DM_W3_06032024

Phạm Ngọc Hải, Lê Thị Minh Anh, Cao Diệu Ly

March 6, 2024

1 Nội dung thực hành

1. Thu thập dữ liệu từ nhiều nguồn website (cả cấu trúc và dữ liệu)

2. Lưu dữ liệu đã thu thập vào MySQL

3. Thực hành nội dung 1 và 2 cho dự án CK

2 Thu thập dữ liệu từ nhiều nguồn website

2.1 Danh sách các nguồn dữ liệu tham khảo

1. https://quotes.toscrape.com/

2. https://parade.com/937586/parade/life-quotes/ && https://parade.com/940913/parade/funny-quotes/

3. https://www.shopify.com/ph/blog/motivational-quotes#

4. https://www.today.com/life/quotes/inspirational-quotes-rcna30589

5. https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week

6. https://www.goodreads.com/quotes/tag/life

7.

```
[ ]: %cd /mnt/DataK/Project
     # !scrapy startproject quotescrawl
```

/mnt/DataK/Project

```
[ ]: %cd quotescrawl/quotescrawl/spiders/
```

/mnt/DataK/Project/quotescrawl/quotescrawl/spiders

```
[ ]: !cat domains.py
```

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


class Domain_1(scrapy.Spider):
    name = "Domain_1_toscrape-css"
    start_urls = [
        'http://quotes.toscrape.com/',
    ]

    def parse(self, response):
        # parse qua được 10 trang và lấy được 100 câu châm ngôn
        for quote in response.css("div.quote"):
            yield {
                'text': quote.css("span.text::text").extract_first(),
                'author': quote.css("small.author::text").extract_first(),
                'tags': quote.css("div.tags > a.tag::text").extract()
            }

        next_page_url = response.css("li.next > a::attr(href)").extract_first()
        if next_page_url is not None:
            yield scrapy.Request(response.urljoin(next_page_url))

class Domain_2(scrapy.Spider):
    name = "Domain_2_toscrape-css"
    start_urls = ['https://parade.com/937586/parade/life-quotes/'] #,
'https://parade.com/940913/parade/funny-quotes/']

    def parse(self, response):
        # for quote in response.css("div.m-page-wrapper is-below-header-ad.p"):
        for p_element in response.css("div.m-page-wrapper.is-below-header-ad
p"):
            quote_text = p_element.css("::text").get().strip()
            yield {
                'text': quote_text
            }

class Domain_5(scrapy.Spider):
    name = "Domain_5_toscrapy-css"
    start_urls = ['https://www.themuse.com/advice/45-inspirational-quotes-that-
will-get-you-through-the-work-week']

    def parse(self, response):
        for quote in response.css("li::text"):
            quote_text = quote.get().strip()  # Lấy nội dung của mỗi phần tử
<li> và loại bỏ các khoảng trắng không cần thiết

            # Kiểm tra xem nội dung có tồn tại không trước khi phân tách
```

```python
        if quote_text:
            # Phân tách nội dung trích dẫn và tác giả bằng cách tách chuỗi
bằng dấu gạch ngang và loại bỏ khoảng trắng ở hai bên
            parts = quote_text.split("-")
            if len(parts) == 2:
                quote_content = parts[0].strip()
                if quote_content[0] == '"':
                    quote_content = quote_content[1:]
                if quote_content[-1] == '"':
                    quote_content = quote_content[:-1]
                author = parts[1].strip()

                yield {
                    'text': quote_content,
                    'author': author
                }
```

```
[ ]: !scrapy list
     %cd ../data/
```

```
Domain_1_toscrape-css
Domain_2_toscrape-css
Domain_5_toscrapy-css
/mnt/DataK/Project/quotescrawl/quotescrawl/data
```

Ở đây chúng ta sẽ bắt đầu crawl từ nhiều domain khác nhau

```
[ ]: !scrapy crawl Domain_1_toscrape-css -o Domain_1_quotes.csv
     !scrapy crawl Domain_5_toscrapy-css -o Domain_5_quotes.csv
```

```
2024-03-06 16:18:13 [scrapy.utils.log] INFO: Scrapy 2.11.1 started (bot:
quotescrawl)
2024-03-06 16:18:13 [scrapy.utils.log] INFO: Versions: lxml 5.1.0.0, libxml2
2.12.3, cssselect 1.2.0, parsel 1.8.1, w3lib 2.1.2, Twisted 23.10.0, Python
3.11.7 (main, Jan 29 2024, 16:03:57) [GCC 13.2.1 20230801], pyOpenSSL 24.0.0
(OpenSSL 3.1.4 24 Oct 2023), cryptography 41.0.7, Platform
Linux-6.7.8-zen1-1-zen-x86_64-with-glibc2.39
2024-03-06 16:18:13 [scrapy.addons] INFO: Enabled addons:
[]
2024-03-06 16:18:13 [asyncio] DEBUG: Using selector: EpollSelector
2024-03-06 16:18:13 [scrapy.utils.log] DEBUG: Using reactor:
twisted.internet.asyncioreactor.AsyncioSelectorReactor
2024-03-06 16:18:13 [scrapy.utils.log] DEBUG: Using asyncio event loop:
asyncio.unix_events._UnixSelectorEventLoop
```

```
2024-03-06 16:18:13 [scrapy.extensions.telnet] INFO: Telnet Password:
2959df26841ba21d
2024-03-06 16:18:13 [scrapy.middleware] INFO: Enabled extensions:
['scrapy.extensions.corestats.CoreStats',
 'scrapy.extensions.telnet.TelnetConsole',
 'scrapy.extensions.memusage.MemoryUsage',
 'scrapy.extensions.feedexport.FeedExporter',
 'scrapy.extensions.logstats.LogStats']
2024-03-06 16:18:13 [scrapy.crawler] INFO: Overridden settings:
{'BOT_NAME': 'quotescrawl',
 'DOWNLOAD_DELAY': 2,
 'FEED_EXPORT_ENCODING': 'utf-8',
 'NEWSPIDER_MODULE': 'quotescrawl.spiders',
 'REQUEST_FINGERPRINTER_IMPLEMENTATION': '2.7',
 'ROBOTSTXT_OBEY': True,
 'SPIDER_MODULES': ['quotescrawl.spiders'],
 'TWISTED_REACTOR': 'twisted.internet.asyncioreactor.AsyncioSelectorReactor',
 'USER_AGENT': 'Mozilla/5.0 (iPad; CPU OS 12_2 like Mac OS X) '
               'AppleWebKit/605.1.15 (KHTML, like Gecko) Mobile/15E148'}
2024-03-06 16:18:13 [scrapy.middleware] INFO: Enabled downloader middlewares:
['scrapy.downloadermiddlewares.robotstxt.RobotsTxtMiddleware',
 'scrapy.downloadermiddlewares.httpauth.HttpAuthMiddleware',
 'scrapy.downloadermiddlewares.downloadtimeout.DownloadTimeoutMiddleware',
 'scrapy.downloadermiddlewares.defaultheaders.DefaultHeadersMiddleware',
 'scrapy.downloadermiddlewares.useragent.UserAgentMiddleware',
 'scrapy.downloadermiddlewares.retry.RetryMiddleware',
 'scrapy.downloadermiddlewares.redirect.MetaRefreshMiddleware',
 'scrapy.downloadermiddlewares.httpcompression.HttpCompressionMiddleware',
 'scrapy.downloadermiddlewares.redirect.RedirectMiddleware',
 'scrapy.downloadermiddlewares.cookies.CookiesMiddleware',
 'scrapy.downloadermiddlewares.httpproxy.HttpProxyMiddleware',
 'scrapy.downloadermiddlewares.stats.DownloaderStats']
2024-03-06 16:18:13 [scrapy.middleware] INFO: Enabled spider middlewares:
['scrapy.spidermiddlewares.httperror.HttpErrorMiddleware',
 'scrapy.spidermiddlewares.offsite.OffsiteMiddleware',
 'scrapy.spidermiddlewares.referer.RefererMiddleware',
 'scrapy.spidermiddlewares.urllength.UrlLengthMiddleware',
 'scrapy.spidermiddlewares.depth.DepthMiddleware']
2024-03-06 16:18:13 [scrapy.middleware] INFO: Enabled item pipelines:
[]
2024-03-06 16:18:13 [scrapy.core.engine] INFO: Spider opened
2024-03-06 16:18:13 [scrapy.extensions.logstats] INFO: Crawled 0 pages (at 0
pages/min), scraped 0 items (at 0 items/min)
2024-03-06 16:18:13 [scrapy.extensions.telnet] INFO: Telnet console listening on
127.0.0.1:6023
2024-03-06 16:18:14 [scrapy.core.engine] DEBUG: Crawled (404) <GET
http://quotes.toscrape.com/robots.txt> (referer: None)
2024-03-06 16:18:16 [scrapy.core.engine] DEBUG: Crawled (200) <GET
```

