

Évaluer l'effet d'une mesure d'aide par méthode de pairage

Webinaire du 3 novembre 2022

Guillaume Loignon

Professeur, département d'éducation et pédagogie, UQAM

Cinq postulats de base pour l'analyse de données d'éducation

1. Chaque décision organisationnelle entraîne un *coût d'opportunité* (Heath, 2014).
2. L'analyse de données d'éducation n'est pas *athéorique* (Mandinach & Schildkamp, 2021).
3. Les concepteurs d'IA tendent à négliger l'expertise de terrain (Sambasivan & Veeraraghavan, 2022).
4. L'IA ne remplace pas le jugement (Mitchell et al., 2019).
5. On peut critiquer sans être spécialiste (Mitchell et al., 2019)

Données de la
recherche
(*evidence*)

**Analyse de
données
d'éducation**

Expertise
locale /
connaissance
du terrain

Données de
l'établissement

La mesure d'aide

Cours de français SP au cégep de Saint-Jérôme, depuis H17

- Semaines 2 à 7: rédaction & réécriture (3 fois).
- Semaines 8 à 15: étude d'œuvres littéraires.

Activités de mise à niveau et
activités favorisant la réussite

Établissements d'enseignement collégial francophones

Enseignement collégial

Question de recherche

Le parcours SP améliore-t-il la persévérance / la diplomation ?

- Relation entre le parcours (SP, régulier, renforcement) et le statut (études en cours, désistement, diplomation)

Trois parcours possibles

SP	Inscription en français 1 dans un groupe SP, pas d'inscription préalable en renforcement.	n = 498
Régulier	Inscription en français 1 régulier (<i>Communication écrite et orale</i>); pas d'inscription préalable en renforcement français.	n = 5153
Renforcement	Inscription en mise à niveau français.	n = 1774

Trois statuts des études

En cours	Inscription confirmée à l'hiver 2022
Désistement	Pas de diplôme obtenu, pas d'inscription ou inscription non-confirmée à l'hiver 2022
Diplomation	Diplôme obtenu au plus tard à l'hiver 2022

Comparaison des proportions

Parcours en français au cégep	Moyenne au secondaire (écart-type)	Statut des études		
		En cours	Désistement	Diplomation
SP	76,1 (8,1)	164 (33%)	120 (24%)	214 (43%)
Régulier	78,3 (7,7)	1242 (24%)	2413 (47%)	1498 (29%)
Renforcement	66,5 (6,1)	304 (17%)	1260 (71%)	210 (12%)
<i>Tous les parcours</i>	<i>74,0 (8,7)</i>	<i>1710 (23%)</i>	<i>3793 (51%)</i>	<i>1922 (26%)</i>

$\chi^2(4) = 492, p < 0,001.$

Méthode de pairage

Méthode des plus proches voisins
(apprentissage machine non-supervisé)

