

Master Degree in Computer Science

Scientific and Large Data Visualization

Francesco Caprari:580154
Academic Year: 2023-2024

Contents

1	Intro	2
1.1	Information Visualization	2
1.1.1	Perchè la visualizzazione?	3
1.2	Fundamentals of 3D Computer Graphics	4
1.3	Scientific Visualization	4
2	Information Visualization	6
2.1	Data	6
2.2	Tipi di attributi	6
2.3	Semantica degli attributi	7
2.4	Grafici Fondamentali (Bivariate data)	7
2.4.1	Bar Charts	7
2.4.2	Histograms	8
2.4.3	Pie Charts	8
2.4.4	Donut Charts	9
2.4.5	Scatter plots	10
2.4.6	Slope Charts	10
2.4.7	Line Charts	11
2.4.8	Area Charts	11
2.4.9	Chropleth maps	11
2.4.10	Symbol maps	12
2.5	Grafici Fondamentali (Attributi Multipli)	13
2.5.1	Stacked Bar Charts	13
2.5.2	Grouped Bar Charts	13
2.5.3	Stacked Line Charts	14
2.5.4	Line Charts Series	14
2.5.5	Bubble Charts	15
2.5.6	Small multiples	15
2.5.7	Symbol Maps	15
2.6	Grafici Fondamentali (Attributi Gerarchici)	16
2.6.1	Sunburst Charts	16
2.6.2	Treempas	16

1 Intro

Information Visualization

- Quando la rappresentazione visiva non è ovvia...
- Percezione visiva, migliori pratiche, dati multidimensionali, grafi e reti

3D Computer Graphics

- Rappresentazione in 3D (mesh poligonali)
- Trasformazione dei dati 3D in immagini generate al computer: rendering, illuminazione, texturizzazione

Scientific Visualization

- Illustrazione grafica dei dati scientifici per estrapolare informazioni sui fenomeni

1.1 Information Visualization

Consiste nella Trasformazione dei dati per estrarre informazioni utili, ci sono diverse definizioni:

- La visualizzazione delle informazioni è l'uso di rappresentazioni visive, computerizzate, interattive, di dati astratti per amplificare la cognizione [S. T. Card, 1999].
- Visualizzare significa rendere visibili e comprensibili certi fenomeni e parti della realtà; molti di questi fenomeni non sono naturalmente accessibili all'occhio nudo, e molti di essi non sono neanche di natura visiva [J. Costa, 1998].
- Visualizzazione delle informazioni vs visualizzazione scientifica: Nella visualizzazione scientifica esiste una relazione naturale tra ciò che viene rappresentato e la sua rappresentazione, mentre nella visualizzazione delle informazioni la relazione è convenzionale.

Ci sono due prospettive diverse nell'interpretare l'Information Visualization:

- **Visualizzazione come scienza applicata:** Il valore di una buona visualizzazione consiste nel permetterci di individuare pattern nei dati e quindi la scienza della percezione dei pattern può fornire una base per decisioni di design [C. Ware, Information Visualization. Perception for design, 2013].
- **Visualizzazione come tecnologia:** Gli infografici sono uno strumento visivo per la comunicazione, la comprensione e l'analisi [A. Cairo, The functional art, 2013]. Le limitazioni funzionali formano: come il design di un oggetto tecnologico deve dipendere dal compito che dovrebbe aiutare la forma grafica dovrebbe essere vincolata dalle funzioni della tua presentazione.

