CCCM IM Handbook

CCCM Cluster

Table of contents

List of Figures

Introduction

> **Warning**
>
> The handbook is in an early draft stage. Most sections are incomplete and structure of the handbook may be subject to change.

> **Note**
>
> This handbook is also available to download in pdf and epub formats (see top left). To use the handbook as an offline mobile app, click "Add to home screen" on Chrome and visit the pages you wish to cache/save.

This handbook was developed with the aim of providing a consolidated guidance document covering the key aspects of Information Management (IM) for Camp Coordination and Camp Management (CCCM). It was created to fill the gaps in guidance specific to information management in CCCM and to document, consolidate and streamline existing IM practices within the sector.[1]

The primary audience for this handbook are information management staff of all levels, who are involved, or plan to be involved in IM for CCCM in either a programmes or cluster capacity. The handbook is also relevant to non IM personnel especially coordinators, acknowledging the importance of understanding both data literacy and CCCM analytical norms outside of the IM function.

The handbook is presented in three parts, broadly reflecting the differing use-cases and audiences.

- **Part 1** provides an overview of the key concepts related to humanitarian information management, which can be applied to all sectors and technical areas in a humanitarian response.
- **Part 2** builds upon the knowledge from part 1, applying it to the role of information management within CCCM programmes.
- **Part 3** focuses on the role of CCCM IM within cluster coordination, with the key responsibilities for each stage of the Humanitarian Programme Cycle (HPC).

[1] Development of the handbook started in late 2021, with the first draft planned for February 2022.

The handbook is available as a website (viewable on desktop or phone) or can be download as a pdf or kindle format for offline viewing. While viewing the web version, the left margin shows the three parts of the handbook, containing each chapter. Some of the larger chapters are split into sections (denoted by a downward-facing arrow). The right margin of the screen are for easy navigation through the content ofe ach chapter/section.

Humanitarian approaches and tools grow and change over time, which is particularly evident in the field of information management. This handbook aims to be a *living document* whose contents will be continually updated to reflect our growing knowledge and best practices, and evolution in our approaches and tools in both IM and CCCM.

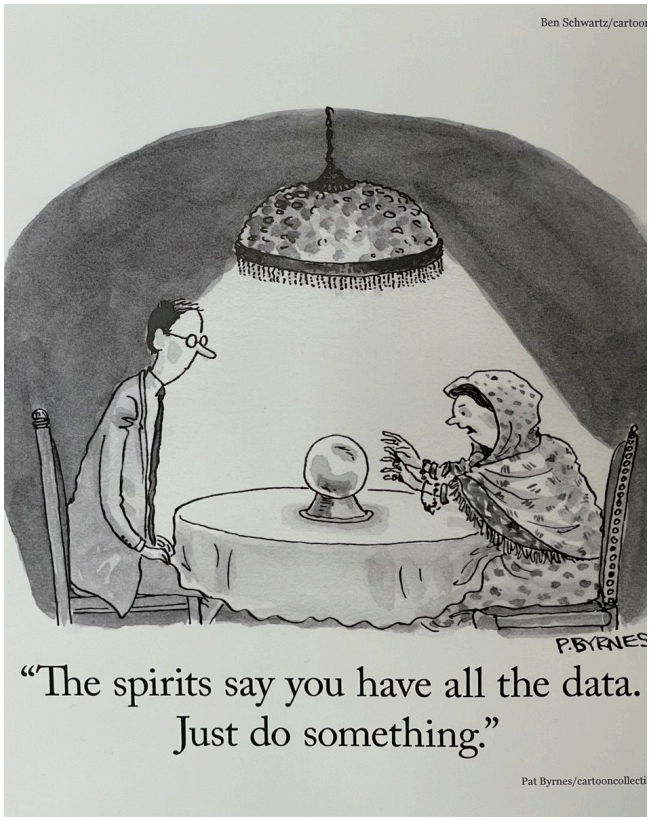


Figure 1: Towards evidence-based decision making

Feedback

If you have any questions, wish to correct any technical or textual mistakes, or wish to suggest improvements to this handbook, plese get in touch with the Global CCCM Cluster Information Management Officers (IMO) Brian Mc Donald - bmcdonald@iom.int or Alisa Ananbeh - ananbeh@unhcr.org

We would like to thank everybody for their support in developing this handbook and hope that you find it accessible and useful

- the CCCM Cluster team

Acknowledgements

The content in this handbook are drawn from three main sources:

- The experience, advice and materials from our CCCM colleagues in the field.
- Guidance materials developed at the global-level from inter-agency platforms including the Global Information Management Working Group and from other Clusters.
- Many excellent information management trainings including: OCHA's Coordinated Assessment and Information Management training (CAIM); and ACAPS's Humanitarian Analysis Program.

Part I

Humanitarian IM

Chapter 1

Data Literacy

> **Warning**
>
> This chapter is in draft stage.

The purpose of this chapter is to introduce the reader to the concept of data, what it is, how it used in humanitarian response and its relevance in the role of information management. This chapter forms an important basis for subsequent chapters as it aims to clearly describe key concepts around data to ensure their clear and shared understanding. This shared vocabulary is vital for the collaboration needed at the various stages of the data's lifecycle. This chapter is primarily aimed at IM's but is also relevant to any humanitarian involved to any degree in evidence-based decision making.[1]

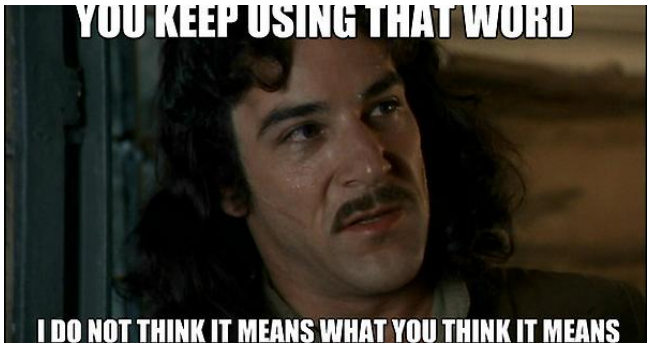


Figure 1.1: A shared understanding of terms is important for engagement around data

[1] Adapted from the FTs Visual Vocabulary. A similar graphics decision tree, based on the type and number of variables, is available at Data-to-Viz.com

1.1 What is data?

Data is the physical representation of information in a manner suitable for communication, interpretation, or processing by human beings or by automatic means.[2] It can be structured or unstructured, can come in many different forms (human-readable or machine-readable) and can come from any number of sources, using any number of methods. While the terms *data*, *information* and *knowledge* are quite often used interchangeably it is helpful to think of information as data integrated into context, and knowledge as a collection of information, processed in a way that provides learning.

1.2 What does it look like?

Data is all around us but is usually messy and unstructured. Processing this information into a structure that can provide sense is at the core of IM. This 'sense-making' can use different approaches - an experienced camp manager may decide to walk into a new camp, walk around it observing it, deciding on what actions they need to prioritize. Another may prefer to set up a list of indicators to measure specific needs in the camp. The approaches are different (and quite often complimentary) but the goal and process are to a large extent the same.

:::{admonition} Exercise Take a look at your desk and choose an object. Describe the attributes of that item. Perhaps you can describe the items colour, its length, its width, its texture, the materials its constructed from or how effective it is for your work. A surprising amount of data can be gathered from even the simplest of objects. :::

To get from messy data to structured that that can be used - by itself, or more commonly in conjunction with other datasets - a degree of organizing, tagging or categorizing must take place. If a survey is used, those categories are determined by the questions asked the type of questions and the response options. When setting these categories it is very important that each person involved with the data - from the person giving the response, the enumerator right up to those whose programmatic decisions it informs - has a clear and common understanding of what and how a concept is captured in these categories.

1.2.1 Formats

Valuable humanitarian data can often start out as paper survey responses, hand written notes (ie. distribution details) or as handwritten notes (such as from Focus Group Discussions). To aid the cleaning, processing and management of this data, digitization may be required. Digital data can be stored in a number of the following formats and is closely linked to the tools used to gather and/or store the data:

[2] Colorbrewer is a good resource for picking color palettes. Datawrapper have a good blog post describing the use of different colour scales.

- **Tabular data:** By far the most common format for humanitarian data, Excel or Comma Separated Value (CSV) files show data as a table where each column represent as variable in your data and each row ideally represents an observation. [3]

- **Relational databases:** Data cant always be represented in a single table. Quite often there is a need to present the data across multiple tables, showing the linkages(relationships) between variables in different table. Relational databases provide an underlying data model for most modern websites and software. An example use for a relational database could be in the recording of trainings, where one table contains rows, each representing a single training while a second table contains the list of participants. The relationships between these two tables could be defined as *each training can contain multiple participants* and *each participant can attend multiple trainings*

- **APIs** To aid the access and transfer of data, it is very common for modern software systems to have an API, in which other websites (or data analysis tools) can request data from the underlying data store. The most common file format for these is called JSON, a semi-human readable format with the advantage over tabular formats in that is can represent messy semi-structured data or complex relationships that would otherwise require a database. [4]

- **Spatial:** Spatial data formats such as .shp, .gpg, .geojson .geotiff or .dem are used to store 2d or 3d spatial data. Most of these formats can display or export to tabular formats.

