HUMANITARIAN ACADEMY · HUMANITAIRES JOBS

Mini MBA AI and Humanitarian Action: working with Claude and Gemini

Course handout · Week 7 : Data analysis and MEAL: spreadsheets, qualitative data and checking calculations

A programme of exactly 3 months (12 weeks, 144 hours) to use the AI assistants Claude (Anthropic) and Gemini (Google) in humanitarian work, in a useful, safe and responsible way: how they work and their limits, writing requests, data protection, ethics and humanitarian principles, reports, proposals, data analysis and MEAL, translation and communication with communities, research and source checking, disinformation and security, productivity and automation, use policy and deployment in an organisation. Corrected exercises, three monthly exams with case studies, and a verifiable diploma from the Humanitarian Academy. Independent programme, with no partnership with Anthropic or Google.

https://humanitairejobs.org/en/academy/mini-mba-ia-claude-gemini/semaine-7/

Free document from the Humanitarian Academy. It may be printed and shared freely for training purposes, without modification.

Week 7 : Data analysis and MEAL: spreadsheets, qualitative data and checking calculations

Objectives

- Use AI to write and explain formulas, clean and structure data, and prepare analyses.
- Analyse qualitative data (focus groups, community feedback) with AI in a rigorous and traceable way.
- Systematically check the calculations and interpretations produced by AI.

What AI brings to MEAL

MEAL teams spend a lot of time on repetitive tasks: cleaning files, writing formulas, producing tables, coding open answers, drafting summaries. AI can speed up these tasks:

- **Formulas and scripts**: writing a spreadsheet formula (XLOOKUP, COUNTIFS, pivot tables), a Python or R script, a query in a data collection tool, and explaining them line by line.
- **File analysis**: Claude and Gemini can read a CSV or Excel file, run calculations in a code environment and produce tables or charts (depending on the version and plan).
- **Qualitative data**: proposing a coding frame, coding answers, grouping by theme, extracting representative quotes.
- **Writing**: turning results into findings, then into recommendations.

But MEAL data is often **personal** (household lists, named surveys). The rule from week 3 applies: only anonymised and aggregated data, or work on a fictitious sample with the same structure.

Formulas and scripts: the fictitious data method

A safe and effective method to automate an analysis without exposing data:

- **1.** Describe the structure of your file (column names, types, examples of invented values) or create 10 fictitious rows with the same structure.
- **2.** Ask for the formula or script: "Write a formula that counts the number of female-headed households that received at least 2 distributions, by district."
- **3.** Test on the fictitious data, where you know the right result.
- **4.** Apply to your real data, **in your spreadsheet or on your computer**.
- **5.** Check with an independent calculation (manual filter, pivot table).

With Gemini in Google Sheets or the AI features of other spreadsheets, the AI acts directly in the file: check that the plan is professional and authorised for this data.

> AI writes the formula, you run it on your data, you check the result. Personal data does not leave your tools.

Checking calculations and statistics

Language models can make mistakes in mental arithmetic, especially on long series. When the assistant actually runs code, the calculations are reliable, but the **interpretation** can be wrong. Essential checks:

- **Denominators**: a percentage of what? Respondents, surveyed households, the target population? Are missing answers excluded?
- **Sample**: a result from 250 surveyed households does not apply to the whole population without an appropriate sampling method. Ask for the margin of error.
- **Mean or median**: for income or expenditure, a few extreme values distort the mean.
- **Correlation is not causation**: "households that received aid have a better food consumption score" does not prove that the aid is the cause without a comparison group.
- **Disaggregation**: sex, age, disability, district; check that the subtotals add up to the total.

Calculation. A simple random sample of 384 households: margin of error of about ±5% at a 95% confidence level for a proportion close to 50%. If the AI states "62% of households in the province" from 60 households chosen by convenience, the conclusion is not representative: this must be said.

