Scrapy startproject project

Cd project

Create spider.py under spider

/html/body/div[1]

/html/head/title

Splash

docker pull scrapinghub/splash

docker run -it -p 8050:8050 --rm scrapinghub/splash

Debug

Create runner.py

from scrapy.cmdline import execute  
  
try:  
 execute(  
 [  
 'scrapy',  
 'crawl',  
 'kb'  
 ]  
 )  
except SystemExit:  
 pass