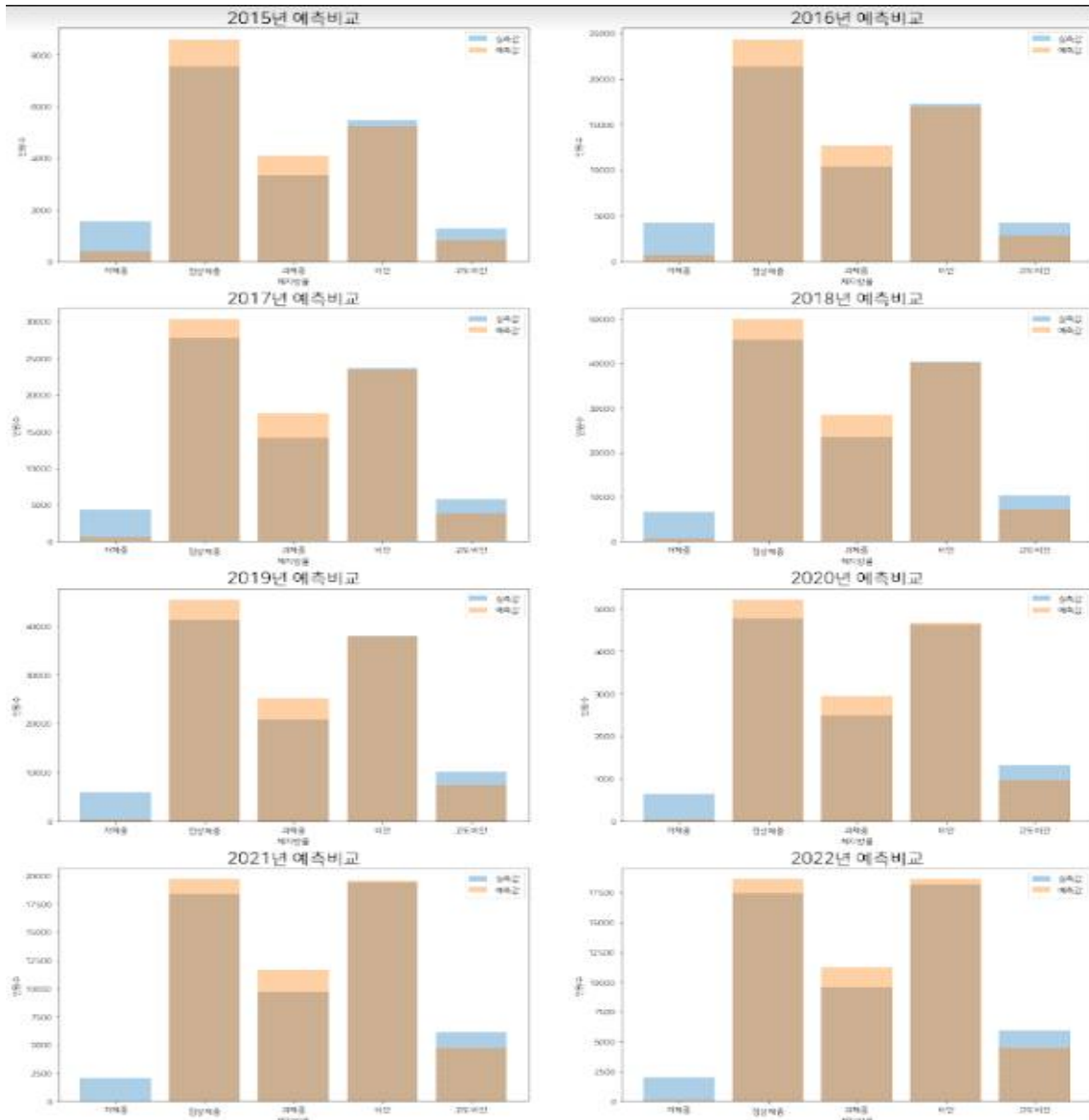


# 1. 체지방률 실측-예측 연도별 차이에 대한 그래프

train\_X = 2015~2022년 데이터, columns = '측정구분', '측정나이', '입력구분', '측정년', '측정월', '측정 회원성별', '신장', '체중', '이완기혈압(최저)', '수축기혈압(최고)', '지역명'