```
http://quotes.toscrape.com/> (referer: None)
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"The world as we have created it is a process of our thinking. It
cannot be changed without changing our thinking."', 'author': 'Albert Einstein',
'tags': ['change', 'deep-thoughts', 'thinking', 'world']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"It is our choices, Harry, that show what we truly are, far more than
our abilities."', 'author': 'J.K. Rowling', 'tags': ['abilities', 'choices']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"There are only two ways to live your life. One is as though nothing
is a miracle. The other is as though everything is a miracle."', 'author':
'Albert Einstein', 'tags': ['inspirational', 'life', 'live', 'miracle',
'miracles']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"The person, be it gentleman or lady, who has not pleasure in a good
novel, must be intolerably stupid."', 'author': 'Jane Austen', 'tags':
['aliteracy', 'books', 'classic', 'humor']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"Imperfection is beauty, madness is genius and it's better to be
absolutely ridiculous than absolutely boring."', 'author': 'Marilyn Monroe',
'tags': ['be-yourself', 'inspirational']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"Try not to become a man of success. Rather become a man of value."',
'author': 'Albert Einstein', 'tags': ['adulthood', 'success', 'value']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"It is better to be hated for what you are than to be loved for what
you are not."', 'author': 'André Gide', 'tags': ['life', 'love']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"I have not failed. I've just found 10,000 ways that won't work."',
'author': 'Thomas A. Edison', 'tags': ['edison', 'failure', 'inspirational',
'paraphrased']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"A woman is like a tea bag; you never know how strong it is until it's
in hot water."', 'author': 'Eleanor Roosevelt', 'tags': ['misattributed-eleanor-
roosevelt']}
2024-03-06 16:18:16 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/>
{'text': '"A day without sunshine is like, you know, night."', 'author': 'Steve
Martin', 'tags': ['humor', 'obvious', 'simile']}
```

2024-03-06 16:18:19 [scrapy.core.engine] DEBUG: Crawled (200) <GET http://quotes.toscrape.com/page/2/> (referer: http://quotes.toscrape.com/)
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': ""This life is what you make it. No matter what, you're going to mess up sometimes, it's a universal truth. But the good part is you get to decide how you're going to mess it up. Girls will be your friends - they'll act like it anyway. But just remember, some come, some go. The ones that stay with you through everything - they're your true best friends. Don't let go of them. Also remember, sisters make the best friends in the world. As for lovers, well, they'll come and go too. And baby, I hate to say it, most of them - actually pretty much all of them are going to break your heart, but you can't give up because if you give up, you'll never find your soulmate. You'll never find that half who makes you whole and that goes for everything. Just because you fail once, doesn't mean you're gonna fail at everything. Keep trying, hold on, and always, always, always believe in yourself, because if you don't, then who will, sweetie? So keep your head high, keep your chin up, and most importantly, keep smiling, because life's a beautiful thing and there's so much to smile about."", 'author': 'Marilyn Monroe', 'tags': ['friends', 'heartbreak', 'inspirational', 'life', 'love', 'sisters']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"It takes a great deal of bravery to stand up to our enemies, but just as much to stand up to our friends."', 'author': 'J.K. Rowling', 'tags': ['courage', 'friends']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': ""If you can't explain it to a six year old, you don't understand it yourself."", 'author': 'Albert Einstein', 'tags': ['simplicity', 'understand']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': ""You may not be her first, her last, or her only. She loved before she may love again. But if she loves you now, what else matters? She's not perfect-you aren't either, and the two of you may never be perfect together but if she can make you laugh, cause you to think twice, and admit to being human and making mistakes, hold onto her and give her the most you can. She may not be thinking about you every second of the day, but she will give you a part of her that she knows you can break-her heart. So don't hurt her, don't change her, don't analyze and don't expect more than she can give. Smile when she makes you happy, let her know when she makes you mad, and miss her when she's not there."", 'author': 'Bob Marley', 'tags': ['love']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"I like nonsense, it wakes up the brain cells. Fantasy is a necessary ingredient in living."', 'author': 'Dr. Seuss', 'tags': ['fantasy']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"I may not have gone where I intended to go, but I think I have ended

up where I needed to be."', 'author': 'Douglas Adams', 'tags': ['life', 'navigation']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"The opposite of love is not hate, it's indifference. The opposite of art is not ugliness, it's indifference. The opposite of faith is not heresy, it's indifference. And the opposite of life is not death, it's indifference."', 'author': 'Elie Wiesel', 'tags': ['activism', 'apathy', 'hate', 'indifference', 'inspirational', 'love', 'opposite', 'philosophy']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"It is not a lack of love, but a lack of friendship that makes unhappy marriages."', 'author': 'Friedrich Nietzsche', 'tags': ['friendship', 'lack-of-friendship', 'lack-of-love', 'love', 'marriage', 'unhappy-marriage']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"Good friends, good books, and a sleepy conscience: this is the ideal life."', 'author': 'Mark Twain', 'tags': ['books', 'contentment', 'friends', 'friendship', 'life']}
2024-03-06 16:18:19 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/2/>
{'text': '"Life is what happens to us while we are making other plans."', 'author': 'Allen Saunders', 'tags': ['fate', 'life', 'misattributed-john-lennon', 'planning', 'plans']}
2024-03-06 16:18:21 [scrapy.core.engine] DEBUG: Crawled (200) <GET http://quotes.toscrape.com/page/3/> (referer: http://quotes.toscrape.com/page/2/)
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"I love you without knowing how, or when, or from where. I love you simply, without problems or pride: I love you in this way because I do not know any other way of loving but this, in which there is no I or you, so intimate that your hand upon my chest is my hand, so intimate that when I fall asleep your eyes close."', 'author': 'Pablo Neruda', 'tags': ['love', 'poetry']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"For every minute you are angry you lose sixty seconds of happiness."', 'author': 'Ralph Waldo Emerson', 'tags': ['happiness']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"If you judge people, you have no time to love them."', 'author': 'Mother Teresa', 'tags': ['attributed-no-source']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"Anyone who thinks sitting in church can make you a Christian must also think that sitting in a garage can make you a car."', 'author': 'Garrison Keillor', 'tags': ['humor', 'religion']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200

http://quotes.toscrape.com/page/3/>
{'text': '"Beauty is in the eye of the beholder and it may be necessary from time to time to give a stupid or misinformed beholder a black eye."', 'author': 'Jim Henson', 'tags': ['humor']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"Today you are You, that is truer than true. There is no one alive who is Youer than You."', 'author': 'Dr. Seuss', 'tags': ['comedy', 'life', 'yourself']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"If you want your children to be intelligent, read them fairy tales. If you want them to be more intelligent, read them more fairy tales."', 'author': 'Albert Einstein', 'tags': ['children', 'fairy-tales']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"It is impossible to live without failing at something, unless you live so cautiously that you might as well not have lived at all - in which case, you fail by default."', 'author': 'J.K. Rowling', 'tags': []}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"Logic will get you from A to Z; imagination will get you everywhere."', 'author': 'Albert Einstein', 'tags': ['imagination']}
2024-03-06 16:18:21 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/3/>
{'text': '"One good thing about music, when it hits you, you feel no pain."', 'author': 'Bob Marley', 'tags': ['music']}
2024-03-06 16:18:23 [scrapy.core.engine] DEBUG: Crawled (200) <GET http://quotes.toscrape.com/page/4/> (referer: http://quotes.toscrape.com/page/3/)
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/4/>
{'text': '"The more that you read, the more things you will know. The more that you learn, the more places you'll go."', 'author': 'Dr. Seuss', 'tags': ['learning', 'reading', 'seuss']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/4/>
{'text': '"Of course it is happening inside your head, Harry, but why on earth should that mean that it is not real?"', 'author': 'J.K. Rowling', 'tags': ['dumbledore']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/4/>
{'text': '"The truth is, everyone is going to hurt you. You just got to find the ones worth suffering for."', 'author': 'Bob Marley', 'tags': ['friendship']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/4/>
{'text': '"Not all of us can do great things. But we can do small things with great love."', 'author': 'Mother Teresa', 'tags': ['misattributed-to-mother-

teresa', 'paraphrased']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"To the well-organized mind, death is but the next great adventure."',
'author': 'J.K. Rowling', 'tags': ['death', 'inspirational']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"All you need is love. But a little chocolate now and then doesn't
hurt."', 'author': 'Charles M. Schulz', 'tags': ['chocolate', 'food', 'humor']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"We read to know we're not alone."', 'author': 'William Nicholson',
'tags': ['misattributed-to-c-s-lewis', 'reading']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"Any fool can know. The point is to understand."', 'author': 'Albert
Einstein', 'tags': ['knowledge', 'learning', 'understanding', 'wisdom']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"I have always imagined that Paradise will be a kind of library."',
'author': 'Jorge Luis Borges', 'tags': ['books', 'library']}
2024-03-06 16:18:23 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/4/>
{'text': '"It is never too late to be what you might have been."', 'author':
'George Eliot', 'tags': ['inspirational']}
2024-03-06 16:18:26 [scrapy.core.engine] DEBUG: Crawled (200) <GET
http://quotes.toscrape.com/page/5/> (referer:
http://quotes.toscrape.com/page/4/)
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/5/>
{'text': '"A reader lives a thousand lives before he dies, said Jojen. The man
who never reads lives only one."', 'author': 'George R.R. Martin', 'tags':
['read', 'readers', 'reading', 'reading-books']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/5/>
{'text': '"You can never get a cup of tea large enough or a book long enough to
suit me."', 'author': 'C.S. Lewis', 'tags': ['books', 'inspirational',
'reading', 'tea']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/5/>
{'text': '"You believe lies so you eventually learn to trust no one but
yourself."', 'author': 'Marilyn Monroe', 'tags': []}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/5/>
{'text': '"If you can make a woman laugh, you can make her do anything."',
'author': 'Marilyn Monroe', 'tags': ['girls', 'love']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/5/>

{'text': '"Life is like riding a bicycle. To keep your balance, you must keep moving."', 'author': 'Albert Einstein', 'tags': ['life', 'simile']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/5/>
{'text': '"The real lover is the man who can thrill you by kissing your forehead or smiling into your eyes or just staring into space."', 'author': 'Marilyn Monroe', 'tags': ['love']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/5/>
{'text': ""A wise girl kisses but doesn't love, listens but doesn't believe, and leaves before she is left."", 'author': 'Marilyn Monroe', 'tags': ['attributed-no-source']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/5/>
{'text': '"Only in the darkness can you see the stars."', 'author': 'Martin Luther King Jr.', 'tags': ['hope', 'inspirational']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/5/>
{'text': '"It matters not what someone is born, but what they grow to be."', 'author': 'J.K. Rowling', 'tags': ['dumbledore']}
2024-03-06 16:18:26 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/5/>
{'text': '"Love does not begin and end the way we seem to think it does. Love is a battle, love is a war; love is a growing up."', 'author': 'James Baldwin', 'tags': ['love']}
2024-03-06 16:18:29 [scrapy.core.engine] DEBUG: Crawled (200) <GET http://quotes.toscrape.com/page/6/> (referer: http://quotes.toscrape.com/page/5/)
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/6/>
{'text': '"There is nothing I would not do for those who are really my friends. I have no notion of loving people by halves, it is not my nature."', 'author': 'Jane Austen', 'tags': ['friendship', 'love']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/6/>
{'text': '"Do one thing every day that scares you."', 'author': 'Eleanor Roosevelt', 'tags': ['attributed', 'fear', 'inspiration']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/6/>
{'text': '"I am good, but not an angel. I do sin, but I am not the devil. I am just a small girl in a big world trying to find someone to love."', 'author': 'Marilyn Monroe', 'tags': ['attributed-no-source']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/6/>
{'text': '"If I were not a physicist, I would probably be a musician. I often think in music. I live my daydreams in music. I see my life in terms of music."', 'author': 'Albert Einstein', 'tags': ['music']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200

http://quotes.toscrape.com/page/6/>
{'text': '"If you only read the books that everyone else is reading, you can only think what everyone else is thinking."', 'author': 'Haruki Murakami', 'tags': ['books', 'thought']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/6/>
{'text': '"The difference between genius and stupidity is: genius has its limits."', 'author': 'Alexandre Dumas fils', 'tags': ['misattributed-to-einstein']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/6/>
{'text': '"He's like a drug for you, Bella."', 'author': 'Stephenie Meyer', 'tags': ['drug', 'romance', 'simile']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/6/>
{'text': '"There is no friend as loyal as a book."', 'author': 'Ernest Hemingway', 'tags': ['books', 'friends', 'novelist-quotes']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/6/>
{'text': '"When one door of happiness closes, another opens; but often we look so long at the closed door that we do not see the one which has been opened for us."', 'author': 'Helen Keller', 'tags': ['inspirational']}
2024-03-06 16:18:29 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/6/>
{'text': '"Life isn't about finding yourself. Life is about creating yourself."', 'author': 'George Bernard Shaw', 'tags': ['inspirational', 'life', 'yourself']}
2024-03-06 16:18:31 [scrapy.core.engine] DEBUG: Crawled (200) <GET
http://quotes.toscrape.com/page/7/> (referer:
http://quotes.toscrape.com/page/6/)
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/7/>
{'text': '"That's the problem with drinking, I thought, as I poured myself a drink. If something bad happens you drink in an attempt to forget; if something good happens you drink in order to celebrate; and if nothing happens you drink to make something happen."', 'author': 'Charles Bukowski', 'tags': ['alcohol']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/7/>
{'text': '"You don't forget the face of the person who was your last hope."', 'author': 'Suzanne Collins', 'tags': ['the-hunger-games']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/7/>
{'text': '"Remember, we're madly in love, so it's all right to kiss me anytime you feel like it."', 'author': 'Suzanne Collins', 'tags': ['humor']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/7/>
{'text': '"To love at all is to be vulnerable. Love anything and your heart will be wrung and possibly broken. If you want to make sure of keeping it intact you

must give it to no one, not even an animal. Wrap it carefully round with hobbies and little luxuries; avoid all entanglements. Lock it up safe in the casket or coffin of your selfishness. But in that casket, safe, dark, motionless, airless, it will change. It will not be broken; it will become unbreakable, impenetrable, irredeemable. To love is to be vulnerable."', 'author': 'C.S. Lewis', 'tags': ['love']}

2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"Not all those who wander are lost."', 'author': 'J.R.R. Tolkien', 'tags': ['bilbo', 'journey', 'lost', 'quest', 'travel', 'wander']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"Do not pity the dead, Harry. Pity the living, and, above all those who live without love."', 'author': 'J.K. Rowling', 'tags': ['live-death-love']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"There is nothing to writing. All you do is sit down at a typewriter and bleed."', 'author': 'Ernest Hemingway', 'tags': ['good', 'writing']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"Finish each day and be done with it. You have done what you could. Some blunders and absurdities no doubt crept in; forget them as soon as you can. Tomorrow is a new day. You shall begin it serenely and with too high a spirit to be encumbered with your old nonsense."', 'author': 'Ralph Waldo Emerson', 'tags': ['life', 'regrets']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"I have never let my schooling interfere with my education."', 'author': 'Mark Twain', 'tags': ['education']}
2024-03-06 16:18:31 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/7/>
{'text': '"I have heard there are troubles of more than one kind. Some come from ahead and some come from behind. But I've bought a big bat. I'm all ready you see. Now my troubles are going to have troubles with me!"', 'author': 'Dr. Seuss', 'tags': ['troubles']}
2024-03-06 16:18:32 [scrapy.core.engine] DEBUG: Crawled (200) <GET http://quotes.toscrape.com/page/8/> (referer: http://quotes.toscrape.com/page/7/)
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/8/>
{'text': '"If I had a flower for every time I thought of you…I could walk through my garden forever."', 'author': 'Alfred Tennyson', 'tags': ['friendship', 'love']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200 http://quotes.toscrape.com/page/8/>
{'text': '"Some people never go crazy. What truly horrible lives they must lead."', 'author': 'Charles Bukowski', 'tags': ['humor']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200

