

# How to set different scrapy-settings for different spiders?

Asked 11 years, 5 months ago

Modified 8 years ago

Viewed 15k times

I want to enable some http-proxy for some spiders, and disable them for other spiders.

20

Can I do something like this?

```
# settings.py
proxy_spiders = ['a1' , b2']

if spider in proxy_spider: #how to get spider name ???
    HTTP_PROXY = 'http://127.0.0.1:8123'
    DOWNLOADER_MIDDLEWARES = {
        'myproject.middlewares.RandomUserAgentMiddleware': 400,
        'myproject.middlewares.ProxyMiddleware': 410,
        'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None
    }
else:
    DOWNLOADER_MIDDLEWARES = {
        'myproject.middlewares.RandomUserAgentMiddleware': 400,
        'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None
    }
```

If the code above doesn't work, is there any other suggestion?

scrapy

Share Improve this question Follow

edited Oct 12, 2013 at 3:48

JavaNoScript

2,413 23 27

asked Oct 11, 2013 at 21:17

Michael Nguyen

1,752 2 19 35

## 5 Answers

Sorted by: Highest score (default)

a bit late, but since release 1.0.0 there is a new feature in scrapy where you can override settings per spider like this:

38

```
class MySpider(scrapy.Spider):
    name = "my_spider"
    custom_settings = {"HTTP_PROXY": 'http://127.0.0.1:8123',
                       "DOWNLOADER_MIDDLEWARES": {'myproject.middlewares.RandomUserAgentMiddleware': 400,
                                                  'myproject.middlewares.ProxyMiddleware': 410,
                                                  'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware':
None}}

class MySpider2(scrapy.Spider):
    name = "my_spider2"
    custom_settings = {"DOWNLOADER_MIDDLEWARES": {'myproject.middlewares.RandomUserAgentMiddleware': 400,

'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None}}
```

Share Improve this answer Follow

answered Dec 11, 2015 at 18:28

user4055746

399 3 5

I get name error "AUTOTHROTTLE\_ENABLED" not defined . Your snippet does not show that you're import settings.py  
– Homunculus Reticulli Nov 14, 2017 at 14:40



There is a new and easier way to do this.

15



```
class MySpider(scrapy.Spider):
    name = 'myspider'

    custom_settings = {
        'SOME_SETTING': 'some value',
    }
```



I use Scrapy 1.3.1

Share Improve this answer Follow



answered Mar 11, 2017 at 18:40



Aminah Nuraini 19.2k 9 97 113



You can add setting.overrides within the spider.py file Example that works:

8

```
from scrapy.conf import settings

settings.overrides['DOWNLOAD_TIMEOUT'] = 300
```



For you, something like this should also work



```
from scrapy.conf import settings

settings.overrides['DOWNLOADER_MIDDLEWARES'] = {
    'myproject.middlewares.RandomUserAgentMiddleware': 400,
    'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None
}
```

Share Improve this answer Follow

edited Sep 15, 2016 at 1:55



Aminah Nuraini 19.2k 9 97 113

answered Aug 28, 2014 at 22:43



Ricky Sahu 24.4k 5 44 32

3 Don't forget this before do that from scrapy.conf import settings – Aminah Nuraini Oct 20, 2015 at 10:38

2 settings.overrides has been deprecated in Scrapy versions greater than 1. Using custom\_settings dictionary in your spider declaration works. – v01d Jan 23, 2016 at 8:33



You can define your own proxy middleware, something straightforward like this:

4

```
from scrapy.contrib.downloadermiddleware import HttpProxyMiddleware

class ConditionalProxyMiddleware(HttpProxyMiddleware):
    def process_request(self, request, spider):
        if getattr(spider, 'use_proxy', None):
            return super(ConditionalProxyMiddleware, self).process_request(request, spider)
```



Then define the attribute use\_proxy = True in the spiders that you want to have the proxy enabled. Don't forget to disable the default proxy middleware and enable your modified one.



Share Improve this answer Follow

answered Oct 14, 2013 at 12:21



R. Max 6,710 1 28 35



Why not use two projects rather than only one?

-2

Let's name these two projects with proj1 and proj2 . In proj1 's settings.py , put these settings:



```
HTTP_PROXY = 'http://127.0.0.1:8123'
DOWNLOADER_MIDDLEWARES = {
    'myproject.middlewares.RandomUserAgentMiddleware': 400,
    'myproject.middlewares.ProxyMiddleware': 410,
    'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None
}
```



In proj2 's settings.py , put these settings:

```
DOWNLOADER_MIDDLEWARES = {
    'myproject.middlewares.RandomUserAgentMiddleware': 400,
    'scrapy.contrib.downloadermiddleware.useragent.UserAgentMiddleware': None
}
```



Share

Improve this answer

Follow

answered Oct 12, 2013 at 3:51

 **JavaNoScript**  
2,413 23 27

That's not what the user wants to do, there are certain cases where you want multiple spiders in same project. – [rajat](#) Sep 3, 2015 at 17:50

Start asking to get answers

Find the answer to your question by asking.

Ask question

Explore related questions

scrapy

See similar questions with these tags.