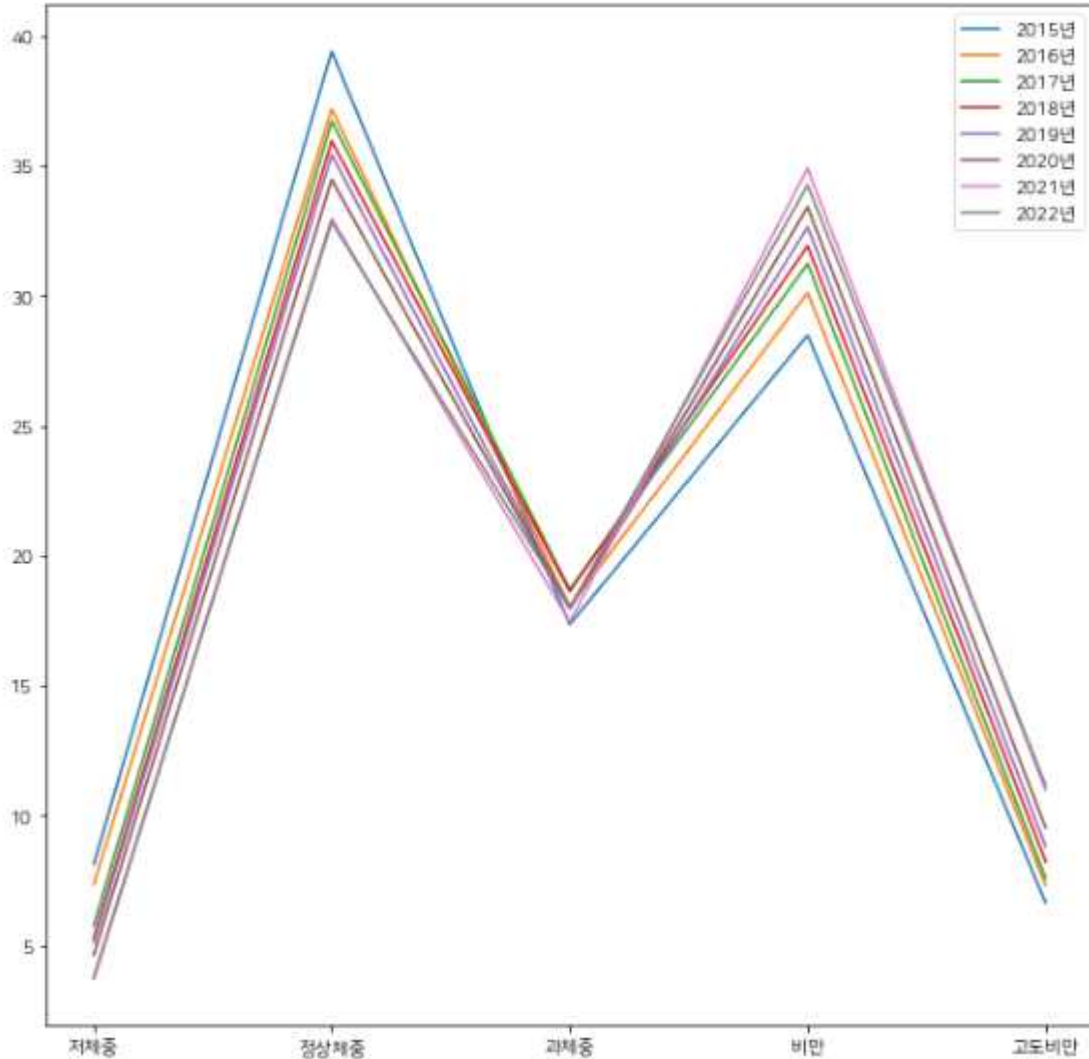
=> 연도별로 예측율의 차이를 기대했으나 train\_X를 전체데이터(15~22)년으로 두고 예측을 진행하여 유의 미한 차이가 나지 않은 듯 싶음,