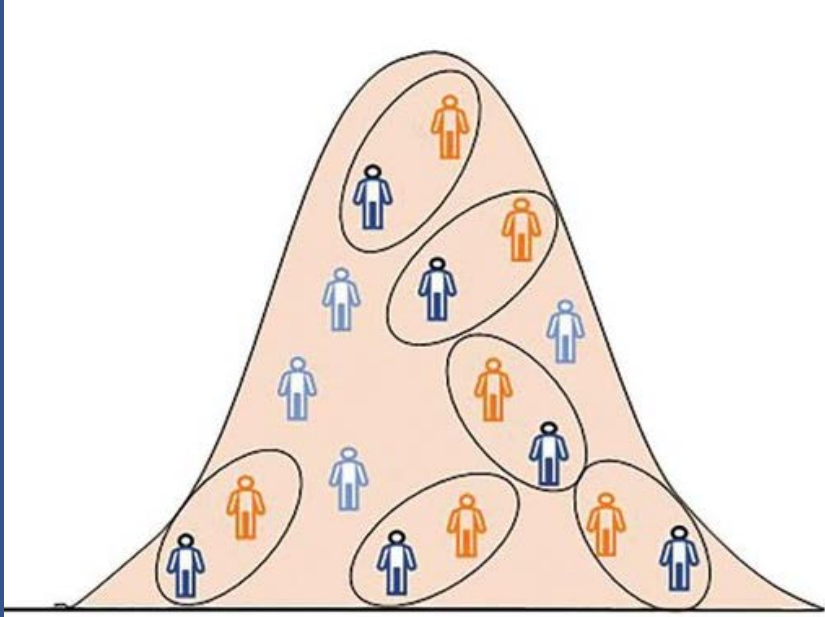


Figure adaptée depuis Chu et al. (2022)

Pairage par: **cohorte**, moy. au
secondaire, programme
d'études, statut Saide, tour
d'admission, genre, langue
maternelle

Méthode de pairage – version geek



Journal of Statistical Software

June 2011, Volume 42, Issue 8.

<http://www.jstatsoft.org/>

MatchIt: Nonparametric Preprocessing for Parametric Causal Inference

Daniel E. Ho
Stanford Law School

Kosuke Imai
Princeton University

Gary King
Harvard University

Elizabeth A. Stuart
Johns Hopkins University

Abstract

MatchIt implements the suggestions of Ho, Imai, King, and Stuart (2007) for improving parametric statistical models by preprocessing data with nonparametric matching methods. **MatchIt** implements a wide range of sophisticated matching methods, making it possible to greatly reduce the dependence of causal inferences on hard-to-justify, but commonly made, statistical modeling assumptions. The software also easily fits into existing research practices since, after preprocessing data with **MatchIt**, researchers can use whatever parametric model they would have used without **MatchIt**, but produce inferences with substantially more robustness and less sensitivity to modeling assumptions. **MatchIt** is an R program, and also works seamlessly with **Zelig**.

Keywords: matching methods, causal inference, balance, preprocessing, R.

Méthode de pairage – version *geek*

```
balance_fr ←  
  MatchIt::matchit(  
    estP4TSP ~ MGS + NoProgAdmission + Saide + NoTour + Sexe + LangueMaternelle,  
    exact = ~ AnSessionDebut, # variable "non-négociable", ici la cohorte  
    data = df_match,  
    method = "nearest", # méthode du plus proche voisin  
    distance = "lasso" # scores de propension estimés par LASSO  
  )
```

Pour chaque élève SP, l'algorithme:

1. Calcule un *score de propension* avec les pairs non-SP
2. Forme une paire (*match*) avec l'élève non-SP ayant le score de propension le plus élevé

Application d'une technique de *bootstrap*

Pour obtenir des intervalles de confiance:

