Slide 84

Se si aggiungesse dell’interattività a questa infografica sarebbe meglio, in questo modo diventerebbe una dashboard.

Una volta capiti quali sono i principi alla base della visualizzazione, si vede come ragionare con i colori e organizzare le cose in modo “occidentale” o centrallizzato → le infografiche e le dashboard sono caratterizzate da diverse visualizzazioni messe insieme.

Si cerca di centrare la dashboard o l’infografica intorno all’argomento principale.

Al centro dei mondiali, interessa chi arriverà in cima al podio finale.

Al centro dell’infografica infatti ci sono i tornamenti che rimangono dopo che sono state eliminate tutte le squadre.

Questa infografica è assai colorata, però l’accostamento dei colori è buono in quanto ci sono buoni contrasti.

A SX si passa col canale del verde, poi come si abbassa il verde si vede comparire il giallo, poi l’arancione, poi il rosso e poi il blu.

Si utilizzano dei colori che si vedono in natura.

Questo accostamento di transizione dal verde al giallo al blu… è visto come più naturale e meno soggetto a contrasti all’interno della nostra percezione.

Nel momento in cui si sa come organizzare l’infografica, ci vuole un attimo a passare alla dashboard.

Le dashboard uniscono l’interattività all’infografica.

La dashboard si può dire che è l’evoluzione interattiva di un’infografica: entrambe combinano testi, visualizzazioni, immagini per attirare l’attenzione.

La dashboard è un’applicazione o una pagina Internet che permette una qualche interattività (a differenza della infografica).

Slide 85

Tufte per formalizzare queste “bugie” all’interno delle visualizzazione, per stabilire quando una visualizzazione è pasticciata ha stabilito queste tre formalizzazioni appunto.

Quando si parla di infografica e di visualizzazione volta a fare una sorta di pubblicità del contenuto che si vuole passare alle persone che vedranno la visualizzazione, allora ad un po’ di estetica bisogna fare attenzione.

Tufte cerca di dare delle regole che permettono di dire quando la visualizzazione passa da un abbellimento a scopo pubblicitario/divulgativo ad un abbellimento che dà fastidio e che impedisce di capire il contenuto che sta dietro la visualizzazione .

Slide 86

*Data ink* → inchiostro deputato alla visualizzazione

*Data ink ratio*: rapporto tra la quantità di inchiostro che si spende per i dati e la quantità di inchiostro che si usa per stampare l’intera visualizzazione.

Quanto è la proporzione dei pixel colorati ch effettivamente si usa per trasmettere un dato reale, un’informazione reale?

Slide 87

Se si fa una visualizzazione così: dove c’è una griglia che dà fastidio, lo sfondo è un colore troppo particolare…

Questa visualizzazione non è *misleading*, dice le cose come stanno, però tutte le informazioni che vuole trasmettere questa visualizzazione sono offuscate da tutte i “ghirigori” presenti attorno, come la griglia, lo sfondo…

Le griglie poi tendono a stare al centro della barra e inoltre creano una falsa percezione dell’area della barra stessa, contro cui il nostro cervello deve combattere.

Tufte dice che il *data ink ratio* fa vedere la proporzione dell’inchiosotro che si usa per mostrare la vera informazione che si vuole passare ai propri utenti.

Questo grafico poteva essere migliore senza *grid* (griglia).

Supposto che l’utilizzo del ***bar chart*** sia un utilizzo pensato, dal momento in cui si sta colorando le barre delle stesso colore, vuol dire che non si vuole mettere in risalto alcuna barra in particolare.

Non si vuole mettere in risalto nemmento un *trend*.

Si poteva pensare di mettere i puntini al posto delle barre, ma si sarebbe persa la percezione di area, di quantità.

Quindi devono per forza essere usate le barre.

[ Era meglio avere delle barre vuote non colorate (in basso) – slide 89.]

Slide 88

1. Tufte dice che innanzitutto bisogna ricordarsi che si vogliono mostarre i dati (non colori)

2. Bisogna massimizzare il *data ink ratio*.

3. Rimuovere il *non data ink*, ossia il *visual clutter*, il rumore visivo → ad esempio nella slide 87 il colore nelle barre era un inchiostro irrilevante, in quanto l’area era già percepita dal contorno della barra stessa.

5. Bisogna fare una visualizzazione e poi bisogna provarla, se non piace → bisogna aggiustarla → bisogna fare delle prove finché non si raggiunge una buona visualizzazione.

*Let’s see in practice* – N.B. APPLICABILE PER IL PROGETTO:

Si parla sempre di grafici a barre perché sono sempre i più semplici da capire.

Ci sono le label, c’è l’asse delle y e la legenda.

Il numero di categorie va bene, non sono più di 5 e quindi la memoria se le ricorda.

C’è la legenda ma c’è anche l’asse delle x .

