

운동 강도에 따른 혈중 간수치 변화

예측 보고서

202355118 김재진

1. 분석 목적

최근 국민들의 건강의식이 증가함에 따라 각자 서로 다른 목적을 가지고 헬스장에 가서 운동하는 사람들이 많아지고 있다. 하지만 헬스장에 가서 하는 고강도 운동을 많이 할 수록 과연 건강에 이로운지 확인하여 보는 것이 목적이다.

국민영양조사 데이터를 이용해 분석함과 더불어 회귀분석과 통계학습과목에서 배운 내용을 이용하여 분석을 함에 목적이 있다

2. 활용 데이터

국민영양조사(2010년~2023년)데이터에서 간 수치(ast, alt), 혈중 크레아틴 수치, 일주일당 고강도 운동 횟수 데이터를 사용할 것이다.

3. 데이터 분석 과정

데이터를 불러오고 분석하는데에는 파이썬을 사용하였다.

아래는 데이터 분석 과정이다.

1) 파이썬을 이용하여 국민영양조사(2010년 ~ 2023년) excel file을 불러온다.

```
data5 = pd.read_excel('nn14_all.xlsx')
data5
```

	mod_d	ID	ID_fam	year	region	town_t	apt_t	psu	sex	age	...	HEI_TVEG	HEI_VEG	HEI_PROTF	HEI_DAIRY	HEI_SFA	HEI_NA	HEI_SWEET	HEI_CHO	H
0	2020.08.31.	A209716213	A2097162	2014	1	1	1	A209	2	30	...	5.000000	5.000000	7.548706	0.0	10.000000	5.971994	10.000000	5.000000	5.
1	2020.08.31.	A209716214	A2097162	2014	1	1	1	A209	2	56	...	5.000000	5.000000	6.769070	0.0	10.000000	6.795513	10.000000	5.000000	5.
2	2020.08.31.	A209716216	A2097162	2014	1	1	1	A209	1	29	...	5.000000	5.000000	8.626120	10.0	10.000000	0.000000	10.000000	5.000000	5.
3	2020.08.31.	A209724716	A2097247	2014	1	1	1	A209	1	65	...	2.499220	3.273753	10.000000	0.0	10.000000	10.000000	10.000000	5.000000	5.
4	2020.08.31.	A209731515	A2097315	2014	1	1	1	A209	2	57	...	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	5.
...
7545	2020.08.31.	P246811414	P2468114	2014	16	1	2	P246	1	30	...	2.833997	3.377718	4.512896	0.0	10.000000	9.903182	10.000000	4.77224	5.
7546	2020.08.31.	P246811415	P2468114	2014	16	1	2	P246	2	55	...	4.973847	3.342060	2.996269	0.0	10.000000	2.504185	5.731145	0.000000	0.
7547	2020.08.31.	P246811416	P2468114	2014	16	1	2	P246	1	58	...	2.935111	1.726862	2.604355	0.0	10.000000	10.000000	10.000000	0.000000	0.
7548	2020.08.31.	P246816515	P2468165	2014	16	1	2	P246	2	63	...	5.000000	5.000000	10.000000	0.0	1.148236	5.254675	10.000000	0.000000	0.
7549	2020.08.31.	P246816516	P2468165	2014	16	1	2	P246	1	71	...	5.000000	5.000000	6.600000	0.0	10.000000	6.234623	10.000000	5.000000	5.

-> 2010년~2023년 데이터중 2014년 데이터를 불러온 코드 및 데이터

2) 먼저 과도한 운동이 간수치에 미치는 영향을 분석하고자 간수치에 관한 데이터를 정제하였다.