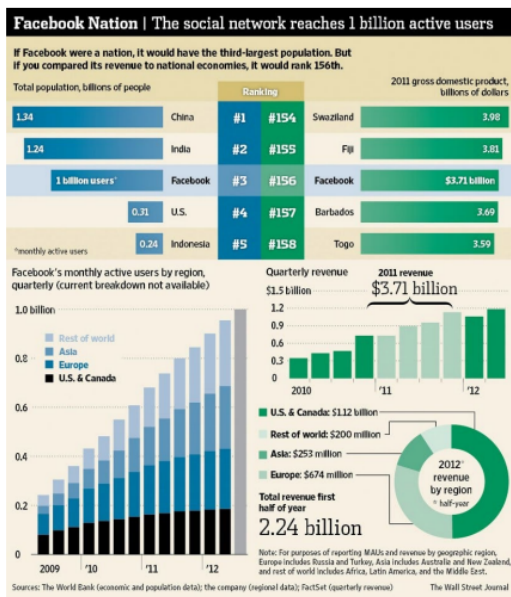


Figure 1: Visualizzazione dati di Netflix

1.1.1 Perché la visualizzazione?

Le visualizzazioni esplicative sono strumenti per presentare informazioni, comunicare dati e messaggi, spiegare qualcosa a qualcun altro, ha tre scopi principali: Explanation, confirmation e exploration.

Siamo una specie visiva. Il sistema visivo umano è molto bravo nell'identificare e analizzare i pattern. Le visualizzazioni sono come artefatti cognitivi (strumenti che gli esseri umani hanno costruito per aiutare a pensare meglio, ad esempio l'abaco). La visualizzazione ha a che fare con la cognizione distribuita (il nostro sistema cognitivo non è esclusivamente composto dal nostro cervello, mente e sensori, ma anche dall'ambiente intorno a noi, che utilizziamo per memorizzare e manipolare informazioni). Le visualizzazioni permettono il trattamento parallelo delle informazioni, anziché sequenziale. Le visualizzazioni esplorative sono strumenti per i lettori per analizzare ciò che viene loro presentato. Molto spesso, l'esito dell'analisi esplorativa non è solo la risposta alle domande originali, ma la generazione di nuove domande.

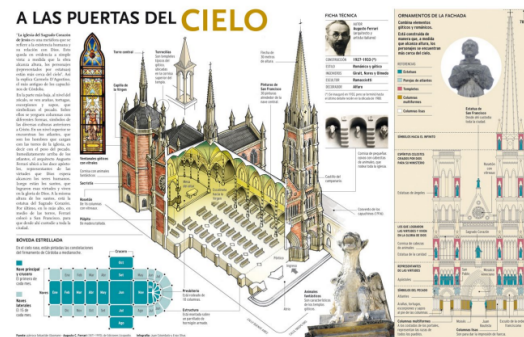


Figure 2: Juan Colombata & Enzo Oliva La Voz del Interior

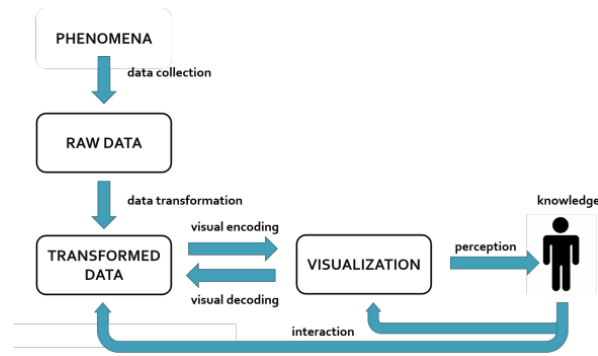


Figure 3: Visualization pipeline

1.2 Fundamentals of 3D Computer Graphics

Animazioni e serious gaming, Progettazione assistita da computer (CAD) e modellazione di prodotti, Patrimonio culturale, Archeologia, Architettura, Fabbricazione digitale e stampa 3D, Biologia e monitoraggio ambientale, Medicina e sistemi di eHealth, Moda del futuro, Sicurezza e difesa...



Figure 4: Esempi di utilizzo dei modelli 3D

1.3 Scientific Visualization

Tecniche informatiche per la generazione di rappresentazioni visuali interattive di dati spazio-temporali acquisiti o simulati (collegamento naturale con il mondo 3D+tempo in cui viviamo). Le illustrazioni a la comunicazione visuale della conoscenza è parte della nostra storia, anche oggi l'incremento della quantità dei dati da analizzare è necessario un approccio grafico per la visualizzazione. La Visualizzazione Scientifica si applica in diversi campi come l'Ingegneria, la medicina E la Scienza. La Visualizzazione Scientifica riguarda anche la definizione di algoritmi efficienti per la manipolazione interattiva ed esplorazione dei dati e delle loro caratteristiche.

