

WEB SCRAPING - WORKSHEET 2

In Q1 to Q10, only one option is correct, Choose the correct option:

1.	Which command method looks through a tag's descendants and retrieves all descendants that match your filters?	
	A) search_all()	B) get_all()
	C) find_all()	D) None of the above
	Answer: C)	
2.	Which operating system among the following	owing is not supported by selenium.
	A) Windows	B) Linux
	C) Unix	D) Solaris
	Answer: C)	,
3.	Which of these languages is NOT supported by the Selenium RC.	
	A) Python	B) Java
	C) ASP	D) C#
	Answer:C)	, -
4.	Which of these commands is used to enter text in text_boxes.	
	A) sentkeys()	B) sent_keys()
	C) keys.send()	D) send_keys()
	Answer:D)	D) sond_negs()
5.	Websites fetched by crawler are indexed and kept in huge database, this process is called as:	
	A) Optimizing	B) Indexing
	C) Crawling	D) None of them
	Answer:B)	
6.	Web Crawler is also called as:	
	A) Search Optimizer	B) Web-manager
	C) Web-spider	D) Linked Directory
	Answer: C)	
7.	Which selenium command among the following makes WebDriver wait for a certain condition to occur before	
	proceeding further with execution.	IIDDADA
	A) verifyTextPresent	B) webdriver_wait
	C) WebDriverWait	D) none of the above
	Answer:C	
8.	Which of the following are commonly used expected conditions in Selenium?	
	A) title_contains	B) staleness_of
	C) verifyelement	D) None of the above
	Answer:A)	
9.	Where xpath is used?	
	A) MS word documents	B) MS excel
	C) XML	D) None of the above
	Answer:XML	
10.	. Which of the following commands is used to find an element in a webpage from the id attribute of its tag.	
	A) find_element_BYid	B) find_element_by_id
	C) find_elements_by_id	D) None of the above
	Answer: B)	



In Q11 to Q13, more than one options are correct, Choose all the correct options:

11. Consider the following line in a HTML code

Lacie

Which of the following are attributes of <a> tag here?

A) class B) href

C) # Lacie # D) link2

Answer:A) & B)

12. Which among the following is/are parsers used in BeautifulSoup?

A) href B) soup_parser

C) lxml D) None of the above

Answer:C)

13. Select the packages you require generally in web-scraping.

A) Selenium B) BeautifulSoup

C) Requests D) B&C only.

Answer: A) & B)



Q14 and Q15 are programming practice questions. Solve it using JUPYTER NOTEBOOK.

14. Write a python program to scrap 10 images of Rayban Sunglasses from flipkart website and save them in a directory.

Github link attached

15. Write a program to scrap 20 mouses' data which includes model name, price from www.amazon.in whose price is less than 500 and make a data frame with 2 columns "model_name"," price" with the scraped data.

Github link attached

