

UE Visualisation

2019-2020

Dr. Antoine Neuraz

AHU Informatique médicale

Hôpital Necker-Enfants malades,
Université de Paris

2 Objectifs principaux

Principes généraux de visualisation

Application à l'aide de R et ggplot2

Organisation des cours

1ère moitié du cours: théorie

2ème moitié du cours: mise en pratique

Evaluation par projet en binôme

Mardi après-midi: description du projet

Soutenance le **2 décembre**

Visualisation



1. Action de rendre visible d'une façon matérielle l'action et les effets d'un phénomène.

2. Présentation visuelle sur un écran, sous forme d'image alphanumérique ou graphique, d'un ensemble d'informations traitées par des moyens informatiques.

Pourquoi visualiser ?

Pourquoi visualiser graphiquement ?

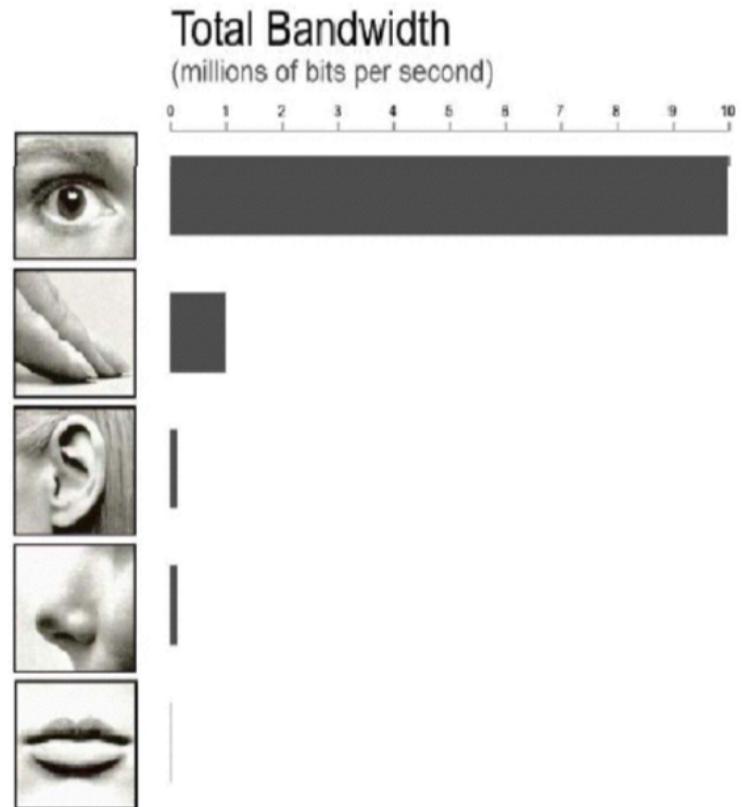
Plus riches : plus d'information en moins d'espace

Rend la **structure plus visible**

Plus accessible

Plus rapide à appréhender

Plus mémorable



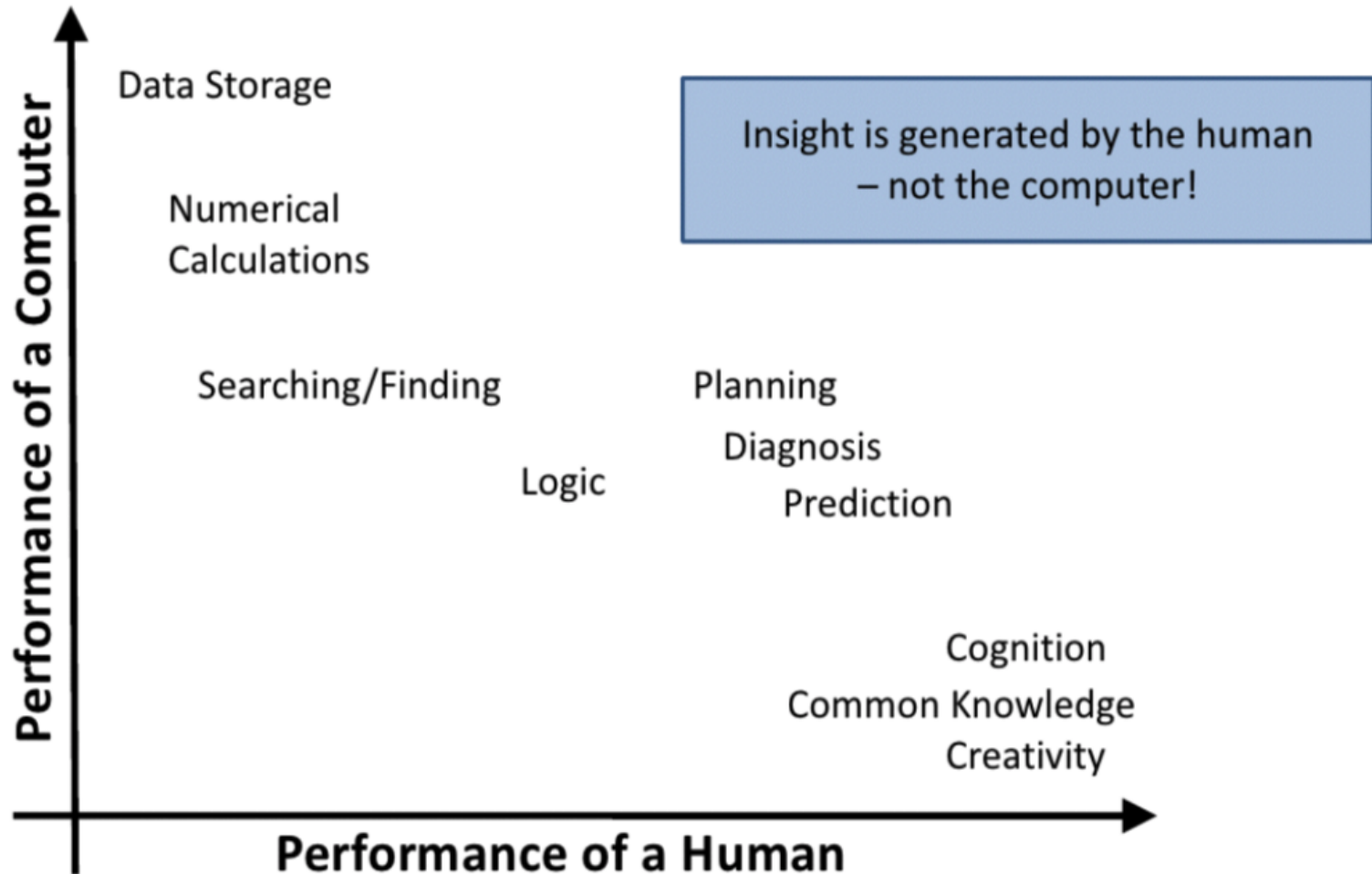
Pourquoi mettre un ordinateur
dans la boucle ?

