
Introducció a la visualització de la informació

PID_00249215

Ignasi Alcalde
Julià Minguillón

Índex

1. La societat de la informació	5
2. El disseny de la informació	7
2.1. Infografia <i>versus</i> visualització	8
2.1.1. Alguns exemples i disciplines implicades	11
3. Una mica d'història	19
4. Avantatges de la visualització	24
5. Estructura d'una infografia	26
6. Tipus d'infografies	28
6.1. Segons la seva composició	28
6.2. Segons la seva finalitat	29
7. Categories de la visualització de la informació	30
8. Eines per a la visualització	36
Resum	38
Bibliografia	39

1. La societat de la informació

En la societat de la informació i en xarxa en la qual vivim avui, la idea que «la informació és poder» està molt estesa, però és parcialment errònia, ja que seria més adequat l'aforisme del filòsof americà David Lewis «**el coneixement és poder; la informació, no**». És factible preguntar-se llavors per què? En part, és senzill d'entendre. Vivim en una societat xarxa en què la Unió Internacional de Telecomunicacions preveu que, a finals del 2015, un total de 3.000 milions de persones estiguin connectades a internet. Això representa que pràcticament el 40% de la població mundial està connectat a internet, la qual cosa implica fluxos ingents de dades i informació (entenent informació com a dades en un context que els aporta significat). En el dia a dia actual, és possible comprovar que els missatges electrònics tenen un creixement exponencial i que les xarxes socials han obert noves vies i potenciat nous mitjans de comunicació: fotografies, vídeos, llibres electrònics, música, etc. El cost d'emmagatzemar i duplicar la informació tendeix a zero amb eines en el núvol, com per exemple Dropbox. El maquinari encarregat de donar suport a tota aquesta infraestructura tecnològica s'ha convertit en una *commodity*, gairebé transparent (en el sentit d'invisible), mentre que el valor s'afegeix amb les dades que es generen. Per tot això és fàcil comprovar que **la informació ens envolta, ens desborda**.

Estem contínuament, doncs, en un tsunami d'informació, la qual cosa es coneix com una sobrecàrrega d'informació o **infoxicació** (combinació d'*informació* i *intoxicació*), que té lloc quan un s'enfronta a més informació de la que és capaç de processar, cosa que pot portar, per exemple, a posposar algunes decisions per voler-ho analitzar tot detalladament o a prendre decisions equivocades sota la pressió de l'estrès informatiu.

La neurologia ja ha demostrat que l'excés d'informació, a més d'alterar la capacitat per a prendre decisions, també pot produir ansietat (síndrome de fatiga per informació) i que els seus efectes psicològics també influeixen sobre el nostre cos. Encara que sembli el contrari, aquesta manera de consumir informació, impulsiva, massiva, esbiaixada i en temps real, comporta noves formes d'ignorància per falta de concreció.

Fent servir com a exemple només les dades numèriques, amb les quals som bombardejats des dels nostres lectors de notícies, piulades o *feeds*, és fàcil adonar-se ràpidament que es rep una quantitat ingent d'informació constantment. Una pantalla amb les dades de la borsa, els resultats de diferents esdeveniments esportius, les dades de la previsió meteorològica, etc. són exemples que inclouen una gran quantitat de dades que són consumides ràpidament. Queda demostrat, doncs, l'aclaparament produït per la sobrecàrrega d'informació o excés de dades. Però es pot absorbir i entendre el que aquestes

signifiquen realment? Aquí és on rau el poder de la visualització d'informació (i de dades), una disciplina realment fascinant l'interès de la qual no ha fet més que començar recentment, tot i que les seves arrels són profundes.

2. El disseny de la informació

Quan es busquen les paraules *infografia* o *visualització de dades* a Google Trends –eina que analitza les tendències de cerca dels internautes–, ràpidament és possible adonar-se del protagonisme que estan prenent aquests dos conceptes, especialment el primer des del 2010, aproximadament. Però què és realment la visualització de la informació?

Es podria definir d'entrada, de manera introductòria, com la presentació i representació de dades que exploten la capacitat de percepció visual humana amb la finalitat d'ampliar el coneixement.

Una visualització explica històries, simplifica, mesura, compara, explora, descobreix, etc. amb l'objectiu de transformar les dades en informació, la informació en coneixement i el coneixement en intel·ligència de negoci, per exemple. Permet obtenir coneixement d'unes dades sense haver de fer càlculs (tot i que en realitat s'està calculant «visualment»).

Les paraules *visualització de la informació* com a terme general parla de la presentació visual del coneixement en qualsevol àrea. En general, es refereix a qualsevol gràfic que mostri i expliqui la informació, siguin dades o text. La visualització en si mateixa consisteix a transformar aquesta informació en imatges que facilitin l'extracció del seu significat. De vegades, aquesta informació és quantitativa i la visualització creada, sovint també interactiva, pot ser anomenada «de dades». En aquest cas, l'objectiu del gràfic és permetre identificar patrons i tendències que serien invisibles si aquestes mateixes dades es presentessin en una taula numèrica, per exemple. Altres vegades la informació és una barreja d'elements qualitatiu, quantitatiu i fins i tot narratiu (per exemple, la descripció d'un procés o esdeveniment); en aquest cas la visualització se sol anomenar infografia.

Com a base, aquesta primera distinció entre visualització i infografia facilita un punt de partida, que pot ser complementat per la visió que en tenen alguns dels experts d'aquestes disciplines. Per exemple, Alberto Cairo, autor del llibre *El arte funcional: Infografía y visualización de información*, indica:

«Visualització és la disciplina que consisteix a transformar dades en informació semàntica, o a crear eines perquè un usuari completi per si sol aquest procés, per mitjà d'una sintaxi de fronteres imprecises i en constant evolució basada en la conjunció de signes de naturalesa icònica (figuratiu) amb altres de naturalesa arbitrària i abstracta (no figuratiu). La visualització de dades i la infografia són formes no només de mostrar unes dades, sinó d'explorar-les, analitzar-les i contextualitzar-les. De raonar-hi.»

A. Cairo (2011). *El arte funcional: Infografía y visualización de información*.

És comprensible, doncs, que en el context del disseny de la informació les dades han d'explicar una història i han de tenir una funcionalitat. Vitaly Friedman, editor de *Smashing Magazine*, n'amplia la definició:

«L'objectiu principal de la visualització de dades és transmetre informació clara i eficaç amb mitjans gràfics. Això no significa que la visualització de dades hagi de ser avorrida per a ser funcional o extremament sofisticada per a ser bella. Per a expressar idees de manera efectiva, tant la dimensió estètica com la funcional han d'anar de la mà, i han de proporcionar determinats punts d'entrada a un conjunt de dades complex per a comunicar-ne els aspectes clau d'una manera més intuïtiva.»

V. Friedman (2008). *Smashing Magazine*.

Per tant, s'han de dominar principalment dues competències bàsiques amb l'objectiu de dissenyar una visualització correctament: la capacitat de síntesi i la capacitat estètica, i totes dues han d'anar combinades correctament. Fernanda Viégas i Martin Wattenberg, considerats els líders del **Big Picture Data Visualization Group de Google** i impulsors de moltes iniciatives basades en la visualització de dades, proposen un punt de matís més:

«La visualització ideal no només ha de comunicar amb claredat, sinó que també ha d'estimular la participació i l'atenció de l'espectador.»

F. Viégas; M. Wattenberg (2006). *Communication-minded Visualization*.

En altres paraules, per a tenir èxit en el disseny d'una visualització no solament cal tenir en compte el suport o el format estètic; la clau és estimular la participació de l'espectador mentre examina la informació presentada.

Extraient les idees clau de cadascuna d'aquestes definicions, es construeix una proposició general, entenent la visualització de la informació com un àmbit multidisciplinari que representa les dades transformant-les en informació semàntica amb mitjans gràfics, combinant la seva funcionalitat i estètica amb simplicitat i estimulants la participació dels usuaris. A més, està composta de dos blocs principals: contingut (informació) i formes i mètodes de presentació (disseny).

La visualització de dades permet **explicar històries, simplificar, mesurar, comparar, explorar, descobrir**, etc., i les dades han de ser compreses de manera efectiva, sense ambigüitats ni manipulacions. L'objectiu de tota bona visualització ha de ser centrar l'atenció del destinatari de la informació en el rellevant i important, de manera que el porti al que en anglès s'anomena *insight* (aquest nivell de comprensió al qual s'arriba quan un exclama «ja ho veig!»).

2.1. Infografia versus visualització

En molts casos, els termes *infografia* i *visualització* de dades es consideren sinònims, però en el món del disseny de la informació signifiquen coses diferents. Actualment, la paraula *infografia* és utilitzada indistintament per diver-

sos col·lectius per a expressar coses diferents. A més, és un recurs gràfic molt emprat en l'àmbit periodístic i educatiu com a forma de presentar informació. Però què és una infografia realment?

Es pot definir **infografia** com una representació més visual que els mateixos textos, en la qual intervenen descripcions, narracions o interpretacions, presentades de manera gràfica normalment figurativa, que poden coincidir o no amb grafismes abstractes o sons.

La infografia neix com un mitjà per a transmetre informació gràficament, que disposa d'un mètode per a **representar la informació icònicament i textualment**, de manera que l'usuari la pugui comprendre sense dificultat, emprant normalment eines informàtiques. En el procés de creació d'una infografia es recull un fet complex i s'explica de manera senzilla perquè pugui ser interpretat amb un simple «cop d'ull».

Les infografies permeten narrar històries, explicar esdeveniments, descriure situacions, exposar processos, etc. Cal destacar, no obstant això, que no qualsevol representació gràfica, amb més o menys text i més o menys atractiva, és una infografia. Perquè es pugui anomenar d'aquesta manera (o més ben dit, perquè sigui una bona infografia) ha de ser fàcil d'interpretar i, alhora, ha d'aportar una gran quantitat d'informació sobre un tema.

D'altra banda, una **visualització** de dades o informació disposa de representacions visuals de valors numèrics, generalment de manera més esquemàtica. Els quadres, les taules i els gràfics són visualitzacions de dades, ja que creen una imatge a partir d'un conjunt donat de dades.

En comparació amb una infografia, la visualització de dades, també anomenada visualització de la informació (ja que posa les dades en el seu context), és l'estudi de la representació visual de dades abstractes (i potser interactives) per a reforçar la cognició humana, que inclou tant dades numèriques i no numèriques com text o informació geogràfica, per exemple. Per tant, es pot deduir que no hi ha pràcticament gaires diferències substancials entre tots dos conceptes, ja que hi ha una naturalesa comuna entre una infografia i una visualització. Això ho explica molt bé Alberto Cairo en la seva obra *El arte funcional: Infografía y visualización de información*, al mateix temps que indica encertadament les subtils diferències conceptuals:

«Alguns especialistes marquen una frontera entre les dues disciplines basada en el fet que, suposadament, la infografia consisteix a presentar informació per mitjà de gràfics estadístics, mapes i esquemes (exposició), mentre que la visualització es basa en la creació d'eines visuals (estàtiques o interactives) que un públic pugui fer servir per a explorar, analitzar i estudiar conjunts complexos de dades. Però pertanyen a un mateix continu en el qual cadascuna ocupa extrems oposats d'una línia. Aquesta és paral·lela a una altra els límits de la qual són definits per les paraules *presentació* i *exploració*.»