Qualitative data: rigour and traceability

For focus group transcripts or community feedback (anonymised):

- **Coding frame**: ask for a proposal based on your evaluation questions, then validate it as a team.
- **Coding**: have the extracts coded with the validated frame, requiring for each theme the number of extracts and two or three exact quotes with their reference (group, anonymised location).
- **Check**: recode a sample yourself (for example 20%) and compare. If agreement is low, revise the frame or the instruction.
- **Nuances**: explicitly ask for minority and contradictory views; AI tends to summarise towards the majority.
- **Quotes**: check that each quote exists word for word in the source; a rephrased quote must not be presented in quotation marks.

NotebookLM or a Claude Project with the transcripts as sources make this work easier, with answers linked to the source passages.

Feedback and complaints mechanisms

Feedback and complaints mechanisms sometimes receive hundreds of messages (phone lines, SMS, boxes, messaging apps). AI can help to:

- propose categories (information, request for assistance, complaint about quality, sensitive complaint);
- summarise monthly trends from aggregated and anonymised data;

- draft validated standard replies.

Strict limits:

- **sensitive complaints** (sexual exploitation and abuse, fraud, security) follow a confidential human channel and never go through an AI assistant;
- automatic classification of a message must never delay the handling of a serious complaint: a human reads incoming messages;
- people must know how their messages are handled.

From results to recommendations

The most delicate step of an evaluation is moving from data to **findings**, then to **recommendations**. AI can propose a clear structure (finding, evidence, conclusion, recommendation, owner, deadline), but:

- each finding must be supported by at least two sources or methods (triangulation);
- recommendations must be realistic for the organisation: budget, capacities, timeline;
- evaluation criteria (for example the OECD DAC criteria: relevance, coherence, effectiveness, efficiency, impact, sustainability) structure the analysis.

Then ask for a critique: "Which conclusions of this report are not sufficiently supported by the data presented?" This adversarial review clearly improves quality.

Exercise: checking a survey analysis

Situation. A post-distribution survey was conducted with 300 households out of 4,500 beneficiaries, selected by random sampling. 270 answered the question on how the money was used. The AI produced: "189 households used the money mainly for food, i.e. 63% of the programme's beneficiaries. The programme therefore improved the food security of 2,835 households."

Task. 1) What is the correct percentage? 2) Is the extrapolation correct? 3) Is the conclusion on food security justified?

Answer. 1) The denominator is the number of respondents to the question: 189 / 270 = 70%, not 189 / 300 = 63%. This must be stated: "70% of the 270 households that answered". 2) As the sample is random, the proportion in the population can be estimated with a margin of error (about ±5% here): "about 70% (± 5 points) of beneficiary households". The figure of 2,835 (63% of 4,500) rests on the wrong percentage and on false precision. 3) No: having spent the money on food does not prove an improvement in food security. Indicators such as the food consumption score or the coping strategies index would be needed, compared with a baseline. Correct wording: "Food is the main declared use of the money; the effect on food security is not measured by this survey."

Key takeaways

- AI writes formulas and scripts; you test them on fictitious data and then run them in your tools.
- Check denominators, sample, mean or median, causality and disaggregation.
- Qualitative data: validated frame, exact quotes, recoding of a sample, minority views.
- Sensitive complaints follow a confidential human channel, never an AI assistant.
- Each finding is triangulated; each recommendation is realistic.

Quiz

1. 189 households out of 270 respondents to a question represent:
   A. 63%
   B. 70%
   C. 189%
   D. 27%

2. The safe method to have a formula written without exposing data is to:
   A. Paste the complete file
   B. Email the file to the provider
   C. Use fictitious data with the same structure
   D. Delete the columns with numbers

3. A complaint of sexual exploitation received on the helpline must be:
   A. Handled through a confidential human channel
   B. Classified by AI first
   C. Summarised by a consumer assistant
   D. Published in the monthly report

4. "Assisted households eat better" without a comparison group proves:
   A. Causation
   B. Certain impact
   C. Efficiency
   D. At best an association, not causation

5. To check qualitative coding done by AI, you:
   A. Trust the result
   B. Recode a sample yourself and compare
   C. Ask for another format
   D. Delete minority views

Quiz answers : 1. B · 2. C · 3. A · 4. D · 5. B