21/11/2025

PANTHÉON SORBONNE
UNIVERSITÉ PARIS 1

ATELIER

La place de la datavisualisation dans les projets de recherche



Adèle Etaix
Benoit Pandolfi

CC BY NC ND

Les problématiques qui guideront cet atelier

Comment éditorialiser la donnée ?

Comment construire une visualisation ?

Comment la datavisualisation peut-elle devenir un outil exploratoire pour produire de nouveaux savoirs ?

DÉFINITION

Utilisation de moyens numériques pour améliorer la compréhension d'informations, de données qui manquerait d'intelligibilité à première vue.

Stéfan Sinclair, Stan Ruecker et Milena Radzikowska, "Information Visualization for Humanities Scholars", MLA Commons, 2013

Contexte :

- Croissance exponentielle de nombre de données
- Milieu entrepreneurial et de l'informatique décisionnelle → aide à la décision
- Engouement pour la visualisation dans le milieu universitaire

➡ Pratiques diffuses et intégrées au monde de la recherche

UN PEU D'HISTOIRE DE LA DATAVISUALISATION

A quand remonte les premières traces de datavisualisation ?

→ un tableau est déjà une visualisation puisque construit un ordre graphique qui induit une visualisation de la donnée

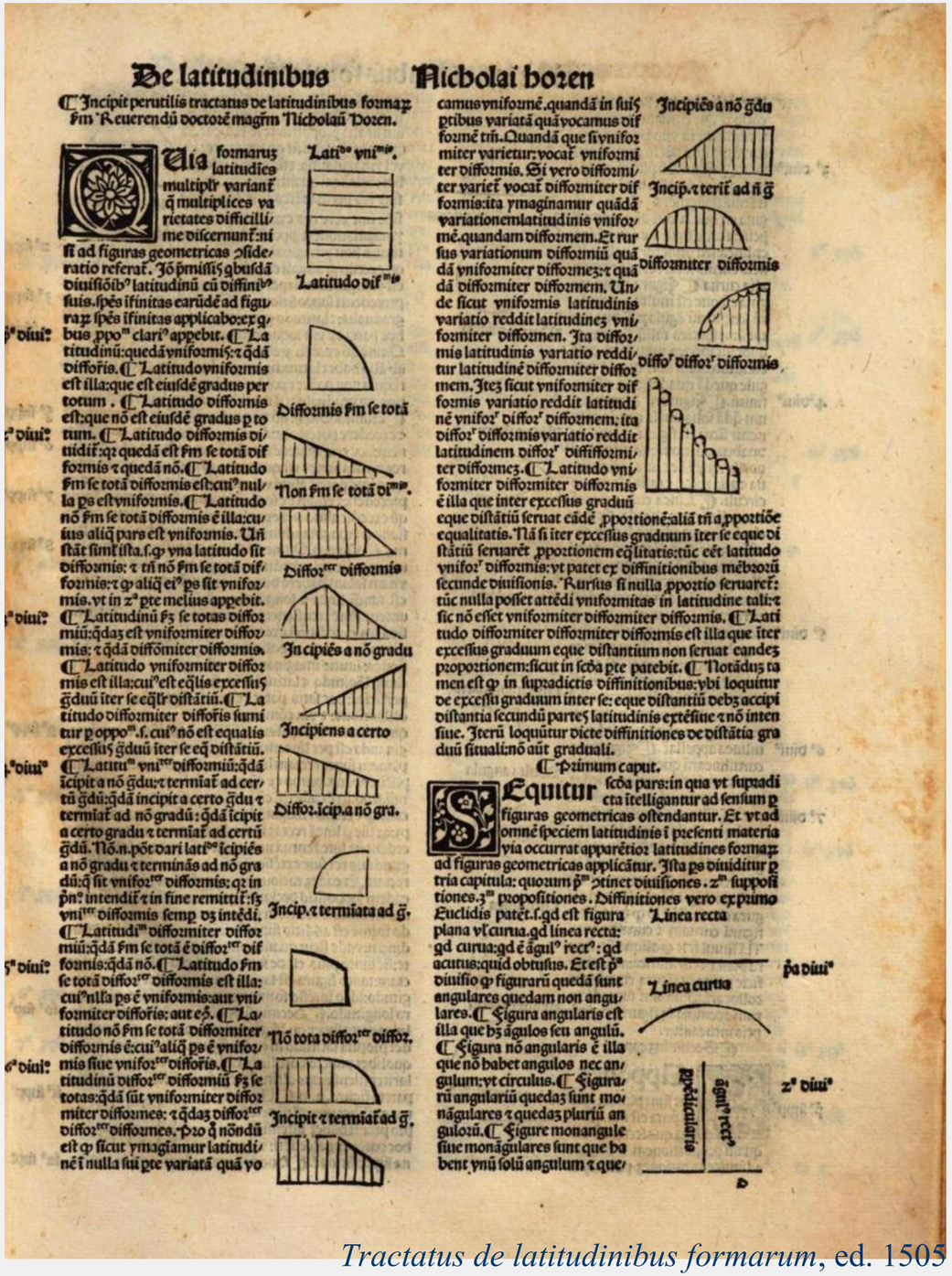
NICOLE ORESME

Vers 1370, les balbutiements de la géométrie analytique

- Représentation sous forme graphique de la relation entre deux variables
- Conception du principe des coordonnées cartésiennes pour la représentation graphique de phénomènes quantitatifs

> « Les propriétés de cette qualité en seront examinées plus clairement et plus facilement dès lors que quelque chose qui lui est semblable est dessiné en une figure plane, et que cette chose, rendue claire par exemple visible, est saisie rapidement et parfaitement par imagination …car l'imagination des figures aide grandement à la connaissance des choses même “

Tractatus de configuratione qualitatum et motuum



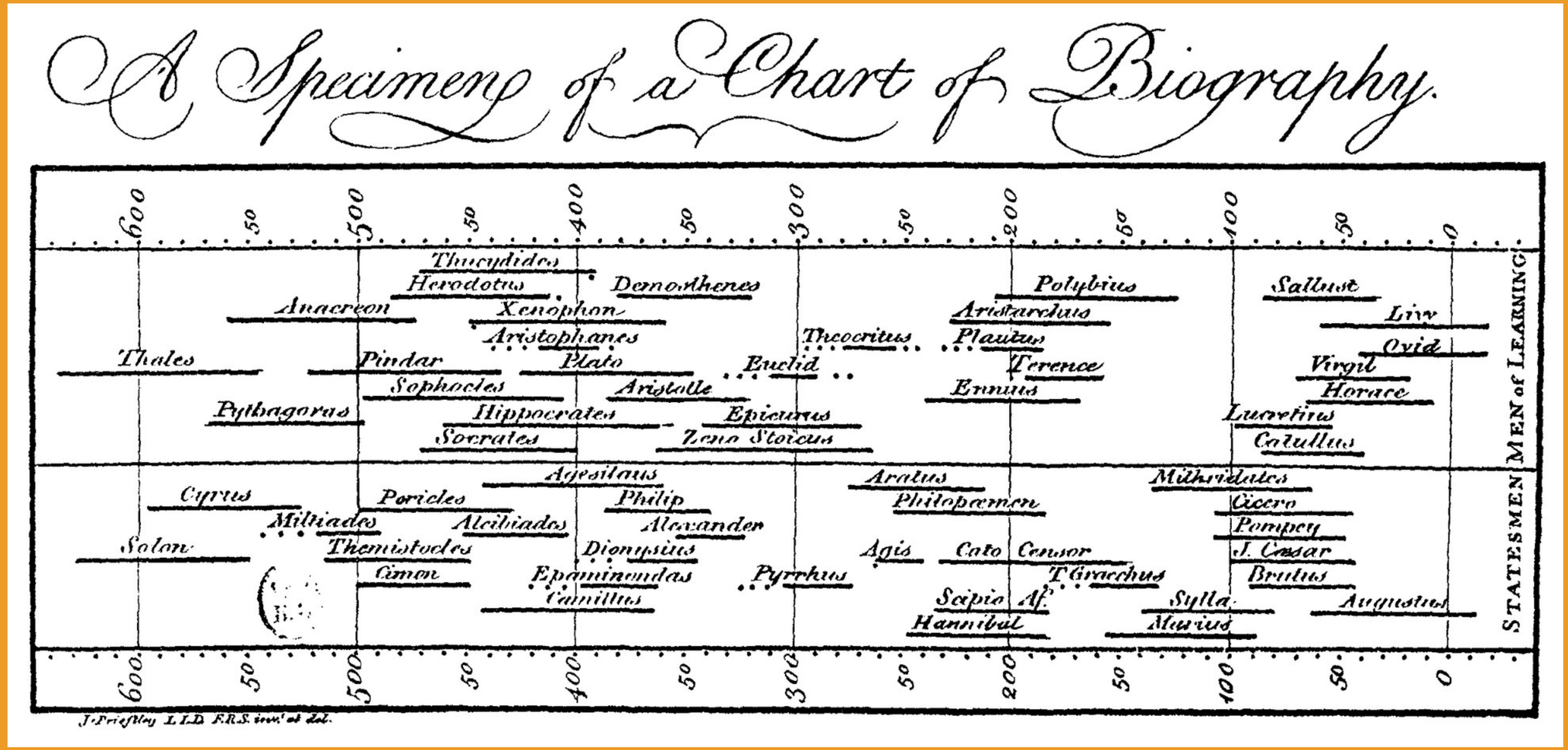
Tractatus de latitudinibus formarum, ed. 1505

JOSEPH PRIESTLEY

Vers 1760, invention de la timeline

Objectif :

Combler le manque de sources à propos de scientifiques et leur position relative les uns par rapport aux autres.