간수치가 증가하는데 영향을 미칠 수 있는 과도한 담배(이번 분석에서는 하루에 1갑(20개비)로 설정함), 과음, 간염 환자는 제외하였다.

```

data1=data1[(data1['BD1_11'] != 5) & (data1['BD1_11'] != 6)] # 과음 인원 제외
data1=data1[(data1['DK8_pr'] != 2)] # B형 간염환자 제외
data1=data1[(data1['DK9_pr'] != 2)] # C형 간염환자 제외
data1=data1[(data1['BS3_2'] <= 20)] # 하루 평균 흡연량이 20개비(1갑)이상은 과도한 흡연으로 판단하고 제외

data2=data2[(data2['BD1_11'] != 5) & (data2['BD1_11'] != 6)]
data2=data2[(data2['DK8_pr'] != 2)]
data2=data2[(data2['DK9_pr'] != 2)]
data1=data1[(data1['BS3_2'] <= 20)]

data3=data3[(data3['BD1_11'] != 5) & (data3['BD1_11'] != 6)]
data3=data3[(data3['DK8_pr'] != 2)]
data3=data3[(data3['DK9_pr'] != 2)]
data1=data1[(data1['BS3_2'] <= 20)]

data4=data4[(data4['BD1_11'] != 5) & (data4['BD1_11'] != 6)]
data4=data4[(data4['DK8_pr'] != 2)]
data4=data4[(data4['DK9_pr'] != 2)]
data1=data1[(data1['BS3_2'] <= 20)]

```

- 과음, 과도한 흡연, 간염환자는 분석 데이터에서 제외하는 코드

3) 데이터정제가 완료 되었으면 고강도 운동 횟수별로 구간을 나눈다.

운동 횟수는 일주일에 1~2회, 3~4회 5~7회로 나누었다.

```

total=pd.concat([data1,data3,data4,data5,data6,data7,data8,data9,data10,data11])
total = total.dropna(subset=['HE_alt', 'HE_ast'])
group_1_2 = total[(total['BE5_1'] >= 2) & (total['BE5_1'] <= 3)][['BE5_1', 'HE_alt', 'HE_ast']]
group_3_4 = total[(total['BE5_1'] >= 4) & (total['BE5_1'] <= 5)][['BE5_1', 'HE_alt', 'HE_ast']]
group_5_7 = total[(total['BE5_1'] == 6)][['BE5_1', 'HE_alt', 'HE_ast']]

```

- 정제 된 데이터를 고강도 운동 횟수별로 나누는 코드

4) 집단이 3개 이므로 ANOVA검정을 시행하여 각 집단의 평균 차이가 유의미 한지 확인하였다.

```

group_1_2_ast = group_1_2['HE_ast']
group_3_4_ast = group_3_4['HE_ast']
group_5_7_ast = group_5_7['HE_ast']
f_statistic, p_value = f_oneway(group_1_2_ast, group_3_4_ast, group_5_7_ast)

p_value

6.836519245913721e-08

group_1_2_alt = group_1_2['HE_alt']
group_3_4_alt = group_3_4['HE_alt']
group_5_7_alt = group_5_7['HE_alt']
f_statistic, p_value = f_oneway(group_1_2_alt, group_3_4_alt, group_5_7_alt)

p_value

0.28539969956737626

```

- ANOVA 검정 코드

5) 다음으로는 과도한 운동이 혈중 크레아틴, 요소질소 농도의 미치는 영향을 분석해보았다.

간 수치 데이터 분석과 마찬가지로 과흡연대상자와 신장 질환이 있는 sample은 제외를 하였다.

```
data1=data1[(data1['DN1_pr'] != 1)]
data1=data1[(data1['DE1_pr'] != 1)]
data1=data1[(data1['BS3_2'] <= 20)]

data2=data2[(data2['DN1_pr'] != 1)]
data2=data2[(data2['DE1_pr'] != 1)]
data1=data1[(data1['BS3_2'] <= 20)]

data3=data3[(data3['DN1_pr'] != 1)]
data3=data3[(data3['DE1_pr'] != 1)]
data1=data1[(data1['BS3_2'] <= 20)]

data4=data4[(data4['DN1_pr'] != 1)]
data4=data4[(data4['DE1_pr'] != 1)]
data1=data1[(data1['BS3_2'] <= 20)]
```