```
http://quotes.toscrape.com/page/8/>
{'text': '"The trouble with having an open mind, of course, is that people will
insist on coming along and trying to put things in it."', 'author': 'Terry
Pratchett', 'tags': ['humor', 'open-mind', 'thinking']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"Think left and think right and think low and think high. Oh, the
thinks you can think up if only you try!"', 'author': 'Dr. Seuss', 'tags':
['humor', 'philosophy']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"What really knocks me out is a book that, when you're all done
reading it, you wish the author that wrote it was a terrific friend of yours and
you could call him up on the phone whenever you felt like it. That doesn't
happen much, though."', 'author': 'J.D. Salinger', 'tags': ['authors', 'books',
'literature', 'reading', 'writing']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"The reason I talk to myself is because I'm the only one whose answers
I accept."', 'author': 'George Carlin', 'tags': ['humor', 'insanity', 'lies',
'lying', 'self-indulgence', 'truth']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"You may say I'm a dreamer, but I'm not the only one. I hope someday
you'll join us. And the world will live as one."', 'author': 'John Lennon',
'tags': ['beatles', 'connection', 'dreamers', 'dreaming', 'dreams', 'hope',
'inspirational', 'peace']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"I am free of all prejudice. I hate everyone equally. "', 'author':
'W.C. Fields', 'tags': ['humor', 'sinister']}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '"The question isn't who is going to let me; it's who is going to stop
me."', 'author': 'Ayn Rand', 'tags': []}
2024-03-06 16:18:33 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/8/>
{'text': '" Classic - a book which people praise and don't read."', 'author':
'Mark Twain', 'tags': ['books', 'classic', 'reading']}
2024-03-06 16:18:35 [scrapy.core.engine] DEBUG: Crawled (200) <GET
http://quotes.toscrape.com/page/9/> (referer:
http://quotes.toscrape.com/page/8/)
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"Anyone who has never made a mistake has never tried anything new."',
'author': 'Albert Einstein', 'tags': ['mistakes']}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
```

{'text': '""A lady's imagination is very rapid; it jumps from admiration to love, from love to matrimony in a moment.""', 'author': 'Jane Austen', 'tags': ['humor', 'love', 'romantic', 'women']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"Remember, if the time should come when you have to make a choice between what is right and what is easy, remember what happened to a boy who was good, and kind, and brave, because he strayed across the path of Lord Voldemort. Remember Cedric Diggory."', 'author': 'J.K. Rowling', 'tags': ['integrity']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"I declare after all there is no enjoyment like reading! How much sooner one tires of any thing than of a book! -- When I have a house of my own, I shall be miserable if I have not an excellent library."', 'author': 'Jane Austen', 'tags': ['books', 'library', 'reading']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"There are few people whom I really love, and still fewer of whom I think well. The more I see of the world, the more am I dissatisfied with it; and every day confirms my belief of the inconsistency of all human characters, and of the little dependence that can be placed on the appearance of merit or sense."', 'author': 'Jane Austen', 'tags': ['elizabeth-bennet', 'jane-austen']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"Some day you will be old enough to start reading fairy tales again."', 'author': 'C.S. Lewis', 'tags': ['age', 'fairytales', 'growing-up']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"We are not necessarily doubting that God will do the best for us; we are wondering how painful the best will turn out to be."', 'author': 'C.S. Lewis', 'tags': ['god']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"The fear of death follows from the fear of life. A man who lives fully is prepared to die at any time."', 'author': 'Mark Twain', 'tags': ['death', 'life']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"A lie can travel half way around the world while the truth is putting on its shoes."', 'author': 'Mark Twain', 'tags': ['misattributed-mark-twain', 'truth']]}
2024-03-06 16:18:35 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/9/>
{'text': '"I believe in Christianity as I believe that the sun has risen: not only because I see it, but because by it I see everything else."', 'author': 'C.S. Lewis', 'tags': ['christianity', 'faith', 'religion', 'sun']]}
2024-03-06 16:18:37 [scrapy.core.engine] DEBUG: Crawled (200) <GET
http://quotes.toscrape.com/page/10/> (referer:

