

# 운동을 하지 않는 이유라는 데이터

Auto Mobile Robot

Exported on 01/04/2024

## Table of Contents

1	서울시 운동을 하지 않는 이유 통계 .....	3
1.1	데이터 .....	3
1.2	데이터 읽기 .....	4
1.3	info 명령 .....	4
1.4	대분류에서 성별만 .....	5
1.5	기간 컬럼을 지우면 .....	5
1.6	inplace 옵션 .....	5
1.7	성별만 다시 저장 .....	6
1.8	한글 대응 .....	6
1.9	운동을 할 충분한 시간이 없어서 .....	6
1.10	그런데 1, 2, 이런 말 말고 남자, 여자라고 되면 좋겠다 .....	7
1.11	다시 한 번 .....	7
1.12	함께 운동을 할 사람이 없어서에 대해서 - 응? 에러가 .....	8
1.13	겉으로 보면 문제가 없지만 .....	8
1.14	데이터 형이 . $\pi\pi$ .....	8
1.15	숫자로 바꾸고 .....	9
1.16	다시 그려보면 .....	9
1.17	좀더 이쁘게 그리기 위해 .....	10
1.18	명령을 잘 모를땐 공식 매뉴얼을 보자 .....	10
1.19	subplots이 return하는 인자가 두 개 .....	11
1.20	fig는 이런 설정을 담당하고 .....	11
1.21	axis는 이렇게 사용된다 .....	12
1.22	양마는 디테일에~ .....	12
1.23	결과~ .....	13

# 1 서울시 운동을 하지 않는 이유 통계

## 1.1 데이터

### 데이터셋

Home &gt; 공공데이터 &gt; 데이터셋



보건

#### 서울시 운동을 하지 않는 이유 통계

○ 통계개요

\* 통계명 : 운동을 하지 않는 이유

\* 통계종류 : 서울서베이의 서울시민 운동을 하지 않는 이유를 제공하는 일반·조사통계

\* 작성목적 : 도시정책지표에 대한 포괄적인 자료구축 및 분석을 실시함으로써 서울의

[전체 설명보기](#)

- <https://data.seoul.go.kr/dataList/10280/C/2/datasetView.do>

#### 미리보기

닫힘 -

Sheet

Chart

Open API

언어

한국어

소수점

기간/

년

2017 년

~

2017 년

※ 데이터에 콤마(,)가 많아 CSV로 변환이 곤란하여, 탭으로 구분하여 TXT로 제공합니다

내려받기(TXT)

내려받기(HWP)

자료분석

조회

단위 : %

기간	대분류	분류	운동을 할 충분한 시간이 없어서
	서울시		49.7
	성별	남자	55.2

- 꼭 Sheet를 선택하고 내려받기

## 1.2 데이터 읽기

```
import pandas as pd
```

```
raw_data = pd.read_csv('./data/report.txt', sep='\t')
raw_data.head()
```

	컬럼 1	2019	2019.1	2019.2	2019.3	2019.4	2019.5	2019.6	2019.7	2019.8	...	2019.42	2019.43	2019.4
0	컬럼 1	서울 시	성별	성별	연령별	연령별	연령별	연령별	연령별	연령별	...	강서구	구로구	금천구
1	컬럼 1	서울 시	남자	여자	10대	20대	30대	40대	50대	60대 이상	...	강서구	구로구	금천구
	운동을													

## 1.3 info 명령

```
raw_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 53 entries, 0 to 52
Data columns (total 8 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   기간                                  53 non-null     int64
1   대분류                                53 non-null     object
2   분류                                  53 non-null     object
3   운동을 할 충분한 시간이 없어서      53 non-null     float64
4   함께 운동을 할 사람이 없어서        53 non-null     object
5   운동을 할 만한 장소가 없어서        53 non-null     float64
6   운동을 싫어해서                      53 non-null     float64
7   기타                                  53 non-null     object
dtypes: float64(3), int64(1), object(4)
memory usage: 3.4+ KB
```

