

VISUALISER SES DONNÉES

Sékolène Albouy

CNRS | Observatoire de Paris

2020/12/03

Plan du cours



Notions générales

Grands principes des
visualisations de données

Voyant Tools

Analyse d'un corpus textuel

Xmind

Réaliser un modèle conceptuel

Tableau

Création de carte,



Barres empilées
(Stacked bar)



Barres
(Bar)



Pyramide des âges
(Mirror bar)



Barres empilées 100%
(100% Stacked bar)



Barres en cascade
(Bar range)



Champs
(Field)



Camembert
(Pie)



Donut
(Doughnut)



Jauge
(Dial arch)



Boîtes à moustache
(Box plot)



Bullet
(Bullet)



Streamgraph
(Streamgraph)



Coordonnées parallèles
(Parallel coordinates)



Courbes
(Line)



Sankey
(Sankey)



Radar
(Radar)



Aire de répartition
(Area range)



Nuage de points
(Scatter)



Réseau relationnel
(Force graph)



Réseau hiérarchique
(radial tree)



Réseau d'accords
(Chord)



Réseau de groupe
(Grouped tree)



Jauge
(Dial speedometer)



Filtres
(Filter)



Carte
(Map)



Choropleth
(Choropleth)



Table
(Table)



Tableau croisé
(Pivot)



Détail
(Detail)



Total: 123
 Σ : 1234
MIN: 7
MAX: 100
Contenu
(Content)



Aire
(Area)



Aires empilées
(Stacked area)



Aires empilées 100%
(100% stacked area)



Pyramide
(Pyramid)



Entonnoir
(Funnel)



Carte proportionnelle
(Tile)

Notions générales

Typologies de visualisation

[Visual poster](#)

Permet de classer en fonction de l'usage

[Dataviz catalog](#)

[Dataviz project](#)

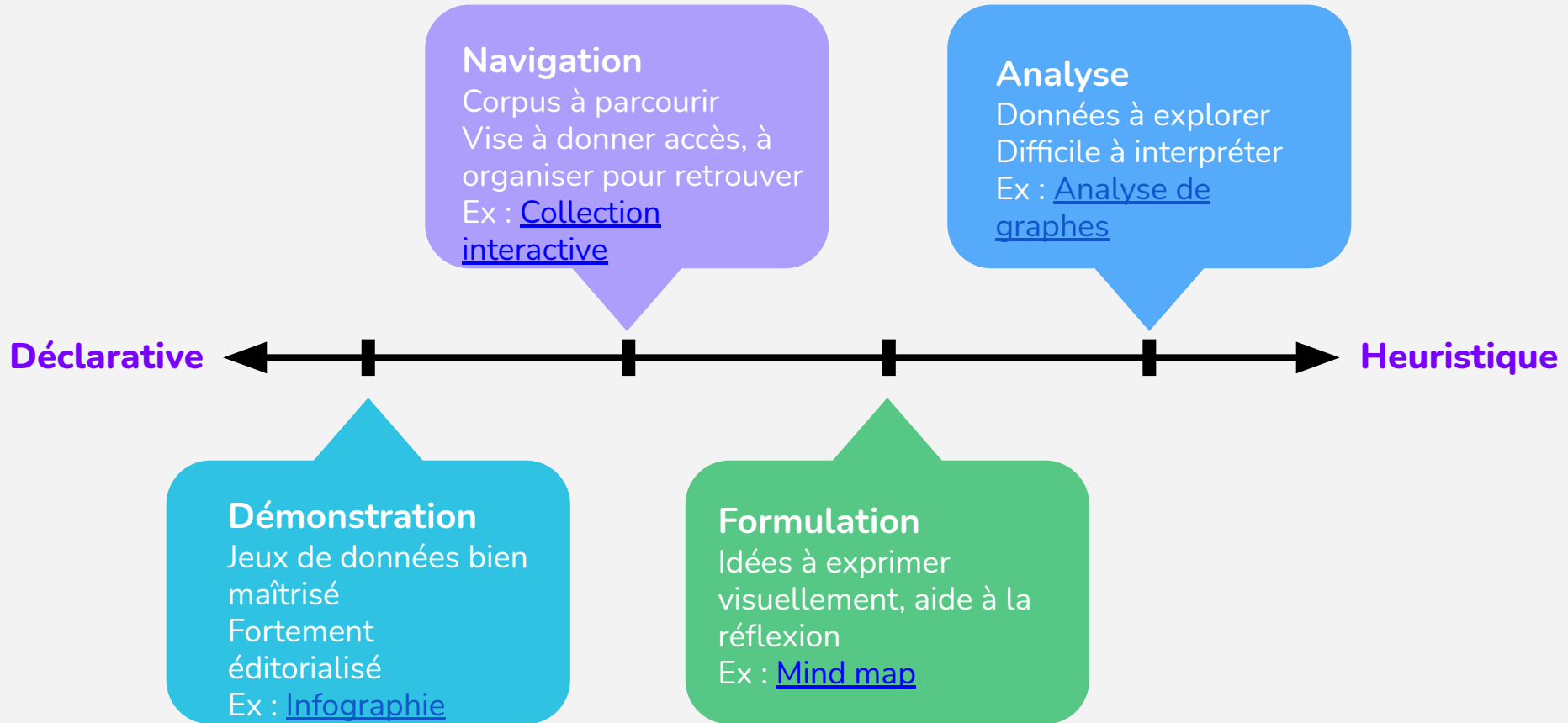
Montrer la diversité des modes de représentation

[Graphic continuum](#)

[Periodic table of visualization](#)

Offre des exemples d'utilisation et des guides de bonnes pratiques

Usages de visualisations



Utilité des visualisations de données

ACCESSIBILITÉ

Montrer les dimensions et les bornes d'un corpus

Aider le lecteur à trouver des points d'intérêt dans les données

COMPRÉHENSION

Fournir un support visuel à la compréhension

Apporter une appréhension globale d'un corpus

Repérer des irrégularités dans les données

ANALYSE

Révéler des corrélations insoupçonnées dans un corpus

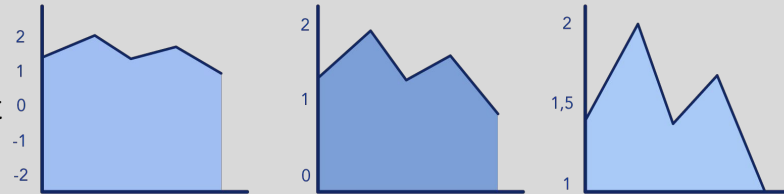
Manipuler et modeler l'information

Souligner le réseau d'interaction entre les données

Un outil, pas une preuve !