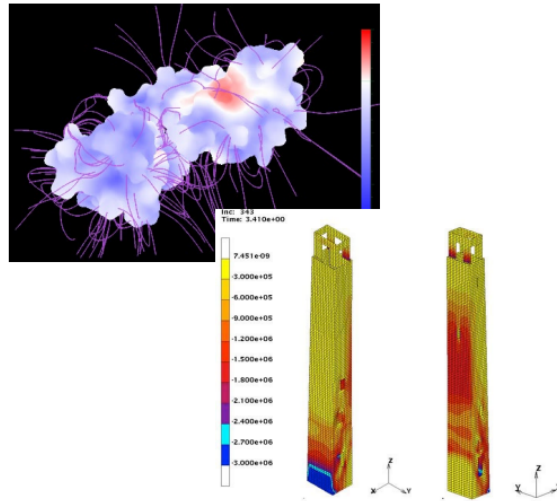


Figure 5: Esempi di Visualizzazione Scientifica

2 Information Visualization

La visualizzazione delle informazioni è un modo per comunicare significativamente i dati e aiutare le persone a dare un senso a grandi quantità di informazioni. Le visualizzazioni come artefatti cognitivi che consentono il trattamento parallelo delle informazioni. Tre scopi principali della visualizzazione:

- **Explanatory:** Le visualizzazioni esplicative sono strumenti per presentare informazioni, comunicare dati e messaggi, spiegare qualcosa a qualcun altro.
- **Exploratory:** Le visualizzazioni esplorative sono strumenti per i lettori per analizzare ciò che viene loro presentato. Molto spesso, l'esito dell'analisi esplorativa non è solo la risposta alle domande originali, ma la generazione di nuove domande.
- **Confirmatory:** Nell'analisi confermativa, le visualizzazioni sono destinate a testare ipotesi.

Bisogna decidere cosa visualizzare

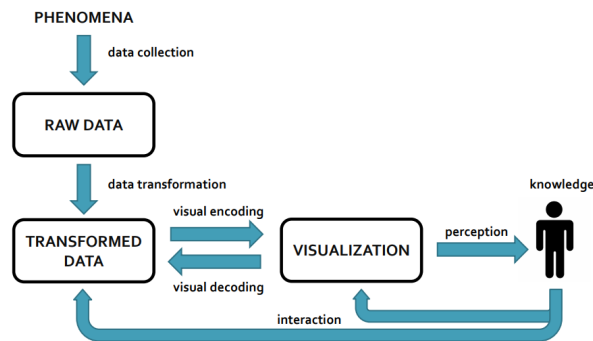


Figure 6: Visualization pipeline

2.1 Data

Informazioni fattuali come misurazioni o statistiche, utilizzate come base per il ragionamento, la discussione o il calcolo. I dati sono collezioni di *items* e *attributes* degli items. Gli *Items* sono gli oggetti/entità che vogliamo visualizzare. Gli *Attributi* sono le proprietà degli oggetti/entità.

Una *Tabella* è una griglia di colonne e righe, dove le righe rappresentano gli item e le colonne gli attributi. Un *network* è una collezione di nodi che rappresentano gli item, connessi tramite dei link; entrambi nodi e link possono avere attributi.

2.2 Tipi di attributi

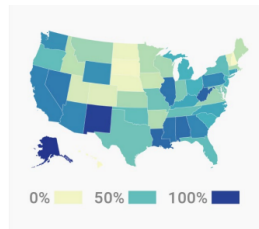
- *Quantitativi:* sono gli attributi dove i valori rappresentano quantità misurate, questi valori possono essere ordinati, ma può essere calcolata anche la distanza tra i valori.
- *Categorici:* sono gli attributi dove i valori descrivono le categorie, possono essere di tre tipi, **Nominali** se non hanno un ordine particolare, **Ordinali** se possono essere ordinati, **Binary** se hanno solo due stati.

