

# 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 InRelease
Get:4 http://security.ubuntu.com/ubuntu bionic-security InRelease [88.7 kB]
Hit:5 https://developer.download.nvidia.com/compute/cuda/repos/ubuntu1804/x86_64 Release
Hit:6 https://developer.download.nvidia.com/compute/machine-learning/repos/ubuntu1804/x86_64 Release
Get:7 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic InRelease [15.9 kB]
Hit:8 http://archive.ubuntu.com/ubuntu bionic InRelease
Get:9 http://archive.ubuntu.com/ubuntu bionic-updates InRelease [88.7 kB]
Hit:11 http://ppa.launchpad.net/cran/libgit2/ubuntu bionic InRelease
Get:13 http://archive.ubuntu.com/ubuntu bionic-backports InRelease [74.6 kB]
Hit:14 http://ppa.launchpad.net/graphics-drivers/ppa/ubuntu bionic InRelease
Get:15 http://security.ubuntu.com/ubuntu bionic-security/universe amd64 Packages [1,391 kB]
Get:16 http://security.ubuntu.com/ubuntu bionic-security/main amd64 Packages [1,929 kB]
Get:17 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic/main Sources [1,733 kB]
Get:18 http://archive.ubuntu.com/ubuntu bionic-updates/main amd64 Packages [2,360 kB]
Get:19 http://ppa.launchpad.net/c2d4u.team/c2d4u4.0+/ubuntu bionic/main amd64 Packages [887 kB]
Get:20 http://archive.ubuntu.com/ubuntu bionic-updates/universe amd64 Packages [2,163 kB]
Fetched 10.7 MB in 4s (2,644 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 31 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 [9,604 kB]
Fetched 9,604 kB in 2s (5,118 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.rcParams['font.family'] = 'NanumBarunGothic'
```

```
print('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 23 07:40 notExercise.xls
drwxr-xr-x 1 root root  4096 Feb 16 16:35 sample_data
```

### ▼ 2) xls 파일 읽어오기

```
import pandas as pd

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

DF.head()
```

	기간	대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
0	2017	서울시	서울시	49.7	8.1	5.8	34.7	1.7
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
3	2017	연령별	10대	55.7	6.1	5.6	32.2	0.3

### ▼ 3) '기간' 열(Column) 삭제

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

DF.tail()
```

		대분류	분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
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
50	강남구		강남구	59.1	5.5	2.6	31.5	1.3

### ▼ 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	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
5	연령별	30대	58.1	5.5	4.6	31.6	0.2
6	연령별	40대	57.7	6.8	5.2	29.4	1
7	연령별	50대	50.1	9.2	7.2	32.2	1.4
8	연령별	60대 이상	27.5	12.8	6.9	46.8	6
9	학력별	중졸 이하	26.1	13.7	6.6	46.2	7.4
10	학력별	고졸 이하	46.3	8.5	6.9	36.5	1.7
11	학력별	대졸 이하	57.4	6.6	4.9	30.7	0.4
12	학력별	대학원 이상	65.3	3.8	0.6	29.6	0.6
13	소득별	100만원 미만	22.7	10.3	3.6	50.7	12.7
14	소득별	100-200만원 미만	34.3	11.5	7.1	41.5	5.6
15	소득별	200-300만원 미만	48.8	12.3	6.8	30.7	1.4
16	소득별	300-400만원 미만	54.9	7.1	6.5	31.2	0.3
17	소득별	400-500만원 미만	56.1	5	5.3	33.0	0.6
18	소득별	500만원 이상	52.3	7.1	5.1	34.9	0.6
19	혼인상태별	기혼	50.7	7.8	5.4	34.9	1.2

## Ⅱ. '성별' 통계

### 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	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

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	운동을 싫어해서	기타
1	남자	55.2	7.9	5.5	29.8	1.6
2	여자	45.0	8.3	6.0	38.8	1.8

3) '분류'를 index로 지정

DF\_G.set\_index('분류', inplace = True)

DF\_G

분류	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
남자	55.2	7.9	5.5	29.8	1.6