Visualiser = traduire

Parmi l'infinité de manières de transcrire visuellement l'information, certaines sont trompeuses



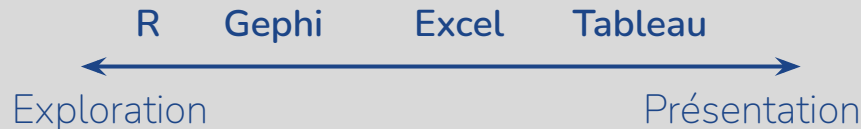
Les données sont biaisées

Les jeux de données sont généralement rassemblés dans un but précis et omettent souvent une partie de l'information

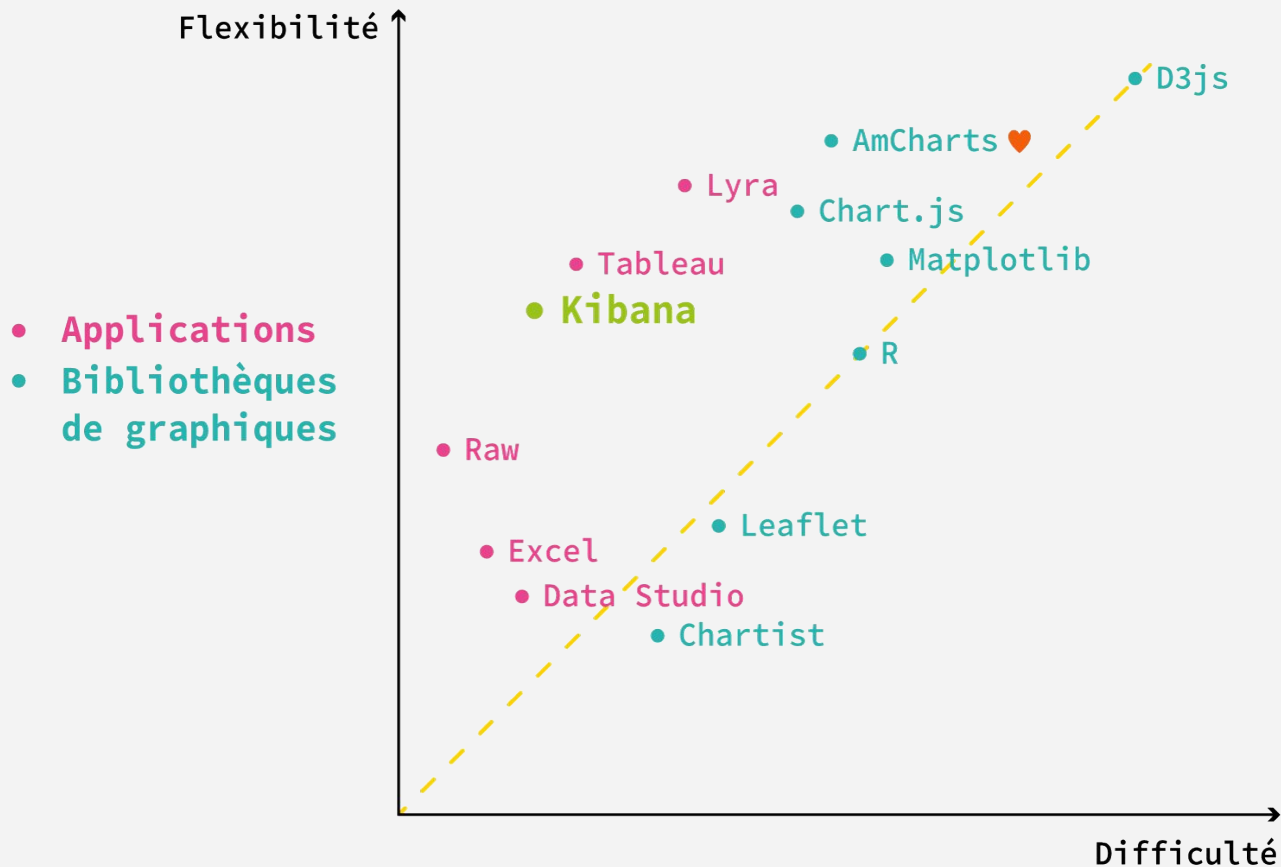


Choisir un outil adapté

Tous les outils de visualisations ne sont pas adaptés pour tout type de visualisation



Technos de *dataviz*



Granularité de la visualisation



Corpus

Point de vue englobant
Moins précis
Création de liens entre les données



Enregistrement

Focus sur un élément
Plus précis
Contextualise la donnée
dans son réseau de
connexion

Processus de conception

Quelle question
?



Processus de conception

Quelle question
?



Quelles données ?

```
tpq + taq  
comment  
public (savoir s'il est possible de montrer  
rice)  
pageMin + pageMax (→ nombre de page  
TABLE TYPE → name  
TABLE TYPE → ASTRO OBJECT → na  
PLACE → placeName  
HIST ACTOR → actorName + actorNam  
viaf  
WORK → title/incipit + tpq + taq  
WORK → HIST ACTOR (array  
actorNameOrig Char + tpq + taq + viaf  
WORK → PLACE → placeName  
PRIM SOURCE → shelfmark + digitall  
imTitle  
PRIM SOURCE → LIBRARY → libraryN  
ni  
EDITED TEXT (array) → editedTextTitle  
EDITED TEXT (array) → HISTORIAN →  
SCRIPT → scriptName + iso15924  
LANGUAGE → languageName + iso639
```

Processus de conception

Quelle question ?



Quelles données ?

```
tpq + taq  
comment  
public (savoir s'il est possible de montrer  
fice)  
pageMin + pageMax (→ nombre de page  
TABLE TYPE → name  
TABLE TYPE → ASTRO OBJECT → na  
PLACE → placeName  
HIST ACTOR → actorName + actorNam  
viaf  
WORK → title/incipit + tpq + taq  
WORK → HIST ACTOR (array  
actorNameOrig Char + tpq + taq + viaf  
WORK → PLACE → placeName  
PRIM SOURCE → shelfmark + digitall  
imTitle  
PRIM SOURCE → LIBRARY → libraryN  
ni  
EDITED TEXT (array) → editedTextTitle  
EDITED TEXT (array) → HISTORIAN →  
SCRIPT → scriptName + iso15924  
LANGUAGE → languageName + iso639
```

Quel outil ?

```
object.keys(this.dish  
g = Object.keys(this  
= this.dishasMap.dat  
  
n series that is dis  
= this.config.serie  
trieve all ids that a  
in the current plac  
tityName = this.conf  
ace.ids[entityName].  
add to the ids prop  
is.dishasMap.mapPins  
timerangeIds.filte  
change the radius a  
to dishasMap.define
```

Processus de conception

Quelle question ?