Passage à l'échelle

Efficiency: réutilisation, diffusion

Qualité et **précision**

Pourquoi mettre un humain dans la boucle ?



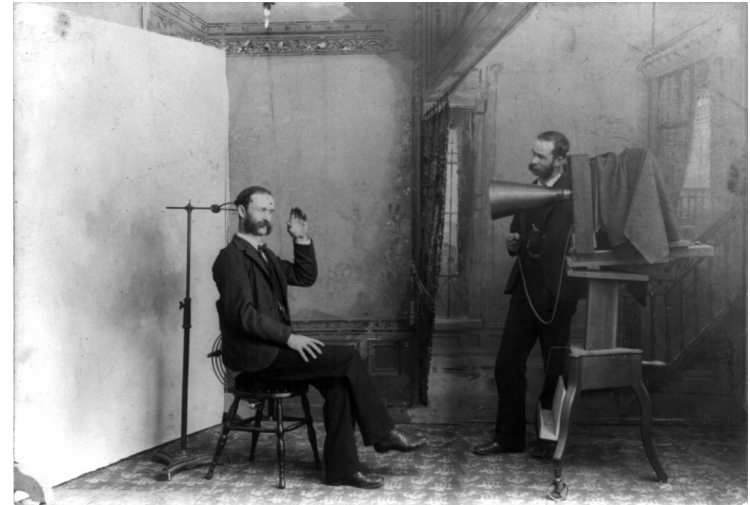
Buts d'une visualisation

Enregistrer
l'information

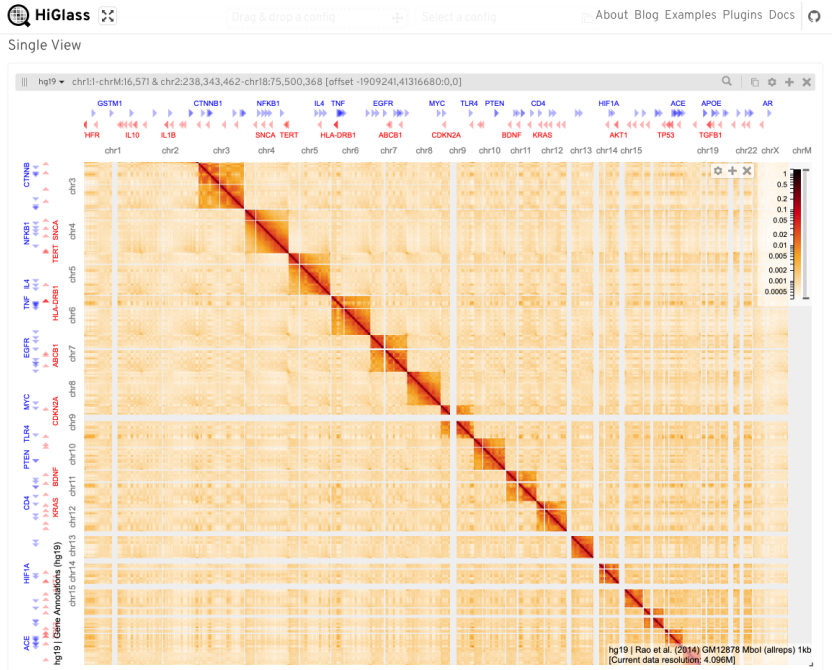
Analyser

Communiquer

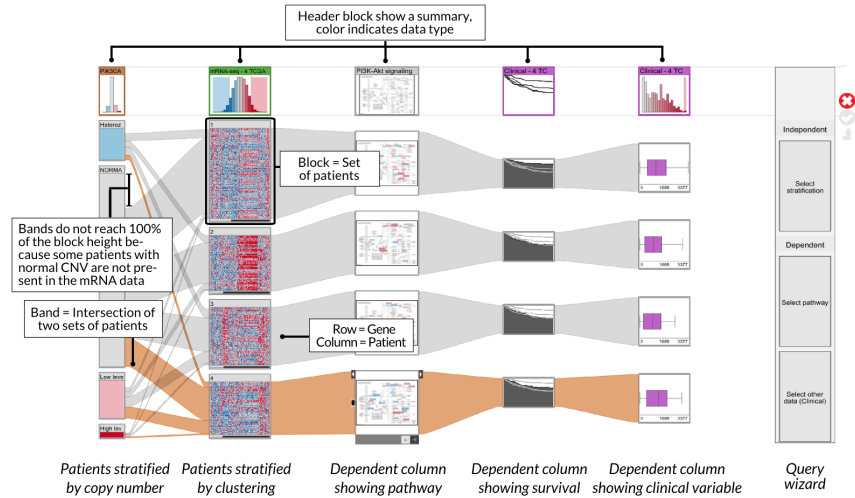
Enregistrer l'information



Analyser



<https://higlass.io/>



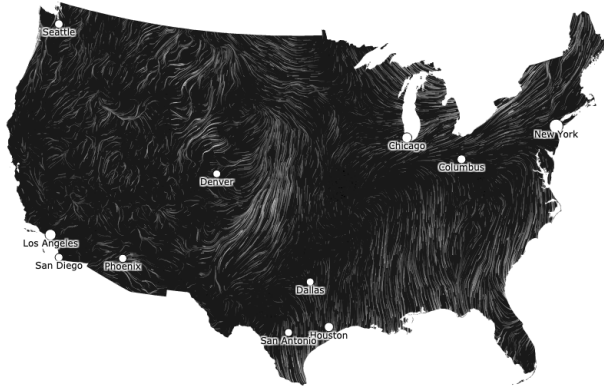
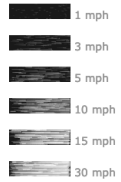
<https://frama.link/stratomex>

