

# 운동을 하지 않는 이유 통계

<http://data.seoul.go.kr>

## ▼ 0. 한글폰트 설치 후 진행

### ▼ 1) 한글 폰트 설치

- 설치 후 '런타임 다시 시작'

```
!apt-get update
!apt-get install -y fonts-nanum
!fc-cache -fv
!rm ~/.cache/matplotlib -rf
```

```
Ign:1 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86\_64 InRelease
Get:2 https://cloud.r-project.org/bin/linux/ubuntu bionic-cran40/ InRelease [3,626 B]
Ign:3 https://developer.download.nvidia.com/compute/machine-learning/repos/ubuntu1804/x86\_64
Get:4 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86\_64 Release [69
Hit:5 https://developer.download.nvidia.com/compute/machine-learning/repos/ubuntu1804/x86\_64
Get:6 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86\_64 Release.gpg
Get:7 http://security.ubuntu.com/ubuntu bionic-security InRelease [88.7 kB]
Get:8 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic InRelease [15.9 kB]
Hit:9 http://archive.ubuntu.com/ubuntu bionic InRelease
Get:10 https://cloud.r-project.org/bin/linux/ubuntu bionic-cran40/ Packages [44.8 kB]
Get:11 http://archive.ubuntu.com/ubuntu bionic-updates InRelease [88.7 kB]
Hit:13 http://ppa.launchpad.net/cran/libgit2/ubuntu bionic InRelease
Ign:14 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86\_64 Packages
Get:14 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86\_64 Packages [
Get:15 http://archive.ubuntu.com/ubuntu bionic-backports InRelease [74.6 kB]
Hit:16 http://ppa.launchpad.net/graphics-drivers/ppa/ubuntu bionic InRelease
Get:17 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic/main Sources [1,725 kB]
Get:18 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 Packages [2,352 kB]
Get:19 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic/main amd64 Packages [883 k
Get:20 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 Packages [2,157 kB]
Fetched 7,987 kB in 3s (2,412 kB/s)
Reading package lists... Done
Reading package lists... Done
Building dependency tree
Reading state information... Done
The following NEW packages will be installed:
  fonts-nanum
0 upgraded, 1 newly installed, 0 to remove and 18 not upgraded.
Need to get 9,604 kB of archives.
After this operation, 29.5 MB of additional disk space will be used.
Get:1 http://archive.ubuntu.com/ubuntu bionic/universe amd64 fonts-nanum all 20170925-1 [9,60
Fetched 9,604 kB in 2s (5,694 kB/s)
Selecting previously unselected package fonts-nanum.
(Reading database ... 146442 files and directories currently installed.)
Preparing to unpack .../fonts-nanum_20170925-1_all.deb ...
```

```

Unpacking fonts-nanum (20170925-1) ...
Setting up fonts-nanum (20170925-1) ...
Processing triggers for fontconfig (2.12.6-0ubuntu2) ...
/usr/share/fonts: caching, new cache contents: 0 fonts, 1 dirs
/usr/share/fonts/truetype: caching, new cache contents: 0 fonts, 3 dirs
/usr/share/fonts/truetype/humor-sans: caching, new cache contents: 1 fonts, 0 dirs
/usr/share/fonts/truetype/liberation: caching, new cache contents: 16 fonts, 0 dirs
/usr/share/fonts/truetype/nanum: caching, new cache contents: 10 fonts, 0 dirs
/usr/local/share/fonts: caching, new cache contents: 0 fonts, 0 dirs
/root/.local/share/fonts: skipping, no such directory
/root/.fonts: skipping, no such directory
/var/cache/fontconfig: cleaning cache directory
/root/.cache/fontconfig: not cleaning non-existent cache directory
/root/.fontconfig: not cleaning non-existent cache directory
fc-cache: succeeded

```

## ▼ 2) 한글 폰트 사용

```

import matplotlib as mpl
import matplotlib.pyplot as plt

mpl.rcParams['axes.unicode_minus'] = False

plt.rc('font', family='NanumBarunGothic')

```

## ▼ I. Data Preprocessing

```

import warnings
warnings.filterwarnings('ignore')

```

### ▼ 1) Colab 파일 업로드

- notExercise.xls

```

!ls -l

total 20
-rw-r--r-- 1 root root 16384 Feb  6 23:12 notExercise.xls
drwxr-xr-x 1 root root  4096 Feb  4 15:26 sample_data

```

### ▼ 2) xls 파일 읽어오기

```

import pandas as pd

df = pd.read_excel('notExercise.xls')

```

```
DF = pd.read_excel('notExercise.xls')
```

```
DF.head()
```

	기간	대 분 류	분 류	운동을 할 충분한 시간 이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서
0	2017	서울시	서울시	49.7	8.1	5.8
1	2017	성별	남자	55.2	7.9	5.1

### 3) '기간' 열(Column) 삭제

```
DF.drop(columns = '기간', inplace = True)
```

```
DF.tail()
```

	대 분 류	분 류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기 타
48	관악구	관악구	50.7	7.1	6.9	33.6	1.7
49	서초구	서초구	59.8	3.4	3.2	32.0	1.5

