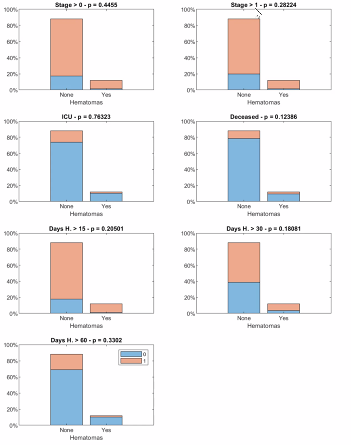
I principi della Gestalt si applicano alle visualizzazioni e possono essere sfruttati per renderle più comprensibili, per far passare meglio le informazioni.

Dobbiamo stare attenti a tutto: ciò non vuol dire rendere semplicissime le visualizzazioni, ma bisogna fare attenzione alla prossimità degli oggetti, colori, forme… un’attenzione mirata alla nostra utenza.

Esempi di bar chart sovrapposti (***stacked***)

In questa visualizzazione si normalizza in modo da avere il 100% delle persone che hanno gli ematomi (*Hematomas*).

Sono state fatte due barre: a SX la percentuale di persone con nessun tipo di ematoma e a DX la percentuale di persone con almeno un tipo di ematoma.

L’azzurro ed il rosso significano sempre Falso e Vero.

Lo ***stacked bar chart*** organizza le barre e dice com’è composta la singola barra.

Stadio 0 → no Covid

Stadio 1 → primo stadio della malattia Covid

Confrontando i due grafici iniziali si vede che le percentuali si mantengono.

Azzurro indica pazienti che non sono stati in terapia intensiva, rosso che sono stati in terapia intensiva.

Azzurro può anche indicare non morti e rosso morti.

Uno *small multiple* (figura sopra) è percettivamente semplice perché una volta spiegato il primo grafico, gli altri si riescono a capire di conseguenza.

Uno ***small multiple*** è una serie di grafici o grafici simili che utilizzano la stessa scala e gli stessi assi, consentendo loro di essere facilmente confrontati.

Utilizza più visualizzazioni per mostrare diverse partizioni di un set di dati.

Il termine è stato reso popolare da Edward Tufte.

Lo ***small multiple*** è utile perché l’utente dve solo capire la modalità del grafico.

Una volta capito il primo grafico a SX (essendo occidentali partiamo a leggere da SX), poi il ragionamento si applica a tutti gli altri grafici.

Se si usa sempre la stessa scala per tutti i sottografici, lo small mutliple non consente solo la comparazione delle singole diverse barre ma permette anche una comparazione tra grafici diversi.

Ognuno di questi ***stacked bar chart*** visualizza una matrice di confusione normalizzata rispetto al numero totale di pazienti.

La legenda in fondo si riferisce a tutti i grafici.

Non è ottimale posizionarla così, sarebbe meglio sistemarla in modo separato, in questo modo:

Messa esternamente a tutti gli *small multiple* rende più chiaro all’utente che quella leganda si riferisce a tutti i grafici.

Per i **principi della Gestalt** quando si guarda la figura si vede che ci sono dei grafici diversi gli uni dagli altri, quindi sia che la legenda la si ponga all’inzio che alla fine della figura, verrà sempre vista.

Non si vuole assolutamente porre la legenda in mezzo in quanto interrompe la visualizzazioe dei grafici.

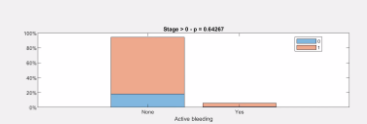
L’ideale è spostare la legenda a SX o a DX.

La rivista per cui è stato richiesto questo ***small multiple*** ha chiesto l’uso di 0 ed 1 (in legenda); ma era meglio mettere True e False.

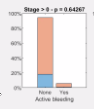
Nella ***caption*** poi si specifica che 1 indica la presenza del sintomo.

L’interattività va bene.

Se si pone la legenda in alto a SX, nella *caption* si dice che la legenda che c’è nel plot in alto a SX si applica a tutti i subplot.