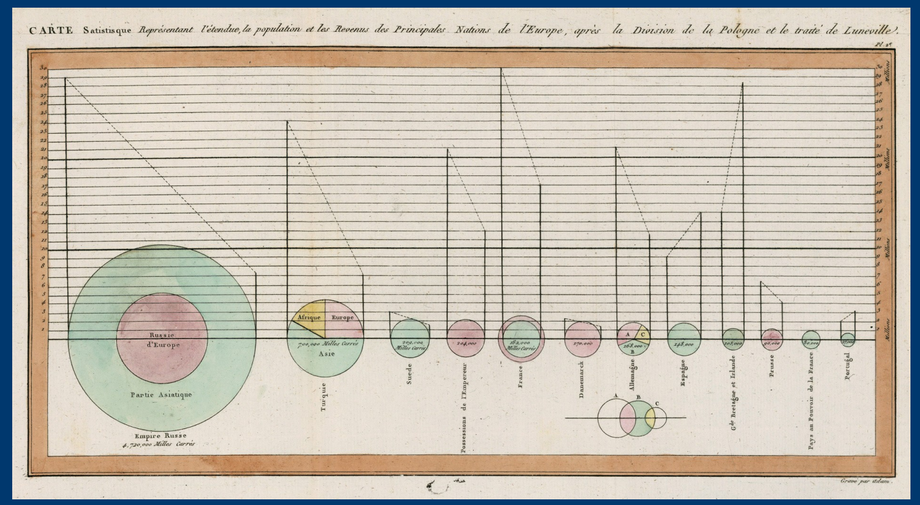


A Chart of Biography, 1765

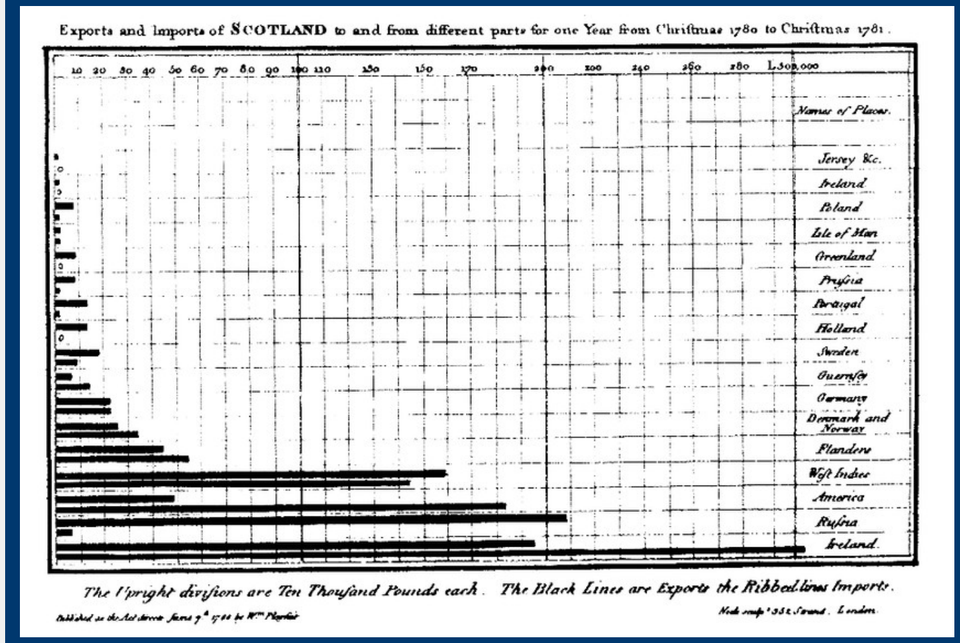
WILLIAN HENRY PLAYFAIR

Au début du XIXème siècle, trois nouveaux types de représentations graphiques

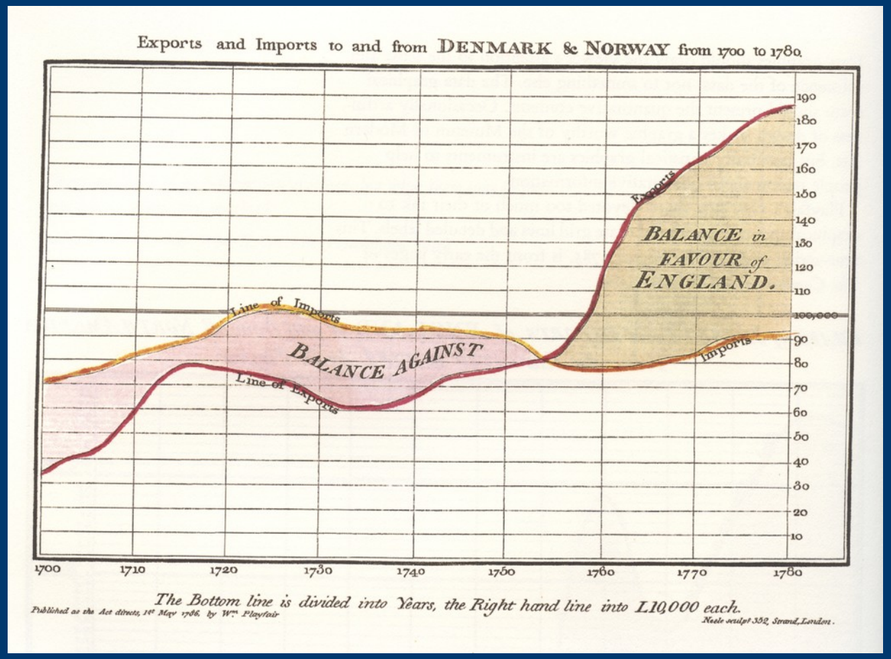
- Série statistique chronologique sous forme de courbe graphique (line chart)
- Diagramme en barres (bar chart)
- Diagramme circulaire (pie chart)



The Statistical Breviary, 1801



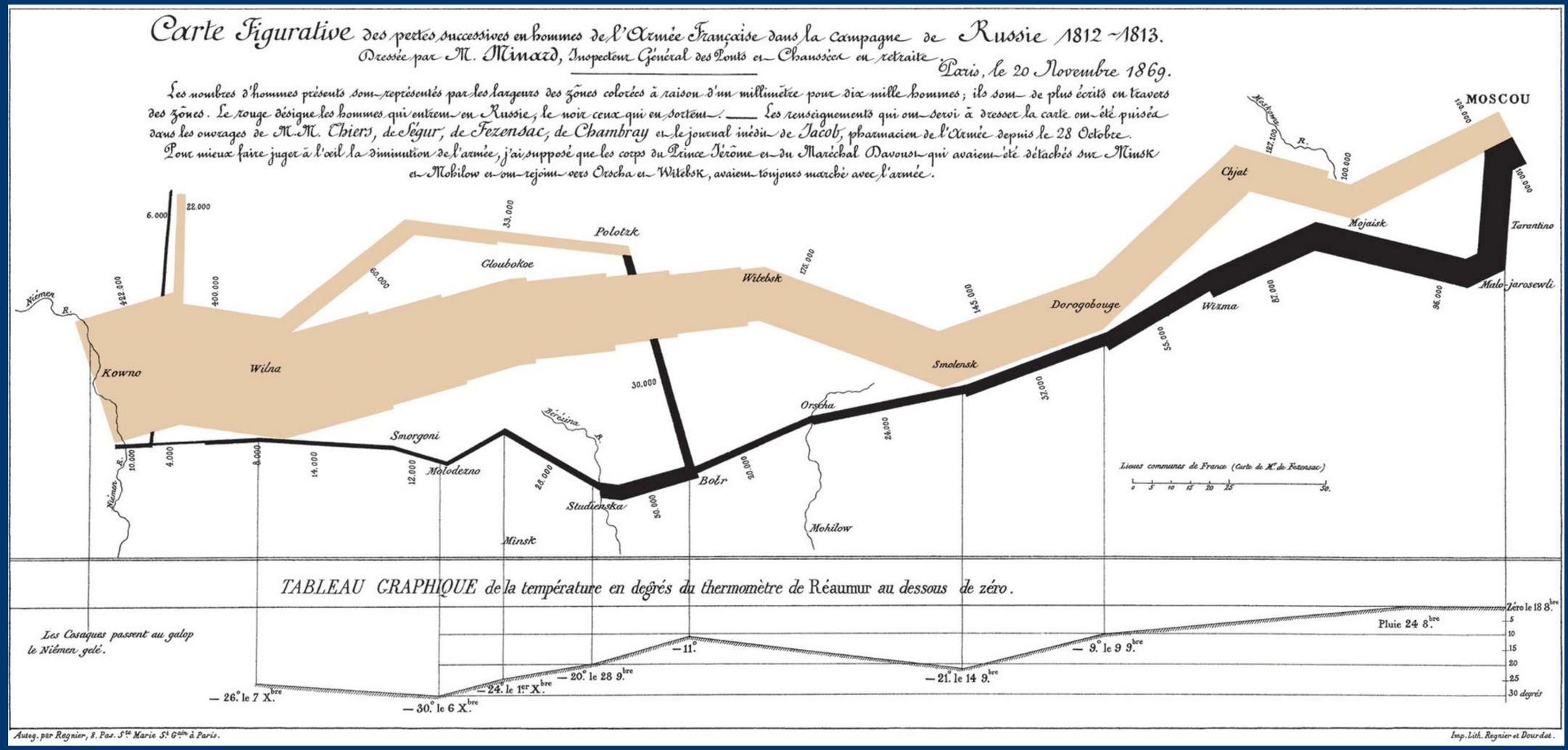
Commercial and political Atlas, 1786



Commercial and political Atlas, 1786

CHARLES JOSEPH MINARD

En 1869, l'invention de la carte de flux



Carte figurative des pertes successives en homme de l'Armée française dans la campagne de Russie en 1812-1813, 1869