- 과흡연자와 신장질환이 있는 sample은 제외하는 코드

6) 데이터 정제 완료 후 고강도 운동 횟수별로 그룹을 나누었다.

```
total=pd.concat([data1,data3,data4,data5,data6,data7,data8,data9,data10,data11])
total = total.dropna(subset=['HE_BUN', 'HE_crea'])
group_1_2 = total[(total['BE5_1'] >= 2) & (total['BE5_1'] <= 3)][['BE5_1', 'HE_BUN', 'HE_crea']]
group_3_4 = total[(total['BE5_1'] >= 4) & (total['BE5_1'] <= 5)][['BE5_1', 'HE_BUN', 'HE_crea']]
group_5_7 = total[(total['BE5_1'] == 6)][['BE5_1', 'HE_BUN', 'HE_crea']]
```

- 운동 횟수별로 그룹화 한 코드

7) 3개의 집단을 ANOVA검정을 진행하였다.

```

group_1_2_BUN = group_1_2['HE_BUN']
group_3_4_BUN = group_3_4['HE_BUN']
group_5_7_BUN = group_5_7['HE_BUN']
f_statistic, p_value = f_oneway(group_1_2_BUN, group_3_4_BUN, group_5_7_BUN)

p_value

1.8909820306526655e-47

group_1_2_crea = group_1_2['HE_crea']
group_3_4_crea = group_3_4['HE_crea']
group_5_7_crea = group_5_7['HE_crea']
f_statistic, p_value = f_oneway(group_1_2_crea, group_3_4_crea, group_5_7_crea)

p_value

4.8572373455191075e-39

```

- ANOVA검정 진행 코드

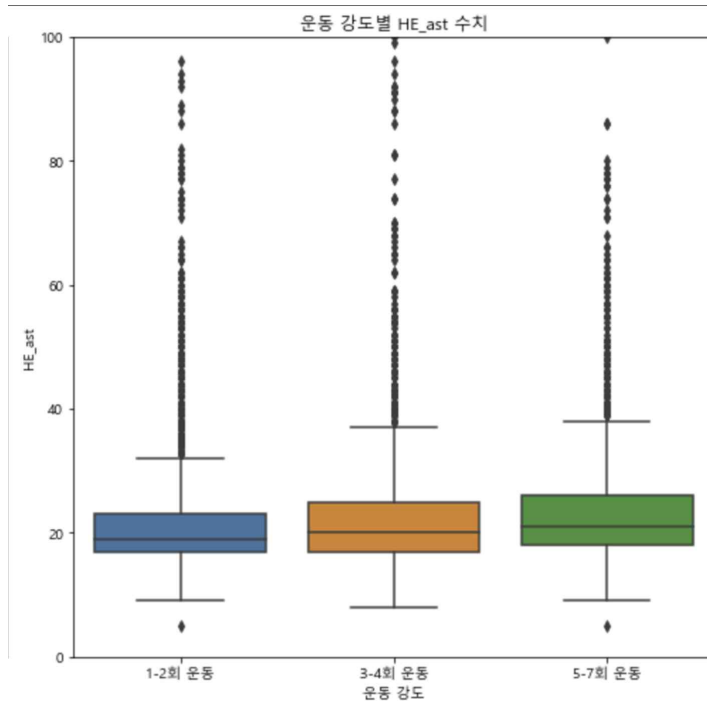
4. 분석결과

먼저 간수치를 이용한 분석 결과 ast수치에서는 p-value가 0.05보다 작아 평균사이의 유의미한 차이가 없다는 귀무가설을 기각하였다. 하지만 alt수치는 평균사이의 유의미한 차이가 없다고 해석할 수 있다. 그렇기에 간수치에서는 운동 횟수를 많이 할수록 ast수치는 증가하는 경향을 보이고 alt수치와는 관계성이 보이지 않는다고 할 수 있다.

