

```
1 import re
```

```
1 # 숫자찾기
```

```
2 example = '2021 11 25 thursday'
```

```
3 print(re.findall('\d', example))
```

```
4
```

```
5 # 숫자 아닌 것 찾기
```

```
6 print(re.findall('\D', example))
```

```
['2', '0', '2', '1', '1', '1', '2', '5']
[' ', ' ', ' ', ' ', 't', 'h', 'u', 'r', 's', 'd', 'a', 'y']
```

```
1 # 이메일 주소 찾기
```

```
2
```

```
3 example = '제 이메일은 hreeeee@yonsei.ac.kr 입니다. hreeeee@gmail.com으로 변경되었습니다.'
```

```
4
```

```
5 print(re.findall('[a-z]+\@[a-z]+\.[a-z]+\.[a-z]+', example))
```

```
['hreeeee@yonsei.ac.kr', 'hreeeee@gmail.com']
```

```
1 # URL 찾기
```

```
2
```

```
3 example = 'https://www.google.com/ 에 네이버 검색하시면 https://www.naver.com/ 이 상단0
```

```
4
```

```
5 print(re.findall('[a-z]+://w+.[a-z]+\.[a-z]+/', example))
```

```
☞ ['https://www.google.com/', 'https://www.naver.com/']
```

```
1 # 파일 찾기
```

```
2
```

```
3 example = '제 바탕화면 보시면 sonata01.png 파일과 avante02.jpg 파일 그리고 socar2021.gif
```

```
4 print(re.findall('[a-z]+\d+\.[a-z]+', example))
```

```
['sonata01.png', 'avante02.jpg', 'socar2021.gif']
```

```
1 sentence = 'He is a ten-year-old boy.'
```

```
2 sentence = re.sub('([^a-zA-Z.,!?])', ' ', sentence)
```

```
3 print(sentence)
```

```
He is a ten year old boy.
```

```
1 sentence = 'First, open the first chapter'
```

```
2 print(sentence.lower())
```

```
3
```

```
4 sentence = 'First, open the first chapter'
```

```
5 print(sentence.upper())
```

```
first, open the first chapter
FIRST, OPEN THE FIRST CHAPTER
```

```

1 def pad_punctuation(sentence, punc):
2     for p in punc:
3         sentence = sentence.replace(p, ' ' + p + ' ')
4
5     return sentence
6
7 sentence = 'Hi, my name is john'
8
9 print(pad_punctuation(sentence, ['.','?','!','(',')']))

```

Hi , my name is john

```

1 corpus = """It is with a kind of fear that I begin to write the history of my li

```

```

1 def cleaning_text(text, punc, regex):
2     text = re.sub(regex, ' ', text)
3     text = text.lower()
4     for p in punc:
5         text = text.replace(p, ' ' + p + ' ')
6     return text
7
8 print(cleaning_text(corpus, ['.','?','!','(',')'], '([a-zA-Z0-9.,!?])'))
9

```

it is with a kind of fear that i begin to write the history of my life . i ha

1