Communiquer

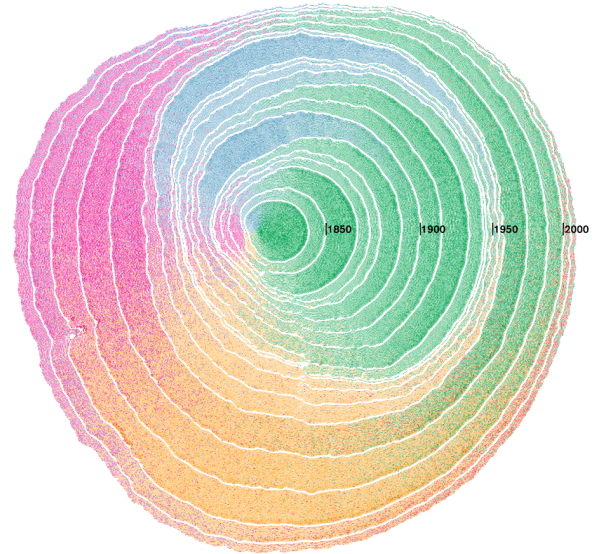
wind map

Nov. 15, 2019
8:43 am EST
(time of forecast download)

top speed: 27.6 mph
average: 7.2 mph



<http://hint.fm/wind/>



<https://frama.link/dendro>

Définition

La visualisation est le processus qui **transforme** les données en **représentation graphique** interactive à des fins d' **exploration**, de **confirmation** ou de **communication**.

Pourquoi ne pas se limiter aux statistiques ?

I		II		III		IV	
<u>x</u>	<u>y</u>	<u>x</u>	<u>y</u>	<u>x</u>	<u>y</u>	<u>x</u>	<u>y</u>
10	8.0	10	9.1	10	7.4	8	6.5
8	6.9	8	8.1	8	6.7	8	5.7
13	7.5	13	8.7	13	12.	8	7.7
9	8.8	9	8.7	9	7.1	8	8.8
11	8.3	11	9.2	11	7.8	8	8.4
14	9.9	14	8.1	14	8.8	8	7.0
6	7.2	6	6.1	6	6.0	8	5.2
4	4.2	4	3.1	4	5.3	19	12.
12	10.	12	9.1	12	8.1	8	5.5
7	4.8	7	7.2	7	6.4	8	7.9
5	5.						6.8

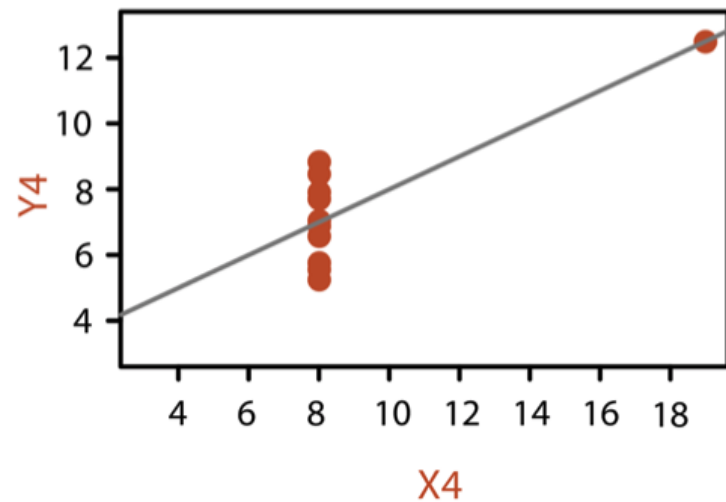
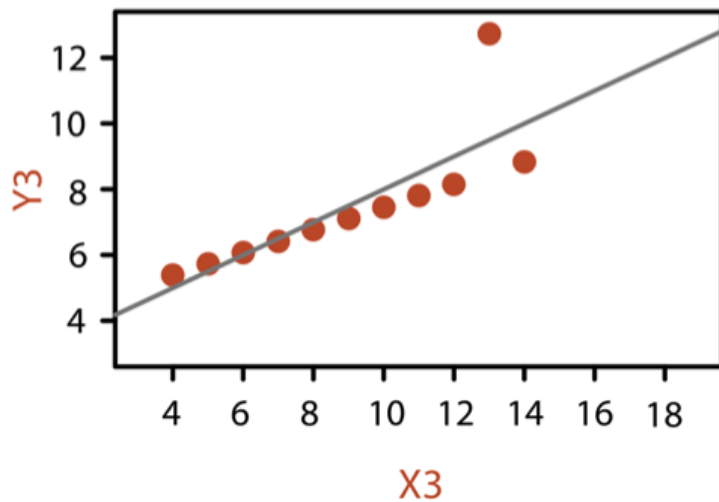
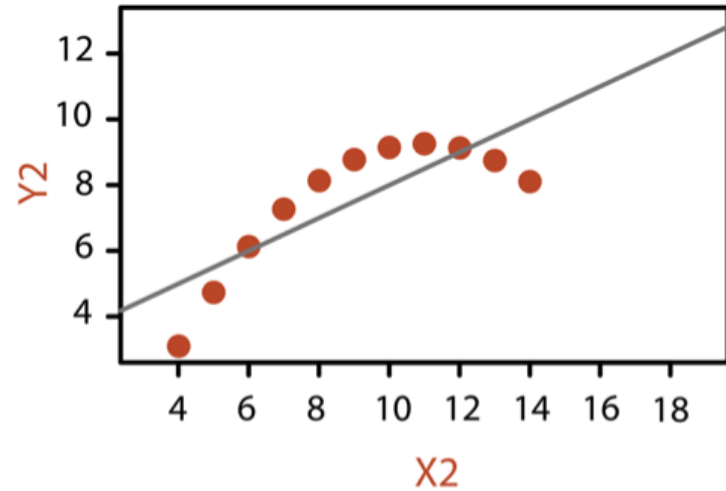
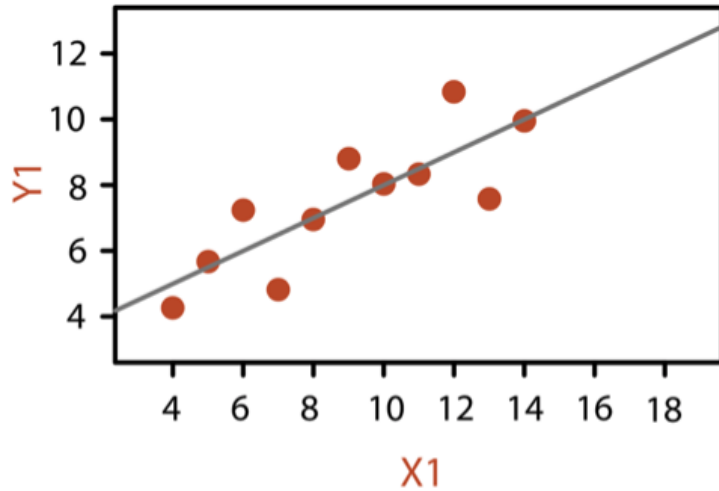
Mean x: 9 y: 7.50

