Alcoholic Anonymous Report

The Hindu Newspaper is second most circulated English-language newspaper in India. Since 2016 Their website publishing the articles on Alcoholic Anonymous which is approximately (~11,825) and in which Meetings Schedules conducted by Alcoholic Anonymous articles is approximately (~10,615). The Hindu Newspaper is my primary choice of resource for the Data Extraction part.

The tools used for data extraction as follows

1. BeautifulSoup
2. Pandas
3. [News-Please](https://github.com/fhamborg/news-please)

News-Please is an open source, easy-to-use news crawler that extracts structured information from almost any news website. It can follow recursively internal hyperlinks and read RSS feeds to fetch both most recent and also old, archived articles. You only need to provide the root URL of the news website to crawl it completely. news-please combines the power of multiple state-of-the-art libraries and tools, such as [scrapy](https://scrapy.org/), [Newspaper](https://github.com/codelucas/newspaper), and [readability](https://github.com/buriy/python-readability). news-please also features a library mode, which allows Python developers to use the crawling and extraction functionality within their own program. Moreover, news-please allows to conveniently [crawl and extract articles](https://github.com/fhamborg/news-please/blob/master/newsplease/examples/commoncrawl.py) from commoncrawl.org.

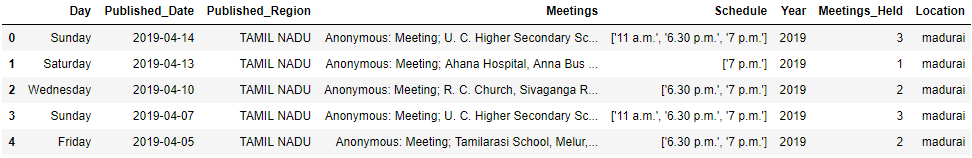
Extracted information

* headline
* lead paragraph
* main content (textual)
* main image
* author's name
* publication date
* language