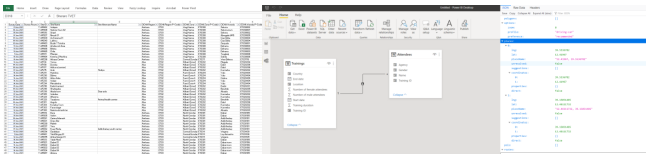


Figure 1.2: Examples of data as tables, a relational database, and as JSON from an travel distance API

1.2.2 Sources

- Common Operating Datasets (CODs)
- HDX
- Internal systems
- Others
- Non traditional sources

[3] For a detailed explanation of RGB and CMYK and how they differ, see here
[4] ACAPS have a great guide on the Use of Colour in Data Display

1.3 Data concepts

1.3.1 Values

Simple data values can be numeric; such as an integer (whole number) or float; boolean, such as a yes/no or true/false; or a string, a sequence of symbols such as a text answer. Compound values are combinations of simple values. Examples include dates, time, or list of values (such as a list of answers to a multiple choice question) [5]

1.3.2 Types of data

Variables (items of data) can spilt into two groups, quantitative (numeric) or categoric (no inherent order). Quantitative variables can either be discrete, meaning they have a finite number of values (eg household size) or continuous, meaning an infinite number of values are possible (eg. a persons height or distance to a health facility)

1.3.3 Scales

Data can be classified under the following 4 scales of measurement:

- **Nominal scales**: Nominal values/variables, sometimes called *categorical values* don't have a numeric value so cannot be added, subtracted or multiplied. They do not have an order. For example, the name of a district that an IDP is from.
- **Ordinal scale**: Contains values that can be put in order. For example, the levels of satisfaction with a training.
- **Interval scale**: Contains ordinal numbers with meaningful divisions. For example, temperature or time.
- **Ratio scale**: Ratio scales have all of the characteristics of interval scales as well as a true zero. For example, a persons height.

1.3.4 Goal/Strategic Objective

A specific end result desired or expected to occur as a consequence, at least in part, of an intervention or activity. It is the higher order objective that will assure national capacity building to which a development intervention is intended to contribute.

[5] For a deeper dive into technical writing, this free Google course is highly recommended

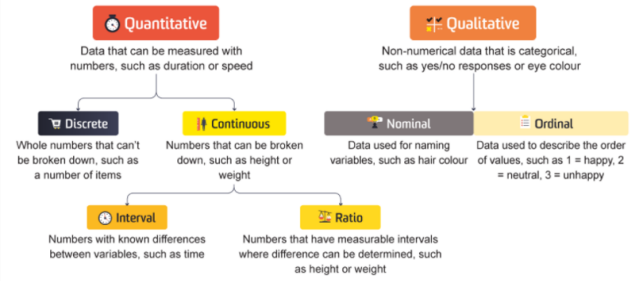


Figure 1.3: Types of data and how they relate

1.3.5 Impact

Impact implies changes in people's lives. This might include changes in knowledge, skill, behaviour, health or living conditions for children, adults, families or communities. Such changes are positive or negative longterm effects on identifiable population groups produced by a development intervention, directly or indirectly, intended or unintended.

1.3.6 Outcome

Outcomes represent changes in the institutional and behavioral capacities for development conditions that occur between the completion of outputs and the achievement of goals.

1.3.7 Outputs

Outputs are changes in skills or abilities and capacities of individuals or institutions, or the availability of new products and services that result from the completion of activities within a development intervention within the control of the organization. They are achieved with the resources provided and within the time period specified.

1.3.8 Activities

Actions taken or work performed through which inputs, such as funds, technical assistance and other types of resources, are mobilized to produce specific outputs.

1.3.9 Inputs

The financial, human, material, technological and information resources used for development interventions.

1.3.10 Indicators

Usually separated into two categories: - Need indicators: a quantitative or qualitative unit of measurement of need which when monitored periodically can be used as a measure of impact. - Response/performance indicators: used to measure outputs or outcomes.

1.3.11 Target

Specifies a particular value that an indicator should reach by to meet agreed standard of service or programme goals. Setting target values for most indicators requires a level of contextualization and can be influences by external factors such as resources available.

1.3.12 Standard

…

1.3.13 Benchmark

Reference point or standard, including norms, against which progress or achievements can be assessed. A benchmark refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably expected to have been achieved in similar circumstances.

1.3.14 Primary data vs secondary data

…

1.4 Information Management tips

The following tips are a collection of commonly encountered issues in humanitarian IM and how to avoid them. The tips are a shortened from their usual form to avoid overlap with other chapters where many of the issues are expanded in more detail.[6]

1.4.1 Use Excel for numerical data

Don't use software such as Word for gathering and analysing numerical data. Likewise, be careful not to use Excel to over visualise how your data is represented. Simple, well structured data is best for analysis and sharing with others.

[6] Adapted from The Chicago Guide to Writing about Numbers, by Jane E. Miller

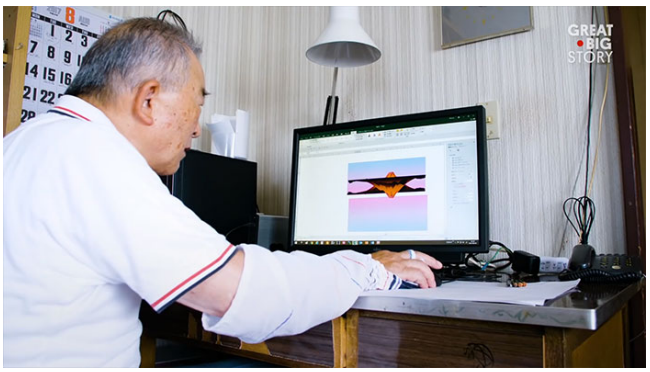


Figure 1.4: Tatsuo Horiuchi uses Excel to paint beautiful Japanese lanscapes. Don't follow Tatsuo

1.4.2 Save often, use versioning, and name files sensibly

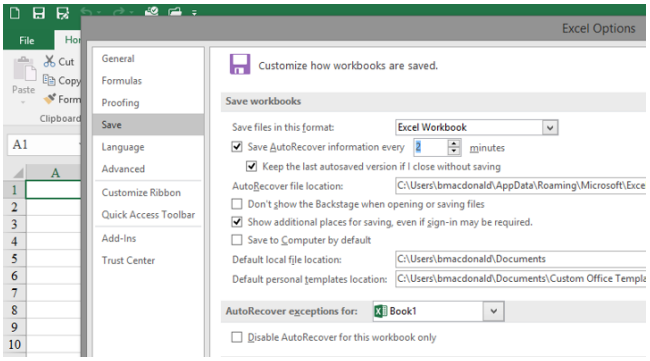


Figure 1.5: Click save!

You don't want to work all day on an analysis to suddenly find that the fil crashed before you had a chance to save it. In Excel, there is an option to have your files autosave at a specified interval.

If saving to OneDrive or Sharepoint, you will see a small downward arrow which allows youto view the file's *Version History*. to view or roll back to previous version of a document. Alternatively you can save multiple versions of the file following certain milestones or use a software versioning software such as Git.

Try have you and your team use a consistent, welll understood naming convention for files. The format I use is as follows: *[year][month(2digists)][date]-[initials]-[version]*. For example 20182307-IMtips-BMD-v1.xlsx *File > Options* to get Excel to save more regularly

1.4.3 Backup your data

What if your computer crashes or is stolen?
What if you need to collaborate on a file?

Make use of Onedrive/Sharepoint. Onedrive is ideal for working documents (you can right-click on a file and share it collaboratively).

1.4.4 Check for existing data, communicate

Networking and communication are important but sometimes overlooked skills for IM. It's important to know what other agencies and clusters are planning in terms of data collection and what challenges they are facing that may be better addressed with a collective approach. Checking for pre-existing data or planning assessments can help avoid duplication of efforts and unnecessarily *reinventing the wheel*.

1.4.5 Use mobile data collection

The use of mobile data collection tools such as Kobo Toolbox support faster and more robust data collection. By enforcing checks on data inputs it reduces input errors, while also removing time consuming and error-prone tasks of manual data entry of paper forms.

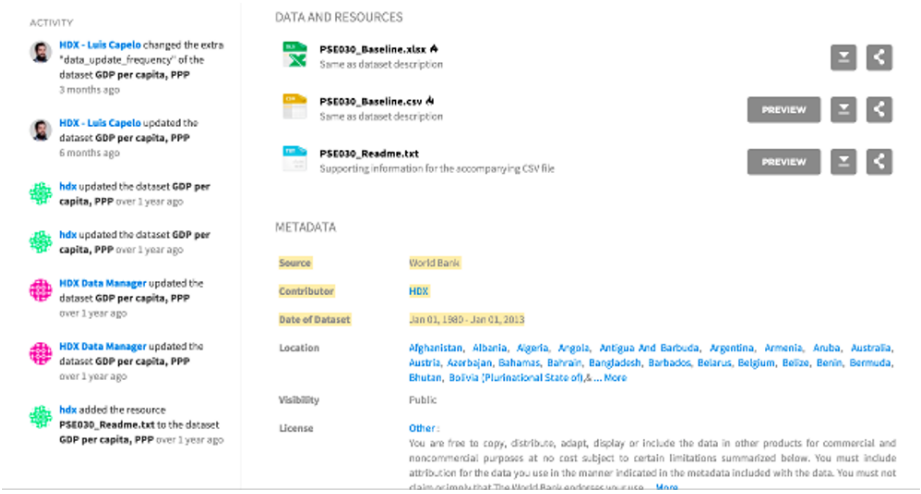
While ideal for surveys/assessments, be careful not to over-fit such tools into scenarios that require *case management* type functionality.

1.4.6 Consistent variable naming

Are we talking about the same thing? The terms/concepts describe in your surveys and data - does everyone have a clear and shared understanding of what they mean? Are you reusing well known and tested terms or are you inventing new ones. [7]

[7] Data journalism put increasing emphasis on the need for a good annotation layer, as can be seen by this article from the Financial Times

1.4.7 Understand meta-data



...and why it is important. Meta-data is data that described data. For example, your survey data files should contain information describing where the data was collected, on which dates, the methodology used and relevant focal point. Including metadata in your datasets is an important habit for IMs, as it encourages reuse of the data and signifies a robust approach to analysis.

1.4.8 Spreadsheets - only one piece of information per cell

	A	B	C	D
1	org	activity	district	
2	Blue Cross	shelter	Kathmandu, Bhaktapur	
3	Camfam	WASH	Kathmandu, Dhading	
4				
5				

	A	B	C	D
1	org	activity	district	
2	Blue Cross	shelter	Kathmandu	
3	Blue Cross	shelter	Bhaktapur	
4	Camfam	WASH	Kathmandu	
5	Camfam	WASH	Dhading	
6				

Storing multiple points of data in a single cell makes many types of analysis very difficult. Where possible try to expand these values onto their own rows (sometimes called “exploding” or “melting”.)