Questo grafico che vediamo più largo in qualche modo modifica la percezione delle differenza tra pazienti che hanno sanguinamenti attivi (*active blooding*) e non.

Percettivamente non sembrano tanti gli azzurri nonostante siano il 90%.

Se la si stringe, percettivamente si fa notare quanto è la maggior presenza di pazienti senza sanguinamenti attivi (azzurri).

Più si allarga l’immagine più si ha una minore percezione dell’avere una barra snella → per noi barra snella vuol dire un elevato valore.

Quanto è quindi il rapporto altezza/larghezza?

Se si hanno le stesse quantità, ad esempio, percentuali su asse y e anche asse x, si usa la stessa scala.

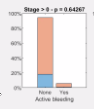
Ma se le unità di misura dei due assi sono diverse, come si decide la scala?

Esempio

Si vogliono visualizzare due classi: i rapporti dell’altezza e del peso della 5° elementare e della 3° elementare.

Sull’asse x si ha un rapporto altezza/peso e anche sull’asse y.

Se si prendono tutti i bambini e si calcola il rapporto altezza/peso per valutare l’andamento del bambino, in questo caso si ha la stessa unità di misura su entrambi gli assi.

In questo caso sull’asse x si hanno due variabili categoriche (si o no) e sull’asse y percentuali.

Studi percettivi dicono che si dovrebbe fare un grafico che faccia in modo che la somma degli angoli di tutte le linee che sono nel grafico è in media di 45 gradi (ossia l’orientazione di tutte le linee).

Se riempire l’altezza e la larghezza della pagina crea grafici troppo magri, si divide in più sottografici, se invece si creano grafici grossi, si restringono.

In ogni caso si agisce sulla base di cosa si vuole far percepire.

Ad esempio se si vuol fare percepire tanti valori, si restringe.

Slide 58

Riusciamo a perceprire l’area, ma la differenza altezza e lunghezza si riesce a percepire meglio soprattutto quando hanno un asse di riferimento comune.

← (Si riesce a percepire la diversa altezza delle due barre, in quanto hanno lo stesso asse di riferimento.)

Se non hanno un comune asse di riferimento sono difficilmente comparabili.

Le aree sono attributi preattentitvi comparabili, in cui conosciamo le differenze, ma non conosciamo le differenza tra aree molto simili.

Comu utilizzare quindi i ***pie chart*** che sfruttano fortemente le aree?

Esempio in slide

Nel ***pie chart*** si hanno tutti i *competitor*.

La compagnia D ha la maggior parte del mercato.

La compagnia G va sicuramente meglio della F, della H, della C, della B e della E.

Ma per quanto riguarda la compagnia A, noi G abbiamo più mercato della A oppure no?

Percettivamente può sembrare che G è più grande.

Slide 61

Ora trasformando il ***pie chart*** in 3D, la G rimane ancora più grande?

Si tracciano due linee rosse sovrapponendole sulle linee gialle per capire che G è più grande,.

Il ***pie chart*** è bello, ma va usato con attenzione.

Va usato quando non si vuole dare in assoluto il valore della quantità, ma si vuole fare vedere solo delle differenze, si hanno poche categorie e le aree sono tra loro molto differenti.

Se si hanno grosse differenze tra le aree e poche aree da visualizzare allora va bene, ma se si hanno già aree simili e tante, non è ottimale, meglio utilizzare un ***bar chart***.

NON usare il 3D perché fa cambiare la percezione di come vediamo le aree!

Aree che prima sembravano più piccole ci sembrano più grosse ed aree che ci sembravano più grosse ci sembrano più piccole.

( Non si poteva usare il pie chart, ma ordinando i valori %?

Sì ma deve essere specificato nella didascalia come sono stati ordinati i dati oppure si mettono i numeri su ogni area → però poi si impone al lettore di ragionare di più (non è ottimale, non gli è subito chiara la visualizzazione), meglio a quel punto un ***bar chart*** – slide seguente.

