# AEM Data Management

*A Data Management Plan created using DMPTool* [*https://dmptool.org/plans/41445*](https://dmptool.org/plans/41445)

Creator: Marissa Marinello

Affiliation: University of Texas at Austin (UT)

Template: Digital Curation Centre (DCC)

Project abstract:

How does the ongoing conflict in Syria impact the Syrian population and neighboring countries? What is the relationship between the severity of the conflict in Syria and the changes in the Syrian refugee population both internally displaced in Syria but mostly outside of Syria into Turkey, Jordan, and Lebanon? How do these populations change over time? (Answered in THA3) What is the relationship between the length of time the conflict continues and the amount of aid provided? Are there changes in the amount of aid over time? (Guess: as the conflict prolongs, aid isn't able to keep up in a sustainable way so it falls off) Do certain kinds of violence lead to increases in aid (attacks against civilians vs attacks against military targets) How long do the aid increases last for? (This can be plotted - violent attacks against amount of aid donated) How does the risk of starvation/the deliberate use of starvation as a method of war interact with the severity of the conflict and the delivery of aid?

### Data Collection

**What data will you collect or create?**

- data that analyzes the ongoing armed conflict in Syria from the begining of the conflict up until now

- data on refugee populations leaving Syria and where they go (Turkey, Lebano, Jordan, or internally displaced)

- data on funding of humanitarian relief efforts and the delivery of said relief to the civilian population in Syria and to host communities supporting large numbers of refugees

- data on food insecurity, starvation, and malnutrition in Syria

- the majority of this data is in CSV or excel format, and includes high volumes of data observations and variables

- data will be analyzed using Python and provided through a repository on github

**How will the data be collected or created?**

- the data collected is all secondary data

- I will combine parts of the data sets for analysis (armed conflict and refugee population changes; armed conflict and humanitarian relief; armed conflict and food insecurity)

- maintain organized codebooks in python and notes and rationale for running codes

- store all data in github and on computer with the same naming (all lowercase, \_ instead of spaces)

### Documentation and Metadata

**What documentation and metadata will accompany the data?**

- some of the datasets come with codebooks or PDF's explaining methodology in research

- I will keep a running data dictionary with variable definitions

- document how data sources were retrieved and provide as much detail as possible as well as links to the data source

- provide rationale for why data was selected

- document my methodology for analaysis

- describe procedural information for running code and analyzing

- describe units of measure

- document any and all assumptions I have so that everything that could be implicit knowledge is made explicit

### Ethics and Legal Compliance

**How will you manage any ethical issues?**

- I want to include a section of ethics and data useage in my final write up; its hard in secondary data analysis to ensure ethics because I am not the one collecting the data - it will be useful to look up the methodolgy of data collection for my data sources to see their ethical standards

- ensure that I obtain any consent for data usage that is necessary

- ensure that I am able to share the data that I use

- I'm not working with any sensitive data at this point so that makes things a bit easier

**How will you manage copyright and Intellectual Property Rights (IP/IPR) issues?**

- include notes from data sources discussing copyright or intellecutal property rights

- do research to make sure that I'm able to use the data I have

### Storage and Backup

**How will the data be stored and backed up during the research?**

- all data will be stored on my computer and on github

- documents will be stored on google drive and on my computer so it is accessible and sharable and changes can be easily tracked

- codebooks will be stored on googe drive and github

**How will you manage access and security?**

N/A

### Selection and Preservation

**Which data are of long-term value and should be retained, shared, and/or preserved?**

- I will preserve my data because I will likely use this data again in the future and others may want it as well.

**What is the long-term preservation plan for the dataset?**

- I will keep this stored long term on github and back it up on an external hard drive.

### Data Sharing

**How will you share the data?**

- I want this data to be very sharable; I need to develop ideas about how to go through with data sharing.

**Are any restrictions on data sharing required?**

- ACLED data has some restrictions that need to be kept in mind.

### Responsibilities and Resources

**Who will be responsible for data management?**

**What resources will you require to deliver your plan?**