Started on	Friday, 28 March 2025, 10:23 AM
State	Finished
Completed on	Friday, 28 March 2025, 10:26 AM
Time taken	3 mins 6 secs
Marks	8.00/15.00
Grade	**53.33** out of 100.00

Question 1

Complete

Mark 0.00 out of 1.00

In Scrapy, what does the start_urls attribute do?

- ○ a. Registers middleware
- ○ b. Automatically triggers the parse() method for each URL
- ◉ c. Stores downloaded pages
- ○ d. Defines URLs for MongoDB

Question 2

Complete

Mark 1.00 out of 1.00

What happens if the next_page selector returns None?

- ○ a. Scrapy restarts from the first page
- ○ b. Scrapy skips the URL
- ◉ c. The crawl ends gracefully
- ○ d. Scrapy throws an error

Question 3

Complete

Mark 1.00 out of 1.00

What is the primary role of the Scrapy Engine?

- ◉ a. Coordinate requests, responses, and items between components
- ○ b. Send data to MongoDB
- ○ c. Render JavaScript-based pages
- ○ d. Manage HTML parsing

Question 4

Complete

Mark 1.00 out of 1.00

What is the purpose of defining an Item class in Scrapy?

○ a.   Automatically maps to MongoDB

◉ b.   Provides structured data containers

○ c.   Stores request headers

○ d.   Replaces the parse() method

Question 5

Complete

Mark 1.00 out of 1.00

What is the result of calling insert_one() in process_item() but forgetting to close the client in close_spider()?

◉ a.   It works, but may cause a memory leak or leave open connections

○ b.   Data will not be written at all

○ c.   MongoDB auto-closes everything safely

○ d.   Scrapy crashes

Question 6

Complete

Mark 1.00 out of 1.00

What will happen if you forget to return the item in process_item()?

○ a.   MongoDB inserts an empty document

○ b.   It crashes

○ c.   Scrapy skips it

◉ d.   Data will not pass to next pipeline or output

Question 7

Complete

Mark 0.00 out of 1.00

Where is the MongoDB connection usually initialized in a Scrapy pipeline?

○ a.   parse()

◉ b.   settings.py

○ c.   open_spider()

○ d.   __init__()

Question 8

Complete

Mark 1.00 out of 1.00

Which CSS selector correctly extracts quote text?

- ⦿ a. span.text::text
- ○ b. div.author
- ○ c. a.tag[href]
- ○ d. div.quote-text::value

Question 9

Complete

Mark 1.00 out of 1.00

Which line extracts all tag texts for a quote?

- ○ a. q.css('a.tag::href').getall()
- ○ b. q.select('tags').all()
- ⦿ c. q.css('div.tags a.tag::text').getall()
- ○ d. q.css('div.tags a.tag').get()

Question 10

Complete

Mark 1.00 out of 1.00

Which method in a Spider is called for each response?

- ⦿ a. parse()
- ○ b. yield()
- ○ c. start_requests()
- ○ d. scrape()

Question 11

Complete

Mark 0.00 out of 1.00

Which of the following can be returned from a Spider's parse() method?

- ○ a. Either Items or Requests
- ○ b. Only Response objects
- ○ c. Only Requests
- ⦿ d. Only Items

Question 12

Complete

Mark 0.00 out of 1.00

Which setting enables a custom pipeline in Scrapy?

- ○ a. ITEM_STORAGE_BACKEND
- ● b. SPIDER_PIPELINE
- ○ c. ITEM_PIPELINES
- ○ d. MONGO_PIPELINE

Question 13

Complete

Mark 0.00 out of 1.00

Why is dict(item) used before inserting into MongoDB?

- ○ a. To remove metadata
- ● b. To convert the response to JSON
- ○ c. MongoDB only accepts strings
- ○ d. Items are not serializable directly

Question 14

Complete

Mark 0.00 out of 1.00

Why might this CSS selector return None? quote.css('div.text::text').get()

- ● a. The site uses JavaScript rendering
- ○ b. You forgot to install BeautifulSoup
- ○ c. The element doesn't exist
- ○ d. Quotes use span.text, not div.text

Question 15

Complete

Mark 0.00 out of 1.00

You run db.quotes.find({ rating: 'Five' }).count() and get 0. What is a likely reason?

- ○ a. Scrapy removes the rating before storing
- ○ b. You used the wrong MongoDB URI
- ○ c. The rating field wasn't scraped properly or pipeline was skipped
- ● d. Ratings are stored as numbers