È stato tolto il background che dava fastidio.

Dato che c’è la legenda perché mettere le label sull’asse x? Meglio rimuovere la legenda, è un extra rimovibile.

Anche Il bordo intorno al grafico introduce un *visual clutter*.

Il grigino dell’ombra per far percepire del 3D non fa capire bene la quantità, meglio rimuoverlo.

Il fatto di rendere grigi i testi è una scelta.

Il testo puramente nero, l’*automatic* *color* che forniscono Word e PowerPoint non è ottimalè, è migliore il grigio scuro che è ugualmente bene percepito dall’occhio umano.

Tufte dice di rimuovere anche l’asse delle y.

Il suggerimento che rende il numero palese e più chiaro, è usare numeri precisi per dare l’informazione dettagliata.

Se ci si è accorti di aver tolto troppo si può tornare indietro.

(In slide figura a DX) il fatto che non ci siano gli assi non crea disturbo e permette di confrontare meglio le diverse frequende delle onde.

Slide 89

*Visual clutter*: tutto quell’inchiostro che si mette in una visualizzazione, che di suo è già fatta bene.

*Chart junk*: uso eccessivo di effetti garfici.

In alto si ha un grafico che dà in milioni le spese della Casa Bianca e del Senato durante le campagne poltiche; sotto si ha lo stesso grafico disegnato però alla “Tufte style” e reso più semplice.

Tufte argomenta il fatto che il grafico sopra è impreciso e complica nella percezione dei numeri.

Argomenta anche che 1977 non ha la stessa base di partenza dell’82 quindi i prezzi dell’82 possono essere percepiti più alti.

Studi percettivi su utenti dicono che si usano molti *chart junk*, se ad esempio la visualizzazione si fa con un meme, gli utenti che hanno visto il meme rispetto ad utenti che hanno visto il grafico vero e proprio a volte si ricordano di più l’informazione passata (proprio grazie al meme).

Slide 90

Tufte dice di non mentire con le scale.

Non usare *chart junk*, effetti grafici troppo eccessivi.

Aggiustare in senso matematico (normalizzare) i dati: non si possono mettere nello stesso plot cose diverse → un grafico sull’età ed un grafico sull’altezza media degli individui poiché le altezza hanno un range totalmente diverso dalle età – si parla di scale completamente diverse.

Se bisogna comparare dati di tipo diverso bisogna normalizzarli.

Slide 91

Il grafico non deve essere ridondante ma autocontenuto: deve avere dei titoli che spiegano bene il contenuto.

Non usare cose *misleading*: non usare nel grafico più dimensioni di quelle che in realtà ci sono nelle variabili.

Slide 92

Il valore nel 1978 comparato al 1985.

I dati reali sono una comparazione tra 18 miglia nel 1978 e 27.5 miglia nel 1985.

Si capisce che 27.5 non è il doppio di 18, mentre 5.3 è 8 volte 0.6 → si stanno visualizzando due segmenti che danno una percezione dela lunghezza, le une diverse dalle altre.

Slide 93

Per formalizzare questa cosa Tufte introduce il *lie factor*.

Slide 94

L’effetto mostrato nel grafico rispetto all’effetto mostrato nei dati.

7.83/0.53 = 14.8 Più è alto o più è basso e più bisogna migliorare le cose.

Slide 96

Stessa cosa vale per la scala delle immagini.

Il trend generale deve essere di 45°, l’orientazione media deve essere di 45°.

Si prende l’immagine come immagine e si applica una decomposizione spettrale, si muovono le componenti in alta frequenza, si prende la curva di trend e si vede qual è la frequenza smussata.

SI ottiene così la frequenza di trend e se si ottiene 45° va bene.

Slide 97

La stessa cosa succede per le tabelle: come pulirle per fare in modo di non inserire troppo *chart junk*?

Slide 99

*Information maps* → sono visualizzaizoni

**McCundless** cerca di mappare tutte le informazioni che trova in giro per il mondo, facendo *scraping* di diversi siti, in una visualizzazione.

Cercare un mapping che porti dall’informzione astratta alla visualizzazione sul computer.

Slide 100

Per fare questo utilizza tecniche classiche di visualizzazione: va a vedere quali sono i soldi spesi, guadagnati ricevuti, donati, in diversi campi e lo fa facendo lo *scraping* da Internet.

È andato a vedere da diversi giornali quali sono i numeri legati ai dollari.

Ha visto cosa succede a questi numeri.

Slide 101

Rosa → donazioni Beige → valori mercato

Con i diversi colori indica l’origine del denaro.

I costi predetti per vari motivi.

Ciò che interessa è che ha mostrato un dato reale.

Poteva visualizzare la stessa cosa con grafici a barre o in altri modi, ma ha scelto la *treemap*.

La inventò **Brayman** per vedere l’occupazione in memoria del suo computer.