A. Cairo (2011). *El arte funcional: Infografía y visualización de información*.

Així doncs, es pot establir que la **infografia** implica la **presentació d'informació** de manera generalment **estàtica**, mentre que la **visualització** de dades té com a eix principal el **component dinàmic** i la **possibilitat d'interacció** i exploració d'aquestes dades, aprofitant les capacitats tecnològiques actuals.

Tot i que no és sempre exactament així, es pot pensar que una infografia és una presentació més senzilla i intuïtiva que no pas una visualització de dades, amb un objectiu més concret, com presentar una idea. Tanmateix, també és possible «introduir» l'usuari en la infografia i permetre-li que en manipuli alguns aspectes.

La figura 1 és un exemple d'infografia creada per Anna Vital, que mostra l'edat en la qual diferents persones van assolir el seu èxit professional, amb l'objectiu de demostrar que mai no és realment tard per a començar. La versió en línia permet canviar la part inferior de la infografia si se selecciona alguna de les boletes que representen diferents franges d'edat. D'aquesta manera, l'usuari receptor de la infografia es pot comparar amb altres persones de la seva mateixa edat que han assolit els seus objectius i respondre la pregunta en qüestió que es planteja des de la seva perspectiva personal.

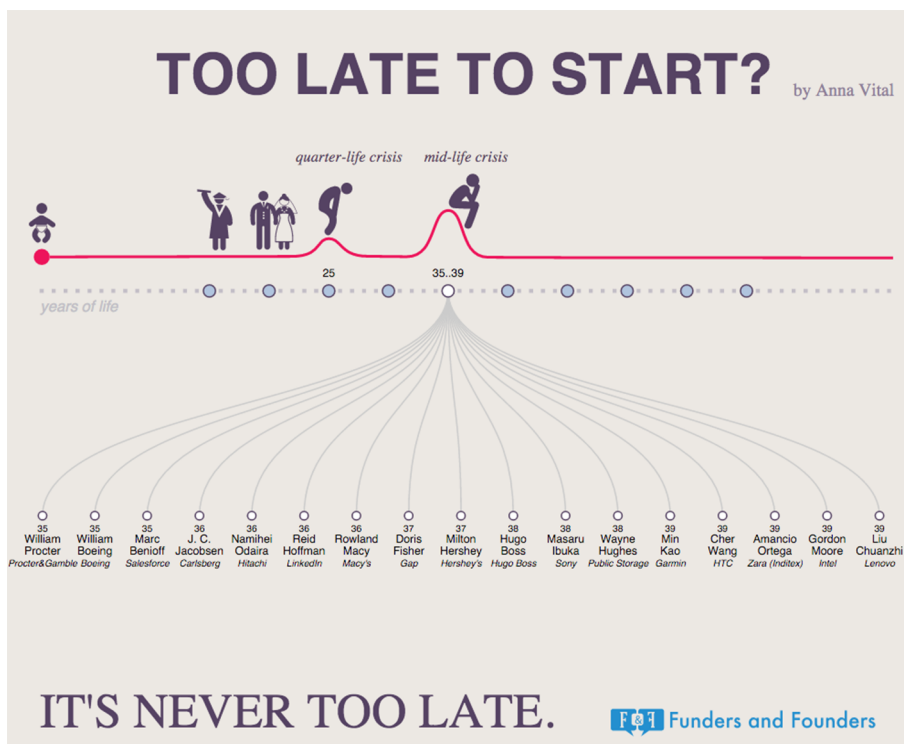


Figura 1. Infografia sobre l'edat en què arriba l'èxit professional
Font: <http://fundersandfounders.com/too-late-to-start-life-crisis/>

D'altra banda, la figura 2 mostra un exemple de visualització de dades que permet accedir conjuntament a 255 gràfics diferents, tots relacionats amb la creació o destrucció d'ocupació en un període de deu anys (2004-2014). A més, aquests gràfics es troben agrupats per temàtica, la qual cosa permet afegir dimensions extres a la visualització, de manera que se simplifica l'accés a un conjunt de dades complex.

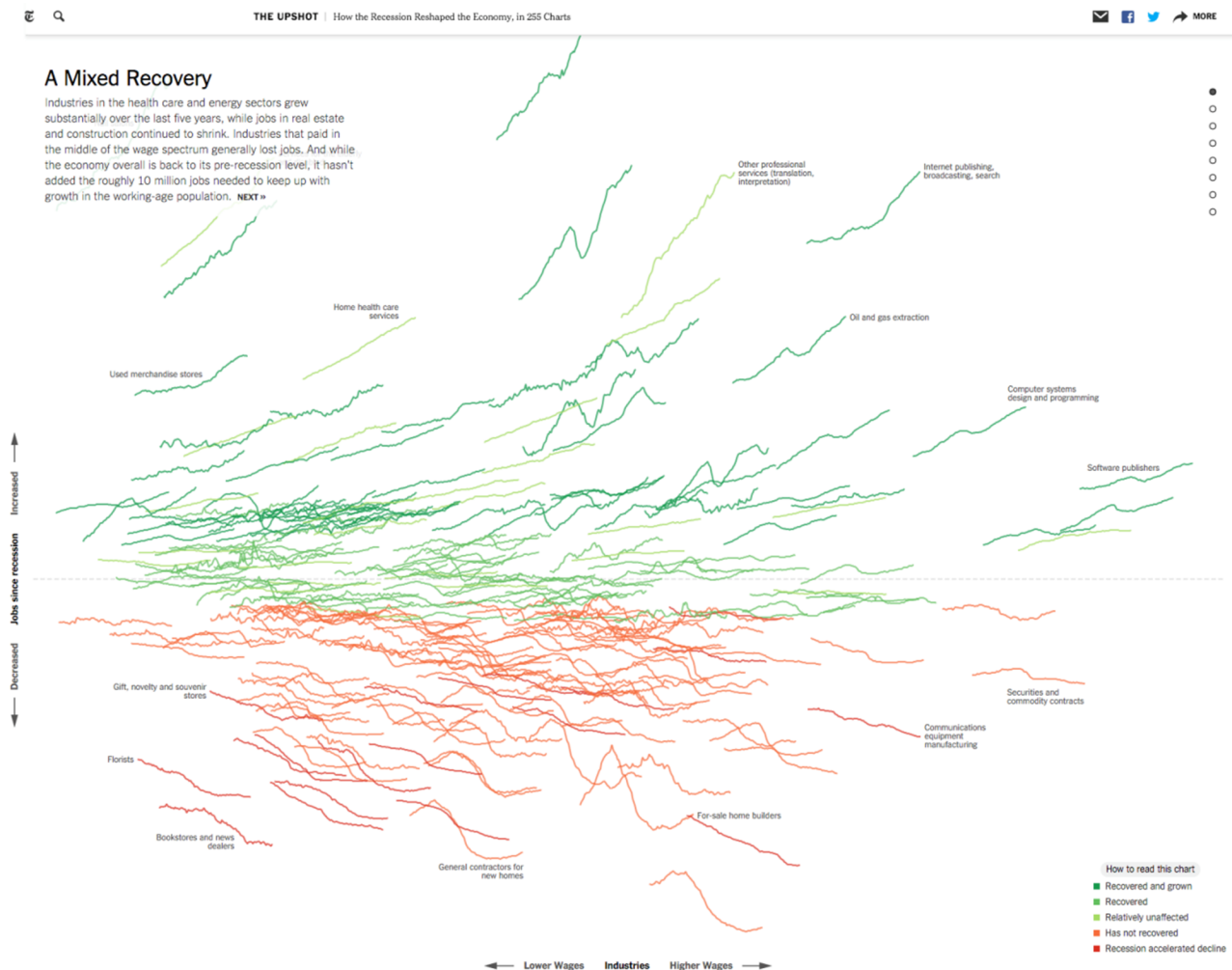


Figura 2. Visualització sobre la creació i pèrdua de llocs de treball
Font: <http://www.nytimes.com/interactive/2014/06/05/upshot/how-the-recession-reshaped-the-economy-in-255-charts.html>

En aquest cas, la diferència del volum de dades mostrat en la visualització i les diferències en la seva naturalesa, la complexitat de la interacció i també el seu objectiu (edat en la qual arriba l'èxit enfront de creació/destrucció d'ocupació en un període) fan que es pensi més en el primer exemple com una infografia, mentre que el segon exemple és més una visualització de dades, encara que la imatge estàtica de resum de la figura 2 també es podria entendre com una infografia que mostra que hi ha sectors que han creat ocupació mentre que d'altres n'han destruït, per exemple.

2.1.1. Alguns exemples i disciplines implicades

Una vegada establert el concepte bàsic de visualització d'informació, passem a observar alguns exemples en línia de com està present en diverses disciplines i àrees de coneixement.

a) Periodisme

La **Fundació Ciutadana Civio**, després d'una petició directa, va treballar amb una base de dades nacional d'incendis forestals, amb dades proporcionades pel sol·licitant, el Ministeri d'Agricultura, Alimentació i Medi Ambient. El projecte de visualització de dades és una selecció de nou històries destacades entre 2001 i 2013. En aquest projecte es poden explorar els incendis dinàmicament sobre un mapa i, alhora, és possible interactuar amb els focs que van consumir 100 hectàrees o més, seleccionar els que més interessin o localitzar els propers a un municipi concret. És una bona eina per a posar en context els incendis, i fins i tot permet baixar les dades per a anàlisis posteriors.