1.4.9 Record data at a granular level and aggregate up

When you have data at a low unit of measurement, for example, the number of people using a specific Complaints and Feedback (CFM) desk, it is straightforward to aggregate that data to a higher unit, for example, the number of people using CFMs in a district. However, be careful of receiving data already aggregated as it is usually not possible to disaggregate it into its component parts. If your analysis depends on having data at a certain unit-level, make sure to have it collected or sent to you in at least the same or lower level of disaggregation. Aggregation may hide or disregard useful data useful for your analysis or quality control.

1.4.10 Learn pivot tables

	Completed distributions								
	Tarpaulin	Blankets	Tents	Household kits (excluding tarpaulins)	Construction Materials	Tool kits and fixings	(blank)	Unconditional Cash	Bu (72
Total	169,257	77,307	767	68,261	1,009	2,803	540	8,584,159	
Dolakha	76,696	55,720	54	15,377	-	-	540	7,787,500	
Kabhrepalanchok	35,103	6,770	625	3,478	1,009	1,144	-	-	
Okhaldhunga	18,697	1,572	68	11,897	-	-	-	-	
Ramechhap	16,907	1,743	-	2,985	-	-	-	-	
Sindhuli	21,854	11,502	20	34,524	-	1,659	-	796,659	
Sindhupalchok	127,717	58,590	4,652	28,238	684	1,700	-	29,032,500	
East									

Pivot tables are a powerful tool for aggregating data in Excel. (also called "groupby" in other software)

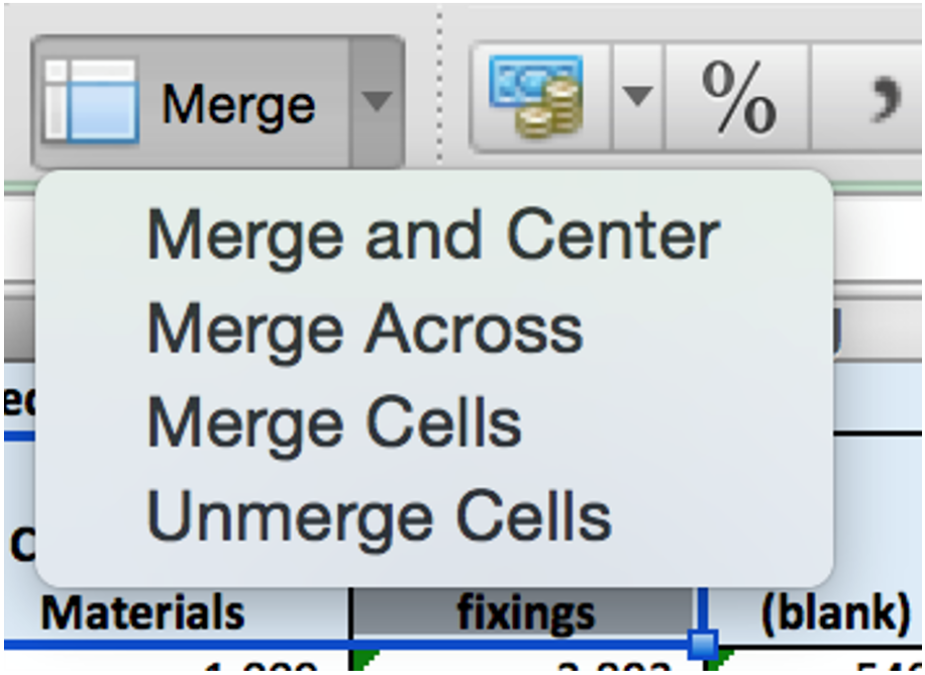
1.4.11 Learn VLOOKUP and Index Match

=VLOOKUP(B3,Sheet1!$A$1:$C$75,3,FALSE)

C	D	E	F	G	H	
Districts	pcode	Dead People	Missing Peop	Injured Peop	Affected Fam	Displa
#adm3	#adm3+code	#affected+de	#affected+m	#affected+injured		
Kathmandu	524 2 05 27	1226	34	2450	432245	
Lalitpur	524 2 05 25	180		3505	35648	
Bhaktapur	524 2 05 26	333	4	227	34192	
Sindhupalch	524 2 05 23	3532	123	3831	69105	
Dhading	524 2 05 30	678	3	838	73842	
Kabhrepalan	524 2 05 24	318	3	1271	77781	
Nuwakot	524 2 05 28	1109	15	1570	74999	
Dolakha	524 2 04 22	177	12	259	57744	
Gorkha	524 3 07 36	449	1	2123	57518	
Rasuwa	524 2 05 29	660	71	752	9422	
Lamjung	524 3 07 37	5		43	13273	
Khotang	524 1 03 12			7	6873	
Ramechhap	524 2 04 21	42		132	41278	
Chitwan	524 2 06 35	10		2	10474	

Figure 1.6: vlookup

1.4.12 Don't merge cells in a spreadsheet



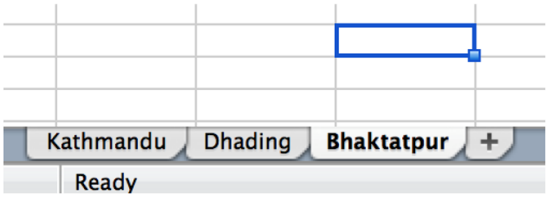
Don't merge cells in a spreadsheet. Tables should contain an equal number of rows and columns. Merging cells breaks pivoting and filtering and goes against rule number 1 (when done for reasons of aesthetics).

1.4.13 Keep data types and names consistent in columns

B	C	D
activity	district	Beneficiary numb
shelter	Kathmandu	320
shelter	Bhaktapur	Two hundred
WASH	Kathmandu	50%
WASH	Dhading	400

Make sure that the spelling (including case) and format of values remain consistent for all values in a column.

1.4.14 Keep all similar data in one sheet



	A	B	C
1	district	org	activity
2	Kathmandu	RedCross	shelter
3	Bhaktapur	red cross	shelter
4	Kathmandu	Red Cross	WASH
5	Dhading	redcross	WASH
6			

Resist the temptation to split large datasets across tabs, using a tab for each region etc. This makes analysis and consolidated storage and management of the data harder. Instead add a column to capture the category names.

1.4.15 Check data relationships

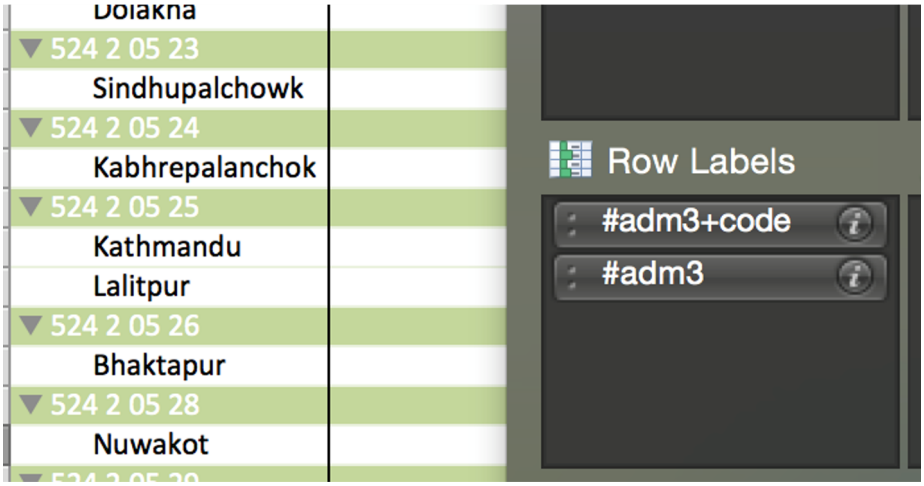
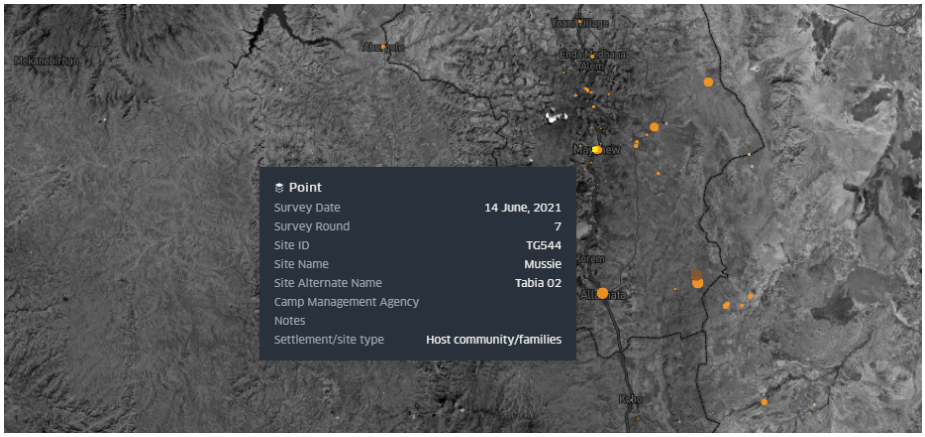


Figure 1.7: check relationaships

1.4.16 Get coordinates for point data



If you have data that represents a point such as a hospital or a camp, instead of just naming it or giving its address you could also collect coordinates for it. There are many possible ways of doing this:

- Using any app on your phone that records GPS
- In a mobile data collection tool such as Kobo Toolbox
- Using a geocoder, which outputs coordinates for given addresses or locations. Nominatim is a free service using Open Street Map data.
- The low tech approach of drawing on paper maps, digitizing later. Fieldpapers

Chapter 2

What is Information Management?

> **Warning**
>
> This chapter is in draft stage.

..

2.1 Profiles

..

2.2 Skills and attitudes

.

2.3 Responsibilities

- The core cluster functions and how they relate to IM.

2.4 IM workflow

.

Figure 2.1: A Data Skills Framework by the ODI

2.4.1 Collection

This section discusses: - Secondary data review - Primary data collection

2.4.2 Processing

.

2.4.3 Analysis

.

2.4.4 Dissemination

.

Chapter 3

Design & Acquire

This chapter is in draft stage.

Any information that doesn't not inform or change an decision is worthless.