Si disegna un rettangolo, si hanno tanti valori da inserire dentro al rettngolo e si divide il rettangolo in tante aree che sono proporzionali ad un valore.

Alla fine quando si sono inseriti tutti i valori, non solo si ha la percezione di ogni singola quantità in quanto si usano delle aree ma si ha anche la percezione di quanto una singola area è diversa rispetto alle altre.

Basti vedere i soldi spesi per la guerra in Afghanistan, sono molto di più rispetto a queli spesi per la Cina ad esempio, quindi l’area è più grande.

Si ha una percezione immediata di quanto è grande un’area rispetto alle altre.

Brayman voleva capire cosa occupava lo spazio di memora del suo computer, l’occupazione delle directory in memoria.

Aveva associato un colore diverso ad ogni directory nel suo caso (magari aveva poche directory).

Il quadratino marrone dei tre mercarti legati a Internet/al commercio di computer è diviso in *Apple market*, *Google market* e *Microsoft market*.

Aree vicine sono in qualche modo connesse tra di loro e aree lontane in qualche modo disconnesse.

Tutte le aree danno il rapporto di quante sono diverse loro rispetto alle altre.

Slide 102

**Shneiderman** è partito dal quadrato (della slide precedente) e ha creato diversi tipi di ***treemap***.

Slide 103

A volte le ***treemap*** sono rappresentate come degli alberi.

Si parte da tutto il computer, da quante directory è composto…finché non si arriva alle foglie.

Si possono fare tante grosse le foglie/i rami quanto è l’occupzione in memoria.

Si può arricchire una ***treemap*** con raggi più spessi.

Slide 104

Ci sono anche le *treemap* radiali.

Slide 105

*Tidy trees*: sono *treemap* in cui le foglie si trovano alla stessa altezza.

Slide 107

SI tratta di *density plot 3D*.

Sono una sorta di istogrammi interpolati in modo da formare *density plot*.

Ogni linea colorata è un density plot.

Vengono anche detti *density area plot*, in quanto composti da aree. Danno più una percezione di quantità.

Si sono interpolate le barre dell’istogramma e si è ottenuto il ***density plot***.

Se si colora l’are sotto il ***density plot*** diventa un ***density area plot***.

Nel momento in cui si impone la *feature* di trasparenza e si vede quando i due grafici si sovrappongono, si riesce a vedere la differenza tra i grafici, ma nel momento in cui però si hanno tanti grafici, questo non è chiaro.

Si chiamano *trails* ossia ***density plot*** disposti in 3d.

Dal momento in cui percepiamo il 3D capiamo che ad esempio il grafico azzurro è più piccolo di quello verde: capiamo la prospettiva → si possono mettere i grafici più piccoli davanti e quelli più grossi in fondo.

Con questo grafico **McCundless** vuole mostare cose di cui gli Americani hanno più paura.

(Ha analizzato dei tweet.)

Slide 108

Se si volevano dare delle informazioni precise e si voleva permettere un confronto preciso, si potevano fare delle linee → ma si rischiava di avere un affollemto nel grafico .

Slide 109

Alternativamente si può usare uno ***small multiple***: si ha un solo asse, si usa la stessa scala e si vedono i diversi ***density plot*** in modo chiaro.

Si hanno diversi ***density plot*** disposti in 3D.

Slide 110

**McCundless** usa i colori in modo intelligente: colori che associamo a cose tranquille per fare vedere quelle paure che per gli Americani sono un po’ meno pericolose.

Il rosso per far vedere quelle paure relate ad omicidi o altro.

Usa tonalità uguali per indicare paure che sono legate allo stesso modo nel mondo reale.

Slide 111

Ha usato il grigio per il *millenium bug*: è stata la paura del 2000.

Oltre il 2000 si pensava che tutti i SW che contemplavano le date, non avrebbero più funzionato (poi così non è stato).

Slide 112

Azzurro e verde indicano delle paure meno importanti.

Slide 113

**McCundless** utilizza questo *frequency trail* per far vedere che tutti gli anni, in certi momenti, c’è la paura dei videogame violenti.

E fa vedere che questa cosa accade spesso a novembre ed aprile.

Slide 114

Nel 2001 si ha un buco perché le paure erano concetrate su altro (i terroristi).

Usa il rosso perché ci sono state tante morti.

Slide 115

Non utilizza l’istogramma ma il ***density lpot*** perché dà più l’idea di trend nel tempo.

Le linee mostrano il trend e mettono meglio in evidenza la presenza di picchi.

Slide 116

Simili ai *frequency trails* sono gli *streamgraphs*.

Slide 117

Uno *streamgraph* è un po’ *misleading*: per quello che vediamo noi pensiamo che siano tante aree una sovrapposta all’altra.