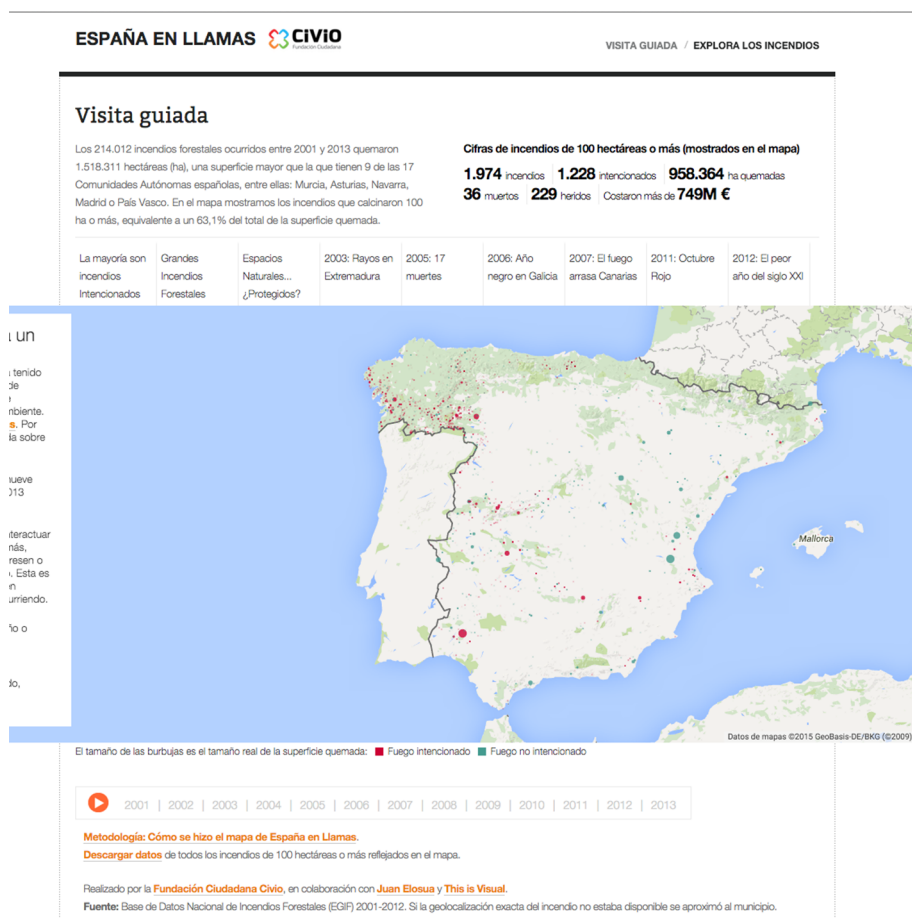


Figura 3. Visualització sobre els incendis a Espanya
 Font: <http://espanaenllamas.es/visita-guiada/>

b) Transport

Aaron Koblin és un artista especialitzat en visualització de dades i narració d'històries amb dades socials i d'infraestructures que fa servir per a examinar les tendències culturals i els patrons emergents que hi estan presents. Es tracta de dades del món real, que han estat generades per diverses comunitats d'usuaris d'algun servei en particular, i les utilitza per a reflexionar sobre les tendències culturals i les relacions canviants entre els éssers humans i la tecnologia. Un dels seus projectes més famosos és l'anomenat patrons de vol, generat amb Processing, un llenguatge de programació de codi obert orientat a la creació de visualitzacions. En el projecte apareixen resultats del trànsit

d'avions sobre Amèrica del Nord (i tot el món) durant un període de 24 hores. Es pot observar que formen una espècie de teranyina (una metàfora utilitzada per l'artista), de vegades multicolor, en una imatge fixa amb tot el lapse de temps de les dades dels vols, amb color aplicat a cada tipus de vol, perquè es puguin veure tots els avions que hi ha al cel en aquest període. L'artista també afegeix dades relatives a l'altura a la qual volen els avions, de manera que en la versió interactiva és possible observar amb tot detall les pistes d'enlairament i aterratge dels aeroports, que queden ressaltades pel gran volum d'avions que passen pel mateix lloc.



Figura 4. Visualització dels vols sobre els Estats Units
Font: <http://www.aaronkoblin.com/work/flightpatterns/index.html>

c) Política

PolitiLines és un projecte de visualització semàntica que permet saber de què van parlar els candidats a les presidencials dels Estats Units del 2012 durant aquell any en diferents debats. Mostra un diagrama interactiu amb tres columnes: temes, paraules i candidats, més una opció per a seleccionar el debat desitjat. Fent clic al damunt de cada ítem es pot obtenir de manera visual la relació entre les tres columnes esmentades i destacar els més importants. És una manera ràpida d'analitzar un text i poder accedir a les diferents parts relacionades. Anàlogament, també permet comparar fàcilment quins temes ha tocat cada candidat, els quals poden resumir el seu programa electoral i ideologia.

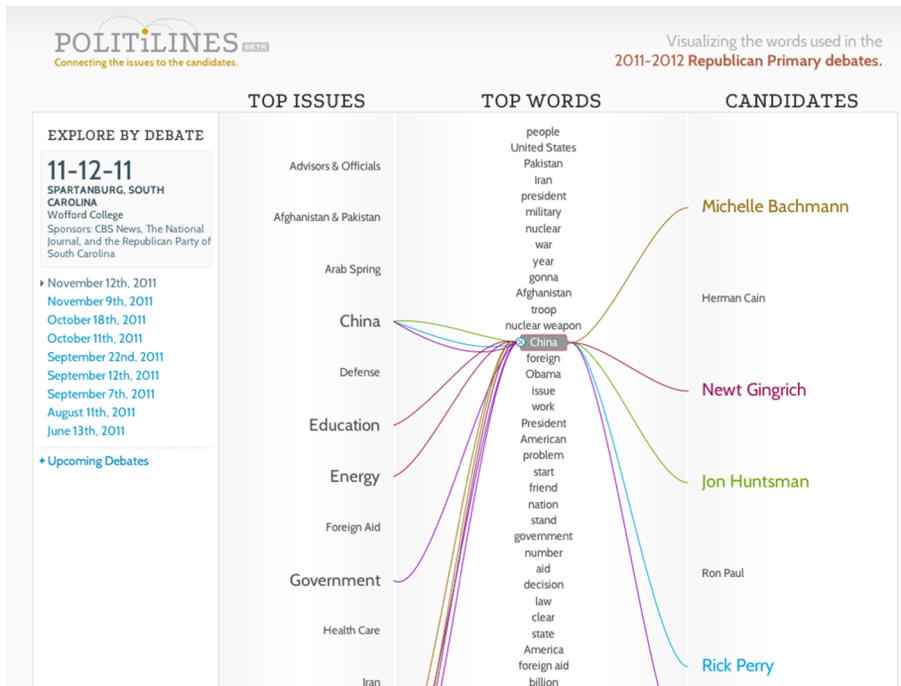


Figura 5. Visualització dels debats presidencials de l'any 2012
Font: <http://politilines.periscopic.com/>

d) Turisme

Espanya és un dels països capdavanters pel que fa al turisme al món. Per a demostrar-ho, **Vizzuality** i **CartoDB** van crear una eina interactiva que permet analitzar la despesa turística a Espanya durant l'estiu del 2014 usant les dades anònimes de 5,4 milions de transaccions de targetes de crèdit proporcionades pel BBVA. Aquesta visualització permet explorar minut a minut com gasten els seus diners els turistes a Espanya i analitza algunes de les ciutats més importants. Disposa d'una interfície de visualització i presentació molt nítida i atractiva, encara que fa servir els colors corporatius del BBVA. La visualització permet descompondre les dades mitjançant diferents pestanyes, que proporcionen múltiples angles en les dades, i és possible combinar mapes de diferents resolucions i línies de temps que s'executen alhora. També és possible aplicar filtres per a veure, per exemple, només els turistes d'una nacionalitat concreta. Amb el *zoom* del mapa principal és possible destriar, fins i tot, quines són les principals artèries comercials de ciutats com Barcelona, per exemple.

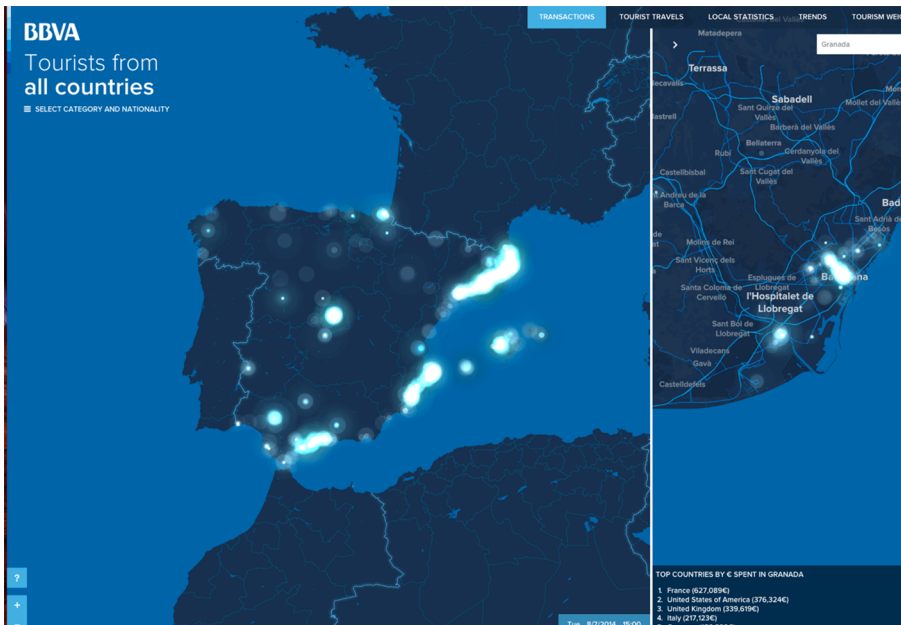


Figura 6. Visualització dels patrons de compra dels turistes
 Font: <http://bbvatourism.vizzuality.com/>

e) Multimèdia

El web actual és un univers d'aplicacions i pàgines web interconnectades que inclou vídeos, fotografies i altres continguts interactius. Però el que no veu l'usuari és com interactuen els navegadors i les tecnologies web per a fer que això sigui possible. Aquesta visualització multimèdia (de fet, una infografia) permet explorar les tecnologies web, com **HTML5**, **CSS3** o **WebGL**, entre altres, i veure com han anat evolucionant al llarg del temps. Les línies de color d'aquesta visualització representen la interacció entre els navegadors existents en cada moment i les tecnologies web presents en cadascun dels navegadors. Per a cada navegador, començant per Mosaic, aparegut el març del 1993, és possible conèixer-ne les diferents versions i l'aspecte mitjançant la superposició de captures de pantalla. D'altra banda, de cada tecnologia web és possible obtenir-ne una breu descripció i la seva disponibilitat en cadascun dels navegadors mostrats.