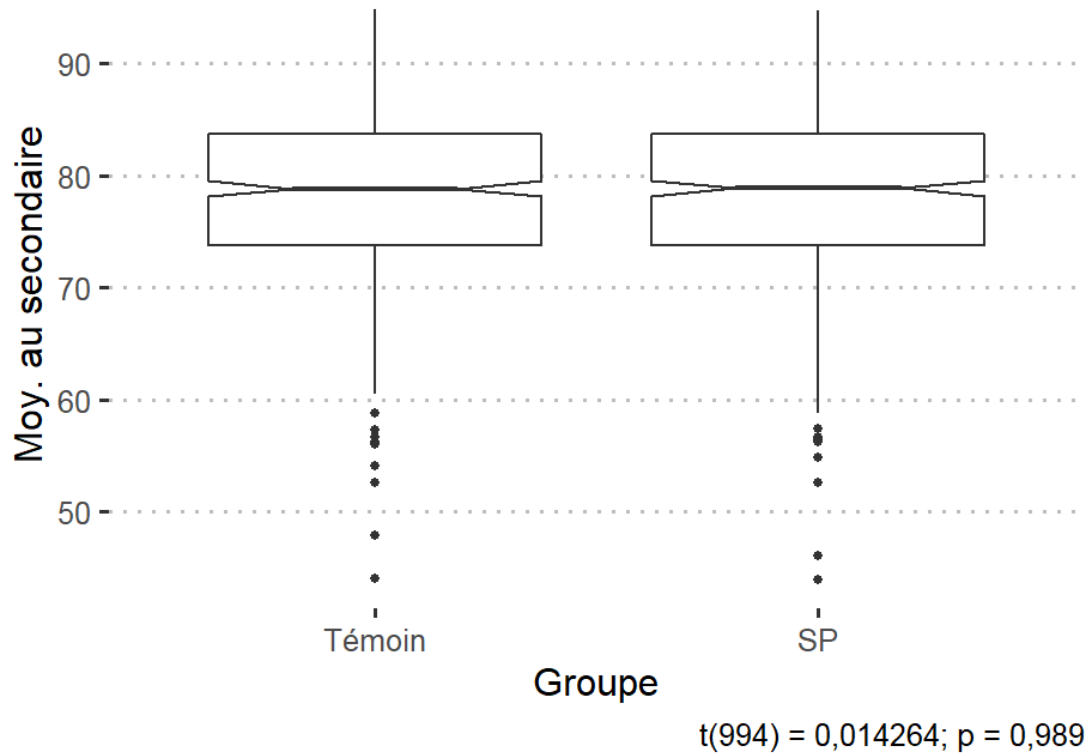
- 1) Tirer N élèves SP, avec remise
- 2) Païrer et calculer les statistiques
- 3) Répéter B fois
- 4) Rapporter la médiane avec les percentiles 2,5 et 97,5.

Une technique souvent utilisée en *data science* !

```
for (i in 1:B) {  
  # B = nombre de répétitions  
  set.seed(i) # pour s'assurer que le résultat varie au fil des répétitions  
  cur_df ← # tirage de N observations, avec remise (bootstrap)  
    df_match %>% slice_sample(n = nrow(df_match), replace = T)  
  cur_result ← boot_matchit(cur_df) # pairage  
  result_boot_fr[i] ← cur_result  
}
```

