**#https://github.com/mGalarnyk/Python\_Tutorials/tree/master/Scrapy**

**#https://www.youtube.com/watch?v=O\_j3OTXw2\_E**

1. Open a terminal (mac/linux) or command line (windows). Navigate to a desired folder (see the image below if you need help) and type

scrapy startproject fundrazr

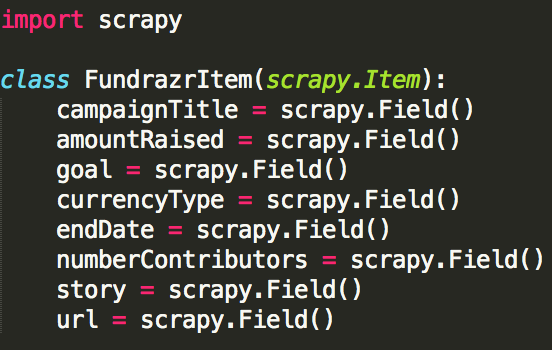
## 2. Itemss

The main goal in scraping is to extract structured data from unstructured sources, typically, web pages. Scrapy spiders can return the extracted data as Python dicts. While convenient and familiar, Python dicts lack structure: it is easy to make a typo in a field name or return inconsistent data, especially in a larger project with many spiders.

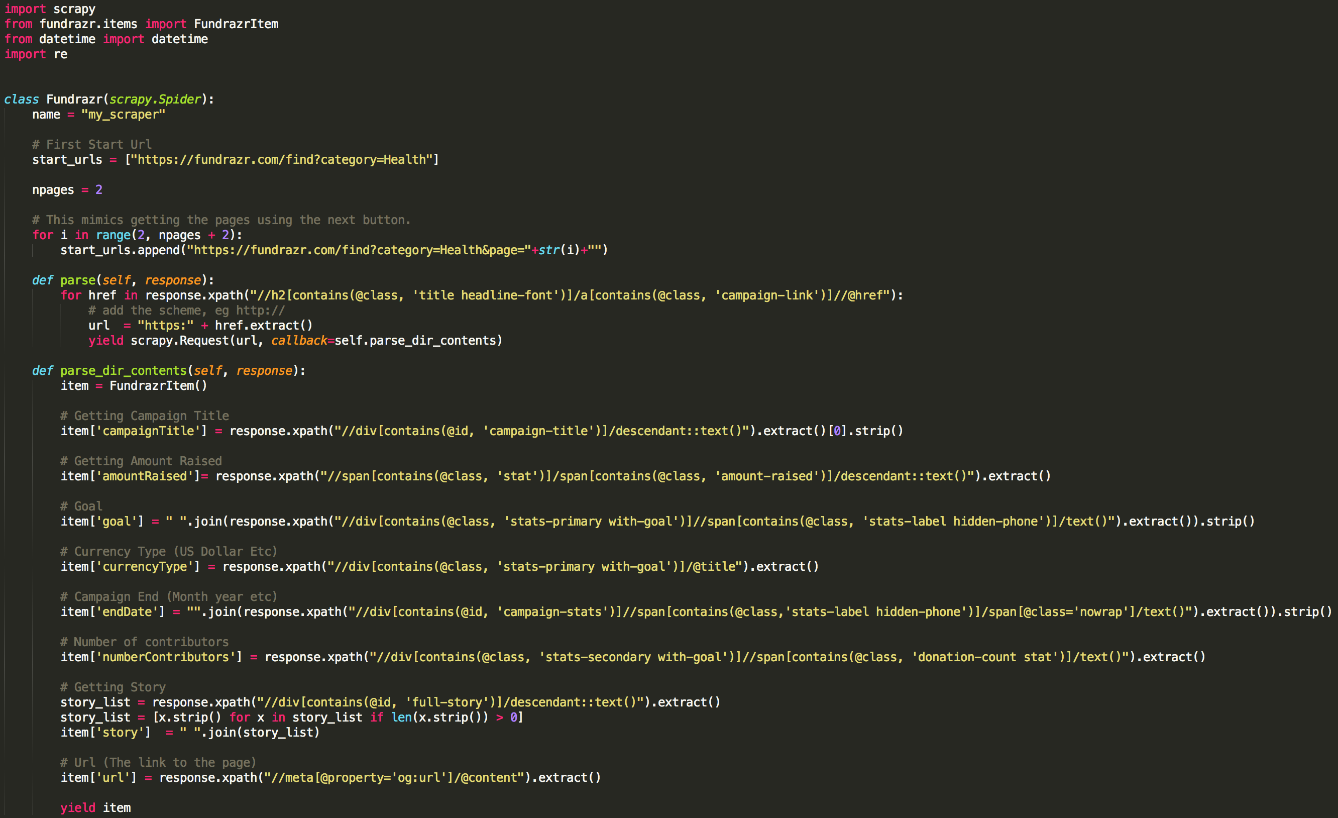
The code for items.py is [here](https://github.com/mGalarnyk/Python_Tutorials/raw/master/Scrapy/fundrazr/fundrazr/items.py).

Save it under the fundrazr/fundrazr directory (overwrite the original iems.py file).

The item class (basically how we store our data before outputting it) used in this tutorial looks like this.

[](https://github.com/mGalarnyk/Python_Tutorials/blob/master/Scrapy/Tutorial_Images/itemsFundrazr.png)

3. Spiders are classes that you define and that Scrapy uses to scrape information from a website (or a group of websites). The code for our spider is below.

[](https://github.com/mGalarnyk/Python_Tutorials/blob/master/Scrapy/Tutorial_Images/fundrazrScrapyCode.png)

Download the code [here](https://raw.githubusercontent.com/mGalarnyk/Python_Tutorials/master/Scrapy/fundrazr/fundrazr/spiders/fundrazr_scrape.py).

Save it in a file named **fundrazr\_scrape.py** under the fundrazr/spiders directory.

4. go to the fundrazr/fundrazr directory and open cmd and type:

scrapy crawl my\_scraper -o MonthDay\_Year.csv

5. If you want to download a larger file (it was made by changing npages = 2 to npages = 450 and adding download\_delay = 2), you can download a bigger file with roughly 6000 campaigns scraped