– Sam L. Savage

- many illustrations start with "data collection" but this should not be the first step
- secondary data review vs primary data
- formulating the questions that need answering.
- create a data analysis plan
- common data sources
- CODs
- methods - KI. FGD
- sampling
- minimal viable information

3.1 Before you start collecting data.

.stop and think. IM guidances sometimes present IM task in a linear fashion that starts with *data collection*. This is problematic. Good analysis requires us to first start with an articulation of what are the overarching questions that need answering and what decisions will be informed by this evidence.

The next step is to take stock of existing data and to develop an analysis plan linking your information needs to you existing data and/or any potential future data collection exercises.

Collection of data has costs - it costs time, it has HR costs, it has financial costs. It is important to understand these costs and viewing them in relation to the value it provides to your decision making. Similarly, when developing surveys, adding questions are no without cost, as for each question that is added, there is a cognitive cost(more questions = higher likelihood of mistakes), a time cost (longer surveys take longer to do) and a quality cost (cleaning and data checks

are more difficult on larger datasets). Understand that 80% of the value of your survey will come from 20% of your survey. As you add additional questions to your survey, the value they add to your decision-making diminishes. [1]

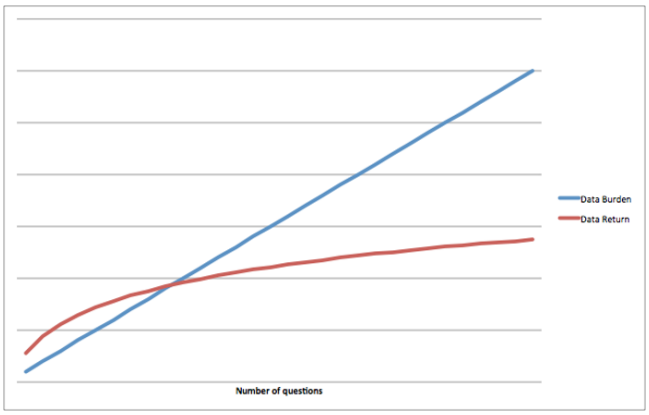


Figure 3.1: As your number of survey questions increase, the return on effort gradually diminishes.

3.2 Methodologies

- KI
- FGD
- Observations
- Non traditional sources
- Representativeness

3.2.1 Sampling

.

[1] Adapted from the FTs Visual Vocabulary. A similar graphics decision tree, based on the type and number of variables, is available at Data-to-Viz.com

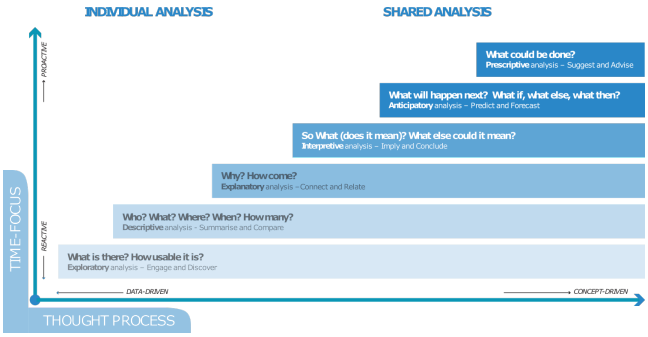
Chapter 4

Analysis

This chapter is in draft stage.

.

4.1 The Analysis Spectrum



.

4.2 Understanding Bias

.

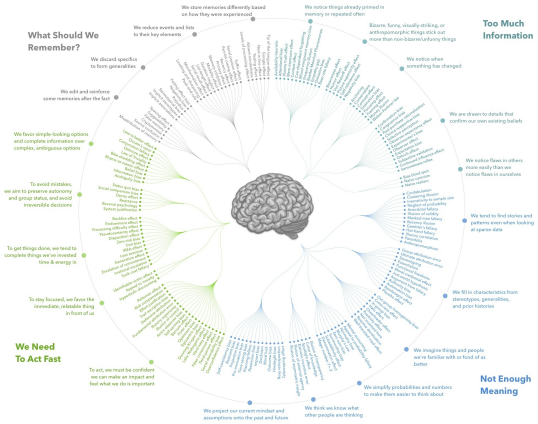


Figure 4.1: 188 Cognitive biases (adapted from Mangoongian 2021) - view larger

Chapter 5

Communication & Visualization

Communication may be the final step in the IMs workflow but is by no means the least important. The best analysis, using the best data collection methods and tools are worthless if the communication of these findings are insufficient to inform decisions. [1]

5.1 How to write about numbers

When preparing to write up an analysis, it is important to first consider the following:

Determine your objectives. Is the intention to inform or update a group on recent activities? Is it to provide insight on a particular topic? Is it to change peoples understanding or decisions on an particular operational issue? Is it to engage with people to gather feedback or to take action?

Identify your target audience. What group or groups are you targeting with the above objectives? You will need to tailor the language(non technical experts may not be familiar with technical language), length (shorter messages may be more suitable for general public consumption) and style (different audiences have different lenses in which they will consume and interpret your message).

There are seven basic principles about writing about numbers: [2]

1. **Establish the context for your facts.** Your text should convey the "who, what, when and where" in which to ground your facts. Don't just assume that the audience has the same contextual understanding.

[1] For a deeper dive into technical writing, this free Google course is highly recommended
[2] Adapted from The Chicago Guide to Writing about Numbers, by Jane E. Miller

2. **Pick simple, plausible examples.** Using examples are a good way to transform abstract numbers to more tangible and relatable to the audiences experiences or understanding. An example of this could be used when describing density of the population of Rohingya refugees in Cox's Bazar, Bangladesh, where comparing the population number and area of the camp can be compared to that of a comparison city familiar to the audience.
3. **Select the right tools and media for the job.** The three basic tools for presenting quantitative information: prose, tables and charts. Choosing the most appropriate tool (or mix of them) and understanding their strengths and weaknesses, is important. Equally important is to use the most appropriate mix of media. Eg. Reports, interactive dashboards, infographics, video, social media, events.
4. **Defining your terms (and be careful with jargon).** Unnecessary use of of acronyms and jargon will likely exclude parts of your audience or cause misunderstanding due to unshared understanding of concepts. If acronyms must be used, it is good practice to show them alongside their long form at the point where they first appear.
5. **Reporting and interpreting.** Describing the numbers around an issue should be support by an explanation of "what does that mean" that explains why that number is important or relevant.
6. **Specify magnitude and direction of an association.** Don't just say "there are more displaced people in camp A than in camp B", provide a number quantifying *how* different it is. When explaining the relationship between variables it is also important to be clear on the direction of that relationship. For example "IDPs in Camp A had a lower number of food complaints compared to the previous month".
7. **Summarize patterns.** Rather than presenting a big table or graph showing the data and letting the viewer figure things out for themselves it is good to summarize and highlight patterns that contribute to the analysis and message.

> Tip
>
> - Tell a story
> - Choose hooks for your audience
> - Say it visually
> - Be transparent with the limitations of your analysis

5.2 Data Visualization

Communicating with visuals can an effective way to communicate a message, either on its own or alongside accompanying text. Good visuals can help engaging the audience and quite often are a good way to convey complex information in a simpler form.

5.2.1 Choosing the right charts

When visualising your data, the choice of chart depends on the quantity and type of data you want to represent; the relationships in that data, and ultimately, whether or not the graph clearly communicates your message.[3]

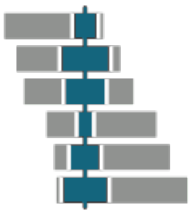
The following is pseudo-decision tree, to support choosing the most appropriate chart type depending on your data and it relationships.

5.2.1.1 Deviation

Emphasize variations (+/-) from a fixed reference point. Typically the reference point is zero but it can also be a target or a long-term average. Can also be used to show sentiment (positive/neutral/negative).

Examples: Showing the number of people entering or exiting a site over a period of time. Showing satisfaction with a component in a training. Demographics pyramid in a site, showing population breakdown by age and gender.

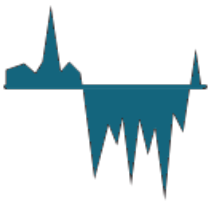
Diverging bar: A simple standard bar chart that can handle both negative and positive magnitude values.



Diverging bar: Perfect for presenting survey results which involve sentiment (eg disagree/neutral/agree).

Spine: Splits a single value into two contrasting components (eg male/female).

[3]Adapted from the FTs Visual Vocabulary. A similar graphics decision tree, based on the type and number of variables, is available at Data-to-Viz.com

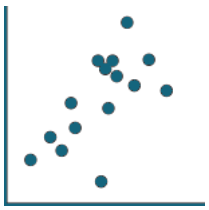


Surplus/deficit filled line: The shaded area of these charts allows a balance to be shown – either against a baseline or between two series.

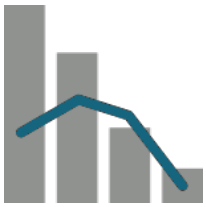
5.2.1.2 Correlation

Show the relationship between two or more variables. Be mindful that, unless you tell them otherwise, many readers will assume the relationships you show them to be causal (i.e. one causes the other).

Examples: Showing the relationships between areas of origin and current location of displacement.



Scatterplot: The standard way to show the relationship between two continuous variables, each of which has its own axis.



Column + line timeline: A good way of showing the relationship between an amount (columns) and a rate (line).

Connected scatterplot: Usually used to show how the relationship between 2 variables has changed over time.

Bubble: Like a scatterplot, but adds additional detail by sizing the circles according to a third variable.

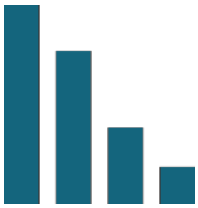
XY heatmap: A good way of showing the patterns between 2 categories of data, less effective at showing fine differences in amounts.

5.2.1.3 Ranking

Use where an item's position in an ordered list is more important than its absolute or relative value. Don't be afraid to highlight the points of interest.

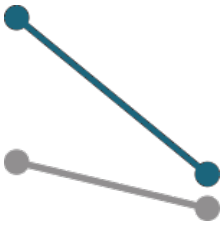
Examples: Comparing indicators of need. Comparing displacement population figures across sites or districts.