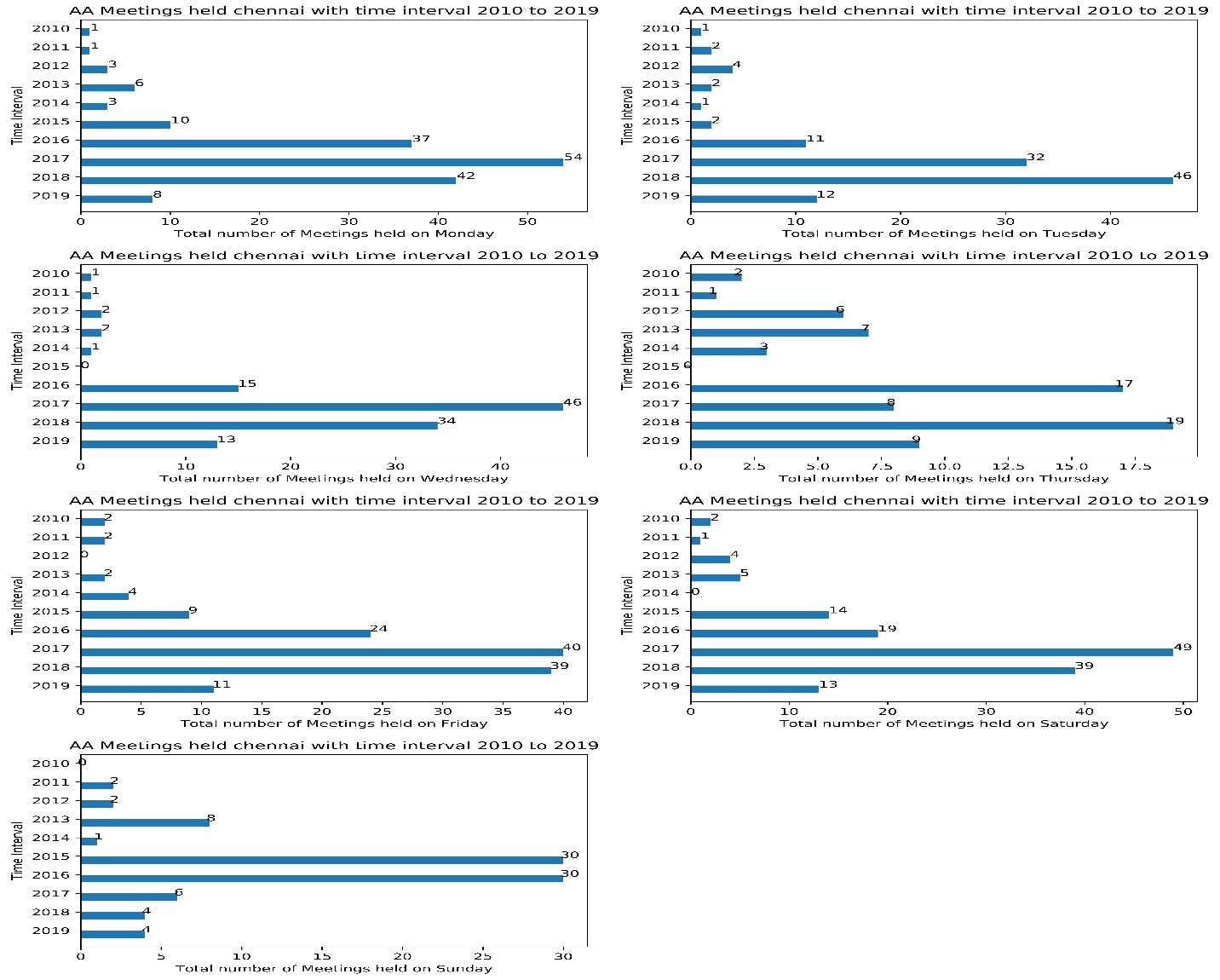
Dependencies Installation

* pip3 install news-please
* pip install beautifulsoup4

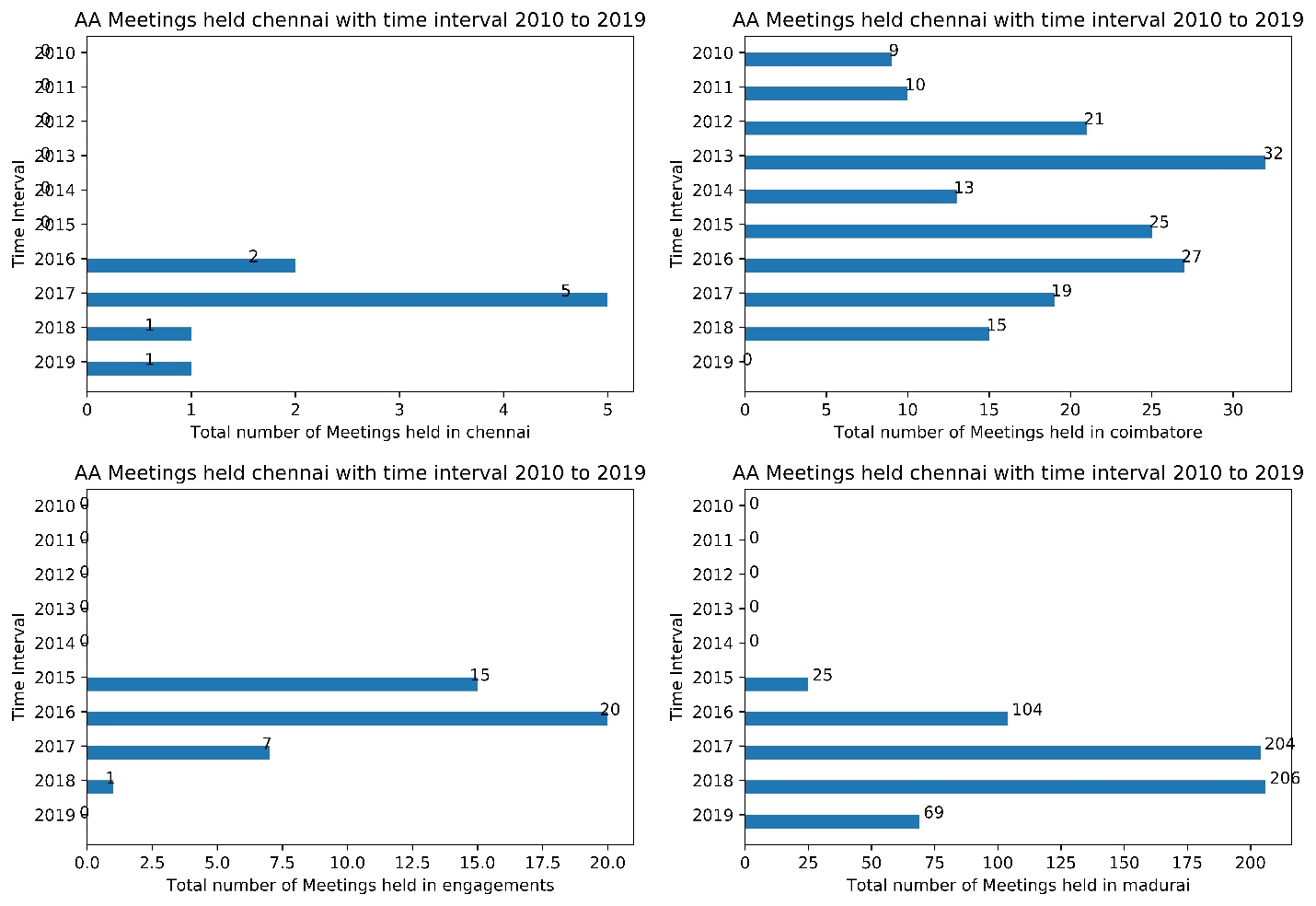
Getting Started

The collected data undergoes data pre-processing and filtering features needed for the Analysing the Alcoholic Anonymous Meetings. The Final Data will like below

Categorizing the Data in terms of time intervals 2016 – 2019. During these three years theirs a continuous decrease in schedule of AA Meetings.

Visualizing total number of meetings scheduled which decreased continuously during weekdays.

Visualizing total number of meetings scheduled which decreased continuously during locations.



Summary:

I don’t find any reason for fall of AA meetups. If we are able gather the main cause behind the continuous fall on Alcoholic Anonymous meetups during these three years. We can come to a conclusion the about the AA treatment in India.

Code and Dataset

* AA\_Meetings\_Report.ipynb (Above Visualization Code)
* Filtering\_AA\_Meetings\_CSV.ipynb (Code for filtering the features required for Data Visualization)
* Meeting\_Schedule.csv (Dataset used for filtering)
* AA\_Meets.csv (Final Dataset)