Variance x: 11 y: 4.122

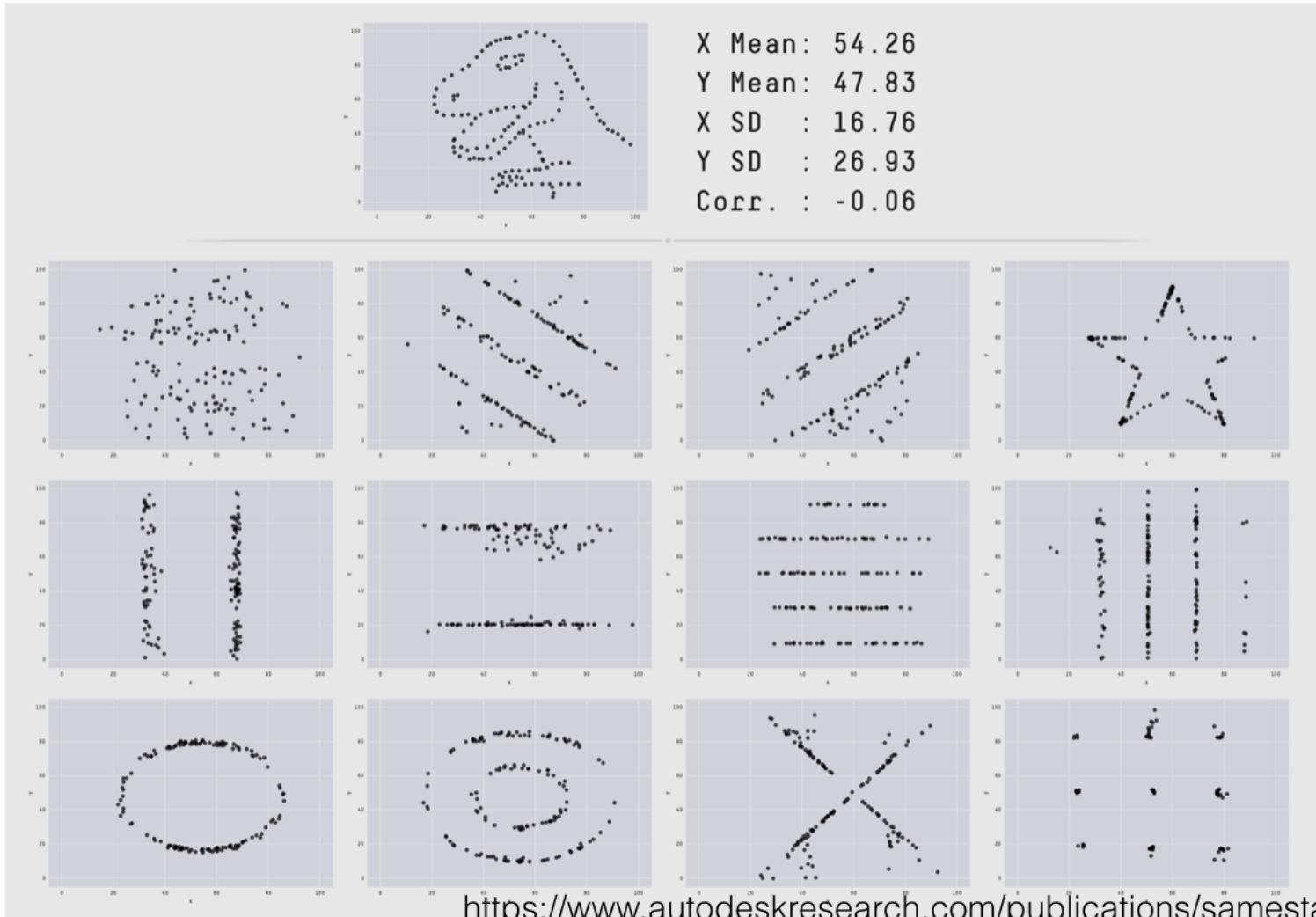
Correlation x - y: 0.816

Linear regression: $y = 3.00 + 0.500x$

Anscombe's quartet



Datasaurus dozen

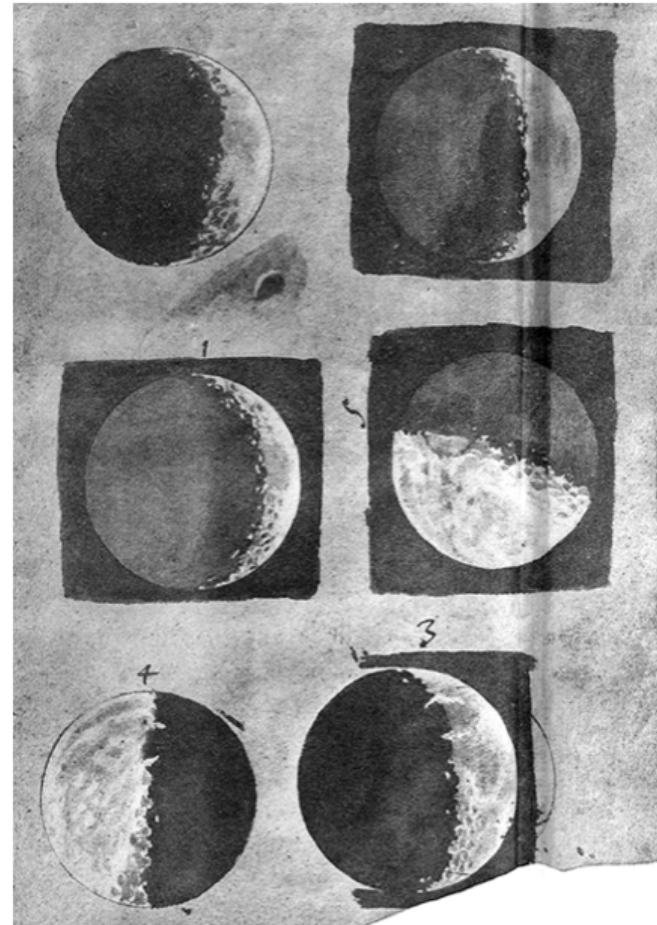


<https://www.autodeskresearch.com/publications/samestats>

Un peu d'histoire: enregistrer