Slide 62

Si poteva quindi fare un grafico a barre, dove si possono plottare le stesse quantità che si erano plottate sul ***pie chart***.

Slide 63

Si usa il ***pie chart*** quando si hanno poche aree.

Si hanno 4 aree ed è una dimostrazione del perché il 3D non va usato.

Quando si progetta in 3D si prende un punto di vista che può essere un asse isometrico.

Si hanno una sorta di assunzioni in cui non è mai definito il punto di vista 3D delle visualizzazioni.

Non si caspice la distorsione data dal 3D.

Non si riesce a capire se l’area azzurina sia più grande o più piccola di quella lilla.

Slide 64

Quella viola percettivamente è più grande. Ma cambiando il punto di vista di come si vede ciascuna area, può far cambiare la percezione che si ha sulle aree.

Attenzione al 3D perché falsa molto le cose, ci fa sembrare le aree più piccole o più grandi quando in realtà sono più grandi o più piccole, rispettivamente.

Slide 65

Un esempio di pessima visualizzazione.

Slide 66

Esempio di ottima visualizzazione, in cui viene utilizzato il 3D in modo intelligente.

Dipende dal tipo di dati che si intende visualizzare.

In questo video l’autore è andato a visualizzare le parole di suo figlio, ha voluto visualizzare la crescita nella capacità del parlato di suo figlio.

Quando c’è un’onda vuole dire che il bambino è stato preso in braccio.

Qui la visualizzazione 3D è fatta bene perché dà idea del movimento delle persone o comunque di cosa sta succedendo all’interno della casa.

Slide 67

Questa visualizzazione in 3D fa vedere quante volte il bambino è passato in dati punti della casa.

Per dire questo picco indica che il bambino ha detto tante volte la parola “acqua” in cucina.

Slide 68

CATEGORIE

Questa visualizzazione è fatta in modo intelligente per quanto riguarda alcuni principi della Gestalt.

Principio della prossimità: sono raggruppate barre relative a gruppi di uno stesso tipo.

Tutte le altre barre sono dolci di altre cmpagnie.

Si usa la prossimità per separare i mesi gli uni dagli altri.

Anche all’interno di ogni mese si usa la prossimità per differenziare la compagnia che si ha sotto osservazione dalle altre compagnie.

Viene utilizzato il principio di similarità: colori diversi in diversi mesi, barre che appartengono a gruppi diversi (percettivamente perché non sono vicine) vengono fatte vedere all’utente come appartenenti allo stesso gruppo in quanto si utilizza lo stesso colore.

Esempio: viola nel mese di luglio indica lo stesso dolce nel mese di agosto.

Si utilizza quindi il principio del ***grouping***, dato dallla prossimità, ed il principio del colore.

Qui si usano però troppe categorie. L’utente non riesce a ricordarsele.

Colori simili (per dire blu e viola) poi possono far pensare che si tratti dello stesso gruppo.

Il fatto che ci siano troppe categorie impedisce al lettore di fare un’associazione che rimanga nel cervello, ossia: vedo il viola so subito che si riferisce a quel dolce.

Se la memoria di lavoro cancella queste associazioni colore-dolce, vuol dire che bisogna andar a rivedere la legenda.

Come si ha una memoria piccola, si ha anche un punto di vista piccolo:

per vedere tutte le categorie bisogna far scorrere lo sguardo dall’alto verso il basso le categorie, in più arrivando alla fine della legenda ci si è già scordati delle categorie sopra, a cosa si riferiscono quei colori.

Inoltre i colori non sono scelti in modo ottimale, sono troppo simili e possono essere confusi tra di loro.

Questa visualizzazione usa dei bar chart raggruppati che sono migliori (degli ***stacked bar chart***).

Si usa una visualizzazione di questo tipo in base all’utenza che si ha.

Tutto va valutato in base alla *audience*.