4) matplotlib 시각화

```
import matplotlib.pyplot as plt

figure, ax = plt.subplots(2, 2, figsize = (10,10))

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

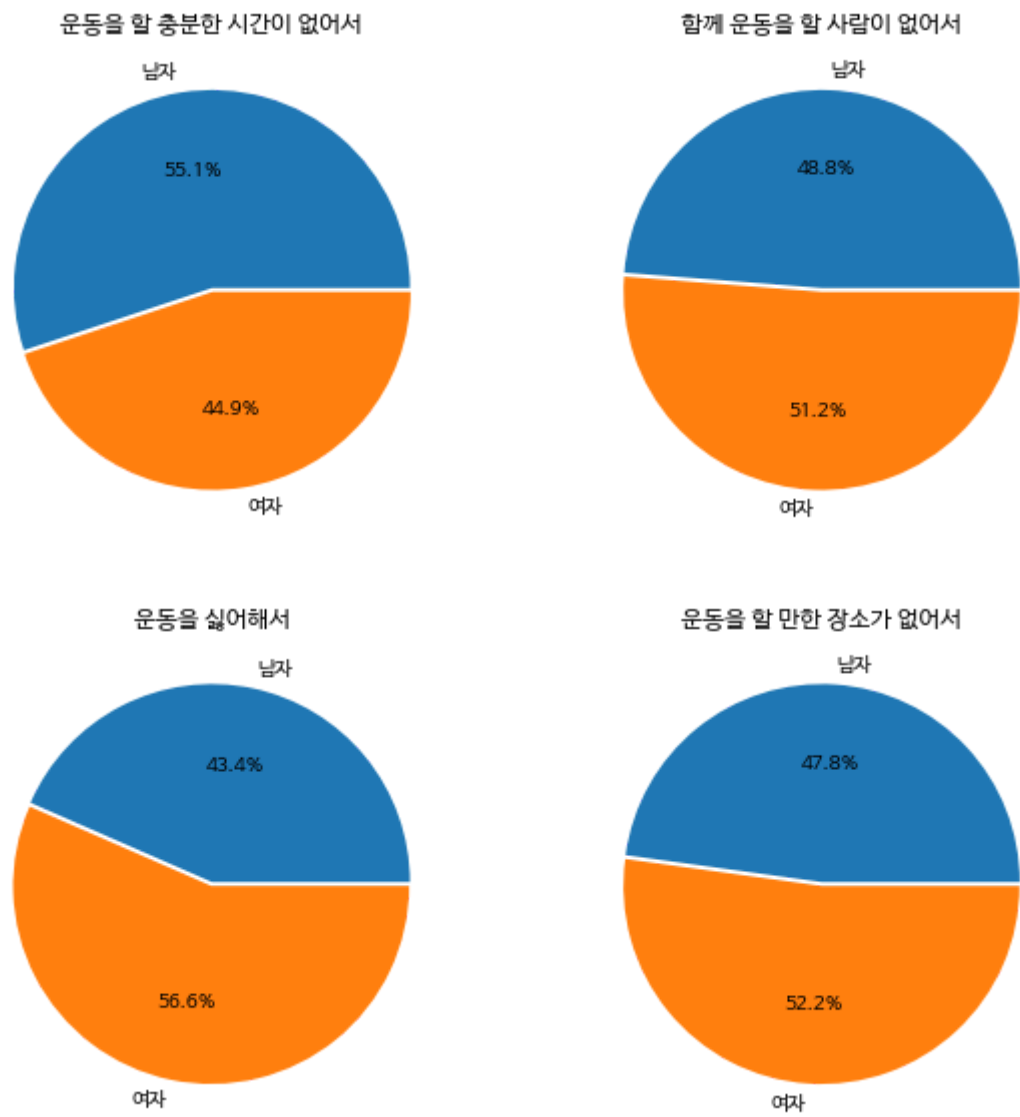
DF_G['함께 운동을 할 사람이 없어서'].plot.pie(explode = [0, 0.02],
                                                    ax = ax[0][1],
                                                    autopct = '%1.1f%%')
ax[0][1].set_title('함께 운동을 할 사람이 없어서')
ax[0][1].set_ylabel('')

DF_G['운동을 싫어해서'].plot.pie(explode = [0, 0.02],
                                   ax = ax[1][0],
                                   autopct = '%1.1f%%')
ax[1][0].set_title('운동을 싫어해서')
ax[1][0].set_ylabel('')
```

```
DF_G['운동을 할 만한 장소가 없어서'].plot.pie(explode = [0, 0.02],
                                                ax = ax[1][1],
                                                autopct = '%1.1f%%')

ax[1][1].set_title('운동을 할 만한 장소가 없어서')
ax[1][1].set_ylabel('')

plt.show()
```



