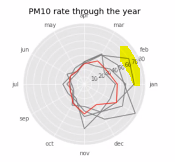
E’ un ***radar plot*** che consente di mettere in evidenza una sequenza di eventi, solitamente descritti temporalmente.

Si supponga di descrivere un evento nel tempo e che si abbia un dataset per cui per ogni mese si ha il numero di promossi in quella sessione, ad esempio.

Si ha un’asse temporale su cui si plotta una quantità.

I raggi del radar plot sono i valori dell’asse x, in ogni istante.

Le linee grigie esprimono i valori di NO2 nei vari anni.

[Linea gialla] si vede che Gennaio e Febbraio hanno lo stesso valore di PM10 in quell’anno.

[ Questi radar plot non sono precisi, sono da sistemare, sono di un gruppo di studenti. ]

Il radar plot consente di far vedere una ripetizione di eventi.

Gli attributi preattentivi (linee, orientazione, punti, colori e forme) permettono una migliore stima delle quantità.

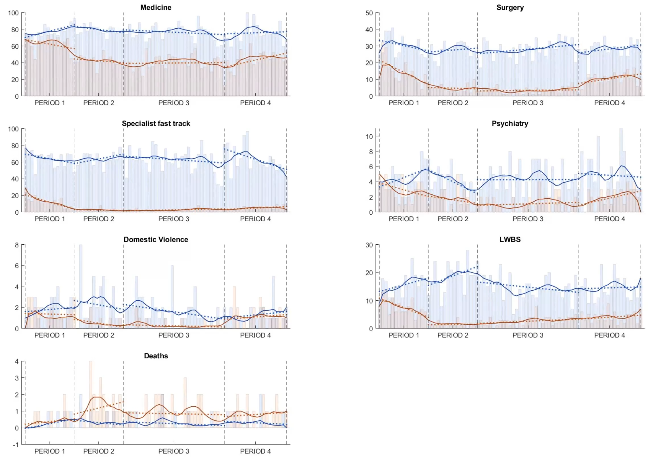
Tendenzialmente l’orientazione più che per riconoscere oggetti diversi serve per vedere ciò che sale e ciò che discende.

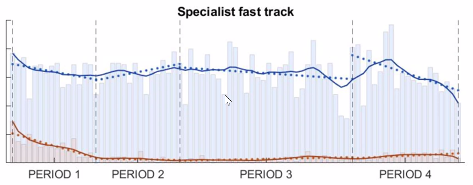
Linee oblique vengono associate a linee che salgono o che scendono.

Per noi lo scendere o il salire di una retta viene associato alla sinistra e alla destra.

Se a sinistra è più in basso e a destra più in alto, si sta salendo.

Esempio





Sono stati confrontati gli accessi al pronto soccorso.

Si tratta di una media del 2019 rispetto al 2020.

La casualità porta ad avere quasi sempre la stessa media.

È stato usato un ***bar chart*** perché dà meglio la dimensione degli ingressi, le aree sottolineano meglio la quantità.

Invece di usare dei puntini o linee per ogni giorno, si sono usate le barre.

Usando una linea si stanno legando i punti tra di loro e quindi si sta dicendo che c’è una connessione tra il punto precedente ed il successivo, si sta dando un *continuum*.

Il trend generale veniva “pasticciato” e non si vedeva usando una linea.

L’occorrenza del singoo giorno è data da una casualità degli eventi.

Le barre sotto per sconettere un giorno rispetto all’altro.

Sono state sfocate per non dare importanza al singolo numero di occorrenze nel giorno.

In giorni in cui si ha un picco è dovuto solo ad una casualità.

SI è imposto ai dati (alle barre) dei singoli giorni, una linea (ottenuta dalla media nei gironi successivi).

Dal punto si prendono i due giorni successivi e precedenti e si fa la media.

Per il bordo a SX si usano i giorni successivi, per il bordo a DX i giorni precedenti.

In modo da fare vedere l’andamente del periodo.

La linea tratteggiata è la linea di trend e fa capire l’andamento.

2019 indicato in rosso.

2020 indicato in blu.

I colori manifestano una sensazione, un sentimento.

Generalmente il rosso indica il pericolo.

Il blu/azzurro è un colore neutro, è un colore associato alla matematica e si usa spesso nei vari grafici.

