

Corrigé des exercices de fin de chapitre

Comprendre les données sociales

Mark Paradis, Ph. D.

19 août 2026

Voici le corrigé détaillé des exercices de fin de chapitre de *Comprendre les données sociales*. Le corrigé des exercices situés à l'intérieur de chaque section se trouve dans le livre lui-même, immédiatement après les exercices auxquels il répond.

Faites l'exercice avant de lire le corrigé. Un corrigé lu en premier n'est qu'un texte de plus ; un corrigé lu en second est une vérification de votre raisonnement, et seul le second enseigne quelque chose.

Annexe A

Corrigé des exercices de fin de chapitre

Ce corrigé accompagne les exercices de fin de chapitre. Le corrigé des exercices situés à l'intérieur de chaque section se trouve dans le livre, immédiatement après les exercices concernés.

Une remarque sur la façon de s'en servir. Les exercices de fin de chapitre demandent presque toujours plus qu'un calcul : ils demandent de choisir une méthode, de justifier ce choix, puis d'interpréter le résultat. Le corrigé montre les trois. Si votre réponse numérique concorde mais que votre justification diffère, c'est la justification qu'il vaut la peine de relire.

Chapitre 1 Introduction aux méthodes quantitatives

Corrigé de l'exercice 1.E1. (a) Le mot qui fait des dégâts, c'est *prouve*. Une étude montrant que les adolescents qui utilisent davantage les médias sociaux déclarent plus de symptômes dépressifs a établi une association. Trois autres explications restent ouvertes : la flèche causale peut aller dans l'autre sens, des adolescents déprimés se repliant sur leur téléphone ; un troisième facteur — perturbation du sommeil, isolement social, situation familiale — peut produire les deux ; ou l'association peut être un artefact de la composition de l'échantillon. Le chapitre 3 traite cela comme il se doit.

(b) Appuyer une affirmation causale exige un devis, et non davantage de données. Une étude longitudinale mesurant les deux variables chez les mêmes personnes sur plusieurs années pourrait au moins établir ce qui est venu en premier. Une expérience assignant au hasard une partie des personnes participantes à un usage réduit serait plus forte encore, puisque la randomisation rend les groupes comparables sur tout, mesuré ou non. Les deux exigent des instruments validés plutôt qu'un seul énoncé

autodéclaré, et les deux exigent un échantillon qui représente une population définie.

(c) Plusieurs enjeux éthiques se présentent. La recherche auprès de personnes mineures exige le consentement parental en plus de l'assentiment de la jeune personne elle-même. Interroger sur des symptômes dépressifs peut causer de la détresse et crée l'obligation de fournir de l'information sur les ressources d'aide. Une expérience restreignant l'accès aux médias sociaux intervient dans la vie sociale des personnes participantes, ce qui demande une justification proportionnée à la connaissance acquise.

Remarquez que (b) et (c) tirent en sens contraire : le devis qui appuierait le mieux l'affirmation causale est aussi celui qui soulève les questions éthiques les plus sérieuses. Cette tension est réelle et reviendra tout au long du livre.

Corrigé de l'exercice 1.E2. *Devis quantitatif.* Question : la fréquence de consommation de contenu politique sur les médias sociaux prédit-elle le vote et d'autres formes de participation chez les 18 à 29 ans ? Données : une enquête auprès d'un échantillon représentatif de ce groupe d'âge, mesurant l'usage des plateformes, l'exposition au contenu politique et plusieurs formes de participation, avec des variables de contrôle démographiques. Force : les résultats se généralisent à la population échantillonnée, et des explications concurrentes comme la scolarité peuvent être prises en compte. Limite : l'enquête saisit ce que les gens déclarent plutôt que ce qu'ils font, et même une association forte n'établira pas la direction.

Devis qualitatif. Question : comment les jeunes adultes comprennent-ils le rapport entre leur vie en ligne et leur action politique ? Données : vingt à trente entrevues en profondeur, échantillonnées de façon raisonnée pour varier les niveaux de participation. Force : ce devis peut mettre au jour des mécanismes et des significations qu'une enquête ne peut pas anticiper — y compris des raisons que personne n'aurait pensé à inscrire dans un questionnaire. Limite : il ne peut pas établir la fréquence d'un motif.