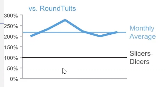
Slide 69

Si poteva usare uno ***small multiple*** con tanti plot, in cui in ogni subplot si vedono due barre nel mese di luglio, …

I lettori dato che sanno pensare e che vogliono fare il confronto, si può considereare l’aglrtimo “S*licersDicers*” come se fosse il 100%.

Se lo scopo è il confronto rispetto ad una determinata categoria, si utilizza quella categoria e si dividono i valori dell’asse per il valore della categoria, così si fanno subito le comparazioni tra una categoria e tutte le altre categorie.

In ogni ***small multiple*** si fa un confronto 1:1.

Principio di continuità: connttenedo i diversi punti nei diversi mesi si dà l’idea di un *trend*.

Si fanno vedere meglio gli andamenti collegando con una linea.

Per sapere in media quanto un dolce (singolo subplot) supera gli altri (altri subplot) si usa la linea di trend.

Se si fosse voluto vedere anchel’andamento degli *SlicersDicers*, non lo si considerava il 100% ogni mese, ma si faceva uno ***small multiple*** in cui si faceva vedere l’andamento dello *SlicersDicers* ed in un altro il dolce.

Si è fatto in modo che la retta di *SlicersDicers* servisse come asse di comparazione.

Si può quindi vedere quali sono gli algoritmi che funzionano bene in base ad un parametro o meno.

Slide 71

Con le visualizzazioni si vuole fare passare un concetto, senza mentire, ossia senza falsare la visualizzazione.

L’importante è che si faccia passare un concetto senza mentire (secondo Tufte).

Il grafico in slide fa vedere come fino a marzo del 2013 c’è stata una campagna per assumere più infermiere.

Se si cerca di dare un valore alle quantità che si stanno percependo, si può vedere un rapporto che è 3.5 (oro) inch contro uno 0.47 inch (azzurro).

Si può notare che il grafico è un grafico che utilizza quello che Tufte chiama il *chart junk*, ossia si fa una sorta di decorazione per dare l’idea che si compone un grafico a barre usando (in questo caso) degli omini.

Ogni omino rappresenta un blocco di infermiere assunte.

L’uso del ***bar chart*** dà una maggior percezione, crea una sorta di area (che abbiamo detto essere percepita di più rispetto ai puntini).

La diversità netta tra aree è molto più percepita rispetto alla dimensione tra puntini.

I puntini sono il numero reale di assunzioni.

Quindi il rapporto tra le aree che si vede è 0.47/3.5.

Il rapporto che quindi si vede è un 13%.

Ciò che si sta facendo in realtà e che si sta provando a passare alla propria utenza è la percentuale di incremento: si sta dicendo se prima eravamo a 0.47 ora siamo a 3.5; quindi la variazione è il delta che c’è tra la barra di SX e di DX.

Quindi l’incremento totale che si vede tra le aree rispetto al 2008 è del 644%.

Slide 70

Prendendo però i numeri reali (i puntini):

se si fa 47500-43147=4353/43147=0,100 → l’incremento reale è del 10%.

Quindi se effettivamente si vede cosa sta succedendo nel grafico si può vedere che si ha un incremento del 91% rispetto allo stato attuale.

Ciò che si percepisce è un 644% però.

Quindi il grafico sta mostrando un incremento del 10% ma percettivamente mostra un incremento nettamente maggiore (644%).

Se si dovesse fare il grafico realre bisogna avere gli stessi rapporti che percettivamente si vedono nel grafico precedente.

Si induce chi vede il grafico a percepire un incremento molto maggiore di quanto in realtà è.

Slide 72

Ciò che dice Tufte è di non ingannare: i grafici devono essere il più possibile realistici.

Inoltre Tufte critica lo “spreco” di inchiostro per far visualizzare le informazioni.

Ciò che lui chiama DATA INK. Vuole che si massimizzi il *data ink ratio*.

*Data ink ratio* vuol dire che bisogna utilizzare la minor quantità di inchiostro comparata alla quantità di incuostro effettivamente utilizzata per visualizzare i dati.

