ACADEMIA HUMANITÁRIA · HUMANITAIRES JOBS

Mini MBA Programas e MEAL: conceber, pilotar e provar o impacto

Material do curso · Semana 7 : Analisar os dados: estatísticas úteis, armadilhas e visualização

Programa de exatamente 3 meses (12 semanas, 144 horas) para conceber, pilotar e avaliar programas humanitários: avaliação das necessidades e severidade, cadeia de resultados e indicadores, focalização e transferências monetárias, planeamento e orçamento, plano MEAL e qualidade dos dados, amostragem, análise, prestação de contas, avaliação, medição do impacto, otimização dos recursos e gestão adaptativa. Exercícios numéricos corrigidos, três exames mensais com estudos de caso e um diploma verificável da Academia Humanitária.

https://humanitairejobs.org/pt/academia/mini-mba-programmes-meal/semaine-7/

Documento gratuito da Academia Humanitária. Pode ser impresso e partilhado livremente para fins de formação, sem modificações.

Semana 7 : Analisar os dados: estatísticas úteis, armadilhas e visualização

Objetivos

- Escolher as estatísticas descritivas adequadas e desagregar para revelar as desigualdades escondidas pelas médias.
- Evitar as armadilhas clássicas: correlação e causalidade, paradoxo de Simpson, denominadores enganadores.
- Analisar dados qualitativos com rigor e apresentar os resultados para que sejam compreendidos e utilizados.

Estatísticas descritivas: média, mediana, dispersão

A **média** é sensível aos valores extremos; a **mediana** (valor que divide a amostra em duas metades) não o é. Para os rendimentos, as despesas, as dívidas ou os tempos de espera, distribuições muito assimétricas, a mediana descreve melhor o agregado familiar típico.

Exemplo. Dívidas de dez agregados (USD): 0, 0, 20, 30, 40, 50, 60, 80, 100, 1 620. Média: 2 000 / 10 = 200 USD. Mediana: (40 + 50) / 2 = 45 USD. Dizer «os agregados têm em média 200 USD de dívidas» descreve um único agregado, não a situação.

A **dispersão** conta tanto como o centro: desvio-padrão, ou amplitude interquartil (entre o 25.º e o 75.º percentil) para as distribuições assimétricas. Dois programas com um tempo de espera médio de 2 horas são muito diferentes se um tiver esperas entre 1 h 30 e 2 h 30, e o outro entre 10 minutos e 7 horas.

Para os indicadores em proporção, apresenta-se o numerador e o denominador: «62 % (248 em 400)». As percentagens calculadas sobre efetivos muito pequenos («50 % dos idosos», ou seja, 2 em 4) devem ser evitadas ou assinaladas.

> Variação em pontos e em percentagem: passar de 20 % para 30 % é um aumento de 10 pontos percentuais, ou seja, um aumento relativo de 50 %. Confundir os dois é o erro mais frequente dos relatórios.

Desagregar para ver as desigualdades

Uma média global satisfatória pode esconder um grupo deixado para trás. A **desagregação** (por sexo e idade, deficiência, estatuto, zona, tipo de agregado) é o meio de verificar que o programa chega a toda a gente.

Exemplo. Taxa de satisfação global: 82 %. Por grupo: agregados sem pessoa com deficiência 86 %; agregados com uma pessoa com deficiência 54 %. O programa não funciona para um grupo inteiro, e o número global escondia-o.

A **tabela cruzada** é a ferramenta de base: um indicador em linha, um grupo em coluna. Para saber se a diferença observada entre dois grupos ultrapassa o que o acaso da amostragem pode produzir, utiliza-se um teste (teste do qui-quadrado para proporções, teste t ou teste não paramétrico para médias) ou comparam-se os intervalos de confiança. Um software calcula-o; o papel do responsável é fazer a pergunta certa e interpretar.

Atenção às **comparações múltiplas**: se se testarem 20 diferenças ao limiar de 5 %, espera-se encontrar uma «significativa» por puro acaso. Definem-se portanto de antemão as comparações principais do plano de análise, e apresentam-se as outras como exploratórias.