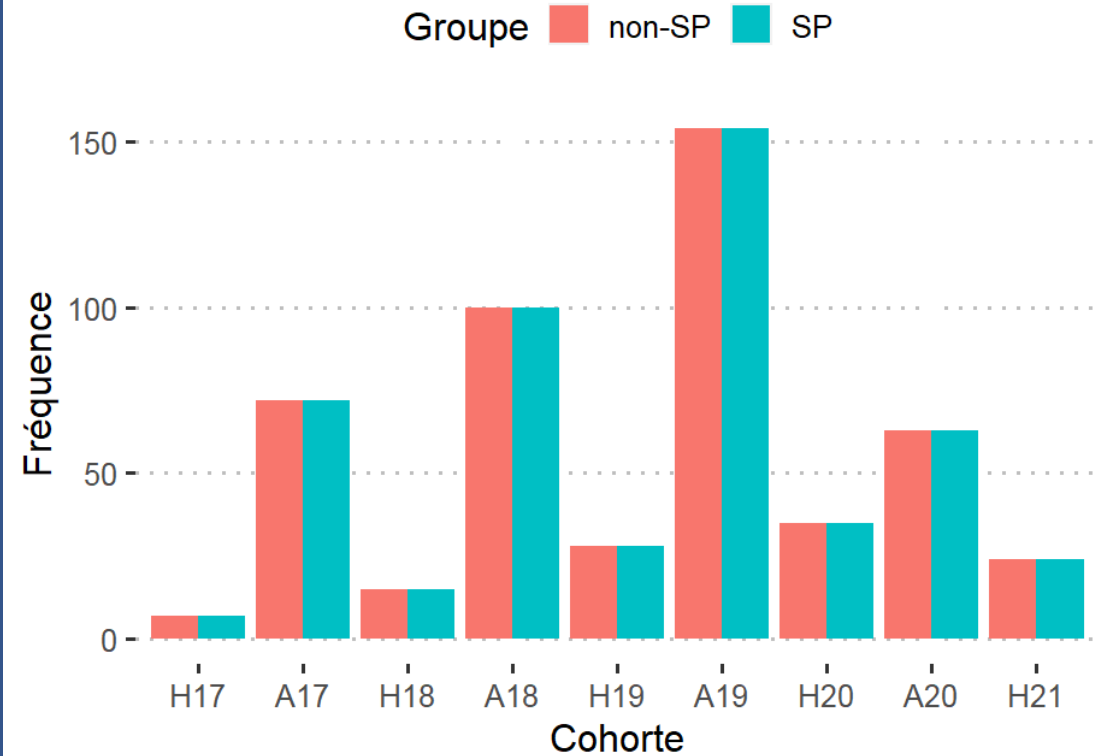
Validation du pairage

Comparaison de la MGS par groupe



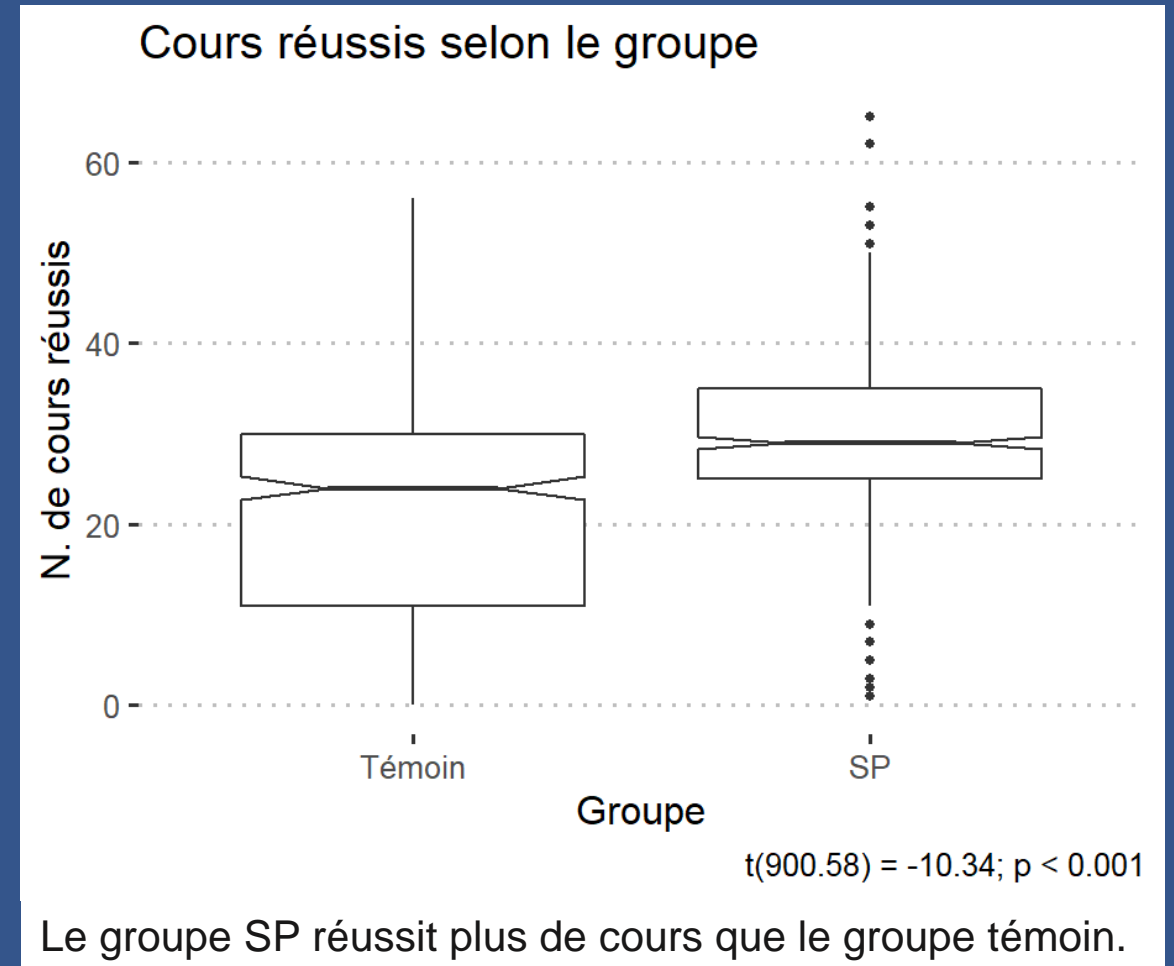
