

La comparaison décide de la marque

Quatre instruments de sécurité alimentaire sur les mêmes 1 955 ménages donnent 7,4 %, 42,9 %, 48,2 % et 51,0 %. Écrit en phrase, un lecteur acquiesce ; dessiné en quatre barres, personne ne passe à côté. Le type de graphique a été décidé par la comparaison, non par préférence.

Pierre Robentz Cassion

Leçon 1 sur 8

Visualisation de données pour les programmes — La marque et la comparaison

Ce que couvre cette leçon

- Partez de la phrase, non de la bibliothèque
- Le graphique des quatre instruments
- Le graphique qui est un tableau
- Où quelque chose s'arrête
- Trois marques à éviter, et par quoi les remplacer
- La barre empilée, qui est parfois juste
- Rapportez-le en entier
- La suite

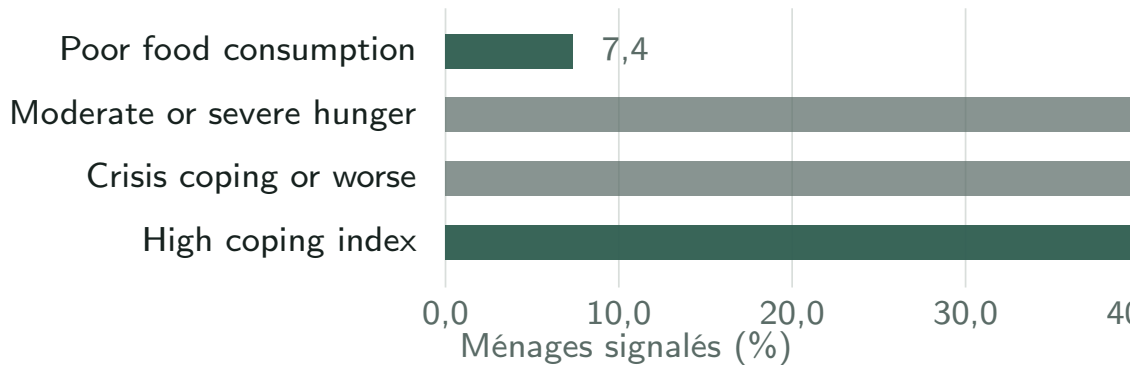
Partez de la phrase, non de la bibliothèque

Ce que la phrase compare	Marque	Pourquoi
Des catégories entre elles	Barres horizontales triées	La longueur est l'encodage le plus fidèlement lu, et les étiquettes tiennent
Une quantité dans le temps	Ligne	La pente est la comparaison ; un diagramme en barres des mois la cache
Une partie contre son tout	Une seule barre, ou le nombre	Un camembert, c'est deux nombres et une légende
Deux variables continues	Nuage de points	La position sur les deux axes
La même comparaison selon les lieux	Petits multiples	Une forme répétée, lue une fois
Où quelque chose s'arrête	Étapes ordonnées	La chute entre les barres est le constat

Partez de la phrase, non de la bibliothèque

- Rien sur cette plateforme n'a encore eu besoin d'une marque hors de cette liste — ce qui est un fait sur les rapports. . .

Le graphique des quatre instruments



Quatre instruments de sécurité alimentaire appliqués aux mêmes 1 955 ménages. La prévalence la plus haute vaut sept fois la plus basse, si bien que l'instrument choisi décide de la charge avant toute analyse.

Le graphique des quatre instruments — En Python

```
import pandas as pd
import matplotlib.pyplot as plt

prevalence = pd.Series({
    "Poor food consumption": 0.074,
    "Moderate or severe hunger": 0.429,
    "Crisis coping or worse": 0.482,
    "High coping index": 0.510,
})

fig, ax = plt.subplots(figsize=(7, 2.6))
ax.barh(prevalence.index, prevalence.values)
ax.invert_yaxis()           # first row at the top
ax.set_xlabel("Households flagged (%")
```

Le graphique des quatre instruments — En R

```
library(ggplot2)

ggplot(prevalence, aes(x = share, y = reorder(instrument, share))) +
  geom_col() +
  labs(x = "Households flagged (%)", y = NULL)
```

Le graphique des quatre instruments

- Horizontales, parce que les étiquettes sont des mots — « Faim modérée ou sévère » pivotée à 45 degrés sous une barre. . .
- Triées, sauf si l'ordre signifie quelque chose — L'ordre alphabétique est un ordre que personne n'a demandé
- Une exception à connaître — Quand les catégories ont leur propre ordre — phases IPC, bandes de PB, mois, niveaux. . .