## 1.4 대분류에서 성별만

```
raw_data[raw_data['대분류'] == '성별']
```

	기간	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
1	2017	성별	남자	55.2	7.9	5.5	29.8	1.6
2	2017	성별	여자	45.0	8.3	6.0	38.8	1.8

## 1.5 기간 컬럼을 지우면

```
raw_data.drop(['기간'], axis='columns').head()
```

	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
0	서울시	서울시	49.7	8.1	5.8	34.7	1.7
1	성별	남자	55.2	7.9	5.5	29.8	1.6
2	성별	여자	45.0	8.3	6.0	38.8	1.8
3	연령별	10대	55.7	6.1	5.6	32.2	0.3
4	연령별	20대	54.8	6.9	5.2	32.9	0.1

## 1.6 inplace 옵션

```
raw_data.drop(['기간'], axis='columns', inplace=True)
raw_data.head()
```

	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
0	서울시	서울시	49.7	8.1	5.8	34.7	1.7
1	성별	남자	55.2	7.9	5.5	29.8	1.6
2	성별	여자	45.0	8.3	6.0	38.8	1.8
3	연령별	10대	55.7	6.1	5.6	32.2	0.3
4	연령별	20대	54.8	6.9	5.2	32.9	0.1

## 1.7 성별만 다시 저장

```
raw_data = raw_data[raw_data['대분류'] == '성별']
raw_data
```

	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
1	성별	남자	55.2	7.9	5.5	29.8	1.6
2	성별	여자	45.0	8.3	6.0	38.8	1.8

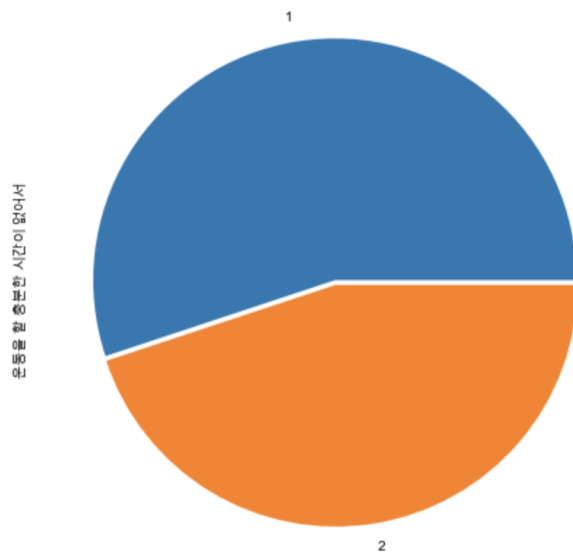
## 1.8 한글 대응

```
from matplotlib import rc

rc('font', family='Arial Unicode MS')
```

## 1.9 운동을 할 충분한 시간이 없어서

```
plt.figure(figsize=(10,8))
raw_data['운동을 할 충분한 시간이 없어서'].plot.pie(explode=[0,0.02])
plt.show()
```



## 1.10 그런데 1, 2, 이런 말 말고 남자, 여자라고 되면 좋겠다

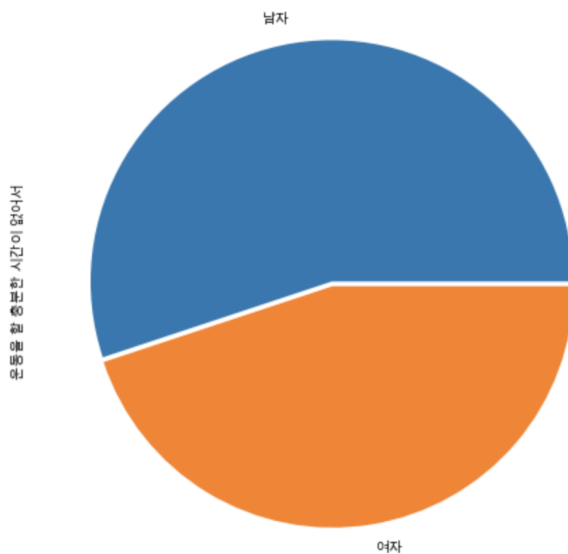
```
raw_data.set_index('분류', inplace=True)
raw_data.head()
```