Histogram: Standard bar charts display the ranks of values much more easily when sorted into order..



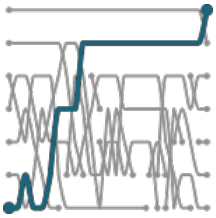
Ordered column: Same as above but more suited to categories of dates or with short labels.

Ordered proportional symbol: Use when there are big variations between values and/or seeing tne differences between data is not so important..



Slope: Perfect for showing how ranks have changed over time or vary between-categories.

Lollipop: Lollipops draw more attention to the data value than standard bar/column and can also show rank and value ef effectively.

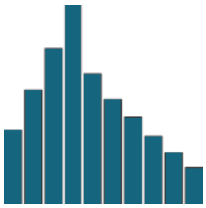


Bump: Effective for showing changing rankings across multiple dates. For large datasets,consider grouping lines using colour.

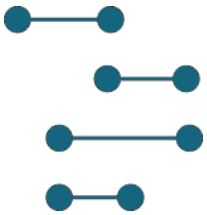
5.2.1.4 Distribution

Show values in a dataset and how often they occur. The shape (or 'skew') of a distribution can be a memorable way of highlighting the lack of uniformity or equality in the data.

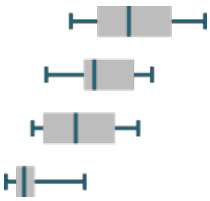
Examples:



Histogram: The standard way to show a statistical distribution - keep the gaps between columns small to highlight the 'shape' of the data.



Dot plot: A simple way of showing the change or range (min/max) of data across multiple categories.



Box plot: Summarise multiple distributions by showing the median (centre) and range of the data.

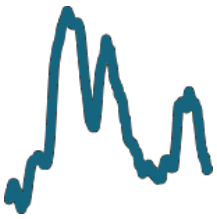
Population pyramid: A standard way for showing the age and sex breakdown of a population distribution; effectively, back to back histograms.

Beeswarm: Use to emphasise individual points in a distribution. Points can be sized to an additional variable. Best with medium sized datasets.

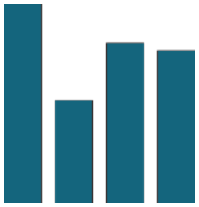
5.2.1.5 Change over time

Give emphasis to changing trends. These can be short (intra-day) movements or extended series traversing decades or centuries: Choosing the correct time period is important to provide suitable context for the reader.

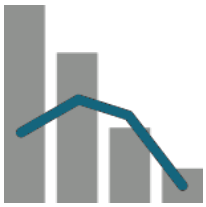
Examples:



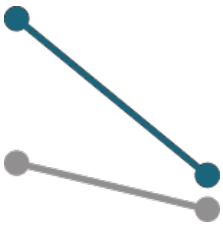
Line: The standard way to show a changing time series. If data are irregular, consider markers to represent data points.



Column: Columns work well for showing change over time - but usually best with only one series of data at a time.

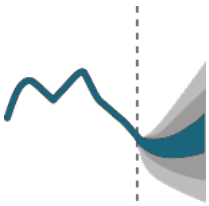


Column and timeline: A good way of showing the relationship over time between an amount (columns) and a rate (line).



Slope: Good for showing changing data as long as the data can be simplifed into 2 or 3 points without missing a key part of story.

Area chart: Use with care – these are good at showingchanges to total, but seeing change in components can be very difficult.

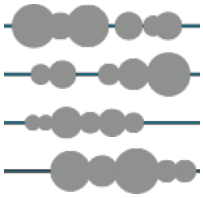


Fan chart (projections): Use to show the uncertainty in future projections - usually this grows the further forward to projection.

Connected scatterplot: A good way of showing changing data for two variables whenever there is a relatively clear pattern of progression.

Calendar heatmap: A great way of showing temporal patterns (daily, weekly, monthly) – at the expense of showing precision in quantity.

Priestley timeline: Great when date and duration are key elements of the story in the data.



Circle timeline: Good for showing discrete values of varying size across multiple categories (eg earthquakes by continent).

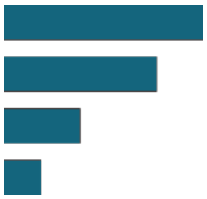
Streamgraph: A type of area chart; use when seeing changes in proportions over time is more important than individual values.

5.2.1.6 Magnitude

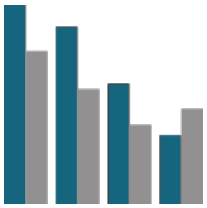
Show size comparisons. These can be relative (just being able to see larger/bigger) or absolute (need to see fine differences). Usually these show a 'counted' number (for example, barrels, dollars or people) rather than a calculated rate or per cent.

Examples:

Streamgraph: A type of area chart; use when seeing changes in proportions over time is more important than individual values.



Bar: See above. Good when the data are not time series and labels have long category names.

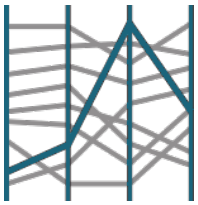


Paired column: As per standard column but allows for multiple series. Can become tricky to read with more than 2 series.

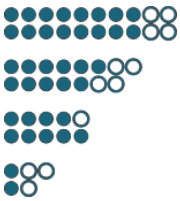
Marimekko: A good way of showing the size and proportion of data at the same time – as long as the data are not too complicated.

Isotype (pictogram): Excellent solution in some instances – use only with whole numbers (do not slice of an arm to represent a decimal).

Radar:A space-efficient way of showing value of multiple variables– but make sure they are organised in a way that makes sense to reader.



Parallel coordinates: A type of area chart; use when seeing changes in proportions over time is more important than individual values.



Grouped symbol: An alternative to bar/column charts when being able to count data or highlight individual elements is useful.

5.2.1.7 Part-to-whole

Show how a single entity can be broken down into its component elements. If the reader's interest is solely in the size of the components, consider a magnitude-type chart instead.

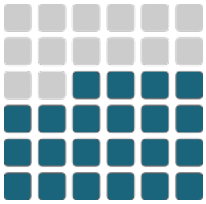
Examples:

Stacked column or bar: A simple way of showing part-to-whole relationships

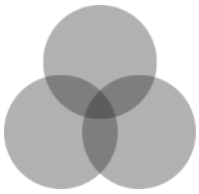
but can be difficult to read with more than a few components.

Radar: Similar to a pie chart – but the centre can be a good way of making space to include more information bout the data (eg total).

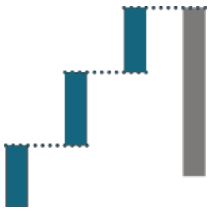
Treemap: Use for hierarchical part-to-whole relationships; can be difficult to read when there are many small segments.



Gridplot: Good for showing % information, they work best when used on whole numbers and work well in small multiple layout form.



Venn: Generally only used for schematic representation.



Waterfall: Can be useful for showing part-to-whole relationships where some of the components are negative.

5.2.1.8 Spatial

Aside from locator maps only used when precise locations or geographical patterns in data are more important to the reader than anything else.

Examples:

Basic choropleth: The standard approach for putting data on a map – should always be rates rather than totals and use a sensible base geography.

Proportional symbol: Use for totals rather than rates – be wary that small differences in data will be hard to see.

Flowmap: For showing unambiguous movement across a map.

Contour map: For showing areas of equal value on a map. Can use deviation colour schemes for showing +/- values

Dot density: Used to show the location of individual events/locations – make sure to annotate any patterns the reader should see.

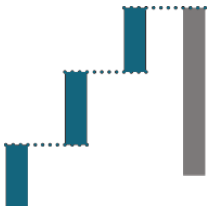
Heatmap: Can be useful for showing part-to-whole relationships where some of the components are negative.

5.2.1.9 Flow

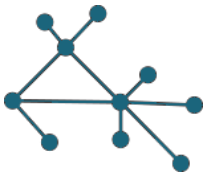
Show the reader volumes or intensity of movement between two or more states or conditions. These might be logical sequences or geographical locations.

Examples:

Sankey: Can be useful for showing part-to-whole relationships where some of the components are negative.



Waterfall: Can be useful for showing part-to-whole relationships where some of the components are negative.

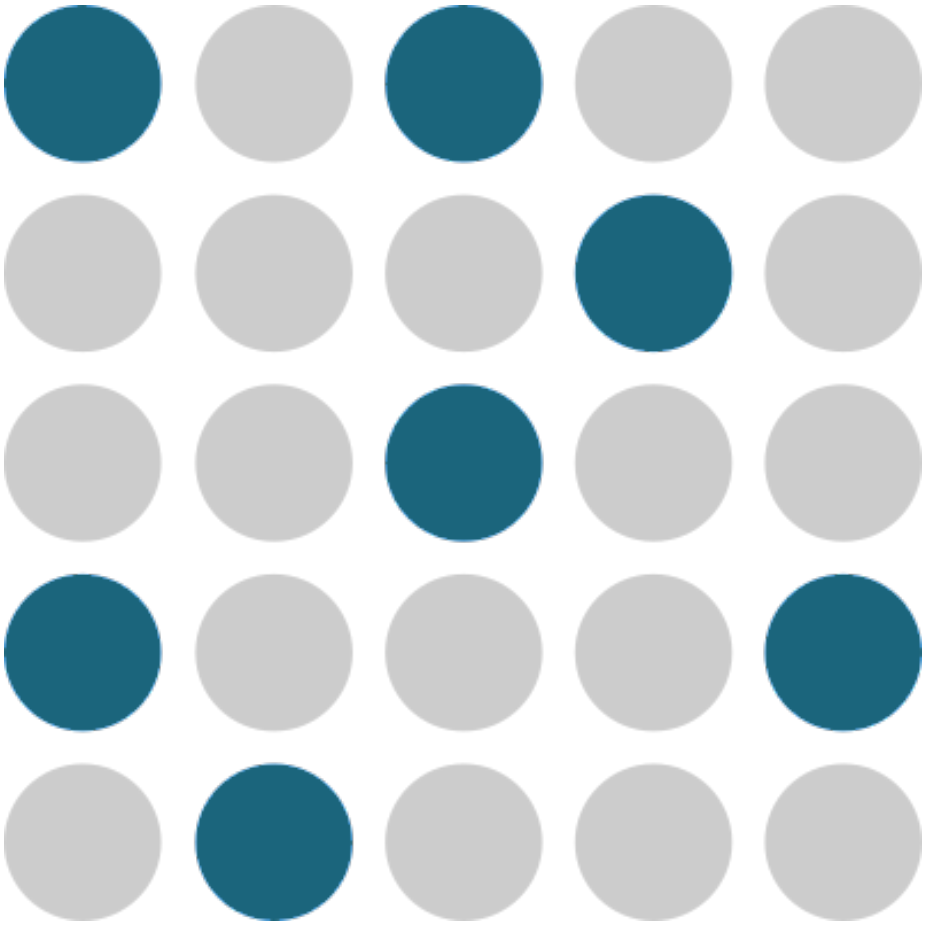


Network: Can be useful for showing part-to-whole relationships where some of the components are negative.