### 4) 23~52번(지역분류) index 삭제

- 23~52 '대분류' 정보 확인

```
DF['대분류'][23:]
```

```

23  지역대분류
24  지역대분류
25  지역대분류
26  지역대분류
27  지역대분류
28  종로구
29  중구
30  용산구
31  성동구
32  광진구
33  동대문구
34  중랑구
35  성북구
36  강북구
37  도봉구

```

```
38     노원구
39     은평구
40     서대문구
41     마포구
42     양천구
43     강서구
44     구로구
45     금천구
46     영등포구
47     동작구
48     관악구
49     서초구
50     강남구
51     송파구
52     강동구
Name: 대분류, dtype: object
```

- 23~52 행(Row) 삭제

```
DF.drop(index = range(23, 53), inplace = True)
```

```
DF
```

	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서
0	서울시	서울시	49.7	8.1	5.8
1	성별	남자	55.2	7.9	5.5

## ▼ II. '성별' 통계

4	연령별	20대	54.8	6.9	5.2
---	-----	-----	------	-----	-----

### ▼ 1) '성별' 데이터 추출

```
DF_G = DF[DF['대분류'] == '성별'].copy()
```

```
DF_G
```

	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서
1	성별	남자	55.2	7.9	5.5

### ▼ 2) '대분류' 열(Column) 삭제

```
DF_G.drop(columns = '대분류', inplace = True)
```

```
DF_G
```

	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서
1	남자	55.2	7.9	5.5

### ▼ 3) '분류'를 index로 지정

```
DF_G.set_index('분류', inplace = True)
```

```
DF_G
```

## ▼ 4) matplotlib 시각화

은

```
import matplotlib.pyplot as plt

figure, ax = plt.subplots(1, 4, figsize = (16,8))

DF_G['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('')

DF_G['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('')

DF_G['운동을 싫어해서'].plot.pie(explode = [0, 0.02],
                                  ax = ax[2],
                                  autopct = '%1.1f%%')

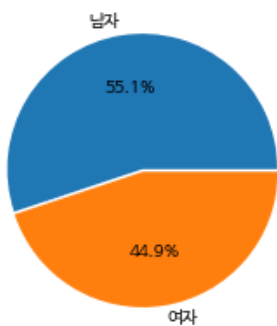
ax[2].set_title('운동을 싫어해서')
ax[2].set_ylabel('')

DF_G['운동을 할 만한 장소가 없어서'].plot.pie(explode = [0, 0.02],
                                                ax = ax[3],
                                                autopct = '%1.1f%%')

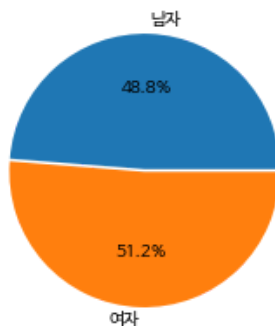
ax[3].set_title('운동을 할 만한 장소가 없어서')
ax[3].set_ylabel('')

plt.show()
```

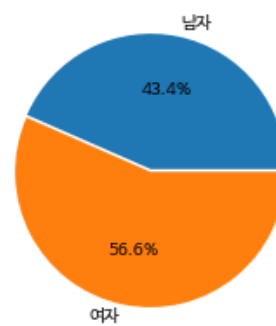
운동을 할 충분한 시간이 없어서



함께 운동을 할 사람이 없어서



운동을 싫어해서



운동을 할 만한 장소가 없어서



## ▼ III. '연령별' 통계

## ▼ 1) '연령별' 데이터 추출

```
DF_A = DF[DF['대분류'] == '연령별'].copy()
DF_A.drop(columns = '대분류', inplace = True)
DF_A.set_index('분류', inplace = True)
```

DF\_A

분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
10대	55.7	6.1	5.6	32.2	0.3
20대	54.8	6.9	5.2	32.9	0.1
30대	58.1	5.5	4.6	31.6	0.2
40대	57.7	6.8	5.2	29.4	1
50대	50.1	9.2	7.2	32.2	1.4