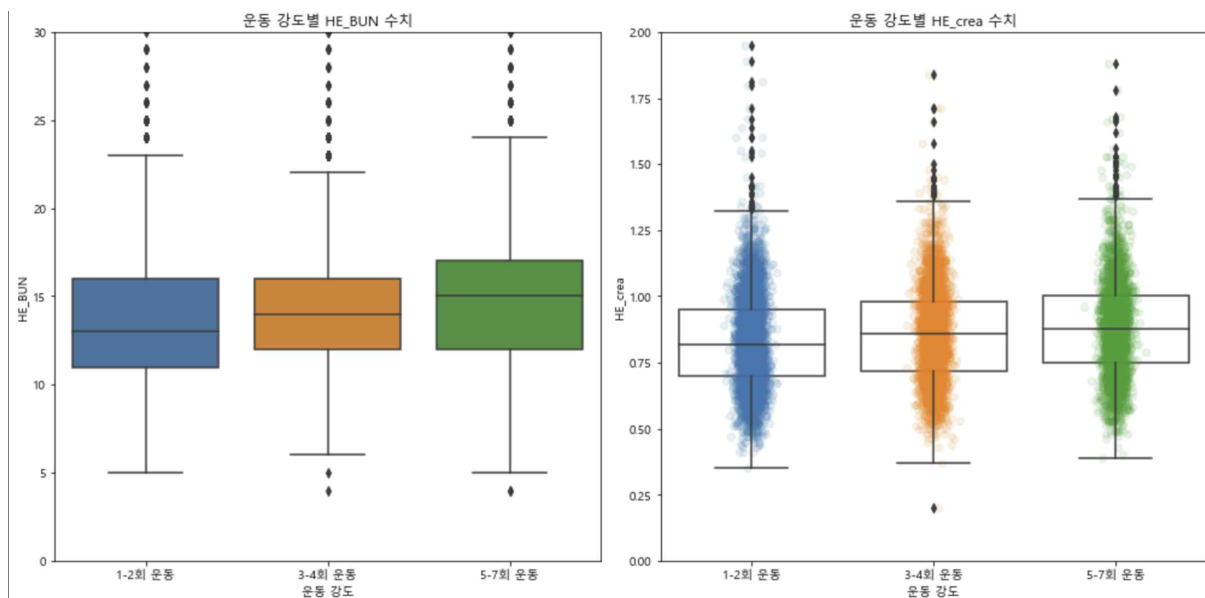
다음으로 혈중 크레아틴과 요소 질소농도를 이용한 분석결과 요소질소(BUN)와 크레아틴(crea) 농도 모두 p-value가 0.05보다 작아 평균값의 차이가 유의미함을 알 수 있다. 그러므로 혈중 크레아틴과 요소질소 농도에서는 운동 횟수를 많이 할 수록 두 수치 모두 증가하는 경향을 보인다고 해석할 수 있다.

5. 결론 및 사후 개선점

간 수치(ast)와 크레아틴, 요소질소 수치(신장건강) 모두에서 유의미한 결과를 보였으나 평균값들은 모두 정상수치 범위 내에서 증가하는 모습이다.



- ast수치를 구간별로 boxplot을 그린 결과 모두 평균이 20에서 조금씩 증가하는 모습을 보임
(ast정상범위는 10~40이다.)



- 각 그룹별 혈중 요소질소(BUN)와 크레아틴(crea)농도를 boxplot으로 그린 결과 평균이 각각의
정상범위인 10~26mg/dL, 0.50~1.4mg/dL 안에서 조금씩 증가하는 경향을 보임

이 결과를 바탕으로 내릴 수 있는 결론은 다음과 같다.

결론: 고강도 운동은 횟수가 많아질 수록 간수치와 혈중 요소질소, 크레아틴농도가 함께 올라간다.

일반적인 사람이라면 이정도 증가는 큰 영향이 없으나 간과 관련된 질환 또는 신장과

관련된 질환이 있는 환자들은 고강도 운동을 할 시에 횡수조절이 필요하다.

사후 개선점:

1. AVOVA검정을 통해 간 수치와 혈중 요소 질소, 크레아틴의 평균변화가 유의미하다는 결론이 나왔지만 각 그룹간 유의미한 차이가 있는지 확인할 수 없었다. 이 그룹간의 차이를 구체적으로 알기 위해 Tukey 검정을 시행할 수 있을듯 하다.
2. 이번 분석에서는 고강도 운동의 횡수로 그룹을 나누어 분석을 하였으나 일주일간 고강도 운동의 전체 시간 또는 횡수 그룹안에서 전체 고강도 운동시간으로 나누어 분석을 진행하면 다른 결과가 나올 수 있을듯 하다.