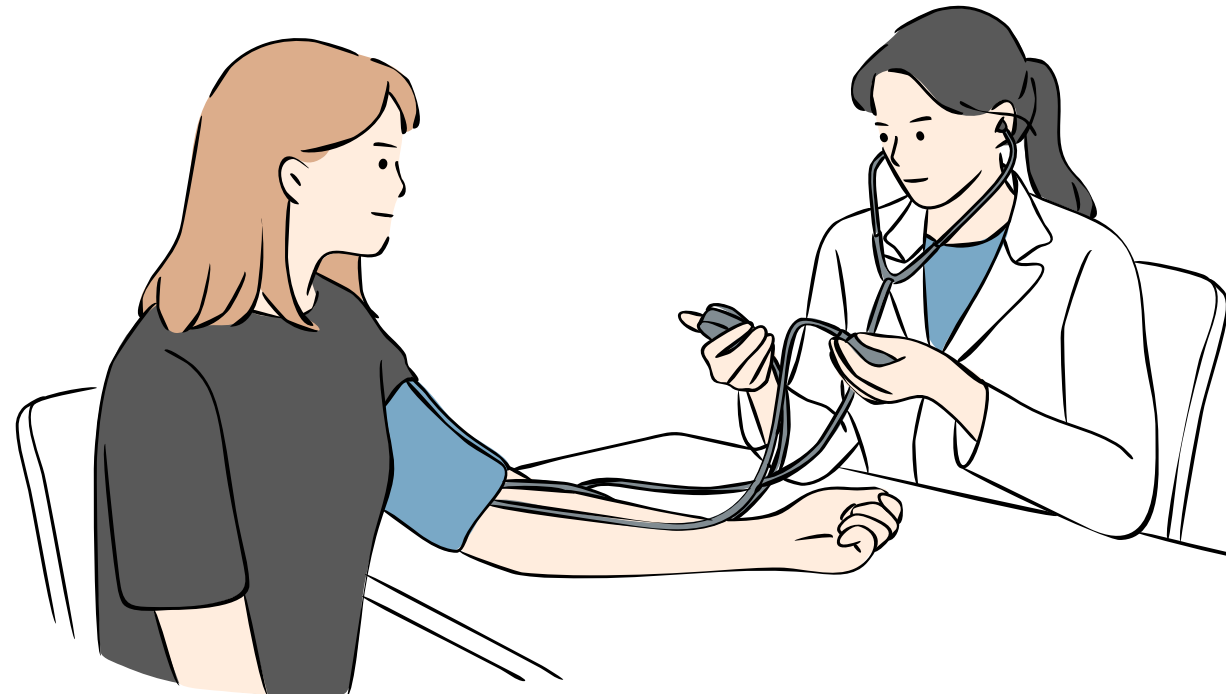

젊은 연령층의 고혈압



일석이조
20202633 김진비
20202638 박선희
20202665 천예은

1. 주제

고혈압

수축기 혈압 140mmHg 이상 또는 이완기 혈압 90mmHg 이상일 때 진단

고령층의 질환

가족력이 강한 질환

1. 주제

고혈압

수축기 혈압 140mmHg 이상 또는 이완기 혈압 90mmHg 이상일 때 진단

고령층의 질환

가족력이 강한 질환

젊은 연령층의 생활 습관

만 12세 이상부터 만 39세 이하의 사람은
앉아서 보내는 시간이 길수록 혈압 수치가 높은가?