FLORENCE NIGHTINGALE

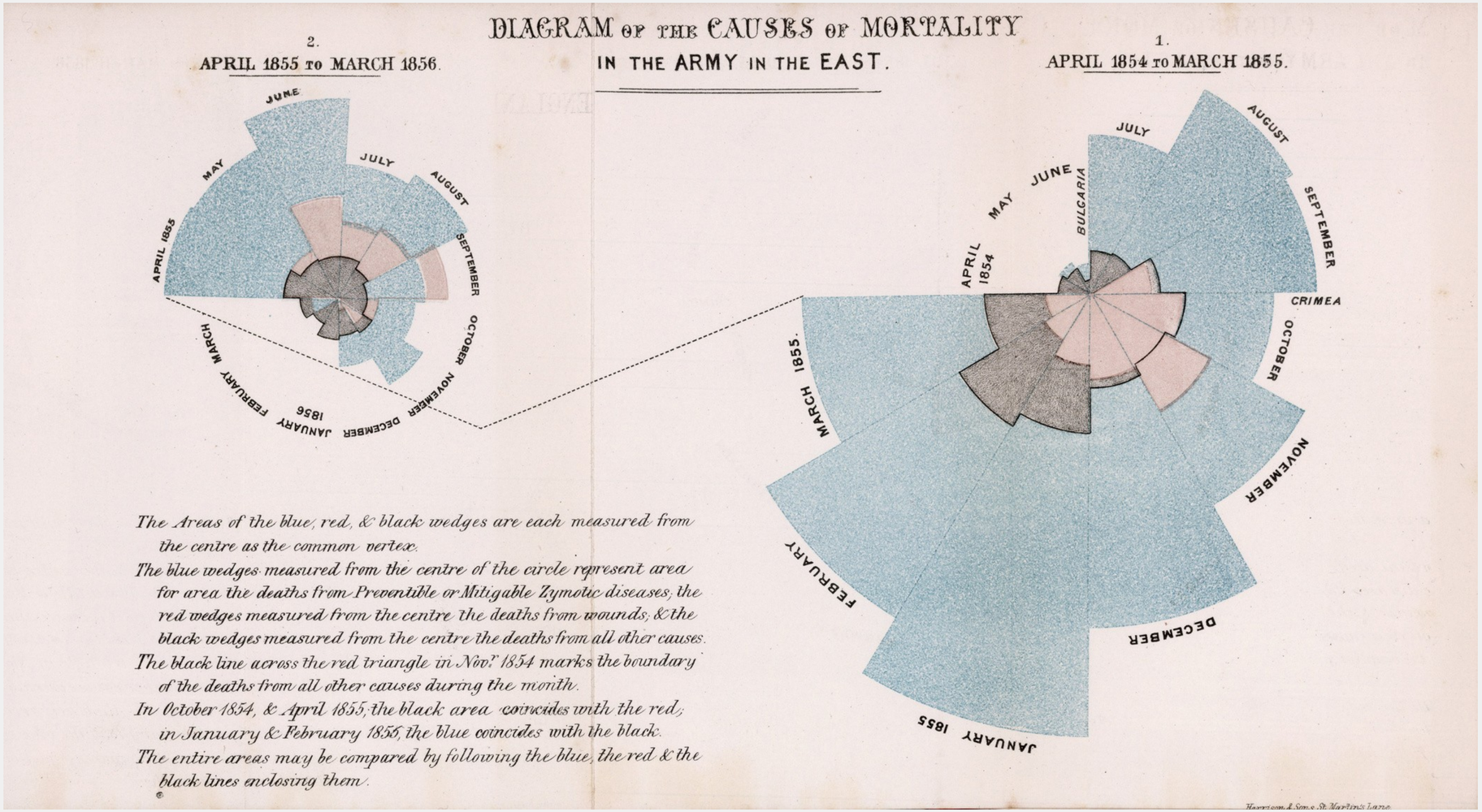
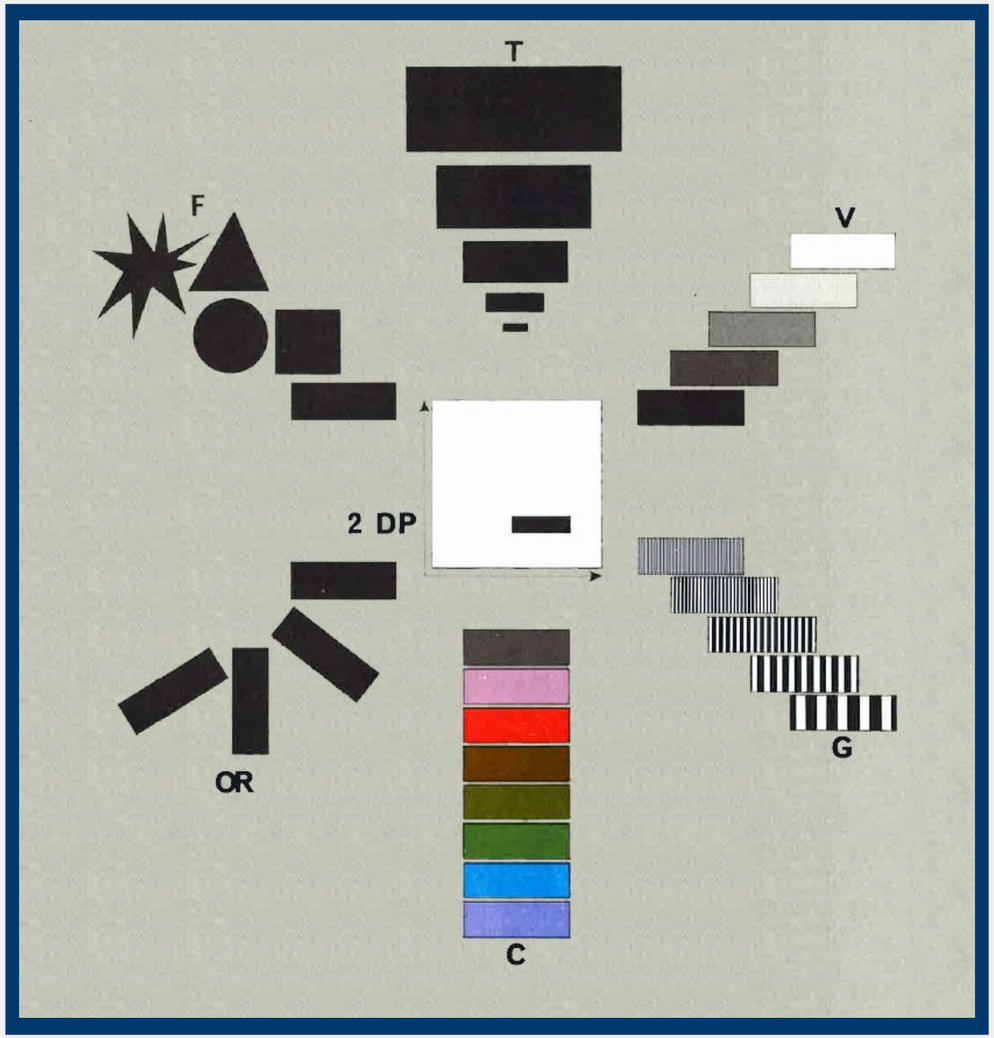


Diagramme des causes de mortalité au sein de l'armée en Orient

Au milieu du XIX, une pionnière de la présentation visuelle de l'information.

JACQUES BERTIN



Sémiologie graphique. Les diagrammes, les réseaux, les cartes, 1967

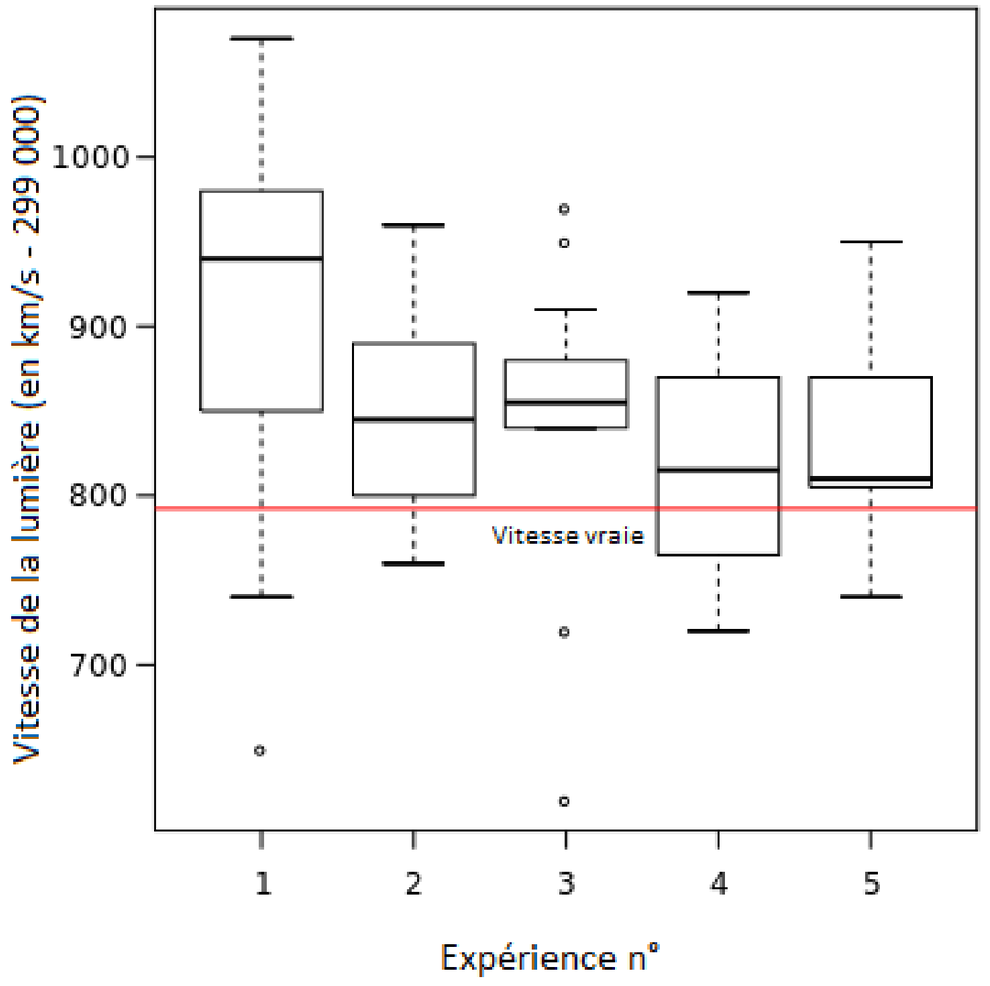
Au milieu du XXème siècle, l'invention de la sémiologie graphique.

- Création d'une doxa cartographique
- Priorité à l'efficacité de la communication (temps d'observation court et moindre coût mental)

JOHN TUKEY

Au XXème siècle...

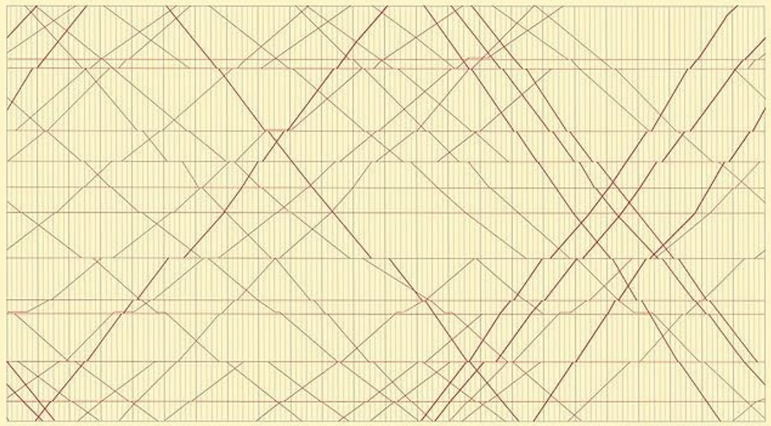
- Distinction entre l’analyse exploratoire des données et l’analyse liée aux tests d’hypothèse
- Présente le principe de la boîte à moustache



EDWARD TUFTE

1980-2000

- Ouvrage fondateur qui formalise les règles de bonne visualisation.
- Pionnier du minimalisme visuel
- Le data-ink ratio
  - Maximiser la proportion d’encre réellement dédiée aux données.
  - Réduire tout artefact graphique non informatif.
- Intégrité graphique et honnêteté des représentations
- Influence durable