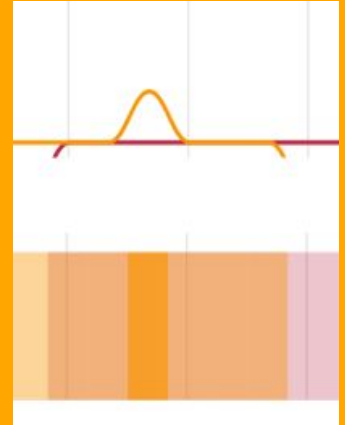
Quelles données ?

```
tpq + taq  
comment  
public (savoir s'il est possible de montrer  
fice)  
pageMin + pageMax (→ nombre de page  
TABLE TYPE → name  
TABLE TYPE → ASTRO OBJECT → na  
PLACE → placeName  
HIST ACTOR → actorName + actorNam  
viaf  
WORK → title/incipit + tpq + taq  
WORK → HIST ACTOR (array  
actorNameOrig Char + tpq + taq + viaf  
WORK → PLACE → placeName  
PRIM SOURCE → shellmark + digitall  
imTitle  
PRIM SOURCE → LIBRARY → libraryN  
ni  
EDITED TEXT (array) → editedTextTitle  
EDITED TEXT (array) → HISTORIAN →  
SCRIPT → scriptName + iso15924  
LANGUAGE → languageName + iso639
```

Quel outil ?

```
object.keys(this.dish  
g = Object.keys(this  
= this.dishasMap.dat  
  
n series that is dis  
= this.config.serie  
trieve all ids that a  
in the current plac  
tityName = this.conf  
ace.ids[entityName].  
add to the ids prop  
is.dishasMap.mapPins  
timerangeIds.filte  
change the radius a
```

Comment visuellement ?



Principes à garder en tête

CLARTÉ

Éviter la surcharge d'information

Utiliser des textes clairs et des couleurs tranchées

Plusieurs visualisations simples > 1 visualisation compliquée

COHÉRENCE

Utiliser une charte graphique uniforme

Garder les mêmes échelles pour comparer

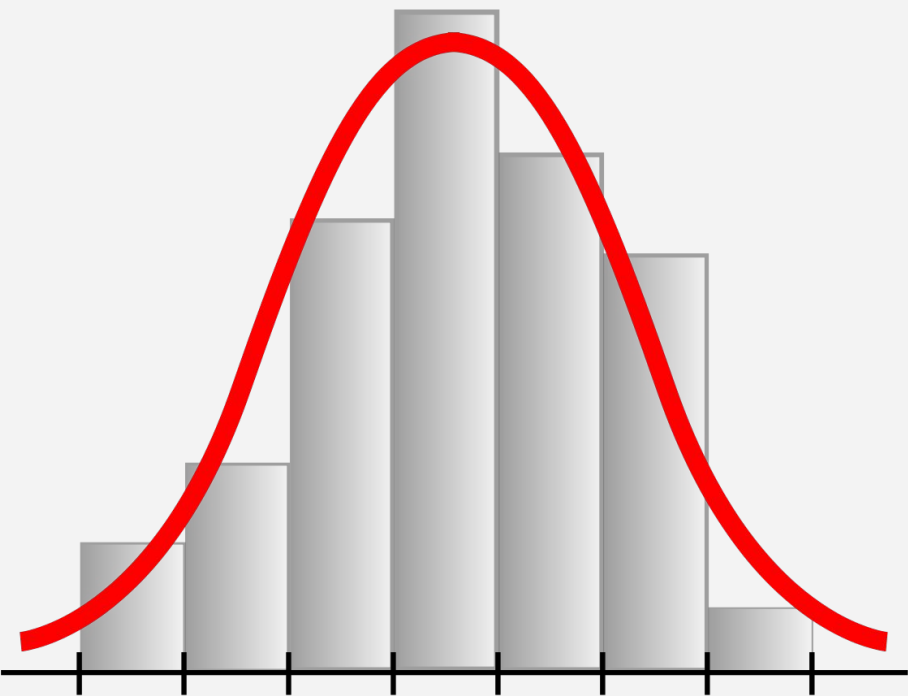
Dénomination et légendes harmonisées

COMMUNICATION

Ne pas tronquer les jeux de données ou les échelles

Utiliser un type de visualisation adapté

Ajouter des textes explicatifs



**Un peu de
statistiques**

Un peu de vocabulaire statistique

individu

Unité d'observation

Ce peut être des personnes, des objets, des textes, des mesures physiques, etc.

variable

Caractéristique des individus

Métadonnées descriptives des éléments étudiés

population
n

Ensemble d'individus

On note souvent sa taille **N**

échantillon
n

Portion d'une population

On note souvent sa taille **n**

effectif

Nombre d'individus

Taille d'une population ou d'un échantillon

fréquence

Échantillon / Population

Rapport d'un effectif particulier d'individus à la taille de la population

Types de variables

QUALITATIVE

Variable **catégories** :
elles expriment une
modalité

Ordinale : peuvent
être ordonnées

Nominale : non
ordonnables

QUANTITATIVE

Variable **numérique** :
elles expriment la
quantité

Manipulable
arithmétiquement

Discrète ou continue

LES AUTRES

Variable **dichotomique** :
valeur booléenne

Variable **temporelle** :
Dates, timestamp

Manières de quantification

MOYENNE

Variables
quantitatives

Moyenne
arithmétique =
somme / effectif

Moyenne
pondérée =
(somme / effectif) /
coefficient

MÉDIANE

Variables
ordinales et
quantitatives

Valeur qui permet
de couper une
population en **2**
parties égales

Répartition en
quartiles, déciles,
centiles, etc.