Viene utilizzato in molti plot scientifici, anche perché il blu è un colore che non dà molto fastidio al nostro occhio, è delicato.

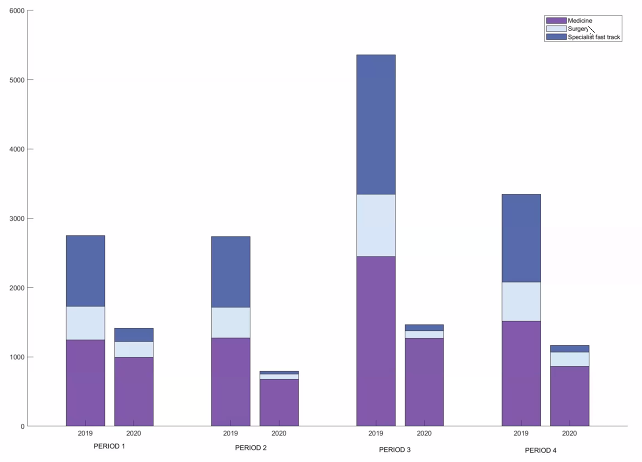
Si sono utilizzati quindi dei ***line plot*** non per il susseguirsi degli eventi nei giorni, ma per far visualizzare la media nei vari giorni.

C’è anche la **linea di trend** nei vari periodi.

I colori sono stati utilizzati per distinguere 2019 e 2020.

Per tutte le specialità si è visto che c’è stato un crollo degli accessi al pronto soccorso.

Esempio



Sono gli accessi nei tre reparti nei due anni.

Si tratta di un ***grouped bar chart***.

E per ogni periodo si sono messi insieme 2019 e 2020.

In ogni periodo si hanno delle barre sovrapposte, per ogni barra si fa vedere la composizione della barra.

Si può fare il confronto sulla barra e si vede quale è la classe che compone di più la barra.

La cosa negativa degli ***stacked bar chart*** è che non si riesce a confrontare bene, a dire quanto è grande, per esempio la classe blu (“specialistica” in legenda) rispetto agli altri periodi,.

Gli *stacked bar chart* non riescono a far confrontare bene le singole classi.

Non tutte le barre hanno un comune inizio.

Gli *stacked bar chart* consentono un confrono assoluto tra le barre prese in toto (con tutte le classi).

Però si fa fatica a confrontare la classe blu (“specialistica”) del periodo 1 del 2020 con la “specialistica” del 2020 del secondo periodo.

Si sono stampati sul’asse y il numero di ingressi medio giornaliero in assoluto.

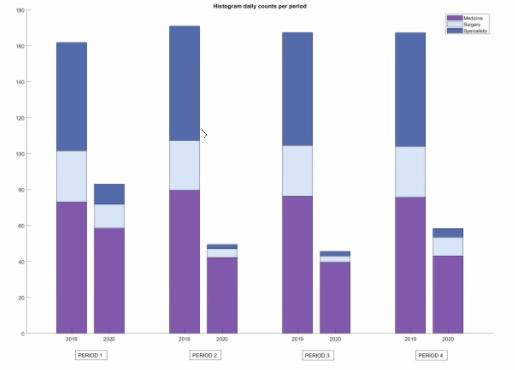
Nel periodo 3 nel 2019 si aveva un ingresso medio di 5400, in TOTALE; per il 2020 di 2004 → si sta facendo quindi la conta assoluta per quel periodo.

Attenzione a ciò che si mostra: se non si dice che si sta facendo la conta assoluta e che nel periodo 3 siano stati considerati più giorni, si potrebbe pensare che nel periodo 3 c’è un numero maggiore di accessi al pronto soccorso, quando in realtà non è così.

( È solo che sono stati considerati più giorni in quel periodo (il terzo). )

Questa cosa dobbiamo farla vedere con la visualizzazione, come si fa?

Si normalizza per numero di giorni, si fa la media giornaliera:



Prima si stava “ingannando” il lettore, si stava facendo pensare che si lavorava di più a marzo (periodo 3).

Quando in realtà dal secondo grafico (questo ↑) si può vedere che non è così.