5.2.2 Visual Design Principles

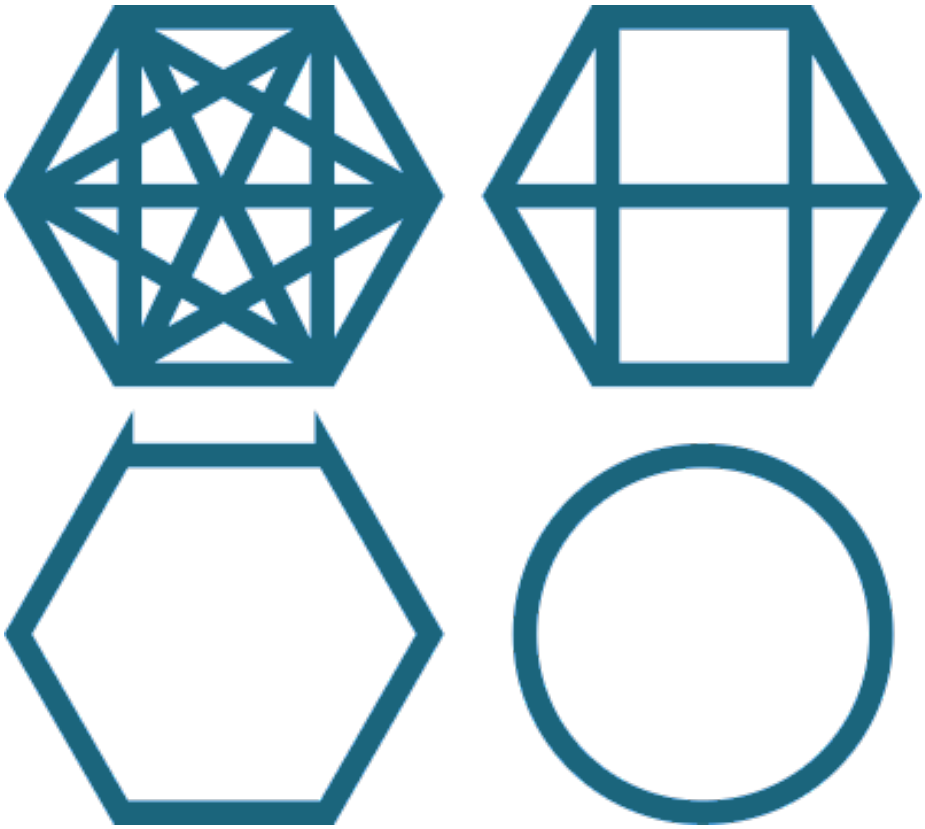
Developed by German psychologists, the Gestalt laws describe how we interpret the complex world around us. They explain why a series of flashing lights appear to be moving. And why we read a sentence like this, *notli ket his ort hat*. Understanding these "laws" can be useful in making sure your message is being conveyed effectively.

5.2.2.1 Law of Similarity



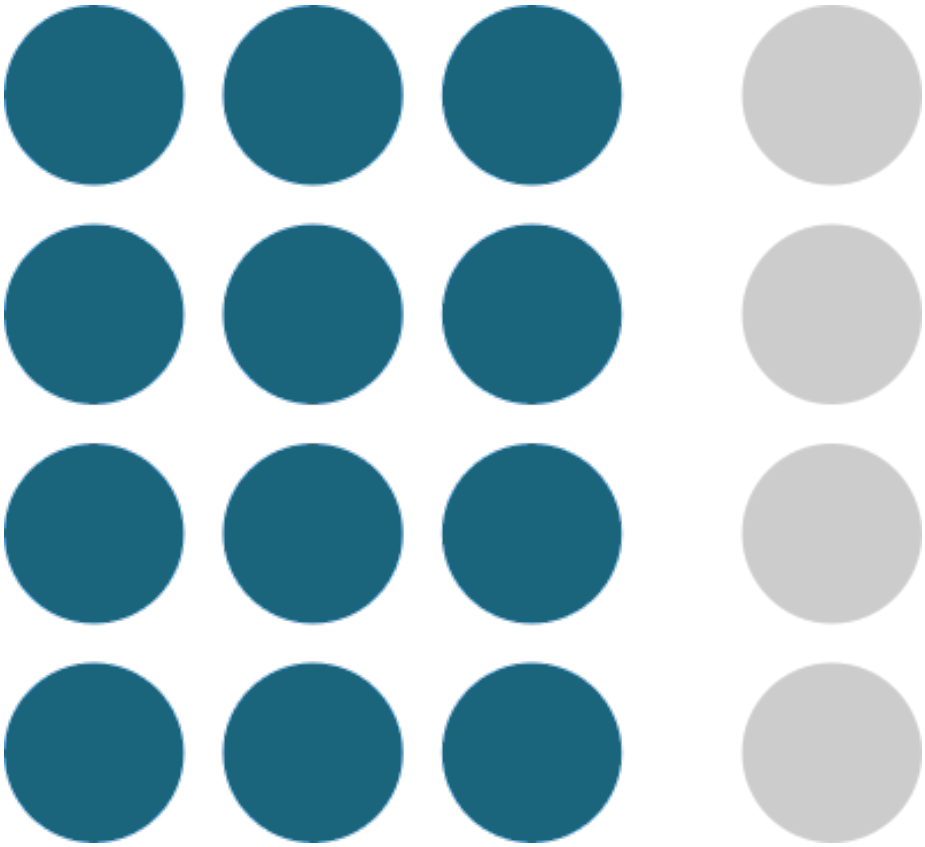
The human eye tends to perceive similar elements in a design as a complete picture, shape, or group, even if those elements are separated. Examples of this could be the use of symbols to signify conflict on a map or the use of colour in dots in a scatter plot that are of the same category.

5.2.2.2 Law of Prägnanz



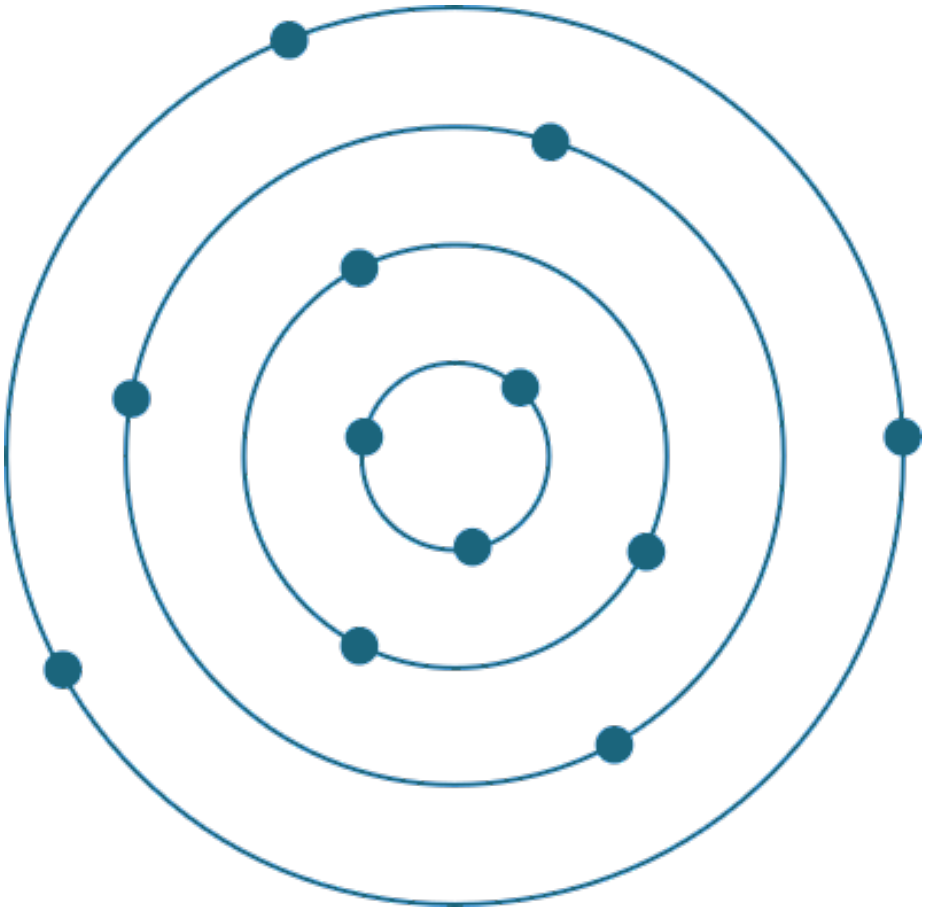
People will perceive and interpret ambiguous or complex images as the simplest form possible, because it is the interpretation that requires the least cognitive effort of us. Charts should aim to be as complex as necessary and as simple as possible to convey their meaning. Edward Tufte calls this the data-ink ratio.