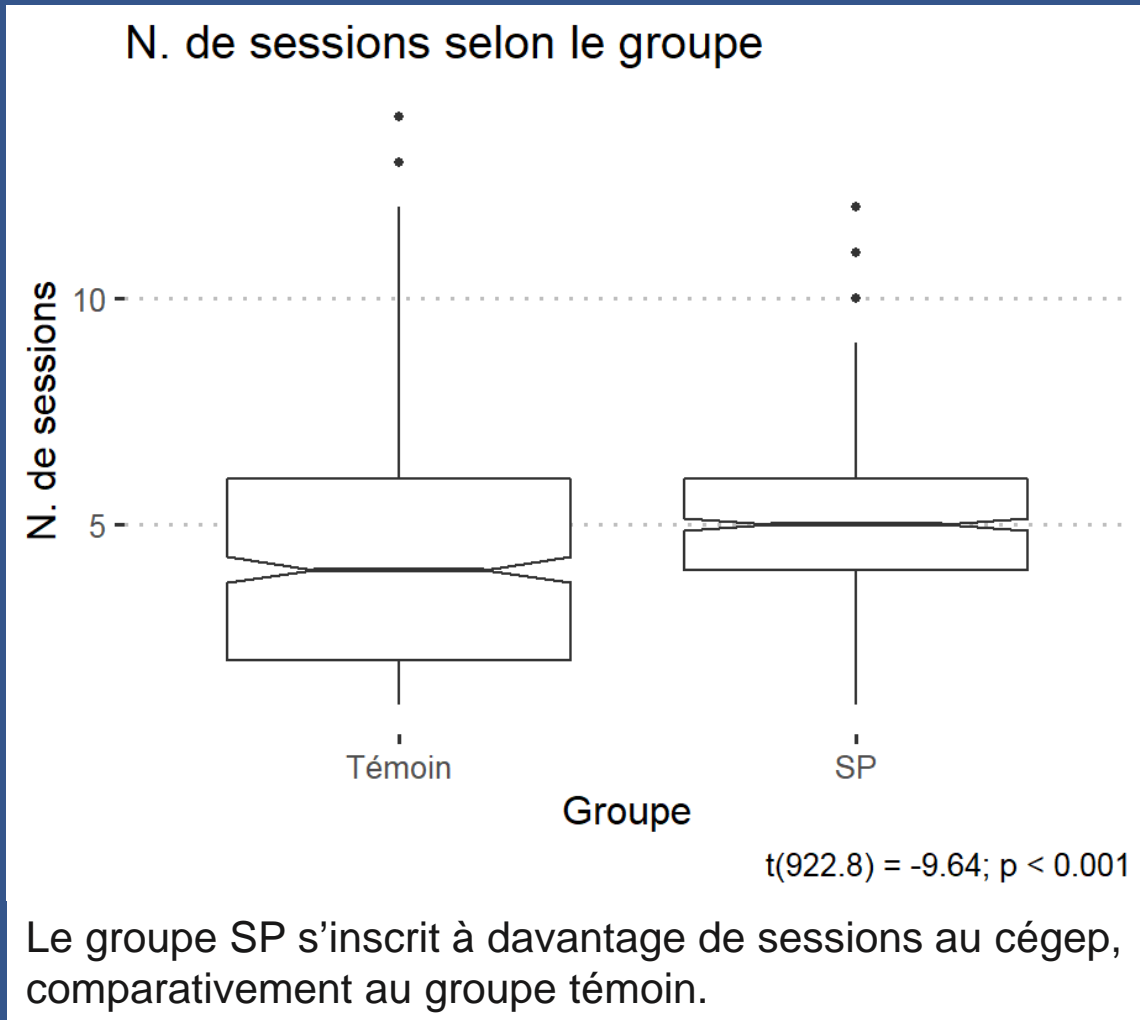
Le groupe témoin et les élèves SP ont une moyenne au secondaire équivalente.

