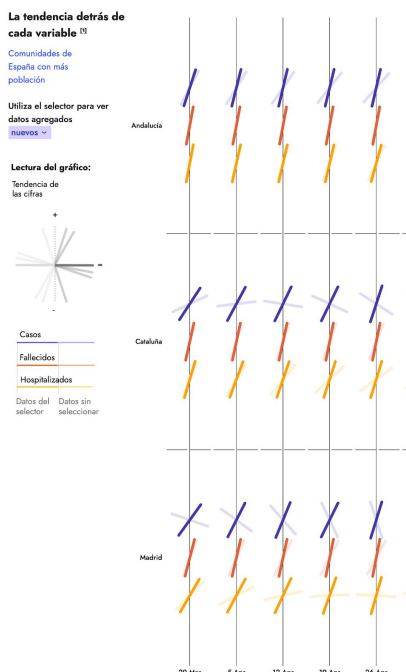


Fig 6

Resultat final

El resultat final pot consultar-se a la següent adreça web:

<https://covisualitza.now.sh/>

Millors pendents d'implementar

Al moment d'entrega del projecte hi ha hagut una sèrie de tasques que en no ser crítiques no s'han prioritzat. No obstant seria interessant implementar-les i es contemplen en un futur.

- Afegir anchor links i millorar comprensió de l'índex de continguts
- Gràfic 1: incorporar més etiquetes per a l'eix x
- Gràfic 2, 4: Revisar valors en l'eix x
- Gràfic 1: destacar millor xifres dels tooltips.
- Donar format al blocs de scroll per tal que si es menciona una dada o variable el color coincideixi amb el del component visual.
- Gràfic 2: refer els 'strings' de cada variable per tal que l'etiqueta categòrica per edats incorpori 'años'.
- Gràfic 2: intercanviar assignació visual entre gènere - variable
- Gràfics verticals: suavitzar tonalitats, espaiar grups.
- Revisió de textos: typos i simplificació.
- Estil de les scroll boxes més ressaltat (alpha background / border)
- Incloure referències.
- Revisar footer

Aprentatge

Aquest projecte ha sigut una gran experiència personal. Doncs malgrat dificultats que he anat trobant i malgrat haver de redirigir el focus del projecte en diverses ocasions, estic satisfet del resultat final i deixa al meu abast una experiència i uns coneixements que superen de molt les expectatives que inicialment tenia.

En termes conceptuals reconec que no tenia clar al principi si seria capaç de traçar un discurs coherent i amb caràcter sobre una temàtica que ha estat sobre-explotada als mitjans darrerament. No obstant les directrius i recursos que se'ns han facilitat durant el curs de postgrau han fet que el discurs s'anés dibuixant sol a mida que hi treballava i m'han permès gaudir molt en l'exploració de l'execució.

També estic content d'haver pogut redirigir dinàmicament el projecte de manera prou versàtil sense que això en condemnés l'execució, apostar pel concepte d'storytelling crec que ha sigut una bona decisió i que, malgrat no he pogut executar altres idees interessants, aquesta en particular em dóna una base d'experiència que pot ser-me molt útil en un futur.

En termes tècnics reconec que els reptes que m'he plantejat són ambiciosos pel meu perfil sense recorregut en programació i que ha escapat la meua capacitat creativa en l'execució alhora que en ocasions m'ha suposat dificultats, inclús, per executar elements senzills. De totes maneres prenc en positiu l'experiència ja que per contra i hi veig un gran potencial per a la visualització de dades i m'ha permès explorar i executar funcionalitats molt interessants que mai hagués imaginar pel meu compte.

Personalment sento que m'hagués agradat poder disposar de més temps i experiència per poder incorporar al projecte més elements de pensament visual i de geodata que considero que encaixaven molt bé en la temàtica però que engrandien massa l'scope del projecte fora de l'abast dels terminis d'entrega.

Agraïments

És precís mencionar el suport del meu tutor Ferran Morales. Per mi és un honor haver pogut comptar amb el mentoratge d'un professional amb una trajectòria com la seva i a qui agraeixo que hagi estat a la meua disposició en tot moment i que m'ha proporcionat feedback constructiu i tangible que han enriquit qualitativament el resultat final d'aquest projecte.

Tanmateix agraeixo a Ignasi Alcalde, per animar-me a col·laborar en la iniciativa del projecte covid19spain.org que ha definit la temàtica del treball aquí exposat. També per haver compartit amb mi els seus coneixements i reputada experiència amb valuosos consells per portar a terme aquest projecte.

Vull també agrair a Javi Cantón i Marta Herrero, així com la resta de companys del curs, que malgrat no han col·laborat directament en aquest projecte, haver compartit la seva passió per la seva professió i els coneixements i experiència amb nosaltres.

És gràcies a l'aportació de tots ells que aquest projecte tingui enguany aquest resultat.