**Instructions:-**

1. Inside SpiderProject-->ScrapInternship-->ScrapInternship , there are 3 main files:-

spider🡪 Intern\_spider = contains sracping code where a scraping spider has been created

pipelines = data fetched by spider is send to pipeline to be stored.

setting = contain different settings of spider bot

1. All the Scraped data is in Scraped\_data folder, url\_id = 44, 57, 144 are not accessible hence no data has been extracted for them.
2. “text\_analysis\_internship.ipynb” contains code for calculating output variables
3. Avg. Sentence length and Avg. Words per Sentence are same as I’m unable to spot the difference.