2.3 Semantica degli attributi

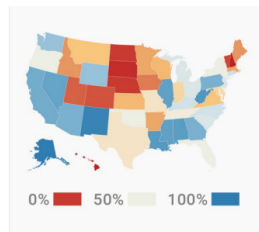
- **Spaziali e temporali:** Esempio la location, latitudine e longitudine o la data di assunzione come attributo temporale.
- **Sequenziali, ciclici divergenti:** Esempio: i mesi dell'anno sono ciclici, la temperatura è un attributo divergente.
- **Gerarchici:** Tipi di prodotto con sottocategorie, un esempio sono i vestiti.

Si deve selezionare la visualizzazione appropriata a seconda del tipo e dalla semantica dell'attributo. Esempio:

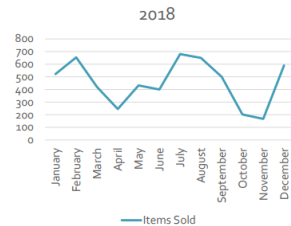
- Good idea: sequential color scheme



- Not a good idea: diverging color scheme



- Good idea



- Not a good idea

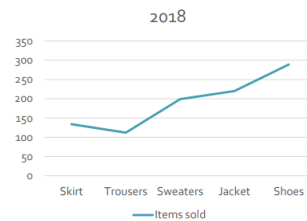


Figure 7: Differenza di Visualizzazione

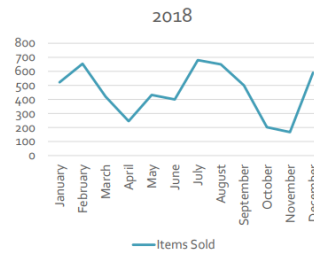
Figure 8: Differenza di Visualizzazione

2.4 Grafici Fondamentali (Bivariate data)

2.4.1 Bar Charts

Visualizza come una quantità misurata si distribuisce tra categorie. Ogni barra rappresenta una categoria e la lunghezza della barra è una quantità misurata in quella categoria. Dati bivariati: nominale/ordinale e quantitativo. Non confondere con gli istogrammi

- Good idea



- Not a good idea

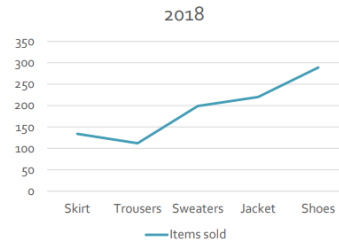


Figure 9: Bar Charts

2.4.2 Histograms

Frequenza degli elementi. Dati bivariati: una variabile indipendente quantizzata in intervalli (blocchi) e una variabile dipendente.

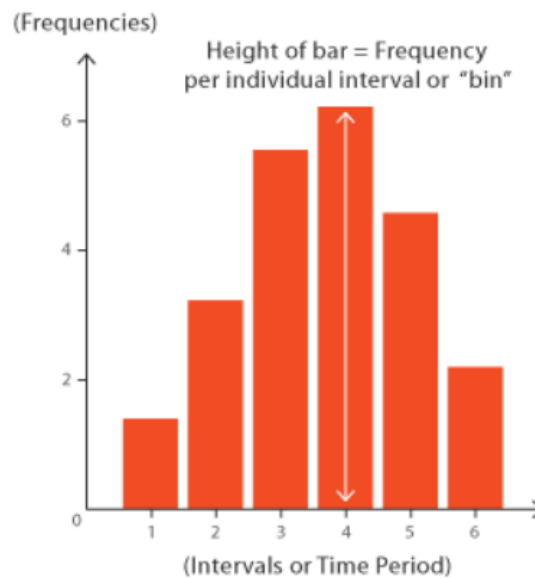
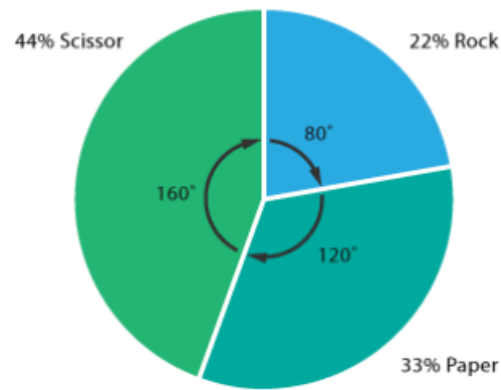


