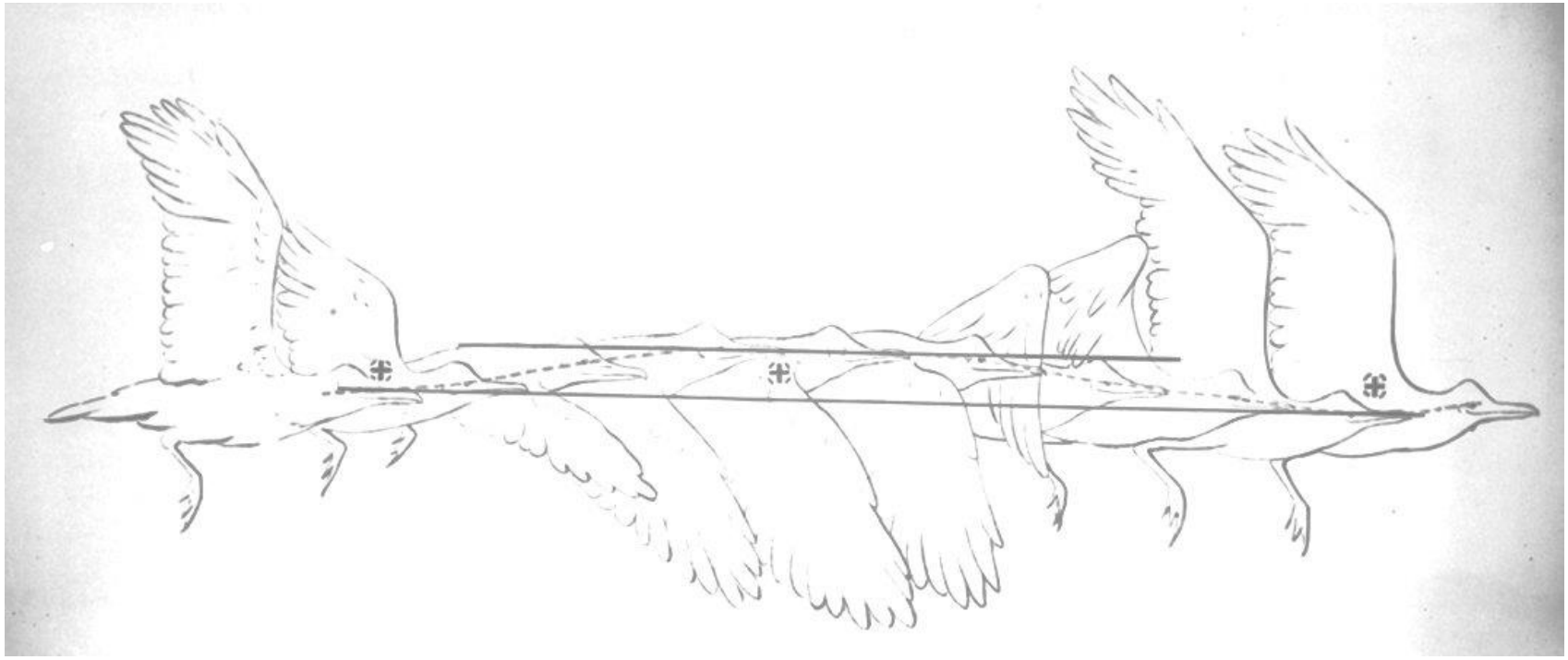
Visualisation d'informations (InfoVis) : Origines de la discipline



E.J Marey [source: salamandre.college-de-france.fr]

Visualisation d'informations (InfoVis) : Origines de la discipline

« raisonner par des moyens visuels (abstraits) sur des données (...) »

Comprendre nos propres données, leur faire dire ce que l'on ne voit pas

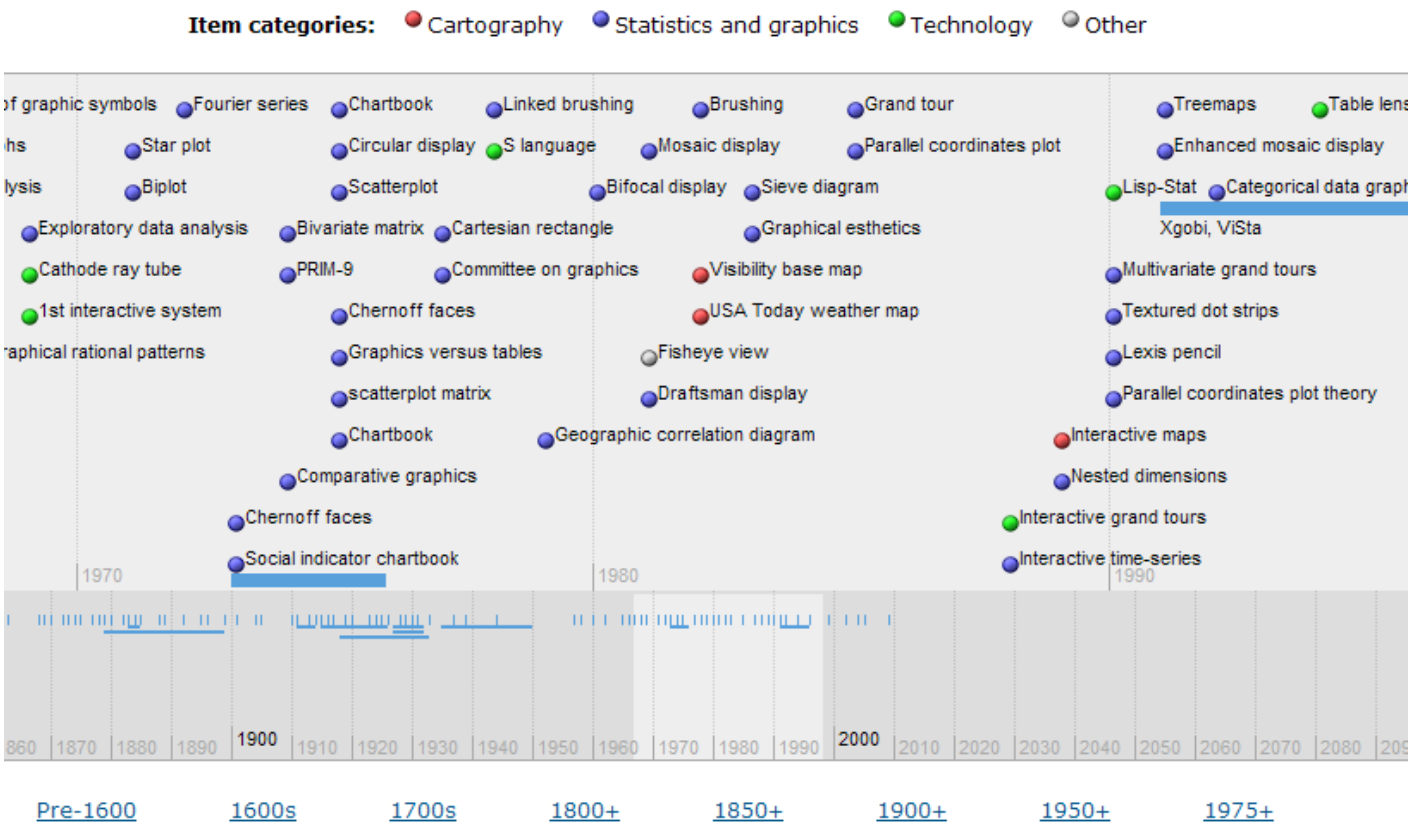
Pas une science, peut-être même pas une sous-discipline des sciences de l'information, mais par contre un jeu de méthodes, de pratiques, de règles* clairement définis en termes d'attendu: **assister l'analyste dans son effort de compréhension, de cognition, de découverte**

[...] Nous visualisons des informations pour raisonner sur des connaissances, pour documenter, communiquer et préserver ces connaissances [...] *

**E. R. TUFTE, Visual explanations, Graphic Press, Cheshire, Connecticut (1997)*

Selon Michael Friendly: Cartographie et statistique

Ici une sélection subjective plutôt orientée données spatiales, temporelles, ou spatio-temporelles



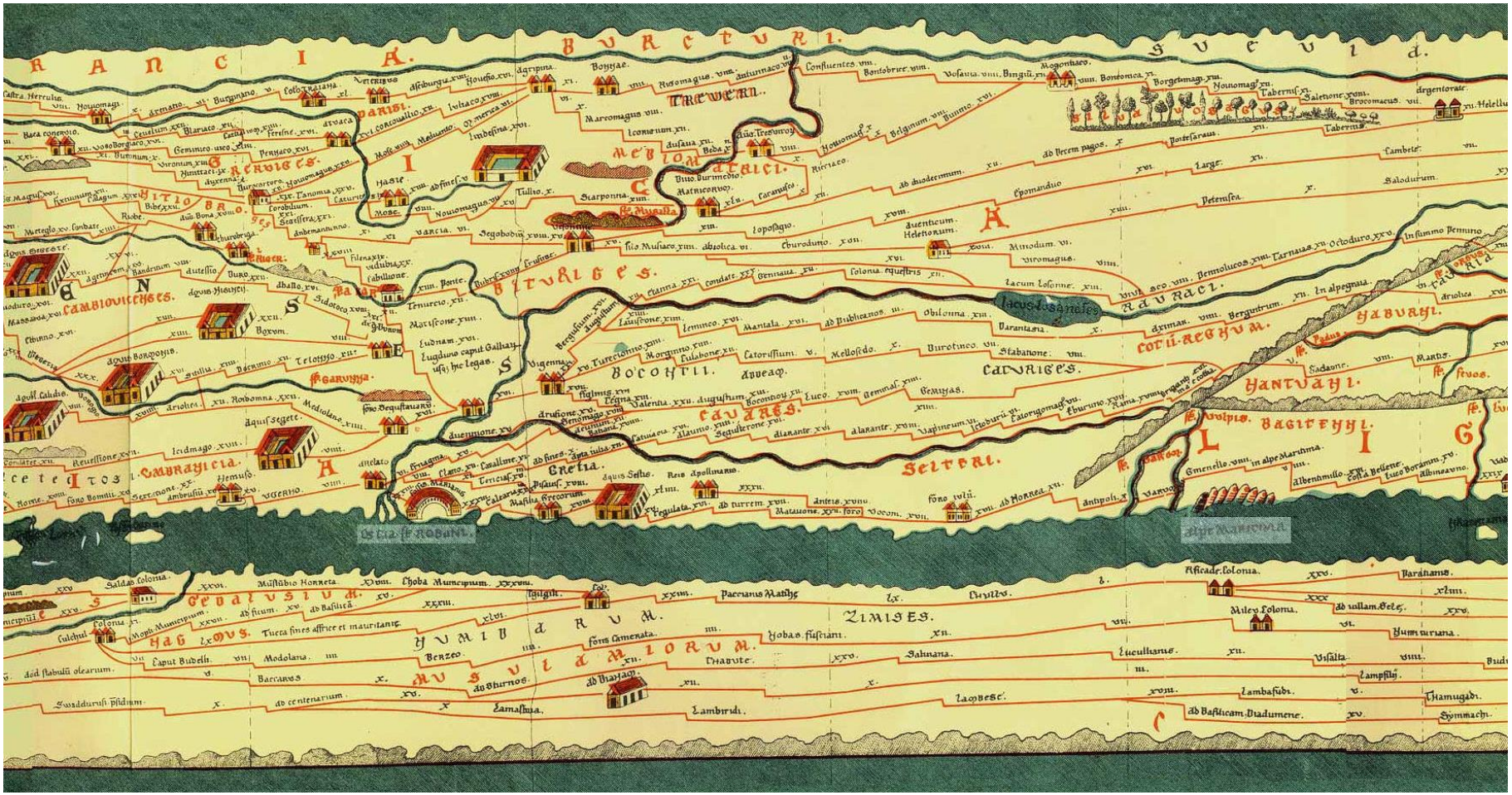
M.Friendly, Milestones project

InfoVis:Wiki; infovis.info; timeviz.net; *etc.*

La carte de Bedolina (2000 : 1000 B.C)
Parcellaire, champs, chemins, habitat et habitants.

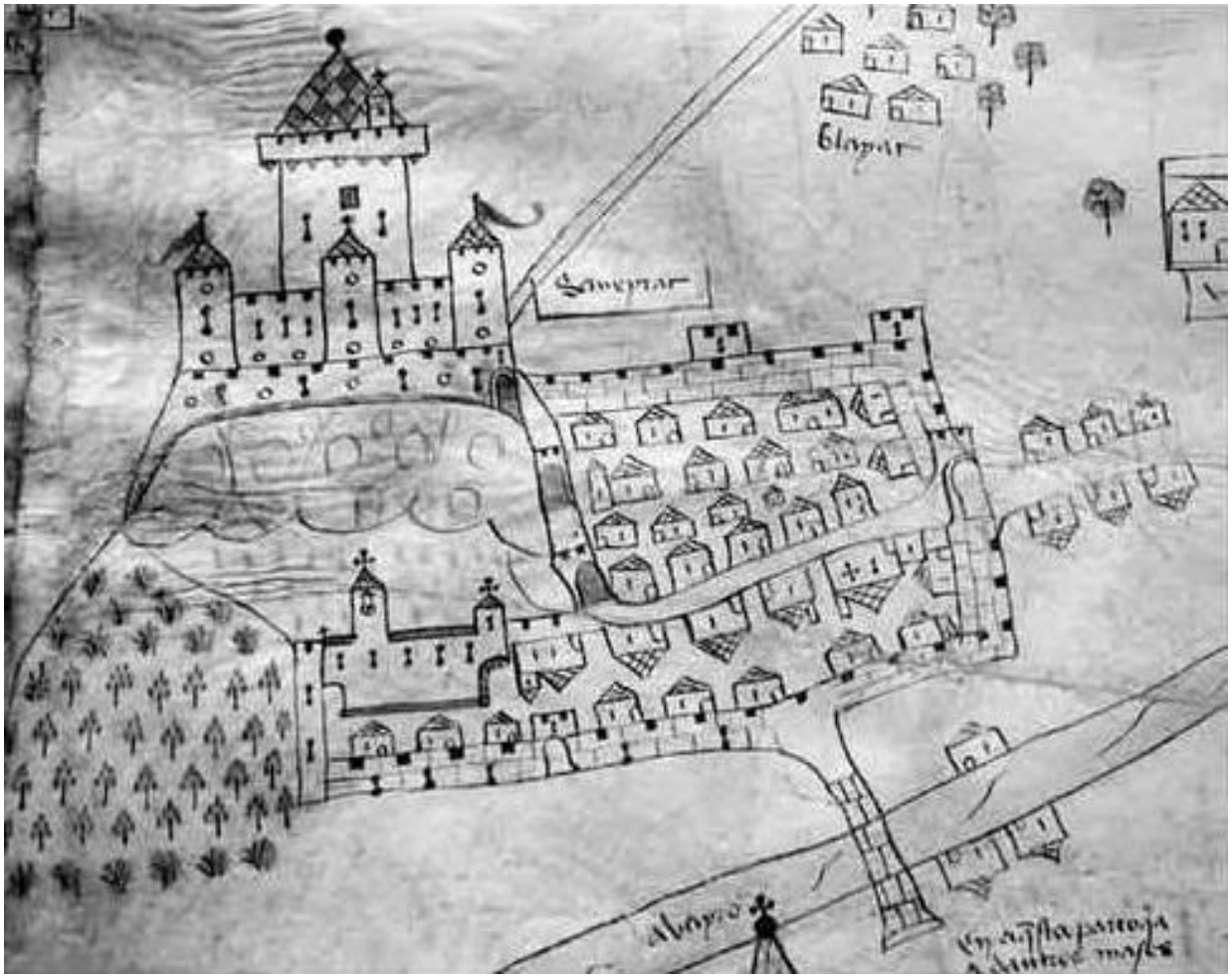
B. Holtzmann (Ed) L'art de l'antiquité 1. Les origines de l'Europe
Editions de la Réunion de musées nationaux, Editions Gallimard 1995]

Table de Peutinger: localiser par des glyphes, et de l'abstraction (jusqu'à un certain point).



http://fr.wikipedia.org/wiki/Table_de_Peutinger

>cartographie



Vue figurée, 1504,
(Ceci ne montre pas le territoire: rivières, châteaux, feux, mais des informations quantitatives sur le territoire.)

vues figurées : cartes confectionnées dans l'ensemble de l'Europe à partir du XIVe siècle pour servir de pièces à conviction devant les tribunaux

http://www.cairn.info/resume.php?ID_ARTICLE=RHIS_093_0621
Entre carte, image et pièce juridique : la vue figurée de la baronnie de Sévérac-le-Château (1504)
Juliette Dumasy

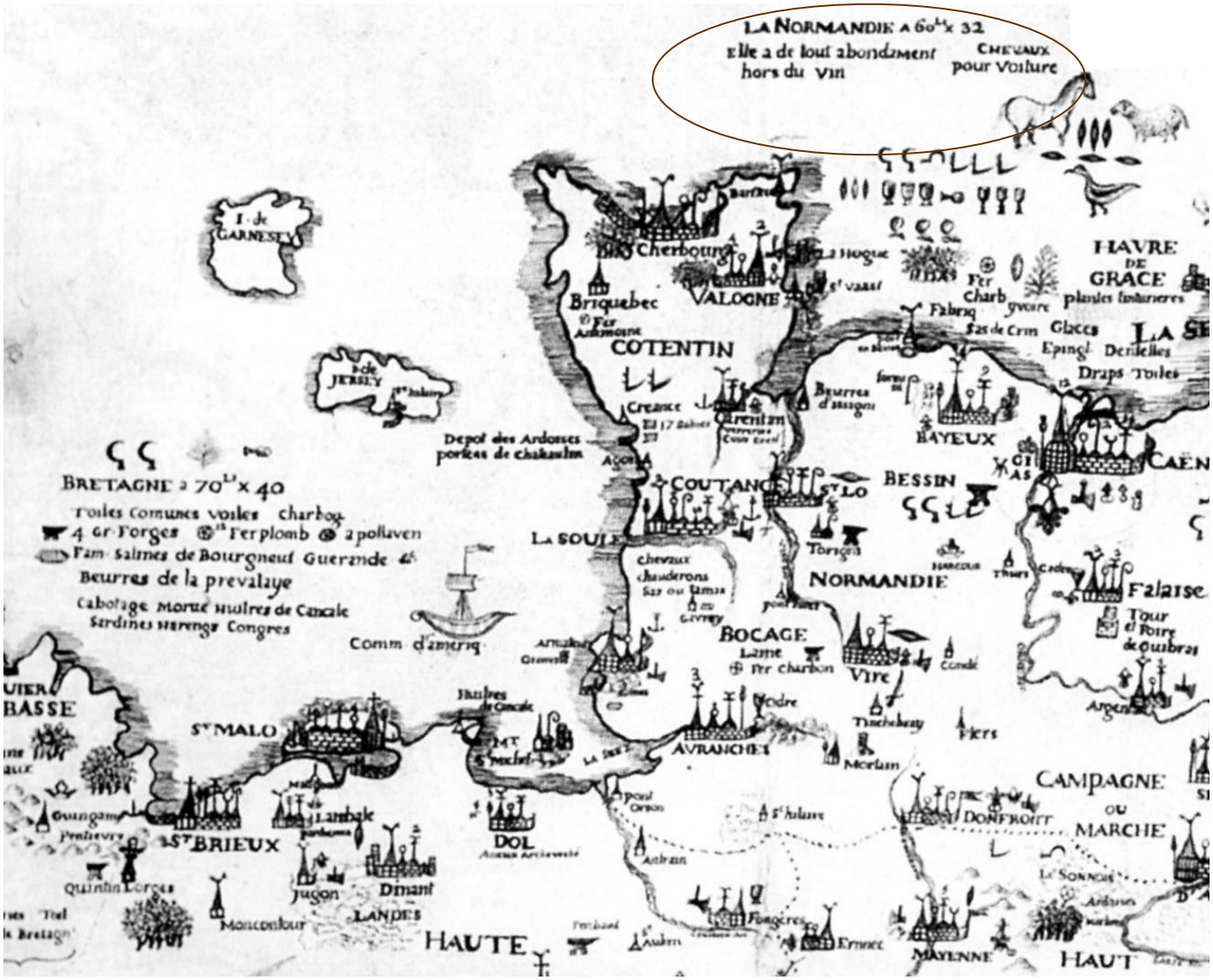
Un mouvement vers l'exactitude (et pas nécessairement vers plus d' informations)

Carte de Cassini

Relevés 1756 - 1789
Publication 1756 -1815.