VARIANCE/ ÉCART-TYPE

Variables
quantitatives

Mesure de
dispersion / de l'
écart à la moyenne

+ c'est élevé →
+ les individus sont
répartis

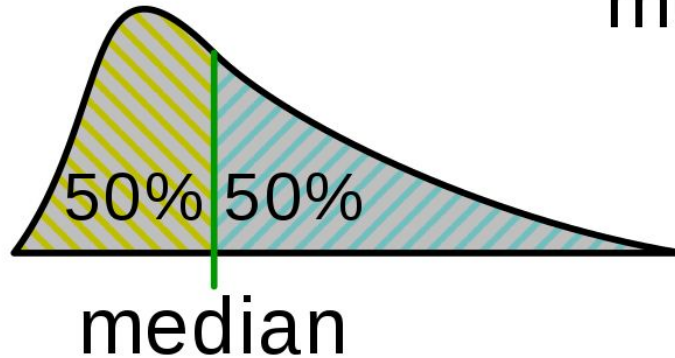
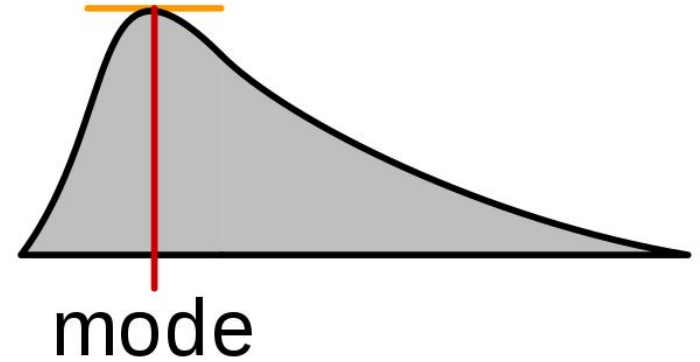
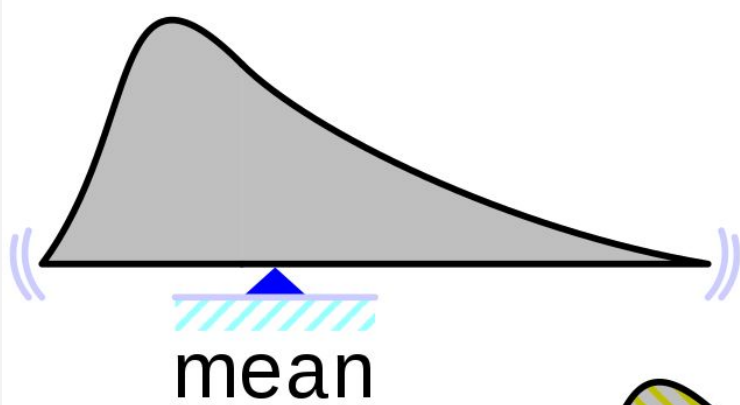
MODE

Tout type de
variables

Valeur **la plus**
représentée

Unimodale /
plurimodale

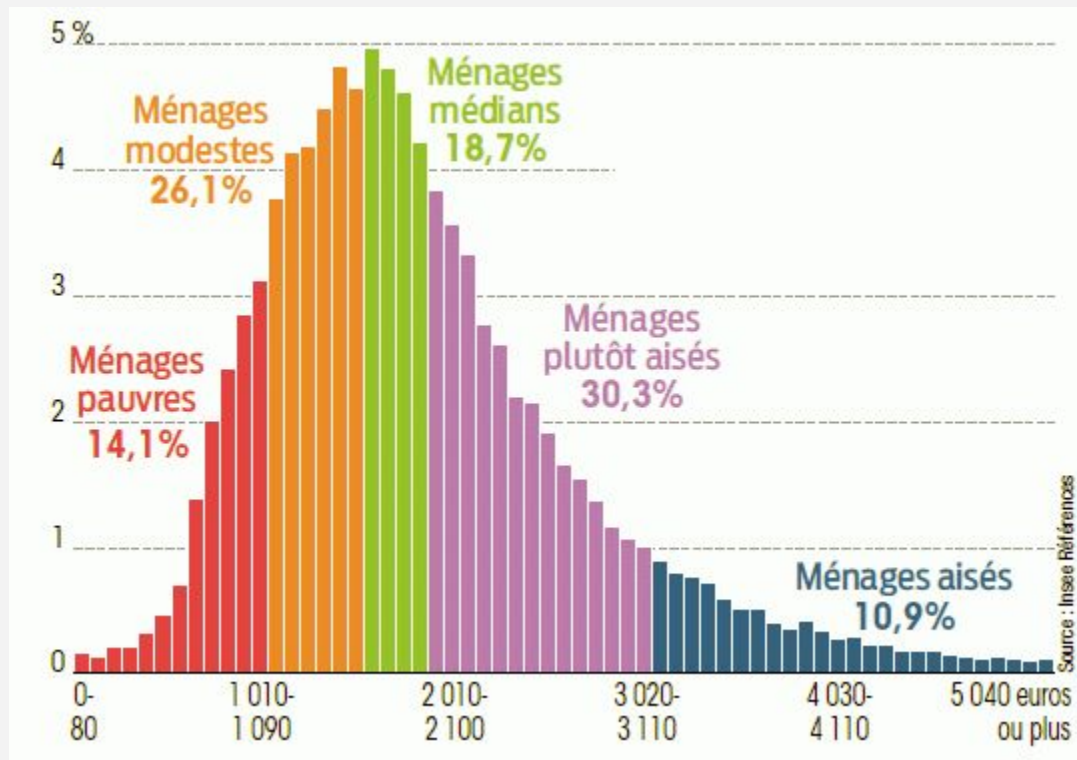
Moyenne / Médiane / Mode



Moyenne \neq Médiane

Salaire *moyen*
en France =
3275€/mois

Salaire *médian*
en France =
1780€/mois



Des principes à garder en tête !

Visualiser pas plus de 3 variables à la fois

Difficile de représenter une information à plus de trois dimensions en restant lisible
(longueur / largeur / taille)

Pondérer les données

La pondération c'est la contextualisation d'une variable, mettre un nombre en regard de la taille de sa population