Figure 10: Histograms

2.4.3 Pie Charts

Mostrare proporzioni e percentuali tra le categorie. Utile per dare un'idea rapida e confrontare una fetta rispetto al totale, non adatto per confronti accurati. Altri svantaggi: numero limitato di valori, occupazione dello spazio.



Data			
Rock	Paper	Scissor	TOTAL
2	3	4	9
To calculate percentages			
$2/9=22\%$	$3/9=33\%$	$4/9=44\%$	100%
Degrees for each "pie slice"			
$(2/9) \times 360 = 80^\circ$	$(3/9) \times 360 = 120^\circ$	$(4/9) \times 360 = 160^\circ$	360°

Figure 11: Pie Charts

2.4.4 Donut Charts

Grafici a torta con l'area centrale tagliata. Maggiore enfasi sulla lunghezza dell'arco rispetto all'area. Più efficienti in termini di spazio.

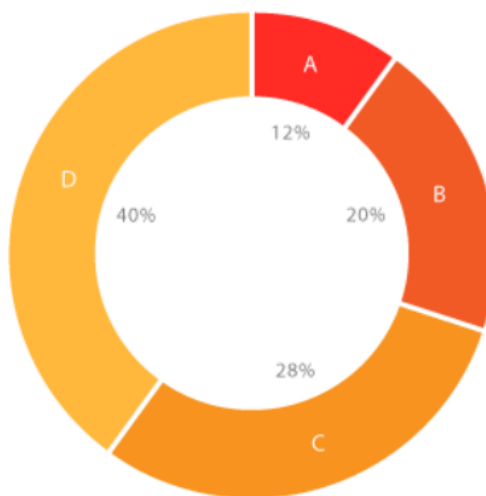


Figure 12: Donut Charts

2.4.5 Scatter plots

Relazione tra attributi: visualizzazione di come una quantità è correlata a un'altra (e analisi di cluster, valori anomali, ecc.). Dati bivariati: due attributi quantitativi indipendenti.

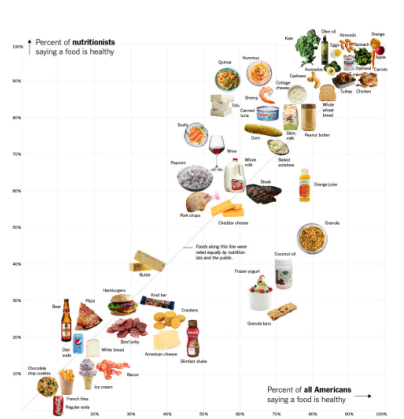


Figure 13: Scatter plots

2.4.6 Slope Charts

Alternativa ai Scatter plots: assi paralleli, ogni elemento è una linea che collega due quantità.