Des unités spatiales d'abord

http://genealogie.orange.fr/v2/services-villes-villages/default.asp



Carte anonyme du XVIIIème siècle – activité agricoles, minières et manufacturières des provinces françaises.

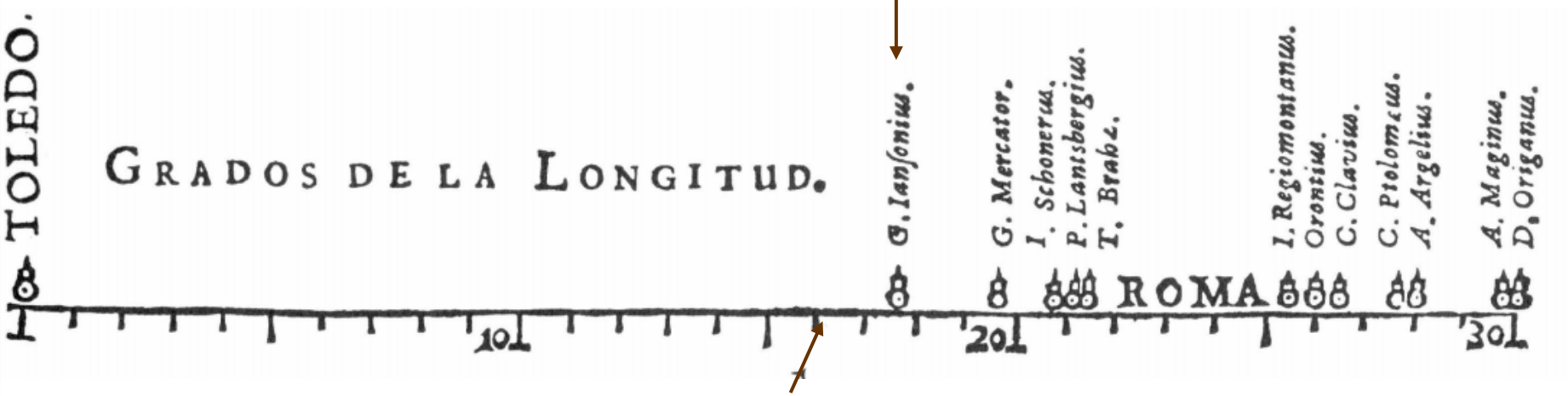
G.Palsky Des chiffres et des cartes. La cartographie quantitative au XIXe siècle CTHS 1996 p36

Première représentation visuelle de données statistiques

12 estimations de la différence en longitude entre Tolède et Rome
(1644 , M.F Van Langren)

Giovanni Antonio Magini (en latin, Maginus) astronome, astrologue, cartographe, mathématicien et universitaire italien ~1582.

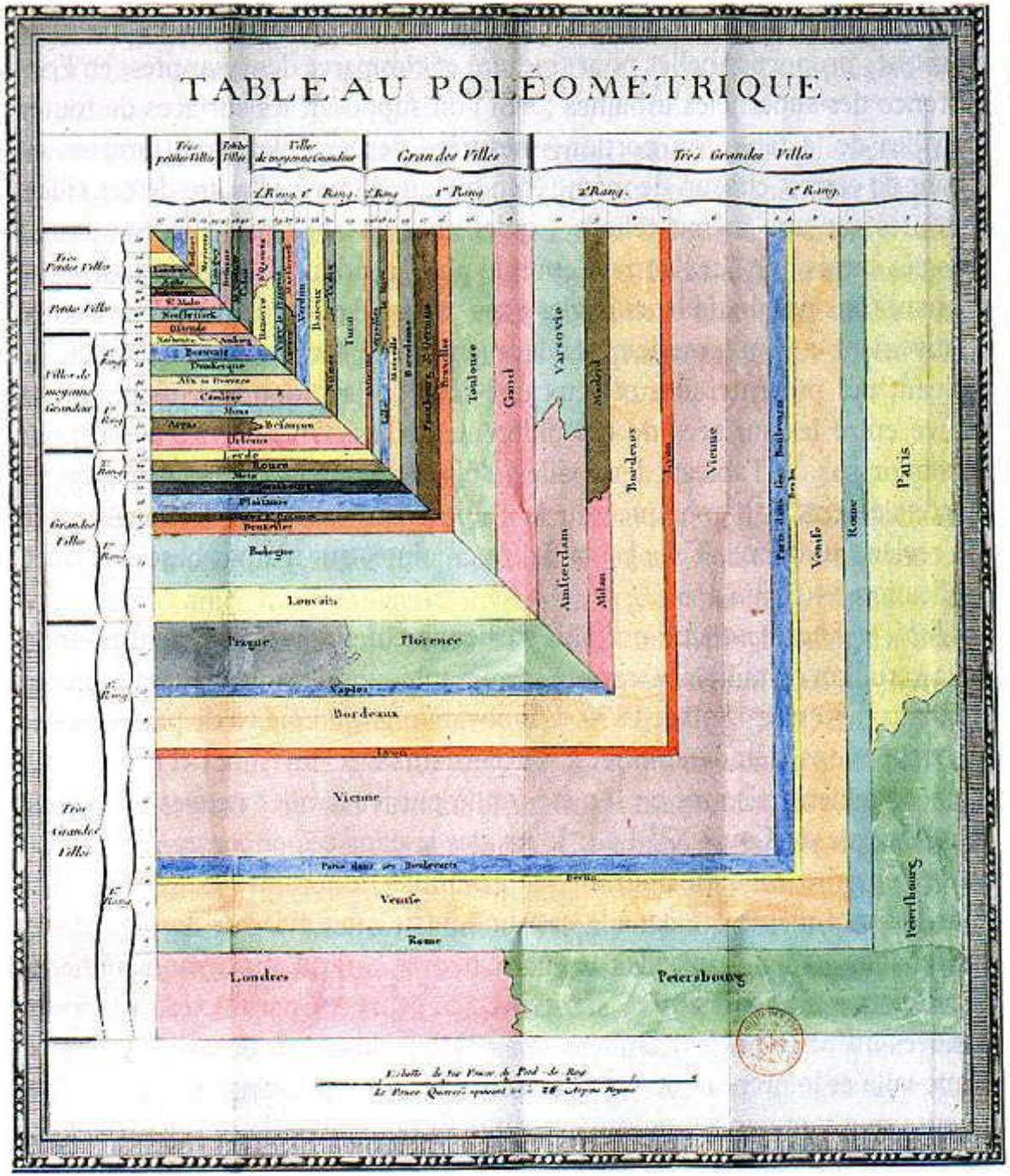
Johannes Janssonius, 1588 - 1664 à Amsterdam, cartographe néerlandais, ~1615.



16°30

* M. Friendly A brief history of data visualisation
http://www.datavis.ca/papers/hbook.pdf

>statistique

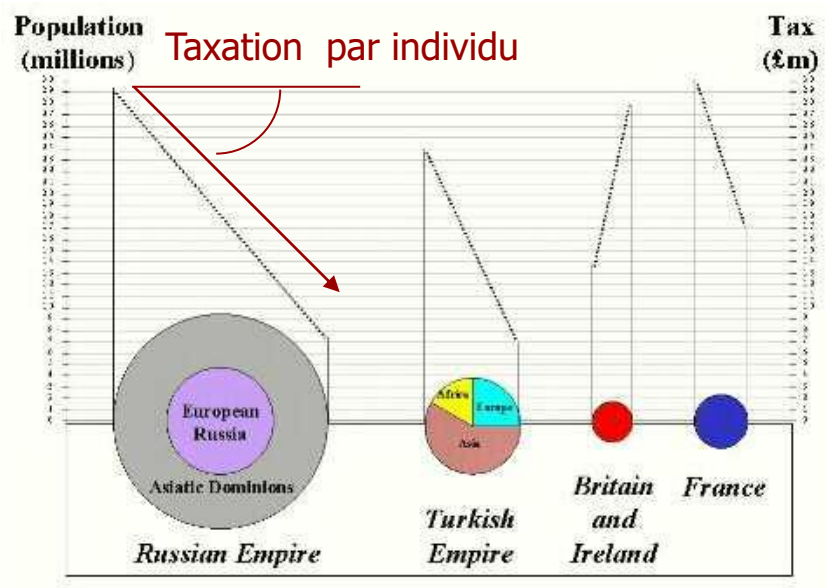


http://euclid.psych.yorku.ca/SCS/Gallery/images/palsky/defourcroy3.jpg

Tableau poléométrique
Charles de Fourcroy 1782

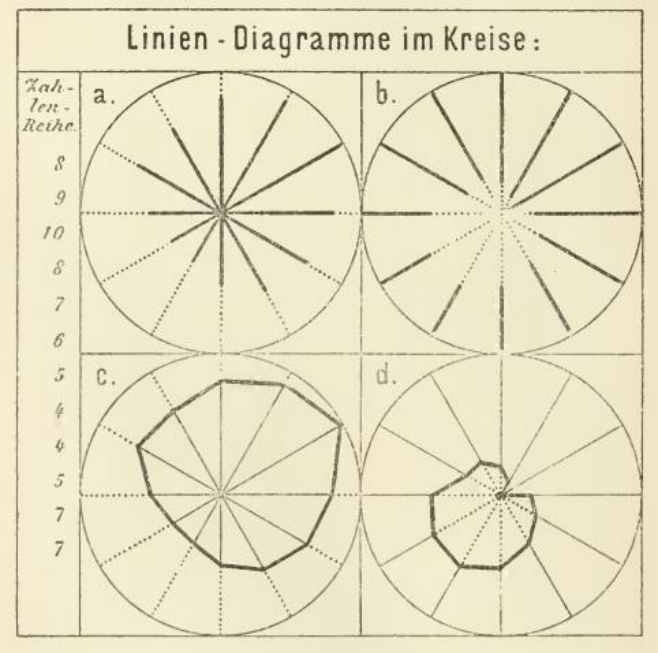
Figures géométriques pour comparer les démographies par superposition
un « tableau graphique » précoce

Des quantités, une classification
Clustering précoce

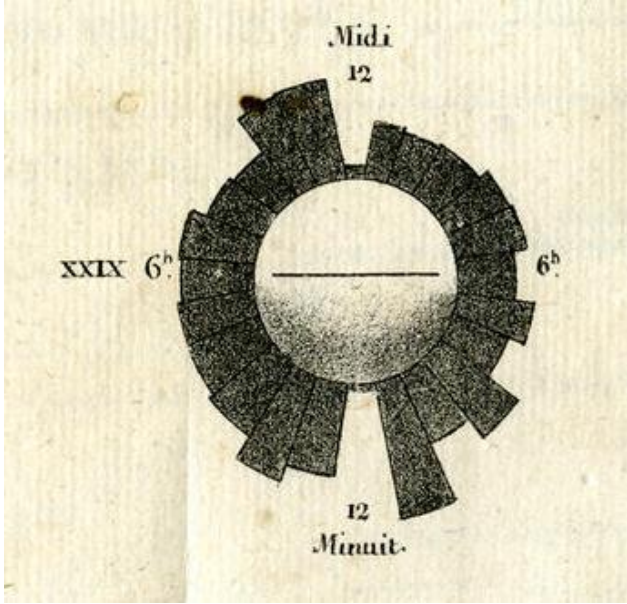


Statistical chart
William Playfair 1801

Et beaucoup d'autres inventions…
cf. ressources en ligne et références bibliographiques

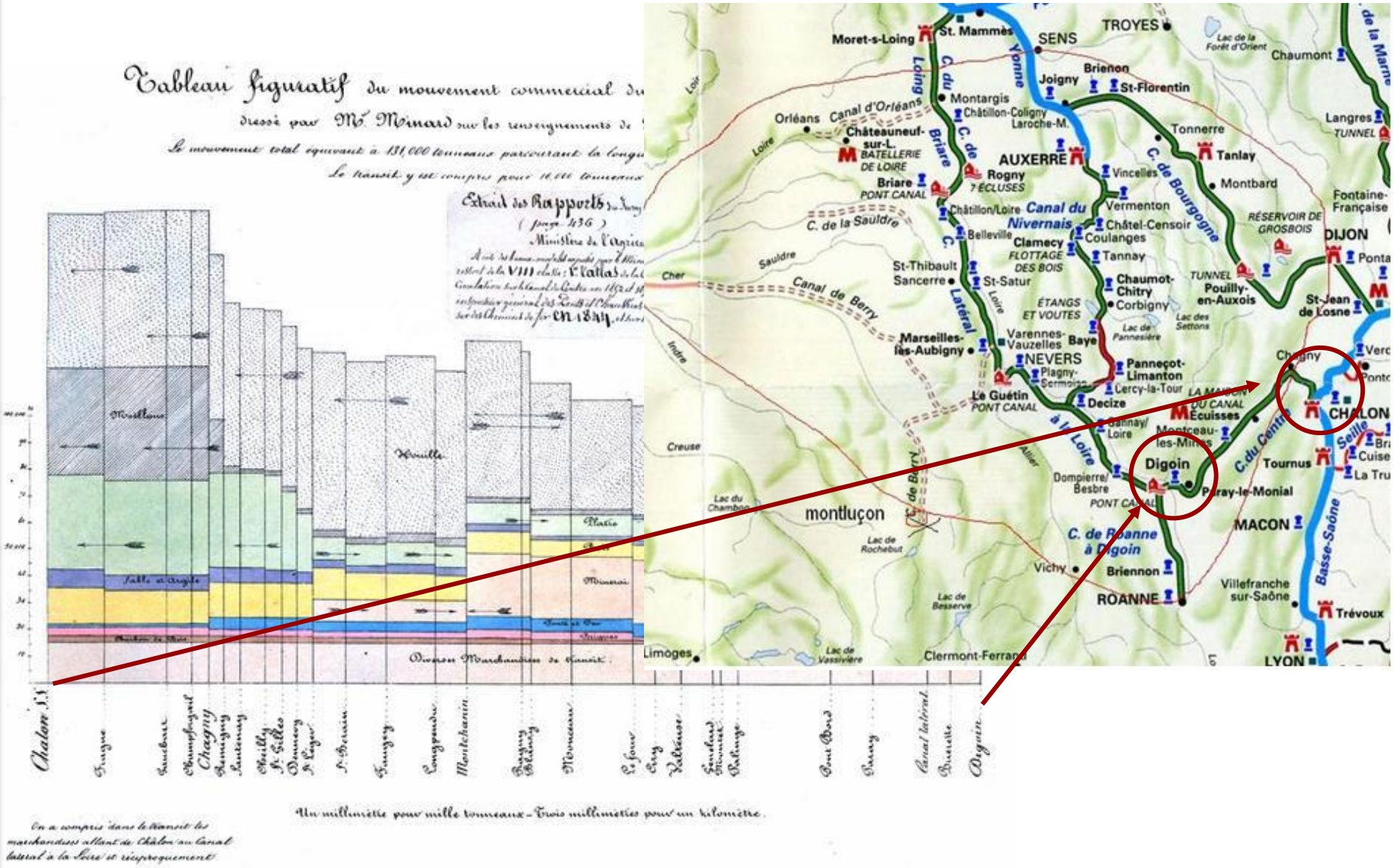


Star plot
Georg von Mayr, 1877



Polar-area charts
André Michel Guerry 1829
Récurrence de phénomènes cycliques

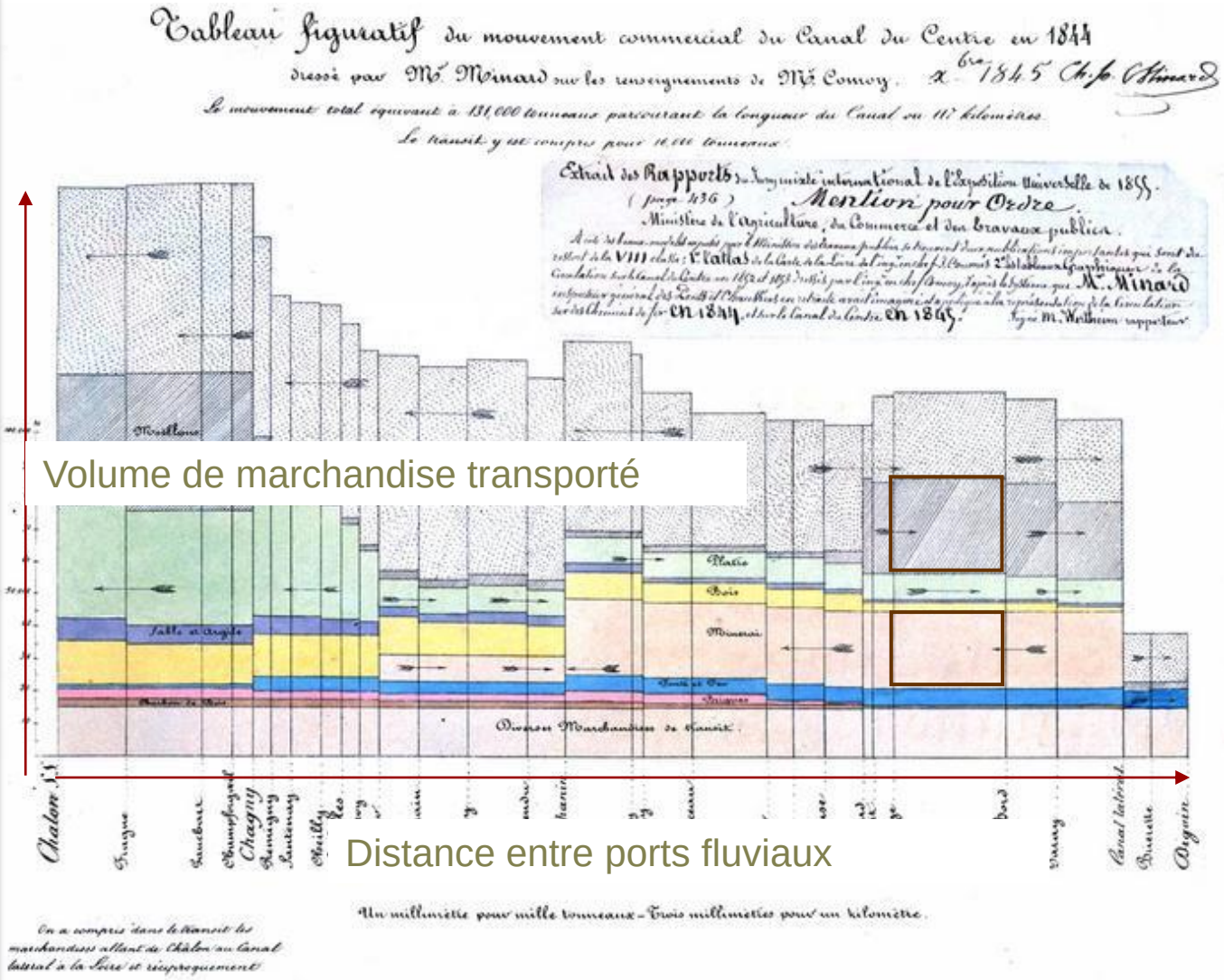
>statistique



http://bordabord.org/news/nouvelle-antenne-pour-les-canaux-du-centre-de-la-france

