Question 1
Not yet answered

Marked out of 1.00

In Scrapy, what does the start_urls attribute do?

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

Question 2
Not yet answered

Marked out of 1.00

What happens if the next_page selector returns None?

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

Question 3
Not yet answered

Marked out of 1.00

What is the primary role of the Scrapy Engine?

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

Question 4
Not yet answered

Marked out of 1.00

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

- ○ a. Stores request headers
- ○ b. Provides structured data containers
- ○ c. Replaces the parse() method
- ○ d. Automatically maps to MongoDB

Question 5

Not yet answered

Marked 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.   Scrapy crashes
- ○ b.   Data will not be written at all
- ○ c.   It works, but may cause a memory leak or leave open connections
- ○ d.   MongoDB auto-closes everything safely

Question 6

Not yet answered

Marked out of 1.00

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

- ○ a.   It crashes
- ○ b.   Data will not pass to next pipeline or output
- ○ c.   MongoDB inserts an empty document
- ○ d.   Scrapy skips it

Question 7

Not yet answered

Marked 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

Not yet answered

Marked out of 1.00

Which CSS selector correctly extracts quote text?

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

Question 9

Not yet answered

Marked out of 1.00

Which line extracts all tag texts for a quote?

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

Question 10

Not yet answered

Marked out of 1.00

Which method in a Spider is called for each response?

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

Question 11

Not yet answered

Marked out of 1.00

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

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

Question 12

Not yet answered

Marked out of 1.00

Which setting enables a custom pipeline in Scrapy?

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

Question 13

Not yet answered

Marked out of 1.00

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

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

Question 14

Not yet answered

Marked out of 1.00

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

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

Question 15

Not yet answered

Marked out of 1.00

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

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