**WEB SCRAPING – WORKSHEET 4**

1. Which of the following functions can be used to get an element from webpage when we know the Name attribute of the element?

Answer: C) find\_element\_by\_name()

2. Which of the following functions can be used when you want to locate an element by tag name?

Answer: C) find\_element\_by\_tag\_name()

3. In what type of Waits, a WebDriver waits for a certain condition to occur before proceeding further with execution.

Answer: B) Explicit wait

4. Which of the following is an expected condition in selenium (python)?

Answer: D) All of the above

5. Which of the following is a disadvantage of html5lib parser in beautiful soup?

Answer: B) Very Slow C) External Pyhton Dependency

6. What are the advantages of using Scrapy over Selenium for web-scraping?

Answer: A) For large data Scrapy is faster than selenium

7. Which of the following is (are) true regarding Scrapy?

Answer: D) both A & B

8. Full form of HTML:

Answer: C) Hyper Text Markup Language

9. Which among the following is the correct syntax for parsing a html page?

Answer: B) soup=BeautifulSoup(html\_doc,’html.parser’)

10. Which among the following is not a valid parser in BeautifulSoup?

Answer: D) “html-xml”

11. Which of the following functions is used to go to the next element in the page?

Answer: D) None of the above

12. Which of the following functions are used to iterate over an element’s siblings that precede it in the tree?

Answer: A) find\_previous\_siblings()

13. Which of the following is an argument used in find\_all() which tells Beautiful Soup to stop gathering results after it’s found a certain number.

Answer: C) limit

14. How would you set the recursive argument in find\_all() if you want Beautiful Soup to consider only the direct children

Answer: B) recursive = False

15. What is the difference between find() and find\_all() in Beautiful Soup?

Answer: find()- It just returns the result when the searched element is found in the page.And the return type will be <class 'bs4.element.Tag'>,that means it gets the first tag of the incoming HTML object that satisfies the condition and returns a function of a label group or a single label.

find\_all()- It returns all the matches (i.e) it scans the entire document and returns all the results and the return type will be <class 'bs4.element.ResultSet'>,that means it gets all the conditions of the incoming HTML object and return it.