The scrapy file toi_spiders.py scrapes the data for one year from Times of India and saves it in a csv. To modify it for a specific year, do the following:

Let's say we want to scrape the data for 2017.

Go to : https://timesofindia.indiatimes.com/archive.cms

The following is displayed:

Then click on January under 2017

The following window opens up

Click on 1 in the calendar.

The following screen shows:
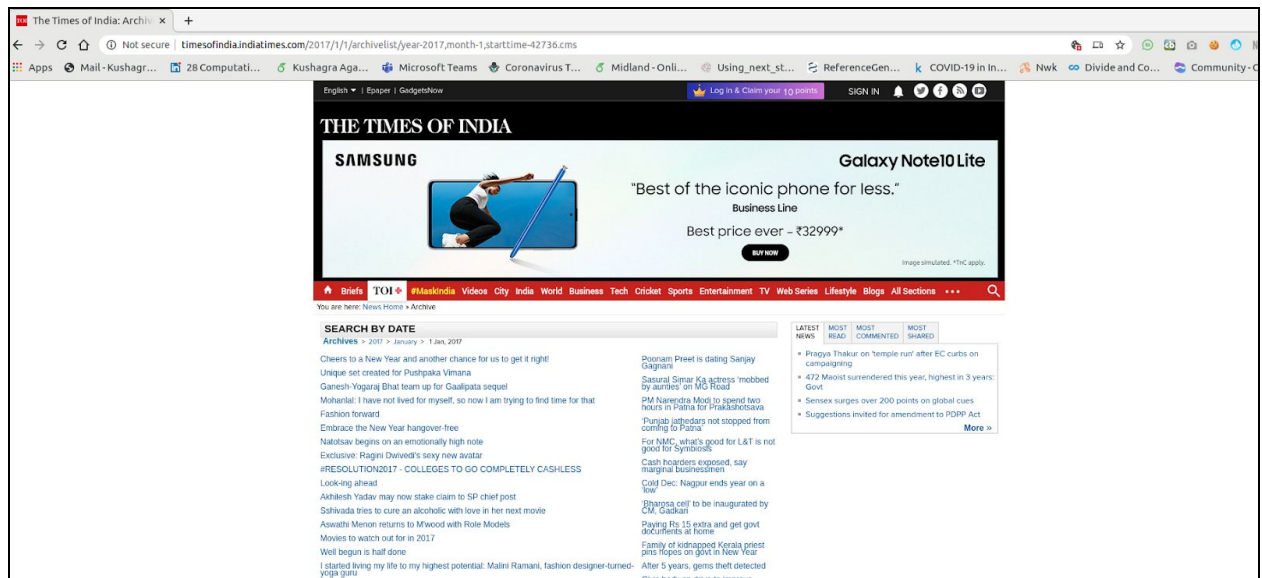
Copy the address of the page:
https://timesofindia.indiatimes.com/2017/1/1/archivelist/year-2017,month-1,starttime-42736.cms

Paste this link in the start_urls in the toi_spider.py file.

```python
import scrapy

class Spider(scrapy.Spider):
  name = "toi_spider"
  start_urls = ["https://timesofindia.indiatimes.com/2009/1/1/archivelist/year-2009,month-1,starttime-39814.cms"]

  def numberofdays(self, month, year):

    x = int(month)
    y = int(year)
    numberdays = 0
    if(x == 1 or x == 3 or x == 5 or x == 7 or x == 8 or x == 10 or x == 12) :
      numberdays = 31
    elif (x == 2):
      if (y % 4 == 0):
        numberdays = 29
      else:
        numberdays = 28
    else:
      numberdays = 30
    return numberdays
```

How to Run the scrapy file?

Install Scrapy: **pip3 install scrapy**
StartProject: **scrapy startproject isb**
Move: **cd isb/isb/spiders**
Paste the toi_spiders.py file in this folder.
Run: **scrapy crawl toi_spider -o out.csv**