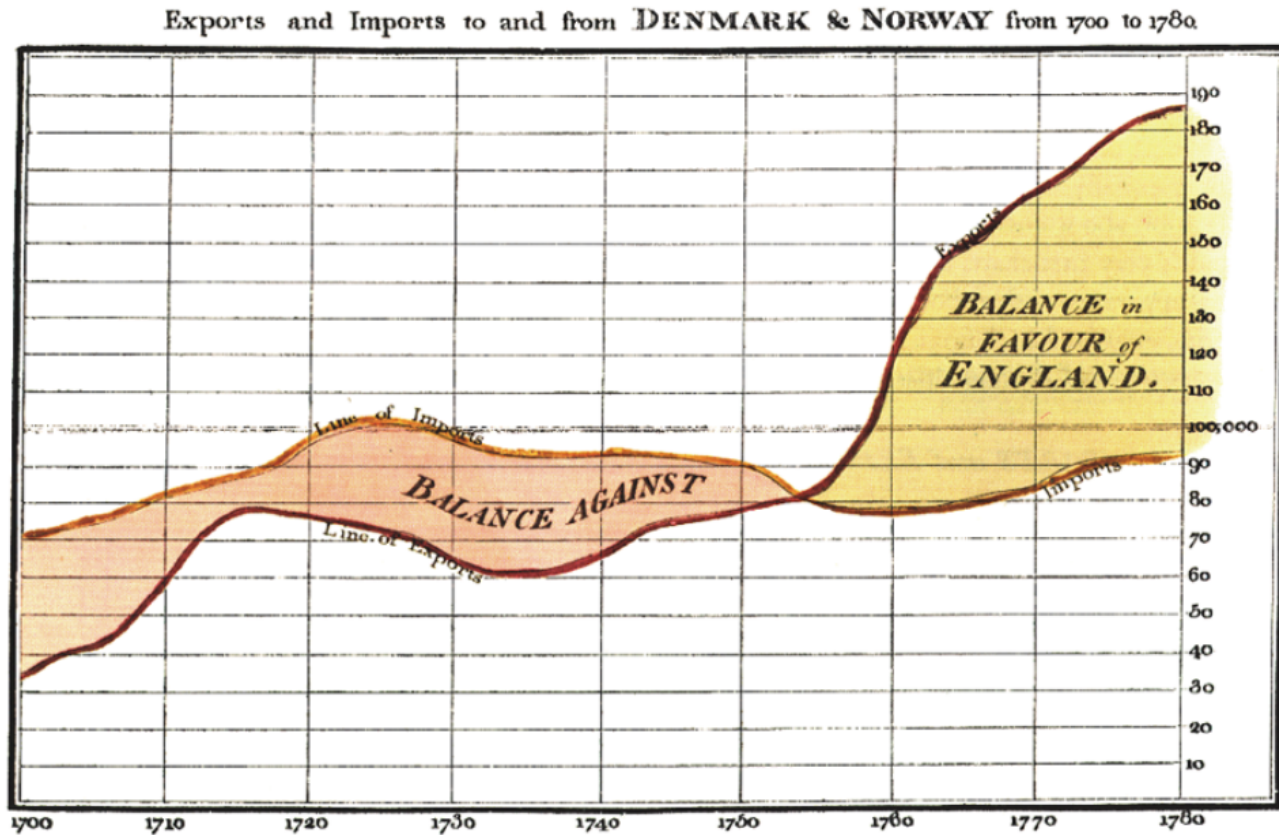


Da Vinci (1500)



Galilée (1616)

Un peu d'histoire: analyser



Playfair(1786)

Un peu d'histoire: trouver des patterns

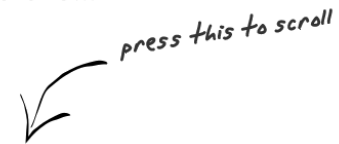


John Snow (1854)

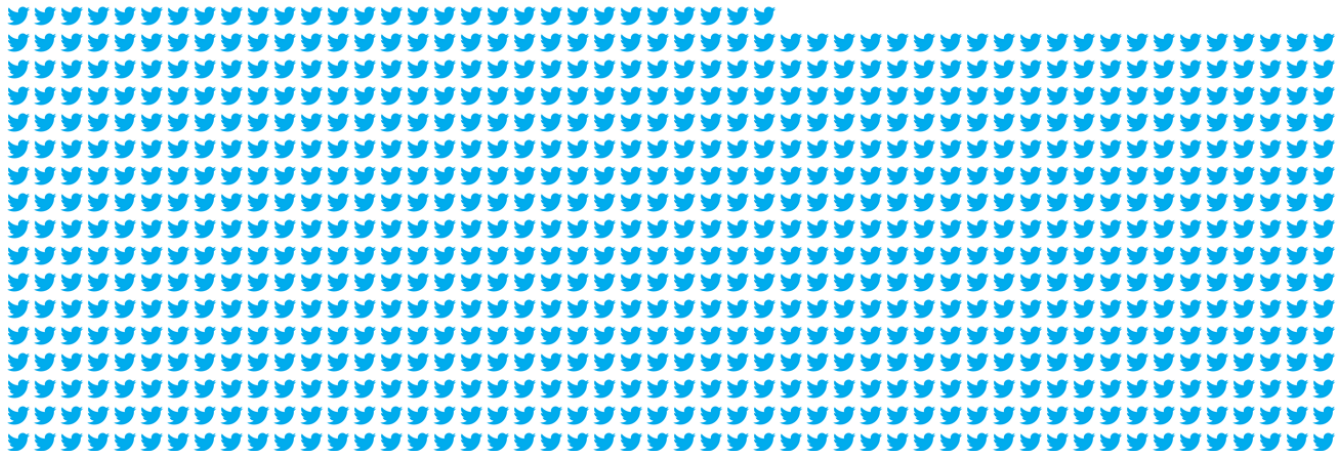
1 second of internet



in 1 second, each and every second, there are...