2. 데이터

데이터

2020년도 질병관리청에서 실시한 국민건강영양조사 결과

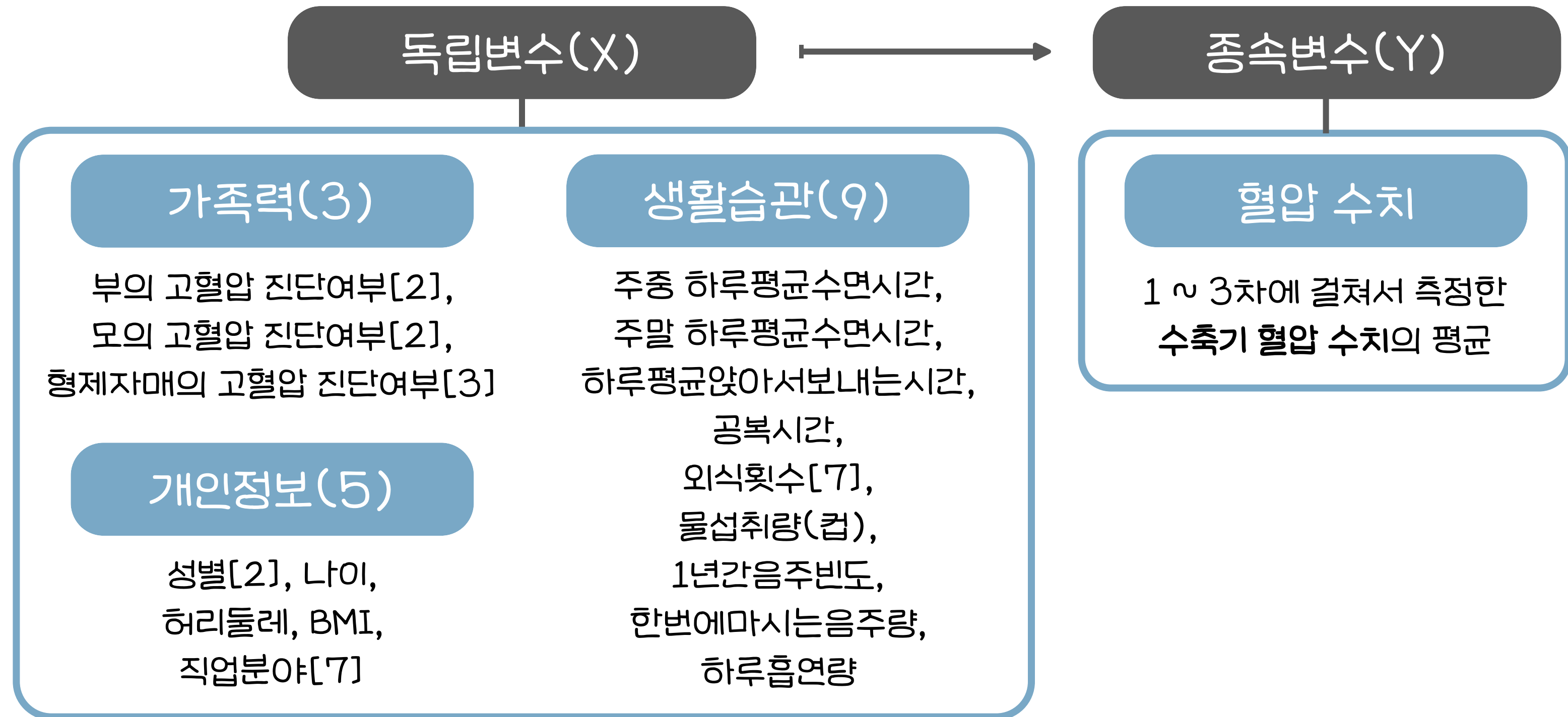
	mod_d	ID	ID_fam	year	region	town_t	apt_t	psu	sex	age	...	LF_S10	LF_S11	LF_S12	LF_S13	LF_S14	LF_S14_1	LF_S15
0	2022.08.10.	A801169401	A8011694	2020.0	1.0	1.0	2.0	A801	1.0	39.0	...	3.0	3.0	3.0	8.0	8.0	8.0	8.0
1	2022.08.10.	A801169402	A8011694	2020.0	1.0	1.0	2.0	A801	2.0	39.0	...	3.0	3.0	3.0	8.0	8.0	8.0	8.0
2	2022.08.10.	A801169403	A8011694	2020.0	1.0	1.0	2.0	A801	1.0	10.0	...	3.0	3.0	3.0	8.0	8.0	8.0	8.0
3	2022.08.10.	A801169404	A8011694	2020.0	1.0	1.0	2.0	A801	1.0	7.0	...	3.0	3.0	3.0	8.0	8.0	8.0	8.0
4	2022.08.10.	A801169405	A8011694	2020.0	1.0	1.0	2.0	A801	1.0	4.0	...	3.0	3.0	3.0	8.0	8.0	8.0	8.0
...
7354	2022.08.10.	R804366601	R8043666	2020.0	8.0	1.0	2.0	R804	2.0	61.0	...	8.0	8.0	8.0	8.0	8.0	8.0	8.0
7355	2022.08.10.	R804368301	R8043683	2020.0	8.0	1.0	2.0	R804	2.0	51.0	...	8.0	8.0	8.0	8.0	8.0	8.0	8.0
7356	2022.08.10.	R804368302	R8043683	2020.0	8.0	1.0	2.0	R804	2.0	24.0	...	8.0	8.0	8.0	8.0	8.0	8.0	8.0
7357	2022.08.10.	R804371701	R8043717	2020.0	8.0	1.0	2.0	R804	1.0	65.0	...	8.0	8.0	8.0	8.0	8.0	8.0	8.0
7358	2022.08.10.	R804371702	R8043717	2020.0	8.0	1.0	2.0	R804	2.0	57.0	...	8.0	8.0	8.0	8.0	8.0	8.0	8.0

