

# ADA Summary: Week 1 Jan 17-23

Peem Lerdputtipongporn

20<sup>th</sup> January 2022

Below are what I have done during Week 1 of the ADA project.

## 1 What I have done

- Update basic libraries (e.g., Pandas, numpy, matplotlib) and ensure they can be used on my Jupyter notebook.
- Create Github repository with initial dataset.
- Perform EDA on each variable and jot down questions I have. Overall, the variables are self-explanatory but there are some weird inputs (see Jupyter notebook)
- Compile a list of questions about the dataset to send to our external collaborators (expected to do by Thursday evening).

## 2 Updates on last week's plan

Not applicable.

## 3 Plans for next week

- Read two of Rema's papers on video classification as well as initial-meeting notes. Identify the project's motivation, dataset, analysis methods, results, and future extensions and jot them down (Deadline: before we meet)
- Refresh myself on basic data manipulation with pandas, numpy, sklearn (Ongoing).
- Perform EDA on other variables. Compile any surprising results to bring up to our collaborators (Ongoing).

## 4 Questions to discuss when we meet

- Python or R? What should I spend time becoming more fluent in. It has been a while since I last used Python, so I'm having some difficulty with EDA, especially how to visualize the dataset. The dataset is also quite messy.
- Set up expectations and regular meeting time and frequency. How should I update you on my progress? When should I contact you (also how) if I encounter problems I cannot resolve.
- When to schedule a meeting with Rema and her colleagues (i.e., until which point have we made enough progress to contact them)? Source code for previous projects they did on this dataset.

- Larry?