http://euclid.psych.yorku.ca/SCS/Gallery/milestone/sec5.html

>statistique



Un grand précurseur:
Charles-Joseph Minard

Tableau figuratif
1844

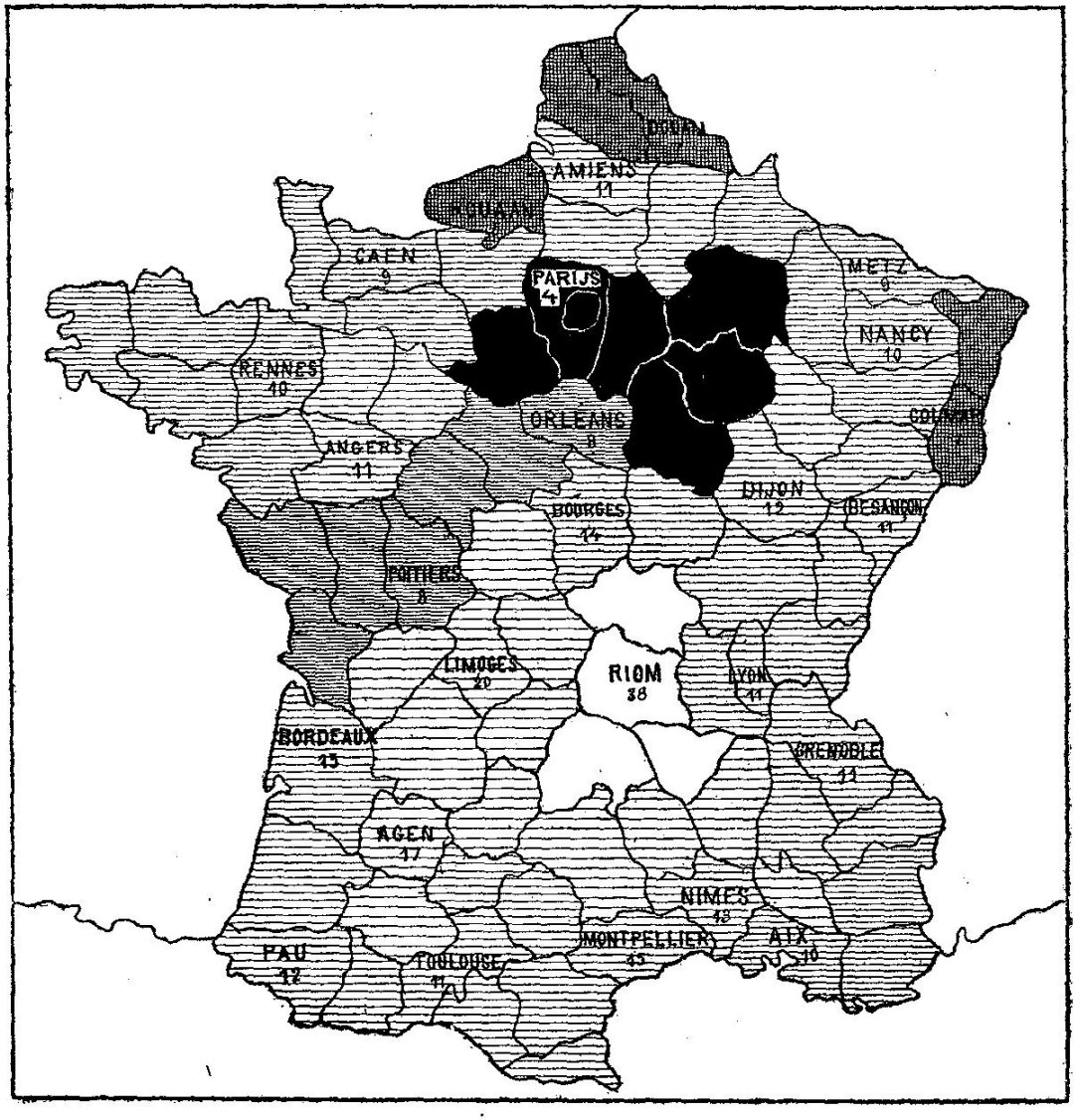
Largeur = distance
Hauteurs ~ quantités
Surface ~ coût

[An early form of the mosaic plot]

Déjà: Faire émerger une information « nouvelle »

http://euclid.psych.yorku.ca/SCS/Gallery/milestone/sec5.html

>cartographie
+ statistique



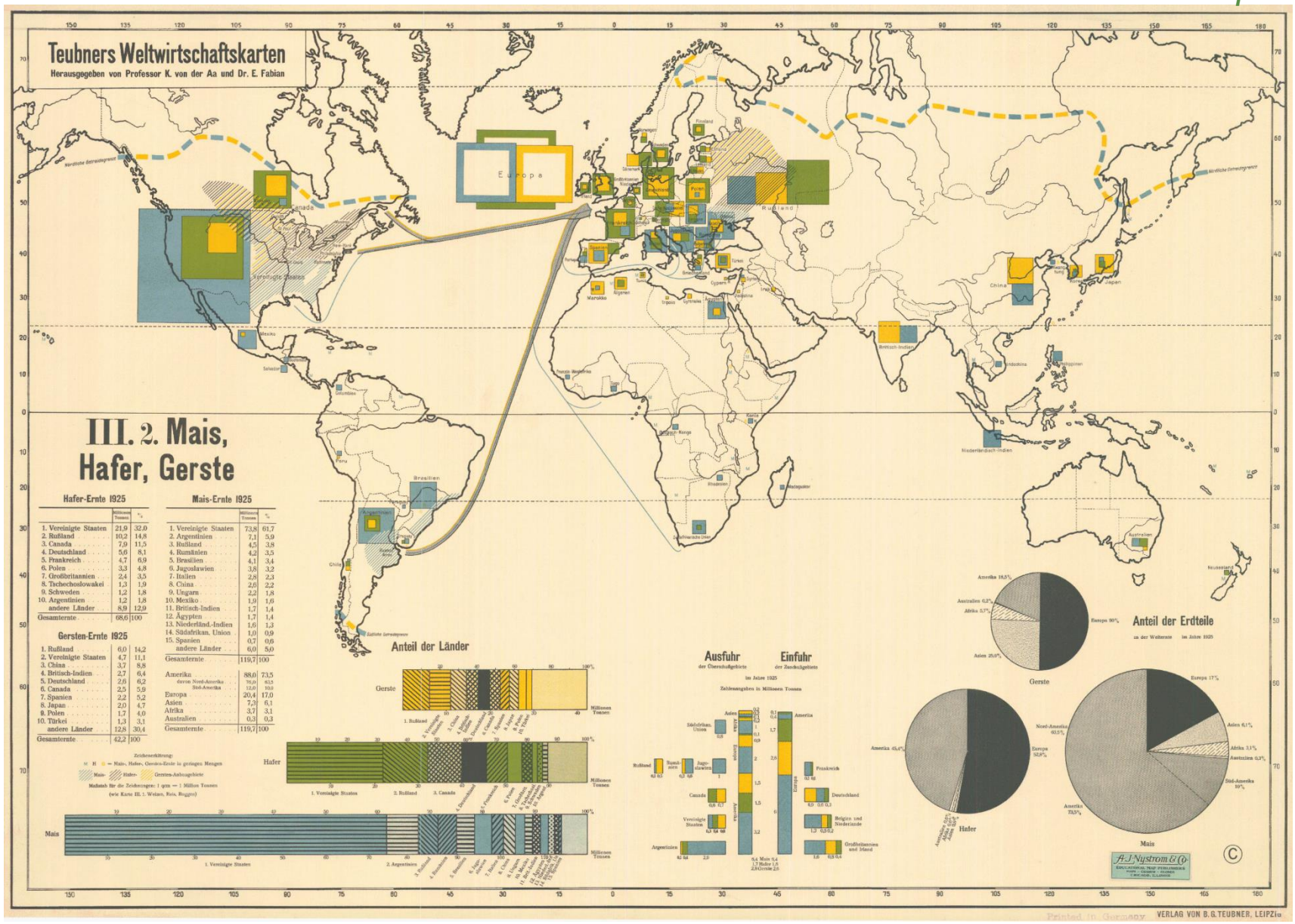
Deux héritages combinés:
« âge d'or des cartes et diagrammes statistiques »

Carte figurative de l'instruction populaire de la France.
Dupin, C. (1826).
Jobard. BNF: Ge C 6588

http://www.InfoVis.info/index.php?words=dupin

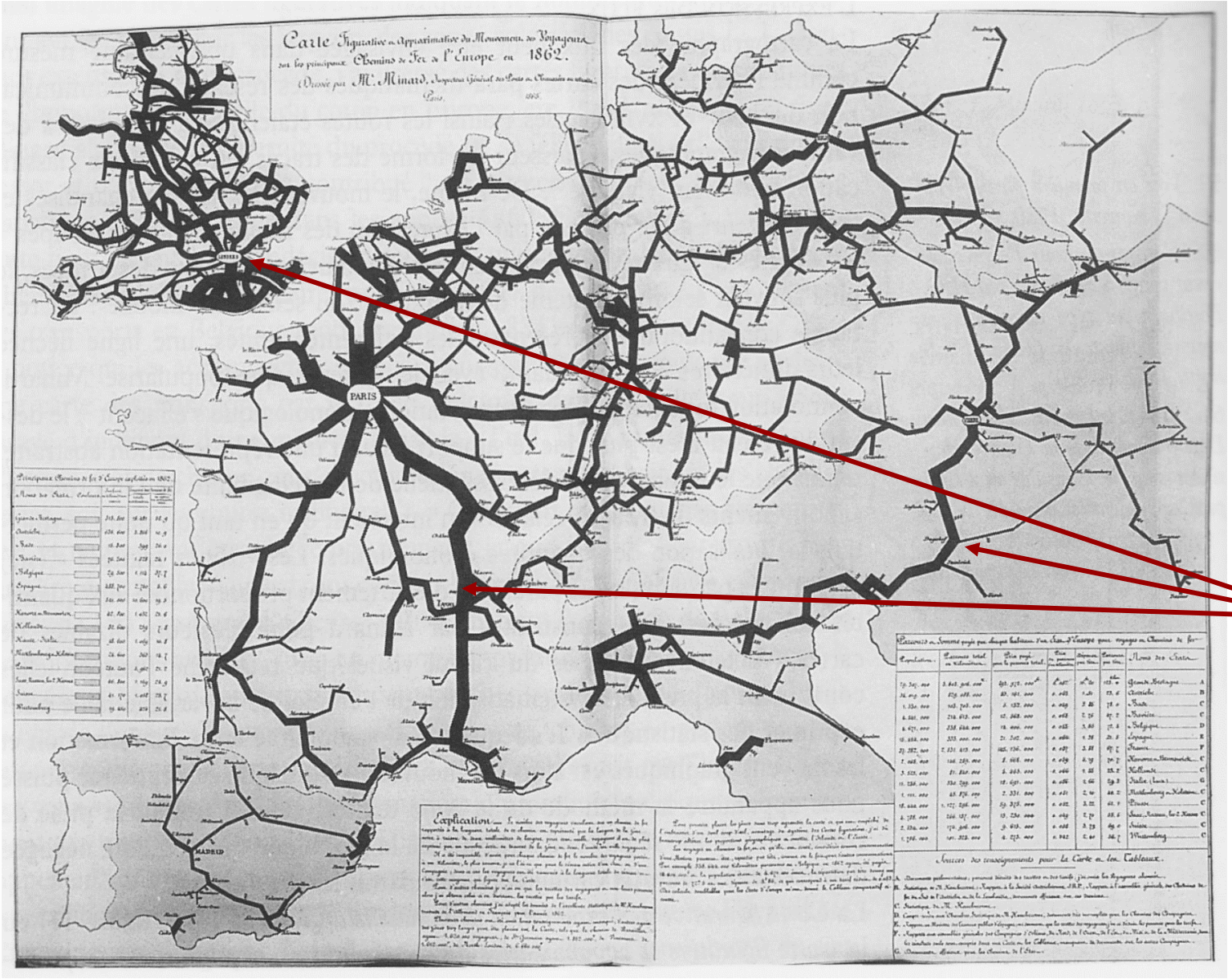
Verlag von B.G. Teubner, Leipzig, 1925.
côte à côte plus qu'intégration

>cartographie
+ statistique



http://www.InfoVis.info/visuals/Individual/MappingGlobalization/Teubners_Weltwirtschaftskarten_III_2.jpg

>cartographie + statistique

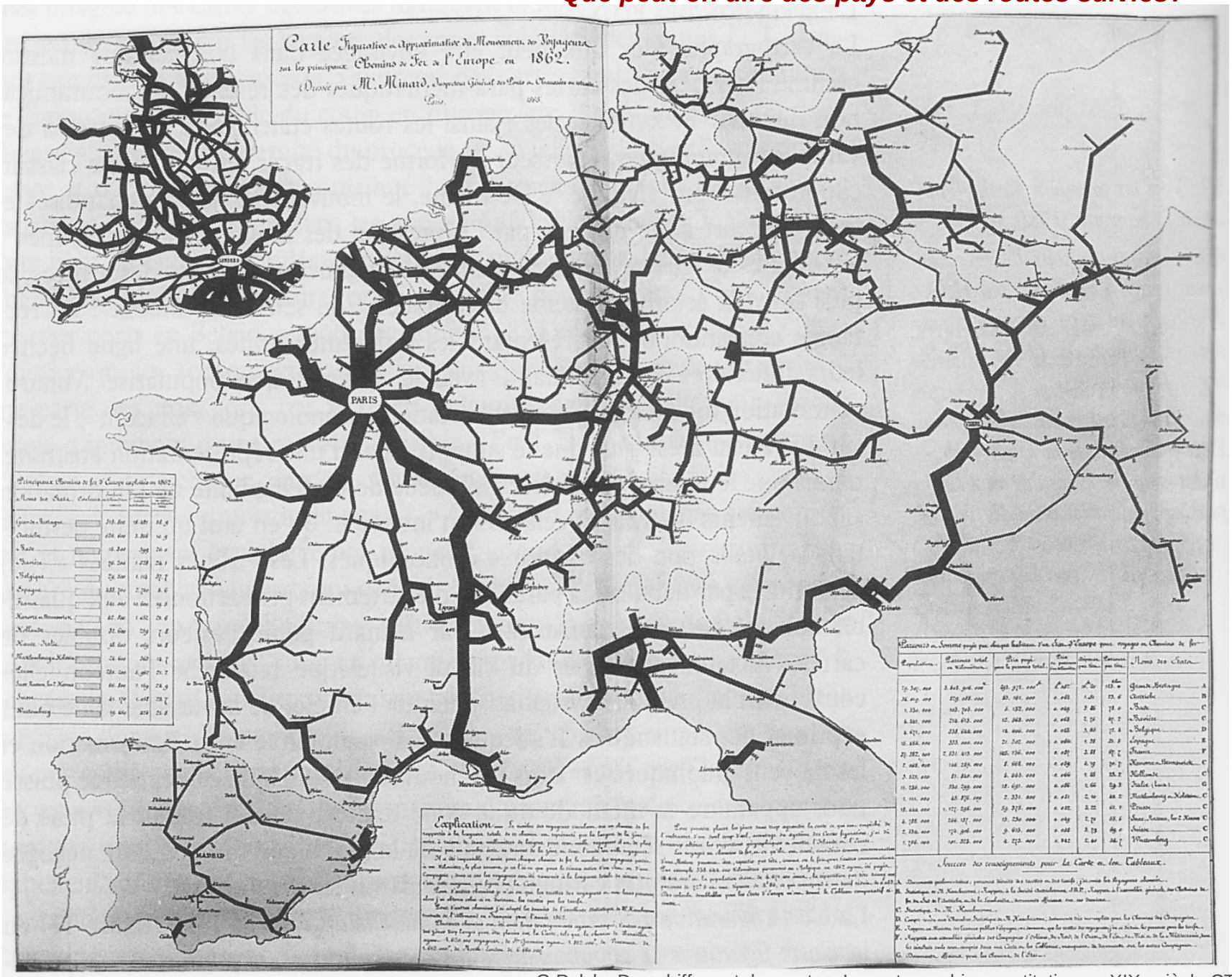


Carte figurative
C.J Minard, 1865

Mouvements de voyageurs en 1862 : largeur trait proportionnel au nombre de voyageurs

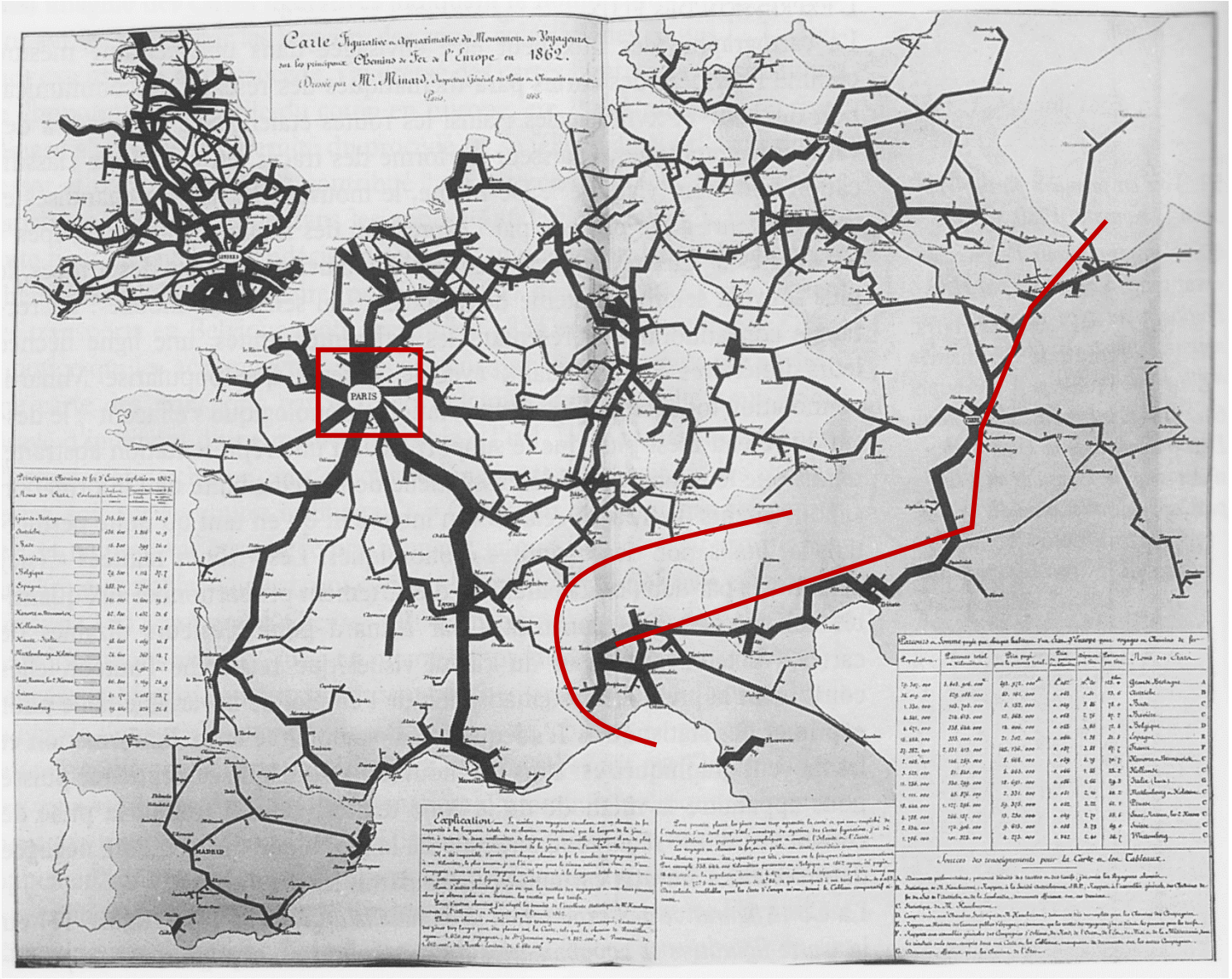
G.Palsky Des chiffres et des cartes. La cartographie quantitative au XIXe siècle CTHS 1996

Que peut-on dire des pays et des routes suivies?