7359 rows × 762 columns



[표1] 원본 데이터

3. 변수 선별



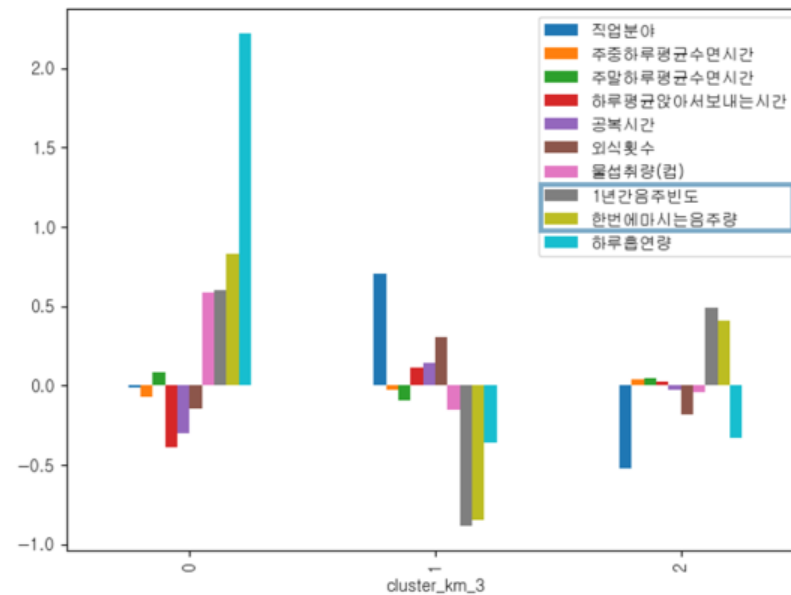
* 변수명[n] : n개의 범주를 가진 범주형 변수

3. 변수 선별

군집분석과 요인분석

개인정보의 **직업분야** 변수와 **생활습관의 9개 변수** 전체(총 10개)를 대상으로,
군집분석과 요인분석을 통해 유의미한 새로운 변수 생성

변수 생성 결과



[그래프1] 군집분석 결과

	0	1	2
직업분야	-0.210843	-0.030618	0.187513
주중하루평균수면시간	-0.071984	0.996782	0.040152
주말하루평균수면시간	-0.050517	0.451355	-0.117994
하루평균아שר보내는시간	-0.177284	-0.065113	-0.135059
공복시간	-0.044507	-0.015201	0.088412
외식횟수	0.136853	-0.076957	1.026385
물섭취량(컵)	0.177587	-0.064229	-0.021109
1년간음주빈도	0.829068	0.020215	0.038485
한번에마시는음주량	0.800138	0.027548	0.017529
하루흡연량	0.323265	-0.043352	-0.048501

