

1 Exercise. NumPy_answer

2

3 -다음 데이터를 이용해서 주어진 문제를 해결하세요.

4

5 import seaborn as sns

6 iris = sns.load_dataset("iris")

7 type(iris)

8 -----

9 pandas.core.frame.DataFrame

10

11 1. iris 데이터에서 처음 다섯 개 행만 출력하시오.

12 iris.head() 또는 iris.loc[0:4,]

13 -----

	sepal_length	sepal_width	petal_length	petal_width	species
0	5.1	3.5	1.4	0.2	setosa
1	4.9	3.0	1.4	0.2	setosa
2	4.7	3.2	1.3	0.2	setosa
3	4.6	3.1	1.5	0.2	setosa
4	5.0	3.6	1.4	0.2	setosa

20

21

22 2. iris 데이터에서 처음 50개 행을 빼내서 temp 변수에 저장하시오.

23 temp = iris.iloc[0:50]

24

25

26 3. 2번에서 선택한 DataFrame의 요약정보를 출력하시오. 모든 열에 대해 요약정보가 출력되어야 한다.

27 temp.describe(include='all')

28 -----

	sepal_length	sepal_width	petal_length	petal_width	species
count	50.00000	50.000000	50.000000	50.000000	50
unique	NaN	NaN	NaN	NaN	1
top	NaN	NaN	NaN	NaN	setosa
freq	NaN	NaN	NaN	NaN	50
mean	5.00600	3.428000	1.462000	0.246000	NaN
std	0.35249	0.379064	0.173664	0.105386	NaN
min	4.30000	2.300000	1.000000	0.100000	NaN
25%	4.80000	3.200000	1.400000	0.200000	NaN
50%	5.00000	3.400000	1.500000	0.200000	NaN
75%	5.20000	3.675000	1.575000	0.300000	NaN
max	5.80000	4.400000	1.900000	0.600000	NaN

41

42

43

44 4. versicolor 종의 데이터만 iris_versicolor 변수에 저장하시오.

45 iris_versicolor = iris.loc[iris.species=='versicolor']

46

47

48 5. iris 데이터의 각 열 평균값을 출력하시오.

49 iris.mean()

50 -----

51 sepal_length 5.843333

```
52 sepal_width    3.057333
53 petal_length   3.758000
54 petal_width    1.199333
55 dtype: float64
```

```
56
```

```
57
```

```
58 6. iris 데이터의 각 열들 사이의 상관계수를 출력하시오.
```

```
59 iris.corr()
```

```
60
```

```
61 7. iris 데이터의 종별 평균을 출력하시오.
```

```
62 iris.groupby(iris.species).mean()
```

```
63 -----
```

```
64          sepal_length sepal_width petal_length petal_width
```

```
65 species
```

```
66 setosa      5.006 3.428 1.462 0.246
```

```
67 versicolor  5.936 2.770 4.260 1.326
```

```
68 virginica   6.588 2.974 5.552 2.026
```