ACADÉMIE HUMANITAIRE · HUMANITAIRES JOBS

Mini MBA Programmes et MEAL : concevoir, piloter et prouver l'impact

Support de cours · Semaine 7 : Analyser les données : statistiques utiles, pièges et visualisation

Programme de 3 mois exactement (12 semaines, 144 heures) pour concevoir, piloter et évaluer des programmes humanitaires : évaluation des besoins et sévérité, chaîne de résultats et indicateurs, ciblage et transferts monétaires, planification et budget, plan MEAL et qualité des données, échantillonnage, analyse, redevabilité, évaluation, mesure d'impact, optimisation des ressources et gestion adaptative. Exercices chiffrés corrigés, trois examens mensuels avec études de cas, et un diplôme vérifiable de l'Académie Humanitaire.

https://humanitairejobs.org/fr/academie/mini-mba-programmes-meal/semaine-7/

Document gratuit de l'Académie Humanitaire. Il peut être imprimé et partagé librement à des fins de formation, sans modification.

Semaine 7 : Analyser les données : statistiques utiles, pièges et visualisation

Objectifs

- Choisir les bonnes statistiques descriptives et désagréger pour révéler les inégalités cachées par les moyennes.
- Éviter les pièges classiques : corrélation et causalité, paradoxe de Simpson, dénominateurs trompeurs.
- Analyser des données qualitatives avec rigueur et présenter les résultats pour qu'ils soient compris et utilisés.

Statistiques descriptives : moyenne, médiane, dispersion

La **moyenne** est sensible aux valeurs extrêmes ; la **médiane** (valeur qui partage l'échantillon en deux moitiés) ne l'est pas. Pour les revenus, les dépenses, les dettes ou les temps d'attente, distributions très asymétriques, la médiane décrit mieux le ménage typique.

Exemple. Dettes de dix ménages (USD) : 0, 0, 20, 30, 40, 50, 60, 80, 100, 1 620. Moyenne : 2 000 / 10 = 200 USD. Médiane : (40 + 50) / 2 = 45 USD. Dire « les ménages ont en moyenne 200 USD de dettes » décrit un seul ménage, pas la situation.

La **dispersion** compte autant que le centre : écart type, ou écart interquartile (entre le 25e et le 75e centile) pour les distributions asymétriques. Deux programmes avec un temps d'attente moyen de 2 heures sont très différents si l'un a des attentes entre 1 h 30 et 2 h 30, et l'autre entre 10 minutes et 7 heures.

Pour les indicateurs en proportion, on présente le numérateur et le dénominateur : « 62 % (248 sur 400) ». Les pourcentages calculés sur de très petits effectifs (« 50 % des personnes âgées », soit 2 sur 4) sont à éviter ou à signaler.

> Variation en points et en pourcentage : passer de 20 % à 30 % est une hausse de 10 points de pourcentage, soit une hausse relative de 50 %. Confondre les deux est l'erreur la plus fréquente des rapports.

Désagréger pour voir les inégalités

Une moyenne globale satisfaisante peut cacher un groupe laissé pour compte. La **désagrégation** (par sexe et âge, handicap, statut, zone, type de ménage) est le moyen de vérifier que le programme atteint tout le monde.

Exemple. Taux de satisfaction global : 82 %. Par groupe : ménages sans personne handicapée 86 % ; ménages avec une personne handicapée 54 %. Le programme ne fonctionne pas pour un groupe entier, et le chiffre global le masquait.

Le **tableau croisé** est l'outil de base : un indicateur en ligne, un groupe en colonne. Pour savoir si la différence observée entre deux groupes dépasse ce que le hasard de l'échantillonnage peut produire, on utilise un test (test du khi-deux pour des proportions, test t ou test non paramétrique pour des moyennes) ou l'on compare les intervalles de confiance. Un logiciel le calcule ; le rôle du responsable est de poser la bonne question et d'interpréter.

Attention aux **comparaisons multiples** : si l'on teste 20 différences au seuil de 5 %, on s'attend à en trouver une « significative » par pur hasard. On définit donc à l'avance les comparaisons principales du plan d'analyse, et l'on présente les autres comme exploratoires.