Ce ne sont pas deux réponses concurrentes à la même question. L'enquête établit qu'un motif existe et à quel point il est répandu ; les entrevues expliquent ce qui s'y passe. Un devis mixte menant les entrevues en premier améliorerait l'enquête, parce qu'il lui fournirait le vocabulaire des personnes concernées pour formuler les questions.

Corrigé de l'exercice 1.E3. Réponse ouverte. Quatre critères pour vérifier votre réponse.

Les trois questions viennent-elles vraiment de disciplines différentes ? Trois variations sur le comportement électoral, c'est une discipline trois fois. Visez des questions qu'une sociologue, un économiste et une psychologue reconnaîtraient chacun comme relevant de leur domaine.

Les questions trouvent-elles réponse dans ce jeu de données ? Une question sur

l'évolution dans le temps exige un jeu de données à plusieurs vagues. Une question d'effet causal ne peut pas être tranchée par des données d'enquête transversales ; dites-le plutôt que d'en réclamer trop.

Les variables sont-elles nommées précisément ? C'est ici que parcourir le dictionnaire des variables rapporte : nommez de vraies variables, ou citez la formulation exacte des questions.

Le plan d'analyse est-il esquissé raisonnablement ? Il n'est pas nécessaire de connaître le nom de la technique. « Comparer la confiance moyenne entre les personnes ayant fait des études postsecondaires et celles qui n'en ont pas fait » est une réponse complète et juste à ce stade-ci.

Corrigé de l'exercice 1.E4. (a) Trois enjeux, au minimum.

Le consentement. Des publications écrites pour ses abonnés n'ont pas été offertes à la recherche. La visibilité publique décrit qui *peut* voir quelque chose, et non à qui cela était destiné ; et des conditions d'utilisation ne sont pas un consentement éclairé.

La sensibilité et les torts. Une information sur la santé mentale porte une charge de stigmatisation et peut avoir des conséquences sur l'emploi, l'assurance et les relations. Les personnes qui publient pendant une catastrophe sont, par définition, en détresse et moins en mesure de considérer jusqu'où leurs mots peuvent voyager.

L'identifiabilité. Une citation textuelle peut généralement être retracée au moyen d'un moteur de recherche, et une poignée d'attributs — tranche d'âge, ville, profession — réduit souvent une population à une seule personne. « Aucun nom n'a été employé » n'est pas une anonymisation.

(b) Mesures d'atténuation : ne rapporter que des tendances agrégées ; paraphraser plutôt que citer ; demander l'examen d'un comité d'éthique de la recherche même si les données sont publiques ; exclure les comptes appartenant à des personnes mineures ; et se demander si la question de recherche pourrait trouver réponse dans des données dont les personnes concernées ont consenti à l'usage.

(c) Les deux réponses se défendent ; c'est le raisonnement qui compte. Un *oui* défendable met en balance le bénéfice public d'une meilleure compréhension de la santé mentale en situation de catastrophe et le faible risque individuel lorsque seuls des agrégats sont diffusés, et il accepte les mesures d'atténuation ci-dessus comme des conditions. Un *non* défendable soutient qu'on ne peut pas se passer du consentement au seul motif qu'il est malcommode à obtenir, et que la vulnérabilité pendant une catastrophe renforce l'obligation plutôt qu'elle ne l'affaiblit. Trancher sans avoir pesé l'autre côté laisse l'exercice à moitié fait.

Corrigé de l'exercice 1.E5. L'argument contient une description juste et une conclusion fausse, et la réponse doit les séparer.

Ce qui est juste. Le postpositivisme accepte bel et bien que la mesure soit

imparfaite et que les valeurs de la personne qui fait la recherche façonnent ce qu'elle étudie. C'est une description exacte, et non une concession arrachée sous la pression.

Où l'argument échoue. La conclusion suppose que, sans certitude, il n'existe aucune base pour comparer des études. Cela ne suit pas. Des études peuvent se comparer sur des bases qui n'ont rien à voir avec la certitude : l'échantillon représente-t-il la population qu'il prétend représenter ? la mesure a-t-elle été validée ? l'analyse a-t-elle été spécifiée avant que les données ne soient vues ? le résultat se reproduit-il ? l'incertitude déclarée correspond-elle à ce que montrent les données ? Une étude dotée d'un échantillon représentatif et d'un instrument validé vaut mieux qu'une étude qui n'a ni l'un ni l'autre, et aucun recours à la certitude n'est nécessaire pour le dire.