```
http://quotes.toscrape.com/page/9/)
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"The truth." Dumbledore sighed. "It is a beautiful and terrible thing,
and should therefore be treated with great caution."', 'author': 'J.K. Rowling',
'tags': ['truth']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"I'm the one that's got to die when it's time for me to die, so let me
live my life the way I want to."', 'author': 'Jimi Hendrix', 'tags': ['death',
'life']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"To die will be an awfully big adventure."', 'author': 'J.M. Barrie',
'tags': ['adventure', 'love']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"It takes courage to grow up and become who you really are."',
'author': 'E.E. Cummings', 'tags': ['courage']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"But better to get hurt by the truth than comforted with a lie."',
'author': 'Khaled Hosseini', 'tags': ['life']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"You never really understand a person until you consider things from
his point of view… Until you climb inside of his skin and walk around in
it."', 'author': 'Harper Lee', 'tags': ['better-life-empathy']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"You have to write the book that wants to be written. And if the book
will be too difficult for grown-ups, then you write it for children."',
'author': "Madeleine L'Engle", 'tags': ['books', 'children', 'difficult',
'grown-ups', 'write', 'writers', 'writing']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"Never tell the truth to people who are not worthy of it."', 'author':
'Mark Twain', 'tags': ['truth']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"A person's a person, no matter how small."', 'author': 'Dr. Seuss',
'tags': ['inspirational']}
2024-03-06 16:18:38 [scrapy.core.scraper] DEBUG: Scraped from <200
http://quotes.toscrape.com/page/10/>
{'text': '"… a mind needs books as a sword needs a whetstone, if it is to keep
its edge."', 'author': 'George R.R. Martin', 'tags': ['books', 'mind']}
2024-03-06 16:18:38 [scrapy.core.engine] INFO: Closing spider (finished)
2024-03-06 16:18:38 [scrapy.extensions.feedexport] INFO: Stored csv feed (100
```

items) in: Domain_1_quotes.csv
2024-03-06 16:18:38 [scrapy.statscollectors] INFO: Dumping Scrapy stats:
{'downloader/request_bytes': 3651,
 'downloader/request_count': 11,
 'downloader/request_method_count/GET': 11,
 'downloader/response_bytes': 22159,
 'downloader/response_count': 11,
 'downloader/response_status_count/200': 10,
 'downloader/response_status_count/404': 1,
 'elapsed_time_seconds': 24.361698,
 'feedexport/success_count/FileFeedStorage': 1,
 'finish_reason': 'finished',
 'finish_time': datetime.datetime(2024, 3, 6, 9, 18, 38, 104523,
tzinfo=datetime.timezone.utc),
 'httpcompression/response_bytes': 108778,
 'httpcompression/response_count': 11,
 'item_scraped_count': 100,
 'log_count/DEBUG': 114,
 'log_count/INFO': 11,
 'memusage/max': 75804672,
 'memusage/startup': 75804672,
 'request_depth_max': 9,
 'response_received_count': 11,
 'robotstxt/request_count': 1,
 'robotstxt/response_count': 1,
 'robotstxt/response_status_count/404': 1,
 'scheduler/dequeued': 10,
 'scheduler/dequeued/memory': 10,
 'scheduler/enqueued': 10,
 'scheduler/enqueued/memory': 10,
 'start_time': datetime.datetime(2024, 3, 6, 9, 18, 13, 742825,
tzinfo=datetime.timezone.utc)}
2024-03-06 16:18:38 [scrapy.core.engine] INFO: Spider closed (finished)
2024-03-06 16:18:39 [scrapy.utils.log] INFO: Scrapy 2.11.1 started (bot:
quotescrawl)
2024-03-06 16:18:39 [scrapy.utils.log] INFO: Versions: lxml 5.1.0.0, libxml2
2.12.3, cssselect 1.2.0, parsel 1.8.1, w3lib 2.1.2, Twisted 23.10.0, Python
3.11.7 (main, Jan 29 2024, 16:03:57) [GCC 13.2.1 20230801], pyOpenSSL 24.0.0
(OpenSSL 3.1.4 24 Oct 2023), cryptography 41.0.7, Platform
Linux-6.7.8-zen1-1-zen-x86_64-with-glibc2.39
2024-03-06 16:18:39 [scrapy.addons] INFO: Enabled addons:
[]
2024-03-06 16:18:39 [asyncio] DEBUG: Using selector: EpollSelector
2024-03-06 16:18:39 [scrapy.utils.log] DEBUG: Using reactor:
twisted.internet.asyncioreactor.AsyncioSelectorReactor
2024-03-06 16:18:39 [scrapy.utils.log] DEBUG: Using asyncio event loop:
asyncio.unix_events._UnixSelectorEventLoop
2024-03-06 16:18:39 [scrapy.extensions.telnet] INFO: Telnet Password:

fb3410a62f125512
```
2024-03-06 16:18:39 [scrapy.middleware] INFO: Enabled extensions:
['scrapy.extensions.corestats.CoreStats',
 'scrapy.extensions.telnet.TelnetConsole',
 'scrapy.extensions.memusage.MemoryUsage',
 'scrapy.extensions.feedexport.FeedExporter',
 'scrapy.extensions.logstats.LogStats']
2024-03-06 16:18:39 [scrapy.crawler] INFO: Overridden settings:
{'BOT_NAME': 'quotescrawl',
 'DOWNLOAD_DELAY': 2,
 'FEED_EXPORT_ENCODING': 'utf-8',
 'NEWSPIDER_MODULE': 'quotescrawl.spiders',
 'REQUEST_FINGERPRINTER_IMPLEMENTATION': '2.7',
 'ROBOTSTXT_OBEY': True,
 'SPIDER_MODULES': ['quotescrawl.spiders'],
 'TWISTED_REACTOR': 'twisted.internet.asyncioreactor.AsyncioSelectorReactor',
 'USER_AGENT': 'Mozilla/5.0 (iPad; CPU OS 12_2 like Mac OS X) '
               'AppleWebKit/605.1.15 (KHTML, like Gecko) Mobile/15E148'}
2024-03-06 16:18:39 [scrapy.middleware] INFO: Enabled downloader middlewares:
['scrapy.downloadermiddlewares.robotstxt.RobotsTxtMiddleware',
 'scrapy.downloadermiddlewares.httpauth.HttpAuthMiddleware',
 'scrapy.downloadermiddlewares.downloadtimeout.DownloadTimeoutMiddleware',
 'scrapy.downloadermiddlewares.defaultheaders.DefaultHeadersMiddleware',
 'scrapy.downloadermiddlewares.useragent.UserAgentMiddleware',
 'scrapy.downloadermiddlewares.retry.RetryMiddleware',
 'scrapy.downloadermiddlewares.redirect.MetaRefreshMiddleware',
 'scrapy.downloadermiddlewares.httpcompression.HttpCompressionMiddleware',
 'scrapy.downloadermiddlewares.redirect.RedirectMiddleware',
 'scrapy.downloadermiddlewares.cookies.CookiesMiddleware',
 'scrapy.downloadermiddlewares.httpproxy.HttpProxyMiddleware',
 'scrapy.downloadermiddlewares.stats.DownloaderStats']
2024-03-06 16:18:39 [scrapy.middleware] INFO: Enabled spider middlewares:
['scrapy.spidermiddlewares.httperror.HttpErrorMiddleware',
 'scrapy.spidermiddlewares.offsite.OffsiteMiddleware',
 'scrapy.spidermiddlewares.referer.RefererMiddleware',
 'scrapy.spidermiddlewares.urllength.UrlLengthMiddleware',
 'scrapy.spidermiddlewares.depth.DepthMiddleware']
2024-03-06 16:18:39 [scrapy.middleware] INFO: Enabled item pipelines:
[]
2024-03-06 16:18:39 [scrapy.core.engine] INFO: Spider opened
2024-03-06 16:18:39 [scrapy.extensions.logstats] INFO: Crawled 0 pages (at 0
pages/min), scraped 0 items (at 0 items/min)
2024-03-06 16:18:39 [scrapy.extensions.telnet] INFO: Telnet console listening on
127.0.0.1:6023
2024-03-06 16:18:40 [scrapy.core.engine] DEBUG: Crawled (200) <GET
https://www.themuse.com/robots.txt> (referer: None)
2024-03-06 16:18:42 [scrapy.core.engine] DEBUG: Crawled (200) <GET
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
```

through-the-work-week> (referer: None)
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Bet on me? Bet I will."', 'author': 'Lizzo'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Before anything else, preparation is the key to success."', 'author':
'Alexander Graham Bell'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Stay afraid, but do it anyway. What's important is the action. You
don't have to wait to be confident. Just do it and eventually the confidence
will follow."', 'author': 'Carrie Fisher'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Make each day your masterpiece."', 'author': 'John Wooden'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Your talent determines what you can do. Your motivation determines
how much you're willing to do. Your attitude determines how well you do it."',
'author': 'Lou Holtz'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"I learned this, at least, by my experiment; that if one advances
confidently in the direction of his dreams, and endeavors to live the life which
he has imagined, he will meet with a success unexpected in common hours."',
'author': 'Henry David Thoreau'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"\u200b\u200bWe will fail when we fail to try."', 'author': 'Rosa
Parks'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Don't count the days, make the days count."', 'author': 'Muhammad
Ali'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"I love to see a young girl go out and grab the world by the lapels.
Life's a bitch. You've got to go out and kick ass."', 'author': 'Maya Angelou'}

2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Without ambition one starts nothing. Without work one finishes
nothing. The prize will not be sent to you. You have to win it."', 'author':
'Ralph Waldo Emerson'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"You've got to get up every morning with determination if you're going
to go to bed with satisfaction."', 'author': 'George Lorimer'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Boss up and change your life / You can have it all, no sacrifice"',
'author': 'Lizzo'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The plan is to fan this spark into a flame."', 'author': 'Lin-Manuel
Miranda'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Real change, enduring change, happens one step at a time."',
'author': 'Ruth Bader Ginsburg'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Light tomorrow with today."', 'author': 'Elizabeth Barrett Browning'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Drench yourself in words unspoken / Live your life with arms wide
open / Today is where your book begins / The rest is still unwritten"',
'author': 'Natasha Bedingfield'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"And now that you don't have to be perfect, you can be good."',
'author': 'John Steinbeck'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Never give up on a dream just because of the time it will take to
accomplish it. The time will pass anyway."', 'author': 'Earl Nightingale'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-

through-the-work-week>
{'text': '"Someday is not a day of the week."', 'author': 'Janet Dailey'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Too late for second-guessing / Too late to go back to sleep / It's
time to trust my instincts / Close my eyes and leap."', 'author': 'Stephen
Schwartz'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"A year from now you may wish you had started today."', 'author':
'Karen Lamb'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Try not. Do, or do not. There is no try."', 'author': 'Yoda'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': 'who has never felt a man's touch and whose heart has turned to stone,
but I don't understand how you could have self-confidence if you don't do the
work… Because confidence is like respect; you have to earn it."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"You get to decide where your time goes. You can either spend it
moving forward, or you can spend it putting out fires. You decide. And if you
don't decide, others will decide for you."', 'author': 'Tony Morgan'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Don't sit down and wait for the opportunities to come. Get up and
make them."', 'author': 'Madam C.J Walker'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The key is not to prioritize what's on your schedule, but to schedule
your priorities."', 'author': 'Stephen Covey'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Imagining something may be the first step in making it happen, but it
takes the real time and real efforts of real people to learn things, make
things, turn thoughts into deeds or visions into inventions."', 'author': 'Mr.
Rogers'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-

through-the-work-week>