	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
분류					
남자	55.2	7.9	5.5	29.8	1.6
여자	45.0	8.3	6.0	38.8	1.8

- 여러 방법이 있지만, index를 바꾸면 그래프를 그리는 옵션이 간결해진다

## 1.11 다시 한 번

```
plt.figure(figsize=(10,8))
raw_data['운동을 할 충분한 시간이 없어서'].plot.pie(explode=[0,0.02])
plt.show()
```



## 1.12 함께 운동을 할 사람이 없어서에 대해서 - 응? 에러가

```
plt.figure(figsize=(10,8))
raw_data['함께 운동을 할 사람이 없어서'].plot.pie(explode=[0,0.02])
plt.show()

245
246     elif is_object_dtype(lvalues.dtype):
--> 247         res_values = comp_method_OBJECT_ARRAY(op, lvalues, rvalues)
248
249     else:

~/opt/anaconda3/envs/tmp2/lib/python3.7/site-packages/pandas/core/ops/array_o
ps.py in comp_method_OBJECT_ARRAY(op, x, y)
    55         result = libops.vec_compare(x.ravel(), y, op)
```

## 1.13 겉으로 보면 문제가 없지만

```
raw_data
```

	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
분류					
남자	55.2	7.9	5.5	29.8	1.6
여자	45.0	8.3	6.0	38.8	1.8

## 1.14 데이터 형이 .πππ.

```
raw_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 2 entries, 남자 to 여자
Data columns (total 5 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   운동을 할 충분한 시간이 없어서      2 non-null     float64
1   함께 운동을 할 사람이 없어서        2 non-null     object
2   운동을 할 만한 장소가 없어서        2 non-null     float64
3   운동을 싫어해서                      2 non-null     float64
4   기타                                  2 non-null     object
dtypes: float64(3), object(2)
memory usage: 96.0+ bytes
```



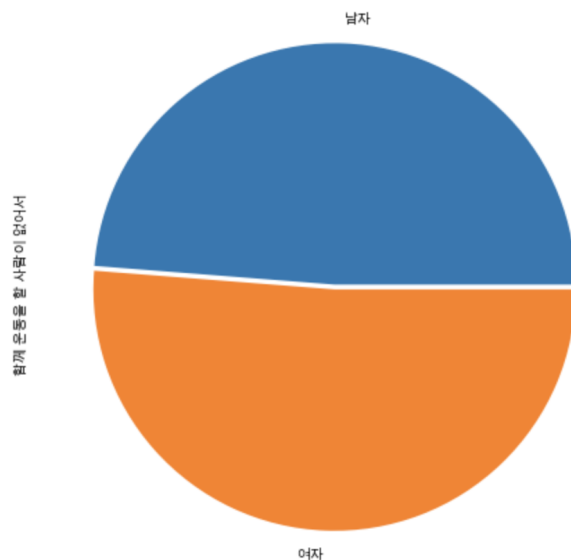
## 1.15 숫자로 바꾸고

```
raw_data['함께 운동을 할 사람이 없어서'] = raw_data['함께 운동을 할 사람이 없어서'].astype('float')
raw_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 2 entries, 남자 to 여자
Data columns (total 5 columns):
#   Column                                Non-Null Count  Dtype
---  -
0   운동을 할 충분한 시간이 없어서      2 non-null     float64
1   함께 운동을 할 사람이 없어서        2 non-null     float64
2   운동을 할 만한 장소가 없어서        2 non-null     float64
3   운동을 싫어해서                      2 non-null     float64
4   기타                                  2 non-null     object
dtypes: float64(4), object(1)
memory usage: 96.0+ bytes
```

