File - E:\scrapy□□□□\□□□□□\quotes\quotes\settings.py

```python
# -*- coding: utf-8 -*-

# Scrapy settings for quotes project
#
# For simplicity, this file contains only settings considered important or
# commonly used. You can find more settings consulting the documentation:
#
#     https://docs.scrapy.org/en/latest/topics/settings.html
#     https://docs.scrapy.org/en/latest/topics/downloader-middleware.html
#     https://docs.scrapy.org/en/latest/topics/spider-middleware.html

# 项目名称
BOT_NAME = 'quotes'

# 爬虫应用路径
SPIDER_MODULES = ['quotes.spiders']
NEWSPIDER_MODULE = 'quotes.spiders'

LOG_LEVEL = 'WARN'  # 警告级别以上的日志信息才打印

# 默认请求头
# USER_AGENT = 'quotes (+http://www.yourdomain.com)'

# 是否遵循爬虫规则/协议
ROBOTSTXT_OBEY = False

# Configure maximum concurrent requests performed by Scrapy (default: 16)
# 下载器总共最大处理的并发请求数,默认值16
# CONCURRENT_REQUESTS = 32

# Configure a delay for requests for the same website (default: 0)
# See https://docs.scrapy.org/en/latest/topics/settings.html#download-delay
# See also autothrottle settings and docs

# 请求延时的秒数
# DOWNLOAD_DELAY = 3

# The download delay setting will honor only one of:
# 每个域名能够被执行的最大并发请求数目，默认值8
# CONCURRENT_REQUESTS_PER_DOMAIN = 16

# 能够被单个IP处理的并发请求数，默认值0，代表无限制，需要注意两点：
#   1、如果不为零，那CONCURRENT_REQUESTS_PER_DOMAIN将被忽略，即并发数的限制是按照每个IP来计算，而不是每个域名
#   2、该设置也影响DOWNLOAD_DELAY，如果该值不为零，那么DOWNLOAD_DELAY下载延迟是限制每个IP而不是每个域
# CONCURRENT_REQUESTS_PER_IP = 16

# Disable cookies (enabled by default)
# 禁用缓存  默认为True
# COOKIES_ENABLED = False

# Disable Telnet Console (enabled by default)
# 是否禁用远程登录控制台
# TELNETCONSOLE_ENABLED = False

# Override the default request headers:
# 请求头
# DEFAULT_REQUEST_HEADERS = {
#   'Accept': 'text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8',
#   'Accept-Language': 'en',
# }
```

```
61
62 # Enable or disable spider middlewares
63 # See https://docs.scrapy.org/en/latest/topics/spider-middleware.html
64 # 爬虫中间件  --SPIDER_MIDDLEWARES
65 # SPIDER_MIDDLEWARES = {
66 #    'quotes.middlewares.QuotesSpiderMiddleware': 543,
67 # }
68
69 # Enable or disable downloader middlewares
70 # See https://docs.scrapy.org/en/latest/topics/downloader-middleware.html
71 # 下载中间件  --DOWNLOADER_MIDDLEWARES
72 # DOWNLOADER_MIDDLEWARES = {
73 #    'quotes.middlewares.QuotesDownloaderMiddleware': 543,
74 # }
75
76 # Enable or disable extensions
77 # See https://docs.scrapy.org/en/latest/topics/extensions.html
78 # 扩展/自定义组件
79 # EXTENSIONS = {
80 #    'scrapy.extensions.telnet.TelnetConsole': None,
81 # }
82
83 # Configure item pipelines
84 # See https://docs.scrapy.org/en/latest/topics/item-pipeline.html
85 # 管道文件配置项
86 # ITEM_PIPELINES = {
87 #    'quotes.pipelines.QuotesPipeline': 300,
88 # }
89
90 # 配置
91 # Enable and configure the AutoThrottle extension (disabled by default)
92 # See https://docs.scrapy.org/en/latest/topics/autothrottle.html
93 # 开启或关闭以下配置的总开关
94 # AUTOTHROTTLE_ENABLED = True
95
96 # The initial download delay  初始下载延迟
97 # AUTOTHROTTLE_START_DELAY = 5
98
99 # The maximum download delay to be set in case of high latencies 最大下载延迟
100 # AUTOTHROTTLE_MAX_DELAY = 60
101
102 # The average number of requests Scrapy should be sending in parallel to each remote server
103 # 每秒并发请求数的平均值，不能高于 CONCURRENT_REQUESTS_PER_DOMAIN或CONCURRENT_REQUESTS_PER_IP，
104 # 调高了则吞吐量增大，强奸目标站点，调低了则对目标站点更加”礼貌“
105 # 每个特定的时间点，scrapy并发请求的数目都可能高于或低于该值，这是爬虫视图达到的建议值而不是硬限
106 # AUTOTHROTTLE_TARGET_CONCURRENCY = 1.0
107
108 # Enable showing throttling stats for every response received:
109 # AUTOTHROTTLE_DEBUG = False  #  DEBUG 调试
110
111
112 # Enable and configure HTTP caching (disabled by default：默认禁止，因为注释掉了)
113 # See https://docs.scrapy.org/en/latest/topics/downloader-middleware.html#httpcache-middleware-settings
114 """
115 1. 启用缓存
116     目的用于将已经发送的请求或相应缓存下来，以便以后使用
117
118     from scrapy.downloadermiddlewares.httpcache import HttpCacheMiddleware
119     from scrapy.extensions.httpcache import DummyPolicy
120     from scrapy.extensions.httpcache import FilesystemCacheStorage
121 """
```

```
122 # 是否启用缓存
123 # HTTPCACHE_ENABLED = True
124 # 缓存超时时间
125 # HTTPCACHE_EXPIRATION_SECS = 0
126 # 缓存保存路径
127 # HTTPCACHE_DIR = 'httpcache'
128 # 缓存忽略的Http状态码
129 # HTTPCACHE_IGNORE_HTTP_CODES = []
130 # 缓存存储的插件
131 # HTTPCACHE_STORAGE = 'scrapy.extensions.httpcache.FilesystemCacheStorage'
132
```