## ▼ 2) matplotlib 시각화

```
import matplotlib.pyplot as plt

figure, ax = plt.subplots(1, 3, figsize = (16,8))

explode_set = [0.02, 0.02, 0.02, 0.02, 0.02, 0.02]

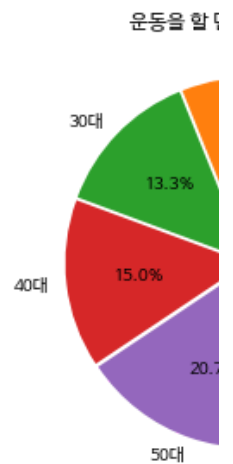
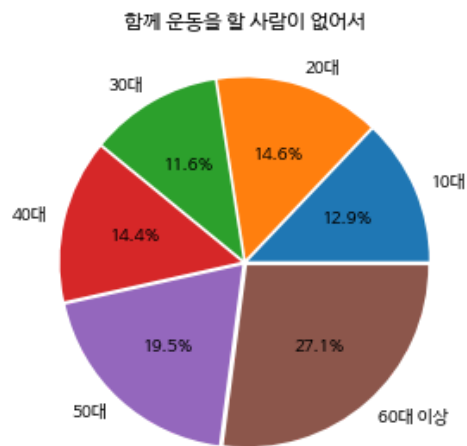
DF_A['운동을 할 충분한 시간이 없어서'].plot.pie(explode = explode_set,
                                                    ax = ax[0],
                                                    autopct = '%1.1f%%')
ax[0].set_title('운동을 할 충분한 시간이 없어서')
ax[0].set_ylabel('')

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

DF_A['운동을 할 만한 장소가 없어서'].plot.pie(explode = explode_set,
                                                    ax = ax[2],
                                                    autopct = '%1.1f%%')
```

```
ax[2].set_title('운동을 할 만한 장소가 없어서')
ax[2].set_ylabel('')

plt.show()
```



## ▼ IV. '학력별' 통계

### ▼ 1) '학력별' 데이터 추출

```
DF_V = DF[DF['대분류'] == '학력별'].copy()
DF_V.drop(columns = '대분류', inplace = True)
DF_V.set_index('분류', inplace = True)

DF_V
```

분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서
중졸 이하	26.1	13.7	6.6
고졸 이하	46.3	8.5	6.9
대졸 이상			

### ▼ 2) matplotlib 시각화

```
import matplotlib.pyplot as plt

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



```

explode_set = [0.02, 0.02, 0.02, 0.02]

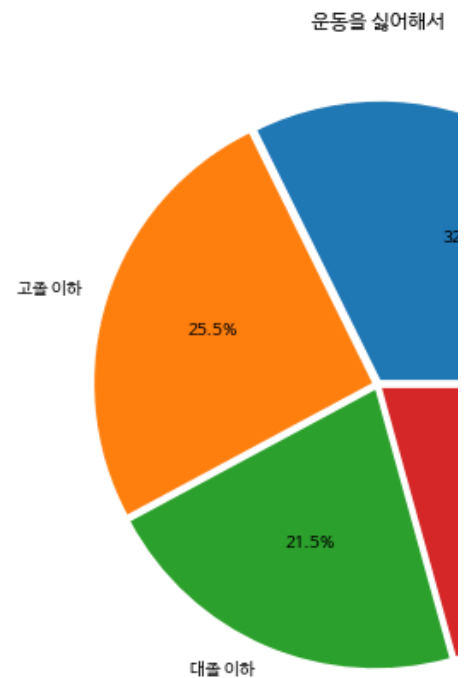
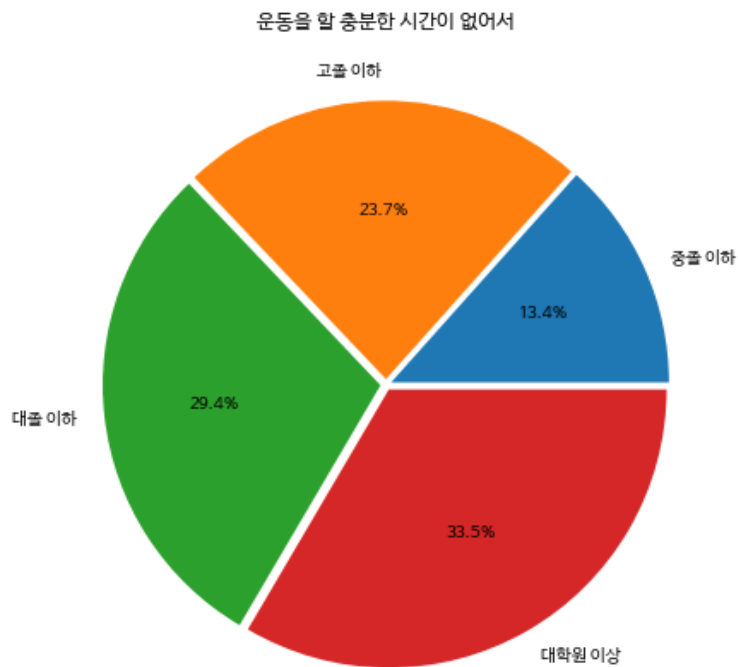
DF_V['운동을 할 충분한 시간이 없어서'].plot.pie(explode = explode_set,
                                                    ax = ax[0],
                                                    autopct = '%1.1f%%')
ax[0].set_title('운동을 할 충분한 시간이 없어서')
ax[0].set_ylabel('')

DF_V['운동을 싫어해서'].plot.pie(explode = explode_set,
                                  ax = ax[1],
                                  autopct = '%1.1f%%')

ax[1].set_title('운동을 싫어해서')
ax[1].set_ylabel('')

plt.show()

```



#

#

#

## The End

#

#

#