**Notes on Meeting with Albert Kettner**

**Friday, October 27, 2023**

# Miscellaneous

* Recommended that we check out [EM-DAT](https://www.emdat.be/)
  + We will need to request access, but might be worth checking out
* Social media mining (WhatsApp + Twitter): [FloodTags](https://www.floodtags.com/)
* Would love to see where our project goes, maybe we can send him something (i.e. a link or report) at the end of the semester

# Information and Background on the Flood Archive

* The Archive was started in the 80s, based on paper sources (BBC, NY Times, etc.)
* The indicators that are part of the data (people impacted, fatalities, etc.) are all based on news sources
  + But there’s a time gap – sometimes flooding takes 2 days, sometimes it takes 2 months – so we might see multiple stories over a period of time about the same flood
  + The flood could have been recorded in the earlier days, and then the numbers could have changed over time, but the Archive wasn’t updated to reflect that
  + Essentially, the archive is based on however well news outlets are informed of each specific flood event
* Data is gathered pretty manually, they do not have any automated process to make it quicker at this moment in time
  + An issue with this is language barriers (when news articles are written in a foreign language)
* Some of the numbers are ballparks and using the shape file drawing that they create of flood events
* Data is updated every few weeks, whenever time allows for the researchers
  + No specific schedule for data updating

# Albert’s Suggestions for Our Project

* Focus on 3k rows that have a severity score first: try to predict how the Flood Archive calculated its severity score using data from our other dataset
* Then, use that prediction model to find severity scores for the 6k rows that the Flood Archive doesn’t already have severity scores for
* The slight problem with this idea: we don’t have much information from the other dataset that could help us predict the flood severity, which is why we turned to the Flood Archive in the first place