{'text': '"When someone tells me 'no,' it doesn't mean I can't do it, it simply means I can't do it with them."', 'author': 'Karen E. Quinones Miller'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"What we fear of doing most is usually what we most need to do."', 'author': 'Ralph Waldo Emerson'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"You may not control all the events that happen to you, but you can decide not to be reduced by them."', 'author': 'Maya Angelou'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"What I've realized is that confidence is a daily journey. We always think that confidence is a destination we must get to, but it's a choice we make. I don't think it's something we ever really get to and stay at. And I think once you know that, it takes the pressure off."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Believe you can and you're halfway there."', 'author': 'Theodore Roosevelt'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"There will always be hurdles in life, but if you want to achieve a goal, you must continue."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Even if you're on the right track, you'll get run over if you just sit there."', 'author': 'Will Rogers'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Never confuse a single defeat with a final defeat."', 'author': 'F. Scott Fitzgerald'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"I'm a survivor / I'm not gonna give up / I'm not gonna stop / I'm gonna work harder"', 'author': 'Destiny's Child'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>

{'text': '"Never allow a person to tell you no who doesn't have the power to say yes."', 'author': 'Eleanor Roosevelt'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Real courage is holding on to a still voice in your head that says, 'I must keep going.' It's that voice that says nothing is a failure if it is not final. That voice that says to you, 'Get out of bed. Keep going. I will not quit.'"', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"I break chains all by myself / Won't let my freedom rot in hell / Hey! I'ma keep running / 'Cause a winner don't quit on themselves"', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Your imagination is your preview of life's coming attractions."', 'author': 'Albert Einstein'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"If you don't design your own life plan, chances are you'll fall into someone else's plan and guess what they have planned for you? Not much."', 'author': 'Jim Rohn'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"You can't make decisions based on fear and the possibility of what might happen."', 'author': 'Michelle Obama'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"It is better to fail in originality than to succeed in imitation."', 'author': 'Herman Melville'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Do not let anyone ever tell you who you are."', 'author': 'Kamala Harris'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"If everything seems under control, you're not going fast enough."', 'author': 'Mario Andretti'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-

through-the-work-week>
{'text': '"There does come a moment when you start saying, 'I want more,' and people look at you a little cross-eyed, because it's loving what you have and also knowing you want to try for more. Sometimes that makes people uncomfortable."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Trust yourself. You probably know more than you think you do… Trust that you can learn anything."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"I dream it / I work hard / I grind 'til I own it"', 'author': 'Beyoncé'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Build your own dreams or someone else will hire you to build theirs."', 'author': 'Farrah Gray'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"If you want to be a true professional, do something outside yourself."', 'author': 'Ruth Bader Ginsburg'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Owning up to your vulnerabilities is a form of strength."', 'author': 'Lizzo'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"Very often a change of self is needed more than a change of scene."', 'author': 'Arthur Christopher Benson'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"I've learned that people will forget what you said, people will forget what you did, but people will never forget how you made them feel."', 'author': 'Maya Angelou'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200 https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-through-the-work-week>
{'text': '"The reality is: Sometimes you lose. And you're never too good to lose. You're never too big to lose. You're never too smart to lose. It happens. And it happens when it needs to happen. And you need to embrace those things."', 'author': 'Beyoncé'}

2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Even if I don't reach all my goals, I've gone higher than I would
have if I hadn't set any."', 'author': 'Danielle Fotopoulis'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Failure is the condiment that gives success its flavor."', 'author':
'Truman Capote'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"I don't harp on the negative because if you do, then there's no
progression. There's no forward movement. You got to always look on the bright
side of things, and we are in control. Like, you have control over the choices
you make."', 'author': 'Taraji P. Henson'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"And how do you know when you're doing something right? How do you
know that? It feels so. What I know now is that feelings are really your GPS
system for life. When you're supposed to do something or not supposed to do
something, your emotional guidance system lets you know. The trick is to learn
to check your ego at the door and start checking your gut instead."', 'author':
'Oprah Winfrey'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Develop success from failures. Discouragement and failure are two of
the surest stepping stones to success."', 'author': 'Dale Carnegie'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '."', 'author': 'Lizzo'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"A dead end street is a good place to turn around."', 'author': 'Naomi
Judd'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The only way to do great work is to love what you do. If you haven't
found it yet, keep looking. Don't settle."', 'author': ''}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>

{'text': '"An unexamined life is not worth living."', 'author': 'Socrates'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Life can only be understood backward, but it must be lived
forwards."', 'author': 'Søren Kierkegaard'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Self-reflection is a much kinder teacher than regret is. Prioritize
yourself by making a habit of it."', 'author': 'Andrena Sawyer'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"I hate that word: 'lucky.' It cheapens a lot of hard work."',
'author': 'Peter Dinklage'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Woo girl, need to kick off your shoes / Got to take a deep breath,
time to focus on you"', 'author': 'Lizzo'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"It takes courage to grow up and become who you really are."',
'author': 'e.e. cummings'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Optimism is a huge asset. We can always use more of it. But optimism
isn't a belief that things will automatically get better; it's a conviction that
we can make things better."', 'author': 'Melinda French Gates'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Let nothing happen, just for a bit, let the minutes toll in the
stunning air, let us lie on our beds like astronauts, hurtling through space and
time."', 'author': 'Olivia Laing'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The more I want to get something done the less I call it work."',
'author': 'Richard Bach'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"There is always light. If only we're brave enough to see it. If only
we're brave enough to be it."', 'author': 'Amanda Gorman'}

2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Someone's sitting in the shade today because someone planted a tree a
long time ago."', 'author': 'Warren Buffett'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"We must be willing to let go of the life we planned so as to have the
life that is waiting for us."', 'author': 'Joseph Campbell'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Don't ever confuse the two, your life and your work. The second is
only part of the first."', 'author': 'Anna Quindlen'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The individual who says it is not possible should move out of the way
of those doing it."', 'author': 'Tricia Cunningham'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"The true meaning of life is to plant trees under whose shade you do
not expect to sit."', 'author': 'Nelson Henderson'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Success is where preparation and opportunity meet."', 'author':
'Bobby Unser'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Success is the sum of small efforts repeated day in and day out."',
'author': 'Robert Collier'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Before I started, I decided that I'd only pursue this career if my
self-worth was dependent on more than celebrity success."', 'author': 'Beyoncé'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"People get scared when you try to do something, especially when it
looks like you're succeeding. People do not get scared when you're failing. It
calms them. But when you're winning, it makes them feel like they're losing or,
worse yet, that maybe they should've tried to do something too, but now it's too
late. And since they didn't, they want to stop you. You can't let them.""',

'author': 'Mindy Kaling'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Success is liking yourself, liking what you do, and liking how you do
it."', 'author': 'Maya Angelou'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Success is a state of mind. If you want success, start thinking of
yourself as a success."', 'author': 'Joyce Brothers'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Don't aim for success if you want it, just do what you love and
believe in and it will come naturally."', 'author': 'David Frost'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Whether I fail or succeed, it's the going for it."', 'author':
'Jonathan Van Ness'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Success only comes to those who dare to attempt."', 'author':
'Mallika Tripathi'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"People rarely succeed unless they have fun in what they are doing."',
'author': 'Dale Carnegie'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '""Always work hard and have fun in what you do because I think that's
when you're more successful. You have to choose to do it.""', 'author': 'Simone
Biles'}
2024-03-06 16:18:43 [scrapy.core.scraper] DEBUG: Scraped from <200
https://www.themuse.com/advice/45-inspirational-quotes-that-will-get-you-
through-the-work-week>
{'text': '"Don't trust any one story of how to become successful."', 'author':
'Mindy Kaling'}
2024-03-06 16:18:43 [scrapy.core.engine] INFO: Closing spider (finished)
2024-03-06 16:18:43 [scrapy.extensions.feedexport] INFO: Stored csv feed (89
items) in: Domain_5_quotes.csv
2024-03-06 16:18:43 [scrapy.statscollectors] INFO: Dumping Scrapy stats:
{'downloader/request_bytes': 650,
 'downloader/request_count': 2,