Choisir une mesure adaptée

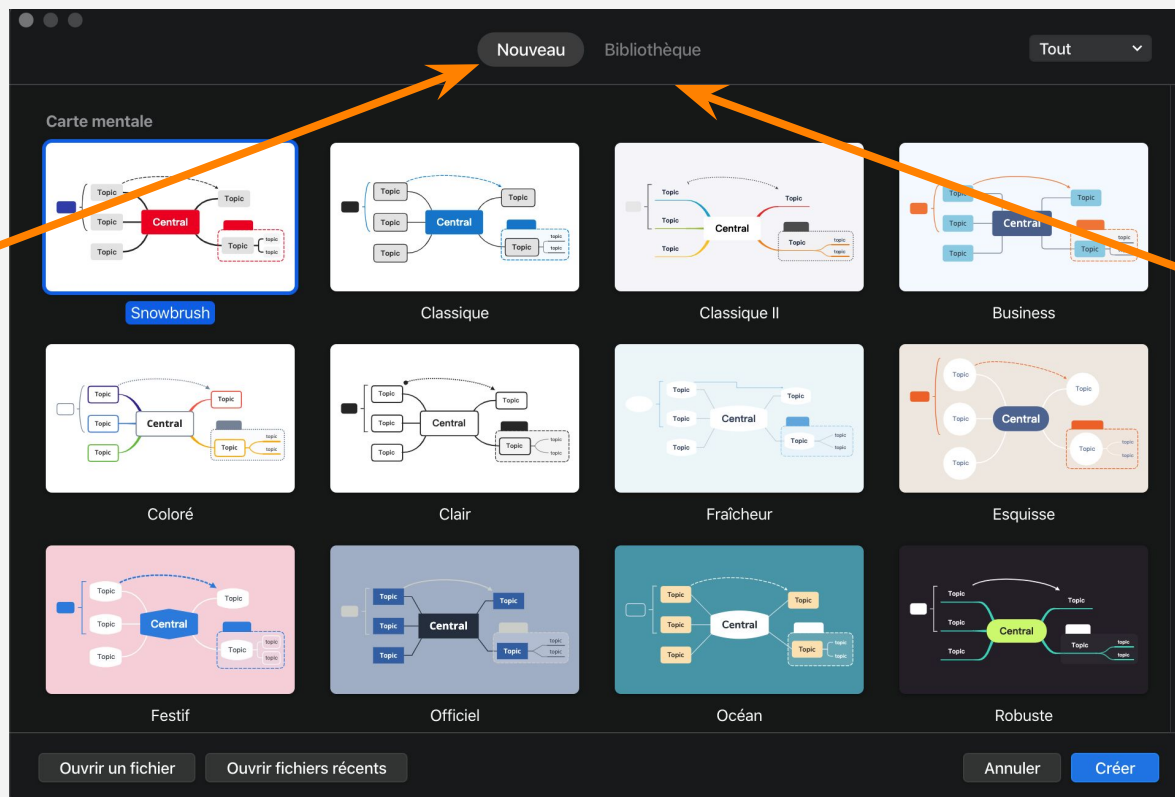
Si on veut mesurer la répartition de certaines variables dans un échantillon, penser à prendre la mesure adaptée pour en décrire la dispersion



Xmind

—

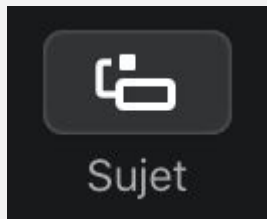
Xmind



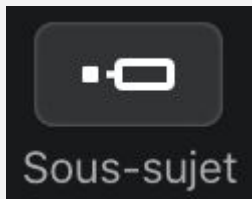
Création d'une nouvelle carte à partir d'un thème

Galerie d'exemples

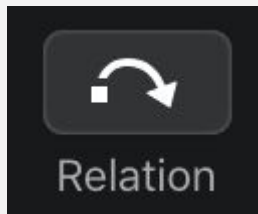
Xmind



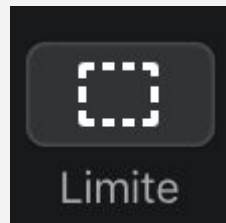
Création d'un sujet
connexe



Création d'un
sous-sujet



Création d'un
sous-sujet

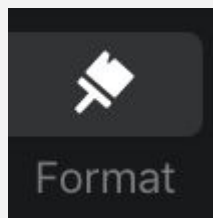


Création d'une
boîte autour de la
sélection



Création d'une
accolade autour de
la sélection

Cliquer/glisser
pour détacher un
sujet



Changer le style
de tout élément

Xmind

- Existe en version mobile (cartes compatible avec la version Desktop)
- Possibilité de lier des fichiers, ajouter des notes
- Le style de tout élément est modifiable
- Export dans différents formats



Voyant Tools

Import

Import par URL
Ou en collant le
texte

Ajouter des textes 

Entrez une URL (ou plusieurs sur des lignes séparées) ou bien copiez tout un texte.

 Ouvrir  Charger  Révéler

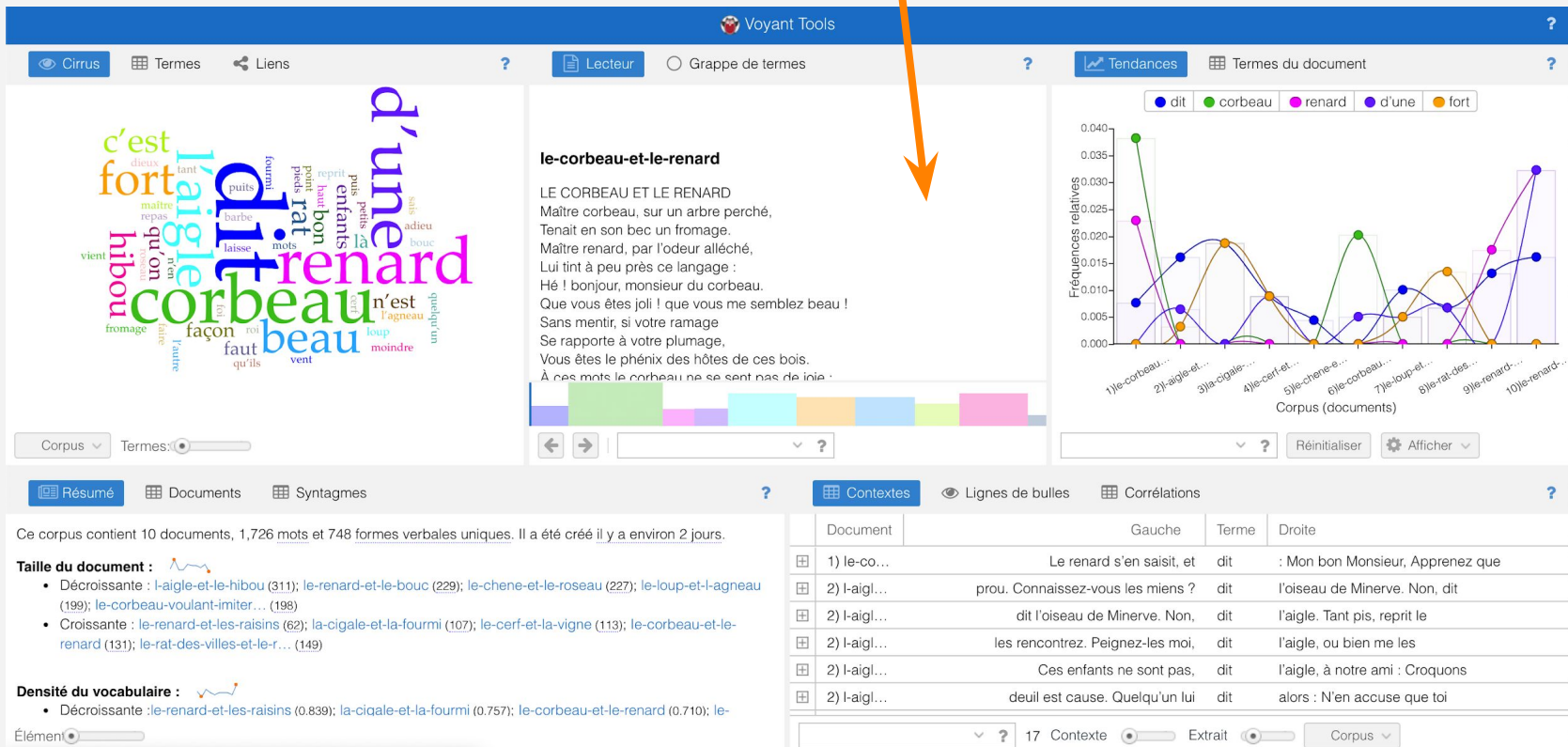
*Voyant Tools est un environnement en ligne de lecture et d'analyse de textes numériques.
Traduction française d'Aurélien Berra*