G.Palsky Des chiffres et des cartes. La cartographie quantitative au XIXe siècle CTHS 1996

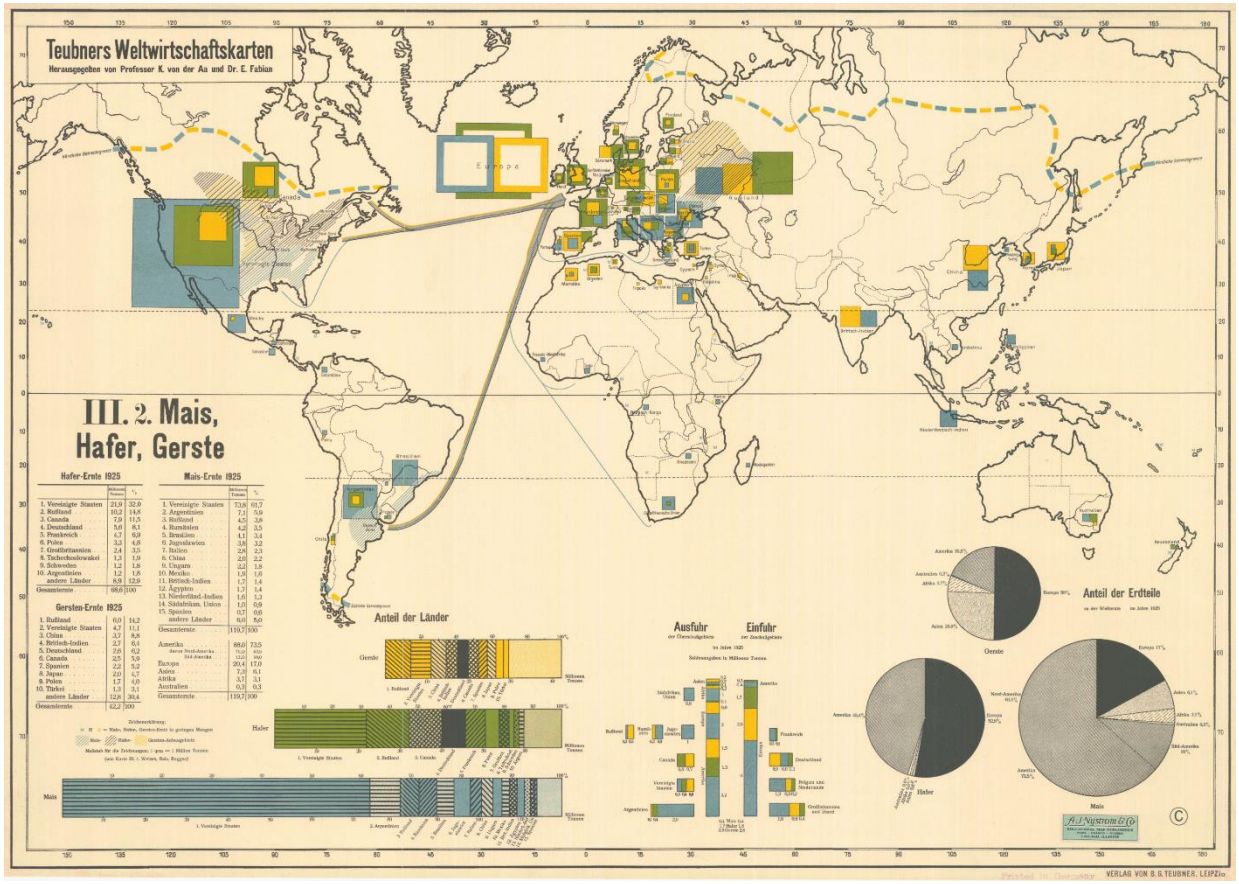
>cartographie
+ statistique



Carte figurative
C.J Minard, 1865

centralisation,
géographie physique,
géographie politique

par le recours à un dispositif visuel, faire comprendre au-delà des données elles-mêmes



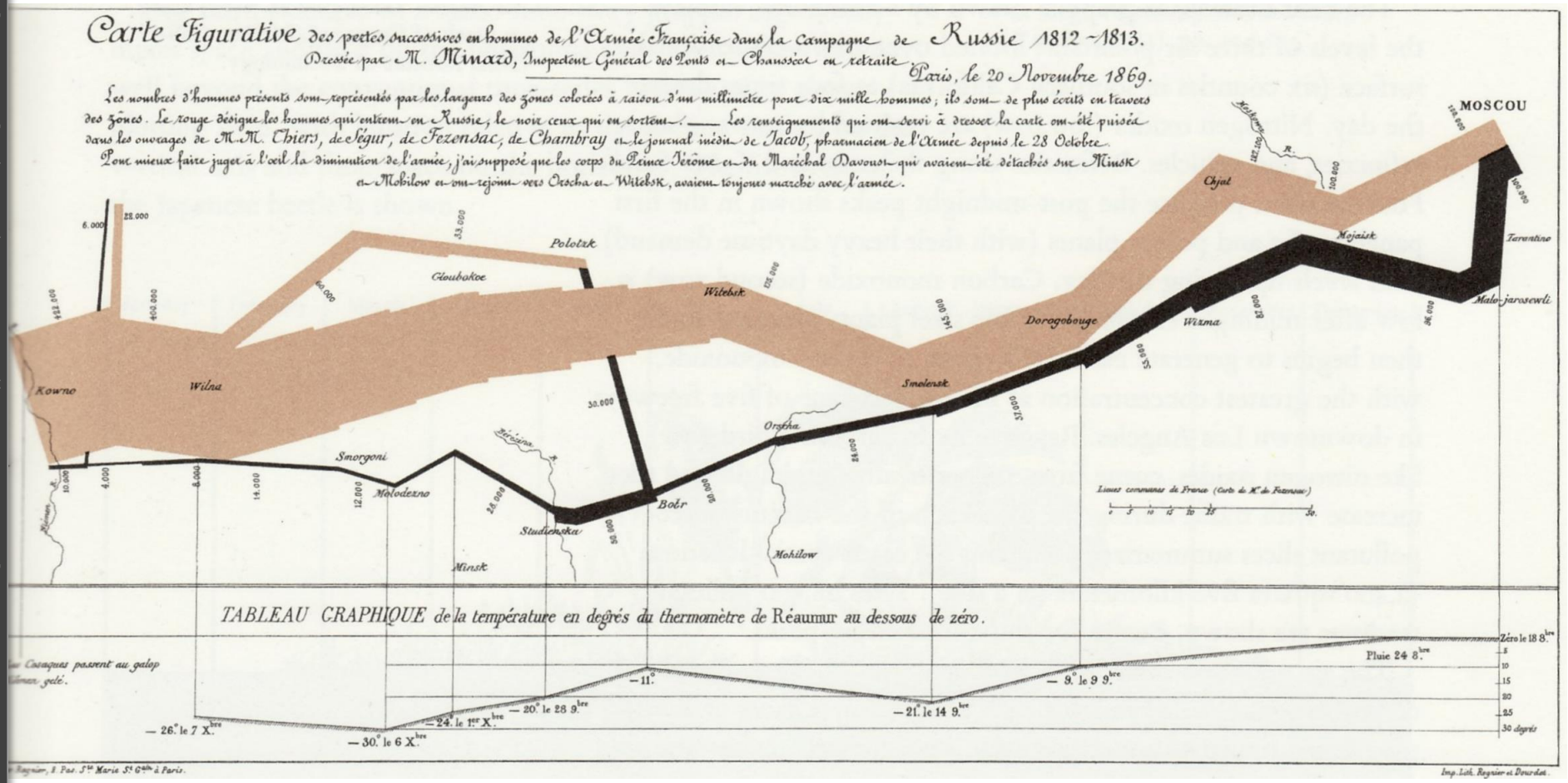
Verlag von B.G. Teubner, Leipzig, 1925.

Cartographie + statistique
> localisations, quantités, mouvement

Et la dimension temporelle?

Et en 1924? Et à quelle vitesse va-t-on de *a* à *b*, en quelle saison, etc?

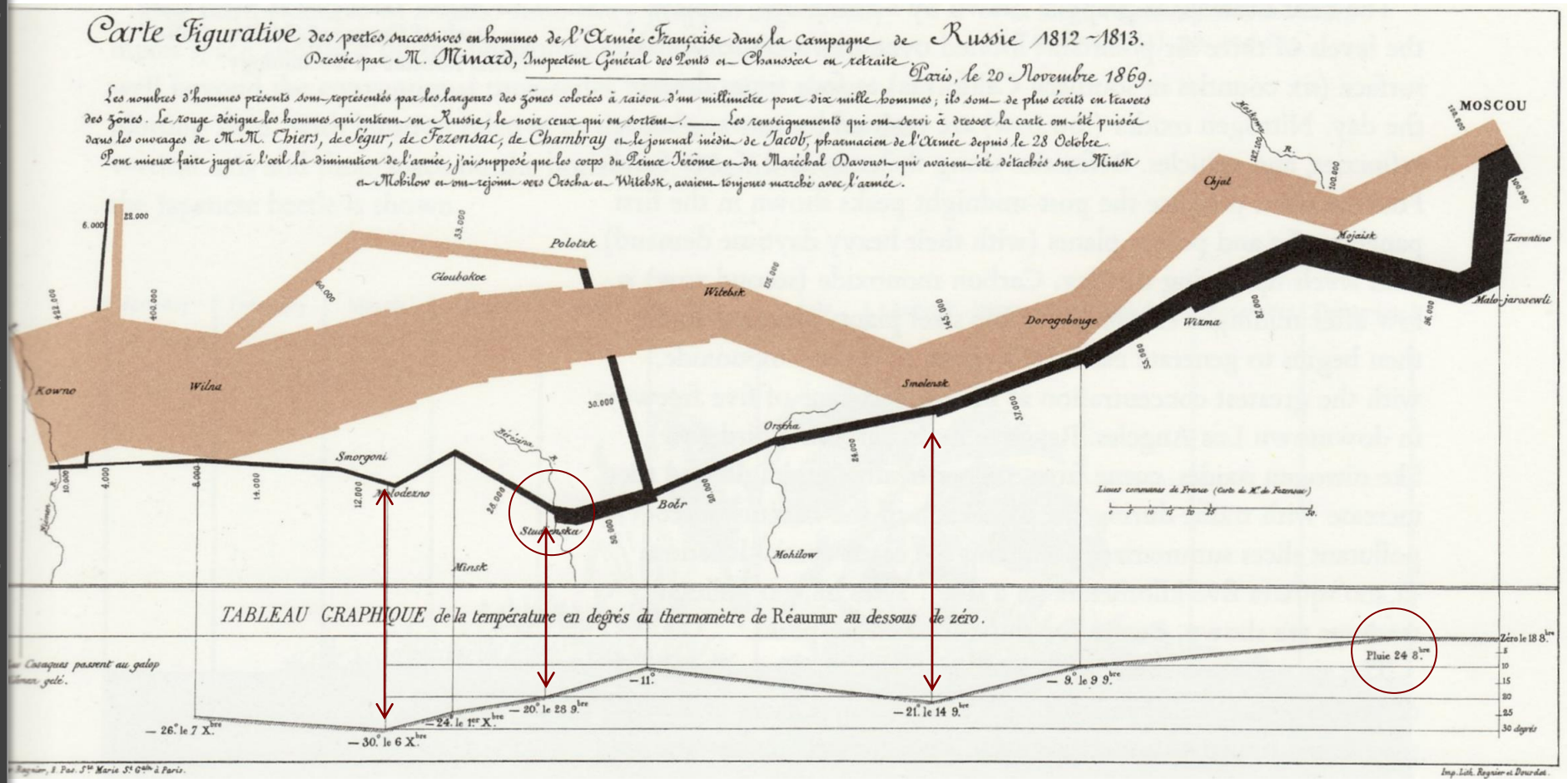
Carte figurative + tableau graphique
C.J. Minard, 1869



« nulle part la marche des armées n'atteint ce degré de brutale éloquence qui semble défier la plume de historien»

E.J Marey La méthode graphique dans les sciences expérimentales, 1878

E.R Tufte, Beautiful Evidence Graphics Press Cheshire 2006, p126

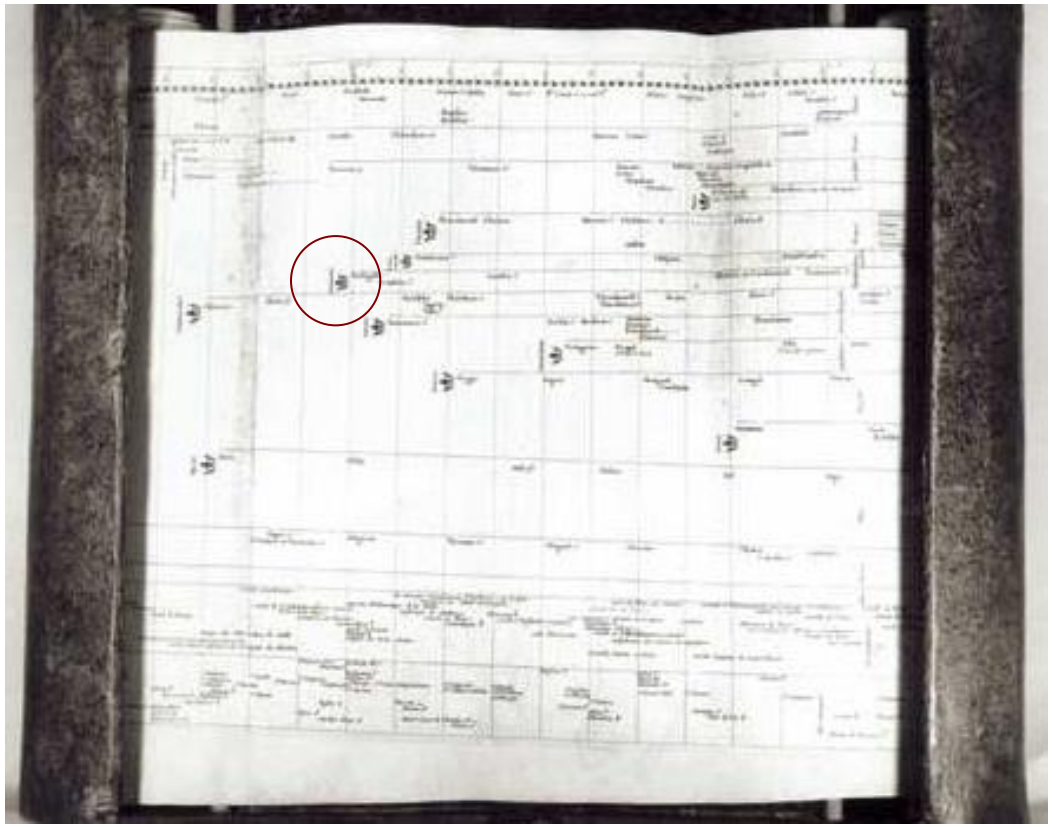


Overview first, zoom and filter, then details on demand

(Ben Shneiderman)

Tout est là pour y arriver

>Le paramètre temps: retour en arrière



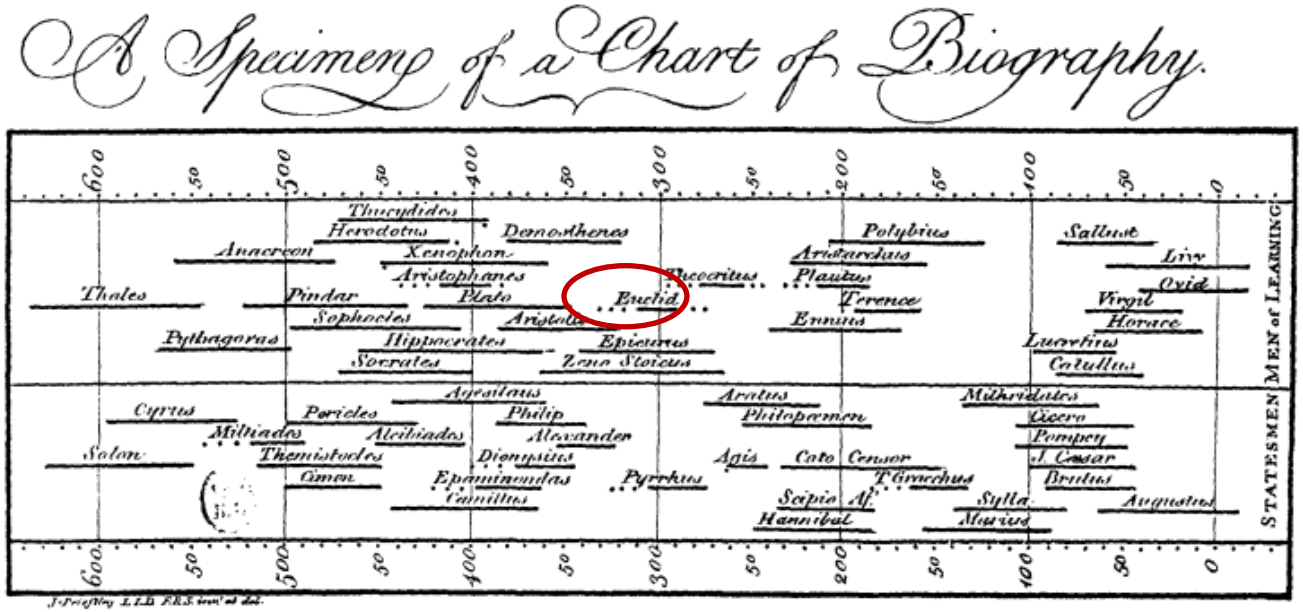
Carte chronologique
Jacques Barbeu-Dubourg
1753

Une ligne de temps annotée comprenant des noms, etdes évènements, groupés par thème, avec des symboles qualitatifs indiquant le « profil » (martyre, tyran, hérétique, noble, etc.) et/ou la profession (peintre, théologien, musicien, moine, etc.)-

*http://www.mediamatic.net/142472/en/carte-chronologique-by-jacques-barbeu-dubourg-1753

>Le paramètre temps: retour en arrière

Modèle dominant: le temps comme linéaire, continu, lecture de gauche à droite



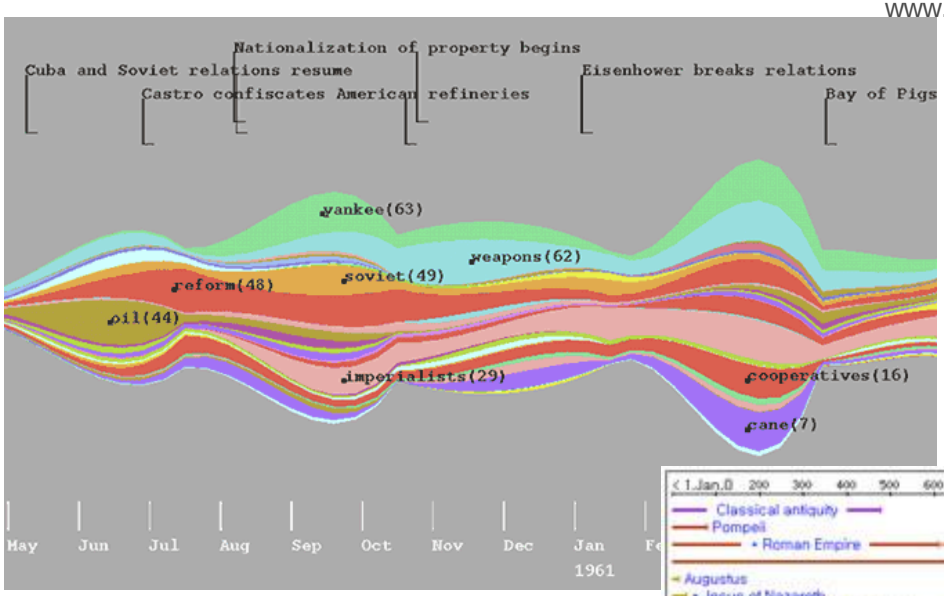
Historical timeline*
Joseph Priestley, 1765

