<ARIMA 모델을 통한 전력소비량 예측>

문제1. Train 데이터는 시간별로 모든항목이 기재되어있었지만, test 데이터 같은 경우 기온, 풍속, 습도, 강수량, 일조는 3시간 단위로 기재되어있었고, 비전기냉방설비운영과 태양광보유는 모두 누락되어 있었다.

* 기온, 풍속, 습도, 강수량, 일조는 연속적으로 변하는 값이기 때문에 선형보간을 통해 결측치를 처리해 주었고, 비전기냉방설비운영과 태양광보유는 train 데이터와 동일하게 하여 결측치를 처리하였다.

문제2. 데이터가 위치가 명시되지 않은 건물의 전력사용량과 기상데이터였기 때문에, 지역이 어디인지 정확히 알 수 없어, 응용기상을 적용할 때 기상청 데이터를 사용할 수 없었다.

* 체감온도와, 불쾌지수를 적용하기 위해서 기상청의 데이터가 아닌, 기존의 기온. 풍속, 습도로 파생변수를 만들어 train 데이터와 test 데이터에 추가한 뒤 전력사용량 예측에 있어서 기상요소의 다양화를 반영할 수 있었다.