```
 'downloader/request_method_count/GET': 2,
 'downloader/response_bytes': 32460,
 'downloader/response_count': 2,
 'downloader/response_status_count/200': 2,
 'elapsed_time_seconds': 3.208085,
 'feedexport/success_count/FileFeedStorage': 1,
 'finish_reason': 'finished',
 'finish_time': datetime.datetime(2024, 3, 6, 9, 18, 43, 84590,
tzinfo=datetime.timezone.utc),
 'httpcompression/response_bytes': 179558,
 'httpcompression/response_count': 2,
 'item_scraped_count': 89,
 'log_count/DEBUG': 94,
 'log_count/INFO': 11,
 'memusage/max': 75595776,
 'memusage/startup': 75595776,
 'response_received_count': 2,
 'robotstxt/request_count': 1,
 'robotstxt/response_count': 1,
 'robotstxt/response_status_count/200': 1,
 'scheduler/dequeued': 1,
 'scheduler/dequeued/memory': 1,
 'scheduler/enqueued': 1,
 'scheduler/enqueued/memory': 1,
 'start_time': datetime.datetime(2024, 3, 6, 9, 18, 39, 876505,
tzinfo=datetime.timezone.utc)}
2024-03-06 16:18:43 [scrapy.core.engine] INFO: Spider closed (finished)
```