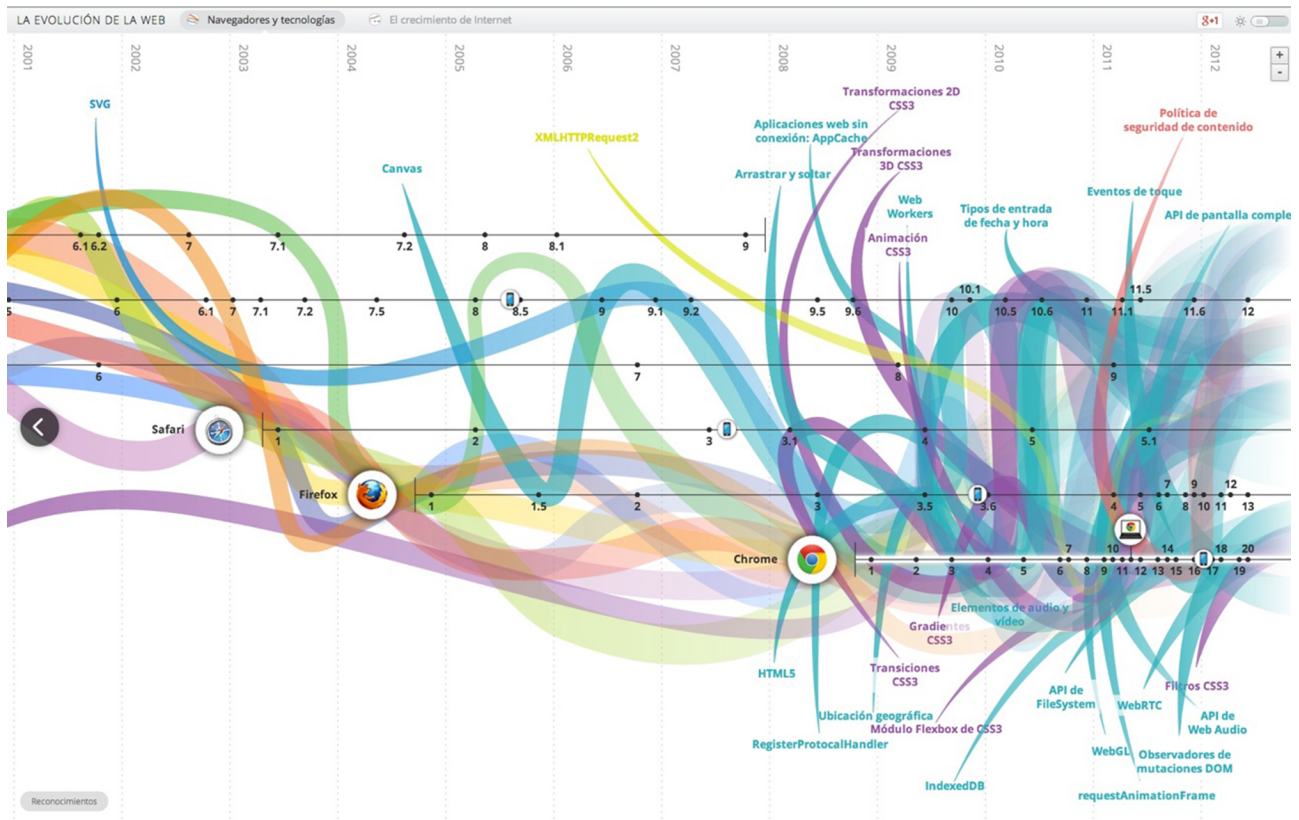


Figura 7. Infografía sobre les tecnologies i navegadors web
 Font: <http://evolutionofweb.appspot.com/>

f) Informes

El projecte de visualització de dades de **PRB Digital** analitza les tendències clau de la població mundial en l'àmbit demogràfic. Permet explorar la població actual i projectar-la per regió i país, i també observar els canvis en la fecunditat total, la mortalitat infantil i l'esperança de vida des del 1970. Es tracta d'un conjunt d'infografies que realment expliquen una història sobre la població del planeta i que combinen diferents tipus de visualitzacions, totes amb cert grau d'interactivitat. Un aspecte molt interessant és la manera d'aconseguir-ho, ja que en el cas de la població mundial (encara que és possible seleccionar un país o continent concret) treballa amb dos gràfics radials (de tipus dònut) concèntrics i connectats entre si, els quals permeten explorar les dades simultàniament en dues dimensions, una de geogràfica per continent (en el cercle interior) o per país (en el cercle exterior), i una altra de temporal, i comparar dades actuals amb una projecció de l'any 2050.

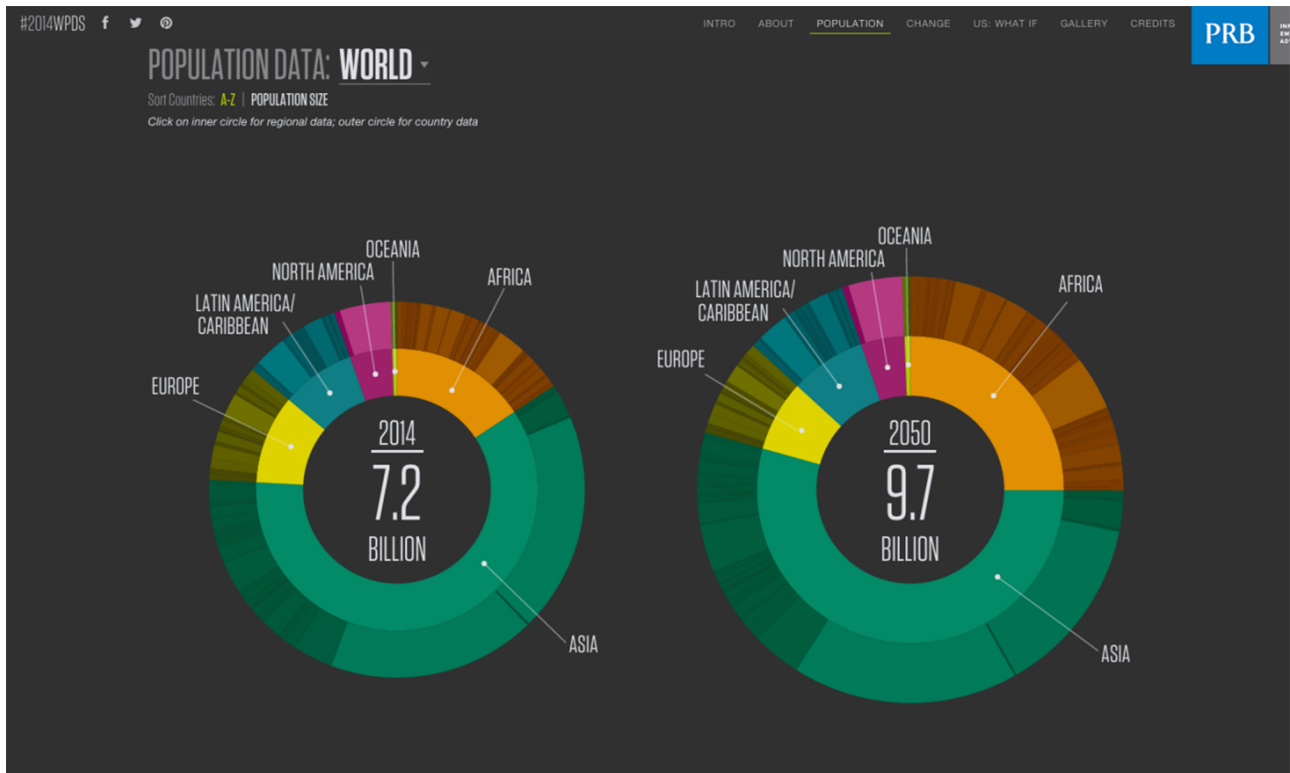


Figura 8. Visualització de dades sobre la població mundial
 Font: <http://visual.ly/2014-world-population-digital-visualization?view=true>

g) Xarxes socials

Mapr és un projecte de visualització de dades geolocalitzades creat per Pimpampum, un col·lectiu que es dedica a impulsar iniciatives amb dades en obert obtingudes de xarxes socials en molts casos. Basant-se en les etiquetes de fotografies que han estat penjades a Flickr i les geodades incloses en aquestes fotografies, aquesta visualització permet posicionar en un mapa les fotografies com a punts basats en la ubicació on van ser fetes. Seguint una senzilla seqüència de passos guiada, es poden introduir paraules de cerca de forma niada, que es faran servir per a anar creant punts de diversos colors segons cada paraula introduïda. Cadascun dels punts és, al seu torn, un element de selecció, i es disposa d'un visor de la fotografia associada a aquest punt i de la seva informació annexa. És destacable que no es visualitza realment cap mapa, sinó que apareix representat pels punts que responen a les paraules clau especificades, tal com mostra l'exemple següent.

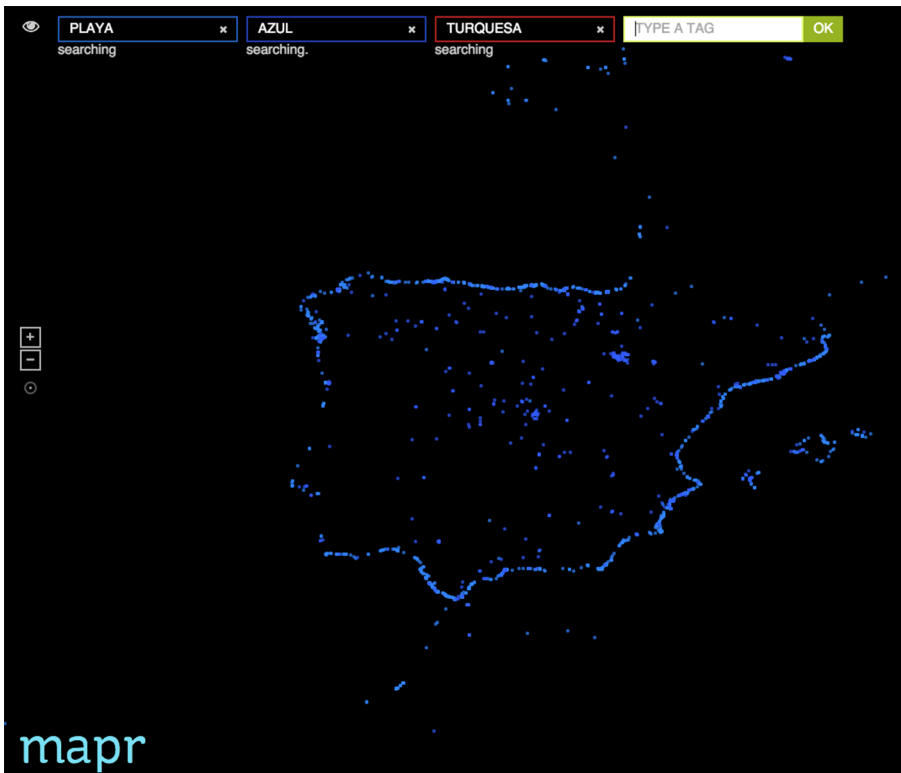


Figura 9. Geolocalització de fotografies penjades a Flickr
Font: <http://www.pimpampum.net/labs/mapr/>

3. Una mica d'història

El disseny de la informació i la seva representació gràfica estan presents actualment en el nostre dia a dia. Estem envoltats de representacions visuals d'informació, gràfics, mapes, icones, gràfics de barres, cartells, diagrames, esquemes, etc. Encara que normalment aquest fet es relaciona amb les noves eines digitals, el concepte de comunicació visual de la informació no és nou. Els éssers humans han dibuixat imatges per a comunicar-se des de fa milers d'anys, des de pictogrames a les parets d'una cova rupestre, passant pels jeroglífics egipcis, fins a ideogrames i iconografia moderna. A l'ésser humà li ha agradat sempre l'ús d'imatges per a comunicar i explicar històries, perquè el cervell humà està dissenyat d'aquesta manera i és molt eficient processant informació mitjançant el sistema visual.

De fet, aquesta cultura visual ja s'integra de manera natural en el dia a dia des de fa uns anys, atès que amb la irrupció de l'ofimàtica, en qualsevol reunió d'empresa en la qual es pretengui presentar informes, per exemple, habitualment es creen petites visualitzacions de dades, encara que sigui un simple gràfic de barres. Però el que és més recent és l'interès que aquest tema desperta i el fet que arriba pràcticament a tots els àmbits, ja que la proliferació de noves eines digitals combinades amb l'ampli espectre actual de fonts d'informació i dades disponibles faciliten que aquesta disciplina arribi a un nombre més gran de professionals de diverses disciplines, que van des del periodisme fins a l'educació, entre altres.