Le point de fond. Le postpositivisme remplace la certitude individuelle par une procédure collective. L'objectivité n'est pas un état qu'une personne atteint en étant sans préjugés ; c'est ce que produisent ensemble la méthode, la transparence, la réplication et la critique par des personnes aux présupposés différents. L'argument traite l'objectivité comme une propriété des individus, ce qui est précisément le présupposé que le postpositivisme abandonne.

Être d'accord avec l'énoncé de départ reste défendable, à condition que la réponse affronte ce contre-argument plutôt que de répéter l'énoncé.

Corrigé de l'exercice 1.E6. Plusieurs réponses se défendent ; c'est la colonne des limites qui vaut la réflexion.

Tranche de revenu (7 catégories). Économie, sociologie. Limite : les tranches font perdre l'information à l'intérieur de chaque intervalle, et la tranche supérieure est ouverte, de sorte que la queue supérieure — là où loge l'inégalité — devient invisible.

Confiance envers le gouvernement (0–10). Science politique. Limite : « gouvernement » est ambigu — le gouvernement en poste, la fonction publique, ou le régime lui-même — et les personnes interrogées peuvent répondre à propos d'objets différents.

Pays de naissance. Sociologie, géographie, démographie. Limite : la variable enregistre où une personne est née et rien sur le moment de son arrivée, son statut ou son attachement ; elle sert souvent d'indicateur indirect pour des choses qu'elle ne mesure pas.

Nombre d'enfants. Démographie, économie, sociologie. Limite : la variable ne dit pas s'il s'agit des enfants nés, des enfants vivants ou des enfants présents dans le ménage — trois nombres différents.

Santé autoévaluée (5 points). Santé publique, psychologie, sociologie. Limite : la variable mesure la santé perçue, et une même condition objective est déclarée différemment selon les cultures, les âges et les attentes.

Identification partisane. Science politique. Limite : les catégories sont propres à chaque pays et instables dans le temps, ce qui rend difficiles la comparaison internationale et la comparaison de longue durée.

Le motif à dégager : chacune de ces limites porte sur ce que la mesure laisse de côté ou brouille, et non sur l'arithmétique appliquée ensuite.

Corrigé de l'exercice 1.E7. Votre réponse dépend des pays que vous avez choisis. Le raisonnement ci-dessous vaut pour n'importe lesquels des trois.

(a) Les écarts seront importants — plusieurs années dans la plupart des cas — et cela tient au devis et non à une erreur.

(b) L'ordre devrait tenir. Un pays dont l'espérance de vie réelle est plus élevée devrait avoir une espérance de vie simulée plus élevée.

(c) Oui, et c'est le but de l'exercice. Les données de ce livre reproduisent la *structure* du monde — quels pays se classent au-dessus de quels autres, et grosso modo comment les relations s'infléchissent — tout en faussant délibérément les *niveaux*.

(d) Trois raisons ; deux suffisent. D'abord, les données simulées ont une vérité connue : comme elles ont été construites plutôt qu'observées, les vraies valeurs de la population sont disponibles, et une estimation peut donc être comparée à la bonne réponse, ce qu'aucun jeu de données réel ne permet. Ensuite, elles sont stables — les chiffres réels sont révisés, ce qui ferait cesser de correspondre les exercices et leur corrigé. Enfin, les licences : reproduire de vrais jeux de données dans un livre vendu entraîne des coûts et des restrictions que les données synthétiques évitent.

Il y en a une quatrième qui mérite d'être ajoutée. Ayant vérifié la distorsion vous-même, vous avez maintenant une raison concrète de vous méfier de tout nombre de ce livre pris comme un fait sur le monde — ce qui est exactement l'habitude que la préface demande.

Corrigé de l'exercice 1.E8. Réponse ouverte. Le but est de lire de près une section méthodologique, une habitude qui vaut la peine d'être prise tôt.