## 1.16 다시 그려보면

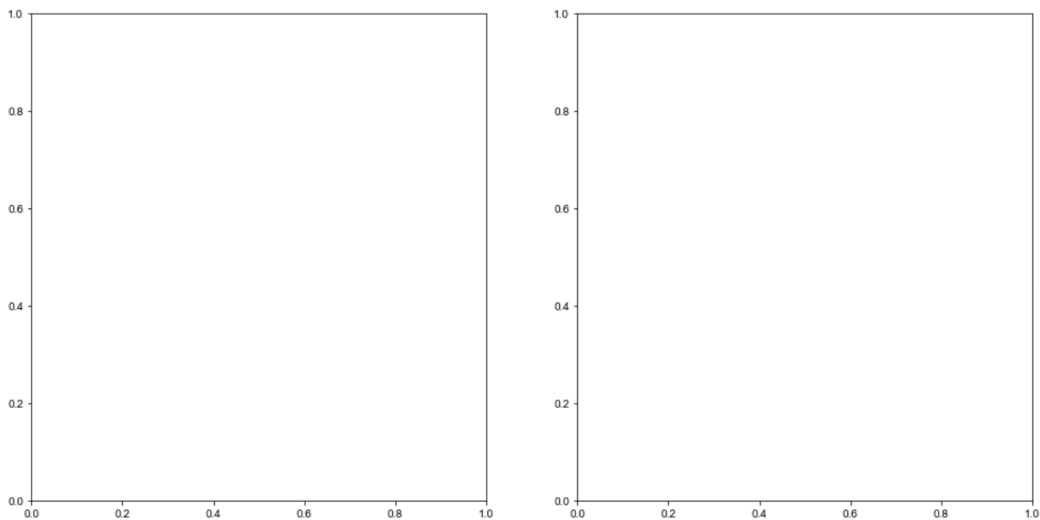
```
plt.figure(figsize=(10,8))
raw_data['함께 운동을 할 사람이 없어서'].plot.pie(explode=[0,0.02])
plt.show()
```



## 1.17 좀더 이쁘게 그리기 위해

```
fig, ax = plt.subplots(1,2, figsize=(16,8))
len(ax)
```

2



## 1.18 명령을 잘 모를땐 공식 매뉴얼을 보자

[Installation](#)
[Documentation](#)
[Examples](#)
[Tutorials](#)
[Contributing](#)

[home](#) | [contents](#) » [API Overview](#) » [matplotlib.pyplot](#) » [matplotlib.pyplot](#) » [matplotlib.pyplot.subplots](#)
[previous](#) | [next](#) | [modules](#) | [index](#)

### matplotlib.pyplot.subplots

```
matplotlib.pyplot.subplots(nrows=1, ncols=1, *, sharex=False, sharey=False, squeeze=True, subplot_kw=None,
                           gridspec_kw=None, **fig_kw)
```

[\[source\]](#)

Create a figure and a set of subplots.

This utility wrapper makes it convenient to create common layouts of subplots, including the enclosing figure object, in a single call.

**Parameters:**

**nrows, ncols**: int, default: 1

Number of rows/columns of the subplot grid.

**sharex, sharey**: bool or {'none', 'all', 'row', 'col'}, default: False

Controls sharing of properties among x (*sharex*) or y (*sharey*) axes:

- True or 'all': x- or y-axis will be shared among all subplots.
- False or 'none': each subplot x- or y-axis will be independent.
- 'row': each subplot row will share an x- or y-axis.
- 'col': each subplot column will share an x- or y-axis.

**Table of Contents**

[matplotlib.pyplot.subplots](#)

- [Examples using matplotlib.pyplot.subplots](#)

**Related Topics**

**Documentation overview**

- [API Overview](#)
  - [matplotlib.pyplot](#)
    - [matplotlib.pyplot](#)
      - [Previous: matplotlib.pyplot.subplots](#)
      - [Next: matplotlib.pyplot.subplots](#)

[Show Page Source](#)

## 1.19 subplots() return하는 인자가 두 개

### Returns:

**fig**: `Figure`

**ax**: `axes.Axes` or array of Axes

ax can be either a single `Axes` object or an array of Axes objects if more than one subplot was created. The dimensions of the resulting array can be controlled with the `squeeze` keyword, see above.