En el passat recent, aquest tipus de disciplina era domini exclusiu dels acadèmics, enginyers i estadístics, però ara s'està democratitzant per la necessitat de gestionar la càrrega cognitiva, és a dir, per la necessitat de tenir la capacitat de discriminar i filtrar la informació que té importància, i també per a entendre com es pot aprofitar al màxim el coneixement circulant emprant una varietat d'eines i tècniques. La pròxima generació de treballadors haurà de desenvolupar les seves pròpies tècniques per a abordar el problema de la sobrecàrrega d'informació mitjançant el filtratge per meritocràcia natural relativa a la importància de les dades, l'etiquetatge col·lectiu o afegint altres metadades al contingut, inclosa la visualització d'informació, per a poder destacar el més rellevant per sobre del «soroll» provinent del volum de dades disponibles.

Com s'ha comentat anteriorment, la necessitat de representació gràfica d'informació es remunta a molts segles enrere i ha estat sempre present en la història de la comunicació humana, encara que només (per a nosaltres actualment, no en el seu moment) fos un simple dibuix aïllat en una pintura rupestre. Per exemple, durant el període comprès entre el 3200 aC i el 400 dC, a Egipte ja existien diversos exemples de com s'utilitzava una tècnica que avui seria reconeguda com a infogràfica per a representar la informació. De mane-

ra similar, es poden destacar la cultura maia i la cultura asteca per les seves grans mostres d'infografies, que expliquen detalladament diversos episodis de la seva història.

Ja al segle XI (concretament l'any 1077), es troba el famós tapís de Bayeux (mostrat en la figura 10), un gran llenç brodat que relata, amb una successió d'imatges amb inscripcions en llatí, els fets previs a la conquesta normanda d'Anglaterra, que va culminar amb la batalla de Hastings.



Figura 10. Tapís de Bayeux
Font: Viquipèdia https://es.wikipedia.org/wiki/Tapiz_de_Bayeux

Es diu que Leonardo da Vinci, un dels grans pintors del Renaixement, anotava les seves idees sobre les recerques que feia en forma de dibuixos explicatius. Per tant, aquest pintor i inventor va utilitzar el concepte d'infografia sense saber-ho (o millor dit, sense conèixer-ne el terme). Un altre exemple semblant és Nicolau Copèrnic, que va desenvolupar un gràfic amb el qual mostrava el canvi del concepte tradicional de l'univers que va donar motiu al desenvolupament de l'astronomia moderna.

Diversos autors sostenen que la infografia moderna té els seus orígens en el segle XVIII, quan els diaris d'aquella època feien grans esforços per a il·lustrar les informacions publicades. Encara que les primeres infografies eren senzilles, en blanc i negre i amb molt poca capacitat d'il·lustració, amb el pas dels anys es van anar desenvolupant noves tècniques i els treballs infogràfics van anar avançant. Ja avançat el segle XIX, cal que ens referim a Charles Minard, destacat enginyer civil francès reconegut pel seu notable treball en el terreny dels gràfics informatius, ja que va ser pioner en l'ús de gràfics estadístics i tècnics. El seu treball més reconegut és «Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813» (mostrat en la figura 11). El gràfic mostra diferents variables en una única imatge bidimensional, entre les quals la situació i direcció de les tropes, i s'hi pot veure com les unitats que les componen es divideixen i reagrupen, i també el minvament de les tropes, per exemple en el pas dels rius o quan hi ha un descens de les temperatures.

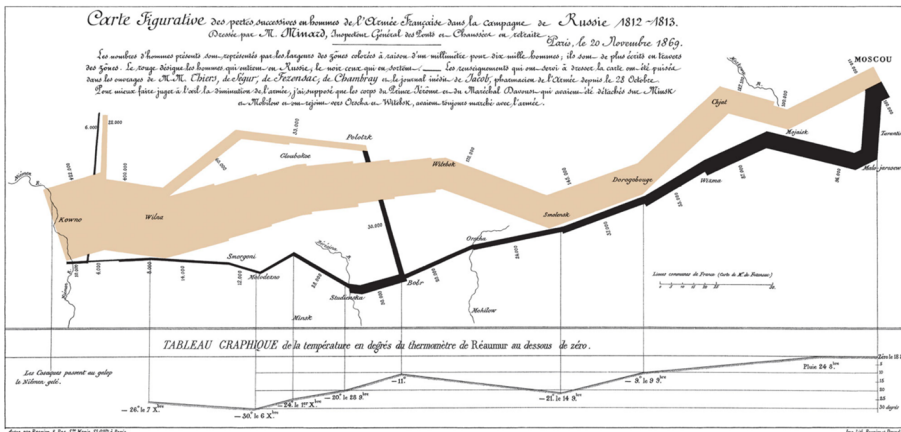


Figura 11. «Carte figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813», de Charles Minard
Font: Viquipèdia https://es.wikipedia.org/wiki/Charles_Joseph_Minard#/media/File:Minard.png

Un altre cas cèlebre del segle XIX és el del metge anglès John Snow, que es podria destacar com el veritable començament de la infografia moderna. El **mapa de Snow (mostrat en la figura 12)** està considerat com un dels primers exemples d'ús d'un mètode geogràfic per a descriure i localitzar els casos d'una epidèmia, i també el seu origen més probable, una cosa no feta mai anteriorment perquè es desconeixien els mecanismes de transmissió de les malalties infeccioses. És un exemple de com una visualització pot ser usada per a validar o refutar una hipòtesi, en aquest cas la transmissió del còlera per l'aigua.

Així, **John Snow (1813-1858)** va ser un metge anglès que va demostrar que els casos de còlera a Londres eren causats pel consum d'aigua contaminada amb materials fecals. El 1854 va cartografiar en un plànol del Soho els pous d'aigua i els casos de còlera, i va localitzar com a causant el pou situat a Broad Street, en ple cor de l'epidèmia, per la qual cosa va recomanar clausurar la bomba d'aigua que l'alimentava, i es va aconseguir així disminuir la proliferació dels casos de còlera en aquesta zona de Londres.

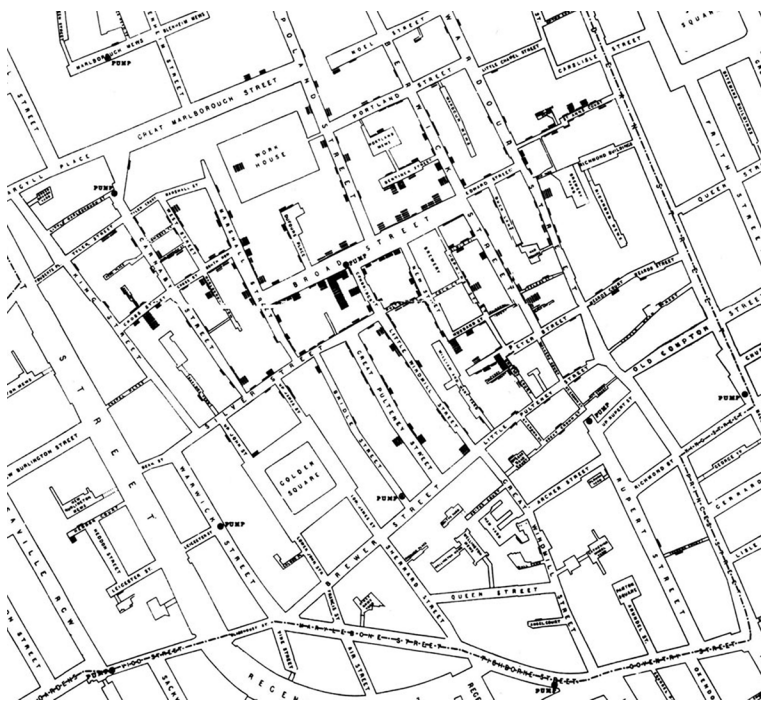


Figura 12. Mapa de John Snow amb la situació dels casos de còlera a Londres
Font: https://es.wikipedia.org/wiki/John_Snow#/media/File:Snow-cholera-map.jpg

Al començament del segle XX, durant la **Primera Guerra Mundial**, es van utilitzar molt els mapes per a informar sobre el que passava als països involucrats, i cada vegada eren més il·lustratius a mesura que avançava el temps. Ja entrats en la segona meitat del segle XX, les infografies eren cada vegada més utilitzades, curiosament sobretot per diaris espanyols que trencaven amb la serietat dels mitjans tradicionals, il·lustrant cada vegada més les seves impressions amb grans fotografies i infografies a color.

La infografia es va veure impulsada en el seu desenvolupament, principalment, a causa de tres grans esdeveniments de la història. El primer va ser el del cas del diari nord-americà USA Today, que va transformar completament el seu disseny i va començar a incloure treballs amb infografies a gran escala, i fins i tot les va començar a publicar en cadascuna de les portades dels seus quaderns. En algunes ocasions va arribar a dedicar tot un quadern únicament a treballs d'aquest tipus. Això va fer que altres mitjans impresos europeus, i sobretot espanyols, adoptessin la mateixa metodologia, de tal manera que fins i tot van sorgir nous mitjans que es van adaptar als treballs amb un estil més infogràfic com a estratègia per a atreure el públic.

El segon esdeveniment, i per a molts el més important, va ser el de l'aparició dels primers ordinadors amb programes digitals i processadors de text, gràcies als quals es podien veure i crear gràfics amb més facilitat, aspecte que és considerat crucial perquè a mesura que anaven avançant aquestes tècniques els lectors tenien més receptivitat cap a aquests continguts i més interès a informar-se en els mitjans que els feien servir.

Un altre dels esdeveniments clau que va impulsar l'ús de la infografia en els mitjans va ser la guerra del **golf Pèrsic** a la dècada dels noranta, ja a la fi del segle XX, ja que els periodistes tenien l'accés denegat al lloc dels fets, per la qual cosa amb prou feines podien fer-ne fotografies i les que publicaven eren molt poc informatives i de significació insuficient. Com a conseqüència d'això, l'enginy periodístic i la necessitat d'informar van buscar camins alternatius per a produir notícies del que estava passant, de manera que la infografia a gran escala i amb informació geogràfica es va convertir en la principal eina per a situar i donar més coneixement dels fets al lector, fins i tot sense poder trepitjar el terreny com en el cas del periodisme tradicional.

A mesura que va anar avançant l'adopció de la tècnica infogràfica en les redaccions periodístiques, les representacions adoptaven característiques noves, els textos eren cada vegada més curts, les imatges més il·lustratives i s'utilitzaven cada vegada més en diferents temes. Avui dia, i encara que segueix sent liderada per periodistes, l'ús de la infografia s'ha anat estenent i s'ha ampliat cap a diversos àmbits, com l'empresarial, el màrqueting, el científic i tecnològic o l'acadèmic (recerca i docència), entre altres.

Per tant, és pot comprovar que la necessitat de visualitzar i explorar la informació emprant gràfics és molt antiga, però que el que sí que és realment nou és l'avidesa per la gestió de la informació i les dades en un entorn digital com l'actual, en el qual tot el que una persona fa durant la seva vida diàriament sovint és registrat i quantificat, cosa que genera un rastre digital de dades enorme. A més, a diferència dels mètodes utilitzats en èpoques anteriors, les tecnologies de la informació i la comunicació (TIC) estan democratitzant l'accés a aquestes dades i facilitant eines digitals per convertir-les en informació i, per tant, en coneixement rellevant. L'aparició d'eines com ara llenguatges de programació específics per a crear visualitzacions i infografies n'és una mostra. Es podria dir, doncs, que la visualització de dades i d'informació no és una disciplina de moda, sinó que ha vingut per a quedar-se.