Import d'un seul
fichier

Import de
plusieurs fichiers

Dashboard

Fenêtres de visualisations
interchangeables



<https://voyant-tools.org/docs/#!/guide>

Fonctionnalités



Export de la
visualisation



Choix d'une autre
visualisation



Configuration de la
visualisation



Documentation &
explications

Cliquer sur un mot
pour filtrer le
dashboard

Réinitialiser

Retirer le filtre

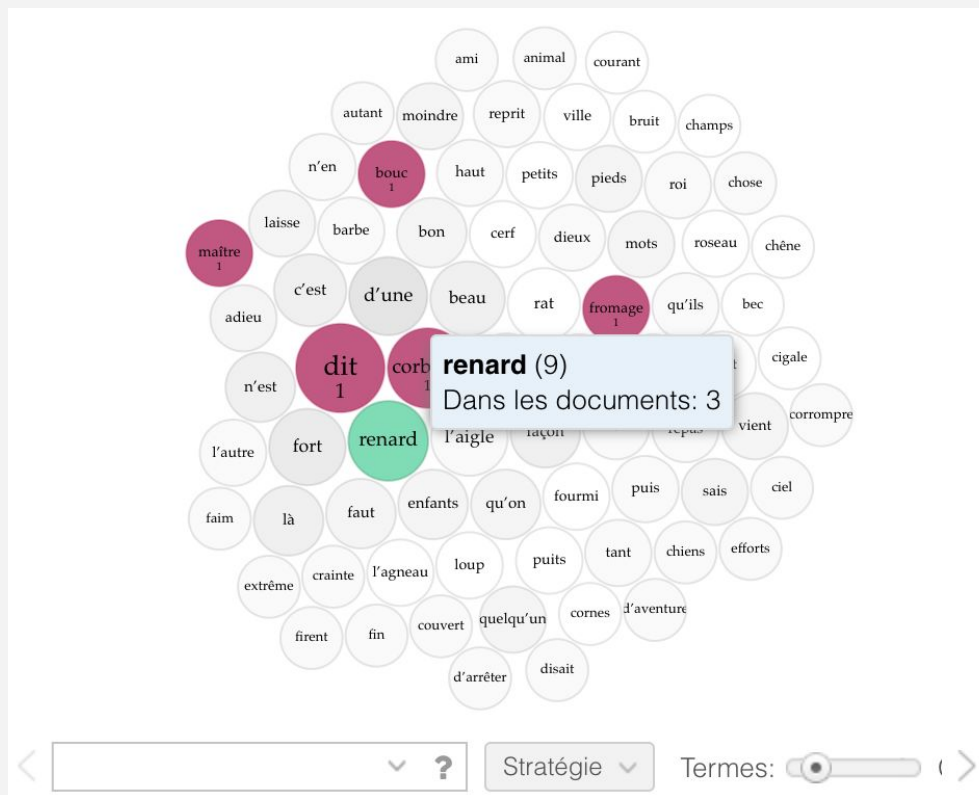
Cirrus



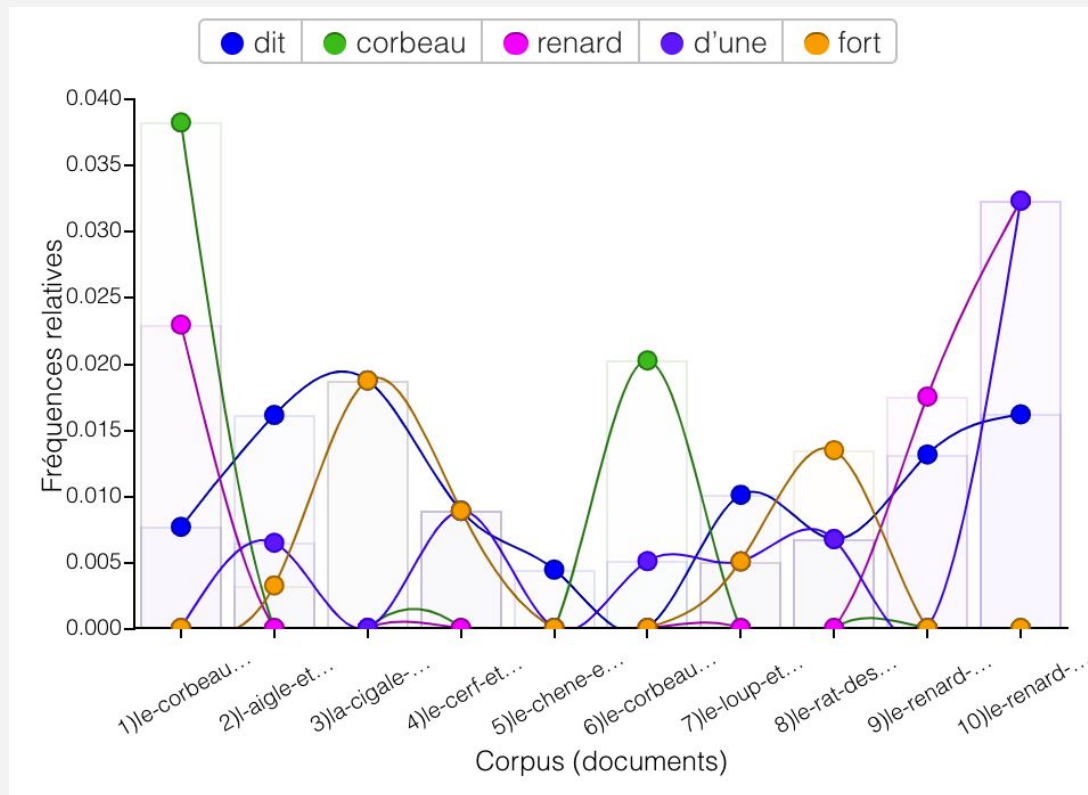
Corpus ▼

Termes:

Grappe de termes



Tendances



Contextes / Arbre de mot

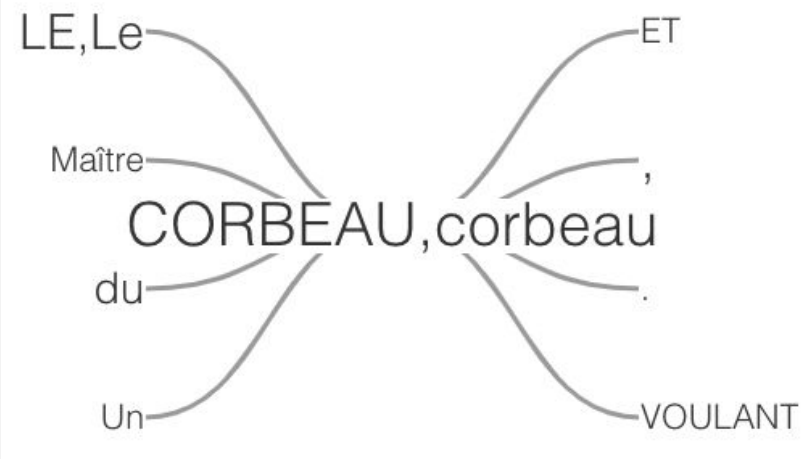
Contextes

Lignes de bulles

Corrélations



	Document	Gauche	Terme	Droite
+	1) le-co...	LE	corbeau	ET LE RENARD Maître corbeau
+	1) le-co...	CORBEAU ET LE RENARD Maître	corbeau	, sur un arbre perché, Tenait
+	1) le-co...	langage : Hé ! bonjour, monsieur du	corbeau	. Que vous êtes joli ! que
			mots le	corbeau ne se sent pas de
			oute. Le	corbeau , honteux et confus, Jura, mais





Tableau

Types de contenus

FEUILLE DE CALCUL

Création d'une
visualisation

Analyse d'un aspect
du corpus

TABLEAU DE BORD

Regroupement de
différentes feuilles de
calcul en une même
page

HISTOIRE

Présentation des
informations et
visualisations sous
forme scénarisée et
dynamique

Vocabulaire

dimension

Variable qualitative

Permet de catégoriser les données

mesure

Variable quantitative

Permet de faire des opérations mathématiques

Possibilité de changer le type d'une variable

Page d'accueil

The image shows the Tableau Public home page with several annotations in French. On the left, a dark blue sidebar contains a 'Connexion' (Connect) section with options like 'Dans un fichier' (In a file) and 'Sur un serveur' (On a server). An orange arrow points from the text 'Basculer page d'accueil / espace de création' (Switch home page / creation space) to the 'Connexion' header. Another orange arrow points from 'Ajout d'un nouveau set de données' (Add a new set of data) to the 'Fichier de données spatiales' (Spatial data file) option. The main content area has a 'Découvrir' (Discover) section with a list of resources, a 'LA STAR DU JOUR' (Star of the Day) feature, and a 'Ressources' (Resources) section. An orange arrow points from the text 'Tutoriel de découverte' (Discovery tutorial) to the 'Découvrir' header. Another orange arrow points from 'Exemple de visualisations' (Example of visualizations) to the 'LA STAR DU JOUR' feature. At the bottom right, an orange button says 'Passer à la version [site:tableau_version] maintenant' (Upgrade to version [site:tableau_version] now).