```
[ ]: !head Domain_1_quotes.csv
     !echo '\n\n'
     !head Domain_5_quotes.csv
```

```
text,author,tags
"The world as we have created it is a process of our thinking. It cannot be
changed without changing our thinking.",Albert Einstein,"change,deep-
thoughts,thinking,world"
""It is our choices, Harry, that show what we truly are, far more than our
abilities."",J.K. Rowling,"abilities,choices"
"There are only two ways to live your life. One is as though nothing is a
miracle. The other is as though everything is a miracle.",Albert
Einstein,"inspirational,life,live,miracle,miracles"
""The person, be it gentleman or lady, who has not pleasure in a good novel,
must be intolerably stupid."",Jane Austen,"aliteracy,books,classic,humor"
""Imperfection is beauty, madness is genius and it's better to be absolutely
ridiculous than absolutely boring."",Marilyn Monroe,"be-yourself,inspirational"
"Try not to become a man of success. Rather become a man of value.",Albert
Einstein,"adulthood,success,value"
"It is better to be hated for what you are than to be loved for what you are
```

not.",André Gide,"life,love"
""I have not failed. I've just found 10,000 ways that won't work."",Thomas A.
Edison,"edison,failure,inspirational,paraphrased"
"A woman is like a tea bag; you never know how strong it is until it's in hot
water.",Eleanor Roosevelt,misattributed-eleanor-roosevelt

text,author
"Bet on me? Bet I will.",Lizzo
""Before anything else, preparation is the key to success."",Alexander Graham
Bell
""Stay afraid, but do it anyway. What's important is the action. You don't have
to wait to be confident. Just do it and eventually the confidence will
follow."",Carrie Fisher
"Make each day your masterpiece.",John Wooden
"Your talent determines what you can do. Your motivation determines how much
you're willing to do. Your attitude determines how well you do it.",Lou Holtz
""I learned this, at least, by my experiment; that if one advances confidently
in the direction of his dreams, and endeavors to live the life which he has
imagined, he will meet with a success unexpected in common hours."",Henry David
Thoreau
"We will fail when we fail to try.",Rosa Parks
""Don't count the days, make the days count."",Muhammad Ali
"I love to see a young girl go out and grab the world by the lapels. Life's a
bitch. You've got to go out and kick ass.",Maya Angelou

3   Cho dữ liệu vào cơ sở dữ liệu SQL

```
[ ]: !pip install sqlalchemy pymysql
```

Requirement already satisfied: sqlalchemy in
/home/harito/venv/py/lib/python3.11/site-packages (2.0.28)
Requirement already satisfied: pymysql in
/home/harito/venv/py/lib/python3.11/site-packages (1.1.0)
Requirement already satisfied: typing-extensions>=4.6.0 in
/home/harito/venv/py/lib/python3.11/site-packages (from sqlalchemy) (4.9.0)
Requirement already satisfied: greenlet!=0.4.17 in
/home/harito/venv/py/lib/python3.11/site-packages (from sqlalchemy) (3.0.3)