	mae	mse	rmse	r2
2015	0.397749	0.483323	0.695214	0.613542
2016	0.398380	0.483583	0.695402	0.614587
2017	0.395090	0.477644	0.691118	0.606448
2018	0.397329	0.485595	0.696846	0.601581
2019	0.387280	0.471234	0.686465	0.618919
2020	0.389290	0.484555	0.696100	0.608612
2021	0.389152	0.476955	0.690619	0.615223
2022	0.390770	0.473718	0.688272	0.618443

	mae	mse	rmse	r2
2015	3.147862	16.939223	4.115729	0.814019
2016	3.184733	17.471345	4.179874	0.808999
2017	3.159582	16.875827	4.108020	0.807531
2018	3.197156	17.389173	4.170033	0.803205
2019	3.148476	16.843724	4.104111	0.813425
2020	3.164546	16.816762	4.100825	0.818950
2021	3.235454	17.526332	4.186446	0.810515
2022	3.217382	17.230129	4.150919	0.811163

체지방률(등급) 오차지표      체지방률(수치) 오차지표  
=> 겹치는부분의 정도를 파악하고 싶어 오차지표를 계산했으나 별 차이 없었음

## 2. 실제 체지방율에 대한 그래프



15년도 기준 정상체중이 가장 높고 그 뒤 16년,17년,18년, 19년, 20년,21년,22년 순으로 오름차순 형태를 띄고 있음, 시간의 흐름에 따라 '정상체중'비율이 줄어든다고 파악할 수 있음

'비만' 항목을 볼 때 21년이 가장 높고 그 뒤 22년, 20년, 19년,18년,17년,16,15년 내림차순으로 내려가는 것을 확인할 수 있음, 시간의 흐름에 따라 '비만'등급이 증가한다고 파악할 수 있음.

'고도비만'항목도 마찬가지로 연도별로 내림차순의 형태를 띄고 있음

	저체중	정상체중	과체중	비만	고도비만
2015	8.130081	39.394413	17.349385	28.476131	6.649990
2016	7.348826	37.175186	18.047034	30.102724	7.326230
2017	5.777972	36.710533	18.696158	31.227208	7.588130
2018	5.281061	35.950132	18.627684	31.907858	8.233265
2019	5.140111	35.415861	17.994686	32.636005	8.813338
2020	4.640589	34.461605	17.956120	33.407910	9.533776
2021	3.727834	32.958217	17.451002	34.889439	10.973509
2022	3.764546	32.837932	17.989099	34.250580	11.157843

연도별 체지방등급 청소년인구율(%)

## 연도별 체지방율 평균 편차

	2015	2016	2017	2018	2019	2020	2021	2022
count	19188.000000	57533.000000	75684.000000	126414.000000	116301.000000	13856.000000	55716.000000	53021.000000
mean	23.820644	24.395752	24.985206	25.092411	25.236240	25.570930	26.019374	26.142193