(a) *Qu'a-t-on mesuré, et comment.* Repérez l'instrument lui-même — une échelle précise, une donnée administrative, une série d'énoncés d'enquête — plutôt que le concept. « Ils ont mesuré le bien-être » ne fait que répéter le nom de la variable ; « ils ont employé une échelle de cinq énoncés portant sur les deux dernières semaines » décrit la mesure.

(b) *Qui a été étudié, et comment sélectionné.* Distinguez la population de l'échantillon, et notez la méthode de recrutement. Les études sont souvent muettes là-dessus ; remarquer ce silence fait partie de la réponse.

(c) *Une chose que l'étude ne peut pas vous apprendre.* Les limites les plus utiles sont celles qui découlent du devis plutôt que du sujet : qu'une étude transversale ne peut pas établir la séquence ; qu'un échantillon tiré d'une seule ville ne peut pas

parler pour un pays ; qu'une mesure autodéclarée ne peut pas distinguer un vrai changement d'un changement dans la disposition à le déclarer.

Une section méthodologique est souvent plus ardue à lire que les résultats, et contient souvent moins de détails qu'on ne s'y attend. Les deux constats méritent d'être faits.

Corrigé de l'exercice 1.E9. 1. *La confiance envers les gouvernements nationaux a-t-elle décliné dans les démocraties riches depuis trois décennies ?* (a) Quantitative. La question porte sur une ampleur et sur une évolution dans le temps, pour de nombreux cas — soit exactement ce à quoi servent les enquêtes à vagues répétées. (b) L'unité d'analyse est le couple pays-année, même si les données sous-jacentes sont recueillies auprès d'individus. (c) L'obstacle est la comparabilité : il faut que la question ait voulu dire la même chose dans chaque pays et à chaque décennie, et la traduction, l'évolution du vocabulaire politique et le changement des modes d'enquête menacent tous cette exigence.

2. *Pourquoi certaines personnes qui se méfient de leur gouvernement continuent-elles néanmoins de voter ?* (a) Qualitative, ou mixte à noyau qualitatif. La question porte sur un raisonnement et sur un sens, et l'éventail des réponses ne peut pas être anticipé assez bien pour rédiger des questions fermées. (b) L'unité est l'individu. (c) L'obstacle est que les gens ne sont pas toujours des narrateurs fiables de leurs propres motifs, et qu'une justification formulée après coup se distingue mal d'un raisonnement réel.

3. *Une campagne d'information publique augmente-t-elle la confiance envers une autorité sanitaire locale ?* (a) Quantitative, et expérimentale si possible — « X augmente-t-il Y » est une question causale, et la randomisation en est la meilleure réponse disponible. (b) L'unité est l'individu, ou peut-être le quartier si la campagne est déployée géographiquement. (c) L'obstacle est la contamination : les personnes du groupe témoin peuvent voir la campagne malgré tout, et la confiance peut bouger pour des raisons sans rapport avec elle.

Le motif à dégager, c'est que la forme de la question, et non le sujet, détermine l'approche. Les trois portent sur la confiance envers le gouvernement ; elles appellent trois devis différents.

Chapitre 2 Concepts statistiques fondamentaux

Corrigé de l'exercice 2.E1. (a) Deux problèmes, et les deux sont des problèmes de vocabulaire.

La médiane n'est pas la moyenne. Dans l'usage courant, « moyenne » évoque la moyenne arithmétique, et dans le cas du revenu les deux diffèrent nettement parce que la distribution est asymétrique à droite — une longue queue de hauts revenus tire

la moyenne au-dessus de la médiane. Rapporter une médiane comme une moyenne invite le lecteur à imaginer une distribution symétrique qui n'existe pas.

« Gagner » n'est pas « revenu total ». Le chiffre du recensement porte sur le revenu total, qui comprend les transferts gouvernementaux, les pensions et les revenus de placement. Les gains n'en sont qu'une partie et, pour bien des gens — personnes retraitées, étudiantes, prestataires — de loin la plus petite.

Un troisième problème passe facilement inaperçu : la population est celle des Canadiennes et des Canadiens de 15 ans et plus, ce qui inclut les élèves du secondaire. « Le Canadien moyen » suggère des adultes en emploi.

