

7월 8일 산보연 보고사항

1) 중분류별 SEX, CAL, ECNY-AGE 의 비율 연도에 따른 변화 시각화!

; 이전과 같은 형태로 그리기!

2) 중분류별 선형 회귀 모형 적합해 forecasting 시행 후 결과 시각화!

; train Data는 2000년 ~ 2014년, validation data는 2015년 ~ 2018년으로 Split

; 반응변수는 "통합 누적 발생률", 설명변수는 "YEAR"로 linear Regression model 적합!

; Performance Criteria는 " $MAPE = \frac{1}{4} \sum_{t=2015}^{2018} \left| \frac{y_{it} - \hat{y}_{it}}{y_{it}} \right|$ (i는 UP2, t는 시점)" 사용 / 나중에 비교할 수 있도록 따로 저장!

; 시각화 시, 실제 통합 누적 발생률과 fitted line 구분되도록, 2014년 기준으로 trend line 구분되도록 시각화!

/ MAPE를 graph 안 text로 표시!