



연습문제

0부터 시작
순서 = 0, 1, 2, 3, ... 10
(-10, -9, -8, -7, ..., 10)

파이썬에서는 0부터 시작.

1. 다음과 같은 파이썬 명령어를 입력한 결과로 알맞은 것은?

library → R
import numpy as np
x = np.arange(-10, 11)
x[2] → 2번째 값 구하기.

3은 1부터
부터 시작!

- ① -10
② -9
③ -8
④ 8

2. 다음 프로그램 중에서 "nex8.head(2)"의 결과는 무엇인가?

import pandas as pd
nex8 = pd.read_csv("c:/data/dataintro/nex8-1.csv", header=0)
nex8.head(2)

위에서부터 몇줄인가. 위에서부터
2줄 추출
↓
1번째가 'header'

- ① 데이터 파일 이름은 nex8이고, 변수 이름을 출력한다.
② 변수 이름은 nex8이고, 2개 케이스를 출력한다.
③ 데이터 파일 이름은 nex8이고, 첫 2개 케이스를 출력한다.
④ 변수 이름은 nex8이고, 첫 2개 관측치를 출력한다.

3. 다음의 프로그램을 통해서 얻을 수 있는 것은?

import pandas as pd
nex8 = pd.read_csv("c:/data/dataintro/nex8-1.csv", header=0)
group_stat_by_sex = nex8.groupby("sex")["salary"].describe()

- ① sex(성별)에 따른 salary 변수에 대한 기술통계량을 구한다.
② sex(성별)에 따른 salary 변수에 대한 분할표를 구한다.
③ sex(성별)에 따른 salary 변수에 대한 히스토그램을 구한다.
④ sex(성별)에 따른 salary 변수에 대한 상자그림을 구한다.

crosstab
hist
boxplot

4. 변수 salary의 줄기-잎 그림을 그렸다. 결과에서 "4 | 1"의 의미는?

```
import stemgraphic
```

```
stemgraphic.stem_graphic(nex8.salary, scale=100)
```

```
10 | 41
  9 | 307
  7 | 20229
  3 | 1057
```

Key: aggr|stem|leaf
4 | 1 = 4.1 x 100 = 409.99999999999994

41: 10개

410: 100개

① 0.41

② 4.1

③ 41

④ 410

5. 다음 파이썬 프로그램에 해당하는 R 명령어로 적합한 것은?

```
import seaborn as sns
```

```
sns.boxplot(x="sex", y="salary", data=nex9)
```

Named Parameter.

① box(nex9.data\$salary ~ nex9.data\$sex)

② box(nex9.data\$sex ~ nex9.data\$salary)

③ boxplot(nex9.data\$salary ~ nex9.data\$sex)

④ boxplot(nex9.data\$sex ~ nex9.data\$salary)

formula 공인한 것.

6. 다음은 데이터를 불러와 분석하는 파이썬 프로그램의 일부이다. 다음 설명 중에서 옳은 것은?

```
import pandas as pd
```

```
nex8 = pd.read_csv("c:/data/dataintro/nex8-1.csv", header=0)
```

```
import stemgraphic
```

```
stemgraphic.stem_graphic(nex8.salary, scale=100)
```

① "c:/data/dataintro/nex8-1.csv"에 저장되어 있는 데이터를 읽어서 변수명을 nex8으로 하였다.

② 데이터는 20개의 관측치로 구성되어 있다.

데이터 파일명.

③ 상자그림을 얻을 수 있다.

④ 줄기-잎 그림을 통해서 연속형 자료의 분포 모양을 파악하고자 할 때 사용된다.

7. 다음은 파이썬 프로그램의 일부이다. 이 프로그램을 통해서 얻는 분석 결과는 무엇인가?

```
import matplotlib.pyplot as plt
plt.scatter(nex8.age, nex8.salary)
```

상정도

① 두 변수 age와 salary 사이의 상관계수를 구한다.

② 두 변수 age와 salary 사이의 산점도를 그린다.

③ 두 변수 age와 salary에 대한 상자그림을 그린다.

④ 두 변수 age와 salary 사이의 단순회귀식을 적합한다.

8. 다음과 같은 파이썬 명령어를 입력한 결과로 알맞은 것은?

```
import numpy as np
x = np.arange(-10,11) x = 2 * np.arange(-5,6)
x[2]
x = 2 * (-5.5)
y = x[x>2] → 4 6 8, 10 → mean(7)
mean = mean(y)
sum((y-mean)^2) (-3 -1 1 3)^2
```

① 10

② 15

③ 18

④ 20

sum(1 1 9) = 20

정답 1 ③ 2 ③ 3 ① 4 ④ 5 ③ 6 ④ 7 ② 8 ④