Si potrebbe sostituire la barra con un pallino (in corrispondenza del 44) → si sarebbe espressa la stessa quantità però con meno inchiostro.

Tufte dice di ridurre l’inchiosto se non puramente utilizzato per visualizzare i dati, quindi ridurre i colori.

Slide 73

Esempi di visualizzazioni inutili.

Gli stessi dati possono essere scritti direttamente nel testo.

Il grafico a SX rende difficile la visualizzazione.

Slide 74

SI tratta di uno ***small multple*** in cui c’è un grafico vuoto.

Si poteva evitare di mettere il 4° grafico e dire che per quell’esperimento si aveva tutto a zero.

Slide 75

Altro esempio in cui si mente.

C’è l’asse x che parte da 0 e in questo non si sta mentendo.

È mostrato il numero di scuole pubbliche ed il numero di scuole private.

Si può mettere in una griglia in cima di tutte le barre per fare vedere esattemente a che punto arriva ciascuna barra.

E’ un po’ inutile avere quelle linee in corrispondenza del centinaio.

Se ci si concentra sul numero di scuole private a partire dal 1930 in realtà è cresciuto tantissimo.

Slide 76

Se si va a vedere questo grafico la prima cosa che si fa è dire che si sta mentendo.

Sì, in realtà si sta mentendo, ma ciò che non interssa è far vedere l’incremento assoluto, ma solamente il trend; non si vuol far vedere la differenza assoluta tra l’inzio e la fine.

Mettendo i singoli puntini (non barre) ci si può permettere di falsare gli assi, poiché se si usano i punti questi permettono di stimolare l’attributo preattentivo di posizione → indica cosa è di più e cosa di meno.

Se si fosse messo un grafico a barre si stimolava l’utente a vedere la quantità che c’era tra lo 0 e la posizione del puntino.

Slide 78

SI vede un altro esempio di menzogna come le chiama Tufte.

Le barre si sostituiscono con i puntini.

Slide 79

Fa vedere le peggiori visualizzazioni mai realizzate.

Slide 81

A SX si hanno scale diverse.

L’infografica di DX è stata considerata come l’infografica peggiore mai fatta e ha problemi con i colori.

Si può pensare che il cerchio di DX (viola) corrisponda al cerchio piccolino di SX in quanto dello stesso colore.

Per noi l’ordine è più importante dell’apparenza di colore, quindi all’inzio si può pensare che si voglia comparare il cerchio rosa a SX con il cerchio viola a DX.

Si tratta di due scale non comparabili.

Non si capisce le due aree in base a cosa sono state fatte.

In realtà la comparazione che si sta facendo percettivamente è fatta tra le aree.

Poi ci sono troppe categorie e troppi colori.

Slide 82

Altra infografica fatta male.

Si tratta di un ***radar plot*** in uno ***stacked donut chart***.

Dai critici è stato considerato come una delle peggiori visualizzazioni in quanto non riesce a far capire cosa vuole mostrare, cosa vuole trasmettere.

Questa visualizzazione non è ben chiara e non è stata ben spiegata.

Va bene fare le visualizzazioni complesse, ma devono avere un senso!

Slide 83

Si tratta di ***donut chart***.

Slide 84

È l’infografica della Coppa del Mondo del 2010 in cui sono visualizzati tutti i tornamenti di gioco e quali sono i possibili finali.

Ciò che si voleva far vedere è quale squadra andrà a giocare con quale altra squadra.

Ache la grandezza di ogni cerchio manifesta il numero di coppe vinte dalla squadra ed il numero di volte che ha partecipato.

C’è una buona spiegazione della grandezza di ogni cerchio.

I cerchi marcati con i punti sono quei paesi per cui ci si aspettava una buona *performance*.

Non c’è alcun tipo di di *marker* o colore che è stato utilizzato inutilmente.

Ogni colore e *marker* è stato utilizzato per un motivo.

Lo stile grafico può rendere una visualizzazione pulita (senza che renda difficile la sua lettura per un utente).