(b) Le chiffre du recensement est une statistique : un résumé calculé à partir des personnes qui ont retourné un questionnaire. Le vrai revenu médian de l'ensemble de la population canadienne est un paramètre — la valeur qu'on obtiendrait en observant tout le monde, ce que personne n'a fait. Même un recensement est un échantillon dans les faits, puisque la couverture est incomplète et que le questionnaire détaillé n'est envoyé qu'à une fraction des ménages. Le chiffre publié estime le paramètre ; il n'est pas le paramètre.

(c) Plusieurs choses. La moyenne à côté de la médiane, ce qui révélerait l'asymétrie. La marge d'erreur. Si le chiffre est avant ou après impôt. S'il porte sur la personne ou sur le ménage — un chiffre par ménage est bien plus élevé et les deux se confondent souvent. Et l'année, puisque les chiffres de revenu sont révisés et que l'inflation rend trompeuse la comparaison de dollars nominaux d'une année à l'autre.

Corrigé de l'exercice 2.E2. 1. Santé autoévaluée. Catégorielle, ordinale ; échelle ordinale ; médiane. Les catégories se classent clairement, mais la distance entre « bonne » et « très bonne » est indéfinie.

2. *Nombre d'enfants de moins de 18 ans.* Numérique, discrète ; échelle de rapport ; moyenne, quoique la médiane soit souvent plus utile parce que la distribution est asymétrique et bornée à zéro par le bas.

3. *Pays de naissance.* Catégorielle, nominale ; échelle nominale ; mode. La seule moyenne disponible.

4. *Indice de démocratie libérale du projet V-Dem.* Numérique, continue ; d'intervalles dans les faits, et cela mérite un commentaire. L'indice va de 0 à 1 et a l'air d'être de rapport, mais c'est un composite de jugements d'experts remis à l'échelle, de sorte qu'un pays à 0,8 n'est pas deux fois plus démocratique qu'un pays à 0,4. La moyenne est conventionnelle et défendable ; le traiter comme une échelle de rapport ne l'est pas.

5. *Avoir voté à la dernière élection.* Catégorielle, nominale et binaire ; échelle nominale ; mode. Notez le cas particulier utile : codée 0/1, sa moyenne est égale à la proportion de personnes ayant voté, de sorte que des outils numériques peuvent s'y

appliquer sans rien enfreindre.

6. *Revenu personnel annuel*. Numérique, continue ; échelle de rapport ; médiane, parce que le revenu est asymétrique à droite et que la moyenne est tirée par la queue supérieure. C'est le point de l'exercice E1 qui revient par l'autre bout.

7. *Religiosité sur une échelle de 1 à 10*. Catégorielle, ordinale à strictement parler ; échelle ordinale ; médiane à strictement parler, moyenne couramment et à condition d'énoncer l'hypothèse d'intervalles égaux.

Le motif à dégager : les items 4 et 7 sont tous deux des échelles couvrant une plage numérique, et tous deux sont traités comme des échelles d'intervalles dans la pratique tout en étant ordinaux en principe. Les items 5 et 6 montrent que la moyenne appropriée dépend de la distribution autant que de l'échelle.

Corrigé de l'exercice 2.E3. (a) Une définition conceptuelle défendable : le capital social est l'ensemble des ressources auxquelles les gens ont accès par leurs réseaux sociaux — les liens, les normes de réciprocité et la confiance qui permettent aux individus et aux groupes d'agir ensemble efficacement. Il vaut la peine de distinguer le capital de *cohésion*, à l'intérieur de groupes serrés, du capital de *liaison*, qui les traverse : les deux peuvent varier indépendamment.

(b) Trois opérationnalisations à trois échelles. *Nominale* : le fait qu'une personne appartienne ou non à une association bénévole (oui/non). *Ordinale* : la fréquence des contacts avec le voisinage sur une échelle en cinq points, de jamais à quotidiennement. *De rapport* : le nombre d'associations bénévoles par millier de personnes résidentes, ou le nombre de personnes à qui la personne interrogée pourrait demander un service important.

