



Jean-Samuel Leboeuf

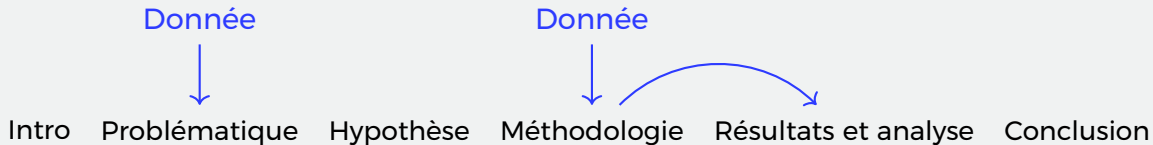
Communiquer efficacement ses
données: un guide pratique

5 mai 2021

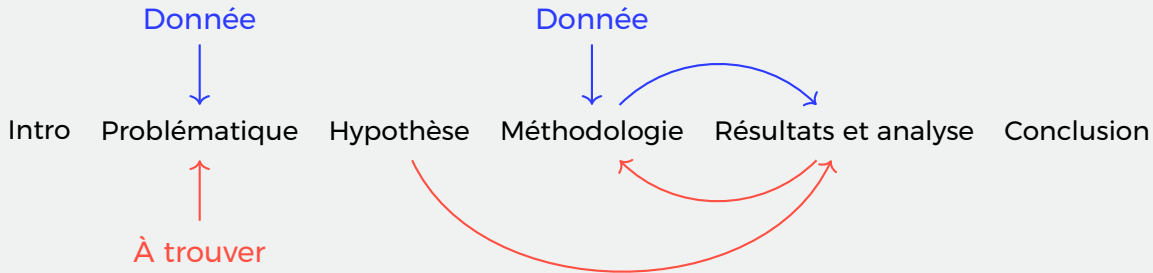
La communication scientifique

Intro Problématique Hypothèse Méthodologie Résultats et analyse Conclusion

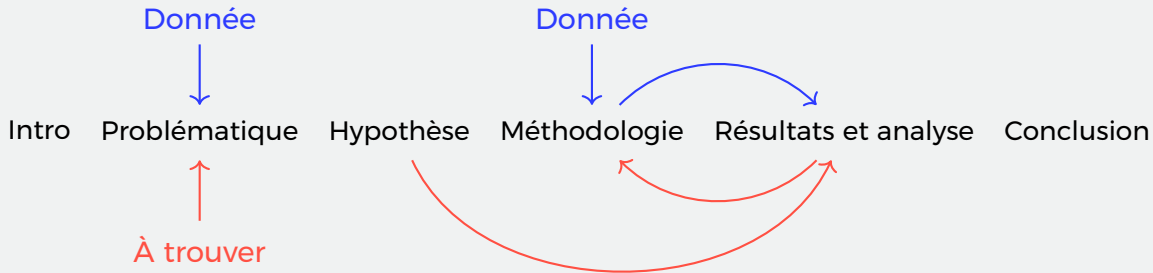
La communication scientifique



La communication scientifique



La communication scientifique



- Résultats **convaincants** et **faciles à comprendre**

Plan

1. Le principe de coût initial négatif
2. Les tableaux
3. Principes premiers des figures
4. Les types de graphiques
5. Les couleurs