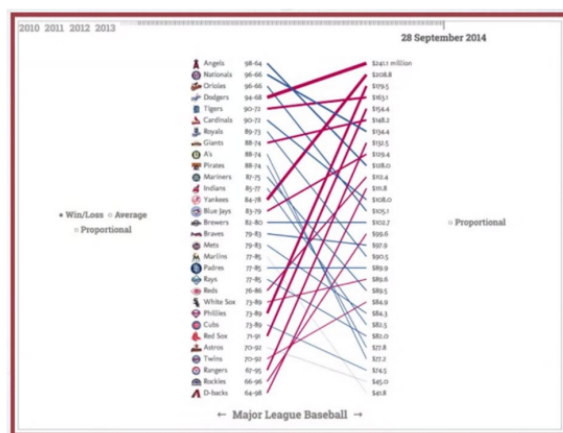


Figure 14: Slope Charts

2.4.7 Line Charts

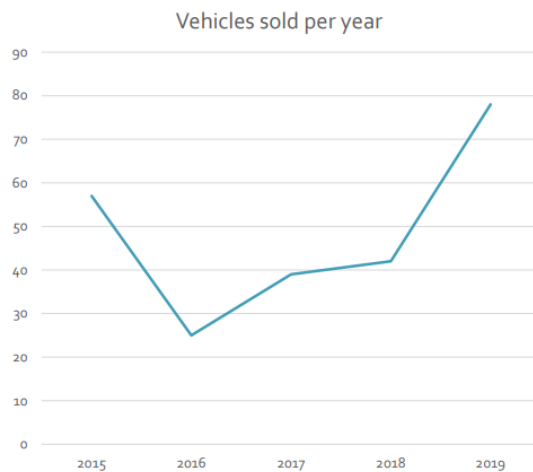


Figure 15: Line Charts

2.4.8 Area Charts

Line Charts dove l'area sotto alla linea è riempita.

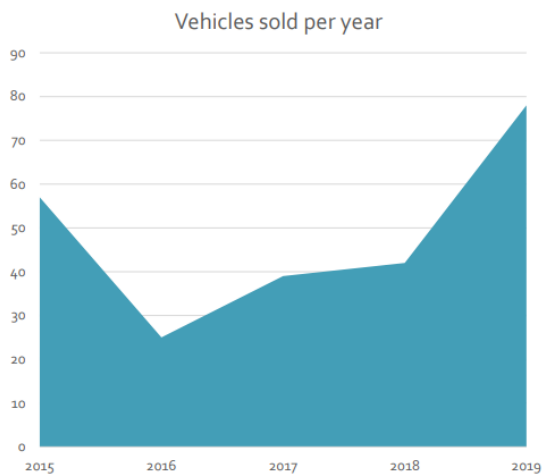


Figure 16: Area Charts

2.4.9 Choropleth maps

Come si distribuisce una quantità nelle diverse aree/geografiche/regioni. Colori, sfumature, pattern sono utilizzati per rappresentare la quantità associata alle aree/regioni. Buona panoramica (ma non confronto accurato). Rischio: confondere l'area geografica con i valori dei dati (prestare attenzione alla normalizzazione).

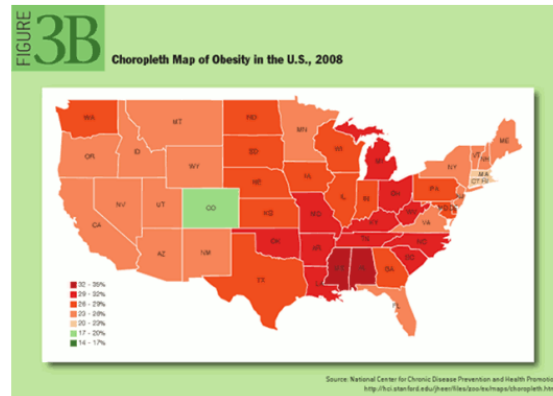


Figure 17: Chropleth maps

2.4.10 Symbol maps

Come si distribuisce una quantità lungo due coordinate spaziali. Un simbolo (spesso un disco o un quadrato) è posizionato in un punto e dimensionato in modo che la sua area sia proporzionale alla quantità associata al punto.