Connexion

Ouvrir

Basculer page d'accueil / espace de création

Ajout d'un nouveau set de données

Découvrir

Ouvrir à partir de Tableau Public

Tutoriel de découverte

Exemple de visualisations

Passer à la version [site:tableau_version] maintenant

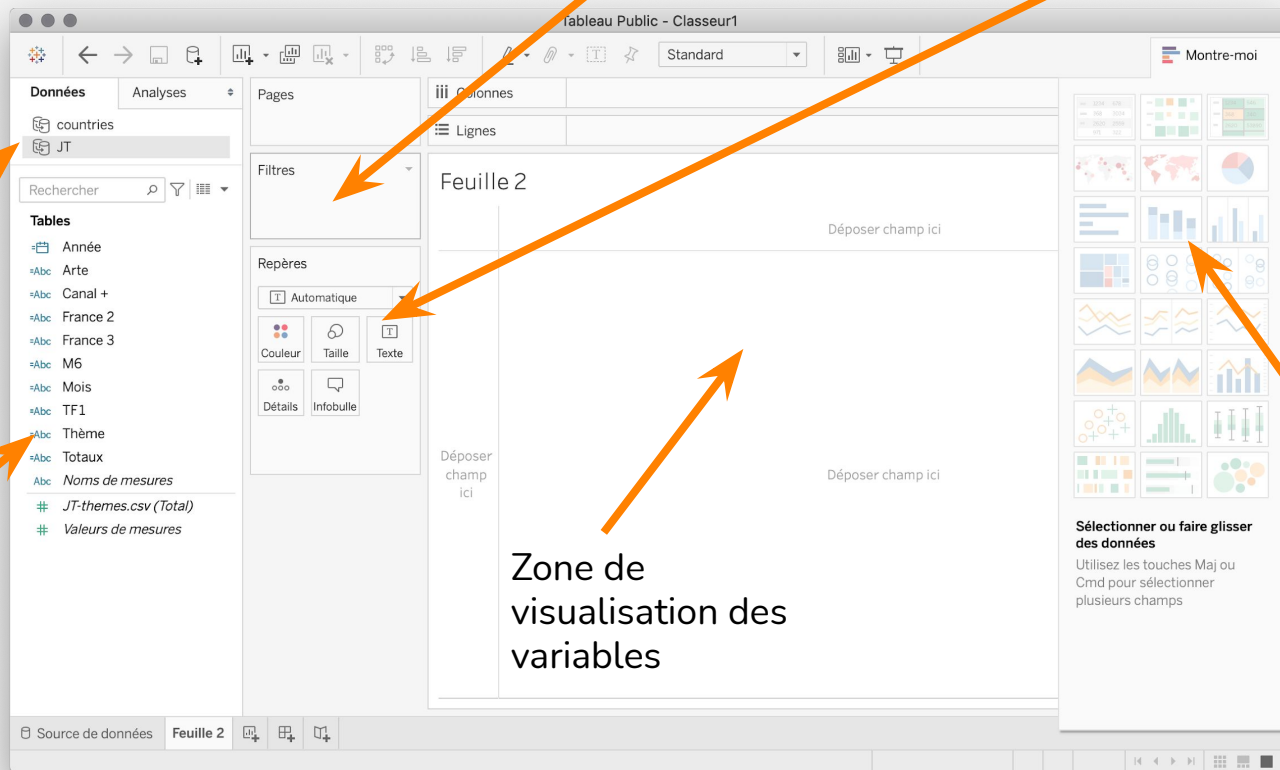
Espace de travail

Option de filtre

Options graphiques

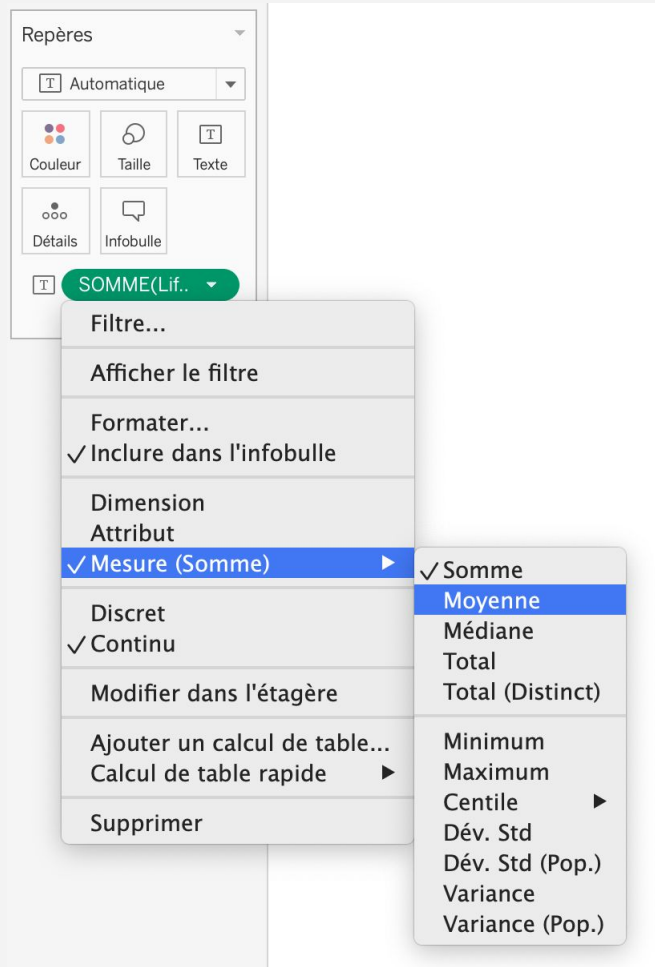
Sources de données

Variables (colonnes)

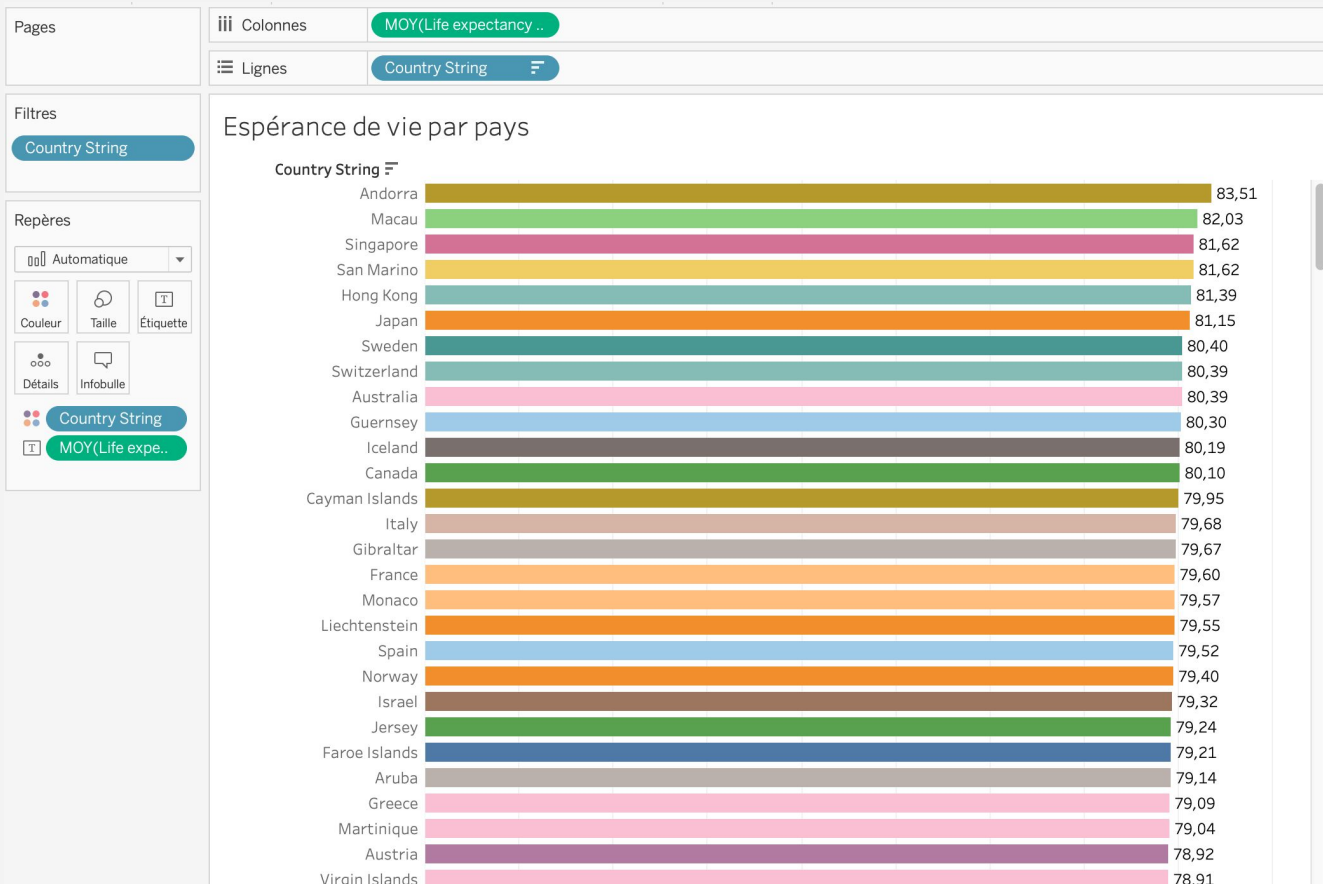


Liste de styles prédéterminés pour les données

Mesures



Dupliquer une variable



Ctrl +
cliquer/glisser