Coût initial négatif

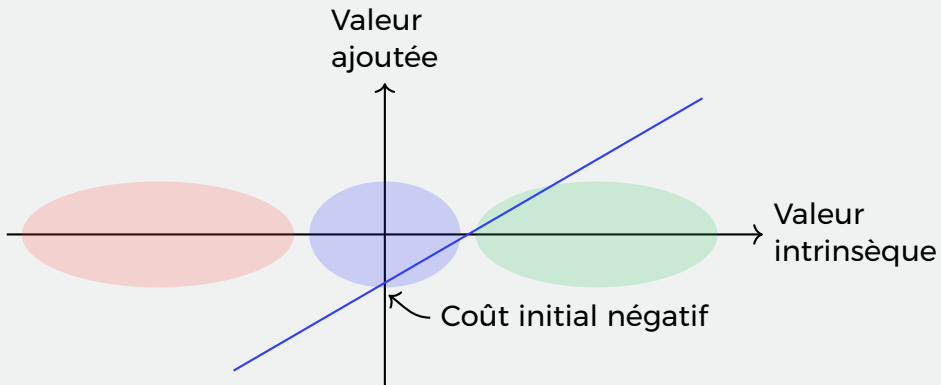
Faux, ambigu, contradictoire

Utile, convaincant

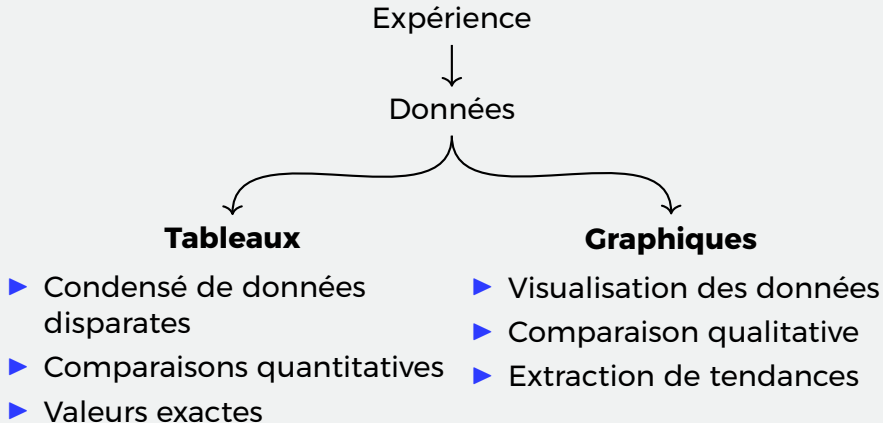


Superflu, redondant, **distayant**

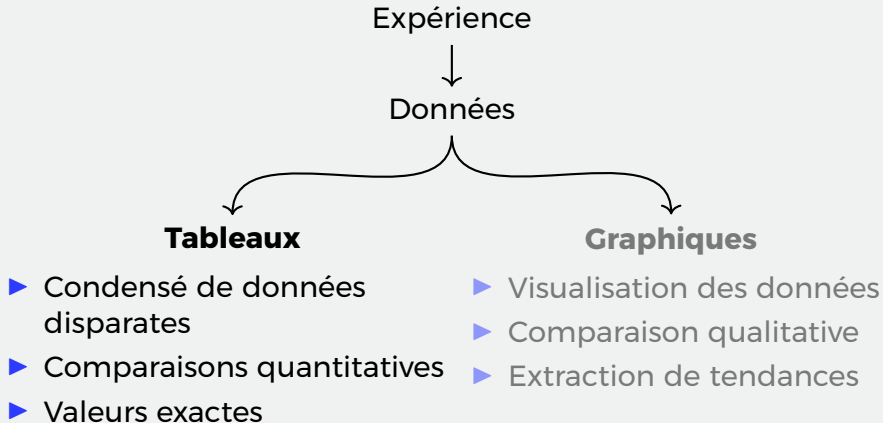
Coût initial négatif



Présentation des résultats



Présentation des résultats



Deux types de tableaux

Tableau de **communication**

Tableau de **compilation**

Deux types de tableaux

Tableau de **communication**

- ▶ Résumé utile des informations
- ▶ Concis et facile à lire
- ▶ Ne contient pas toutes les informations (coût initial négatif)
- ▶ Met en évidence des différences ponctuelles
- ▶ Dans le **corps** d'un document

Tableau de **compilation**

Deux types de tableaux

Tableau de **communication**

- ▶ Résumé utile des informations
- ▶ Concis et facile à lire
- ▶ Ne contient pas toutes les informations (coût initial négatif)
- ▶ Met en évidence des différences ponctuelles
- ▶ Dans le **corps** d'un document

Tableau de **compilation**

- ▶ Contient toutes les informations recueillies
- ▶ Se veut une référence
- ▶ Longs et plates à lire
- ▶ Dans l'**annexe** d'un document

Les tableaux de communication

ResNet-50 (He et al., 2016)	76.9
ViT-B-16 (Dosovitskiy et al., 2021)	77.9
ResNet-50 (RGB+FF)	73.5
ViT-B-16 (RGB+FF)	76.7
Transformer (64x64)	57.0
Perceiver	76.4

Table 1. Top-1 validation accuracy (in %) on ImageNet. Methods shown in **red** exploit domain-specific grid structure, while methods in **blue** do not. The first block reports standard performance from