Vie de 2,000 personnage connus, de -1200 à 1750, comparaisons quantitatives par la longueur des barres

Déjà , des doutes visualisés

*http://euclid.psych.yorku.ca/SCS/Gallery/images/priestley.gif

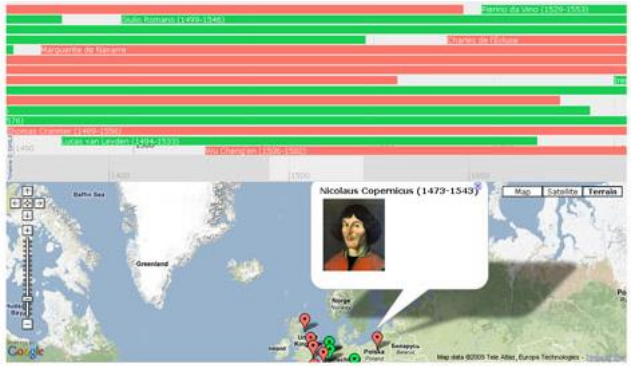
www.ifs.tuwien.ac.at/~silvia/wien/vu-InfoVis/references/havre-ieeeInfoVis00.pdff

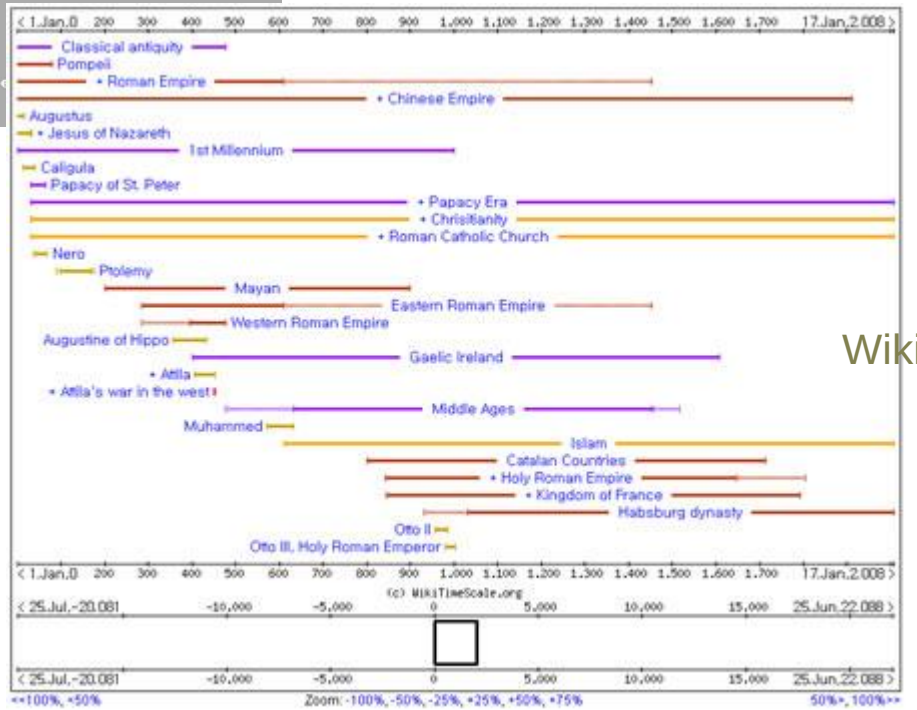


themeRiver

Modèle (toujours) dominant - Plus d'inventivité graphique, mais même dimension du paramètre temps exploitée

Google TimeMap





WikiTimeScale

Simile Timeline

Simile

http://www.computus.org/journal/?p=8

http://www.computus.org/journal/?p=7

http://www.computus.org/journal/?p=1206

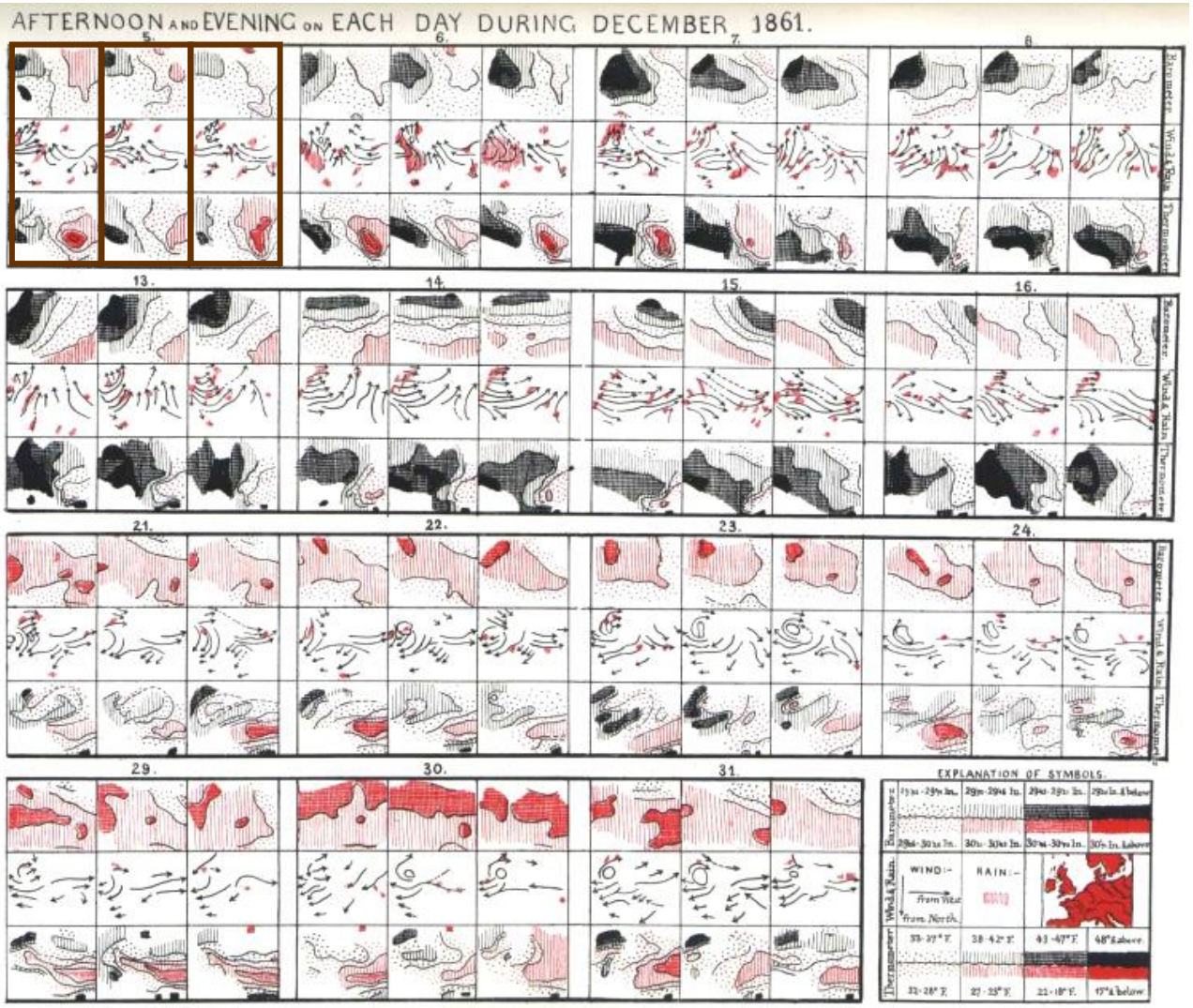
Le paramètre temps

Alternatives (dès l'âge d'or)

Matin

Après-midi

Soir



A series of weather maps from the *Meteorographica*.

Les weather charts de Francis Galton

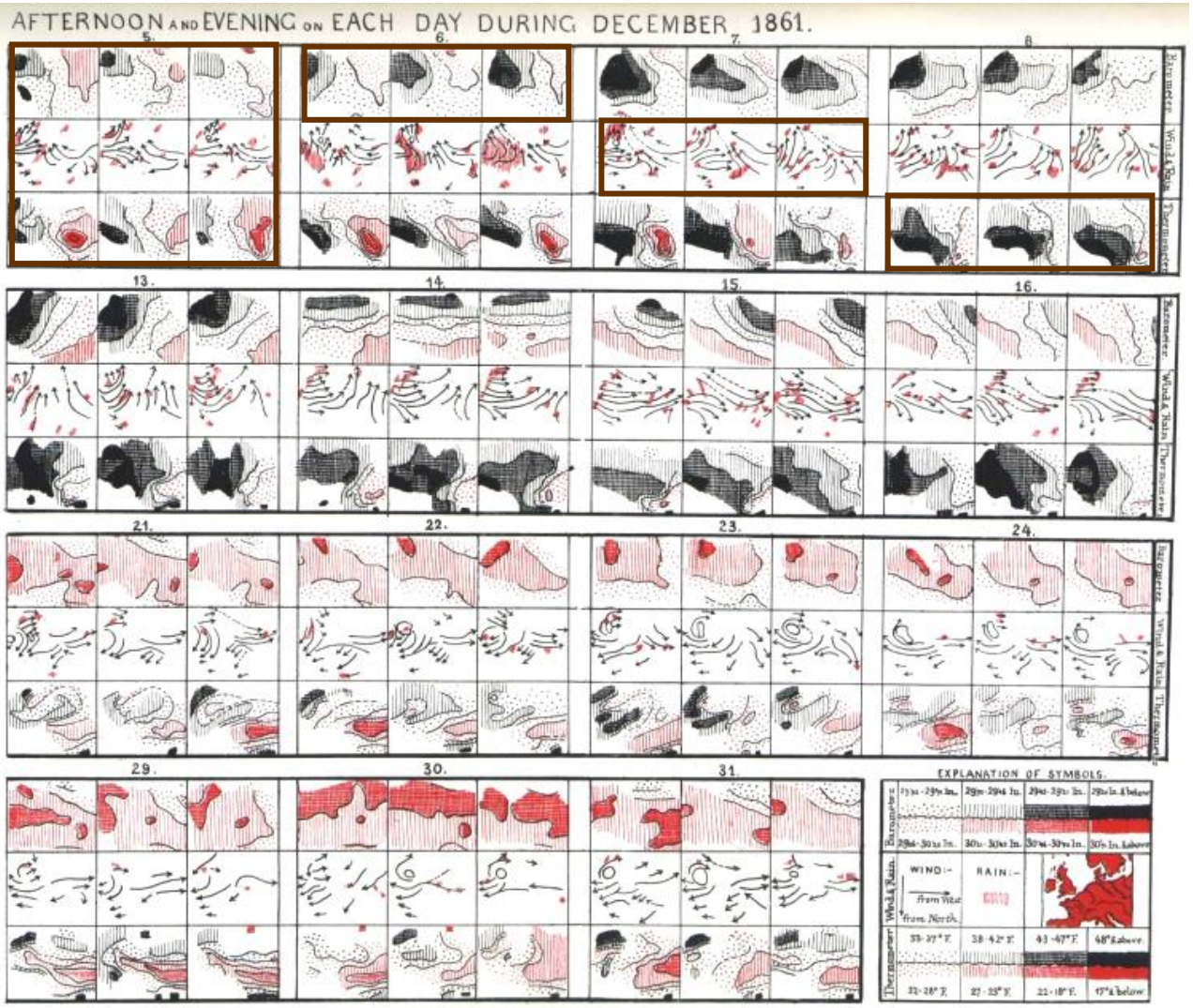
Observation systématique de 3 paramètres sur un mois: pression (atm) direction des vents et température - Décembre 1861.

Francis Galton, pub. 1863.
in M.Friendly, A Brief History of Data Visualization
Multivariate weather charts, arranged as small multiples

Le paramètre temps

Alternatives (dès l'âge d'or)

Jour



A series of weather maps from the *Meteorographica*.

pression barométrique

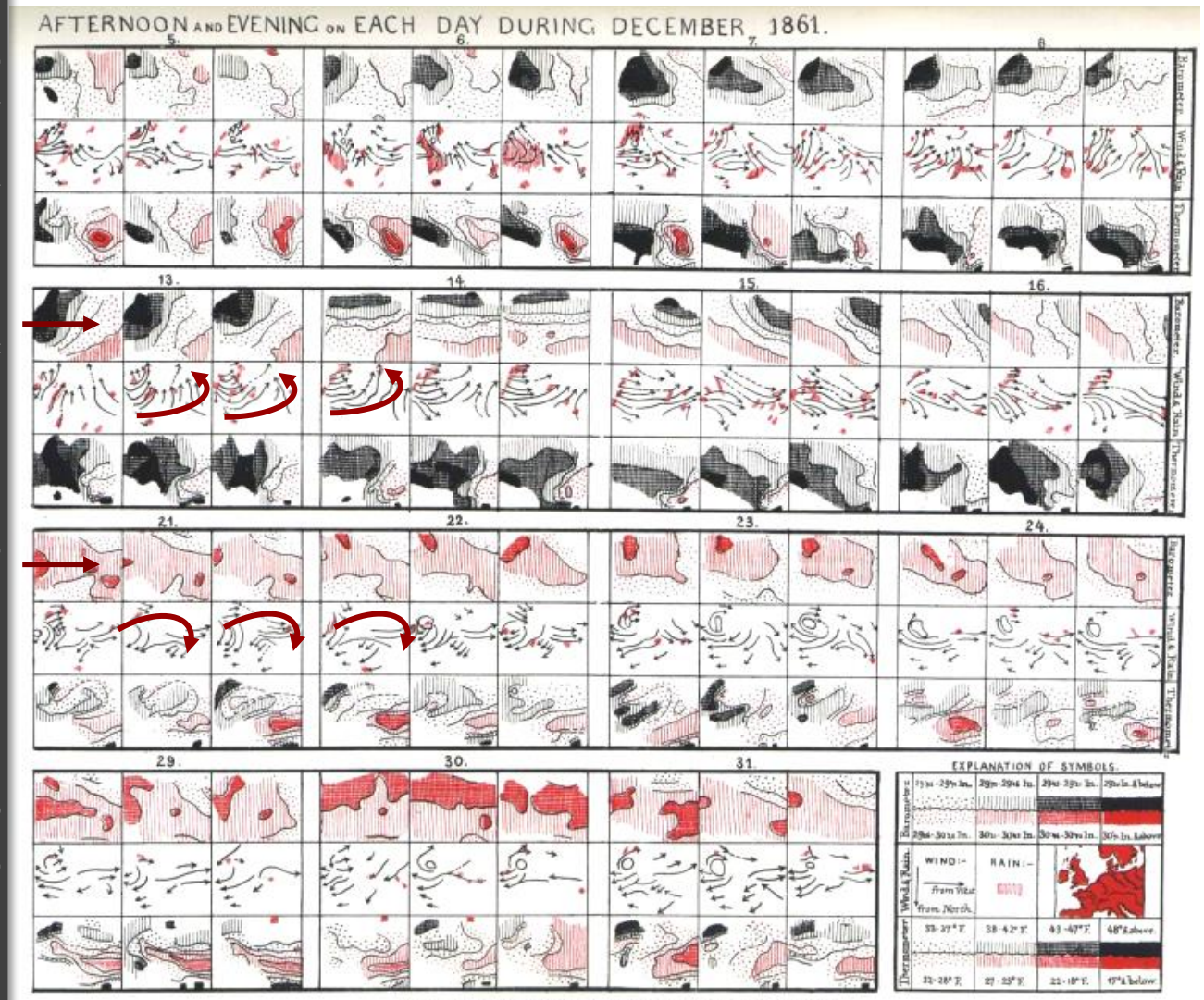
vent et pluie

température

Francis Galton, pub. 1863.
in M.Friendly, A Brief History of Data Visualization
Multivariate weather charts, arranged as small multiples

Le paramètre temps

Alternatives (dès l'âge d'or)



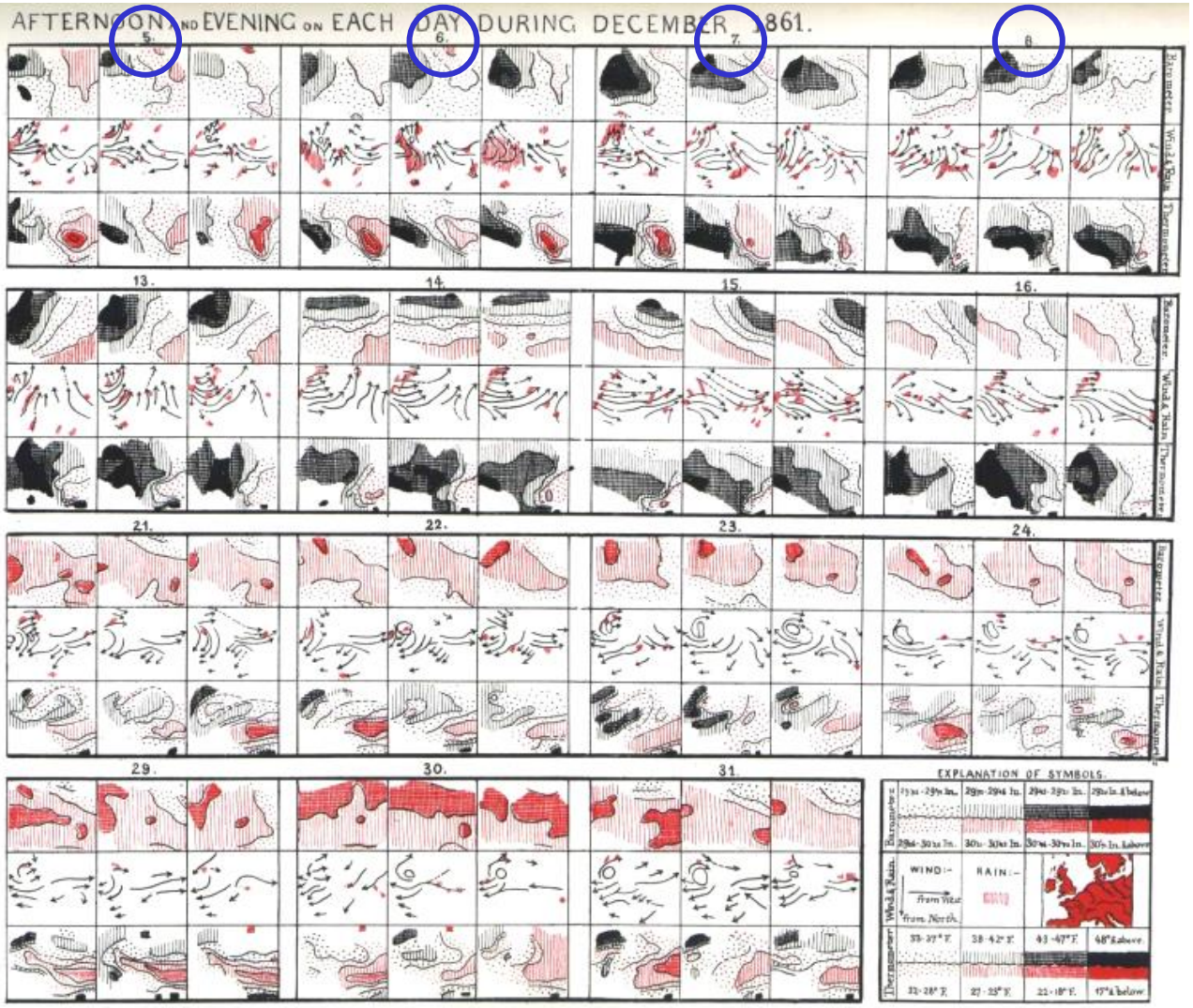
A series of weather maps from the *Meteorographica*.

La démonstration graphique : motif "anti-cyclonique" (CCW - sens contraire aiguilles d'une montre) des vents dans les régions de basse pression, rotations CW des vents dans les régions de haute pression.

Francis Galton, pub. 1863.
in M.Friendly, A Brief History of Data Visualization
Multivariate weather charts, arranged as small multiples

Le paramètre temps

Alternatives (dès l'âge d'or)



A series of weather maps from the *Meteorographica*.