Typical idioms for handling the return value are:

```
# using the variable ax for single a Axes
fig, ax = plt.subplots()

# using the variable axs for multiple Axes
fig, axs = plt.subplots(2, 2)

# using tuple unpacking for multiple Axes
fig, (ax1, ax2) = plt.subplot(1, 2)
fig, ((ax1, ax2), (ax3, ax4)) = plt.subplot(2, 2)
```

The names `ax` and pluralized `axs` are preferred over `axes` because for the latter it's not clear if it refers to a single `Axes` instance or a collection of these.

## 1.20 fig는 이런 설정을 담당하고

### Parameters:

**figsize**: 2-tuple of floats, default: `rcParams["figure.figsize"]` (default: [6.4, 4.8])

Figure dimension (width, height) in inches.

**dpi**: float, default: `rcParams["figure.dpi"]` (default: 100.0)

Dots per inch.

**facecolor**: default: `rcParams["figure.facecolor"]` (default: 'white')

The figure patch facecolor.

**edgecolor**: default: `rcParams["figure.edgecolor"]` (default: 'white')

The figure patch edge color.

**linewidth**: float

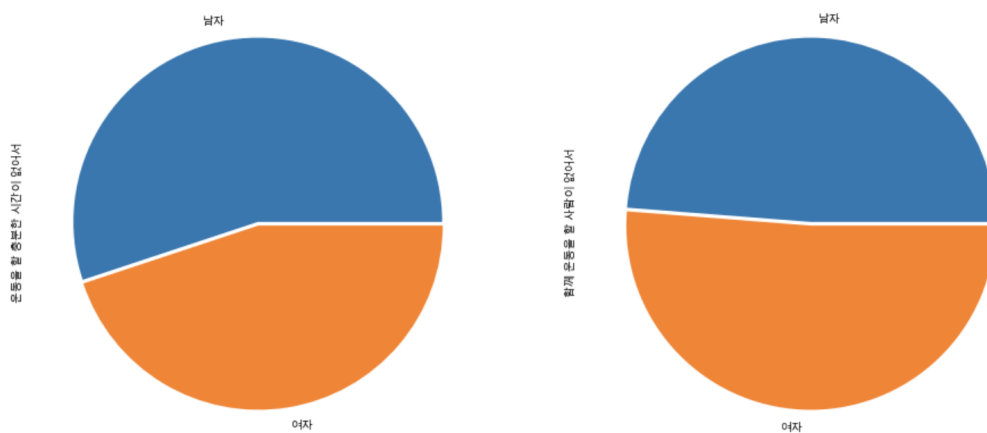
The linewidth of the frame (i.e. the edge linewidth of the figure patch).

## 1.21 axis는 이렇게 사용된다

```
fig, ax = plt.subplots(1,2, figsize=(16,8))

raw_data['운동을 할 충분한 시간이 없어서'].plot.pie(explode=[0,0.02], ax=ax[0])
raw_data['함께 운동을 할 사람이 없어서'].plot.pie(explode=[0,0.02], ax=ax[1])

plt.show()
```



## 1.22 양마는 디테일에~

```
f, ax = plt.subplots(1,2, figsize=(16,8))

raw_data['운동을 할 충분한 시간이 없어서'].plot.pie(explode=[0,0.02],
                                                         ax=ax[0], autopct='%1.1f%%')
ax[0].set_title('운동을 할 충분한 시간이 없어서')
ax[0].set_ylabel('')

raw_data['함께 운동을 할 사람이 없어서'].plot.pie(explode=[0,0.02],
                                                         ax=ax[1], autopct='%1.1f%%')
ax[1].set_title('함께 운동을 할 사람이 없어서')
ax[1].set_ylabel('')

plt.show()
```

1.23 결과~