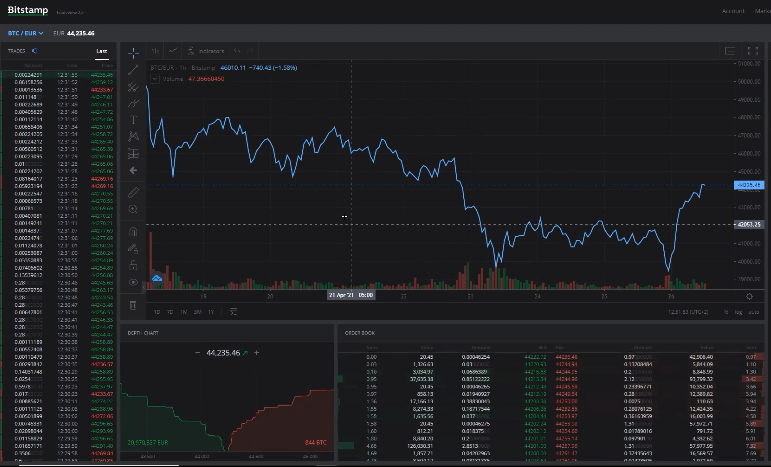
Tufte ha formalizzato questo inganno (il trattamento dei numeri che si fa prima di visualizzare e anche come si visualizza).

Ciò che si vuol far vedere con questa visulizzazione è come i ***line plot*** (quello su “*Specialist fast track*”) permettono di far vedere gli andamenti nel tempo rispetto ad una variabile che è consecutiva (sull’asse x).

Lo ***stacked bar chart*** invece può far vedere anche una variabile per cui i valori non sono consecutivi.

I ***bar chart*** sono slegati dall’asse, mentre i ***line chart*** uniscono.

**DASHBOARD**: *bitstamp.net*



Al centro viene mostato l’elemento più importante, ossia l’andamento del Bitcoin.

Il rosso indica gli ordini di vendita ed il verde gli ordini di acquisto.

Il brutto di questa dashboard è che dà per scontato che il lettore sappia già cosa vogliano dire i colori, quando in realtà non è stato indicato da nessuna parte il loro significato.

**DEPTH CHART**

Sull’asse delle x il prezzo e sull’asse delle y il volume.

Al centro, quando si incontrano le due linee, è dove si incontrano domanda e offerta e avviene quindi l’ordine (di acquisto o di vendita).

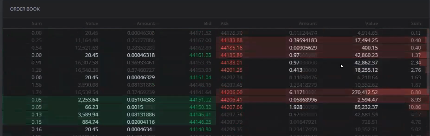
<https://www.bitstamp.net/market/tradeview>

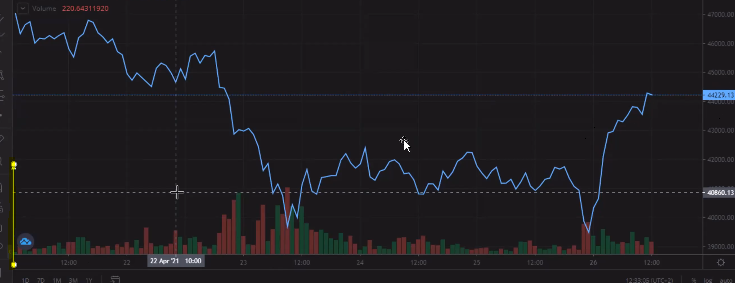
Quando i due grafici (verde e rosso) sono separati è perché la gente chiede troppo oppure offre troppo poco, quindi non avviene alcun ordine.

È stato mantenuto il rosso per le vendite, mentre il verde è associato agli acquisti.

Usare gli stessi colori per diversi grafici, che non c’entrano nulla l’uno con l’altro, non è una buona idea:

per dire il lettore se si studia una legenda di un grafico in cui il blu indica i guariti rosso i morti e verde i dimessi, vedendo un altro grafico con questi tre colori, associa di nuovo questi colori a guariti, morti e dimessi. → Ma non è corretto, questa volta i colori si riferiscono ad altro.

Si possono vedere le formattazioni condizionali: quando si superano certi valori, si evidenziano di verde o di rosso a seconda della categoria che si osserva.

Esistono i multiple axis:

In questa dashboard se ne poteva aggiungere uno a SX che indicava i valori dei ***bar chart***, mentre a DX (come già c’è in azzurro) indica i diversi valori del ***line plot***.