<u>temps ordonné</u>, discret (chaque jour est précédé / suivi d'un autre jour, succession d'unités temporelles sur un pas arbitraire)

<u>temps orienté-points</u> (points temporels en principe « sans épaisseur » - pas d'informations sur ce qui se situe entre deux points).

granularité
(mois ; jour ; jour /3)

Francis Galton, pub. 1863.
in M.Friendly, A Brief History of Data Visualization
Multivariate weather charts, arranged as small multiples

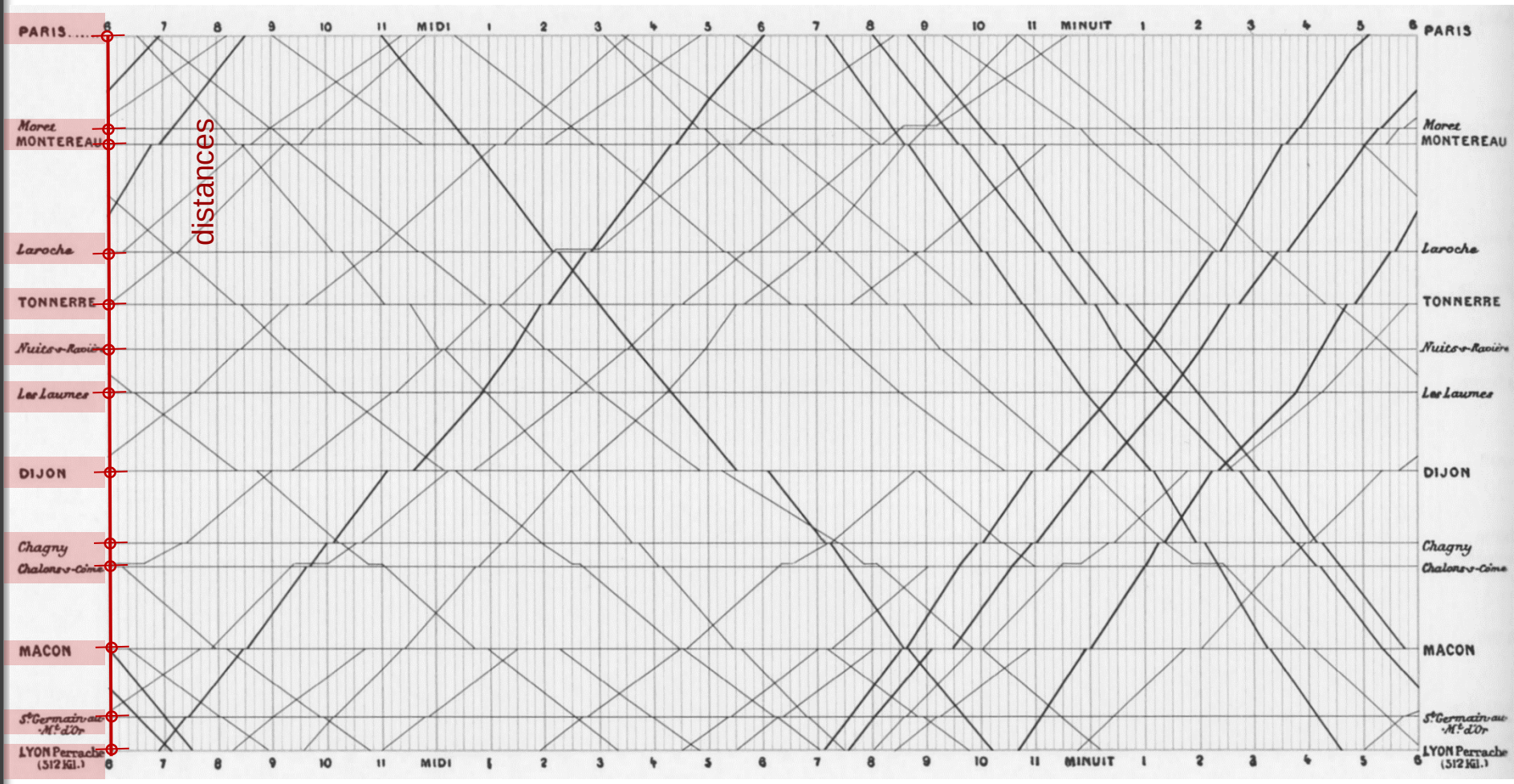
Le paramètre temps

temps continu, orienté intervalles

Horaire D' E.J Marey (1885)

24 heures

gares



Train schedule

E.J. Marey 1885

E.R Tufte The visual display of quantitative information ,

Graphic Press, Cheshire 2001

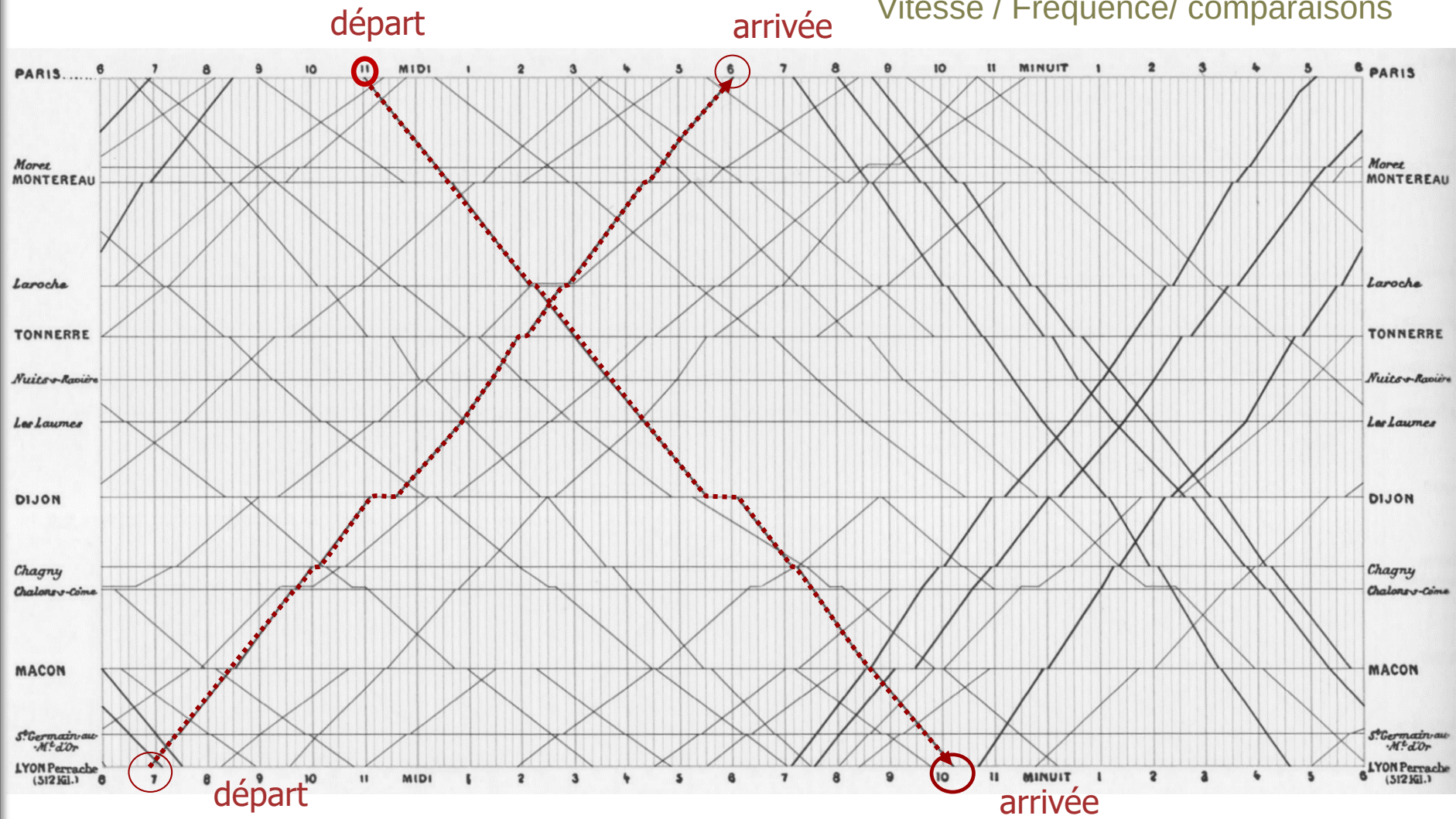
J.Y Blaise I.Dudek CNRS – [MASTER SIS 2016-2017 "Ingénierie et Multimédia" 2025] InfoVis: fondamentaux et applications aux données spatio-historiques [J.Y Blaise I.Dudek]

Le paramètre temps

temps continu, orienté intervalles

Horaire D' E.J Marey (1885)

Vitesse / Fréquence/ comparaisons



Train schedule

E.J. Marey 1885

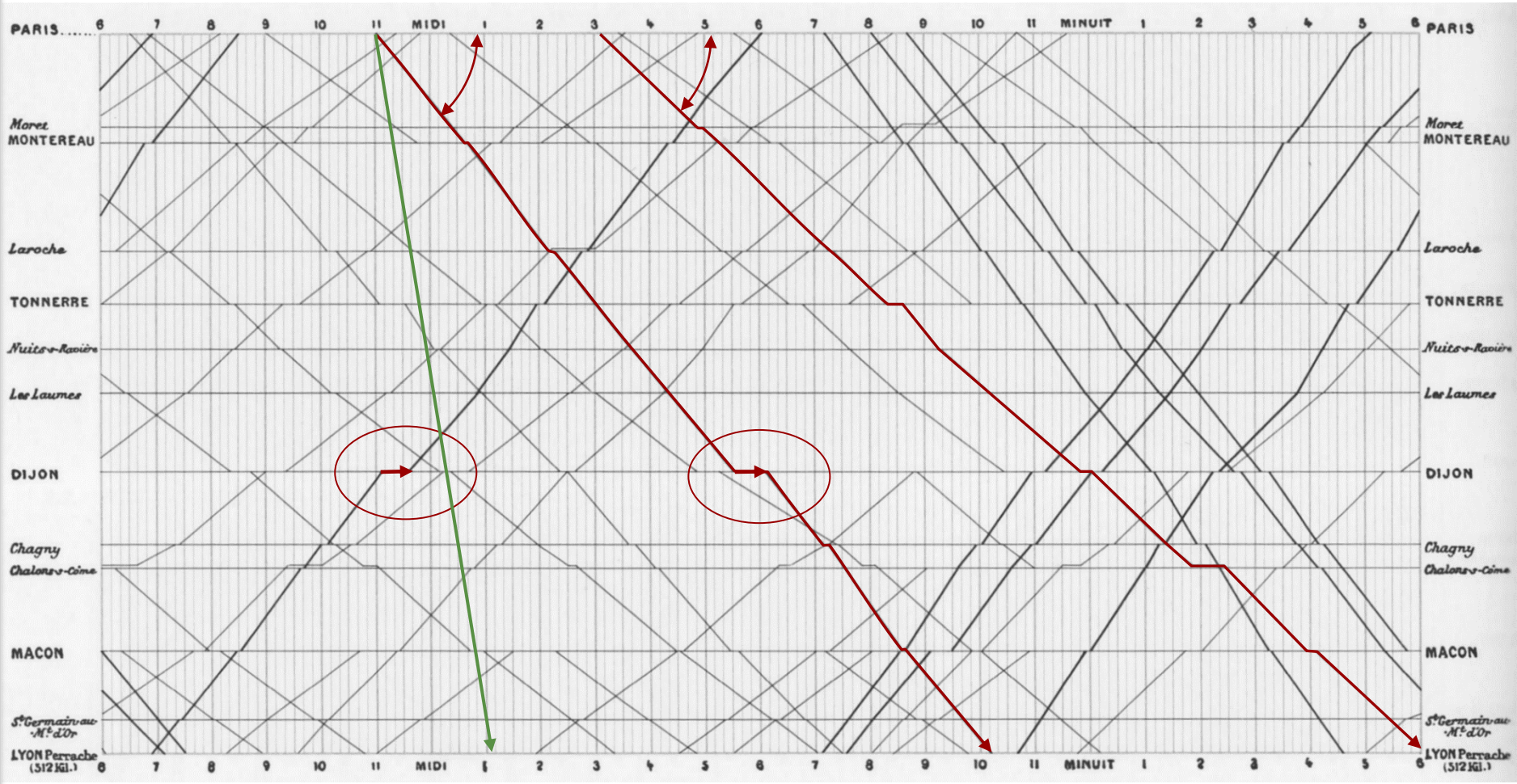
E.R Tufte The visual display of quantitative information ,

Graphic Press, Cheshire 2001

Le paramètre temps

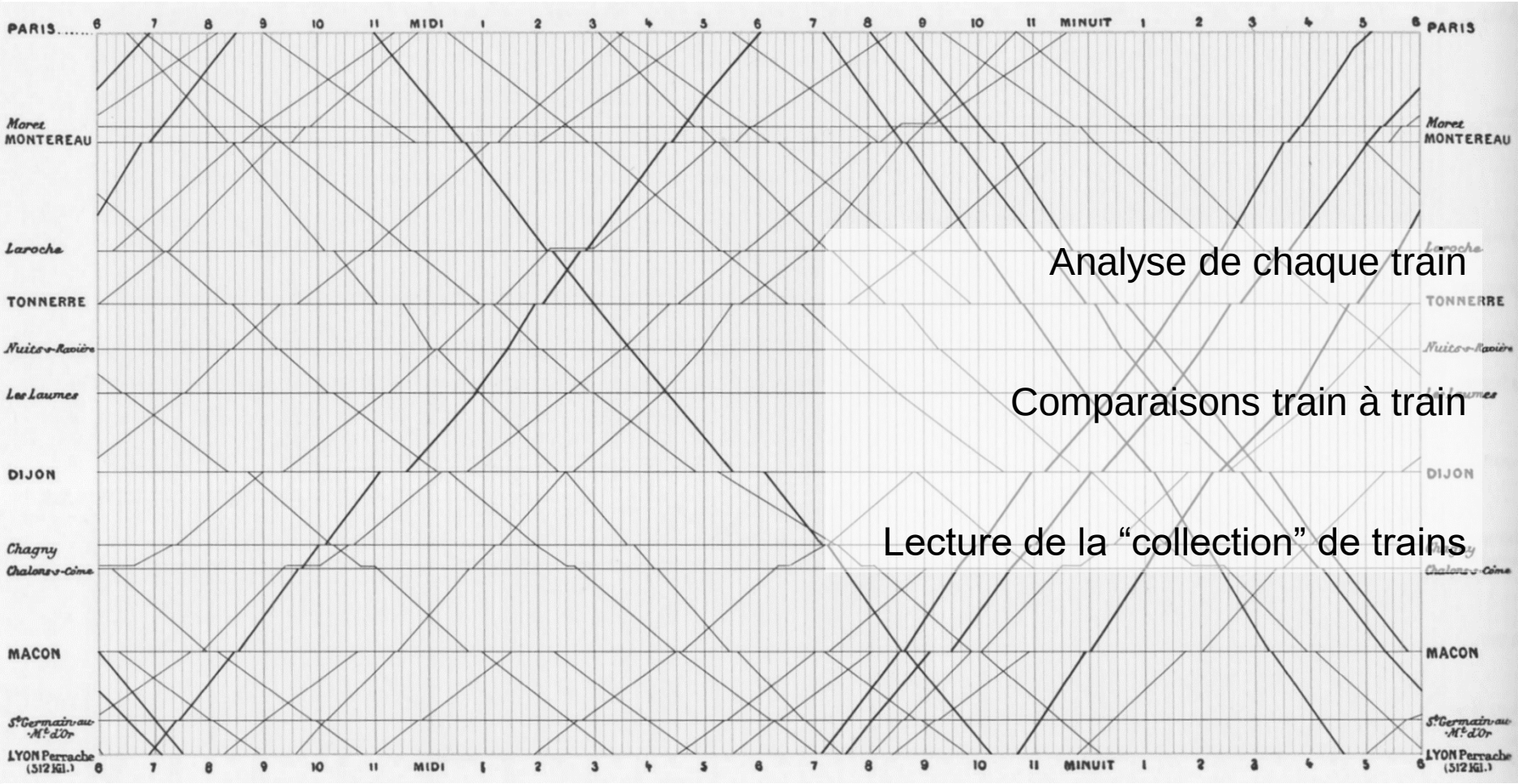
tgv

angle= vitesse



temps escale

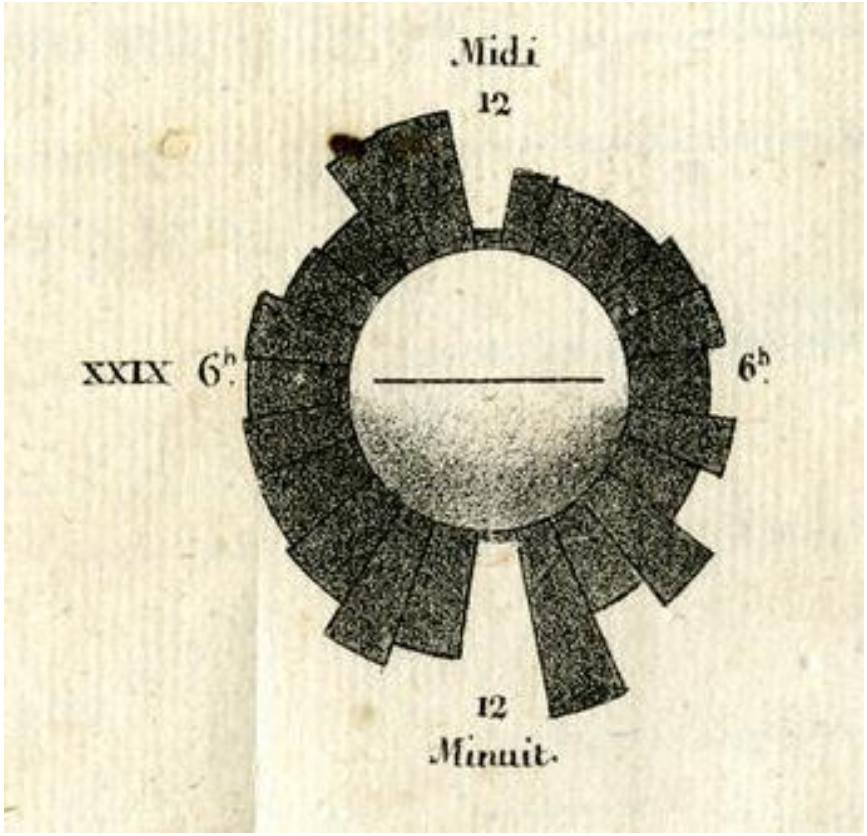
Train schedule
E.J. Marey 1885
E.R Tufte The visual display of quantitative information ,
Graphic Press, Cheshire 2001



Règle de Tufte > ***1+1 = 3***
horaires + distances =
horaires + distances + comparaison des vitesses

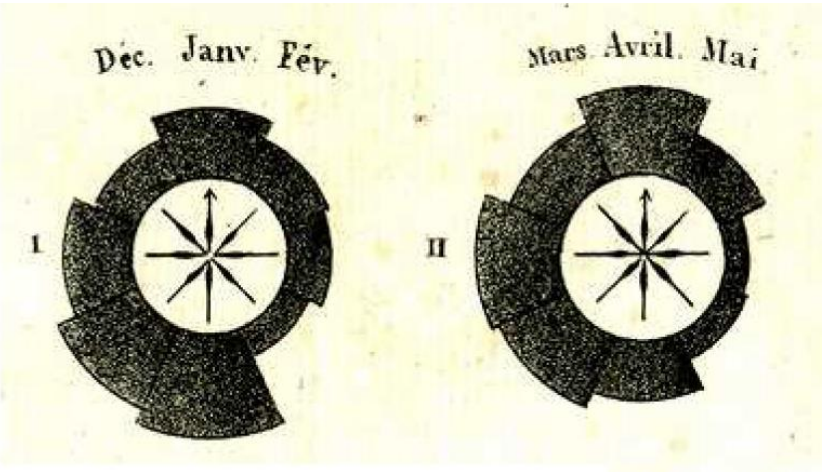
Train schedule
E.J. Marey 1885
E.R Tufte The visual display of quantitative information ,
Graphic Press, Cheshire 2001

