

```
1 Lab. HTML file 다루기
2
3 1. urllib.request module 이용하기
4 1)URL로 접속하여 HTML page 화면에 출력하기
5     from urllib.request import urlopen
6
7     url = 'https://www.python.org'
8     with urlopen(url) as f:
9         doc = f.read().decode()
10        print(doc)
11
12 2)URL로 접속하여 HTML page 파일로 저장하기
13     from urllib.request import urlopen
14
15     url = 'https://www.python.org'
16     with urlopen(url) as f:
17         doc = f.read().decode()
18         with open('pythonhome.html', 'w') as h:
19             h.writelines(doc)
20
21     print('Save Successfully')
22
23 3)Web image file로 저장하기
24     from urllib.request import urlopen
25
26     imgurl = 'https://www.python.org/static/img/python-logo@2x.png'
27     imgname = imgurl.split('/')[-1]
28     try:
29         with urlopen(imgurl) as f:
30             with open(imgname, 'wb') as h:
31                 img = f.read()
32                 h.write(img)
33         print('Save Successfully.')
34     except Exception as e:
35         print(e)
36
37
38 2. HTTPError를 이용하여 error 처리하기
39     from urllib.request import urlopen
40     from urllib.error import HTTPError
41
42     def download(url):
43         print('downloading:', url)
44         try :
45             html = urlopen(url).read().decode('utf-8')
46         except HTTPError as e:
47             print('Download error:', e.reason)
48             html = None
49         return html
50
51
```

```
52 download('http://httpstat.us/500')
53 -----
54 Downloading: http://httpstat.us/500
55 Download error: Internal Server Error
56
57 download('http://www.samsung.com')
58 -----
59 Downloading: http://www.samsung.com
60 Download error: Forbidden
61
62
63 3. requests module 이용하기
64 1)Python HTTP for Humans'
65 2)http://docs.python-requests.org/en/master/
66 3)Install
67     pip install requests
68
69
70     import requests
71     url = 'https://www.naver.com'
72     naver = requests.get(url)
73     print(naver.text)
74
75 4)pprint module 이용하기
76     -대량의 data를 보기 쉽게 표시해주는 표준 module
77     -pprint(prettyprint)
78
79     import requests
80     import pprint
81     url = 'https://www.naver.com'
82     naver = requests.get(url)
83     pprint.pprint(naver.text)
```