(c) Menaces à la validité. Le dénombrement des adhésions compte les organisations formelles et rate les réseaux informels — qui sont la forme principale de capital social dans certaines communautés, de sorte que la mesure les sous-estimera systématiquement. La fréquence de contact autodéclarée est sujette à l'erreur de rappel et à la désirabilité sociale. La densité d'organisations mesure l'offre plutôt que l'usage : un quartier peut compter de nombreuses associations auxquelles peu de gens adhèrent.

(d) La densité d'organisations est la plus faisable à partir de données secondaires. Du côté canadien, on peut penser à la liste des organismes de bienfaisance enregistrés de l'Agence du revenu du Canada, agrégée par secteur postal, et aux cycles de l'Enquête sociale générale de Statistique Canada sur l'identité sociale et le don, qui portent l'adhésion associative et le bénévolat. N'importe quelle source précise, réelle et accessible fera l'affaire ; « des données gouvernementales » n'en est pas une.

Corrigé de l'exercice 2.E4. (a) La variation relative est la différence divisée par la

valeur de départ.

$$\frac{6000 - 2000}{2000} \times 100 = \frac{4000}{2000} \times 100 = 200 \%$$

L'affirmation tient. Notez qu'une hausse de 200 % signifie que le chiffre a triplé, et non doublé — une confusion courante.

(b) Les deux ont raison, pour des questions différentes, et une réponse qui en choisit une sans dire de quelle question il s'agit est incomplète. Le chiffre absolu de 4000 décès supplémentaires est ce qui détermine la charge pesant sur les bureaux de coroner, les services de traitement et les familles — c'est l'ordre de grandeur pertinent pour les politiques publiques. Le taux pour 100 000 est ce qui rend la variation comparable dans le temps et d'un territoire à l'autre, parce qu'il retire l'effet de la croissance démographique. Une chercheuse qui demande si l'épidémie s'est aggravée a besoin du taux ; une agence qui budgète sa réponse a besoin du dénombrement.

(c) Taux pour 100 000 :

$$2015 : \frac{2000}{35\,000\,000} \times 100\,000 = 5,71$$

$$2020 : \frac{6000}{38\,000\,000} \times 100\,000 = 15,79$$

Le taux est passé de 5,71 à 15,79 pour 100 000, une hausse d'environ 176 % plutôt que de 200 %.

La croissance démographique explique donc une partie de la hausse annoncée, mais une petite partie seulement. L'interprétation est ainsi *nuancée plutôt que renversée* — et c'est la réponse honnête. Conclure que le calcul du taux « démolit » le chiffre de 200 % serait une surcorrection. L'épidémie s'est aggravée considérablement selon l'une comme l'autre des deux mesures.

Corrigé de l'exercice 2.E5. (a) La variable est ordinale en principe : rien ne garantit que le pas de 2 à 3 représente le même accroissement de confiance que le pas de 7 à 8. Employer la moyenne suppose des intervalles égaux.

Savoir si c'est approprié dépend de ce qui est affirmé. Pour un écart aussi faible — 5,4 contre 5,1, trois dixièmes de point sur une échelle de dix — l'hypothèse fait un vrai travail, et la réponse dépendrait aussi de la taille de l'échantillon et de la dispersion, dont ni l'une ni l'autre n'est donnée. Une réponse défendable dit ceci : la moyenne est conventionnelle et acceptable comme résumé, mais un écart de 0,3 ne devrait pas être rapporté comme un résultat sans mesure d'incertitude, et les distributions devraient être comparées directement.

(b) Les menaces à la validité comprennent la désirabilité sociale, puisque

exprimer sa méfiance envers le gouvernement peut sembler inconvenant à certaines personnes ; l'ambiguïté de ce à quoi renvoie « le gouvernement fédéral » — le gouvernement en poste, l'institution ou la fonction publique — des personnes différentes répondant alors à propos d'objets différents ; et la possibilité que les hommes et les femmes emploient l'échelle différemment, ce qui produirait un écart mesuré sans écart réel de confiance. Cette dernière menace est la plus sérieuse ici, parce qu'elle fabriquerait exactement le résultat rapporté.

(c) Non. « 60 % plus » est une affirmation de rapport, et les rapports exigent un vrai zéro. Sur cette échelle, 0 signifie « aucune confiance » comme *étiquette*, et non comme absence d'une quantité mesurable, et l'échelle est au mieux d'intervalles. L'énoncé défendable est que les moyennes provinciales diffèrent de 2,6 points sur une échelle de dix — une différence, non un multiple.