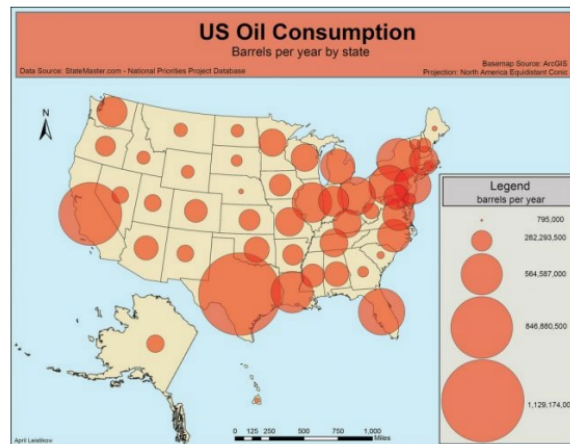


Figure 18: Symbol maps

2.5 Grafici Fondamentali (Attributi Multipli)

2.5.1 Stacked Bar Charts

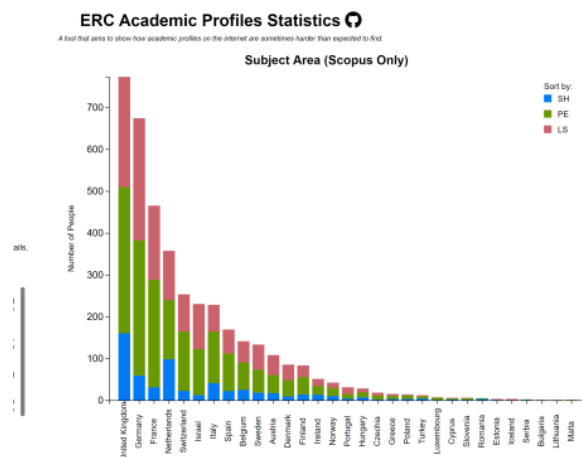


Figure 19: Stacked Bar Charts

2.5.2 Grouped Bar Charts

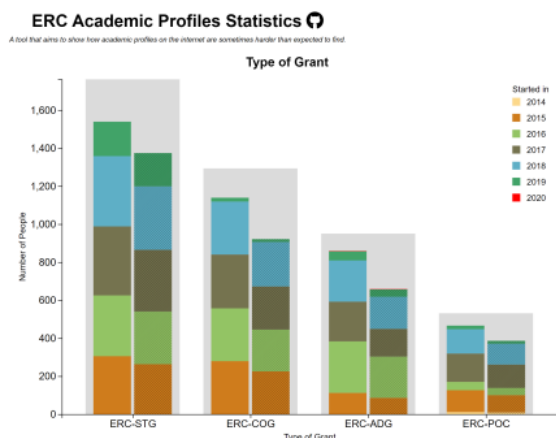


Figure 20: Grouped Bar Charts

2.5.3 Stacked Line Charts

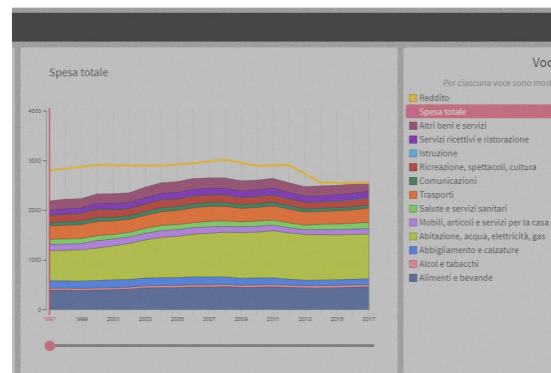


Figure 21: Stacked Line Charts

2.5.4 Line Charts Series

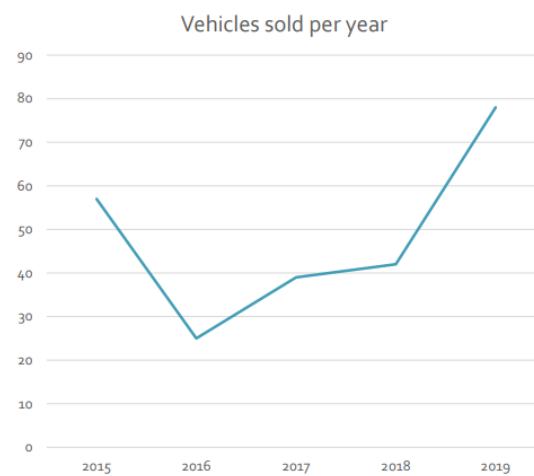


Figure 22: Line Charts Series