Correlação, causalidade e paradoxo de Simpson

Duas variáveis que evoluem em conjunto não provam que uma causa a outra. Existem sempre três explicações alternativas: o **acaso**, um **fator comum** (confundidor) e a **causalidade inversa**.

Exemplo. Os agregados que participaram nas formações agrícolas têm melhores rendimentos. Mas são também os que têm mais terras e ferramentas, e que estavam mais motivados para se inscrever: a formação talvez tenha um efeito, mas a comparação simples não permite separá-lo dessas diferenças iniciais. Medir um efeito exige um contrafactual (semana 10).

O **paradoxo de Simpson** ocorre quando uma tendência presente em cada grupo se inverte quando se agrupam os dados, devido a uma repartição desigual entre grupos.

	Centro A	**Centro B**
Casos graves	60 curados em 100 (60 %)	110 curados em 200 (55 %)
Casos moderados	270 curados em 300 (90 %)	170 curados em 200 (85 %)
Total	330 em 400 (82,5 %)	280 em 400 (70 %)

Aqui, A é melhor em cada grupo e no total: não há paradoxo. Mas se B recebesse sobretudo casos moderados e A sobretudo casos graves, B poderia parecer melhor no total sendo pior em cada grupo. Regra: comparar desempenhos apenas com **composição comparável**, ou desagregando pelo fator que altera a dificuldade.

Analisar os dados qualitativos

As entrevistas e as discussões de grupo explicam o «porquê» por detrás dos números. A sua análise segue um método, sem o qual só se retêm as citações que confirmam o que já se pensava.

- **Transcrição ou notas detalhadas** no próprio dia, com as palavras das pessoas.
- **Codificação**: atribuem-se códigos (temas) às passagens. Parte-se de códigos ligados às perguntas de avaliação, e acrescentam-se os que emergem.
- **Matriz**: temas em linha, grupos (mulheres, homens, jovens, deslocados, comunidade de acolhimento) em coluna. Vê-se assim que temas regressam em que grupos.

- **Procura ativa de contraexemplos**: os casos que contradizem a tendência geral são muitas vezes os mais instrutivos.
- **Saturação**: para-se a recolha quando novas entrevistas já não trazem temas novos.

Não se apresentam os resultados qualitativos em percentagens, mas pode indicar-se a frequência: «este tema surgiu em 9 dos 12 grupos de discussão, incluindo todos os grupos de mulheres». Uma citação ilustra um resultado, não o prova.

> Métodos mistos: o número diz «30 % dos agregados não levantam a sua transferência», o qualitativo diz «o ponto de levantamento fica a 3 horas de caminho e os agentes muitas vezes não têm liquidez». Os dois em conjunto levam a uma decisão.

Visualizar para fazer compreender

Um gráfico deve responder a uma pergunta em poucos segundos. Regras práticas:

- **Barras horizontais** para comparar categorias, ordenadas por valor; **linhas** para as evoluções no tempo; **mapas** para as diferenças entre zonas; evitar os gráficos circulares para além de três categorias e os efeitos 3D.
- O eixo das barras começa em **zero**; caso contrário, uma diferença de 2 pontos parece enorme.
- Um **título que diz a mensagem** («A cobertura baixou nas zonas norte desde março») em vez de um título neutro («Cobertura por zona»).
- Poucas cores, uma cor de destaque para o que conta, e uma referência para a meta.
- A fonte, a data e a dimensão da amostra no fundo do gráfico.

Para um painel de controlo de direção, limita-se a 6 a 10 indicadores-chave, com, para cada um, o valor atual, a meta, a tendência e um código de cores cujos limiares estão definidos. Um painel de 40 indicadores não é lido por ninguém.

Num mapa, representam-se **taxas** e não números absolutos quando as zonas têm populações diferentes: uma grande cidade terá sempre mais casos do que uma aldeia.

Do resultado à recomendação

Uma análise útil segue uma progressão em três tempos: **constatação** (o que dizem os dados), **interpretação** (porquê, com os dados qualitativos e o contexto) e **recomendação** (o que é preciso fazer, por quem, quando). As recomendações são pouco numerosas, precisas e realistas, e cada uma decorre de uma constatação.