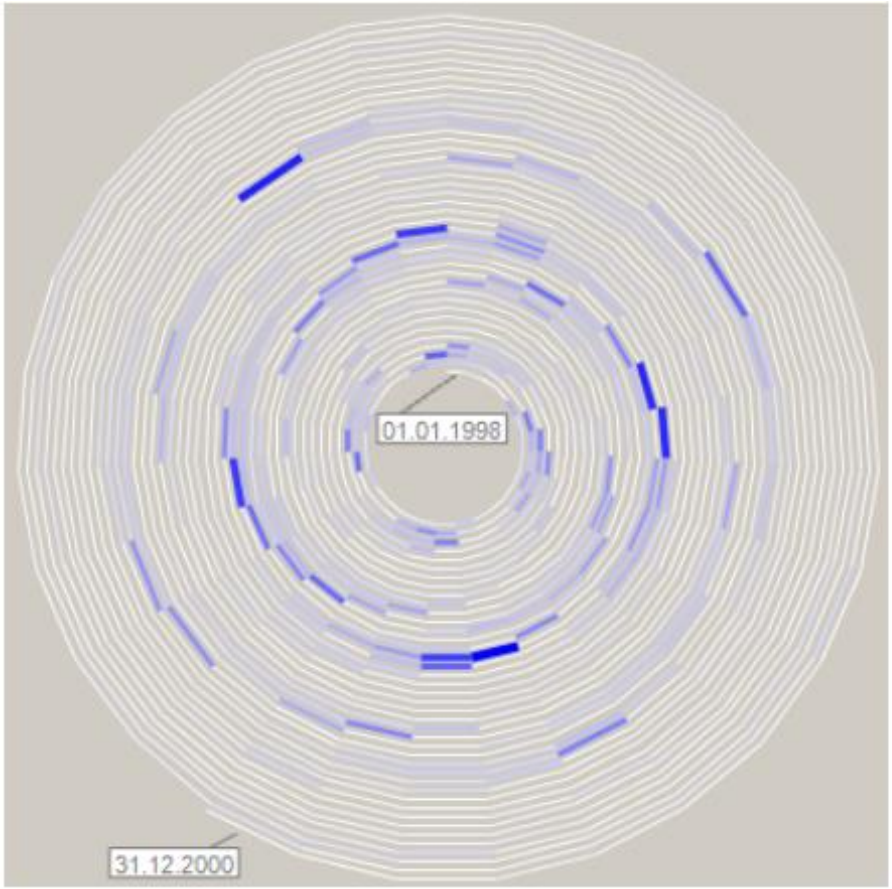
Visualisation de cycles



polar-area charts
André Michel Guerry 1829

récurrence de phénomènes cycliques (direction des vents dominants par exemple)





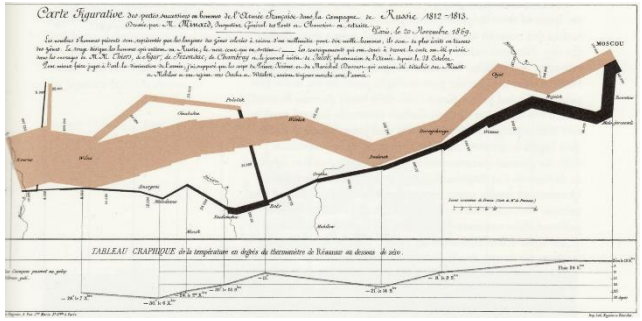
3 ans de cas de grippe,
un cycle = 27 jours

Spiral graph
http://www.informatik.uni-rostock.de/~ct/pub_files/Aigner08TimeVis.pdf
Visual Methods for Analyzing Time-Oriented Data
W.Aigner, S.Miksch, W.Müller, H.Schumann, and C.Tominski
IEEE Transactions on Visualization and Computer Graphics, Vol. 14, No. 1, 2008.

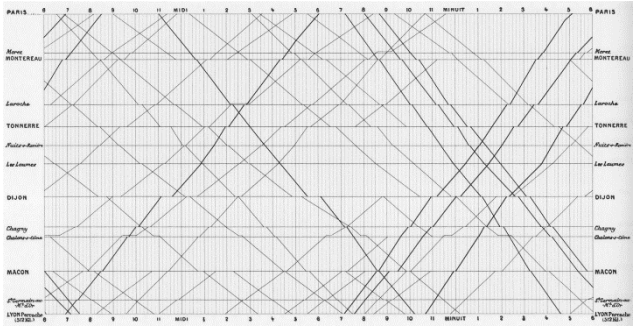
[extracts of original legend] Visual representations of a time-oriented dataset describing the number of influenza cases over a period of three years – left: SpiralGraph encoding 27 days per cycle (improperly parameterized – periodic pattern is hard to see
right: SpiralGraph encoding 28 days per cycle (properly parameterized – periodic pattern stands out).

Que retenir de cet héritage?

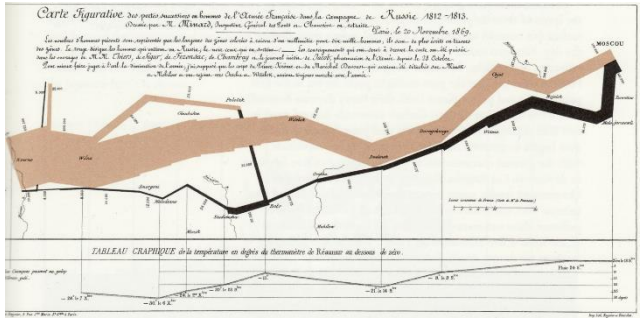
D’abord une approche :
apprendre à re-considérer, à ré-analyser un jeu de données, y rechercher motifs et exceptions.

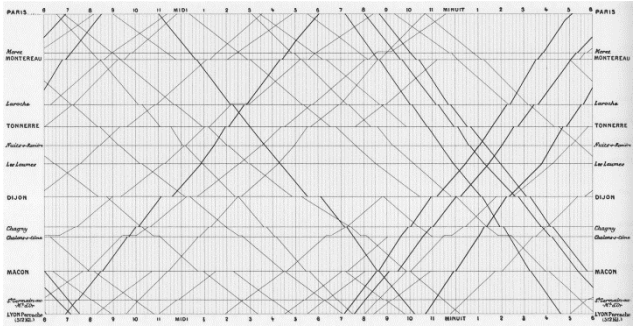


Des outils visuels variés – mais un fond de méthodes et de pratiques cadrés: *put the data first*.



Des solutions pour prendre en compte *différentes* dimensions du paramètre temps.





Que retenir de cet héritage?

Une « bonne visualisation » n'est pas affaire de technologie, mais de créativité et de rigueur.

Une « bonne visualisation » est un outil de découverte, si nous « n'apprenons » rien que nous n'ayons déjà su alors il fallait faire autre chose.

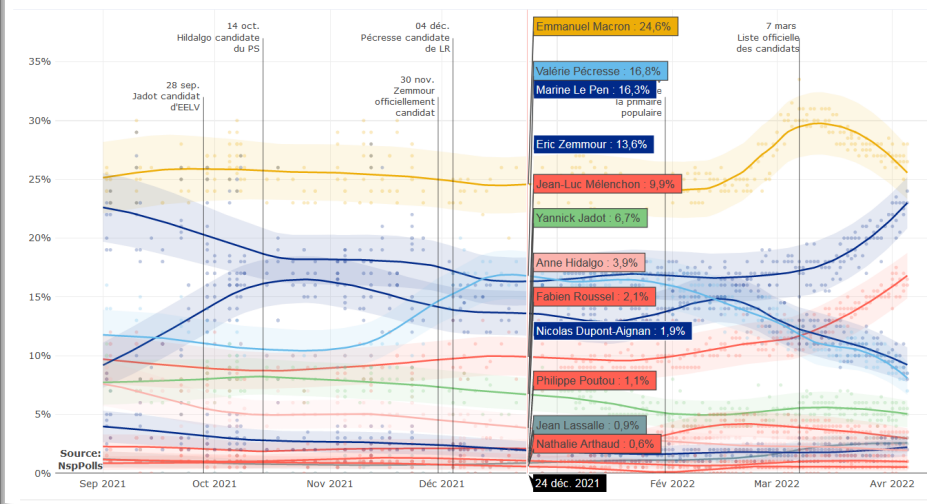
Ce que la « technologie » apporte :

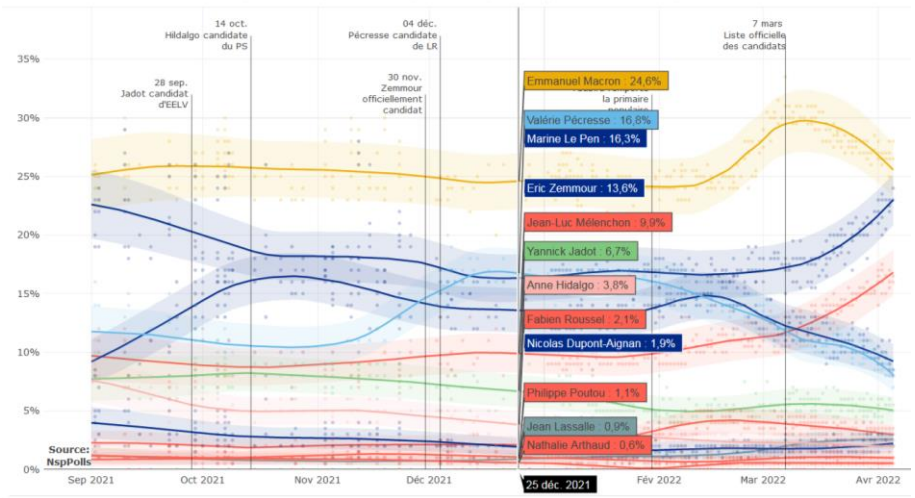
- pré-traitement de données massives
- interaction utilisateur
- Mise à jour / production dynamique

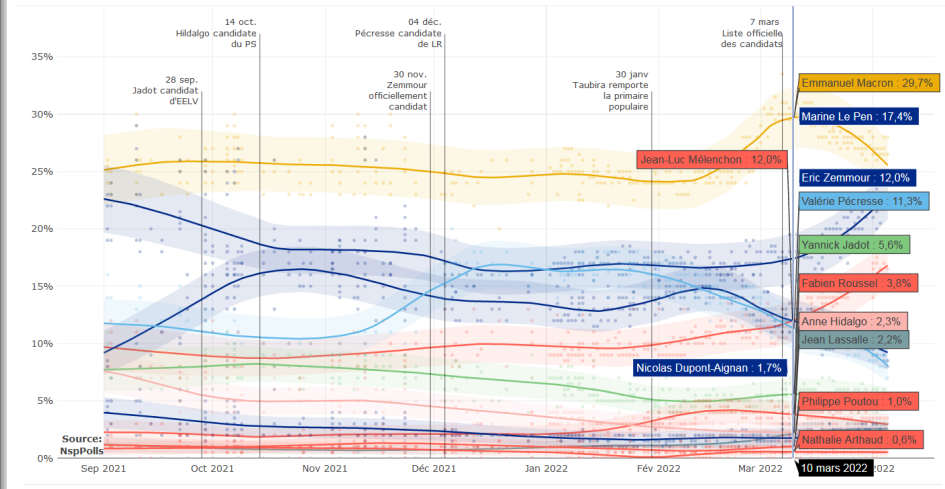
Ce qu'elle n'apporte pas : esprit critique

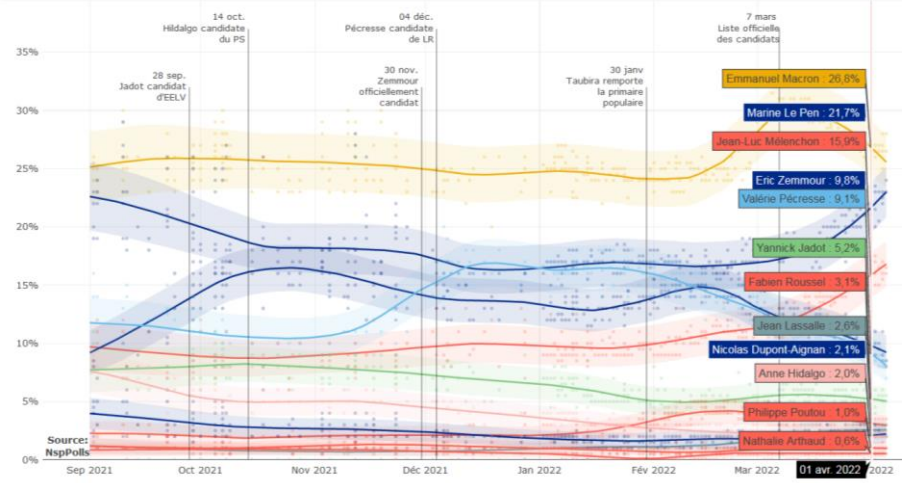
Ce qu'elle coûte : difficulté à trouver une juste balance entre l'automatisation et la qualité

Plus interactif, mise à jour automatisable, mais « mieux? »





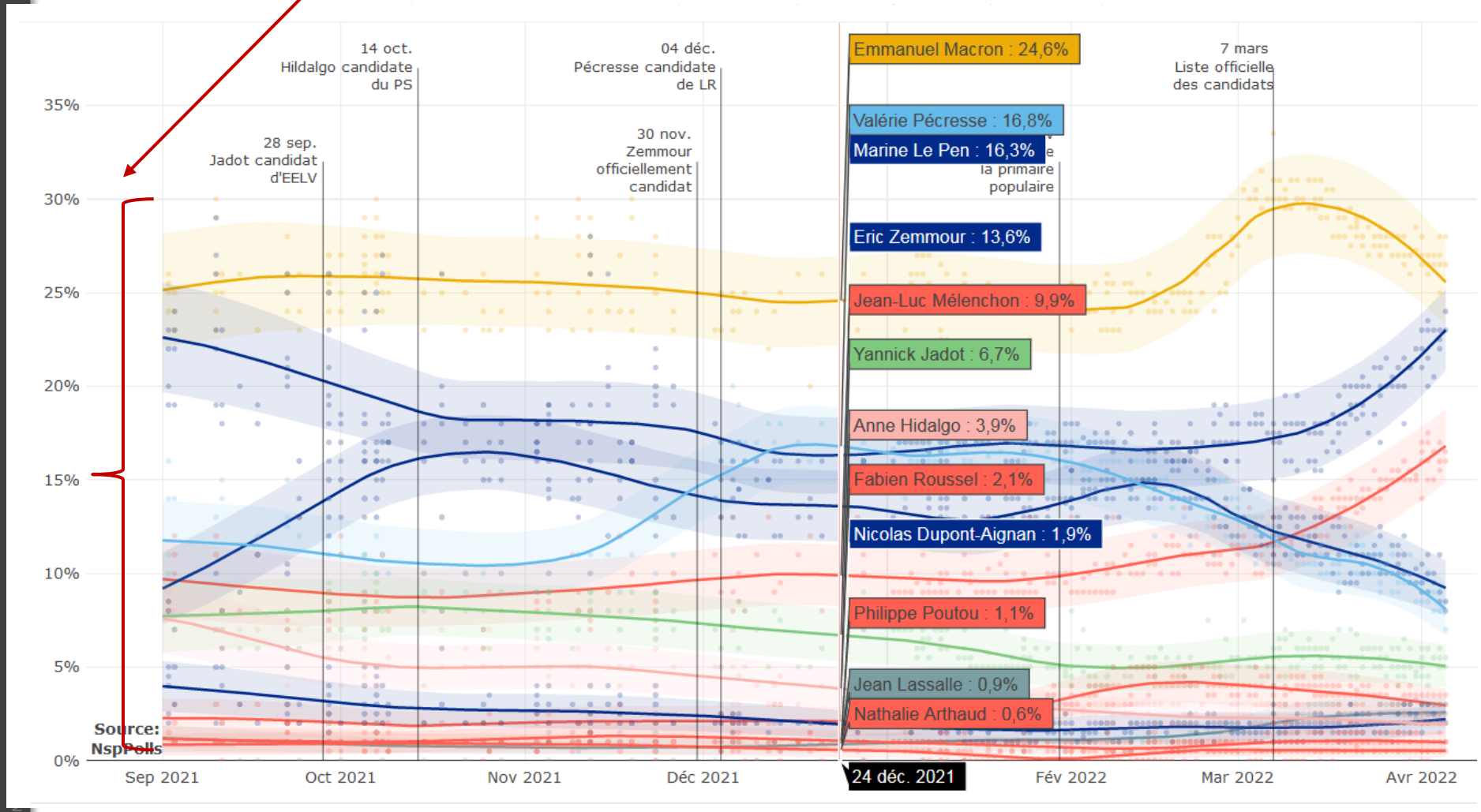




Journal « Le Télégramme », visualisation en ligne pendant les élections présidentielles de 2022
Curseur temporel interactif (ici 4 jours montrés)

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

Une ligne par candidat ou candidate

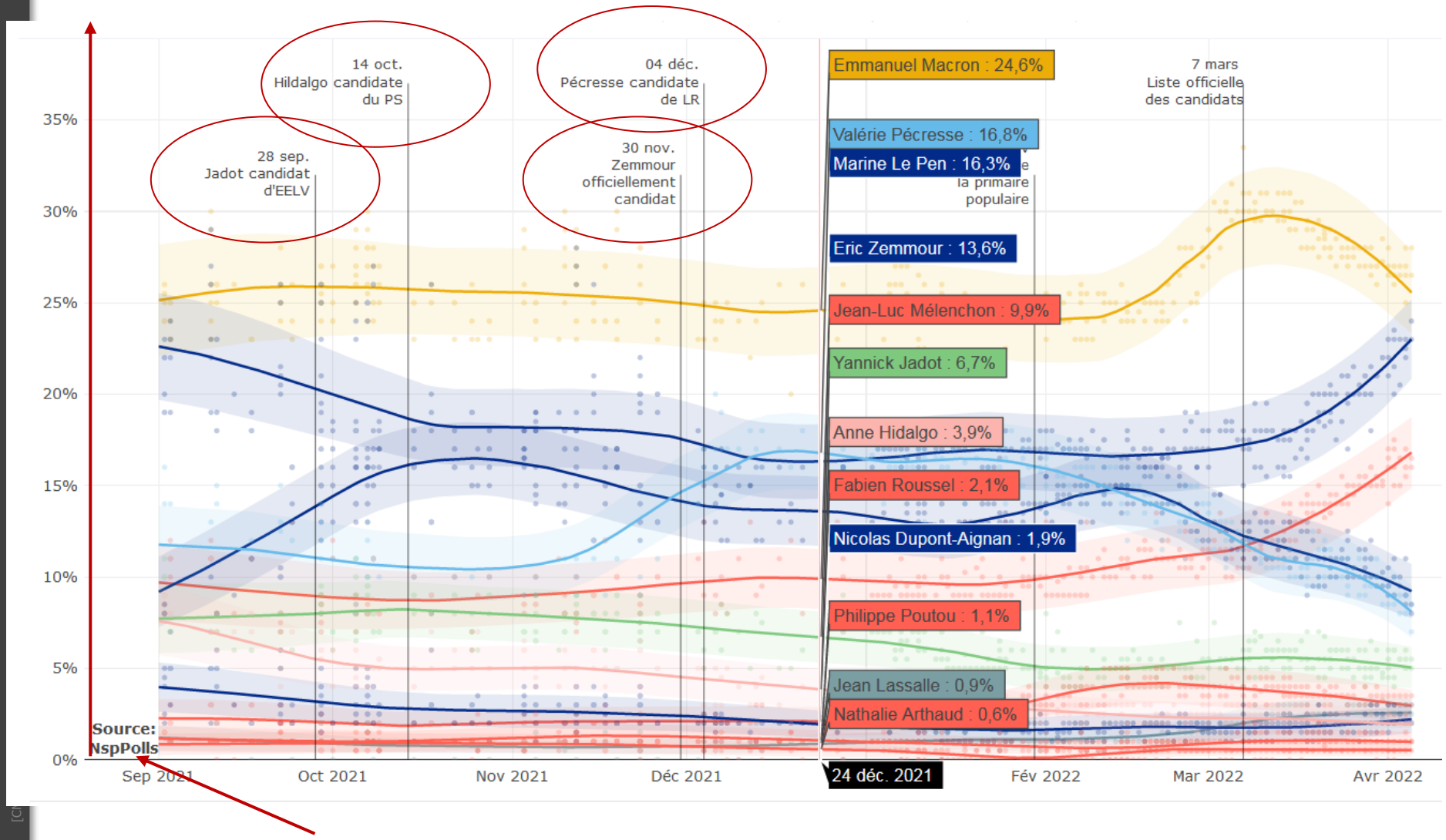


Temps (temps discret, jour par jour), curseur interactif permettant de lire les valeurs d'un jour

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

Pourcentage - mais de quoi? Intentions de vote ?