```
[ ]: !pwd
```

/mnt/DataK/Project/quotescrawl/quotescrawl/data

```
[ ]: import pandas as pd
     from sqlalchemy import create_engine

     # Đọc dữ liệu từ các tệp CSV
```

```python
df1 = pd.read_csv('Domain_1_quotes.csv')
df2 = pd.read_csv('Domain_5_quotes.csv')

# Kiểm tra số lượng cột trong DataFrame df1
if len(df1.columns) == 2:
    # Thêm cột "tags" và gán giá trị null
    df1['tags'] = None
if len(df2.columns) == 2:
    # Thêm cột "tags" và gán giá trị null
    df2['tags'] = None

# Kết hợp với DataFrame từ tệp CSV thứ hai
combined_df = pd.concat([df1, df2], ignore_index=True)

# Tạo kết nối đến cơ sở dữ liệu MySQL
engine = create_engine('mysql+pymysql://root:@localhost/quotes_db', echo=True)

# Lưu DataFrame vào bảng MySQL
combined_df.to_sql('quotes', con=engine, if_exists='replace', index=False)
```

```
2024-03-06 16:18:50,552 INFO sqlalchemy.engine.Engine SELECT DATABASE()
2024-03-06 16:18:50,553 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,556 INFO sqlalchemy.engine.Engine SELECT @@sql_mode
2024-03-06 16:18:50,558 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,560 INFO sqlalchemy.engine.Engine SELECT
@@lower_case_table_names
2024-03-06 16:18:50,562 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,565 INFO sqlalchemy.engine.Engine BEGIN (implicit)
2024-03-06 16:18:50,571 INFO sqlalchemy.engine.Engine DESCRIBE
`quotes_db`.`quotes`
2024-03-06 16:18:50,572 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,578 INFO sqlalchemy.engine.Engine DESCRIBE
`quotes_db`.`quotes`
2024-03-06 16:18:50,579 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,585 INFO sqlalchemy.engine.Engine SHOW FULL TABLES FROM
`quotes_db`
2024-03-06 16:18:50,587 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,590 INFO sqlalchemy.engine.Engine SHOW FULL TABLES FROM
`quotes_db`
2024-03-06 16:18:50,591 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,621 INFO sqlalchemy.engine.Engine SHOW CREATE TABLE `quotes`
2024-03-06 16:18:50,623 INFO sqlalchemy.engine.Engine [raw sql] {}
2024-03-06 16:18:50,628 INFO sqlalchemy.engine.Engine
DROP TABLE quotes
2024-03-06 16:18:50,629 INFO sqlalchemy.engine.Engine [no key 0.00150s] {}
2024-03-06 16:18:50,650 INFO sqlalchemy.engine.Engine
CREATE TABLE quotes (
        text TEXT,
```

```
        author TEXT,
        tags TEXT
)
```

```
2024-03-06 16:18:50,652 INFO sqlalchemy.engine.Engine [no key 0.00193s] {}
2024-03-06 16:18:50,674 INFO sqlalchemy.engine.Engine INSERT INTO quotes (text,
author, tags) VALUES (%(text)s, %(author)s, %(tags)s)
2024-03-06 16:18:50,676 INFO sqlalchemy.engine.Engine [generated in 0.00499s]
[{'text': '"The world as we have created it is a process of our thinking. It
cannot be changed without changing our thinking."', 'author': 'Albert Einstein',
'tags': 'change,deep-thoughts,thinking,world'}, {'text': '"It is our choices,
Harry, that show what we truly are, far more than our abilities."', 'author':
'J.K. Rowling', 'tags': 'abilities,choices'}, {'text': '"There are only two ways
to live your life. One is as though nothing is a miracle. The other is as though
everything is a miracle."', 'author': 'Albert Einstein', 'tags':
'inspirational,life,live,miracle,miracles'}, {'text': '"The person, be it
gentleman or lady, who has not pleasure in a good novel, must be intolerably
stupid."', 'author': 'Jane Austen', 'tags': 'aliteracy,books,classic,humor'},
{'text': '"Imperfection is beauty, madness is genius and it's better to be
absolutely ridiculous than absolutely boring."', 'author': 'Marilyn Monroe',
'tags': 'be-yourself,inspirational'}, {'text': '"Try not to become a man of
success. Rather become a man of value."', 'author': 'Albert Einstein', 'tags':
'adulthood,success,value'}, {'text': '"It is better to be hated for what you are
than to be loved for what you are not."', 'author': 'André Gide', 'tags':
'life,love'}, {'text': '"I have not failed. I've just found 10,000 ways that
won't work."', 'author': 'Thomas A. Edison', 'tags':
'edison,failure,inspirational,paraphrased'}  … displaying 10 of 571 total
bound parameter sets …  {'text': '"Always work hard and have fun in what you
do because I think that's when you're more successful. You have to choose to do
it."', 'author': 'Simone Biles', 'tags': None}, {'text': '"Don't trust any one
story of how to become successful."', 'author': 'Mindy Kaling', 'tags': None}]
2024-03-06 16:18:50,714 INFO sqlalchemy.engine.Engine COMMIT
```

[ ]: 571

Như vậy đã thêm được 380 (ra 571 ở trên vì chạy lại code 2 lần) dữ liệu bản ghi có các thông tin
(quotes content, author, tags) thu thập nhiều nguồn domain khác nhau trên Internet và tổng hợp
lại vào CSDL SQL.

Sau đây là hình ảnh dữ liệu thu thập được:

```python
from IPython.display import Image

# Hiển thị hình ảnh
Image(filename='sql_database.png')
```

[ ]:

Showing rows 0 - 24 (380 total, Query took 0.0003 seconds.)

`SELECT * FROM `quotes``

☐ Profiling [Edit inline] [Edit] [Explain SQL] [Create PHP code] [Refresh]

text	author	tags
"The world as we have created it is a process of o...	Albert Einstein	change,deep-thoughts,thinking,world
"It is our choices, Harry, that show what we truly...	J.K. Rowling	abilities,choices
"There are only two ways to live your life. One is...	Albert Einstein	inspirational,life,live,miracle,miracles
"The person, be it gentleman or lady, who has not ...	Jane Austen	aliteracy,books,classic,humor
"Imperfection is beauty, madness is genius and it'...	Marilyn Monroe	be-yourself,inspirational
"Try not to become a man of success. Rather become...	Albert Einstein	adulthood,success,value
"It is better to be hated for what you are than to...	André Gide	life,love
"I have not failed. I've just found 10,000 ways th...	Thomas A. Edison	edison,failure,inspirational,paraphrased
"A woman is like a tea bag; you never know how str...	Eleanor Roosevelt	misattributed-eleanor-roosevelt
"A day without sunshine is like, you know, night."	Steve Martin	humor,obvious,simile
"This life is what you make it. No matter what, yo...	Marilyn Monroe	friends,heartbreak,inspirational,life,love,sisters
"It takes a great deal of bravery to stand up to o...	J.K. Rowling	courage,friends
"If you can't explain it to a six year old, you do...	Albert Einstein	simplicity,understand

4 Thực hành nội dung 1, 2 cho dự án cuối kỳ

Chủ đề cuối kỳ nhóm em đã có dữ liệu (dữ liệu ảnh y tế chất lượng cao) nên không dùng crawl từ trang web khác.

Việc tiếp theo bọn em làm với bộ dữ liệu đã có này là tiền xử lý, lọc nhiễu và gán nhãn dữ liệu cho các vùng ảnh cần quan tâm trong ảnh thu thập được để training model phân loại sau.