4. Avantatges de la visualització

Els avantatges de la visualització de la informació són múltiples i se situen en diversos camps, especialment els relacionats amb la manipulació de dades, com són el científic i tecnològic. L'any 2001, **Edward R. Tufte**, un dels pioners de la visualització moderna de dades i professor emèrit de la Universitat de Yale, va explicar clarament aquest fenomen quan va dir:

«De tots els mètodes d'anàlisi per a comunicar informació estadística, el gràfic de dades ben dissenyat és generalment el més simple i al mateix temps el més poderós.»

E. R. Tufte (2001). *The Visual Display of Quantitative Information*.

Aquesta nova manera de comunicar permet divulgar en el camp gràfic, i el seu avantatge principal consisteix a facilitar la comprensió i l'aprenentatge, a més de la retenció cognitiva del contingut. A continuació es descriuen algunes de les raons bàsiques del seu èxit.

- **Més bona gestió de la sobrecàrrega informativa**

Els humans del segle XXI són éssers *informívors*, terme encunyat per **George Armitage Miller**, un dels fundadors de la psicologia cognitiva el 1956, amb la finalitat de descriure el comportament dels éssers humans alhora de recollir i consumir informació. Diàriament cal processar una immensa quantitat de dades i informació, incloses notícies, anuncis, correus electrònics, converses, missatges de text, piulades, llibres, cartells, presentacions, vídeos, etc.; una llista innombrable. I per descomptat, tot el relacionat amb la cerca i gestió d'informació a internet. El repte és principalment el de filtrar el «soroll», centrar-se en la informació pertinent i recordar la informació important.

Estudis recents revelen que, actualment, la informació que es genera a internet a escala mundial es duplica cada dos anys. No és estrany, doncs, que les persones experimentin estrès informacional i una sobrecàrrega d'informació. A causa d'aquesta sobrecàrrega, cal ser més efectius i eficaços a l'hora de cridar l'atenció, i crear una infografia o visualització de dades convincent pot permetre sobresortir del «soroll» que es genera a internet. Es tracta, doncs, de crear contingut atractiu, rellevant i sintètic, que activi la curiositat de la persona i la'n faci partícip.

- **Millora de la comprensió**

L'ésser humà és una criatura molt visual. La capacitat d'interpretar ràpidament la informació visual és molt més gran que la de processar paraules escrites o encara més gran que en el cas de dades numèriques. A més, la majoria dels ciutadans del segle XXI disposa de períodes d'atenció cada vegada més curts a causa de la sobrecàrrega informativa, per la qual cosa es tendeix a «escanejar» el contingut (el que comunament es diu una lectu-

ra en diagonal) en lloc de llegir el text detalladament. Això implica una baixa capacitat de retenció lectora, ja que segons estudis fets recentment, les taxes de retenció de les persones són molt baixes depenent de la informació que llegeixen, ja que s'estima que només recorden el 20% del que llegeixen. Es recorda més la informació que s'ha vist que la que s'ha llegit, per la qual cosa el disseny visual de la informació facilita la comprensió d'informació complexa.

Feta servir com a eina d'aprenentatge, la visualització de la informació permet als estudiants desenvolupar funcions cognitives d'ordre superior, ja que utilitzant-la correctament, han d'avaluar, analitzar i, finalment, representar gràficament la informació, de manera que s'adquireixen noves competències que es consideren bàsiques per a aquest començament del segle XXI. D'altra banda, si el que s'ha de fer és gestionar grans volums de dades, la seva visualització permet una primera aproximació per a comprendre-les, ja que permet fer emergir i percebre propietats que no havien estat previstes o detectar problemes relatius a les dades; per exemple, errors en la seva recollida (absència) i en la seva validesa (valors extrems).

- **Promou compartir**

Si una persona ha de triar entre una cosa que és visualment atractiva i fàcil de digerir, la triarà preferentment abans d'altres formats com el text escrit, que requereix més temps i energia. D'altra banda, la naturalesa visual de l'espècie humana és una de les raons per les quals a les persones ens encanta compartir a les xarxes socials una infografia i visualització de dades que considerem interessant. Així ho demostra l'adopció generalitzada de la infografia per les agències de comunicació i màrqueting per a crear l'anomenat *link building* i màrqueting viral, ja que una bona infografia permet a una audiència concreta consumir més informació complexa en menys temps.

Aquesta mateixa visió provinent de l'àmbit del màrqueting es pot utilitzar amb diferents finalitats, siguin socials, comunitàries (conscienciació, prevenció o ensenyament), comercials (venda de productes, funcionalitats de productes o serveis), informatives, estadístiques o institucionals.

5. Estructura d'una infografia

En el moment de crear una infografia és clau determinar-ne la finalitat; per exemple, si es tracta d'explicar un concepte o una tecnologia, d'analitzar dades estadístiques, de mostrar el resum d'un procés, la descripció d'un producte o servei, la història d'una empresa, etc. També és important recaptar la informació, a part de la d'elaboració pròpia de la qual es disposi, filtrar-la i descobrir el punt clau que es vol comunicar. En aquest sentit cal recordar que l'excés d'informació genera desconcentració i frustració, de manera que s'assoleixen objectius adversos. També cal tenir en compte que la infografia pot disposar d'elements textuais i gràfics i combinar-los (icones i pictogrames, fotografies). Com a part final es construirà la infografia seguint una estructura que constarà normalment de tres seccions ben diferenciades: la introducció, el cos i la conclusió.

- La primera secció és la introducció, i és on s'introdueix el lector en el tema sobre el qual tracta la infografia. Pretén respondre la pregunta de per què hauria d'interessar al lector explorar-la. És on s'explica al públic objectiu que aquesta infografia conté quelcom rellevant específicament pensat per a ell. Es compon normalment d'un títol i un breu paràgraf textual, i estableix les bases de la informació que es disposarà a continuació en el cos de la infografia, on posteriorment es desenvoluparà amb més detall.
- Seguidament, el cos és la secció que dona sentit a la infografia, i de totes tres és la més important. Perquè una infografia sigui efectiva i significativa per a l'audiència a la qual va dirigida ha de contenir alguna informació desconeguda, nova i rellevant. El cos és, en general, el bloc visual dominant de la infografia que conté aquesta informació. L'aspecte més visual i iconogràfic de la infografia normalment es col·loca en aquesta secció, ja que és el punt principal i el que ha de recordar el lector.
- Tota infografia té un tancament en la seva part inferior, anomenat conclusió. En aquesta secció es troba el resum o, també, la moralitat, el missatge clau que posa el coronament final al discurs visual que articula la infografia. En cas que la infografia tingui una clau comercial, aquesta secció ha d'incitar a l'acció, com ara comunicar-se amb el creador, comprar un producte, etc. Si és una infografia educativa, la conclusió ha de ser el concepte clau que es vol transmetre, de manera resumida i efectiva.



Figura 13. Exemple d'infografia amb una estructura de capçalera (1), cos (2) i peu (3).

6. Tipus d'infografies

En el disseny d'informació visual i infogràfica es disposa d'un espectre molt ampli de criteris o dimensions per a definir i classificar una infografia. Es podrien seguir diverses pautes, però en aquest cas, per a donar una orientació senzilla, ens centrarem en dos criteris principals: la tipologia d'infografia segons la seva composició i segons la seva finalitat.

6.1. Segons la seva composició

Les tipologies d'infografies principals segons la seva composició són les següents: cronològica, comparativa, diagrama de flux, descriptiva, seqüencial, estadística, geogràfica i jeràrquica, i es descriuran a continuació. Òbviament, una infografia pot combinar més d'una tipologia, encara que sol haver-n'hi una que predomina sobre la resta. Hi ha serveis com VENNGAGE que permeten crear infografies de les tipologies esmentades.

- **Cronològica:** també anomenada línia de temps, mostra la informació en una línia temporal i en descriu l'evolució en el temps. És molt utilitzada en les descripcions biogràfiques d'algun personatge o marca, ja que permet introduir-hi el context històric.
- **Comparativa:** compara dos o més valors o ítems utilitzant diferents recursos gràfics, normalment confrontant-los i destacant-ne les diferències.
- **Diagrama de flux:** visualitza un diagrama d'activitats que mostra el flux de control general d'un procés o activitat, esquemàticament, per blocs o amb diagrames.
- **Descriptiva:** descriu els aspectes fonamentals d'un determinat objecte, utilitzant símils visuals com un espejament o l'ús de seccions i perspectives.
- **Seqüencial:** representa una història en passos i en etapes, sense fer tant d'èmfasi en l'aspecte temporal. Normalment s'utilitza per a visualitzar els passos d'un procés.
- **Estadística:** mostra mètriques, vendes, estudis de mercat, enquestes, tendències, resultats, basant-se en propietats com ara mitjanes, variàncies, etc.
- **Geogràfica:** situa la informació en un lloc, posicionant-la geogràficament mitjançant l'ús de mapes, croquis de carrers o qualsevol concepte amb una estructura espacial subjacent.

- **Jeràrquica:** estableix un ordre de la informació segons un procés de classificació jeràrquica que consisteix a separar un conjunt d'objectes en classes i subclasses, seguint generalment una estructura en arbre.

6.2. Segons la seva finalitat

D'altra banda, un altre enfocament per a caracteritzar les infografies pot ser la seva classificació d'acord amb l'objectiu que es persegueix amb la seva utilització. Les principals tipologies d'infografies segons la seva finalitat es descriuen a continuació.

- **Periodística:** es tracta d'una infografia narrativa, atès que és una espècie d'article visualitzat. És ideal per a situacions en les quals el pes numèric no sigui la clau i el tema es pugui refermar amb les imatges utilitzades.

El paradigma de visualització moderna d'aquest tipus és el reportatge en línia «The Russia left behind», d'Ellen Barry, que fa servir diversos elements gràfics per a explicar una història.

- **Didàctica:** és la infografia científica, que mostra i ajuda a descriure processos de producció o ensenya a utilitzar algun producte i a comprendre el funcionament de qualsevol element mitjançant esquemes. En general, contribueix millor que qualsevol altre material didàctic a la interpretació del que s'hi mostra.

Per exemple, el Jet Propulsion Laboratory de la NASA utilitza moltes infografies d'aquest tipus amb finalitats educatives, com conscienciar sobre el canvi climàtic. En l'àmbit científic, la conferència SIGGRAPH és el lloc on acadèmics que treballen amb visualitzacions de dades científiques presenten els seus resultats de recerca.

- **Informativa:** és qualsevol infografia que presenta dades o informació rellevant, de manera contundent. Aquesta tipologia pot contenir els trets o aspectes esmentats anteriorment barrejats o combinats entre si; persegueix un objectiu publicitari i ajuda a disposar les dades o la informació rellevant eficientment.

Per exemple, es pot destacar la visualització de la crisi dels refugiats sirians, que mostra cruament, mitjançant dades, l'horror de la guerra en aquest país.

