

NumPy (4) 기술 통계

```
"""
데이터의 개수(count)
평균(mean, average)
분산(variance)
표준 편차(standard deviation)
최댓값(maximum)
최소값(minimum)
중앙값(median)
사분위수(quantile)
"""
```

```
import numpy as np

"""통계: sum, mean, median, std, var
   최대/최소: min, max, argmin, argmax"""

x = np.array([18, 5, 10, 23, 19, -8, 10, 0, 0, 5, 2, 15, 8,
              2, 5, 4, 15, -1, 4, -7, -24, 7, 9, -6, 23, -13])

print( x.sum(), np.sum(x) )      #합
print( x.mean(), np.mean(x) )   #평균
print( np.median(x) )           #중앙값
print( x.var(), np.var(x) )     #분산
print( x.std(), np.std(x) )     #표준 편차

""" """

x = np.array([1, 3, 2])
print( x.min(), np.min(x) )     #최소값
print( x.max(), np.max(x) )     #최대값
print( x.argmin(), np.argmin(x) ) #최소값의 인덱스
print( x.argmax(), np.argmax(x) ) #최대값의 인덱스
```

실행결과

```
125 125
4.8076923076923075 4.8076923076923075
5.0
115.23224852071006 115.23224852071006
10.734628476137871 10.734628476137871
1 1
3 3
0 0
1 1
```

```

"""연산의 대상이 2차원 이상인 경우에는
axis=0인 경우는 열 연산
axis=1인 경우는 행 연산
"""

x = np.array([[1, 1], [2, 2]])

print("전체합", x.sum())
print(x.sum(axis=0)) # 열 합계
print(x.sum(axis=1)) # 행 합계

```

실행결과

전체합 6

[3 3]

[2 4]

```

"""사분위수
데이터를 가장 작은 수부터 순서대로 정렬하였을 때
1/4, 2/4, 3/4 위치에 있는 수. 각각 1사분위수, 2사분위수, 3사분위수 라고 함
2사분위수는 중앙값과 같다.
"""

print(np.percentile(x, 0) ) #최소값
print(np.percentile(x, 25) ) # 1사분위 수
print(np.percentile(x, 50) ) # 2사분위 수
print(np.percentile(x, 75) ) # 3사분위 수
print(np.percentile(x, 100) ) # 최댓값

```

실행결과

1.0

1.0

1.5

2.0

2.0