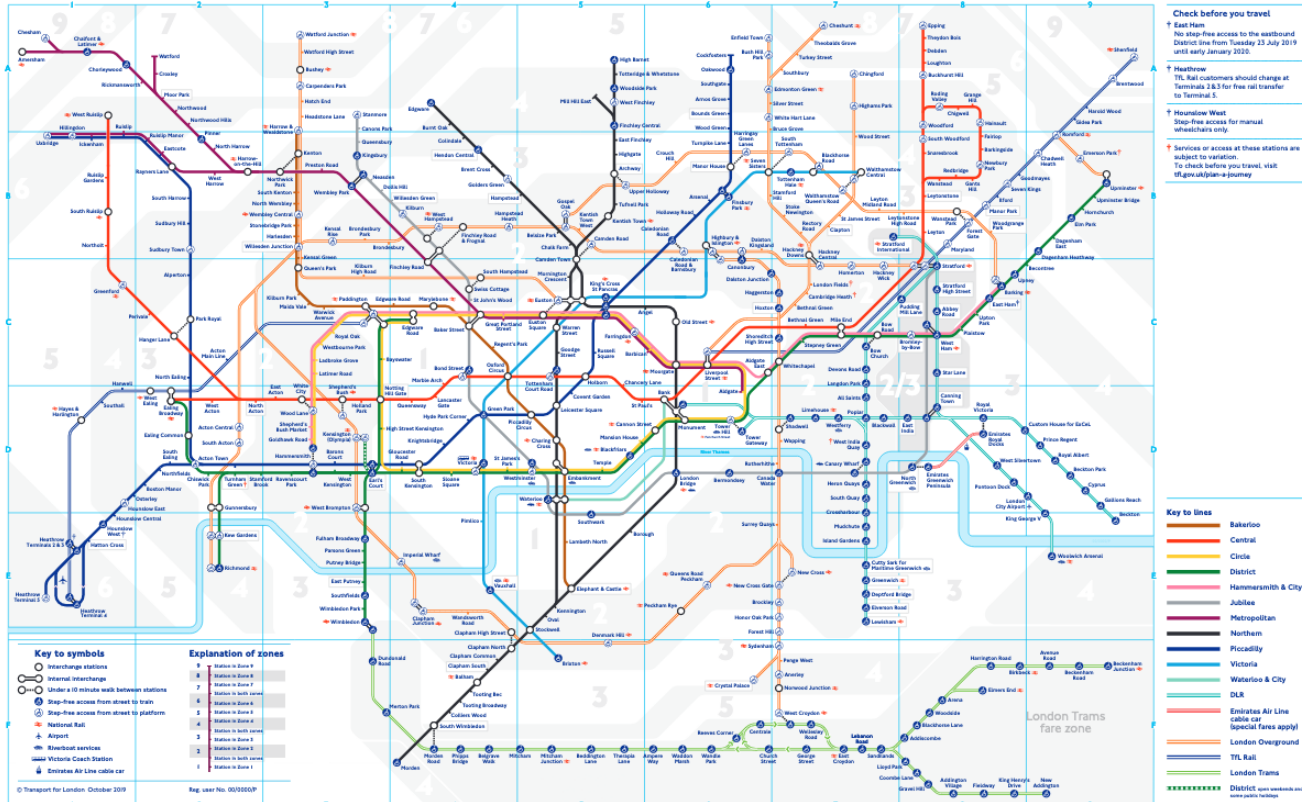
8,729 Tweets sent in 1 second  126,717 Tweets



London Tube 1933



London Tube today



MAYOR OF LONDON

tfl.gov.uk

24 hour travel information
0343 222 1234*

Sign up for email updates
tfl.gov.uk/emailupdates

[@TFLTravelAlerts](https://twitter.com/TFLTravelAlerts)



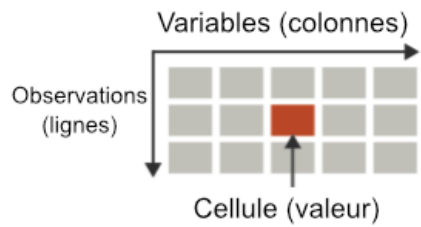
TRANSPORT FOR LONDON
EVERY JOURNEY MATTERS

*Network images may apply. See TfL.gov.uk for details.

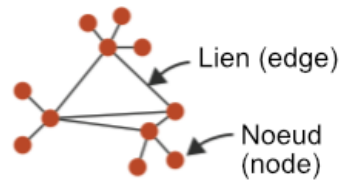
Online maps are strictly for personal use only. To license the Tube map for commercial use please visit tfl.gov.uk/licensing

Types de datasets

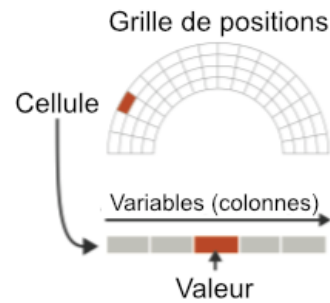
→ Tableaux



→ Réseaux



→ Champs



→ Spatial



Caractéristiques des données de tableau

Autres caractéristiques des données

Liens : relation entre 2 entités (observations, noeuds)

Positions (données spatiales)

Grilles (grids) : stratégie d'échantillonnage de données continues

Types de variables (attributs)

→ Discret

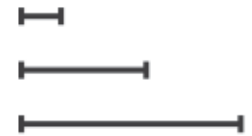


→ Ordonné

→ Ordinal



→ Quantitatif



Egalité



Comparaison



Arithmétique



Variables quantitatives

Intervalles = zéro arbitraire

ex: Dates, positions géographiques

→ pas de comparaison directe (cf °C et °F)

→ seulement des différences

Ratios = zéro absolu

ex: longueur, masse

→ ratios et proportions

Comment peut-on représenter visuellement 2 nombres,
par exemple:

4 et 8

Marques et échelles

Marques et échelles

Marques = geometry

Représente les **observations** (items) ou les liens

Echelles = scales

Change l'**apparence** des marques en fonction de **variables** (attributs)

Marques pour observations

Éléments géométriques de base

➔ Points



0D

➔ Lignes



1D

➔ Aires



2D

Marques 3D: Volume (rarement utilisé)

Marques pour les liens

→ Englobement



→ Connexion



Echelles (= variables visuelles)

