

Un gain de sept points qui était l'année scolaire

La littératie a monté de 7,0 points entre le passage initial et le final, et une demande de financement l'appelle l'effet de la cantine scolaire. Les écoles sans cantine ont monté de 7,6 points. Tout le gain, et un peu plus, appartient à l'année et non aux repas.

Pierre Robentz Cassion

Leçon 1 sur 8

Méthodes d'évaluation d'impact — Le contrefactuel

Ce que couvre cette leçon

- Une question, trois réponses
- Le contrefactuel, énoncé en une phrase
- Pourquoi le nombre avant-après est si convaincant
- Trois choses que contient une évolution avant-après
- Quand une comparaison avant-après est légitime
- Rapportez-le en entier
- La suite

Une question, trois réponses — En Python

```
import pandas as pd
import numpy as np

assessment = pd.read_csv("learning-assessment-2024.v1.csv")
assessment["pct"] = assessment["raw_score"] / assessment["items"]
literacy = (assessment[assessment["domain"] == "literacy"]
            .pivot_table(index="student_id", columns="assessment_round",
                          values="pct").dropna())

roster = pd.read_csv("school-roster-2024.v1.csv").drop_duplicates(
    "student_id", keep="last")
d = literacy.join(roster.set_index("student_id"), how="inner")

print(d.groupby("feeding_programme")[["baseline", "endline"]].mean().round(4))
print(f"overall gain: {(d['endline'] - d['baseline']).mean():+.4f}")
```

Une question, trois réponses — En R

```
library(dplyr)
```

```
d |> summarise(baseline = mean(baseline), endpoint = mean(endpoint),  
              .by = feeding_programme)
```

Une question, trois réponses

	Initial	Final	Gain	n
Avec cantine	57,75 %	64,41 %	+6,66 pts	398
Sans cantine	54,67 %	62,29 %	+7,62 pts	187
Tous les élèves	56,76 %	63,73 %	+6,97 pts	585

Une question, trois réponses

- « Le programme a produit un gain de 7,0 points. » — Avant-après, apparié, IC à 95 % +6,1 à +7,9, $t = 15,2$
- « Les écoles avec cantine ont 2,1 points d'avance. » — Coupe transversale finale, IC à 95 % -1,2 à +5,4
- « Le gain a été inférieur de 0,96 point dans les écoles avec cantine. » — Chaque groupe contre son propre point de...

Le contrefactuel, énoncé en une phrase

La littératie dans les écoles avec cantine a monté de 6,66 points **par rapport à ce qui se serait passé sans le programme.**

Le contrefactuel, énoncé en une phrase

- Cette seconde proposition est le contrefactuel, et il n'est jamais observé — Les mêmes enfants, la même année, sans. . .

Le contrefactuel, énoncé en une phrase

Affirmation	Contrefactuel supposé	Est-ce plausible ?
Avant-après	Les mêmes enfants n'auraient pas changé du tout	Non — une année scolaire a eu lieu
Coupe transversale finale	Les deux groupes auraient été identiques	Non — ils différaient de 3,1 points au départ
Différences de différences	Les deux groupes auraient changé d'autant	Discutable, et la leçon 3 est cette discussion

Le contrefactuel, énoncé en une phrase

- Nommer le contrefactuel est toute la méthode — Une fois écrit, l'affirmation avant-après s'effondre sans la moindre. . .

Pourquoi le nombre avant-après est si convaincant — En Python

```
gain = d["endline"] - d["baseline"]
se = gain.std(ddof=1) / np.sqrt(len(gain))
print(f"{gain.mean():+.4f}  95% CI [{gain.mean() - 1.96*se:+.4f}, "
      f" {gain.mean() + 1.96*se:+.4f}]  t = {gain.mean()/se:.1f}")
```

Pourquoi le nombre avant-après est si convaincant — En R

```
t.test(d$endline - d$baseline)
```

Pourquoi le nombre avant-après est si convaincant

- $t = 15,2$ — Toutes les barrières du cours de statistiques passent
- Aucune de ces vérifications ne peut détecter un groupe de comparaison manquant — et c'est ce qu'il faut emporter de. . .

Trois choses que contient une évolution avant-après

- Le passage du temps — Les enfants grandissent et sont enseignés
- La régression vers la moyenne — Ceux qui ont été sélectionnés parce qu'ils étaient les plus mal lotis paraîtront mieux. . .
- Tout le reste de ce qui est arrivé — Une sécheresse, un mouvement de change, le programme d'une autre agence dans le. . .

Trois choses que contient une évolution avant-après — En Python

```
by_school = d.groupby(["school_id", "feeding_programme"])[  
    ["baseline", "endline"]].mean()  
by_school["gain"] = by_school["endline"] - by_school["baseline"]  
print(by_school.groupby("feeding_programme")["gain"].describe()[  
    ["mean", "min", "max"]].round(4))
```

Trois choses que contient une évolution avant-après — En R

```
by_school |> summarise(mean = mean(gain), min = min(gain), max = max(gain),  
                        .by = feeding_programme)
```

Trois choses que contient une évolution avant-après

- Chacune des 24 écoles a progressé — de +1,2 point à +12,2

Quand une comparaison avant-après est légitime

- Quand le contrefactuel est réellement plat et que vous pouvez le montrer — Trois passages ou plus avant le programme. . .
- Quand le résultat ne peut pas changer tout seul — Une latrine ne se construit pas elle-même
- Quand vous faites du suivi, non de l'évaluation — « La présence est passée de 84 % à 88 % » est une affirmation vraie. . .
- La formule à surveiller est « grâce à » — C'est là qu'un nombre de suivi devient une affirmation d'impact, et elle est. . .

Rapportez-le en entier — Exemple

Literacy outcomes, baseline to endline 2024

All students	56.8% to 63.7%	+6.97 points	95% CI +6.1 to +7.9
With feeding	57.8% to 64.4%	+6.66 points	
Without feeding	54.7% to 62.3%	+7.62 points	

The overall change is reported as a monitoring result and is not attributed to the feeding programme. Schools without the programme gained slightly more, so the difference-in-differences estimate of the programme effect is -0.96 points (95% CI -3.5 to +1.6, clustered on 24 schools).

The before-after change assumes children would have learned nothing over a school year. That counterfactual is not defensible here and the comparison group is what replaces it.

- Le dernier paragraphe fait trois lignes et c'est la différence entre un rapport de suivi et une affirmation d'impact —...

- L'estimation en différences de différences ci-dessus repose sur le fait que les deux groupes d'écoles sont comparables sur ce qui compte.

Lire la leçon complète, avec le code exécutable

```
https://data-analysis.cassion.dev/fr/cours/impact-evaluation-methods/  
ie-01-the-gain-that-was-the-school-year
```