É preciso também indicar o **nível de confiança**: forte (várias fontes concordantes, amostra suficiente), médio, fraco (uma só fonte, amostra pequena). Uma direção pode decidir com base numa evidência média em emergência, mas deve saber que o está a fazer.

Por fim, a **apresentação** dos resultados faz-se antes do relatório definitivo: uma sessão com a equipa e os parceiros permite corrigir os erros de interpretação e construir a adesão às

recomendações. Os resultados são também partilhados com as comunidades, sob uma forma adaptada (reunião, cartaz, mensagem de voz).

Exercício: analisar uma monitorização pós-distribuição

Situação. Monitorização pós-distribuição de 400 agregados. Tempo de espera: média 2 h 40, mediana 1 h 20. Satisfação global: 78 %. Agregados chefiados por uma mulher (120): 60 % satisfeitos; chefiados por um homem (280): 85,7 % satisfeitos. A pontuação média de consumo alimentar passou de 31 para 38; a proporção de agregados com um consumo aceitável passou de 30 % para 45 %. Nas discussões de grupo, 8 grupos de mulheres em 10 referem a hora da distribuição (depois das 15 h) e o trajeto de regresso.

Trabalho pedido. 1) O que diz a diferença entre a média e a mediana do tempo de espera? 2) Exprima a variação do consumo aceitável em pontos e em percentagem. 3) Que conclusão e que recomendação retira sobre a satisfação?

Correção. A média muito superior à mediana indica que a maioria espera pouco, mas que um grupo espera muito tempo: é preciso identificar que locais ou que dias. Consumo aceitável: +15 pontos percentuais, ou seja, +50 % em termos relativos (não se pode atribuir esta mudança apenas ao programa sem contrafactual). Satisfação: verificação, $0{,}3 \times 60 + 0{,}7 \times 85{,}7 = 78$ %. A diferença de 26 pontos entre agregados chefiados por uma mulher e por um homem é importante e explicada pelos grupos de discussão: horários tardios e trajeto de regresso perigoso. Recomendação: distribuições de manhã, fila prioritária, pontos de distribuição mais próximos, e depois verificação na próxima monitorização.

A reter

- Mediana em vez de média para as distribuições assimétricas (rendimentos, dívidas, tempos de espera).
- Uma variação de 20 % para 30 % é +10 pontos, ou seja, +50 % em termos relativos.
- Desagregar revela os grupos deixados para trás que a média esconde.
- Correlação não é causalidade; comparar desempenhos com composição comparável.
- Análise qualitativa: codificação, matriz por grupo, procura de contraexemplos; uma citação ilustra, não prova.

Quiz

1. Despesas de 5 agregados: 10, 15, 20, 25, 430 USD. Média e mediana:
   A. Média 100, mediana 20
   B. Média 20, mediana 100
   C. Média 25, mediana 20
   D. Média 100, mediana 25

2. Uma cobertura passa de 40 % para 50 %. Isto é:
   A. +10 % em termos relativos
   B. +10 pontos percentuais, ou seja, +25 % em termos relativos
   C. +50 pontos
   D. +25 pontos

3. Testam-se 40 diferenças entre grupos ao limiar de 5 %. Quantos resultados «significativos» se podem esperar apenas por acaso?
   A. Nenhum
   B. Cerca de 2
   C. Cerca de 20
   D. Exatamente 5

4. Num mapa que compara distritos com populações muito diferentes, deve representar-se:
   A. Os números absolutos de casos
   B. Taxas (por exemplo, por 1 000 habitantes)
   C. Apenas os nomes dos distritos
   D. Os orçamentos das ONG

5. Um tema surge em 9 dos 12 grupos de discussão. A apresentação correta é:
   A. 75 % da população pensa isto
   B. Este tema surgiu em 9 dos 12 grupos, indicando quais
   C. Toda a gente pensa isto
   D. Este resultado é estatisticamente significativo

Respostas do quiz : 1. A · 2. B · 3. B · 4. B · 5. B