⌚ Position

→ Horizontal



→ Vertical



→ Both



⌚ Couleur



⌚ Forme



⌚ Inclinaison



⌚ Taille

→ Longueur



→ Aire

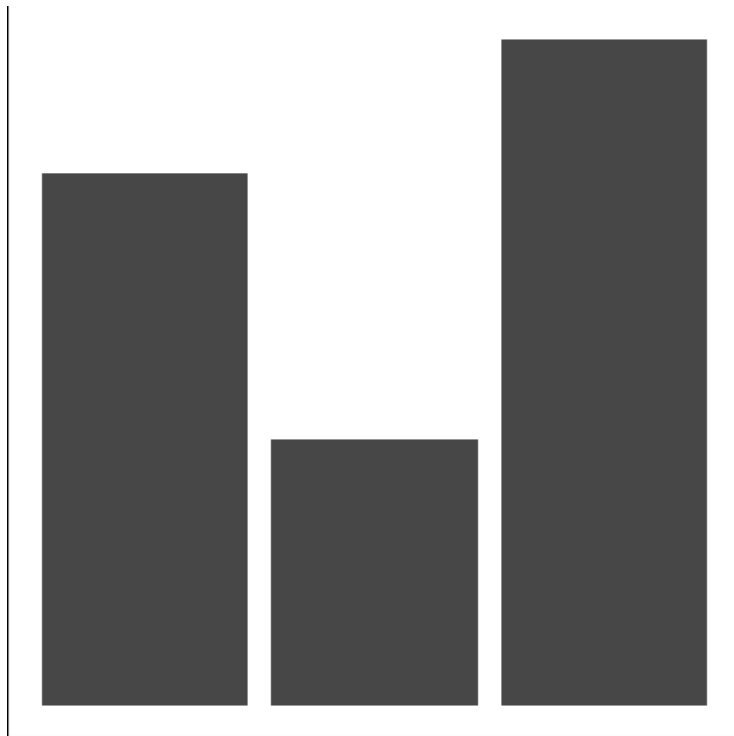


→ Volume



Contrôle l'apparence proportionnellement ou en fonction de variables

Utiliser les marques et les échelles



Marque: ligne

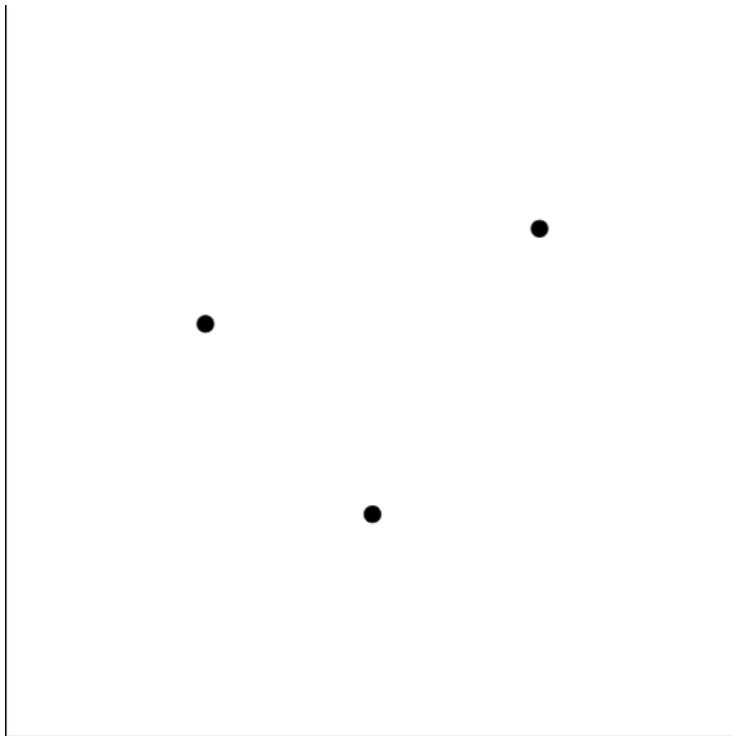
Echelles: longueur et position

Variables

1 variable quantitative

1 variable qualitative

Utiliser les marques et les échelles



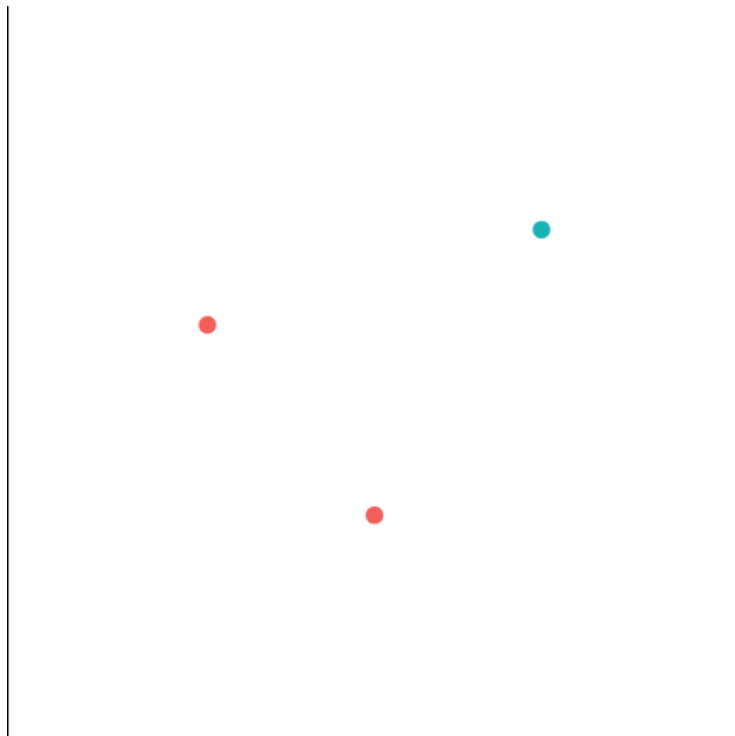
Marque: point

Echelles: position x et position y

Variables

2 variables quantitatives

Utiliser les marques et les échelles



Marque: point

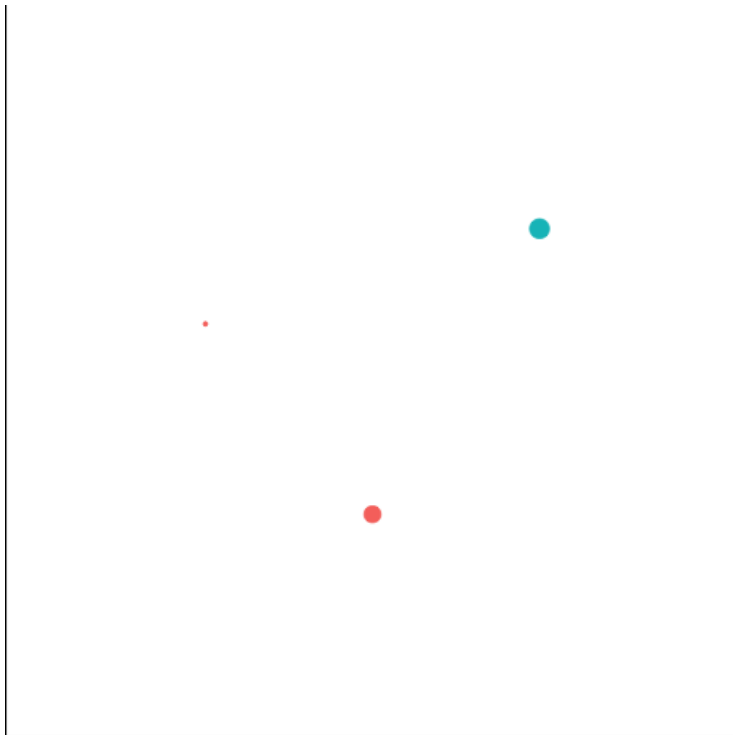
Echelles: position x et position y

Variables

2 variables quantitatives

1 variable qualitative