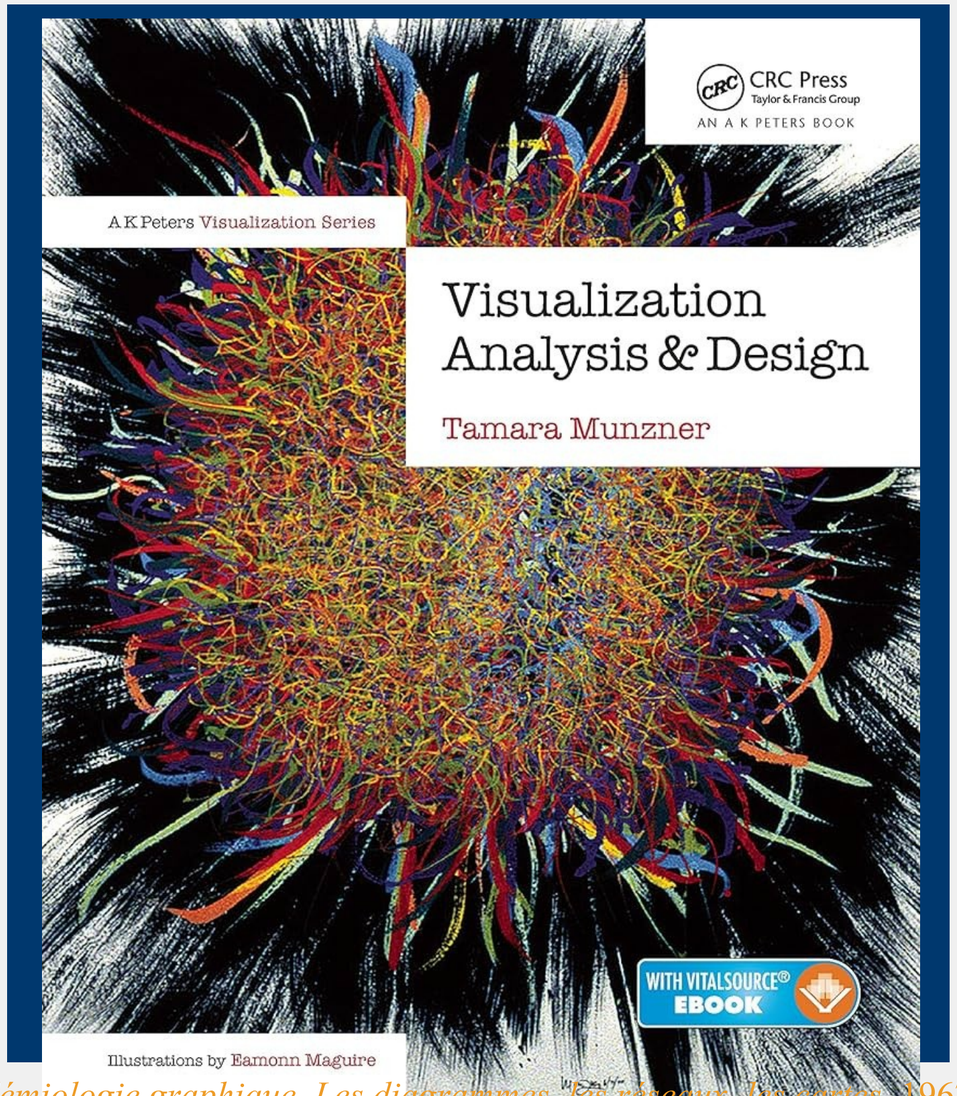
SECOND EDITION

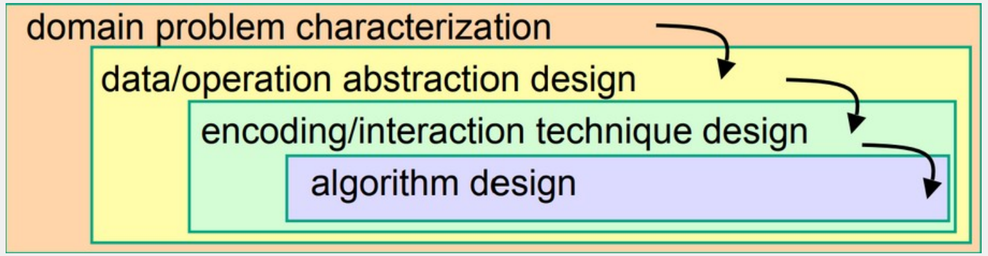
The Visual Display of Quantitative Information

EDWARD R. TUFTE

Tufte, Edward R. (1983). The Visual Display of Quantitative Information (2nd ed.). Cheshire, Connecticut, US: Graphics Press LLC.

TAMARA MUNZNER





2000 - 2020

- Le “Modèle de Munzner” (Nested Model for Visualization Design)
  - Domain problem & data (comprendre le problème)
  - Data/operation abstraction (choisir l’abstraction)
  - Visual encoding & interaction (choisir la forme)
  - Algorithm design (implémentation)

Sémiologie graphique. Les diagrammes, les réseaux, les cartes, 1967

AUJOURD'HUI

Big Data, temps réel & visualisation augmentée

- **Explosion du Big Data**
- **Visualisation interactive et web-native**
- **Data storytelling**
- **Visual Analytics**
- **IA & automatisation**

Nouvelles pratiques

- **Dataviz en temps réel** (monitoring, IoT, data pipelines).
- **Design systémique** : palettes contrôlées, guidelines.
- **Éthique de la data viz** : biais, choix de représentation, transparence.
- **Accessibilité :** couleurs perceptibles, narration inclusive, alternatives textes.
- **Hybridation Dataviz × IA × UX :** assistants de visualisation, exploration guidée.

Figures contemporaines

- **Ben Fry & Casey Reas** : Processing = fondation du creative coding moderne.
- **Mike Bostock** : créateur de D3.js, figure majeure de la visualisation web.
- **Hannah Fairfield, Amanda Cox** : New York Times, pionnières du data journalism moderne.
- **Alberto Cairo** : théoricien de la datavisualisation contemporaine, The Functional Art.

IEEE VIS
International Conference on Visualization

COMMENT ÉDITORIALISER UNE DONNÉE ?

Quand construire une datavisualisation ?

Quels risques dans la production d’une datavisualisation ?

Quels sont les différents types de visualisations ?

QUAND CONSTRUIRE UNE DATAVISUALISATION ?

Données quantitatives : se prêtent facilement aux représentations visuelles, on y pense en premier lieu

Or, en SHS : pluralité des interprétations, des hypothèses, des prismes d'étude. Donc la mobilisation de datavisualisation est également possible pour les données textuelles, qualitatives (notamment avec la numérisation croissante des documents et l'exploration automatisé de textes).

Quelle question de recherche ? Quelle problématique technique ? Méthodologique ?

2013-2017: mots les plus fréquents dans les discours de F.Hollande



Mots les plus fréquents dans les discours de François Hollande entre 2013 et 2027,
C. Cottet, 2023

QUELQUES RISQUES LIÉS À LA PRODUCTION D'UNE DATAVISUALISATION

Réification des données : simplification extrême des données puisque données déjà travaillées donc première interprétation et choix.

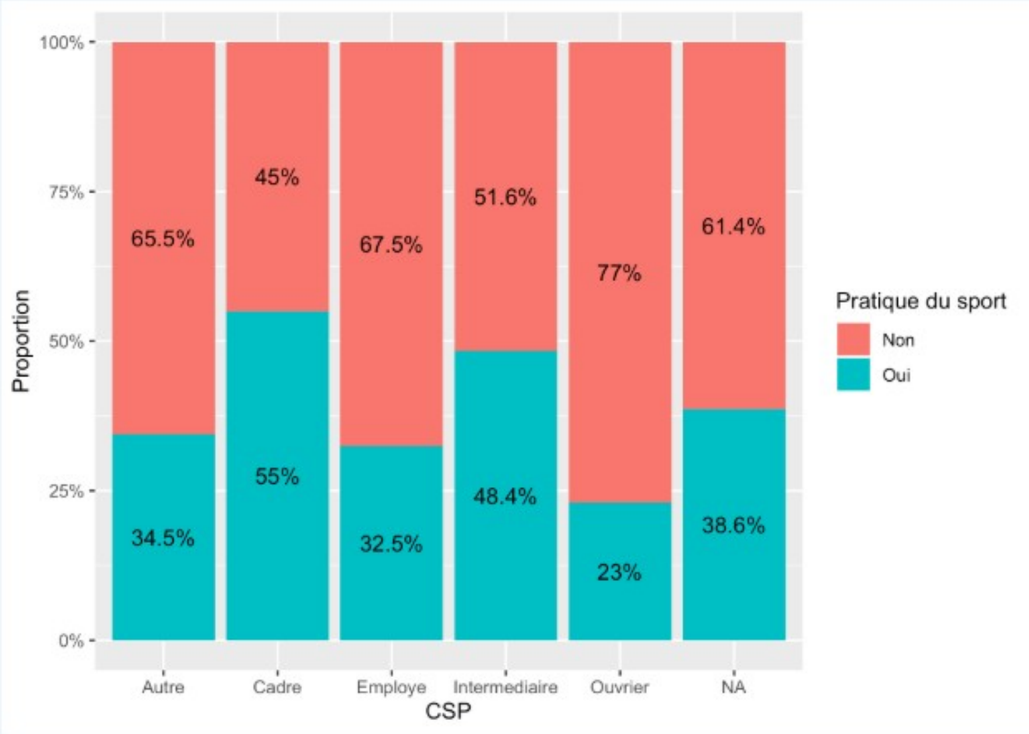
Conflit entre réalisme scientifique et exigences esthétiques de la visualisation : raisonnement analytique de la démarche scientifique vs pyschologie cognitive.

La tentation du storytelling : logique d'accroissement des connaissances et d'éclairage des phénomènes sociaux.

→ Prendre en compte le continuum "Matériaux - Information - Connaissance"

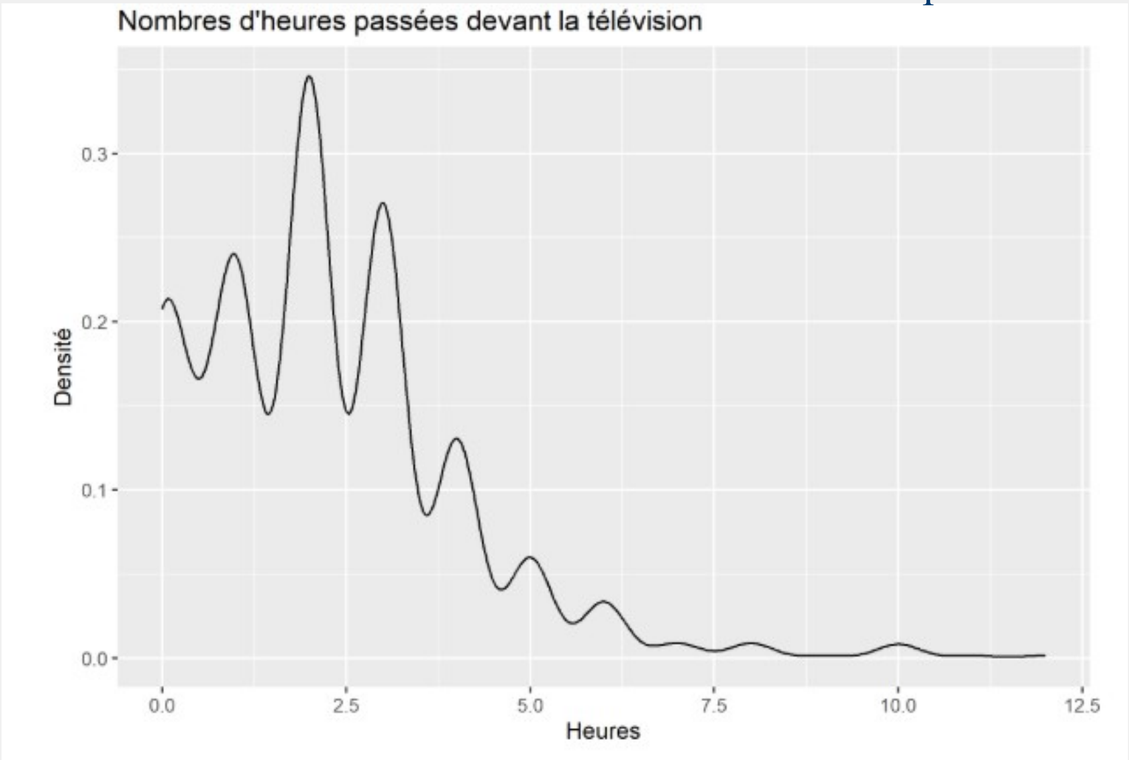
LES TYPES DE DATAVISUALISATION (1)

Le diagramme en barre (ou bar charts)
→ comparer les données par catégories



Pratique du sport selon la classe socio-professionnelle,
J. Larmarange et al., analyse-R