▼ 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)
```

	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
분류					
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
60대 이	37.5	12.8	6.8	46.8	6

▼ 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

	운동을 할 충분한 시간이 없어서	함께 운동을 할 사람이 없어서	운동을 할 만한 장소가 없어서	운동을 싫어해서	기타
분류					
중졸 이하	26.1	13.7	6.6	46.2	7.4
고졸 이하	46.3	8.5	6.9	36.5	1.7
대졸 이상	27.7	77.8	86.5	17.3	90.9

▼ 2) matplotlib 시각화

```
import matplotlib.pyplot as plt
```

```
import matplotlib.pyplot as plt

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

explode_set = [0.02, 0.02, 0.02, 0.02]

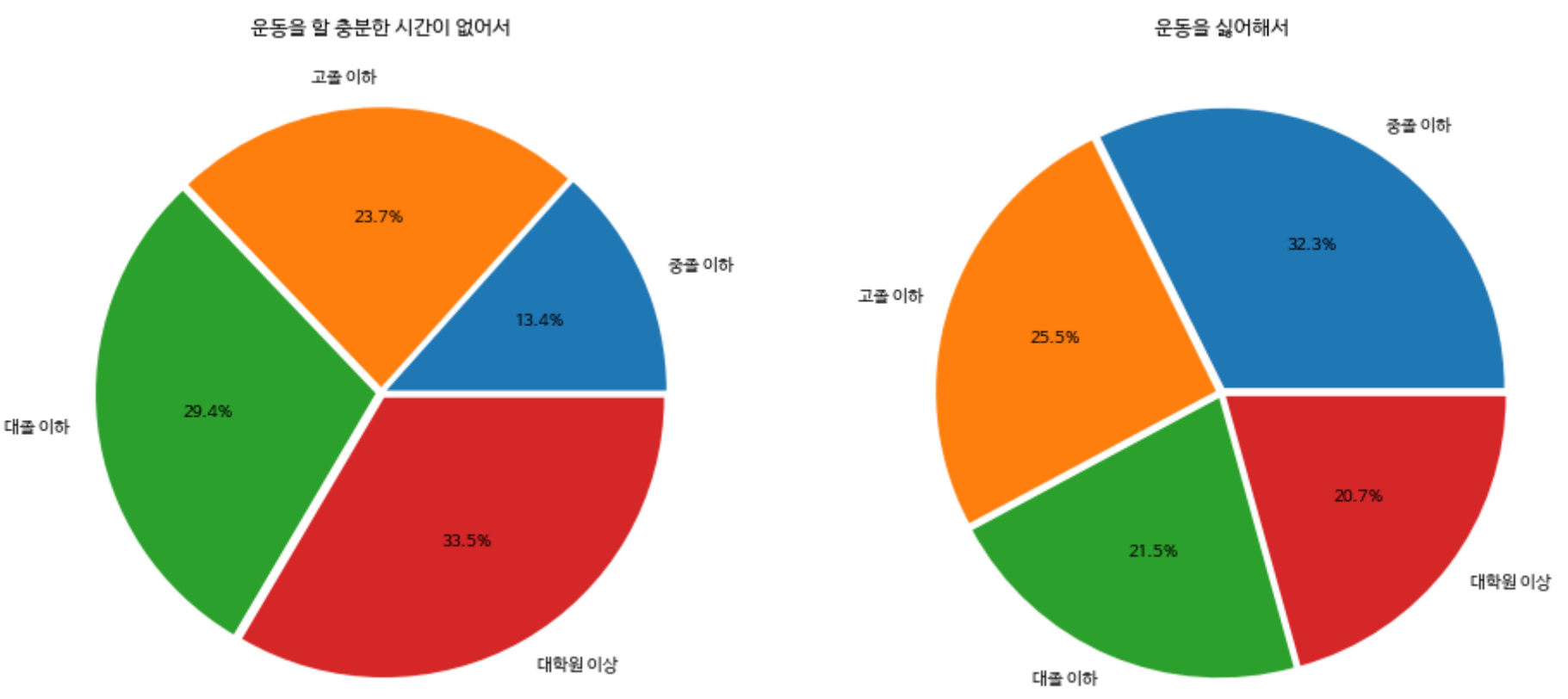
DF_V['운동을 할 충분한 시간이 없어서'].plot.pie(explode = explode_set,
                                                    ax = ax[0],
                                                    autopct = '%1.1f%%')

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

#  
#  
#