Effectifs par cohorte et par groupe



Nombre d'élèves équivalent par cohorte

Résultats – effet SP



Résultats – effet SP

Groupe	Statut des études, %		
	En cours	Désistement	Diplomation
SP	32,8 [29.6, 36.55]	23.8 [19.95, 28.56]	43.3 [38.62, 47.3]
Témoin	22.9 [18.98, 26.56]	42.3 [36.69, 46.31]	34.9 [30.54, 39]
<i>Différence</i>	<i>14.9 [5.2, 10.3]</i>	<i>-18.4 [-13.6, -23,4]</i>	<i>8.4 [2.8, 13.3]</i>

Note. Différences entre les groupes vérifiées par méthode de *bootstrap* et statistiquement significatives à un seuil de $p < 0,001$. Intervalles de confiance à 95% entre crochets.

Conclusions

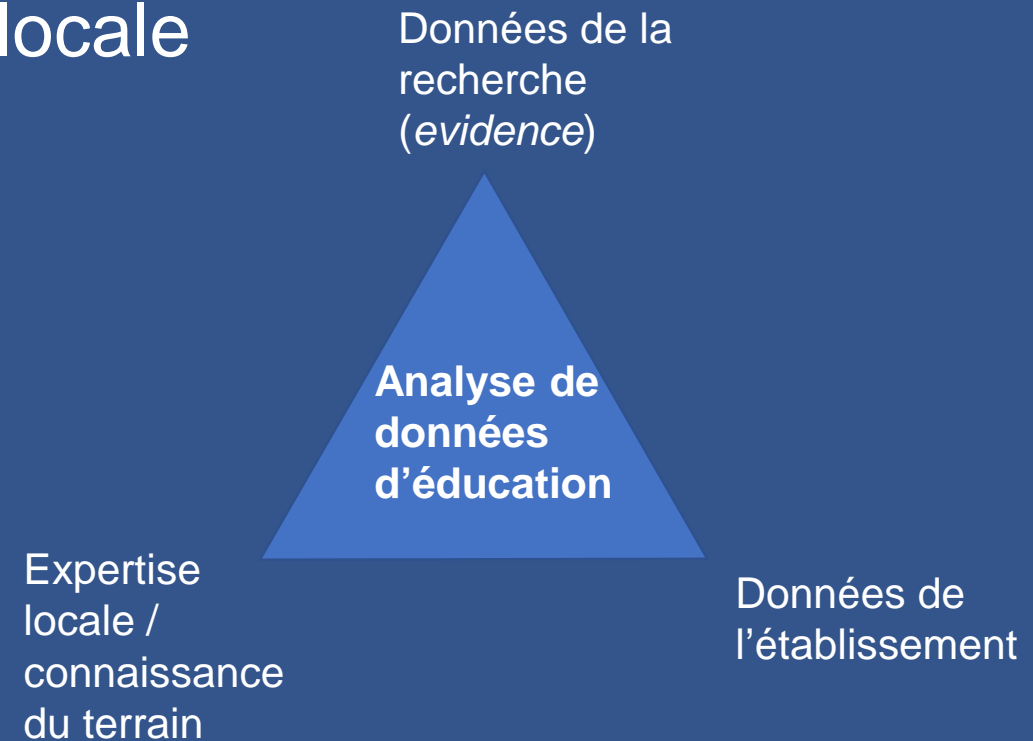
- Les élèves du parcours SP demeurent plus longtemps au collège et, en moyenne, réussissent davantage de cours.
- En contrôlant pour les principaux facteurs, les élèves ayant suivi le cours SP se désistent moins; la différence est d'environ 18 points de pourcentage (entre 14 et 23 points, 19 fois sur 20).
- De plus, les élèves SP obtiennent davantage le diplôme que des élèves équivalents n'ayant pas fait le parcours SP; la différence est d'environ 8 points de pourcentage (entre 3 et 13 points, 19 fois sur 20).