Si tratta di grafici totalmente scorrelati tra di loro e che vengono sovrapposti.

Invece del nero come background o altri colori, l’ideale sarebbe utilizzare un grigio chiaro per coloro che hanno problemi come la dislessia…

Slide 27

A seconda della cultura, i colori rappresentano cose diverse.

Ad esempio per la cultura occidentale il viola rappresenta distinzione, mentre per altre culture è un colore negativo.

Attenzione ai colori saturati che stimolano i coni ed i bastoncelli e attirano l’attenzione.

Se si vogliono utilizzare colori saturati, devono essere però utilizzati su elementi su cui si vuole attirare l’attenzione.

Quando non si vuole dare particolare attenzione a delle informazioni, l’ideale è utilizzare dei colori naturali, non troppo saturati, ossia questi:  e non questi (assai saturati): 

Perché in natura i colori saturati non esistono.

Inoltre i colori saturati se utilizzati in grande quantità danno fastidio, sono utili quando se ne utilizza uno solo.

Quando ad esempio in un bar chart, la barra saturata è una sola.

Si utilizzano i colori saturati quando vogliamo mettere in evidenza UNA informazione e staccarla dal resto.

Slide 28

Se niente è più importante di altro, se non si vuole far risaltare nulla rispetto al resto, si utilizzano tutti i colori non saturati.

Attenzione che più gli oggetti sono più piccolini più il colore diventa ininfluente.

In uno ***scatter plot***, utilizzando i puntini, è inutile utilizzare i colori, perché non si distinguerebbero.

Quindi si potrebbe pensare di utilizzare dei colori saturati nettamente distinguibili, ma si rischierebbe di dar fastidio.

Il problema non c’è se si utilizza un ***bar chart***, la distinzione dei colori è ben evidente.

Bisogna fare attenzione ad utilizzare dei colori che anche i daltonici possano vedere.

In questo c’è in aiuto *colorbrewer*.

“*Nature of your data*”:

• Sequenziale → se si vuol far vedere una scala di colori (ad esempio il numero di morti Covid mostrati su una mappa), si mostrano delle informazioni consecutive, sequenziali, ordinabili.

Se si visualizzano delle quantità, sono ordinabili.

Con l’opzione “***colorblind safe***” selezionata non si hanno problemi, in quanto i daltonici riescono a distinguere l’intensità del colore fornito da *colorbrewer*.

• Divergente → se si passa a scale divergenti, ossia ordinali, che in qualche modo esprimono valori divergenti tra di loro (tipo negativo e positivo) si fa passare il colore da meno saturato al più saturato e poi si usa un altro colore.

• Qualitative → si hanno classi che sono completamente sconnesse tra di loro, non comparabili, quindi categoriche.

Slide 29

Nel grafico sotto, si può vedere come un colore anche non saturato (l’azzurro in questo caso) può essere utilizzato per mettere in evidenza un’informazione che interessa e staccarla da altre informazioni che sono meno importanti e che non si vogliono mettere in risalto (si utilizza il grigio solitamente – nei paesi occidentali).

Slide 30

I colori servono anche per mettere in evidenza momenti/elementi particolari.

I colori ad intensità molto alta per mettere in evidenza zone che comunque si vogliono far vedere ma che momentaneamente non interessano.

Si rendono più saturate zone che si vogliono far vedere.

Si tratta di uno ***stacked (sovrapposto) density plot*** che mostra aree geografiche diverse e la povertà in aree geografiche diverse.

Si hanno i dollari per giorno sull’asse x e sull’asse y il numero di persone che accumulano quei dollari per giorno.

Sono stati messi in evidenza determinati paesi asiatici.

Come lo ***stacked bar chart***, lo ***stacked density plot*** permette poco un confronto tra diverse nazioni (a meno che non ci siano grandi differenze – si vede bene che l’India è molto più povera di noi).

Si fa a fatica a confrontare valori diversi della Germania, ad esempio.

Questo grafico permette maggiormente di vedere l’andamento globale.

Il ***density plot*** è a tutti gli effetti un’istogramma un po’ smussato, si fanno medie consecutive delle barre di un istogramma.

L’istogramma spiega la distibuzione di una variabile.

Come viene fatto un density plot? Si decidono quali sono le nazioni da posizionare sull’asse x.

Fa vedere come le diverse nazioni contribuiscono.