Tiré de l'article « Perceiver : General Perception with Iterative Attention » par DeepMind (Google)

Les tableaux de communication

ResNet-50 (He et al., 2016)	76.9
ViT-B-16 (Dosovitskiy et al., 2021)	77.9
ResNet-50 (RGB+FF)	73.5
ViT-B-16 (RGB+FF)	76.7
Transformer (64x64)	57.0
Perceiver	76.4

Table 1. Top-1 validation accuracy (in %) on ImageNet. Methods shown in **red** exploit domain-specific grid structure, while methods in **blue** do not. The first block reports standard performance from

Tiré de l'article « Perceiver : General Perception with Iterative Attention » par DeepMind (Google)

Table 1. Top-1 validation accuracy on ImageNet. Blablabla. . .

Model	Top-1 (%)		Uses grid
	No FF	FF	
ResNet-50	76.9	73.5	yes
ViT-B-16	77.9	76.7	yes
Transformer	—	57.0	no
Perceiver	—	76.4	no

Conseils généraux

- ▶ Étiquettes des colonnes avec unités et abbréviations limitées
- ▶ Tableau \neq grille :
 - ▶ Lignes horizontales : début, fin et entre les sections seulement
 - ▶ JAMAIS de lignes verticales : ajuster l'espace entre les colonnes
- ▶ Mettre en évidence une seule chose à la fois
- ▶ Titre *au-dessus* du tableau avec explications
- ▶ Éviter les répétitions
- ▶ Citer dans le titre ou le texte
- ▶ Éviter les couleurs
- ▶ Référer au tableau explicitement dans le texte
- ▶ Aligner les chiffres à droite, le texte à gauche

Tableaux professionnels avec LaTeX

Les packages :

- ▶ `booktabs` : Redéfinit entièrement les tableaux ([documentation](#))
- ▶ `multicol` et `multirow` : Fusionner des cellules

Outil d'aide à la génération de tableaux :

- ▶ <https://www.tablesgenerator.com/>
- ▶ `python2latex` (Création du code LaTeX à partir de `numpy` arrays)

Commandes LaTeX

Les colonnes :

- ▶ `@{\hskip<distance>}` pour ajuster la distance entre deux colonnes

Le titre :

- ▶ `\caption{<titre>}` pour le titre
- ▶ `\vspace{<distance>}` pour ajuster la distance entre le titre et le tableau

Les lignes :

- ▶ `\toprule` et `\bottomrule` pour les lignes du haut et du bas
- ▶ `\midrule` pour les lignes entre sections
- ▶ `\cmidrule(lr){<col début>-<col fin>}` pour les lignes partielles

Les cellules fusionnées :

- ▶ `\multirow{<n rangées>}{*}[<y shift>]{<contenu>}` pour les rangées
- ▶ `\multicolumn{<n col>}{c}{<contenu>}` pour les colonnes