Utiliser les marques et les échelles



Marque: point

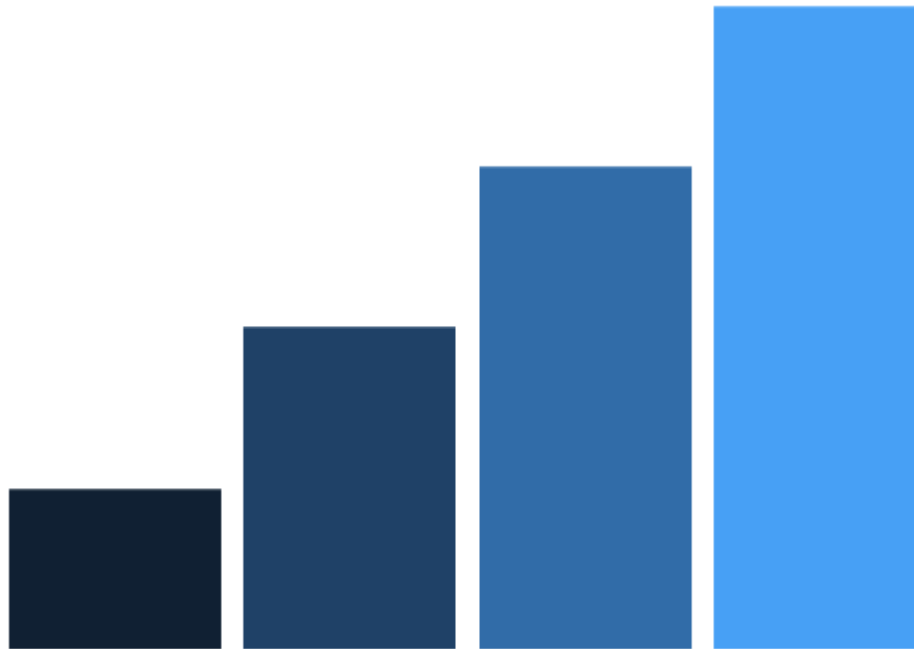
Echelles: position x et position y

Variables

3 variables quantitatives

1 variable qualitative

Encodage redondant



Longueur, position et Luminosité

Toutes les échelles ne naissent pas
égales

Echelles de magnitudes

Combien ?

Echelles d'identité

Quoi? Où?

Toutes les échelles ne naissent pas égales

Echelles de magnitudes

Combien ?

Position

Longueur

Saturation

...

Données quantitatives

Echelles d'identité

Quoi? Où?

Forme

Teinte

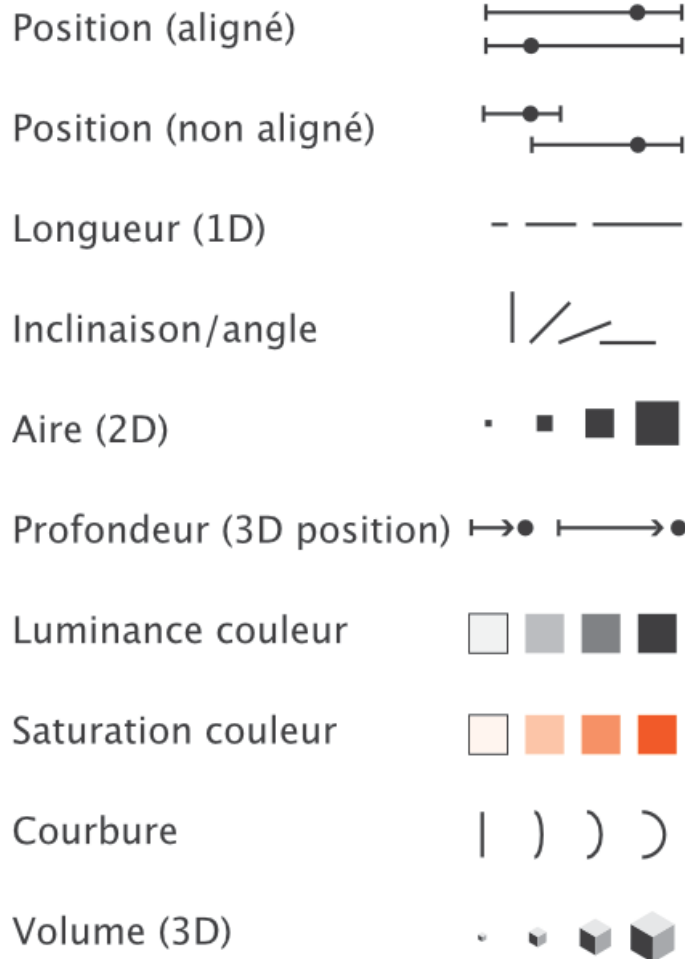
Région spatiale

...

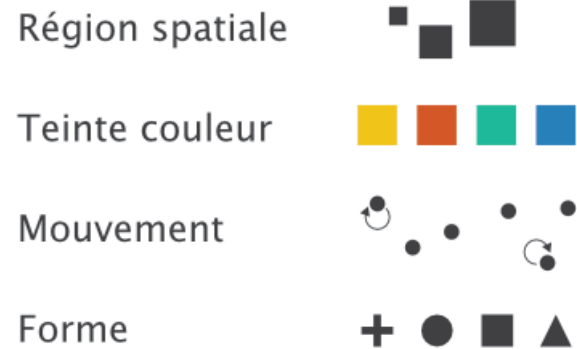
Données qualitatives

Echelles - efficacité

→ Echelles de Magnitude



→ Echelles d'identité



Plus ↑
Efficiency
↓ Moins

Same