C'est dans les plans...

- Créer une infrastructure d'analyse de données pour simplifier l'application de ce type de méthodes
- Recherche partenariale
- Développer et valoriser l'expertise locale

Questions ?

loignon.guillaume@uqam.ca



Fiche sommaire

Détails des analyses	<ul style="list-style-type: none">• Développé par: G. Loignon, 2022• Données: cohortes de H2017 à A2021 du cégep de Saint-Jérôme, plus inscriptions et diplomation pour H2022• Type d'analyse: constitution <i>post hoc</i> d'un groupe témoin• Méthode: pairage par plus proche voisin, implémentation dans MatchIt pour R, scores de propension calculés avec technique de LASSO
Utilisation	<ul style="list-style-type: none">• But: Vérifier l'effet d'une mesure d'aide• Connaissances produites pour: direction, CP, profs de français
Variables	<ul style="list-style-type: none">• Dépendante: statut des études (en cours, désistement, diplôme)• Indépendante: parcours en français (SP, régulier, renforcement)• Confondantes (utilisées pour l'appariement): cohorte, MGS, programme d'études, statut Saide, genre, tour d'inscription, langue maternelle
Indicateurs de performance	<ul style="list-style-type: none">• Validation de l'appariement avec test <i>t</i> et <i>boxplots</i>• Stabilité vérifiée avec méthode de <i>bootstrap</i>
Analyses quantitatives	Statistiques sur le statut des études comparant les élèves SP aux élèves non-SP; on rapporte médiane et IC 95% produits par <i>bootstrap</i> .
Données d'entraînement	Ne s'applique pas (méthode d'apprentissage non-supervisé)
Considérations éthiques	<ul style="list-style-type: none">• Données anonymisées par la technicienne (identifiant encrypté)• Pas de réidentification possible à partir des résultats diffusés

Fiche adaptée depuis Mitchell et al. (2019)

Bibliographie

- Chu, J. J., Shamsunder, M. G., Yin, S., Rubenstein, R. R., Slutsky, H., Fischer, J. P., & Nelson, J. A. (2022). Propensity scoring in plastic surgery research: an analysis and best practice guide. *Plastic and Reconstructive Surgery Global Open*, 10(2).
- Heath, J. 2014. Cost-Benefit Analysis as an Expression of Liberal Neutrality. Dans *The Machinery of Government*.
https://www.academia.edu/9811886/Cost_Benefit_Analysis_as_an_Expression_of_Liberal_Neutralit
[Y](#)
- Ho, D., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1–28.
<https://doi.org/10.18637/jss.v042.i08>
- Mandinach, E. B., & Schildkamp, K. 2021. Misconceptions about data-based decision making in education : An exploration of the literature. *Studies in Educational Evaluation*, 69, 100842.
<https://doi.org/10.1016/j.stueduc.2020.100842>
- Mitchell, M., Wu, S., Zaldivar, A., Barnes, P., Vasserman, L., Hutchinson, B., Spitzer, E., Raji, I. D., & Gebru, T. 2019. Model Cards for Model Reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*, 220-229. <https://doi.org/10.1145/3287560.3287596>
- Romele, A. 2022. Images of Artificial Intelligence : A Blind Spot in AI Ethics. *Philosophy & Technology*, 35(1), 4. <https://doi.org/10.1007/s13347-022-00498-3>
- Sambasivan, N., & Veeraraghavan, R. 2022. The Deskillling of Domain Expertise in AI Development. *CHI Conference on Human Factors in Computing Systems*, 1-14.
<https://doi.org/10.1145/3491102.3517578>
- Von Davier, A. A., Wong, P. C., Polyak, S., & Yudelson, M. 2019, July. The argument for a Data Cube for large-scale psychometric data. In *Frontiers in Education*, 471.