Example

```
\usepackage{booktabs}

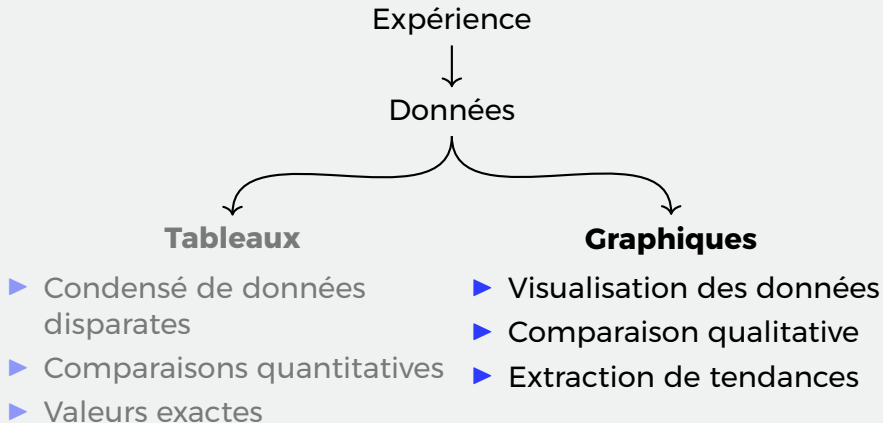
\begin{table}
\caption{Top-1 validation accuracy on ImageNet.}
\label{tab:top1_imagenet}
\vspace{3pt}

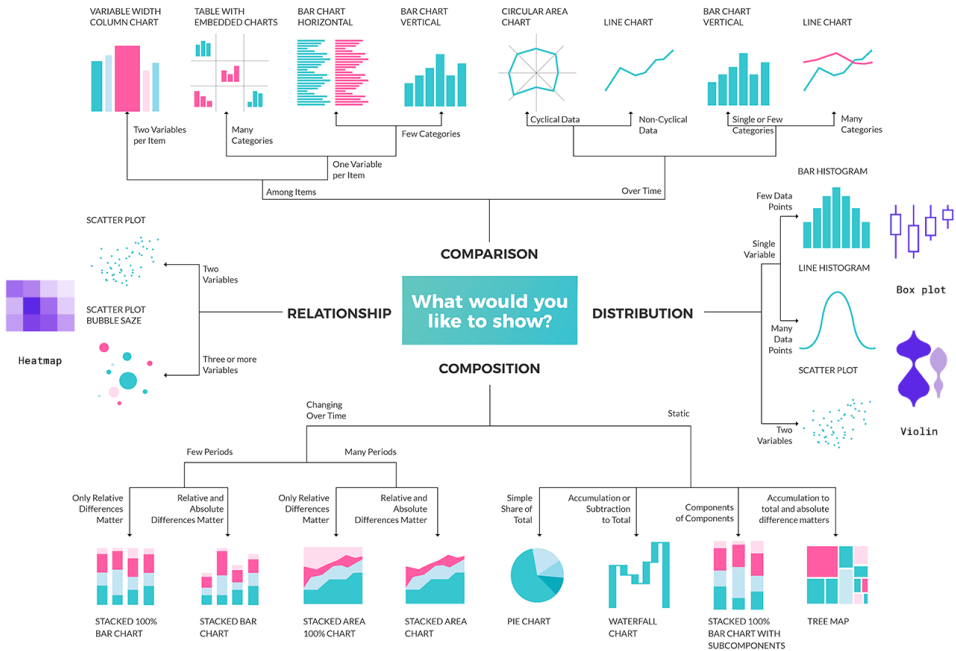
\begin{tabular}{l@{\hspace{5pt}}r@{\hspace{10pt}}r@{\hspace{5pt}}c}
\toprule
\multirow{2}{*}{[-2pt]{Model}}
& \multicolumn{2}{c}{Top-1 (\%)}
& \multirow{2}{*}{[-2pt]{Uses grid}}
\\
& No FF & FF & 
\\
\cmidrule{r}{2-3}
& No FF & FF & \\
\midrule
ResNet-50 & 76.9 & 73.5 & yes\\
ViT-B-16 & \textbf{77.9} & 76.7 & yes\\
Transformer & --- & 57.0 & no\\
Perceiver & --- & 76.4 & no\\
\bottomrule
\end{tabular}
\end{table}
```

Top-1 validation accuracy on ImageNet.

Model	Top-1 (%)		Uses grid
	No FF	FF	
ResNet-50	76.9	73.5	yes
ViT-B-16	77.9	76.7	yes
Transformer	—	57.0	no
Perceiver	—	76.4	no

Les graphiques





Principes premiers d'une figure

1. Représentation fidèle, non-trompeuse et sans distorsions
2. Représentation claire et évidente
3. Représentation facile à analyser et à comparer