## 연도별 체지방율(등급)의 평균값 변화

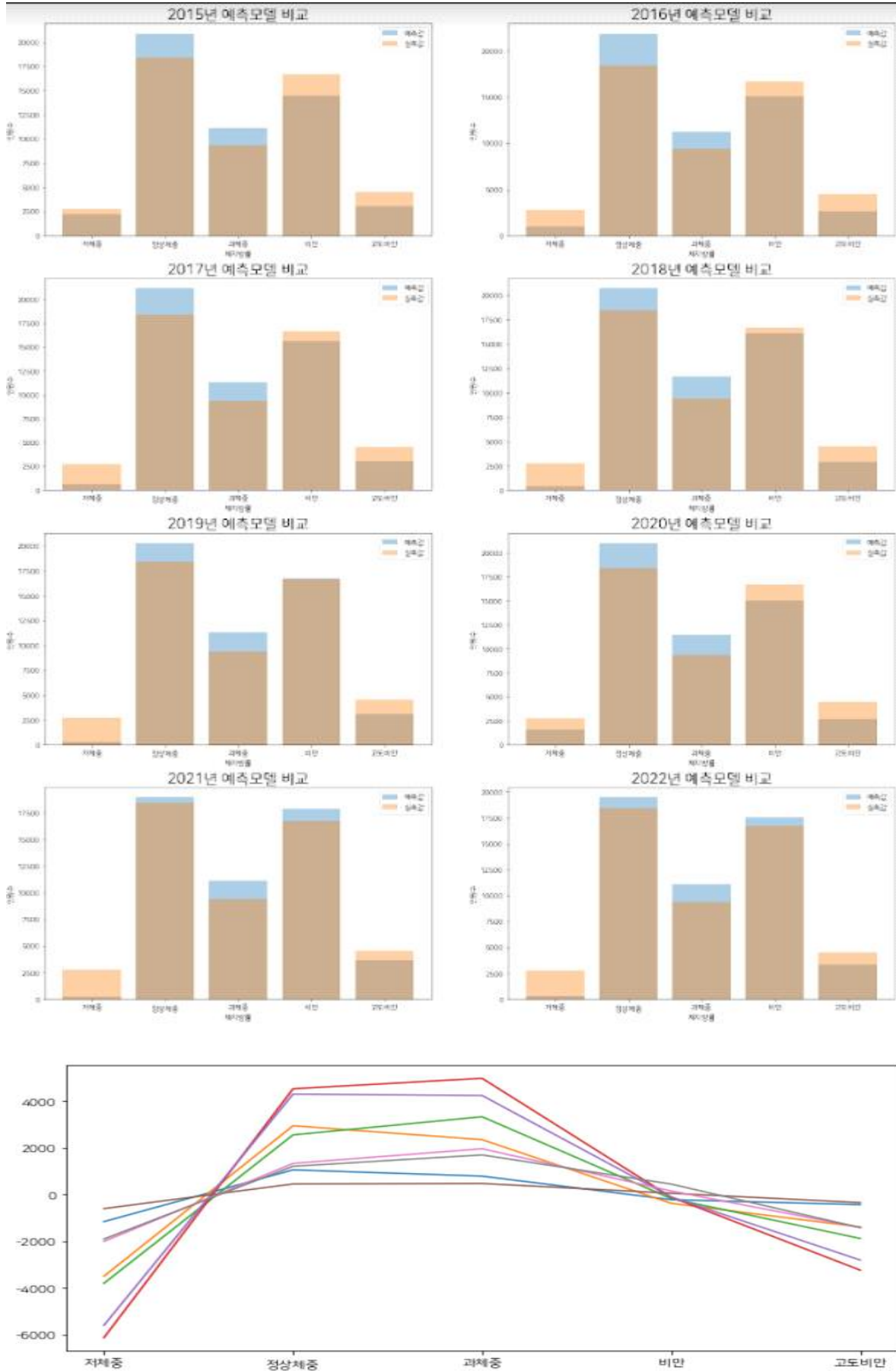
	저체중	정상체중	과체중	비만	고도비만
2015	9.953693	16.650583	25.930929	32.554450	40.344374
2016	9.583522	16.789938	25.905161	32.565995	40.558652
2017	10.055495	17.062754	25.944642	32.713471	40.513542
2018	9.941756	16.855768	25.704896	32.516650	40.617143
2019	9.699657	16.732299	25.703956	32.505606	40.596442
2020	9.475988	16.660428	25.538586	32.603651	41.030961
2021	9.488734	16.549374	25.356405	32.372733	40.931843
2022	9.660220	16.793749	25.459096	32.460416	40.922488

원래는 전체적으로 평균값이 올라가길 기대했으나....

정상체중,과체중,비만 은 범위가 정해져있으므로 값이 많이 변화하면 표본에서 이탈하므로 평균값에 변화가 없는 것이 맞음

저체중과 고도비만도 실질적인 저체중과 고도비만의 인구가 늘어났더라도 고도비만인구 중 낮은 체지방율(ex.36)을 가진 새로운 인원이 많이 유입되면 오히려 평균치가 낮아지는 효과를 가지므로 크게 변화가 없음

## 7. 연도별 모델로 동일한 측정값에 대해 비교



20년 모델이 가장 오차가 적게 나옴