Centre for Humanitarian Data - Data Literacy Team

Data Foundations Quiz

Data Terminology

1. What is the difference between incidence and prevalence?
2. What is mean, median and mode? Please describe a situation in which median would be a better measure of central tendency than mean.

Data Responsibility

Data responsibility: a set of principles, processes and tools that support the safe, ethical and effective management of data in humanitarian response.

1. Please list the documents related to data responsibility that apply to your work or your office's work.
2. How would you go about seeking guidance if you found someone gained unauthorised access to your office's sensitive data?

Planning

1. Your head of office has a request from a government authority for the number of IDPs, and needs these for a meeting in two weeks. What are 4 probing questions you would ask your head of office to inform the data you collect?
2. Your head of office wants to know the number of people affected in a recent flood. Before you start data collection, what are four main factors you should consider?
3. What data would you need to report on the following indicator to be used in a Humanitarian Response Plan? *Percentage of IDPs that receive NFI kits, Shelter kits and Camp safety messages.*
4. How would you improve the indicator above?

Collecting

1. Create an Excel Spreadsheet to collect the following information from people that attend a general coordination meeting for sector focal points from all organizations - organization, focal point, sector, contacts, location. No need to enter data, as the users will enter this into the Spreadsheet when they attend the meeting. Note that this spreadsheet will be used to report to your head of office the diversity and good attendance of the meetings. Standardize the data entry for users as much as possible. Save the file with your name and collection form and either send to peterson1@un.org or insert a link to the file below.
2. What are the five key pieces of metadata you like would to know about this dataset [Insert Link]?

Processing

1. Identify the errors/anomalies in this dataset [Insert Data].
2. Compare the two datasets here [Insert Data] and identify the differences for triangulating the numbers.
3. You have received a data file by a flashdisk and have put it on your hard drive. You open the file and see that it needs to be cleaned. Before starting to clean the data, what should you do to the file?

Analysis

1. What entries would you question in the dataset and why?
2. If you average the number of NFIs, with the number of NFIs distributed at Al-Ikhlas as a "-" and then again with the number as a "0", what is the difference?
3. What is the range of the dataset excluding the outliers?
4. Round the individual numbers to the 10 and sum them. What is the difference between this total and the total without rounding?
5. Make a bar chart with the two datasets at this link [Insert Data]. Save the file with your name and barchart and either send to peterson1@un.org or insert a link to the file below.

Communicating

1. What are three points that stand out to you in the barchart you created above?