Le graphique qui est un tableau — En Python

```
coverage = pd.Series({"Measles first dose": 0.87, "Measles second dose": 0.61})
```

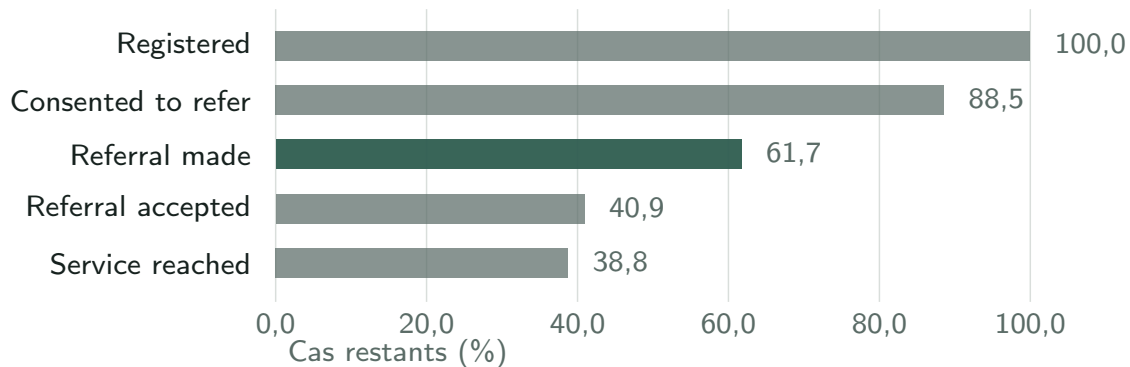
Le graphique qui est un tableau — En R

Two numbers. A chart of two numbers is a chart of nothing.

Le graphique qui est un tableau

- Deux ou trois nombres ont leur place dans une phrase — « La couverture passe de 87 % à 61 % entre la première et la... »
- Un graphique gagne sa place vers cinq catégories et cesse d'être lu vers vingt-cinq — En dessous de cinq, écrivez la... »

Où quelque chose s'arrête



Orientations de protection survivant à chaque étape, en part des 1 850 cas enregistrés. La perte la plus forte se situe entre le consentement et l'émission de l'orientation, décision du travailleur social et non défaillance du service

Où quelque chose s'arrête — En Python

```
stages = pd.Series({
    "Registered": 1.000, "Consented to refer": 0.885, "Referral made": 0.617,
    "Referral accepted": 0.409, "Service reached": 0.388,
})
losses = -stages.diff().dropna()
print(losses.round(3))
print(f"largest drop: {losses.idxmax()} ({losses.max():.1%})")
```

Où quelque chose s'arrête — En R

The drop is the finding; the bars are how you show it.

Où quelque chose s'arrête

- La plus forte chute vaut 26,8 points, entre le consentement et l'émission de l'orientation — Ni au service, ni à...
- Barres dans l'ordre des étapes, jamais triées — La séquence est le sens

Trois marques à éviter, et par quoi les remplacer

- Le camembert — Les angles se lisent moins fidèlement que les longueurs, et un camembert de plus de trois parts a besoin. . .
- Le graphique à double axe — Deux séries sur deux échelles, et le point de croisement est décidé par l'endroit où vous. . .
- Le quoi que ce soit en 3D — La profondeur n'ajoute aucune donnée et déforme chaque longueur qu'elle touche

Trois marques à éviter, et par quoi les remplacer — En Python

The one line of matplotlib configuration this course insists on.

```
plt.rcParams.update({"axes.spines.top": False, "axes.spines.right": False})
```

Trois marques à éviter, et par quoi les remplacer — En R

```
theme_set(theme_minimal(base_size = 11))
```

La barre empilée, qui est parfois juste

- Les barres empilées sont lisibles pour le segment du bas et illisibles au-dessus — Chaque segment sauf le premier part. . .
- Empilez quand le total compte et que la composition est secondaire — une charge de cas ventilée par catégorie, où la. . .
- N'empilez jamais des pourcentages qui ne totalisent pas 100 — Les quatre instruments ci-dessus se recouvrent : un. . .

Rapportez-le en entier — Exemple

Food security instruments, 1,955 households, 2024

Four instruments were applied to the same households and are shown as a sorted horizontal bar chart. The bars are prevalences, not components of a total: households can be flagged by more than one instrument, so the bars do not sum and are not stacked.

Poor food consumption	7.4%
Moderate or severe hunger	42.9%
Crisis coping or worse	48.2%
High coping index	51.0%

The chart's claim is the spread, not any single bar. 76.3% of households are flagged by at least one instrument and 3.4% by all four.

- Le second paragraphe est celui qui empêche de mal lire le graphique — et il fait deux phrases

- Chaque barre de cette leçon est une proportion estimée sur un échantillon, et aucune ne porte d'intervalle.

Lire la leçon complète, avec le code exécutable

`https://data-analysis.cassion.dev/fr/cours/data-visualisation/
viz-01-the-comparison-decides-the-mark`