Gli *streamgraph* a tuti gli effetti prendono diversi ***density plot*** e li impognono uno sopra l’altro.

Dove cominciano i valori del giallo, cominciano poi i valori del verde.

Non è che cominciano insieme i valori (come accade per i *density plot*).

PROBLEMA: sono stati usati colori simili.

La vicinanza di colori verdi rende difficile la distinzione.

A volte per evitare di superimporre troppe aree e colori, si ribaltano tra di loro le aree e si passa da ciò che sono i ***monodirectional streamgraph*** agli *streamgraph* veri e propri che si creano attorno all’asse delle x ( in basso nella slide).

Questi utlimi, in alcuni ambiti, fanno vedere la crescità.

Il violetto chiaro a SX è lo stesso che si trova a DX? No, però persone che hanno difficoltà nel vederli, potrebbero percepirli come uguali.

La cosa difficile di questo tipo di grafico è che è difficile confrontare le diverse quantittà.

Si sono molto diffusi nel periodo del Covid.

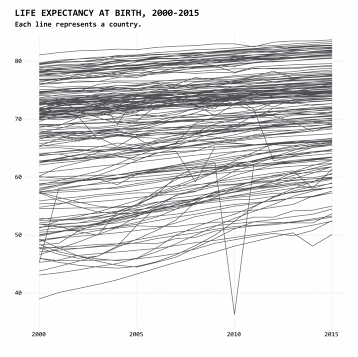
Slide 118

Se si vuole studiare l’andamento di una data quantità.

Slide 121

Effettivamente qual è il grafico che si può usare?

**Nathan Yau** (un visualizzatore) dice che uno stesso dataset si può visualizzare in tanti modi diversi.

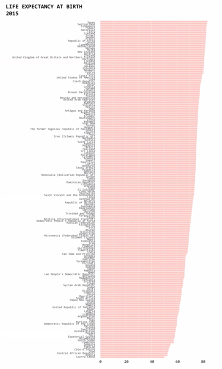


Voleva far vedere che per tutti i paesi c’è un trend in crescita.

Lo sa che non è una visualizzazione chiara avere tutte le linee disposte così.

Immagine che contiene testo

Descrizione generata automaticamente Così fa concentrare sugli specifici andamenti di ogni paese.

Alternativamente si può fare il grafico a barre dove per oni nazione si pone una barra.

Qui si ha un trend particolare, si vede che sono ordinate dal più alto al più piccolo.

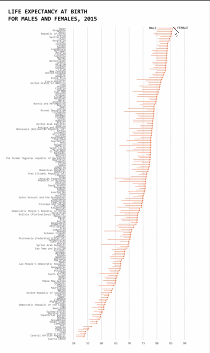
Quando le categorie però non sono ordinabili (come in questo caso) si deve dare un modo per ordinare, così da dare modo al lettore il modo di vedere subito quale paese si trova in che posizione (in base all’andamento).

Sotto la Danimarca si ha un crollo delle diverse nazioni.



Si potrebbe ordinare le barre dalla A alla Z, ma si perde l’idea del sorting.

Questa visualizzazione è bella ma non dà un’idea globale.

Mostra l’aspettativa di vita degli uomini rispetto alle donne:

c’è una freccia che va da SX a DX (la donna ha una maggior aspettativa rispetto all’uomo) o viceversa (la donna ha un’aspettativa minore dell’uomo).

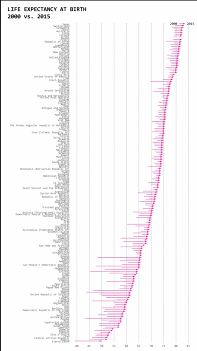
Dove non c’è la freccia è perché per quei paesi non si è riuscito a stabilire le aspettative di vita per donne e uomini.

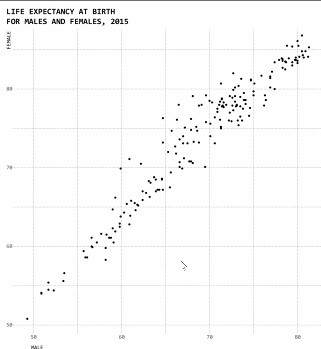
Immagine che contiene tavolo

Descrizione generata automaticamenteSi può andare a plottare uno ***small multiple*** in cui si colora come area la differenza tra uomini e donne: più è differente l’area, più le donne vivono di più rispetto agli uomini.

Si stanno usando sempre viusalizzazioni semplici ma che trasmettono informazioni diverse.

Un grafico simile con la differenza tra 2000 e 2015 è questo:





Con lo ***scatter plot***: aspettative delle donne rispetto agli uomini, ogni pallino è una nazione.

Poi si possono anche avere i:

• **Dot plot**, viene anche detto ***heatmap*** → fa vedere il cambiamento negli anni delle aspetttative di vita.

• **Violin plot**

…