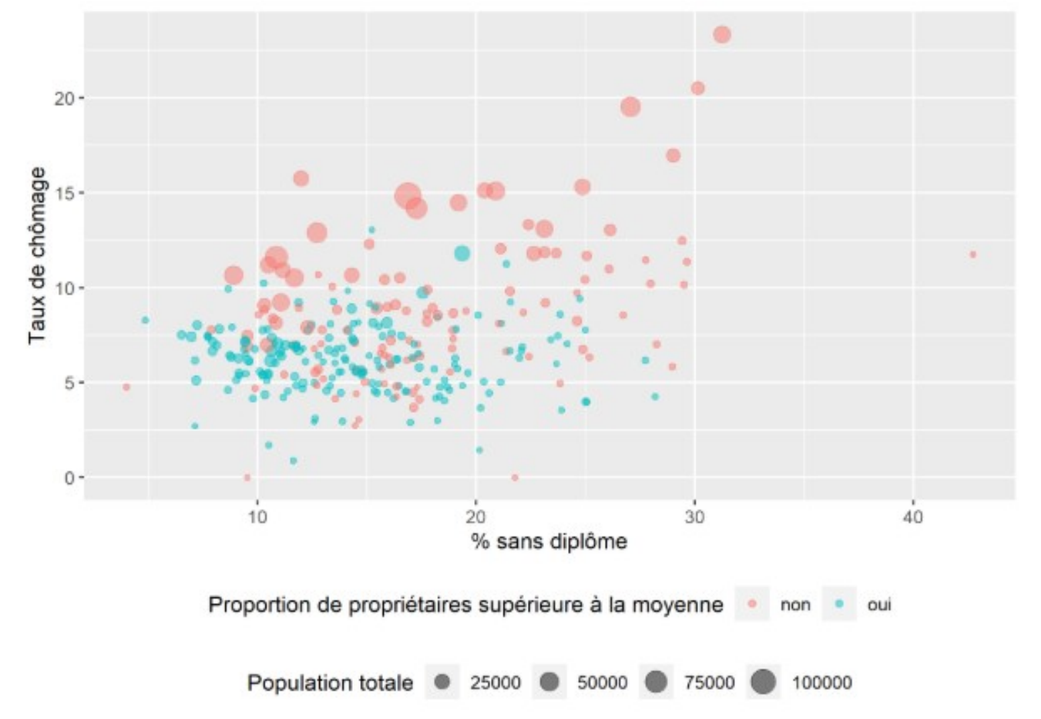
La courbe de densité (ou line graph)
→ visualiser des tendances à travers le temps



Nombre d'heures passées devant la télévision,
J. Larmarange et al., analyse-R

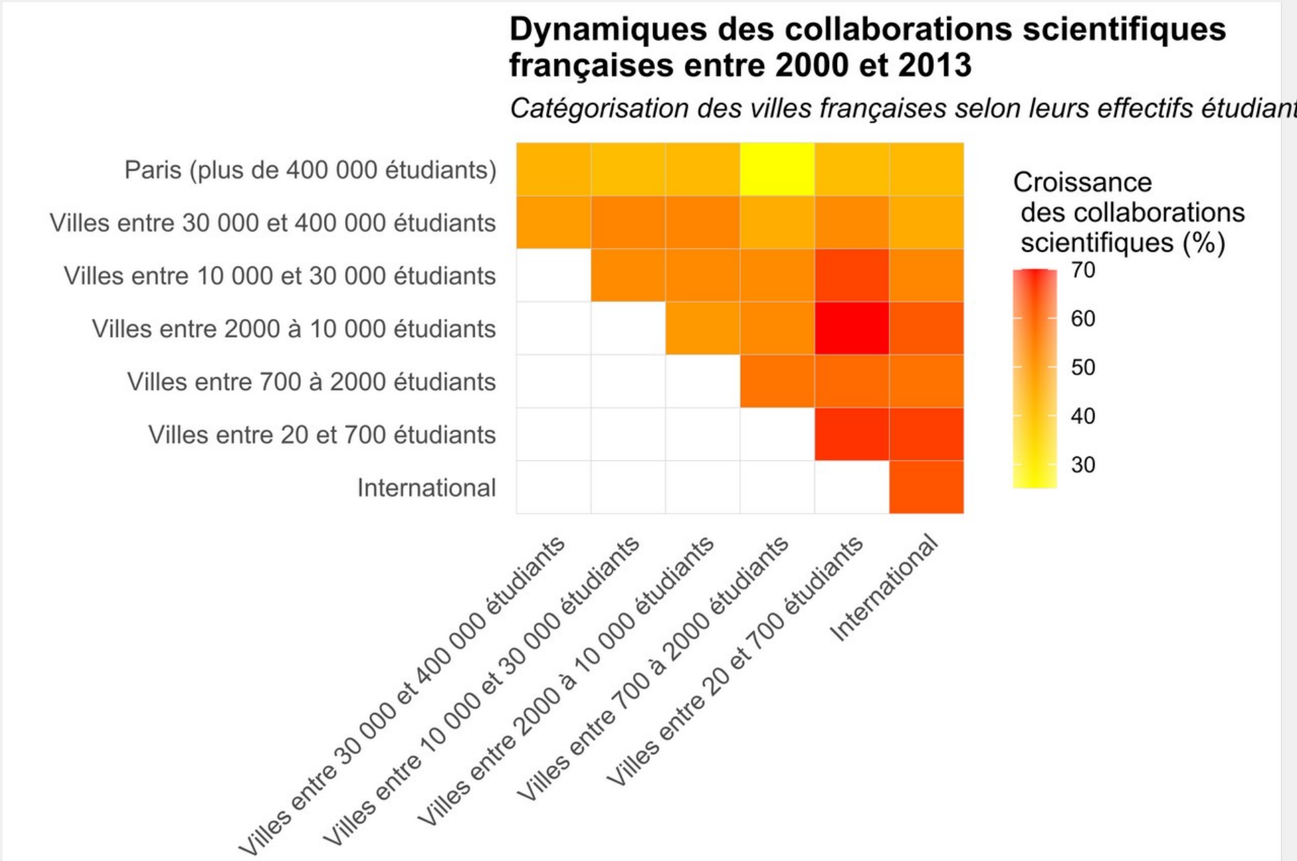
LES TYPES DE DATAVISUALISATION (2)

Le nuage de points (ou scatter plot)
→ visualiser des relations entre variables



Taux de chômage en fonction du diplôme,
J. Larmarange et al., analyse-R

La carte de chaleur (ou heat map)
→ visualiser la distribution des valeurs dans une matrice ou une table

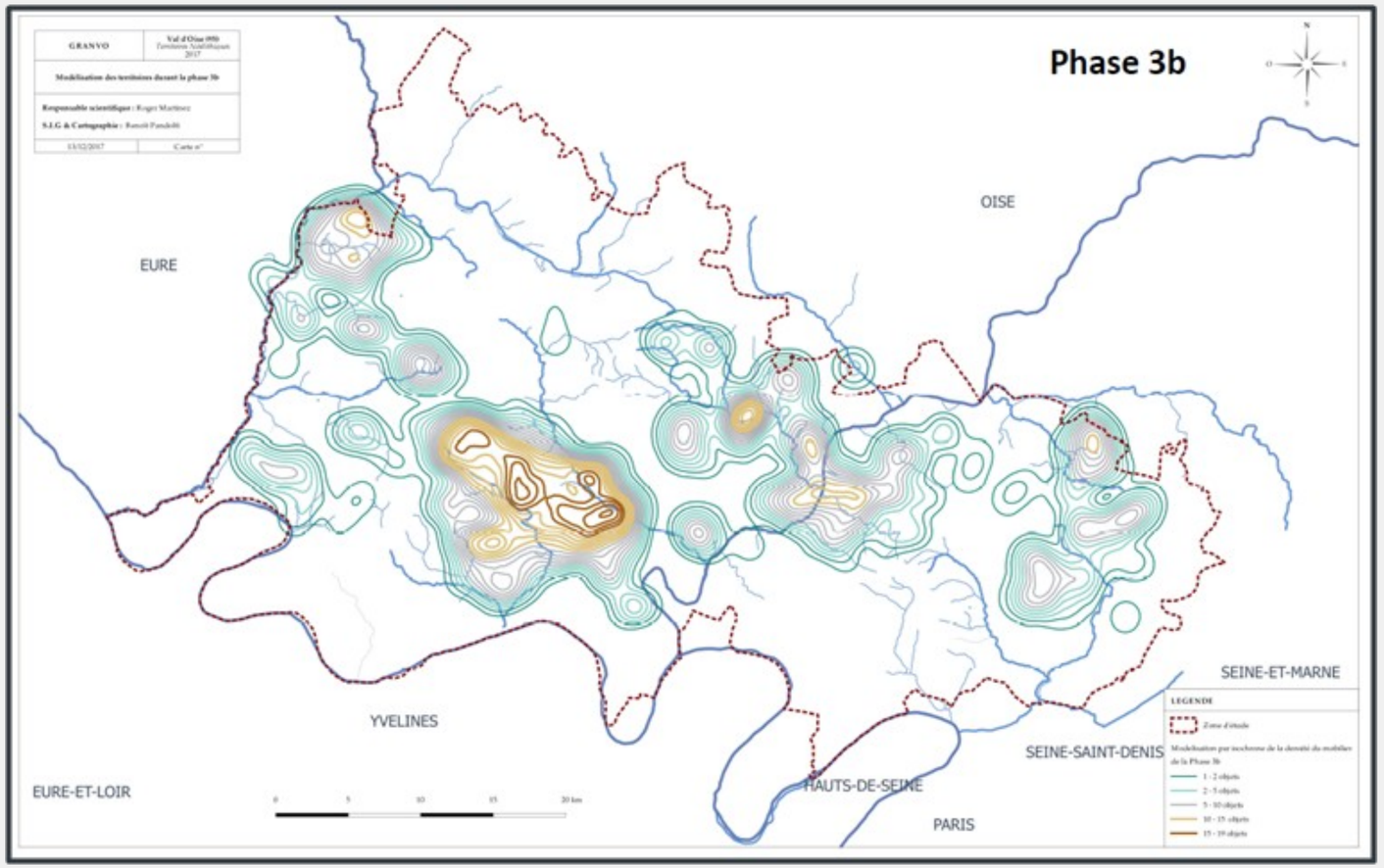


Clef de lecture : Entre 2000 et 2013, les collaborations entre les villes de 2000 à 10 000 étudiants et l'international ont augmenté de plus de 60 %.
Sur la même période, les collaborations scientifiques de Paris avec l'international n'ont augmenté que de 40 %,
et celles des 11 villes entre 30 000 et 400 000 étudiants et l'international, de 45 %.
Source: Science Citation Index Expanded (articles, recensions et lettres) ;
Nbre. d'inscriptions principales d'étudiants du sup. public en 2014, sans doubles comptes (données ouvertes du MESR).
Comptage entier fractionné à l'agglomération urbaine, moyenne mobile sur 3 ans.
Données du projet Labex SMS NETSCIENCE (UMR LISST, 2017). Réalisation : MM, UMR Géographie-cités, 2020.