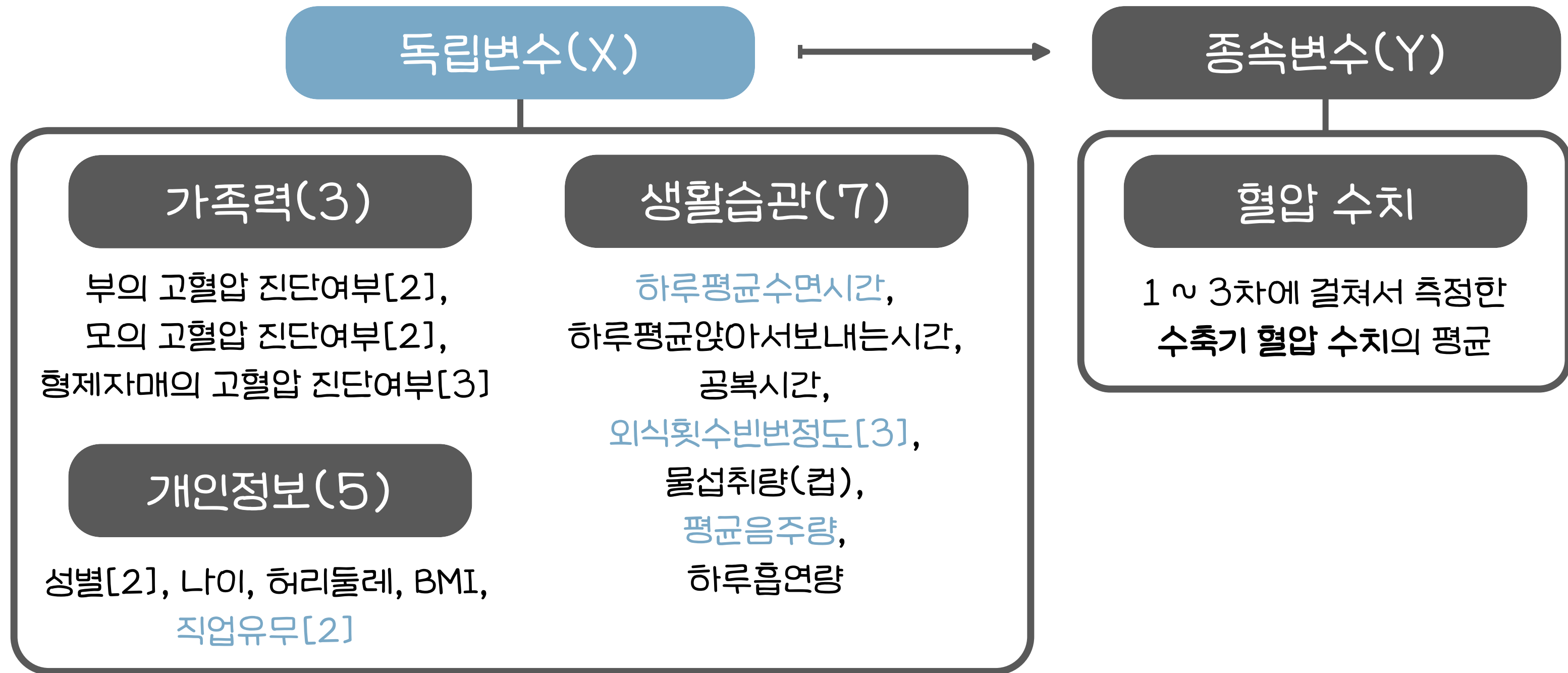
[표2] 요인분석 결과

	직업유무	하루평균수면시간	하루평균아שר보내는시간	공복시간	외식횟수빈번정도	물섭취량(컵)	평균음주량	하루흡연량
0	1	9.0	10.0	13.0	2	10.0	6.5	0.0
1	1	0.0	11.0	13.0	1	1.0	7.0	0.0
2	0	4.0	13.0	15.0	2	4.0	8.5	0.0
...
966	1	30.0	5.0	12.0	1	3.0	9.0	5.0

963 rows × 8 columns

[표3] 생활습관 군집화 결과 (8개 변수)

