EDA하면서 내 생각 정리

1. 건물별로 시간과 요일에 따른 전력사용량 분석 (단순 kmeans 적용)

* other : 기타건물이라 그런지 진짜 중구난방, 모델링 할 때 가장 어려워보임
* public : 16번을 제외하고 모두 비슷한 형태, 16번은 주말에도 오픈하는 공공시설로 보임
* university : 대체적으로 유사한 형태, public보다 약 2시간정도 오른쪽으로 밀려있는 형태 (늦은 시간에 오픈, 늦은 시간에 마감을 의미)
* data center : 대체적으로 유사한 형태, 매일 10시부터 자정까지 사용량 높음
* department store : 서로 유사한 정도가 가장 큼, 9시오픈 20시마감, 주말 사용량 높음이 매우 뚜렷하게 보임
* hospital : 서로 매우 유사, 45,50번은 토요일에 오전 진료를 하는 것으로 보임
* commercial : 전체적으로는 유사, 54를 제외한 나머지 건물은 유사한 성격을 띄고 있을 것으로 보이고, 그 중 57은 주말에도 오픈, 54는 월요일 제외 오픈으로 아예 다른 성격의 건물로 추정
* apartment : 백화점과 마찬가지로 서로 유사한 정도가 매우 큼, 사람들이 일을 하는 평일 낮에는 전력소비량이 많지 않고, 주말 낮에는 쉬기 때문에 올라감, 저녁 시간에는 평일과 주말 모두 매우 높음, 자는 시간인 밤과 새벽은 전력소비량 거의 없음
* research : 71, 75 제외 모두 유사한 형태, 평일 낮 시간에만 많고, 그 외에는 거의 없다고 봐도 무방, 71번은 토요일 10시와 14시만 낮고 그 외에는 어느 정도 고른 형태를 띄고 있음, 75번은 월요일이 상대적으로 낮고, 주말에도 연구를 하는 것으로 보임
* knowledge : 81 제외 모두 유사한 형태, 81번은 다른 센터들에 비해 오픈 시간이 1시간정도 늦어보이고, 대신 마감 시간이 23시정도로 매우 늦으며, 주말에 오히려 좀 더 많은 전력소비 발생
* mart : 거의 다 유사한 형태, 91번은 운영시간이 좀 더 짧음
* hotel : 거의 다 유사한 형태, 아침 시간부터 밤시간까지 평일 주말 모두 고른 형태를 띄고 있음
* 종합 : public, commercial, research, knowledge 등은 유사한 형태, 그 외 건물들은 각 건물들에 특성에 맞는 분포가 나타남, 하지만 각 건물 타입 내에서도 특이한 건물들이 존재함
* 클러스터링을 9개로 해보고 비교해볼예정

1. TimeSeriesKMeans 사용해서 클러스터링 수행