Corrélation, causalité et paradoxe de Simpson

Deux variables qui évoluent ensemble ne prouvent pas que l'une cause l'autre. Trois explications alternatives existent toujours : le **hasard**, un **facteur commun** (confondant) et la **causalité inverse**.

Exemple. Les ménages qui ont participé aux formations agricoles ont de meilleurs rendements. Mais ce sont aussi ceux qui ont plus de terres et d'outils, et qui étaient plus motivés à s'inscrire : la formation a peut-être un effet, mais la comparaison simple ne permet pas de le séparer de ces différences initiales. Mesurer un effet exige un contrefactuel (semaine 10).

Le **paradoxe de Simpson** survient quand une tendance présente dans chaque groupe s'inverse quand on regroupe les données, à cause d'une répartition inégale entre groupes.

	Centre A	**Centre B**
Cas sévères	60 guéris sur 100 (60 %)	110 guéris sur 200 (55 %)
Cas modérés	270 guéris sur 300 (90 %)	170 guéris sur 200 (85 %)
Total	330 sur 400 (82,5 %)	280 sur 400 (70 %)

Ici, A est meilleur dans chaque groupe et au total : pas de paradoxe. Mais si B recevait surtout des cas modérés et A surtout des cas sévères, B pourrait paraître meilleur au total tout en étant moins bon dans chaque groupe. Règle : comparer des performances seulement à **composition comparable**, ou en désagrégeant par le facteur qui change la difficulté.

Analyser les données qualitatives

Les entretiens et les discussions de groupe expliquent le « pourquoi » derrière les chiffres. Leur analyse suit une méthode, sans laquelle on ne retient que les citations qui confirment ce que l'on pensait déjà.

- **Transcription ou notes détaillées** le jour même, avec les mots des personnes.
- **Codage** : on attribue des codes (thèmes) aux passages. On part de codes liés aux questions d'évaluation, et l'on ajoute ceux qui émergent.
- **Matrice** : thèmes en ligne, groupes (femmes, hommes, jeunes, déplacés, hôtes) en colonne. On voit ainsi quels thèmes reviennent dans quels groupes.

- **Recherche active des contre-exemples** : les cas qui contredisent la tendance générale sont souvent les plus instructifs.
- **Saturation** : on arrête la collecte quand de nouveaux entretiens n'apportent plus de thèmes nouveaux.

On ne présente pas les résultats qualitatifs en pourcentages, mais l'on peut indiquer la fréquence : « ce thème est apparu dans 9 des 12 groupes de discussion, dont tous les groupes de femmes ». Une citation illustre un résultat, elle ne le prouve pas.

> Méthodes mixtes : le chiffre dit « 30 % des ménages ne retirent pas leur transfert », le qualitatif dit « le point de retrait est à 3 heures de marche et les agents manquent souvent de liquidités ». Les deux ensemble mènent à une décision.

Visualiser pour faire comprendre

Un graphique doit répondre à une question en quelques secondes. Règles pratiques :

- **Barres horizontales** pour comparer des catégories, triées par valeur ; **courbes** pour les évolutions dans le temps ; **cartes** pour les différences entre zones ; éviter les camemberts au-delà de trois catégories et les effets 3D.
- L'axe des barres commence à **zéro** ; sinon une différence de 2 points paraît énorme.
- Un **titre qui dit le message** (« La couverture a baissé dans les zones nord depuis mars ») plutôt qu'un titre neutre (« Couverture par zone »).
- Peu de couleurs, une couleur d'accent pour ce qui compte, et un repère pour la cible.
- La source, la date et la taille d'échantillon en bas du graphique.

Pour un tableau de bord de direction, on limite à 6 à 10 indicateurs clés, avec pour chacun la valeur actuelle, la cible, la tendance et un code couleur dont les seuils sont définis. Un tableau de bord de 40 indicateurs n'est lu par personne.

Sur une carte, on représente des **taux** et non des nombres absolus quand les zones ont des populations différentes : une grande ville aura toujours plus de cas qu'un village.