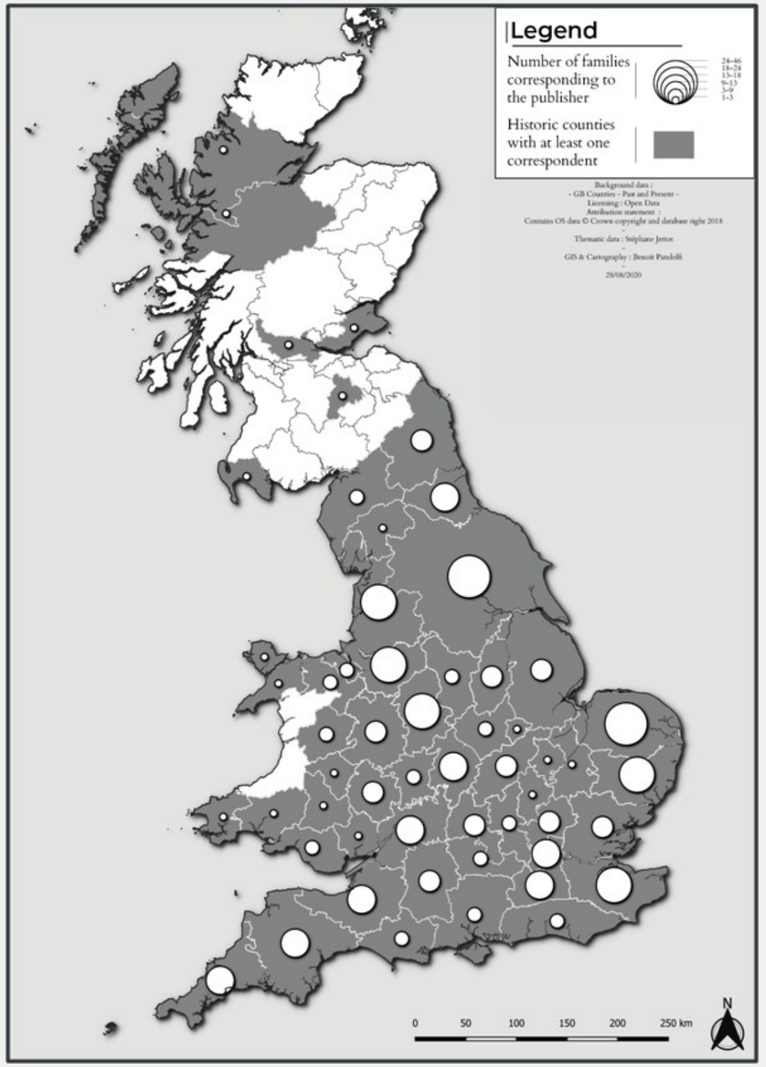
LES TYPES DE DATAVISUALISATION (3)

La carte

⇒ L’enjeu : spatialiser



Haches polies au territoires - GRANVO 2017

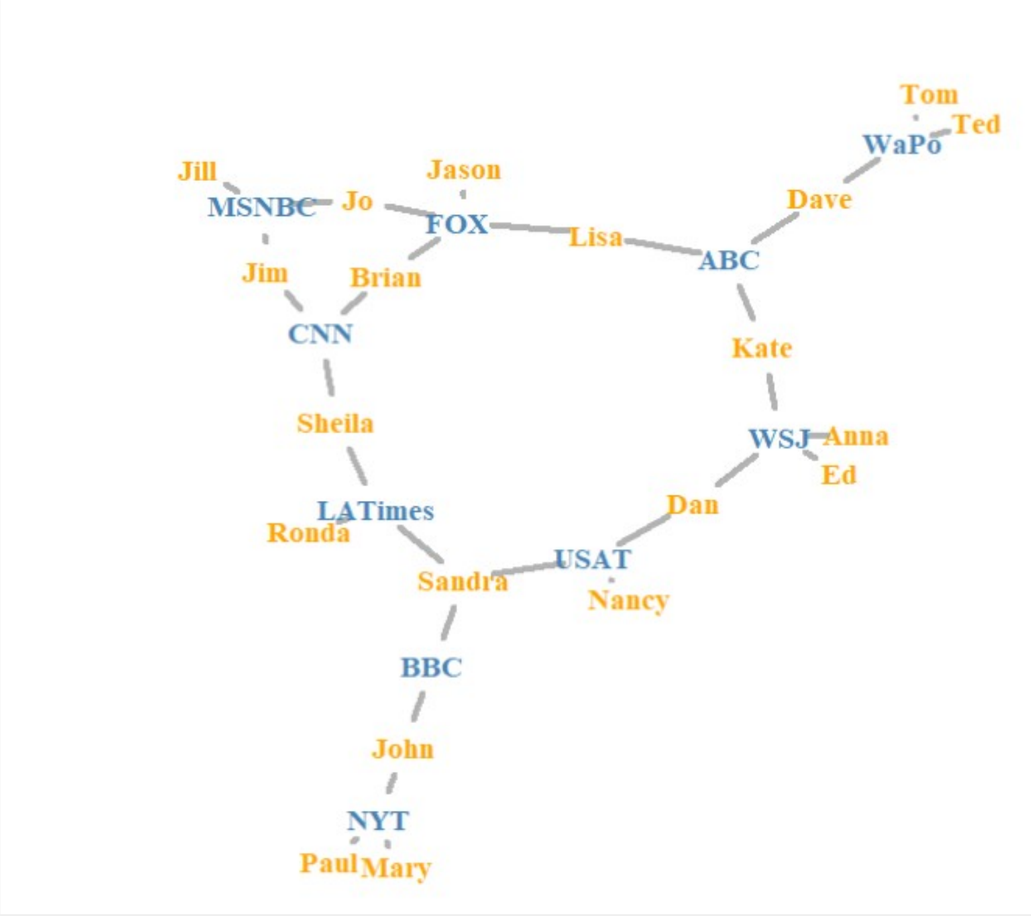


Stephane Jettot, An outdated impulse? Genealogical practices and Enlightenment values, avec J.-P. Zuniga, University Studies in the Enlightenment, Oxford University Studies in the Enlightenment, Liverpool, Voltaire Foundation, 2021, 320p.

LES TYPES DE DATAVISUALISATION (3)

Le diagramme en réseau (ou network diagram)

→ visualiser des relations (liens) entre des entités (nœuds) sous forme de réseaux

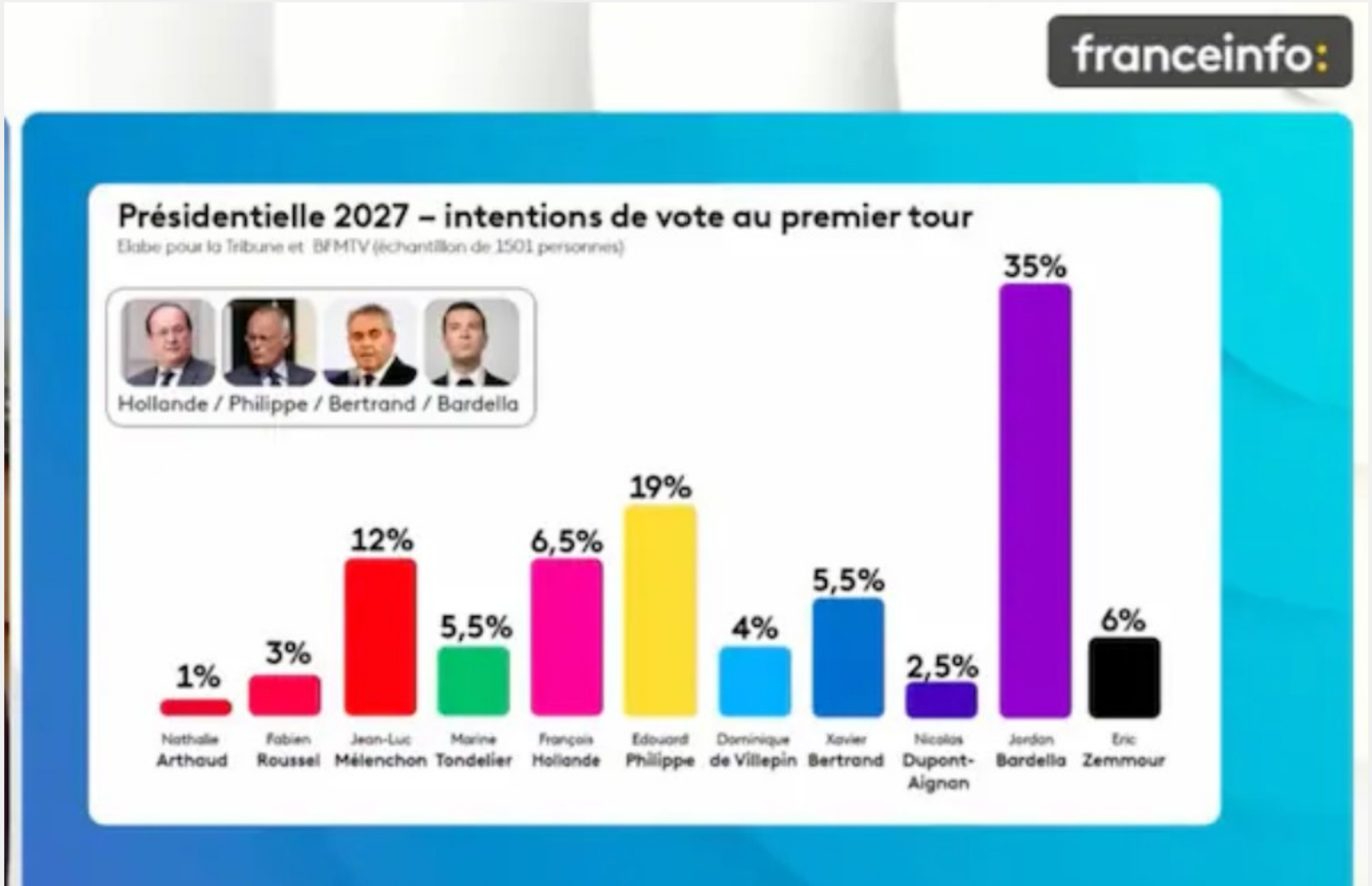


Médias et utilisateurs,

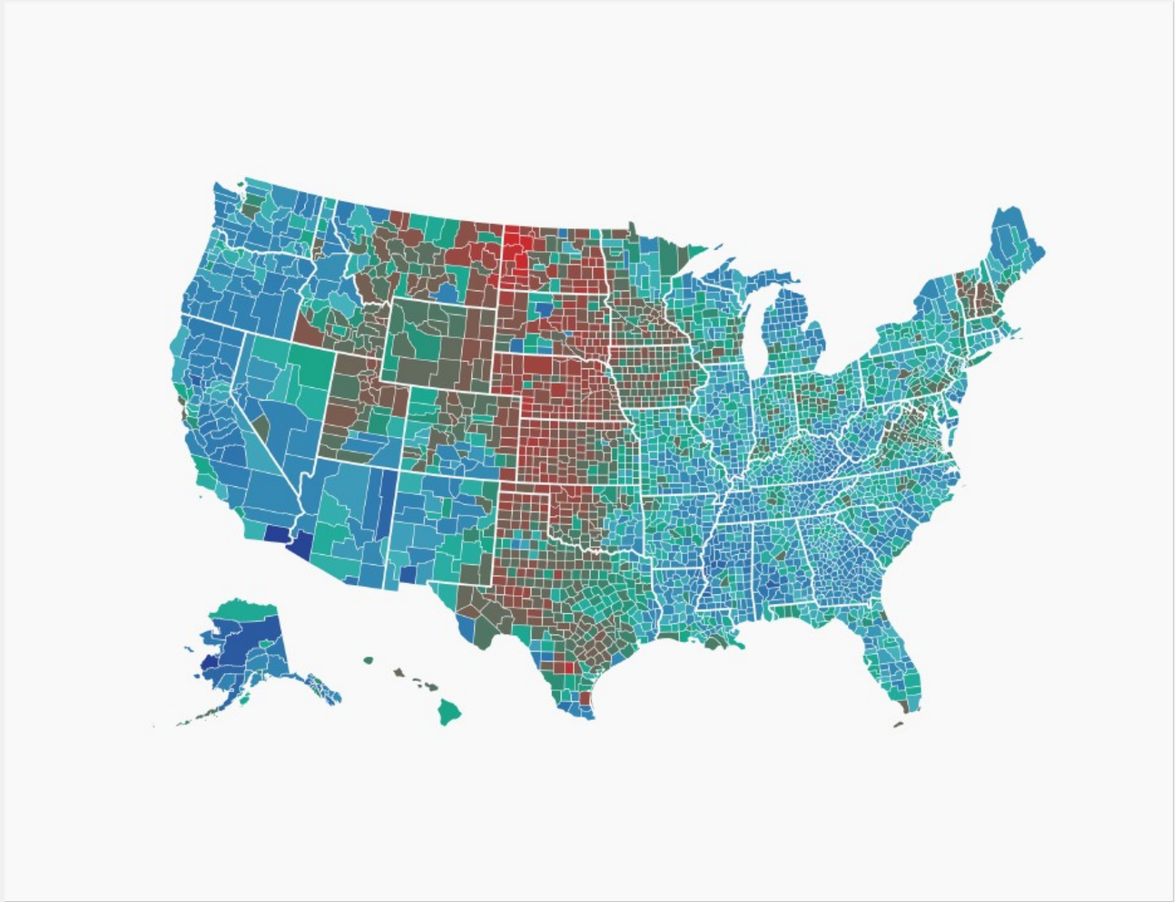
K. Ognyanova, Static and dynamic network visualisation with R, 2025

⇒ L'enjeu : choisir la bonne technique pour les données et le message à transmettre.