3. 변수 선별



* 변수명[n] : n개의 범주를 가진 범주형 변수

4. 분석

회귀분석

'앉아서보내는시간' 변수를 포함한 15개의 독립변수가
종속변수인 혈압 수치에 미치는 영향을 회귀분석을 통해 파악

	독립변수															종속변수
	HE_HPfh1	HE_HPfh2	HE_HPfh3	sex	age	HE_wc	HE_BMI	BE8_1	HEfst	N_WAT_C	smoke	L_OUT	occp_yn	Sleep	BP16	HE_sbp
0	0.0	0.0	0.0	1.0	39.0	85.8	24.185489	10.0	13.0	10.0	0.0	2	1	9.0	6.5	116.666667
1	0.0	0.0	0.0	2.0	39.0	68.0	17.935939	11.0	13.0	1.0	0.0	1	1	0.0	7.0	96.000000
2	0.0	0.0	0.0	1.0	18.0	63.5	16.398308	13.0	15.0	4.0	0.0	2	0	4.0	8.5	133.333333
3	0.0	0.0	0.0	2.0	39.0	68.5	18.940652	10.0	3.0	2.0	0.0	1	1	4.0	7.0	104.666667
5	0.0	0.0	0.0	2.0	26.0	82.0	22.864261	12.0	15.0	5.0	0.0	1	1	20.0	9.0	121.333333
...
962	1.0	0.0	0.0	1.0	31.0	79.2	21.916303	12.0	18.0	10.0	20.0	1	1	2.0	6.0	104.666667
963	1.0	0.0	0.0	1.0	39.0	82.4	23.381186	13.0	15.0	4.0	0.0	2	1	9.0	8.0	114.000000
964	0.0	0.0	0.0	2.0	30.0	97.6	29.159847	10.0	15.0	5.0	0.0	2	1	3.0	5.0	120.000000
965	0.0	0.0	8.0	2.0	26.0	59.4	17.570804	12.0	12.0	2.5	0.0	2	1	6.0	7.0	104.666667
966	0.0	0.0	0.0	2.0	24.0	75.5	26.129927	5.0	12.0	3.0	5.0	1	1	30.0	9.0	103.333333

963 rows × 16 columns

4. 분석

1차 모델 결과

	coef	std err	t	P> t	[0.025	0.975]
Intercept	84.6730	4.206	20.134	0.000	76.420	92.926
HE_HPfh1[T.1.0]	3.6610	0.721	5.075	0.000	2.245	5.077
HE_HPfh2[T.1.0]	4.0315	0.886	4.553	0.000	2.294	5.769
HE_HPfh3[T.1.0]	-0.5181	2.489	-0.208	0.835	-5.403	4.367
HE_HPfh3[T.8.0]	-1.5299	1.361	-1.124	0.261	-4.201	1.141
sex[T.2.0]	-6.4185	0.830	-7.737	0.000	-8.046	-4.791
L_OUT[T.1]	-0.2276	1.220	-0.186	0.852	-2.622	2.167
L_OUT[T.2]	-0.1805	1.218	-0.148	0.882	-2.571	2.210
occp_yn[T.1]	-1.2023	0.712	-1.688	0.092	-2.600	0.195
age	0.0077	0.052	0.148	0.882	-0.094	0.109
HE_wc	0.0637	0.086	0.743	0.458	-0.105	0.232
HE_BMI	0.8466	0.218	3.881	0.000	0.418	1.275
BE8_1	-0.1362	0.089	-1.529	0.127	-0.311	0.039
HE_fst	-0.2040	0.118	-1.727	0.084	-0.436	0.028
N_WAT_C	0.2096	0.096	2.189	0.029	0.022	0.398
smoke	-0.0530	0.063	-0.848	0.397	-0.176	0.070
Sleep	0.1659	0.042	3.926	0.000	0.083	0.249
BP16	0.6658	0.253	2.631	0.009	0.169	1.162

해석

변수의 회귀 계수가 0에 가까워
설명력이 떨어지는 독립변수에 대해 제거가 필요.