5.2.2.3 Law of Proximity



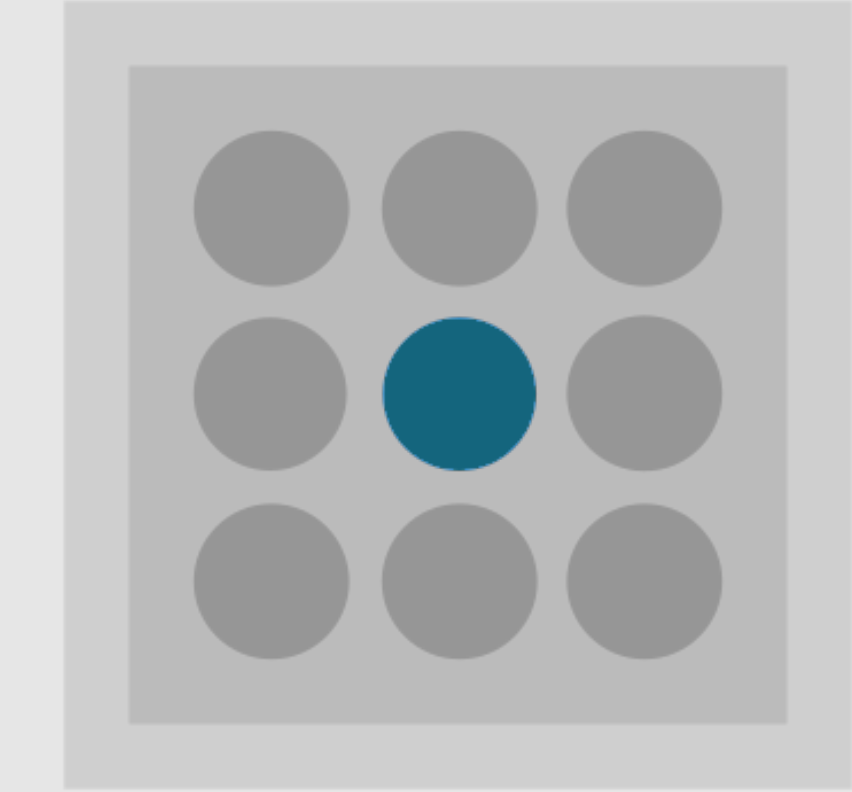
Objects that are near, or proximate to each other, tend to be grouped together. An example of this could be a grouped bar chart where for the funding for each year is grouped by donor.

5.2.2.4 Law of Continuity



Elements that are visually connected are perceived as more related than elements with no connection. This principle is visible when using a line graph to connect point values.

5.2.2.5 Law of Common Region



Elements tend to be perceived into groups if they are sharing an area with a clearly defined boundary. This law is perhaps most commonly used in maps, where administrative boundaries are shown with solid or dashed lines.

When presenting static charts a useful tip is to use annotation to guide the viewer through the graph, to put the data in context and to highlight the key relevant facts. [4]

5.2.3 Use of Colour

When choosing colours in your charts its important to understand possible local significance that may be associated to a specific colour. For instance, in one country a colour may signify good luck, whereas in a different country, the same colour could be associated with a non-state armed group.[5]

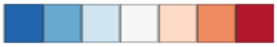
[4] Data journalism put increasing emphasis on the need for a good annotation layer, as can be seen by this article from the Financial Times

[5] ACAPS have a great guide on the Use of Colour in Data Display

Where possible, special attention should be taken to ensure that chart remain readable when printed in gray scale and that they are colour blind safe, meaning that the chart should not be confusing for people with red-green colour blindness (an estimated 8% of men and 0.4% of women).

Adding to the previous description of the role of color in perception, the use of colour in scales, particularly maps, typically takes one of the following three forms.[6]

1.  **Continuous(sequential) scales** used to show values going from low to high. Eg. population density per district.

2.  **Diverging scales** which visualize difference from a norm, such as this example showing location in St Vincent that showed both net inward and outward movements of people following the eruption of a volcano.

3.  **Categorical(qualitative) scales** used to distinguish different (non numeric) objects eg. a map using different shades for different countries.

Two of the most common ways to respresent colour are RGB and CMYK. RGB, commonly used on websites can be shown as a hex number or RGB number. For printed materials where colour accuracy is important, CYMK is typically used. Not all software supports the CMYK colour space, so if color accuracy is important you may want to use an Adobe tool such as Illustrator or In Design to apply finishing touches to print materials.[7]

5.3 Presenting

Having a great data collection system, doing great analysis and creating effective visuals don't necessarily lead to informing or changing decision by themselves. An important skill for IMOs that is often overlooked is the importance of verbal communication and presentation skills, be in in an in-person context such as a Cluster meeting or as is becoming more common, web-based calls. These meeting offer an important window of opportunity where, if communicated clearly and in a convincing manner a good analysis can meet it's objectives.

The following video is an example of effective communication, where the speaker shows a clear understanding of his audience(s), succinctly describes the context, the cause, the call for action (giving specific examples) and the urgency and scale.

When presenting slides, consider the following:

[6] Colorbrewer is a good resource for picking color palettes. Datawrapper have a good blog post describing the use of different colour scales.

[7] For a detailed explanation of RGB and CMYK and how they differ, see here

- **Only one idea per slide** Having multiple ideas presented will distract your audience and confuse your key message.
- **Explain your point, then show slide.** Your audience can interpret either the visuals on screen or your spoken message. It is very difficult to both at the same time.
- **Speaker is the star, not the slides.** The slides exist to aide the communication of the speaker, not to distract from it.
- **Never read from the slides.** It portrays a lack of preparedness and dilutes the communication rather than complimenting it.
- **Keep your hands free to move.** Not verbal expression can help the audience relate to the message and can help emphasise key messages.
- **Tell a story to drive home your message** Conveying your message through a narrative is a powerful way to introduce your audience with your key points, for them to engage with the topic and to remember it.
- **Use photos and drawings on slides.** Photos can help bring an emotive human element into otherwise abstract messages. Effective visuals can communicate concepts that would be much harder to explain through written or spoken word alone.
- **Face your audience, not your slides.** You are trying to convince, your audience, not the slides.
- **Avoid complexity.** Unnecessary complexity is a barrier for comprehension and can cause your audience to disengage with the topic.
- **Rehearse, rehearse, rehearse.**

Chapter 6

Data Responsibility

This chapter is in draft stage.

This section describes data security; data protection

Here's my note!

Outline - Main resources: IASC Data Responsibility in Humanitarian Action; ICRC Handbook on Data Protection in Humanitarian Action; IOM Data Protection Manual - Explain what is personal information - Examples of types of data in CCCM - Data protection principles

IASC Data Responsibility in Humanitarian Action

ICRC Handbook on Data Protection in Humanitarian Action

IOM Data Protection Manual

Chapter 7

Tools

7.1 test

The choice of software used to create CCCM products can be influenced by factors such as personal preference/familiarity, experience - some IMs may prefer more advanced tools, time constraints and budget. The following list of software is not a prescriptive list, rather a list of tools preferred by the authors

7.1.1 Microsoft Excel

.

7.1.2 QGIS/ArcGIS

.

7.1.3 Inkscape/Adobe Illustrator

.

7.1.4 Microsoft PowerBI

.

7.1.5 Microsoft Forms/Google Forms

.

7.1.6 Kobo toolbox

.

7.2 Advanced tools

.

Part II

CCCM Cluster IM

Chapter 8

The role of IM in the Cluster

> **Warning**
>
> Part 3 of the IM handbook is in a zero draft phase. Both the structure and content are likley to change significantly

> **Recommended reading**
>
> - Humanitarian Population Figures - IMWG
> - IASC Reference Module for the Implementation of the Humanitarian Programme Cycle- IASC
> - Responsibilities of Cluster/Sector Leads and OCHA in Information Management - IASC (2008)

The role of Information Management in the CCCM Cluster at country-level is best viewed through the lens of the **six core functions**:[1]

1. To support service delivery by:
   - Providing a platform that ensures service delivery is driven by the Humanitarian Response Plan and strategic priorities.
   - Developing mechanisms to eliminate duplication of service delivery.

The IM role supports service delivery by ensuring information systems are in place to understand where CCCM partners and activities are conducted, where there are gaps or overlaps and how those activities are meeting the overall objectives and targets decided upon by the cluster.

[1] From the IASC Reference Module for Cluster Coordination at Country Level, revised July 2015

2. To inform the HC/HCT's strategic decision-making by:
    - Preparing needs assessments and analysis of gaps (across and within clusters, using information management tools as needed) to inform the setting of priorities.
    - Identifying and finding solutions for (emerging) gaps, obstacles,duplication and cross-cutting issues.
    - Formulating priorities on the basis of analysis.

By collecting and analysing CCCM needs to inform; an understanding of gaps; barriers for the response; where and whom requires prioritization.

3. To plan and implement cluster strategies by:
    - Developing sectoral plans, objectives and indicators that directly support realization of the overall response's strategic objectives.
    - Applying and adhering to common standards and guidelines.
    - Clarifying funding requirements, helping to set priorities, and agreeing cluster contributions to the HC's overall humanitarian funding proposals.

To support the setting of measurable objectives and indicators, the categorization of activities and cost estimations for these activities or projects.

4. To monitor and evaluate performance by:
    - Monitoring and reporting on activities and needs.
    - Measuring progress against the cluster strategy and agreed results.
    - Recommending corrective action where necessary.

To manage systems that monitor collective progress against the set targets and objected.

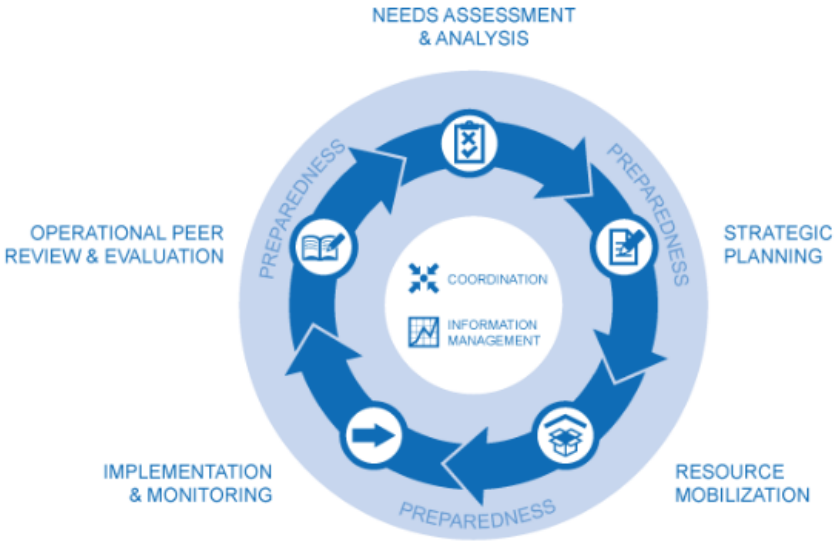
5. To build national capacity in preparedness and contingency planning.