- **Empresarial:** orientada a les empreses amb fort component comercial o de comunicació interna, està dirigida a exposar dades estadístiques, processos, presentació de productes, estudis de mercat, etc.

Un exemple és el conjunt de categories de paraules clau més usades a Google i el seu impacte en el negoci fet per Google mitjançant la venda d'anuncis en línia.

7. Categories de la visualització de la informació

En aquest cas, per a donar una orientació sobre les diverses categories de la visualització d'informació i dades s'analitzen amb un altre enfocament, seguint els criteris estadístics que es persegueixen o segons el seu origen i tipologia. A continuació es descriuen algunes de les categories més comunes:

- **Descriptiva**

Es tracta d'utilitzar una visualització per a descriure gràficament les característiques o els atributs que determinen una població, a partir del conjunt de dades disponible, bo i permetent a l'usuari fer una anàlisi descriptiva visual de cada atribut o de les relacions entre aquests. Això inclou els gràfics típics fets servir en estadística descriptiva (histogrames, funcions de distribució, diagrames de caixa, etc.), però també altres tècniques com, per exemple, els núvols d'etiquetes, tal com mostra la figura 14, on les aparicions d'usuaris de Twitter durant un debat televisiu en plena campanya electoral apareixen amb una grandària proporcional al nombre d'aparicions.

(1) També anomenada georeferenciada: <https://es.wikipedia.org/wiki/georeferenciación>



Figura 14. Núvol d'etiquetes a partir de les aparicions a Twitter

- **Tendències**

El concepte de tendència, en un sentit general, és un patró de comportament dels elements d'un entorn particular durant un període de temps. Un típic exemple n'és l'evolució de la piràmide de població, tal com mostra la figura 15.

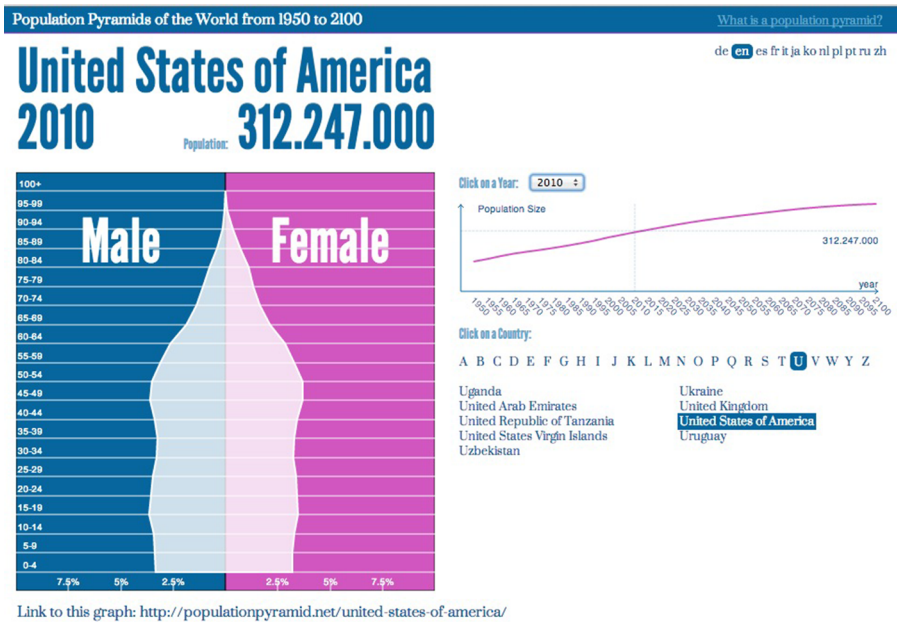


Figura 15. Exemple de piràmide de població
Font: <http://populationpyramid.net/>

- **Patrons**

Un patró abstracte és una forma o model els elements del qual es repeteixen d'una manera previsible, aportant informació sobre la seva estructura interna. En la figura 16 es mostra un detall de l'anàlisi dels patrons dels recorreguts fets pels corredors d'una ciutat, en aquest cas **Washington DC**.

Where People Run in Major Cities

BY NATHAN YAU / POSTED TO DATA UNDERLOAD / TAGS: R, RUNNING, EXERCISE



Figura 16. Patrons dels recorreguts dels corredors a Washington DC.
Font: <http://flowingdata.com/2014/02/05/where-people-run/>

- **Anomalies**

L'anomalia en les dades és una característica definida subjectivament que s'assigna a les dades que es considera que mostren unes condicions rares o disfuncionals. L'exemple mostrat en la figura 17 visualitza les anomalies de temperatura dels últims cinquanta anys als Estats Units i fa èmfasi, en aquest cas, en els valors extrems.

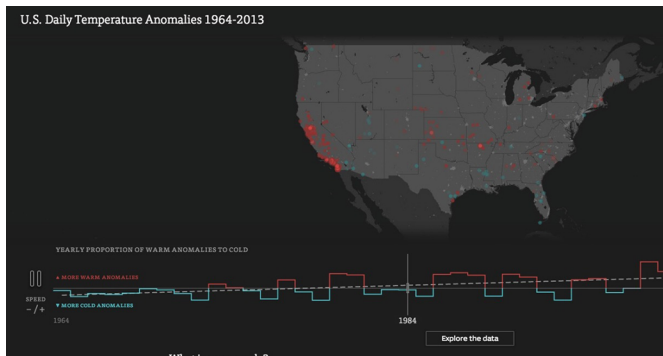


Figura 17. Mapa d'anomalies en les temperatures
Font: <http://blog.enigma.io/climate-map-announcement/>

• Connexions

Aquest tipus de visualització es refereix a com s'estableixen les relacions de connexió entre variables principalment categòriques. L'exemple mostrat en la figura 18 presenta els membres de l'Institut de Crèdit Oficial i les seves relacions segons una ontologia preestablerta, en aquest cas els càrrecs i les dependències dins d'aquesta entitat. Es tracta d'una visualització interactiva que permet manipular el gràfic que es visualitza.



Figura 18. Mapa de relacions entre persones i institucions de l'Institut de Crèdit Oficial
Font: <http://quienmanda.es/organizacions/ico>

• Correlacions

La correlació és un instrument de l'estadística que determina el grau de relació lineal o dependència (proporcionalitat) que hi ha entre dues variables. És a dir, es pretén determinar si els canvis en una de les variables influeixen en els canvis de l'altra, sempre en la mateixa direcció. En cas que això succeeixi, esquemàticament es diu que les variables estan corre-

lacionades o que hi ha una correlació entre les variables. En aquest cas, l'exemple mostrat en la figura 19 intenta analitzar la correlació entre la criminalitat i la desocupació als Estats Units.

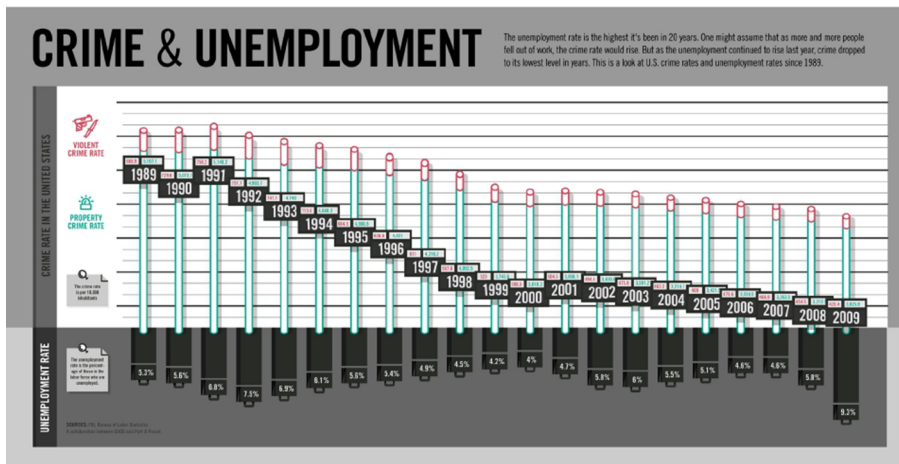


Figura 19. Correlació entre criminalitat i desocupació
Font: <http://awesome.good.is/transparency/web/1006/crime-and-unemployment/flat.html>

Estadísticament, la correlació entre dues variables és una mesura que quantifica aquesta proporcionalitat emprant un coeficient de correlació. És important recordar que una correlació no implica causalitat entre les dues variables, ja que pot existir una tercera variable oculta que sigui la que determina aquesta correlació o ser completament espúria. Uns exemples divertits d'aquesta diferència entre els conceptes de correlació i de causalitat es poden trobar al web <http://www.tylervigen.com/spurious-correlations>, tal com mostra la figura 20, que relaciona dos conceptes completament independents.

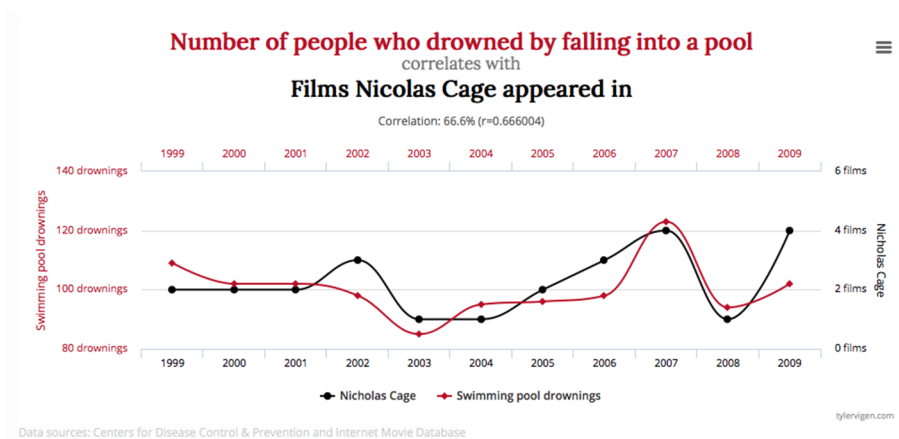


Figura 20. Relació (falsa) entre el nombre de persones ofegades en una piscina i el nombre de pel·lícules en les quals apareix Nicolas Cage
Font: <http://www.tylervigen.com/spurious-correlations>

- **Comparacions**

De vegades cal comparar dues o més variables quantitatives depenent d'un o més grups per a observar diferències entre aquests. En aquest cas, l'exemple mostrat en la figura 21 permet comparar el patrimoni dels diputats i classificar-lo per diversos criteris, entre els quals el partit polític a què pertanyen.