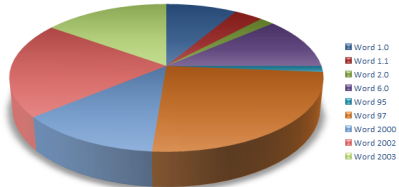
Représentation fidèle

- ▶ Ne donne pas de fausses impressions

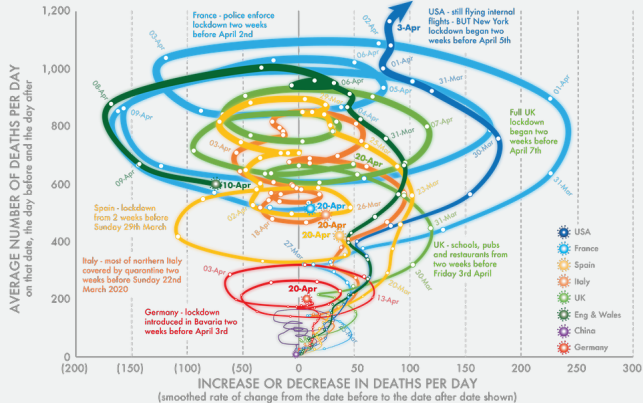
Same Data, Different Y-Axis



Microsoft Word Features By Version Added



Représentation claire



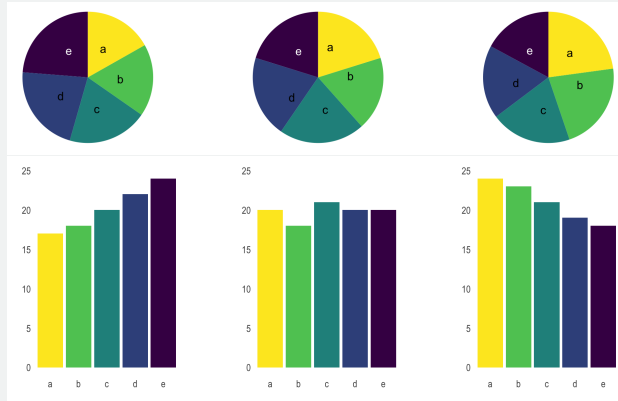
DannyDarling.org. Illustration by Kirsten McClure @orpheuscol

Représentation facile à analyser

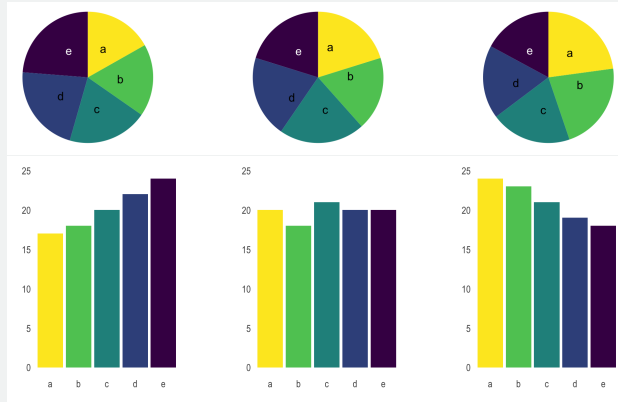
Représentation facile à analyser



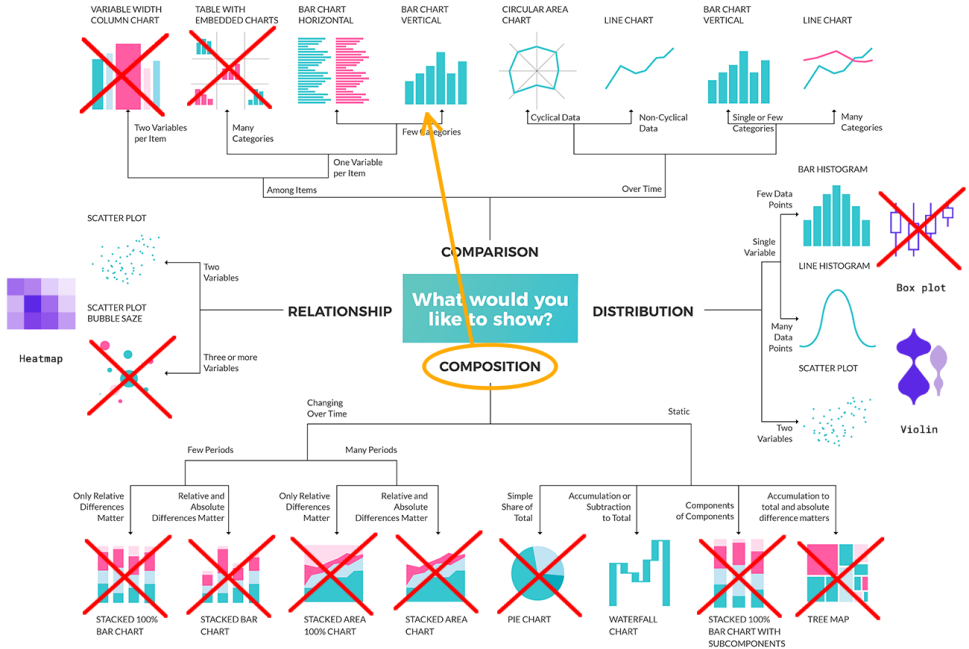
Représentation facile à analyser



Représentation facile à analyser



- ▶ Le cerveau est poche pour comparer des aires et des distances
- ▶ Le cerveau est bon pour comparer des positions



