

Lag Variable 추가

문제점 : 1~8주를 각각 다른 컬럼으로 설정할 경우, 시계열 데이터의 특성인 time lag를 고려할 수 없게 된다.

Rolling function을 사용하여 변수를 생성

1~8주의 데이터에 대해 window를 sliding하며 대푯값(mean, std..)을 계산한다.

이를 통해 같은 window내에 존재하였다면 lag와 관계없이 같은 대푯값을 취하여, lag를 완화시킬 수 있다. (시도한 대푯값 : mean, std, sum, diff, beta coefficient of simple linear regression)

