

# Python 데이터 분석과 이미지 처리

나동빈

# Pandas의 연산과 함수

## 데이터 프레임의 Null 여부 확인

```
import numpy as np
import pandas as pd

word_dict = {
    'Apple': '사과',
    'Banana': '바나나',
    'Carrot': '당근',
    'Durian': '두리안'
}

frequency_dict = {
    'Apple': 3,
    'Banana': 5,
    'Carrot': np.nan,
    'Durian': 2
}

importance_dict = {
    'Apple': 3,
    'Banana': 2,
    'Carrot': 1,
    'Durian': 1
}
```

```
word = pd.Series(word_dict)
frequency = pd.Series(frequency_dict)
importance = pd.Series(importance_dict)

summary = pd.DataFrame({
    'word': word,
    'frequency': frequency,
    'importance': importance
})

print(summary)
print(summary.notnull())
print(summary.isnull())
summary['frequency'] = summary['frequency'].fillna('데이터 없음')
print(summary)
```

# Pandas의 연산과 함수

## 시리즈 자료형의 연산

```
import pandas as pd

array1 = pd.Series([1, 2, 3], index=['A', 'B', 'C'])
array2 = pd.Series([4, 5, 6], index=['B', 'C', 'D'])

array1 = array1.add(array2, fill_value=0)
print(array1)
```

# Pandas의 연산과 함수

## 데이터 프레임 자료형의 연산

	1차	2차			1차			1차	2차
동빈	8점	5점	+	동빈	5점	=	동빈	13점	5점
길동	9점	7점		길동	7점		길동	16점	7점
				순신	8점		순신	8점	NaN

# Pandas의 연산과 함수

## 데이터 프레임 자료형의 연산

```
import pandas as pd

array1 = pd.DataFrame([[1, 2], [3, 4]], index=['A', 'B'])
array2 = pd.DataFrame([[1, 2, 3], [4, 5, 6], [7, 8, 9]], index=['B', 'C', 'D'])

print(array1)
print(array2)

array1 = array1.add(array2, fill_value=0)
print(array1)
```

# Pandas의 연산과 함수

## 데이터 프레임 집계 함수

```
import pandas as pd

array1 = pd.DataFrame([[1, 2], [3, 4]], index=['A', 'B'])
array2 = pd.DataFrame([[1, 2, 3], [4, 5, 6], [7, 8, 9]], index=['B', 'C', 'D'])

array1 = array1.add(array2, fill_value=0)
print(array1)
print("컬럼 1의 합:", array1[1].sum())
print(array1.sum())
```

# Pandas의 연산과 함수

## 데이터 프레임의 정렬 함수

```
import numpy as np
import pandas as pd

word_dict = {
    'Apple': '사과',
    'Banana': '바나나',
    'Carrot': '당근',
    'Durian': '두리안'
}

frequency_dict = {
    'Apple': 3,
    'Banana': 5,
    'Carrot': 1,
    'Durian': 2
}

importance_dict = {
    'Apple': 3,
    'Banana': 2,
    'Carrot': 1,
    'Durian': 1
}
```

```
word = pd.Series(word_dict)
frequency = pd.Series(frequency_dict)
importance = pd.Series(importance_dict)

summary = pd.DataFrame({
    'word': word,
    'frequency': frequency,
    'importance': importance
})

summary = summary.sort_values('frequency', ascending=False)
print(summary)
```