Guide général de style

Les **axes** :

- ▶ En bas et à gauche seulement
- ▶ Étiquettes avec unité
- ▶ Discrétisation modérée (5 à 10)
- ▶ Police de taille similaire au texte
- ▶ Grille aérée et pâle; **pas** de fond

Le **titre** :

- ▶ Préférentiellement sous la figure
- ▶ Numérotée et référencée dans le texte
- ▶ Explications spécifiques non redondantes

Guide pour les courbes

- ▶ Éviter les légendes si possibles
 - ▶ Sinon, lister les entrées en ordre d'apparition
- ▶ Couleurs pour différencier les courbes
- ▶ Lignes pleines, discontinues ou pointillées pour différencier les familles de courbes
- ▶ 8 courbes maximum par diagrammes

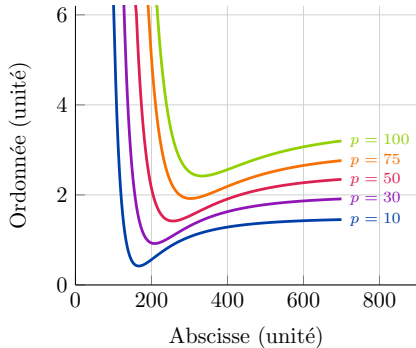


Figure 1. Exemple d'un diagramme à courbes. La légende est remplacée par une étiquette placée à la fin des courbes, ce qui facilite la lecture du graphique.

Guide pour les nuages de points

- ▶ Formes et couleurs pour distinguer les catégories
- ▶ Opacité pour les données superposées
- ▶ Pour une troisième variable : utiliser une carte de couleurs, pas la taille des points

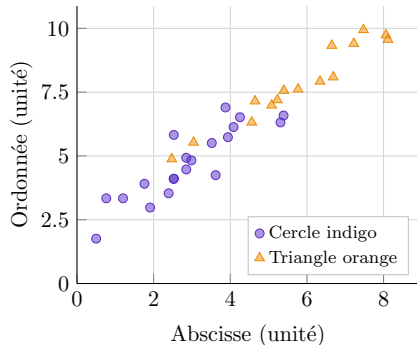


Figure 1. Exemple d'un nuage de point avec deux types de points. De la transparence est utilisée pour voir les points superposés.

Guide pour les diagrammes à bandes

- ▶ Repères verticaux non-nécessaires
- ▶ Axe vertical commence à l'origine
- ▶ Données près de la bande
- ▶ Bandes en ordre croissant
- ▶ Bandes à l'horizontal si les étiquettes ne rentrent pas (ne pas pivoter le texte)

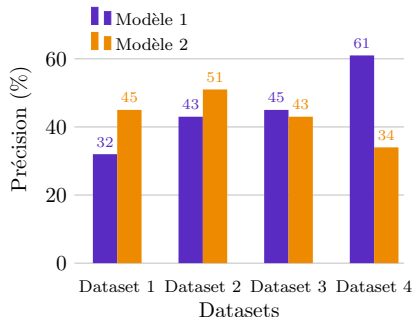


Figure 1. Exemple d'un diagramme à bandes avec des catégories.

Guide pour les « heatmaps »

- ▶ Se lit comme une matrice : origine en haut à droite
- ▶ Carte de couleurs :
 - ▶ Séquentielle = intensité
 - ▶ Divergente = déviation
 - ▶ Considérer discrétiser les couleurs
- ▶ Blanc pour les données non pertinentes
- ▶ Annotations pour les données pertinentes

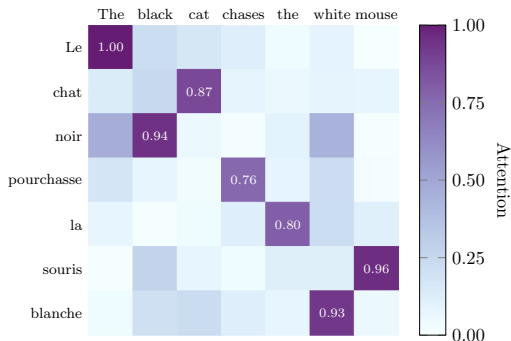


Figure 1. Exemple d'une « heatmap » pour l'attention d'un réseau pendant la traduction. Seuls les points d'intérêts sont annotés.

