Parsing HTML data with jsoup library:

Jsoup lib:java library for working with html code used to extract so many info from html code.

Steps:

1.add jsoup maven dependency in the project.

2.get the response as string🡪String Response = RestAssured.*given*().when().get("http://localhost:8085").asString();

3.Create object of type Document which is provided by jsoup library.

Document doc=Jsoup.*parse*(Response);//import it to **import** org.jsoup.nodes.Document;

Parsing the response and storing it in doc variabke of type Document.

**public** **class** jsoupexamples {

//form authentication to login

//session filter to get session id that has to be passes to get details in the redirected page

SessionFilter filter;

@BeforeMethod

**public** **void** init()

{

RestAssured.*baseURI*="http://localhost:8085";

filter=**new** SessionFilter();

FormAuthConfig form=**new** FormAuthConfig("/login", "uname", "pwd");

*given*()

.filter(filter)

.auth()

.form("user", "user",form)

.when()

.get();

}

//jsoup.parse parses the string response and returns in the document type.

@Test

**public** **void** extracttitleofthepage()

{

String Response = RestAssured.*given*().when().get("http://localhost:8085").asString();

Document doc=Jsoup.*parse*(Response);

System.***out***.println("Title of page is:"+doc.title().toUpperCase());

}

@Test

//since ,it returns multiple elementsbytag,it is stored in elements

**public** **void** extractelementsbytag()

{

String Response = RestAssured.*given*().when().get("http://localhost:8085").asString();

Document doc=Jsoup.*parse*(Response);

Elements element=doc.getElementsByTag("label");

**for**(Element e :element)

{

System.***out***.println(e);

}

//System.out.print(element);

}

@Test

**public** **void** extractelementsbyid()

{

String Response = RestAssured.*given*().when().get("http://localhost:8085").asString();

Document doc=Jsoup.*parse*(Response);

Element element=doc.getElementById("command");

System.***out***.println("elements under form tag:"+element.text());

}

@Test

**public** **void** extractlinks()

{

String Response = RestAssured.*given*().when().get("http://localhost:8085").asString();

Document doc=Jsoup.*parse*(Response);

Elements links=doc.select("a[href]");

System.***out***.println("links :"+links.text());

}

//here,we are extracting all the emailids with the help of css selector

In order to assert the values,we are storing all the elements in array list and then asserting from the list.

@Test

**public** **void** extractemails()

{

String res=RestAssured.*given*()

.sessionId(filter.getSessionId())

.when()

.get("/student/list")

.asString();

Document doc=Jsoup.*parse*(res);

Elements element=doc.select("table tbody tr td:nth-child(4)");

System.***out***.println("Size of table is :"+ element.size());

//to assert tthis,collect all this in arrylist and then assert it from there

List<String> lstactualvalues=**new** ArrayList<String>();

**for**(Element e:element)

{

//System.out.println("email ids of table is :"+ e.text());

lstactualvalues.add(e.text());

}

//Assert.assertEquals(lstactualvalues,hasItem("tincidunt.dui@ultricessit.co.uk"));

*assertThat*(lstactualvalues, *hasItem*("tincidunt.dui@ultricessit.co.uk"));

}