Corrigé de l'exercice 2.E6. (a) Numérique et continue en apparence, quoique bâtie à partir de composantes discrètes. L'échelle est d'intervalles au mieux. Elle va de 0 à 100 avec un zéro apparent, mais ce zéro est le bas d'un schéma de cotation construit plutôt qu'une absence de liberté, et les composantes sous-jacentes sont des jugements d'experts.

(b) Non. Une affirmation de rapport exige un vrai zéro, et un score de 0 sur cet indice ne signifie pas l'absence complète d'une quantité mesurable appelée liberté — il signifie le score le plus bas que l'instrument peut attribuer. Rien ne garantit non plus qu'un écart de dix points veuille dire la même chose en haut de l'échelle qu'en bas.

(c) Plusieurs choix incorporent des jugements, et en nommer un seul répond à la question. Additionner 25 questions pondérées également affirme que chacune contribue également à la liberté — qu'une restriction donnée de la liberté de presse importe exactement autant qu'une restriction équivalente de la liberté de réunion. La sélection des 25 questions détermine ce qui compte comme liberté tout court, et les libertés civiles et politiques y figurent alors que la sécurité économique n'y figure pas. Et la cotation de 0 à 4 de chaque question est attribuée par des analystes qui appliquent des critères demandant interprétation.

(d) L'inquiétude porte sur la validité de construit, et plus précisément sur le fait que la liberté et la démocratie ne sont pas le même construit. Freedom House mesure les libertés civiles et les droits politiques ; un pays pourrait tenir des élections véritablement concurrentielles tout en restreignant la liberté de presse, ou protéger les libertés sous un gouvernement que personne n'a élu. Employer l'un comme mesure de l'autre suppose qu'ils voyagent ensemble, ce qui est une affirmation empirique et non une affirmation de définition. Une critique pourrait ajouter que l'indice est produit par un seul organisme, dont la perspective institutionnelle est connue, et qu'il ne devrait donc pas être la seule mesure sur laquelle une étude

s'appuie.

Corrigé de l'exercice 2.E7. Classements pour le tableau du chapitre 1.

Tranche de revenu (7 catégories). Catégorielle, ordinale ; échelle ordinale ; médiane.

Confiance envers le gouvernement (0–10). Ordinale en principe, traitée comme d'intervalles dans la pratique ; médiane à strictement parler, moyenne par convention.

Pays de naissance. Catégorielle, nominale ; échelle nominale ; mode.

Nombre d'enfants. Numérique, discrète ; échelle de rapport ; moyenne ou médiane, la médiane étant préférable vu l'asymétrie.

Santé autoévaluée (5 points). Catégorielle, ordinale ; échelle ordinale ; médiane.

Identification partisane. Catégorielle, nominale ; échelle nominale ; mode.

L'exercice est conçu pour qu'on y revienne, et la question utile est donc de savoir ce qui a changé. Au chapitre 1, il s'agissait de repérer les *limites* de chaque mesure — ce qu'elle laisse de côté ou brouille. Ici, il s'agit de repérer ce qu'on peut légitimement en *calculer*. Les deux sont liés : la limite de la tranche de revenu (distances indéfinies, tranche supérieure ouverte) est exactement ce qui la rend ordinale et exclut la moyenne.

Corrigé de l'exercice 2.E8. 1. La population est celle des personnes résidant dans la municipalité — ou, plus défendable vu le devis, celle des personnes qui fréquentent la bibliothèque. L'échantillon est formé des 600 personnes interrogées. La statistique est le chiffre de 32 %. Le paramètre qu'elle prétend estimer est la vraie variation du bien-être moyen dans la population, laquelle est inconnue.

2. Toute opérationnalisation défendable convient, pourvu qu'elle soit précise et que son échelle soit énoncée. Une échelle de bien-être validée à plusieurs énoncés, administrée avant et après, produisant un score ordinal ou quasi d'intervalles. Ou une mesure de rapport : le nombre d'heures par semaine passées en activité sociale hors du domicile. Ou nominale : le fait que la personne déclare avoir quelqu'un à appeler en cas d'urgence.