Guide pour les violons

- ▶ « Boîte à moustaches ++ »
- ▶ Permet d'identifier les modes
- ▶ Peuvent être groupés comme les diagrammes à bandes
- ▶ Normaliser l'**aire**, pas la largeur
- ▶ Pas de palette séquentielle

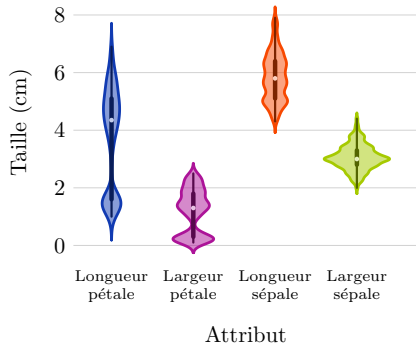


Figure 1. Exemple d'un diagramme à violons pour la distribution des attributs du jeu de données *iris*. Les lignes intérieures correspondent à un *box plot*.

En résumé

- ▶ Favoriser les graphiques qui reposent sur le **positionnement**
- ▶ Ne pas surcharger le graphiques
- ▶ Annoter sur la figure plutôt qu'utiliser une légende
- ▶ Soyez constant dans le style entre les figures

Références utiles :

- ▶ [Guide de visualisation de données de Material](#)

Outils

LaTeX :

- ▶ Nativement avec les packages TikZ et PGFPLOTS
- ▶ À partir de Python avec python2latex ou tikzplotlib

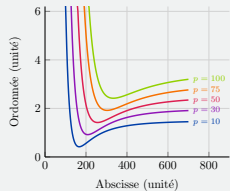
Python :

- ▶ matplotlib avec seaborn
 - ▶ PDF seulement, pas de PNG ou JPEG
 - ▶ Attention à la taille du texte
 - ▶ Les paramètres par défaut sont **mauvais**

Les couleurs

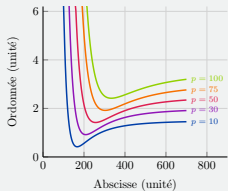
Les couleurs

Esthétique

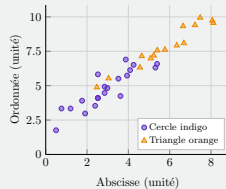


Les couleurs

Esthétique

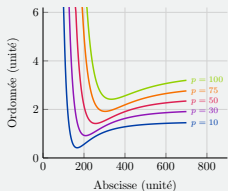


Différencier des **catégories**

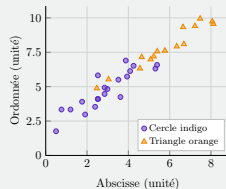


Les couleurs

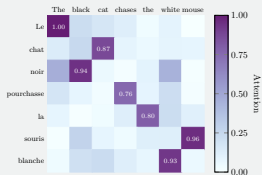
Esthétique



Différencier des **catégories**

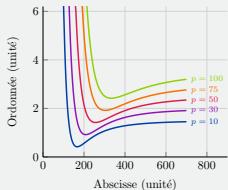


Représenter un **spectre**

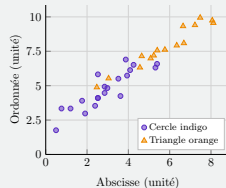


Les couleurs

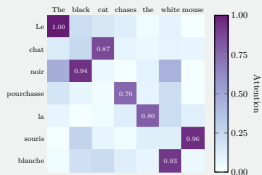
Esthétique



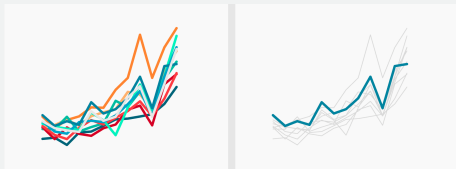
Différencier des **catégories**



Représenter un **spectre**



Mettre en **évidence**

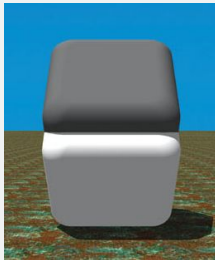


Limitations

- ▶ Difficile de comparer des couleurs semblables

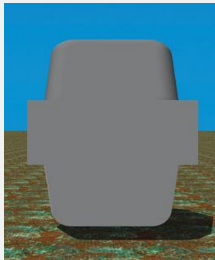
Limitations

- ▶ Difficile de comparer des couleurs semblables
- ▶ Perception affectée par les couleurs environnantes