2.5.5 Bubble Charts

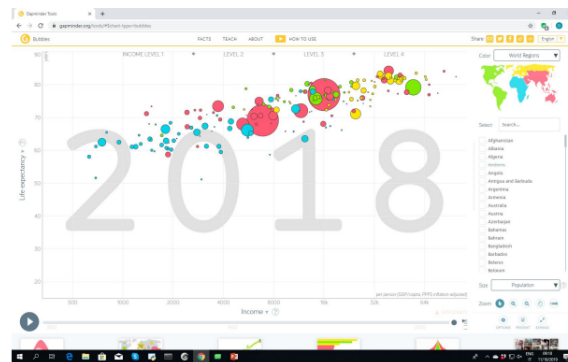


Figure 23: Bubble Charts

2.5.6 Small multiples

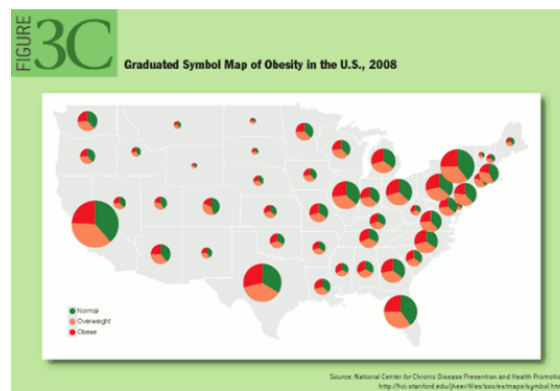


Figure 24: Small multiples

2.5.7 Symbol Maps

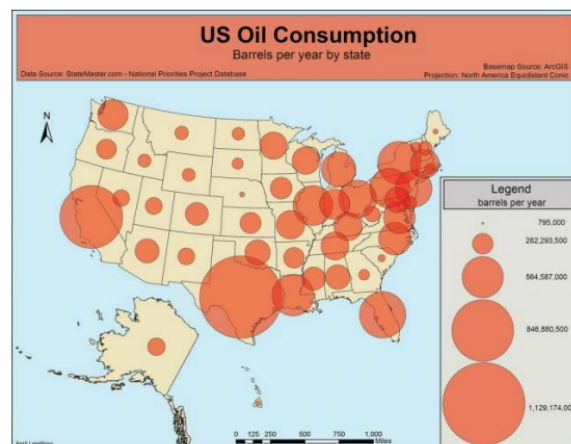


Figure 25: Symbol Maps

2.6 Grafici Fondamentali (Attributi Gerarchici)

2.6.1 Sunburst Charts

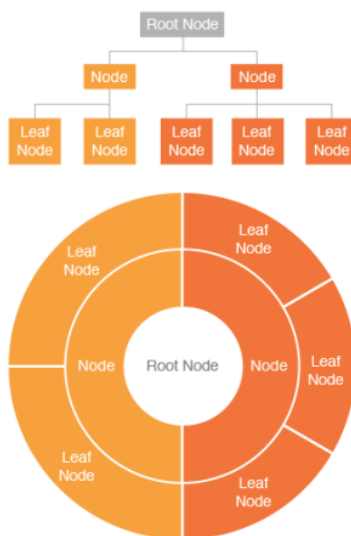


Figure 26: Sunburst Charts

2.6.2 Treempas

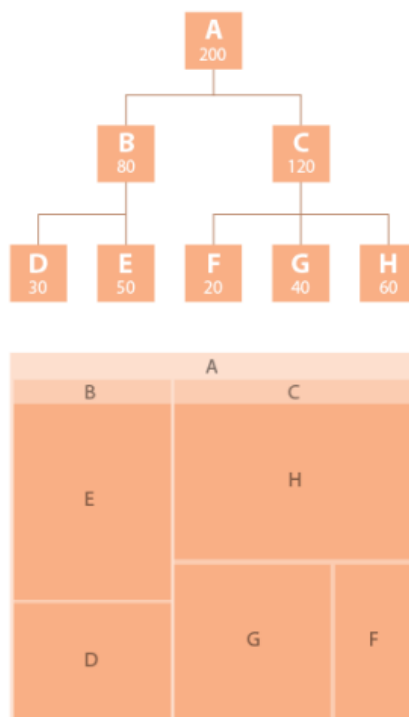


Figure 27: Treempas