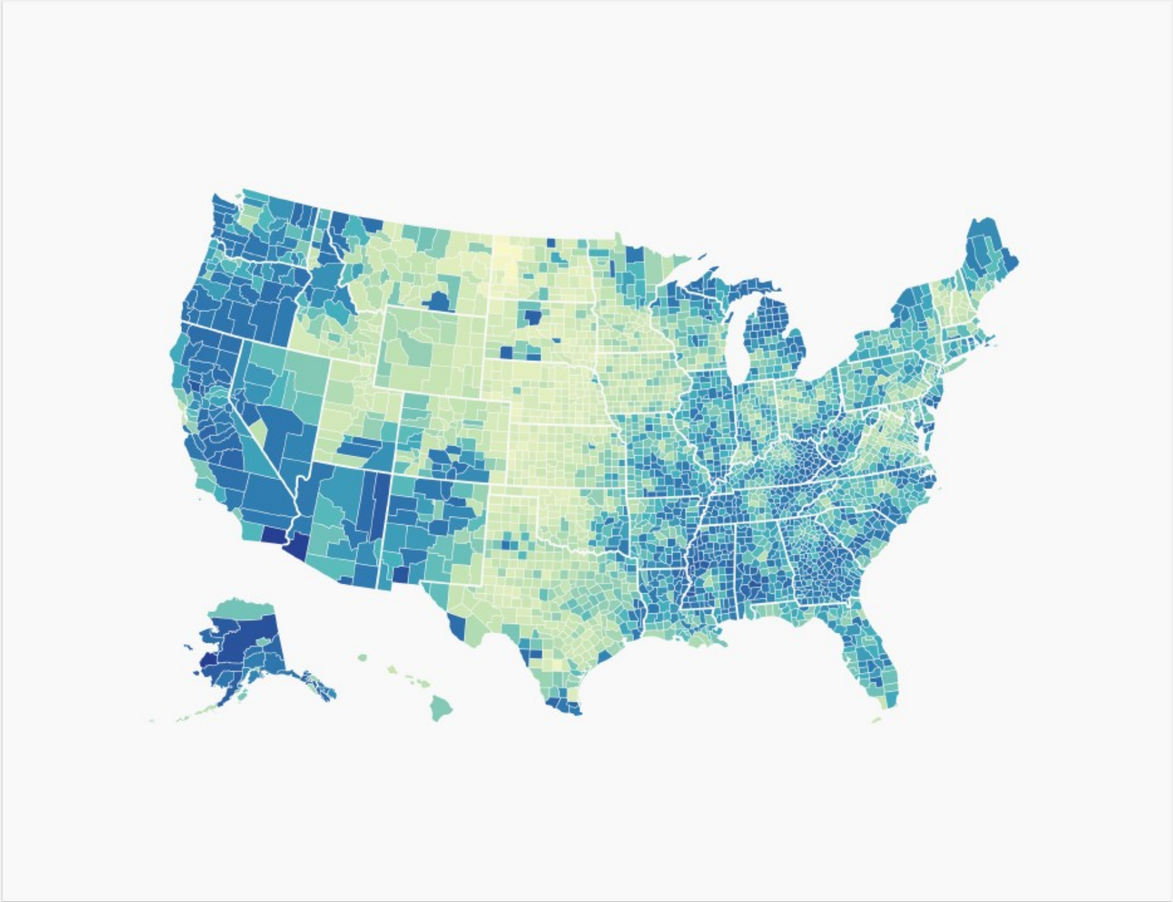
ÉDITORIALISER



ÉDITORIALISER

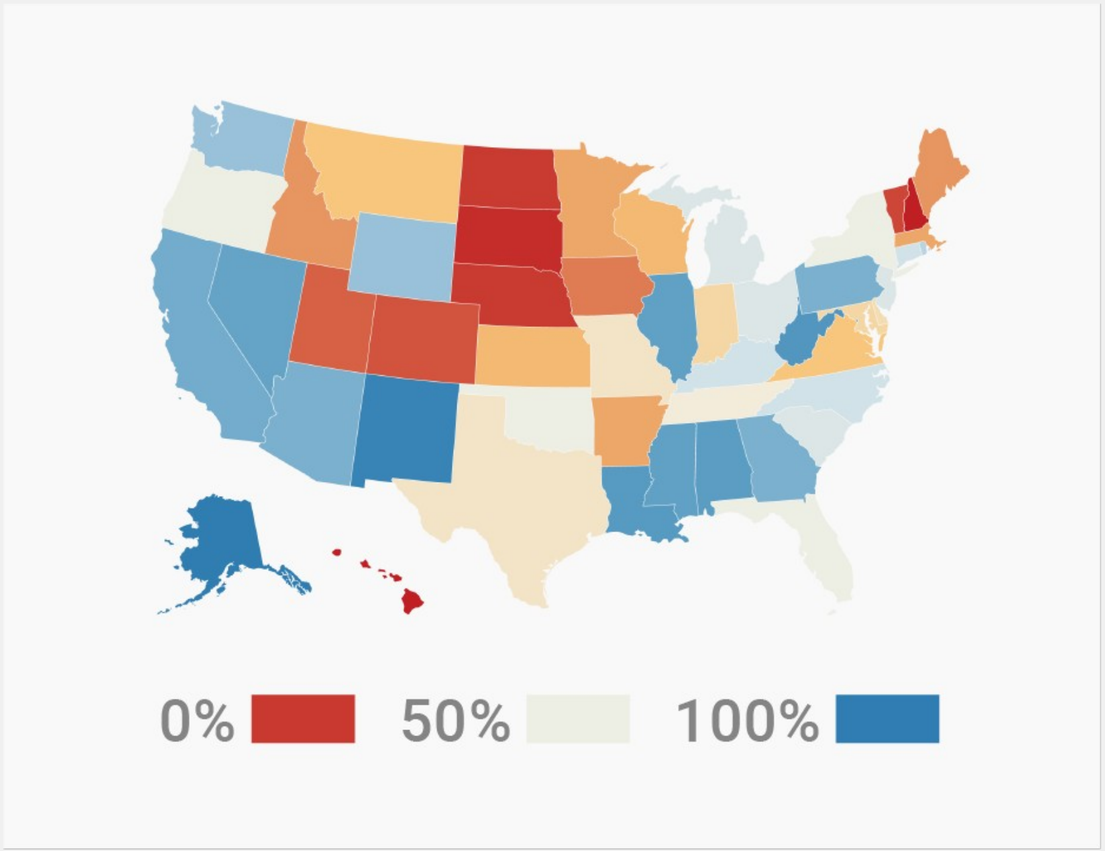


NOT IDEAL

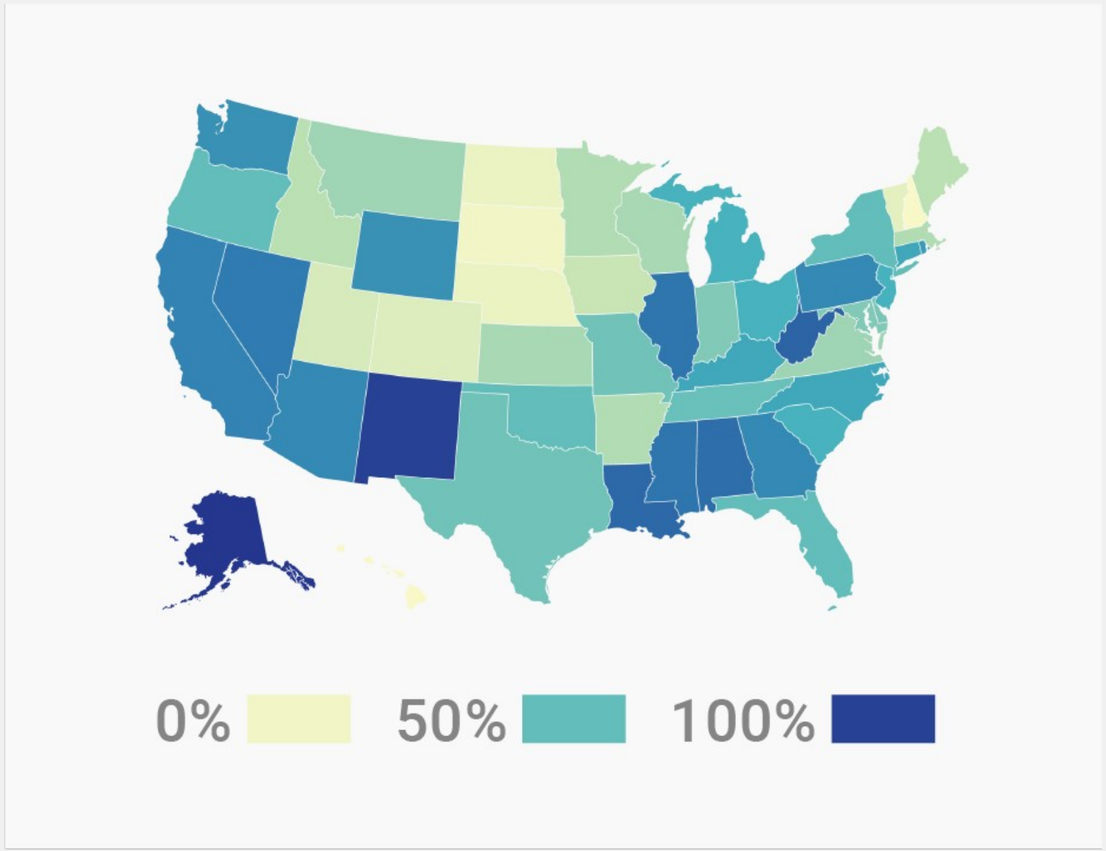


BETTER

ÉDITORIALISER



NOT IDEAL

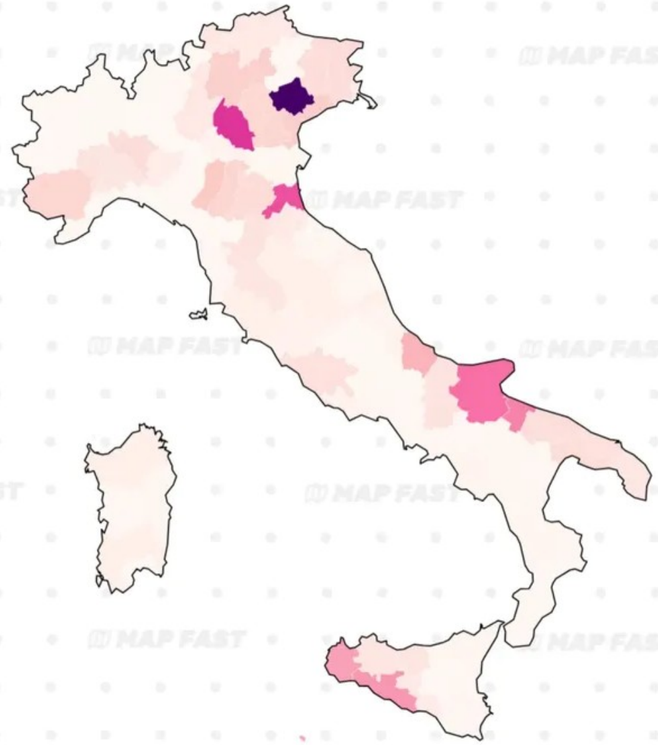


BETTER

ÉDITORIALISER

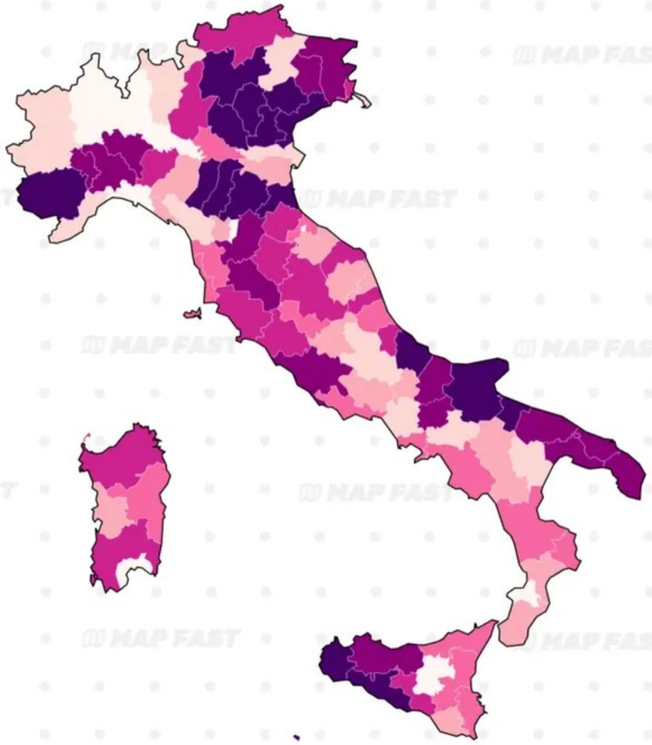
Linear ❌

Italian Wine Production, 2023



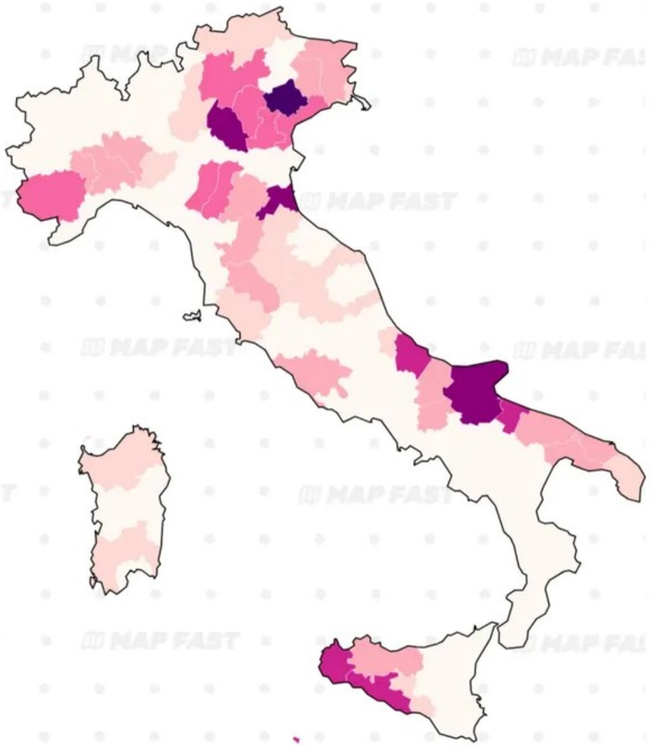
Quantile ❌

Italian Wine Production, 2023



Fisher-Jenks ✅

Italian Wine Production, 2023



ÉDITORIALISER



A *Rzeczpospolita* 19.02.1992.



B *Życie Warszawy* 30.12.1999.

CONSTRUIRE UNE DATAVISUALISATION

Les outils de la datavisualisation.

comment choisir une datavisualisation ?

La sémiologie graphique

LES OUTILS DE DATAVISUALISATION

Tableurs (Spreadsheets) Les outils les plus courants

- Capacités de dataviz de base : diagrammes en barres, lignes, histogrammes, nuages de points…
- Faciles à prendre en main, idéals pour des analyses rapides.
- Exemples : Excel, Google Sheets, LibreOffice Calc.

Outils spécialisés en datavisualisation Pour des visualisations plus avancées et interactives

- Tableaux de bord, filtres interactifs, cartes, diagrammes de réseaux, heatmaps…
- Adaptés à l’exploration, à la publication et au partage.
- Exemples : Tableau, Power BI, Datawrapper, RAWGraphs, Flourish.

LES OUTILS DE DATAVISUALISATION

Bibliothèques de programmation (R / Python / JavaScript) Pour des visualisations sur mesure, très flexibles

- Permettent un contrôle précis de l'esthétique, des interactions, de l'intégration dans des pipelines de recherche.
- Requiert une expertise technique.
- Exemples : ggplot2 (R), matplotlib / seaborn / plotly (Python), D3.js / Vega-Lite (JavaScript).

Outils de géovisualisation (Spatial) Pour les données géographiques ou spatialisées