Du résultat à la recommandation

Une analyse utile suit une progression en trois temps : **constat** (ce que disent les données), **interprétation** (pourquoi, avec les données qualitatives et le contexte) et **recommandation** (ce qu'il faut faire, par qui, quand). Les recommandations sont peu nombreuses, précises et réalistes, et chacune découle d'un constat.

Il faut aussi indiquer le **niveau de confiance** : fort (plusieurs sources concordantes, échantillon suffisant), moyen, faible (une seule source, petit échantillon). Une direction peut décider sur une preuve moyenne en urgence, mais elle doit savoir qu'elle le fait.

Enfin, la **restitution** des résultats se fait avant le rapport définitif : une séance avec l'équipe et les partenaires permet de corriger les erreurs d'interprétation et de construire l'adhésion aux

recommandations. Les résultats sont aussi partagés avec les communautés, sous une forme adaptée (réunion, affiche, message vocal).

Exercice d'application : analyser un suivi post-distribution

Situation. Suivi post-distribution de 400 ménages. Temps d'attente : moyenne 2 h 40, médiane 1 h 20. Satisfaction globale : 78 %. Ménages dirigés par une femme (120) : 60 % satisfaits ; dirigés par un homme (280) : 85,7 % satisfaits. Le score moyen de consommation alimentaire est passé de 31 à 38 ; la proportion de ménages avec une consommation acceptable est passée de 30 % à 45 %. Dans les discussions de groupe, 8 groupes de femmes sur 10 citent l'heure de distribution (après 15 h) et le trajet du retour.

Travail demandé. 1) Que dit l'écart entre la moyenne et la médiane du temps d'attente ? 2) Exprimez la variation de la consommation acceptable en points et en pourcentage. 3) Quelle conclusion et quelle recommandation tirez-vous sur la satisfaction ?

Corrigé. La moyenne très supérieure à la médiane indique que la plupart attendent peu, mais qu'un groupe attend très longtemps : il faut identifier quels sites ou quels jours. Consommation acceptable : +15 points de pourcentage, soit +50 % en relatif (on ne peut pas attribuer ce changement au seul programme sans contrefactuel). Satisfaction : vérification, 0,3 × 60 + 0,7 × 85,7 = 78 %. L'écart de 26 points entre ménages dirigés par une femme et par un homme est important et expliqué par les groupes de discussion : horaires tardifs et trajet de retour dangereux. Recommandation : distributions le matin, file prioritaire, points de distribution plus proches, puis vérification au prochain suivi.

À retenir

- Médiane plutôt que moyenne pour les distributions asymétriques (revenus, dettes, temps d'attente).
- Une variation de 20 % à 30 % est +10 points, soit +50 % en relatif.
- Désagréger révèle les groupes laissés pour compte que la moyenne cache.
- Corrélation n'est pas causalité ; comparer des performances à composition comparable.
- Analyse qualitative : codage, matrice par groupe, recherche des contre-exemples ; une citation illustre, elle ne prouve pas.

Quiz

1. Dépenses de 5 ménages : 10, 15, 20, 25, 430 USD. Moyenne et médiane :
   A. Moyenne 100, médiane 20
   B. Moyenne 20, médiane 100
   C. Moyenne 25, médiane 20
   D. Moyenne 100, médiane 25

2. Une couverture passe de 40 % à 50 %. C'est :
   A. +10 % en relatif
   B. +10 points de pourcentage, soit +25 % en relatif
   C. +50 points
   D. +25 points

3. On teste 40 différences entre groupes au seuil de 5 %. Combien de résultats « significatifs » peut-on attendre par hasard seul ?
   A. Aucun
   B. Environ 2
   C. Environ 20
   D. Exactement 5

4. Sur une carte comparant des districts de populations très différentes, il faut représenter :
   A. Les nombres absolus de cas
   B. Des taux (par exemple pour 1 000 habitants)
   C. Seulement les noms des districts
   D. Les budgets des ONG

5. Un thème apparaît dans 9 des 12 groupes de discussion. La présentation correcte est :
   A. 75 % de la population pense cela
   B. Ce thème est apparu dans 9 des 12 groupes, en précisant lesquels
   C. Tout le monde pense cela
   D. Ce résultat est statistiquement significatif

Corrigé du quiz : 1. A · 2. B · 3. B · 4. B · 5. B