Limitations

- ▶ Difficile de comparer des couleurs semblables
- ▶ Perception affectée par les couleurs environnantes



Limitations

- ▶ Difficile de comparer des couleurs semblables
- ▶ Perception affectée par les couleurs environnantes
- ▶ Les moniteurs et les imprimantes donnent des rendues différents

Limitations

- ▶ Difficile de comparer des couleurs semblables
- ▶ Perception affectée par les couleurs environnantes
- ▶ Les moniteurs et les imprimantes donnent des rendues différents
- ▶ Sort mal en tons de gris

Limitations

- ▶ Difficile de comparer des couleurs semblables
- ▶ Perception affectée par les couleurs environnantes
- ▶ Les moniteurs et les imprimantes donnent des rendues différents
- ▶ Sort mal en tons de gris
- ▶ Pire encore... tous les humains ne perçoivent pas les couleurs de la même manière

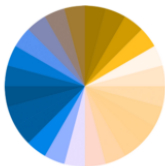


Le daltonisme

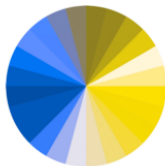
Environ un homme sur 10 est daltonien.



Regular vision



Deuteranopia



Protanopia



Tritanopia



Monochromacy



Palettes qualitatives

Utilité	Catégories
Propriétés	1 teinte par catégorie
Faiblesses	Daltonisme, tons de gris
Mitigations	Varier la luminosité, annoter

Continue



Discrète



python2latex **holi**

Palettes qualitatives

Utilité

Catégories

Propriétés

1 teinte par catégorie

Faiblesses

Daltonisme, tons de gris

Mitigations

Varier la luminosité, annoter

Continue



Discrète



python2latex holi

Deuteranopia



Protanopia



Tritanopia



Monochromacy



Palettes séquentielles

Utilité

Intensité

Propriétés

1 teinte; le pâle indique une faible intensité

Faiblesses

Difficile à analyser, affecté par les couleurs environnantes

Mitigations

Échelle perceptuellement linéaire, annoter, discrétiser

Continue



Discrète



Palettes séquentielles

Utilité	Intensité
Propriétés	1 teinte; le pâle indique une faible intensité
Faiblesses	Difficile à analyser, affecté par les couleurs environnantes
Mitigations	Échelle perceptuellement linéaire, annoter, discrétiser

Continue



Discrète



Autres exemples

viridis



magma



Palettes divergentes

Utilité	Déviation
Propriétés	2 teinte; le pâle indique le neutre
Faiblesses	Difficile à analyser, affecté par les couleurs environnantes
Mitigations	Échelle perceptuellement linéaire, annoter, discrétiser

Continue



Discrète



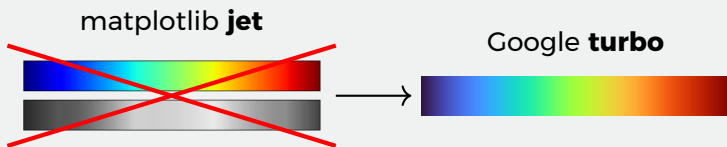
Palettes divergentes

Utilité	Déviation
Propriétés	2 teinte; le pâle indique le neutre
Faiblesses	Difficile à analyser, affecté par les couleurs environnantes
Mitigations	Échelle perceptuellement linéaire, annoter, discrétiser

Continue



Discrète



En résumé

Moyen utile, mais plus ou moins fidèle. Pour mitiger les faiblesses :

- ▶ Complémenter l'information (ex. annotations, formes)
- ▶ Palettes « colorblind-safe »
- ▶ **Luminosités** perceptuellement différentes

Outils :

- ▶ ColorBrewer pour des palettes éprouvées
- ▶ Coblis : Simulateur de daltonisme pour vérifier vos figures

Références :

- ▶ seaborn
- ▶ Google turbo

Bonne session !