3. « A augmenté de 32 % » applique une opération de rapport à une échelle qui ne la porte pas. Passer de 5,0 à 6,6 sur une échelle de 0 à 10 est une hausse de 1,6 point ; appeler cela 32 % suppose que l'échelle a un vrai zéro, ce qui n'est pas le cas. Un meilleur compte rendu donne les valeurs de départ et d'arrivée ainsi que l'écart en points, et fournit une mesure d'incertitude — par exemple : « le bien-être moyen est passé de 5,0 à 6,6 points sur une échelle de dix ».

4. Fidélité : une mesure à énoncé unique, administrée une fois avant et une fois après, est vulnérable à l'humeur, à la météo et à l'ordre des questions, de sorte qu'une variation apparente peut n'être que du bruit. Validité : les échelles de bien-être sont sujettes à la désirabilité sociale, et les personnes qui savent que le programme est

évalué peuvent déclarer une amélioration qu'elles ne ressentent pas — surtout si ce sont les responsables du programme qui les interrogent.

5. Distribuer l'enquête à l'intérieur de la bibliothèque restreint l'échantillon aux personnes qui la fréquentent, et aux plus assidues d'entre elles. La statistique ne peut alors rien dire de la population résidente en général, et ne peut pas appuyer l'affirmation que le programme a amélioré le bien-être collectif — tout au plus que le bien-être a monté chez des personnes qui utilisaient déjà le service. C'est un problème de sélection, et aucune augmentation de la taille de l'échantillon ne le corrige. Le chapitre 5 le traite au complet.

Ce qu'il faut remarquer : l'item 5 mine l'item 1 — la population nommée dans l'affirmation et la population dont l'échantillon peut parler sont deux populations différentes.

Corrigé de l'exercice 2.E9. La réponse dépend de la vague que vous consultez et des chiffres actuels de Statistique Canada. C'est le raisonnement qui compte ici, non les villes en particulier.

(a) Dans `cma_panel`, le taux de criminalité simulé le plus élevé revient à Winnipeg, et le plus bas à l'une des RMR québécoises. Lisez-le dans le tableau plutôt que de le supposer.

(b) Sur l'indice de gravité de la criminalité réel, les RMR aux scores les plus élevés se trouvent constamment dans l'Ouest — Kelowna, Lethbridge, Winnipeg, Regina, Saskatoon — et les plus bas sont les RMR québécoises, Toronto se trouvant elle aussi près du bas.

(c) Globalement oui aux extrémités, et imparfaitement au milieu. Winnipeg en haut et les RMR québécoises en bas, cela concorde. Trouver un écart n'est pas une erreur : la préface promet que les ordres suivent la réalité *globalement* et non exactement, et les niveaux simulés sont délibérément faux d'un bout à l'autre.

(d) L'indice de gravité de la criminalité pondère chaque infraction par la peine moyenne qui lui est imposée, de sorte qu'un homicide y contribue bien davantage qu'un vol à l'étalage. Ce que cela *ajoute*, c'est une distinction qu'un simple taux ne peut pas faire : deux villes ayant des taux d'incidents identiques mais des compositions d'infractions différentes s'y trouvent correctement séparées. Ce que cela *masque*, c'est la pondération elle-même — l'indice est un composite (§2.6.6), et ses pondérations encodent la pratique en matière de peines, laquelle reflète des décisions de poursuite et des décisions judiciaires plutôt que le seul préjudice. Un changement du droit pénal en matière de peines change l'indice sans qu'aucun changement ne survienne dans la criminalité.

Le comparer à un simple taux n'a donc rien d'évident, parce que les deux sont des construits différents sur des échelles différentes. L'indice de gravité est un indice normalisé à une année de base, et non un dénombrement par habitant ; une ville peut

se classer plus haut sur l'un et plus bas sur l'autre. C'est une question de validité de mesure plutôt qu'un problème de données, et la nommer comme telle est l'objet de l'exercice.

Aucun chapitre français n'est encore traduit. Ce corrigé se remplira à mesure que les chapitres paraîtront ; les quatorze fichiers qui l'alimentent existent déjà et attendent leur contenu.