4. 분석

최종 선별 변수

	독립변수										종속변수
	HE_HPfh1	HE_HPfh2	sex	occp_yn	HE_BMI_log	BE8_1	HE_fst	N_WAT_C	Sleep	BP16	HE_sbp
0	0.0	0.0	1.0	1	3.185753	10.0	13.0	10.0	9.0	6.5	116.666667
1	0.0	0.0	2.0	1	2.886806	11.0	13.0	1.0	0.0	7.0	96.000000
2	0.0	0.0	1.0	0	2.797178	13.0	15.0	4.0	4.0	8.5	133.333333
3	0.0	0.0	2.0	1	2.941310	10.0	3.0	2.0	4.0	7.0	104.666667
5	0.0	0.0	2.0	1	3.129575	12.0	15.0	5.0	20.0	9.0	121.333333
...
962	1.0	0.0	1.0	1	3.087231	12.0	18.0	10.0	2.0	6.0	104.666667
963	1.0	0.0	1.0	1	3.151932	13.0	15.0	4.0	9.0	8.0	114.000000
964	0.0	0.0	2.0	1	3.372793	10.0	15.0	5.0	3.0	5.0	120.000000
965	0.0	0.0	2.0	1	2.866239	12.0	12.0	2.5	6.0	7.0	104.666667
966	0.0	0.0	2.0	1	3.263081	5.0	12.0	3.0	30.0	9.0	103.333333

961 rows × 11 columns

최종 모델 결과

	coef	std err	t	P> t	[0.025	0.975]
Intercept	32.1035	6.533	4.914	0.000	19.283	44.924
HE_HPfh1[T.1.0]	3.7081	0.701	5.289	0.000	2.332	5.084
HE_HPfh2[T.1.0]	3.4335	0.864	3.974	0.000	1.738	5.129
sex[T.2.0]	-6.5413	0.670	-9.765	0.000	-7.856	-5.227
occp_yn[T.1]	-1.3219	0.638	-2.071	0.039	-2.574	-0.069
HE_BMI_log	24.7858	1.862	13.310	0.000	21.131	28.440
BE8_1	-0.1241	0.085	-1.464	0.144	-0.290	0.042
HE_fst	-0.2135	0.115	-1.864	0.063	-0.438	0.011
N_WAT_C	0.2040	0.093	2.185	0.029	0.021	0.387
Sleep	0.1484	0.040	3.705	0.000	0.070	0.227
BP16	0.6430	0.245	2.627	0.009	0.163	1.123

4. 분석

최종 모델 결과

	coef	std err	t	P> t	[0.025	0.975]
Intercept	32.1035	6.533	4.914	0.000	19.283	44.924
HE_HPfh1[T.1.0]	3.7081	0.701	5.289	0.000	2.332	5.084
HE_HPfh2[T.1.0]	3.4335	0.864	3.974	0.000	1.738	5.129
sex[T.2.0]	-6.5413	0.670	-9.765	0.000	-7.856	-5.227
occp_yn[T.1]	-1.3219	0.638	-2.071	0.039	-2.574	-0.069
HE_BMI_log	24.7858	1.862	13.310	0.000	21.131	28.440
BE8_1	-0.1241	0.085	-1.464	0.144	-0.290	0.042
HE_fst	-0.2135	0.115	-1.864	0.063	-0.438	0.011
N_WAT_C	0.2040	0.093	2.185	0.029	0.021	0.387
Sleep	0.1484	0.040	3.705	0.000	0.070	0.227
BP16	0.6430	0.245	2.627	0.009	0.163	1.123

해석

'앉아서보내는시간' 변수의 회귀계수에 대한 p-value 값이 0.144로 '혈압 수치'에 유의미한 영향을 끼치지 않음.

4. 분석

젊은 연령층의 생활 습관

만 12세 이상부터 만 39세 이하의 사람은
앉아서 보내는 시간이 길수록 혈압 수치가 높은가?

결론

만 12세 이상부터 만 39세 이하의 사람의 앉아서 보내는 시간은
혈압 수치에 유의미한 영향을 끼치지 않으므로 그렇지 않다.

감사합니다