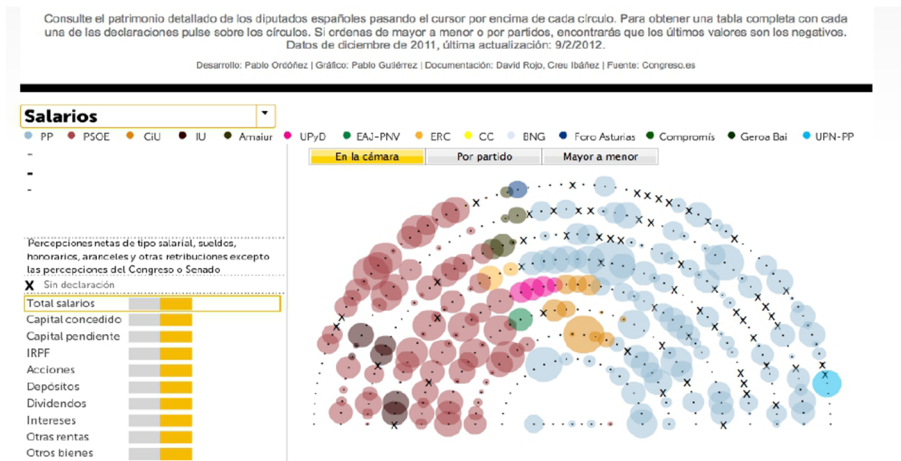


Figura 21. Comparació de salaris entre diputats espanyols per partit polític

- **Jerarquia**

La jerarquia és el criteri que permet establir un ordre de superioritat o de subordinació entre persones, institucions o conceptes. El *treemap* o mapa en arbre que es mostra en l'exemple de la figura 22 permet acotar l'espai de representació i dividir-lo en una jerarquia de rectangles, mostrant-ne alhora els atributs (que habitualment determinen la grandària de cada rectangle) i identificant propietats d'aquesta jerarquia mitjançant l'ús dels colors. En aquest cas, el projecte **maPo** sorgeix de crear i donar una nova dimensió a les paraules *mapa* i *política*.

Aquest projecte planteja una lectura alternativa de les compareixences que s'han dut a terme al Congrés dels Diputats en relació amb el Projecte de llei de transparència, accés a la informació pública i bon govern. Aquest mapa interactiu permet identificar ràpidament les temàtiques més rellevants que han tractat els compareixents, comparar les seves prioritats, comprendre els diferents enfocaments del problema i compartir aquests descobriments.

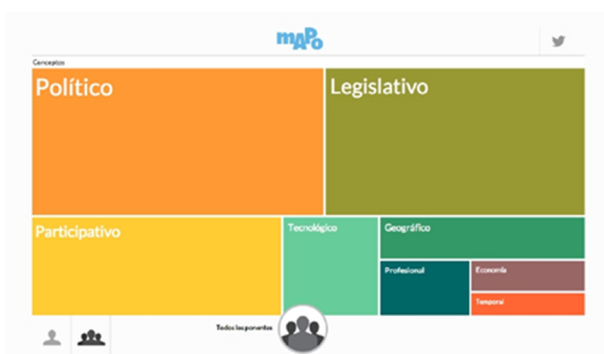


Figura 22. Mapa jeràrquic entre conceptes presents en les compareixences al Congrés

- **Localitzacions**

La localització permet visualitzar informació geoposicionada¹, és a dir, que disposa de dades de longitud i latitud (i en alguns casos també d'altura o profunditat). L'exemple mostrat en la figura 23, part del projecte BIG TIME BCN, és un mapa interactiu de la ciutat de Barcelona en el qual és possible

comparar l'edat de cadascuna de les seves edificacions i descobrir-ne la història i herència cultural.



Figura 23. Mapa de Barcelona d'acord amb l'etapa històrica de cada àrea
 Font: <http://bigtimebcn.300000kms.net/>

- **Processos**

Un procés és un conjunt d'activitats mútuament relacionades que, en interactuar entre si, transformen elements d'entrada i els converteixen en resultats. En el cas mostrat en la figura 24 s'analitza el procés de viralització d'un rumor a les xarxes socials, en aquest cas Twitter.



Figura 24. Visualització del procés de viralització d'un rumor a Twitter

8. Eines per a la visualització

Es descriuen aquí breument un conjunt d'eines que poden ser usades per a la visualització de dades o per a crear infografies. Aquesta llista no pretén ser exhaustiva, atès que cada dia apareixen noves eines en aquest àmbit, sinó proporcionar un punt d'entrada per a començar a crear visualitzacions i infografies.

- **R**

R és un entorn complet per a l'anàlisi estadística que incorpora, de base, eines per a crear gràfics a partir de les dades que s'estan analitzant amb l'objectiu de visualitzar el resultat d'una anàlisi estadística descriptiva. Ara bé, el veritable poder de R és la possibilitat d'estendre les seves funcionalitats mitjançant *packages*, mòduls específics que afegeixen noves característiques a l'entorn, tant per al càlcul estadístic com algorismes per a mineria de dades, incloses també llibreries per a la visualització avançada:

- **ggplot2**: és un paquet que complementa les opcions bàsiques de R per a la visualització de dades; amaga alguns detalls necessaris per a la creació d'un gràfic, canviant la idea tradicional d'un gràfic únic que inclou tota la informació (dades, eixos, llegendes, etc.) per un nou paradigma basat en una descomposició per capes, en què cada element o grup d'elements que formen un gràfic es troba en una capa diferent, i cada capa és definida amb un llenguatge comú que en simplifica la reutilització entre diferents gràfics, de manera que es poden manipular tots els seus elements més eficientment.
- **ggmap**: és un paquet per a la creació de mapes o, en general, per a la visualització de dades espacials per sobre de mapes estàtics provinents de diferents fonts, entre els quals **Google Maps** i **Stamen Maps**. De fet, seguint la filosofia de **ggplot2**, aquest paquet permet l'addició de capes orientades a dades geogràfiques fent servir la mateixa estructura i llenguatge.

- **D3.js**

Una altra opció diferent per a la visualització de dades és **D3.js**, una llibreria **Javascript** creada per **Mike Bostock** que explota les possibilitats de diferents tecnologies web. Ofereix un marc coherent per a una gran diversitat d'estils i dissenys de visualitzacions fent servir una estètica refinada i molt atractiva. **D3** és l'acrònim de *data-driven documents*; és a dir, a partir d'un o més fitxers de dades en diferents formats i que es poden carregar dinàmicament, es genera la visualització desitjada (entre un ampli repertori de tipus disponibles), que pot incloure elements d'interactivitat, per la qual cosa la mateixa visualització es converteix en la interfície d'accés a les dades. Aquesta llibreria ha estat i és usada pels principals mitjans de

comunicació, *The New York Times* o la cadena de televisió MTV, entre altres; i hi ha multitud d'exemples disponibles per a la seva reutilització.

- **Infogr.am**

Es tracta d'una eina en línia que es pot fer servir gratuïtament (encara que està basada en un model de semipagament (*freemium*) per a crear infografies a partir de plantilles i que incorpora diversos recursos, que inclouen exemples i tutorials per a la seva utilització des del primer moment. Les infografies creades poden ser emmagatzemades com a imatges o compartides en diferents xarxes socials.

- **Processing**

Una altra opció radicalment diferent de les anteriors és Processing, un llenguatge de programació ideat expressament per a la creació de gràfics mitjançant l'ús de gràfics bàsics primitius (punt, línia, cercle, etc.). Processing està orientat a la creació d'un mitjà visual interactiu de manera senzilla, especialment per a perfils no tecnològics, per la qual cosa ha estat adoptat ràpidament per artistes, dissenyadors, arquitectes, etc. gairebé sense coneixements previs de programació o molt limitats. Processing va ser creat el 2001 per Ben Fry i Casey Reas, tots dos estudiants en aquell moment del Media Lab del MIT, dins del grup de recerca de John Maeda sobre estètica i computació. Com a llenguatge de programació que és, Processing és una eina de propòsit general, que pot ser usat per a crear qualsevol tipus de visualització, encara que això pot requerir un cert temps segons la seva complexitat.

- **Gephi**

Per a la visualització (i anàlisi) de xarxes existeix Gephi, un entorn de codi obert basat en Java molt utilitzat en contextos acadèmics per a la manipulació d'aquest tipus d'estructures. Està preparat per a visualitzar xarxes de grandària mitjana, de fins a 50.000 nodes i 500.000 relacions. Inclou diferents algorismes per a la visualització de la xarxa i també per al càlcul de diversos indicadors (centralitat, modularitat, comunitats, etc.).

- **Wordle i Tagxedo**

Hi ha moltes eines que poden ser usades per a generar núvols d'etiquetes, fins i tot sense haver d'instal·lar cap programari, com és el cas de les dues esmentades, Wordle (basada en Java) i Tagxedo (basada en Silverlight). A partir d'un conjunt de paraules o d'un URL, se n'analitza el contingut i es genera un núvol d'etiquetes on cada paraula és representada amb una grandària proporcional a la seva freqüència d'aparició. És possible determinar la forma i el color del núvol d'etiquetes, normalment utilitzant una paleta d'estils i colors. També sol ser possible eliminar paraules no desitjades individualment o, per exemple, les més comunes d'un idioma determinat (preposicions, articles, etc.). El resultat pot ser emmagatzemat com una imatge o es pot compartir en línia.

Resum

Al llarg d'aquest mòdul s'ha explorat la visualització de la informació com un procés de destil·lació de coneixement des de les dades originals. Una visualització explica històries, simplifica, mesura, compara, explora, descobreix, etc., i proporciona així un accés més senzill a les dades subjacents, al mateix temps que permet obtenir coneixement d'aquestes dades gràcies a la seva presentació o exploració visual.

Dins d'aquest continu que és la visualització de la informació, se n'han descrit les dues grans àrees, que són la infografia i la visualització de dades. La primera està indicada per a la presentació normalment estàtica de la informació o de dades, i és de caràcter més aviat narrativa. La segona està pensada per a l'exploració interactiva d'un conjunt normalment massiu de dades, en la qual l'usuari ja coneix mínimament el context.

A mode de cloenda, com bé indica el prestigiós periodista **David McCandless** en el seu llibre *Information is Beautiful*, quan es crea una visualització, de fet es tracta de «**dissenyar per a la comprensió**».

Bibliografia

Alcalde, I. (2015). *Visualización de la información de los datos al conocimiento*. Editorial UOC, SL («Manuales»).

Cairo, A. (2012). *The Functional Art: An introduction to information graphics and visualization*. New Riders.

Evergreen, S. D. (2013). *Presenting data effectively: Communicating your findings for maximum impact*. SAGE Publications.

Krum, R. (2013). *Cool infographics: effective communication with data visualization and design*. John Wiley & Sons.

McCandless, D. (2009). *Information is beautiful*. Londres: Collins.

Meirelles, I. (2013). *Design for Information: An Introduction to the Histories, Theories, and Best Practices Behind Effective Information Visualizations*. Rockport Publishers.

Tufte, E. R.; Graves-Morris, P. R. (1983). *The visual display of quantitative information* (vol. 2, núm. 9). Cheshire, CT: Graphics Press.

Wickham, H. (2009). *ggplot2: elegant graphics for data analysis*. Springer Science & Business Media.

Yau, N. (2013). *Data points: visualization that means something*. John Wiley & Sons.

Enllaços i recursos recomanats

<https://cran.r-project.org/>

<http://visual.ly/>

<http://www.davidmccandless.com/>

<http://www.informationisbeautiful.net/>

<http://www.edwardtufte.com/tufte/>

<http://www.visualisingdata.com/>

<http://flowingdata.com/category/visualization/>

<http://flowingdata.com/category/visualization/infographics/>

<http://maps.stamen.com/>