- Cartes thématiques, flux, maillages, analyses spatiales.
- Exemples : QGIS, ArcGIS, Kepler.gl.

Outils de visualisation pour la recherche (science / big data) Pour des volumes élevés ou des données complexes

- Exploration de graphes, traitement interactif de grandes séries temporelles, visual analytics.
- Exemples : Gephi (réseaux), ParaView (scientifique), Observable (infovis interactive).

COMMENT CHOISIR UNE DATAVISUALISATION ?

Il n’y a pas une méthodologie unique pour choisir une datavisualisation

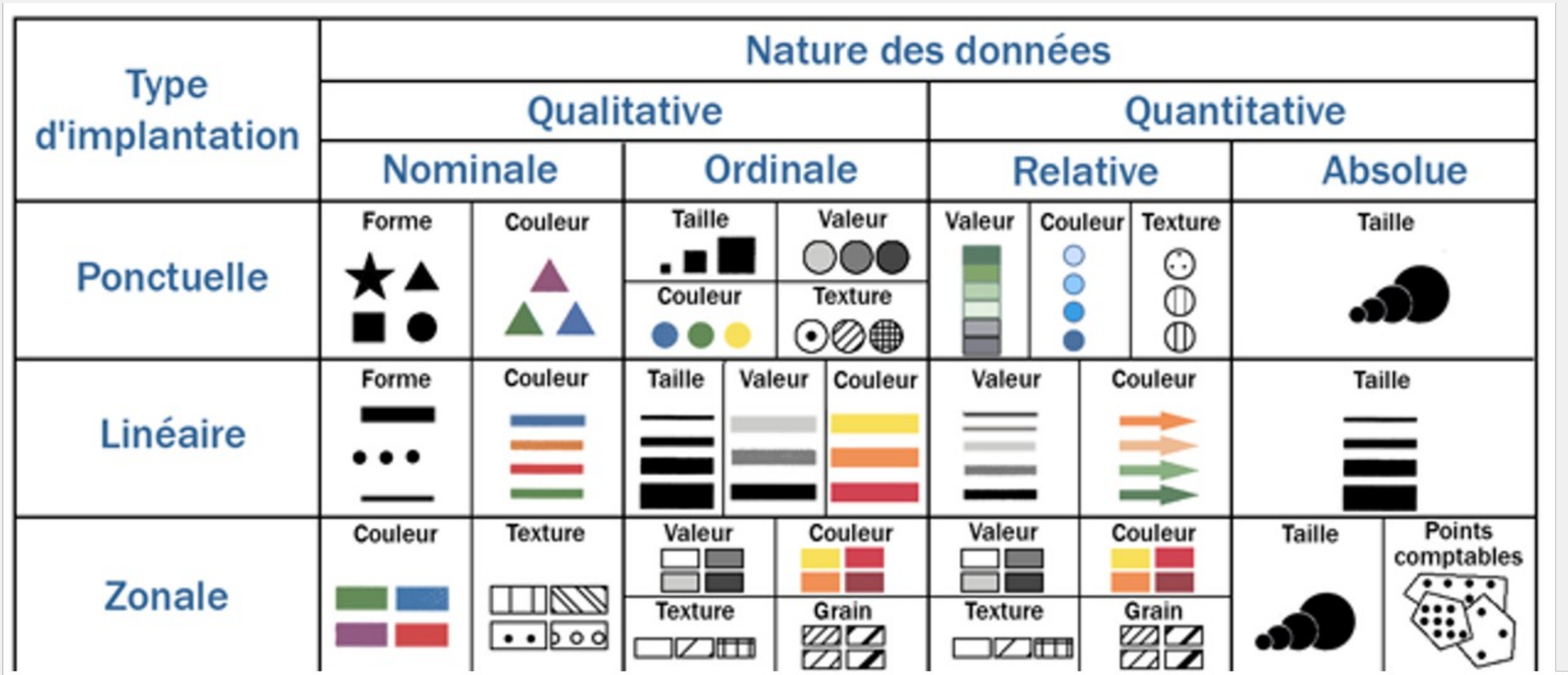
Approche	Point de départ	Avantage	Limite
Orientée tâche	Ce que fait l’utilisateur	Facile, pédagogique	Réduit parfois la richesse des données
Orientée données	Structure des données	Solide pour éviter les erreurs	Peu narrative
Orientée modèle (Nested Model)	Processus	Robuste, académique	Complexe à vulgariser
Orientée récit	Message / discours	Idéale en SHS et journalisme	Peut biaiser si mal contrôlée
Orientée exploration	Ce qu’on veut découvrir	Fait pour la recherche	Nécessite des outils avancés

LA SÉMIOLOGIE GRAPHIQUE

7 variables permettant de coder l'information :

- **Position (la plus puissante)**
- **Taille**
- **Valeur (clarté / intensité)**
- **Texture / grain**
- **Couleur (teinte)**
- **Orientation**
- **Forme**

Ces variables permettent de représenter différentes natures de données : quantitative, ordinale, catégorielle.



Source : C. Zanin et M.-L. Tremelo (2003), Savoir faire une carte. Aide à la conception et à la réalisation d'une carte thématique. Belin, Paris.

LA DATAVISUALISATION UN OUTIL EXPLORATOIRE ?

Révélateur d’hypothèse (patterns émergents, anomalies).

Structure le dialogue (itérations, tests).

Produire de la science (exploratoire → interprétation → méthode).

CC BY NC ND