

2017년 2학기 공업수학 기말고사 시험지

이름:

학번:

모듈 임포트

코드를 실행하기 위해 필요한 모듈들이다.

```
In [1]: from __future__ import division, print_function
import numpy as np
import pandas as pd
from datetime import datetime as dt
```

넘파이 어레이

아래 모양의 어레이를 생성하기 위해 reshape 함수를 이용한다.

$$\begin{bmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{bmatrix}$$

```
In [2]: a = np.arange(1, 12, 2)
b = a.reshape(3,2)
b
```

```
Out[2]: array([[ 1,  3],
               [ 5,  7],
               [ 9, 11]])
```

추가로 아래 코드를 실행하자.

```
In [3]: b[2,1] = 3
```

문제 1

(1) a의 값은 어떻게 변경되었는지 설명하라.

(2) 위와 같은 현상이 발생하는 이유를 설명하라.

(3) 위와 같은 현상을 피하기 위해서는 어떻게 해야 하는지 설명하라.

넘파이 어레이 인덱싱/슬라이싱

아래 코드로 생성된 어레이를 이용하는 문제이다.

```
In [4]: a = np.arange(6) + np.arange(0, 51, 10)[: , np.newaxis]  
a
```

```
Out[4]: array([[ 0,  1,  2,  3,  4,  5],  
               [10, 11, 12, 13, 14, 15],  
               [20, 21, 22, 23, 24, 25],  
               [30, 31, 32, 33, 34, 35],  
               [40, 41, 42, 43, 44, 45],  
               [50, 51, 52, 53, 54, 55]])
```

a 어레이를 이용하여 예를 들어 아래 모양의 어레이를 생성할 수 있다.

$$\begin{bmatrix} 20 & 22 \\ 40 & 42 \end{bmatrix}$$

```
In [5]: a[2::2, :4:2]
```

```
Out[5]: array([[20, 22],  
               [40, 42]])
```

문제 2

(1) a 어레이에 인덱싱과 슬라이싱을 이용하여 아래 모양의 어레이를 생성하라.

$$\begin{bmatrix} 20 & 21 & 22 \\ 40 & 41 & 42 \end{bmatrix}$$

(2) a 어레이에 인덱싱과 슬라이싱을 이용하여 아래 모양의 어레이를 생성하라.

$$\begin{bmatrix} 5 & 15 & 25 & 35 & 45 & 55 \end{bmatrix}$$

.

(3) 마스크 인덱싱을 사용하여 아래 결과가 나오도록 하라.

```
array([ 0,  3, 12, 15, 21, 24, 30, 33, 42, 45, 51, 54])
```

.

정수 인덱싱

정수 인덱싱을 사용하여 아래 결과가 나오도록 할 수 있다.

```
array([ 12, 23, 34])
```

```
In [6]: a[(1, 2, 3), (2, 3, 4)]
```

```
Out[6]: array([12, 23, 34])
```

(4) 정수 인덱싱을 사용하여 아래 모양의 어레이가 나오도록 하라.

$$\begin{bmatrix} 30 & 32 & 35 \\ 50 & 52 & 55 \end{bmatrix}$$

.

데이터 분석

오늘 사용할 데이터는 다음과 같다.

- 미국 51개 주(State)별 담배(식물) 도매가격 및 판매일자: Weed_price.csv

아래 그림은 미국의 주별 담배(식물) 판매 데이터를 담은 Weed_Price.csv 파일을 엑셀로 읽었을 때의 일부를 보여준다. 실제 데이터량은 22899개이며, 아래 그림에는 5개의 데이터만을 보여주고 있다.

- 주의: 1번줄은 테이블의 열별 목록(column names)을 담고 있다.
- 열별 목록: State, HighQ, HighQN, MedQ, MedQN, LowQ, LowQN, date

	A	B	C	D	E	F	G	H
1	State	HighQ	HighQN	MedQ	MedQN	LowQ	LowQN	date
2	Alabama	339.06	1042	198.64	933	149.49	123	2014-01-01
3	Alaska	288.75	252	260.6	297	388.58	26	2014-01-01
4	Arizona	303.31	1941	209.35	1625	189.45	222	2014-01-01
5	Arkansas	361.85	576	185.62	544	125.87	112	2014-01-01
6	California	248.78	12096	193.56	12812	192.92	778	2014-01-01

문제 3

(1) 아래 코드를 설명하라.

```
In [7]: prices_pd = pd.read_csv("data/Weed_Price.csv", parse_dates=[-1])
```

.

(2) 아래 코드를 설명하라.

```
In [8]: prices_pd.sort_values(['State', 'date'], inplace=True)
```

.

(3) 아래 코드를 설명하라.

```
In [9]: prices_pd.fillna(method='ffill', inplace=True)
```

.

(4) 아래 코드를 설명하라.

```
In [10]: california_pd = prices_pd[prices_pd.State == "California"].copy(True)
```

.

(5) 아래 코드를 설명하라.

```
In [11]: ca_sum = california_pd['HighQ'].sum()  
ca_count = california_pd['HighQ'].count()  
ca_sum / ca_count
```

```
Out[11]: 245.3761247216035
```

.

(6) 아래 코드를 설명하라.

```
In [12]: def getYear(x):  
         return x.year  
  
year_col = prices_pd.date.apply(getYear)  
prices_pd["year"] = year_col
```

.

(7) 아래 코드를 설명하라.

[illegible]