To disseminate the above skills, systems and processes to national and local levels to ensure national-level preparedness on systems and skills relevant to CCCM.

6. To support robust advocacy by:
    - Identifying concerns, and contributing key information and messages to HC and HCT messaging and action.
    - Undertaking advocacy on behalf of the cluster, cluster members, and affected people.

Evidence-based advocacy, to inform strategic decision making and to represent the challenges, needs and achievements of the cluster.

8.1 The Humanitarian Program Cycle



The humanitarian programme cycle (HPC) is a coordinated series of actions undertaken to help prepare for, manage and deliver humanitarian response. It consists of five elements coordinated in a seamless manner, with one step logically building on the previous and leading to the next. Successful implementation of the humanitarian programme cycle is dependent on effective emergency preparedness, effective coordination with national/local authorities and humanitarian actors, and information management. [2]

8.2 Cross-cutting themes

8.2.1 Data Responsibility

8.2.2 Knowledge Management

8.2.3 Preparedness

8.3 Useful Conceptual Aids

8.3.1 The Onion Model of Population Categories

8.3.2 Profile Framework / CCCM Site Typology

[2] More details on HumanitarianResponse.info

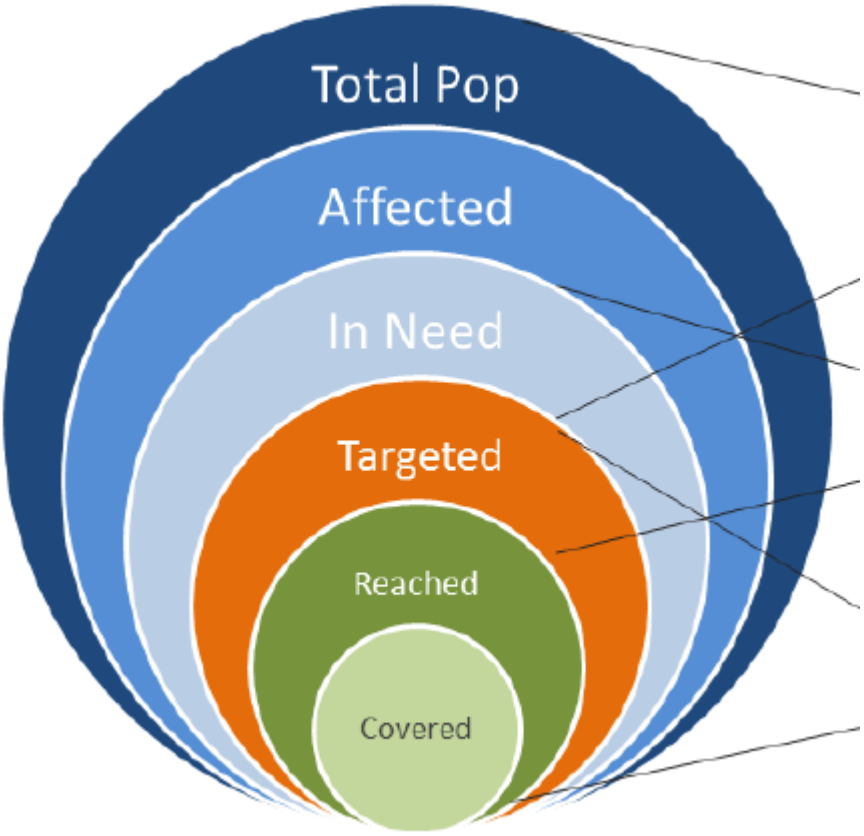


Figure 8.1: Onion Model, Source: IASC

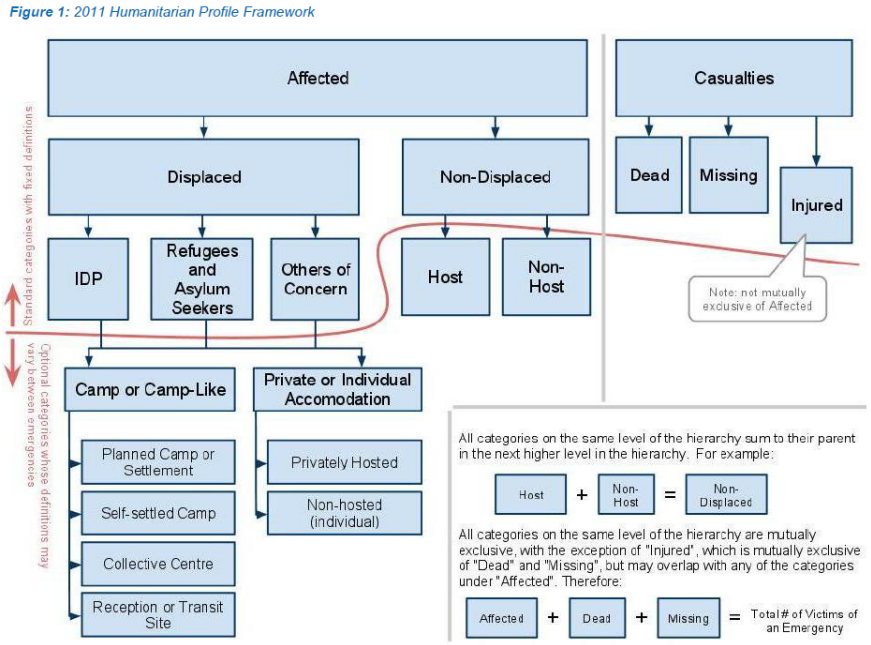


Figure 8.2: Humanitarian Profile Framework, Source: IASC

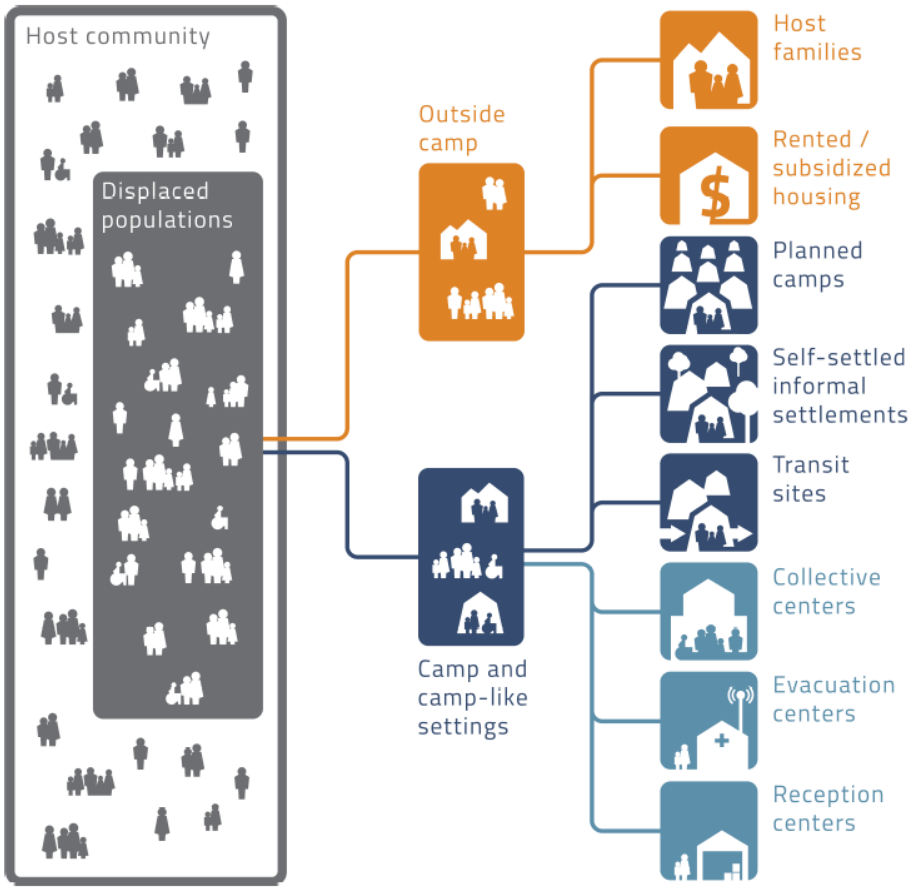


Figure 8.3: CCCM Site TYpology, Source: Minimum Standards for Camp Management

Chapter 9

Needs Assessment & Analysis

Recommended reading

- Humanitarian Needs Assessment: The Good Enough Guide - ACAPS
- JIAF 1.1 Guidance

9.1 CCCM Sectoral analysis

a

9.1.1 Setting your needs analysis parameters

a

9.1.2 People in Need calculation

a

9.1.3 Severity of needs in CCCM

a

9.2 Inter-sectoral analysis

a

9.3 Humanitarian Needs Overview & Flash Appeals

a

Chapter 10

Strategic Planning

10.1 Strategic Objectives

.

10.2 Indicators

.

10.3 Activities

.

10.4 Targets

10.5 Humanitarian Response Plan

.

Chapter 11

Resource Mobilization

11.1 Costing

.

Chapter 12

Implementation & Monitoring

Recommended reading

- Measuring and aggregating population figures for planning and monitoring - OCHA
- Minimum Standards for Camp Management MSCM

12.1 People Reached

.

12.2 People Covered

.

Chapter 13

Operational Peer Review & Evaluation

13.1 Cluster Coordination Performance Monitoring

.

Part III

CCCM Operations IM

Chapter 14

IM in CCCM Operations

This part of the handbook will be drafted following the completion of the Part 1 - Humanitarian IM and Part 2 - CCCM Cluster IM of the handbook