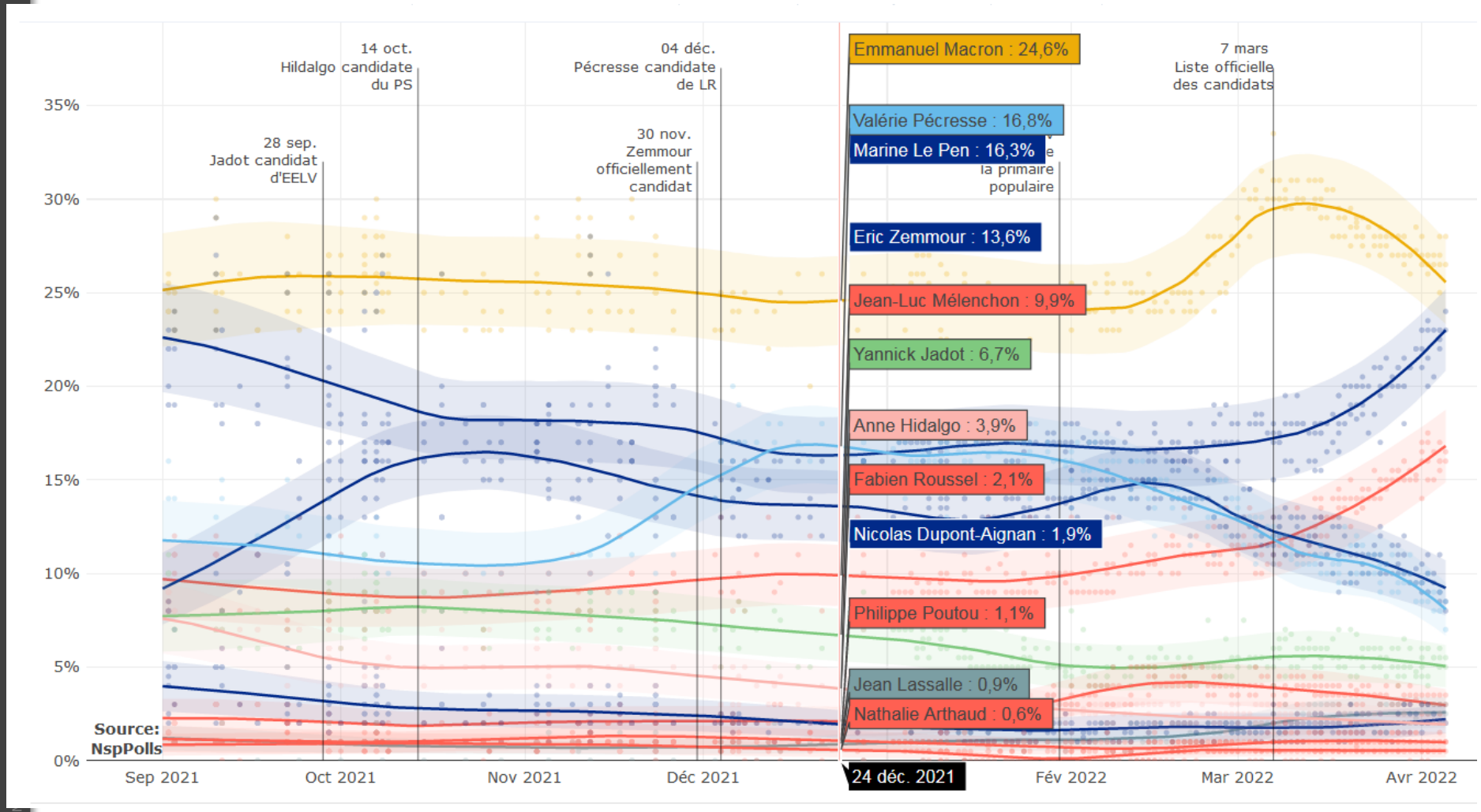
Alors que ce ne sont pas des « candidats » au sens propre?



Premier souci : la question n'est pas énoncée, la source n'est qu'un nom de firme

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

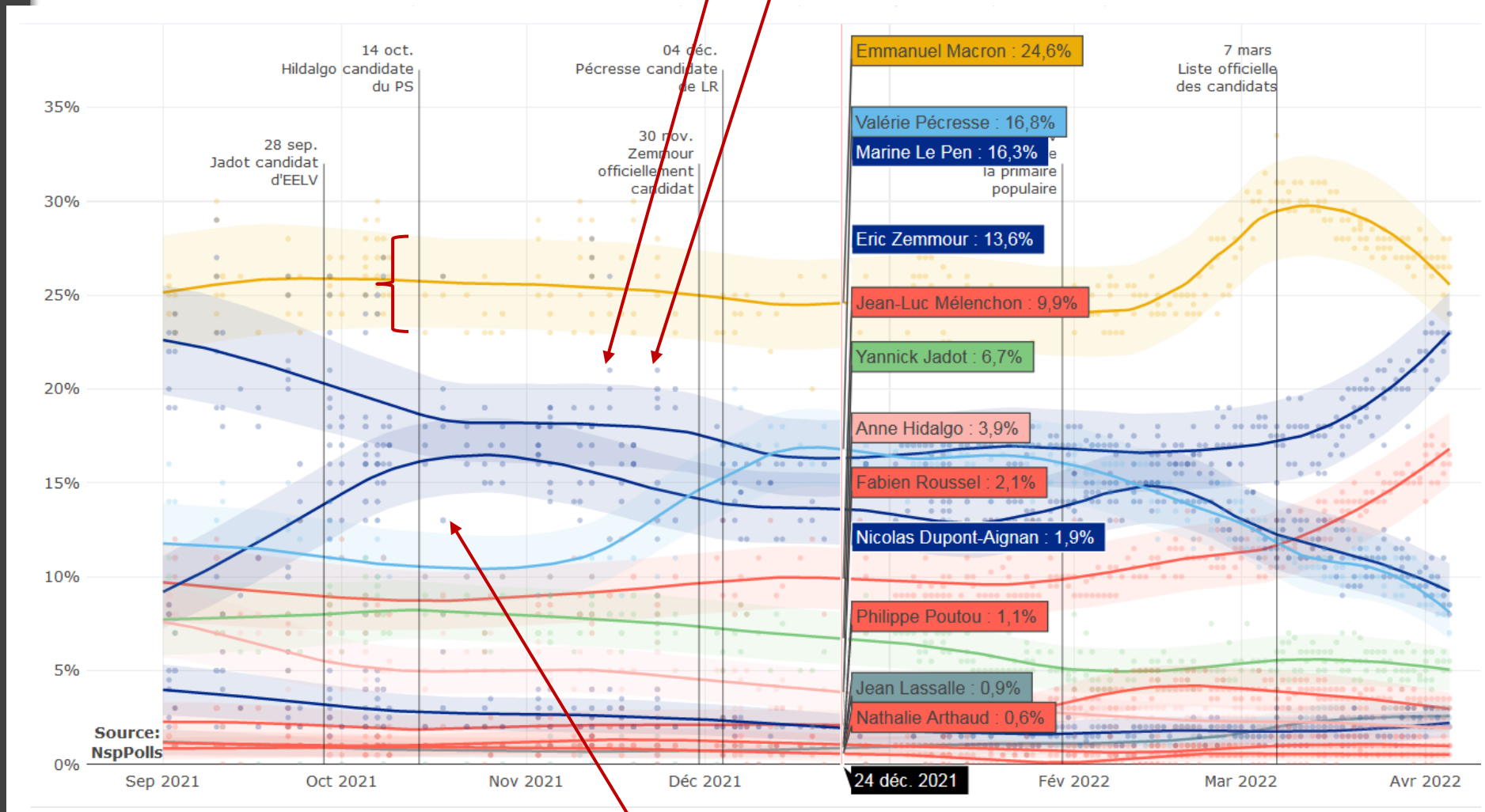
Choix des couleurs ? Sans doute en lien avec une connaissance supposée préalable ? Votre avis?



Deuxième souci : ce code couleur « supposé connu » impacte la lisibilité / des couleurs ressortent mieux que d'autres

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

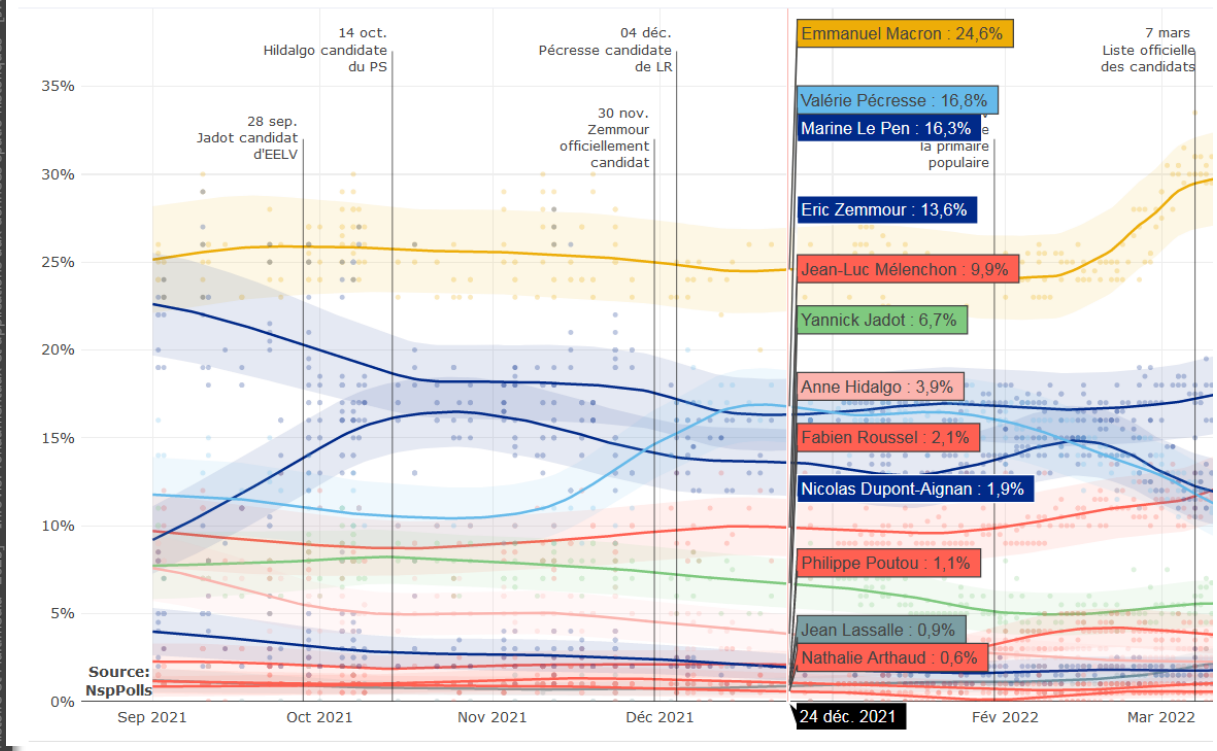
donnée de base sans doute ces points, traduite par une forme « ligne »
Fait mal lire les amplitudes

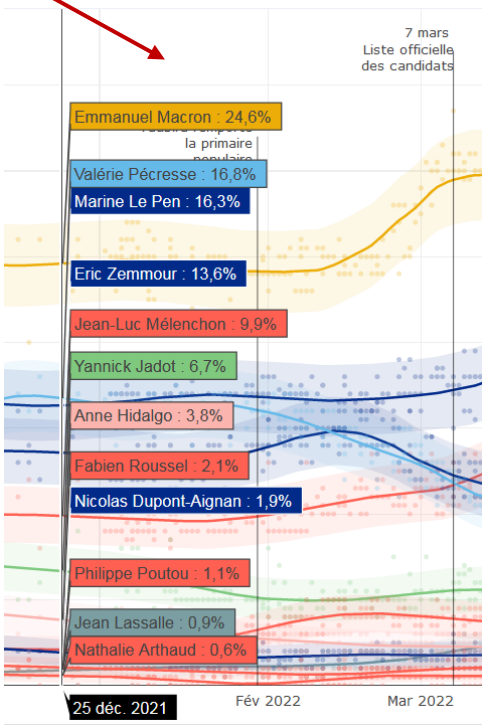


Troisième souci : comme lecteur on ne comprend pas comment se fait le passage de « points » vers une courbe moyenne, en particulier parce que des points sont « hors cadre » et non attribuables à l'un ou l'autre.
Donc on nous demande de croire sans comprendre, sans pouvoir raisonnner

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

Distribution verticale des étiquettes : pas de rapport avec la « valeur »

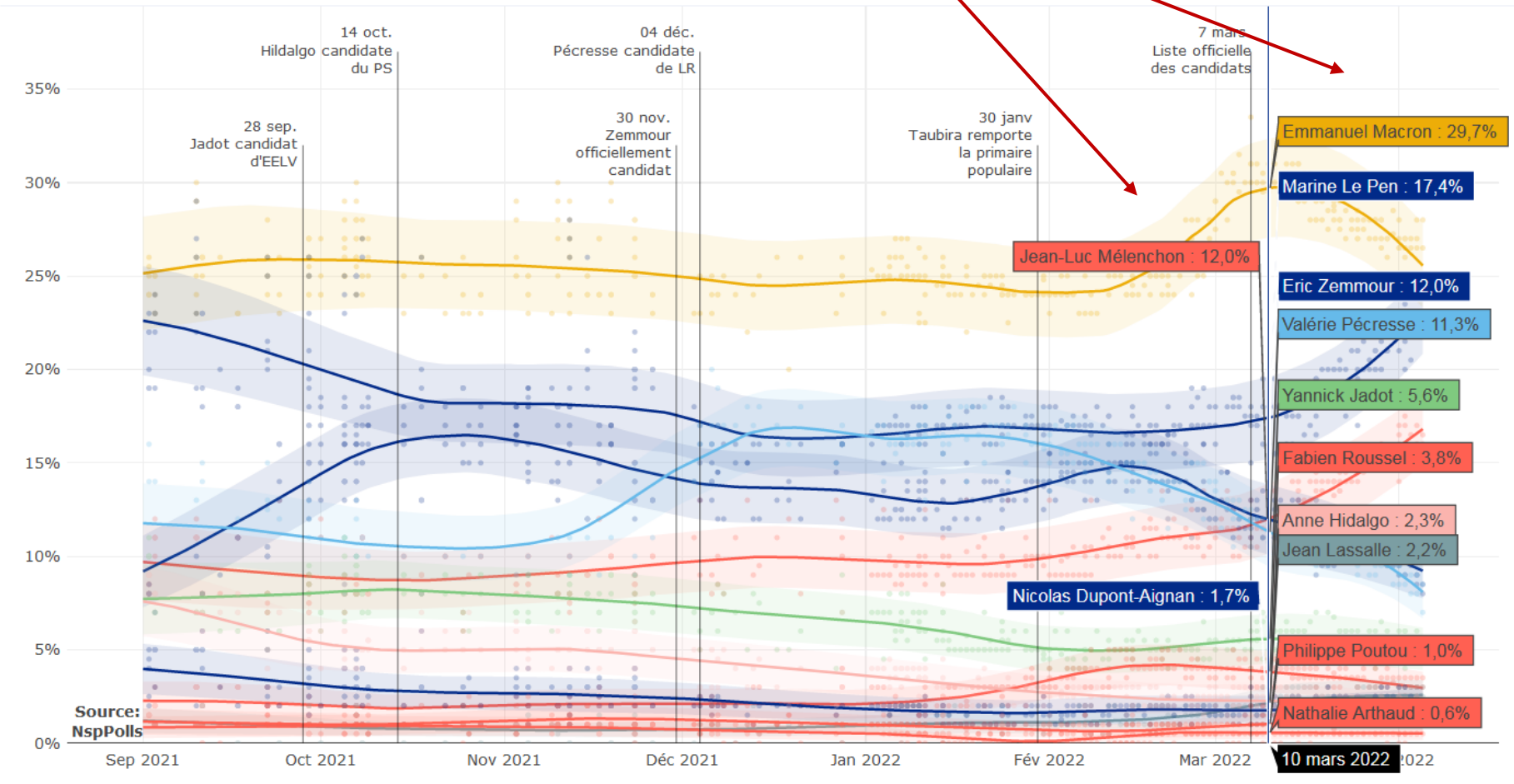




Quatrième souci : la position verticale des étiquettes, qui est en soi une indication visuelle, ne semble pas porter une quelconque information mais être le résultat d'un algorithme lambda

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

Distribution des étiquettes de part et d'autre de la barre verticale correspondant au jour choisi



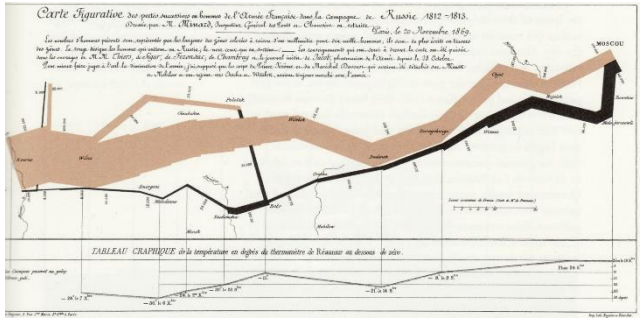
Cinquième souci : le dispositif souligne (visuellement) deux items, non pas par choix du concepteur de la visualisation mais faute d'espace (noms trop longs ..)

https://www.letelegramme.fr/elections/sondages/sondages-election-presidentielle-2022.php

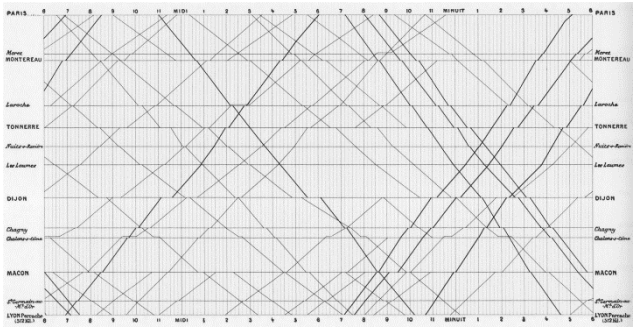
[CMI "Ingénierie en Histoire et Multimédia" 2025] InfoVis: fondamentaux et applications aux données spatio-historiques [J.Y Blaise I.Dudek]

Que retenir de cet héritage?

Une « bonne visualisation » n’est pas affaire de technologie, mais de créativité et de rigueur.



Le passage à une solution informatisée doit être compris comme une juste balance à trouver entre l’automatisation (efficacité algorithmique, reproductibilité) et la qualité (lisibilité, efficacité du dispositif en tant que vecteur de raisonnement)



Le fait qu’une solution « fonctionne » sur le plan informatique n’est pas un résultat suffisant: il faut s’astreindre à une lecture critique

A suivre: quelques règles / bonnes pratiques