Test_0825

💡 테스트 공통 조건

- compare_model() 사용 X : et 지정 단일 모델 사용
- 클러스터 통일
- 변수 유지(work_time, …)

💡 테스트 항목 : 변수 추가

▼ 전력소비량 관련 파생변수 추가

- 추가 변수
  - 건물별, 요일별, 시간대별 평균
  - 건물별, 시간대별 평균
  - 건물별, 시간대별 표준편차

▼ 변수 중요도

▼ 테스트 : pycaret

💡 테스트 항목 : 변수 선택

▼ feature importance 0인 변수 제외

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean','sy_en','date_sin','ess_capacity','building_number','work_time','weekday_cos',
 'building_type_Research Institute','cooling_area','holiday','temperature_mean','hour_std',
 'building_type_University','solar_power_capacity','particular','total_area','discomfort','hour_mea
n',
 'discomfort_mean','hour_cos','CDH_mean','weekday_sin','CDH','ma_dis_5','ma_dis_3','pcs_capacity',
 'low_day','hour_sin','humidity','temperature_3','sensory_temperature','windspeed','discomfort_3','te
mperature']
```

- Cluster 1

```
['weekday_hour_mean','date_sin','hour_mean','discomfort','work_time','weekday_sin','total_area','CDH_
mean',
 'building_number','temperature_mean','hour_std','discomfort_mean','cooling_area','humidity','ma_dis_
5',
 'weekday_cos','holiday','ma_dis_3','CDH','hour_sin','sensory_temperature','temperature','discomfort_
3',
 'hour_cos','temperature_3','windspeed','solar_power_capacity']
```

- Cluster 2

```
['weekday_hour_mean','building_number','hour_mean','date_sin','ma_dis_5','cooling_area','discomfor
t','hour_sin',
```

```
 'total_area','hour_std','CDH','hour_cos','humidity','holiday','discomfort_mean','CDH_mean','weekday_
cos',
 'weekday_sin','temperature_mean','sensory_temperature','temperature_3','ma_dis_3','work_time','winds
peed',
 'discomfort_3','temperature']
```

- Cluster 3

```
['weekday_hour_mean','date_sin','holiday','hour_std','weekday_sin','hour_mean','total_area','discomfo
rt',
 'ma_dis_5','low_day','CDH','building_type_Commercial','ma_dis_3','cooling_area','building_number',
 'sensory_temperature','building_type_Hospital', 'temperature_mean','CDH_mean','weekday_cos','discomf
ort_mean',
 'solar_power_capacity','work_time','sy_en','ess_capacity', 'hour_cos','sy_ey','hour_sin','humidit
y','windspeed',
 'temperature_3','discomfort_3','temperature']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0

  `SMAPE : 4.08768`

- Cluster 1

  `SMAPE : 4.67567`

- Cluster 2

  `SMAPE : 2.31264`

- Cluster 3

  `SMAPE : 2.93476`

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et

  `SMAPE : 4.09256`

- Cluster 1 : et

  `SMAPE : 4.69697`

- Cluster 2 : Catboost

  `SMAPE : 2.19484`

- Cluster 3 : et

  `SMAPE : 2.94628`

▼ Shap 결과 전부 사용

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean', 'date_sin', 'CDH', 'temperature_mean', 'building_number', 'hour_std', 'hour_mean',
 'discomfort_mean', 'discomfort', 'holiday', 'CDH_mean', 'ma_dis_3', 'hour_sin', 'total_area', 'cooling_area',
 'weekday_sin', 'weekday_cos', 'ma_dis_5', 'solar_power_capacity', 'temperture']
```

- Cluster 1

```
['weekdat_hour_mean', 'date_sin', 'discomfort', 'temperature_mean', 'discomfort_mean', 'hour_std', 'CDH_mean',
 'hour_mean', 'CDH', 'building_number', 'humidity', 'total_area', 'cooling_area', 'ma_dis_5', 'temperature',
 'hour_sin', 'weekday_cos', 'ma_dis_3', 'weekday_sin', 'sensory_temperature']
```

- Cluster 2

```
['weekday_hour_mean', 'hour_mean', 'building_number', 'date_sin', 'ma_dis_5', 'discomfort', 'CDH', 'hour_sin',
 'hour_std', 'humidity', 'total_area', 'cooling_area', 'discomfort_mean', 'weekday_cos', 'hour_cos', 'holiday',
 'CDH_mean', 'temperature_3', 'temperature_mean', 'ma_dis_3']
```

- Cluster 3

```
['weekday_hour_mean', 'date_sin', 'holiday', 'hour_mean', 'hour_std', 'CDH', 'sensory_temperature', 'discomfort',
 'hour_sin', 'weekday_sin', 'weekday_sin', 'temperature_mean', 'ma_dis_5', 'total_area', 'building_number',
 'CDH_mean', 'discomfort_mean', 'ma_dis_3', 'cooling_area', 'weekday_cos', 'humidity']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0

  `SMAPE : 4.08973`

- Cluster 1

  `SMAPE : 4.66345`

- Cluster 2

  `SMAPE : 2.27170`

- Cluster 3

  `SMAPE : 2.94844`

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et

  `SMAPE : 4.07536`

- Cluster 1 : et

  `SMAPE : 4.65229`

- Cluster 2 : CatBoost

  `SMAPE : 2.12017`

- Cluster 3

  `SMAPE : 2.95855`

▼ Shap 결과 일부 사용

▼ Shap 결과 내 맘대로 제거

▼ 사용 변수

- Cluster 0

```
['weekday_hour_mean', 'date_sin', 'CDH', 'temperature_mean', 'building_number',
 'hour_std', 'hour_mean', 'discomfort_mean', 'discomfort', 'holiday', 'CDH_mean',
 'ma_dis_3', 'hour_sin', 'total_area', 'cooling_area', 'cooling_area','weekday_sin',
 'weekday_cos', 'ma_dis_5', 'solar_power_capacity']
```

- Cluster 1

```
['weekday_hour_mean', 'date_sin', 'discomfort', 'temperature_mean', 'discomfort_mean',
 'hour_std', 'CDH', 'building_number', 'humidity', 'total_area', 'ma_dis_5', 'temperature',
 'hour_sin', 'weekday_cos', 'ma_dis_3']
```

- Cluster 2

```
['weekday_hour_mean', 'hour_mean', 'building_number', 'date_sin', 'ma_dis_5',
 'discomfort', 'CDH', 'hour_sin', 'hour_std', 'humidity', 'total_area']
```

- Cluster 3

```
['weekday_hour_mean', 'date_sin', 'holiday', 'hour_mean', 'hour_std', 'CDH',
 'sensory_temperature', 'discomfort', 'hour_sin', 'weekday_sin', 'temperature_mean',
 'ma_dis_5', 'total_area', 'building_number', 'CDH_mean', 'discomfort_mean']
```

▼ 테스트 : Pycaret - et 지정 모델

- Cluster 0

  `SMAPE : 4.05002`

- Cluster 1

  `SMAPE : 4.73512`

- Cluster 2

  `SMAPE : 2.43135`

- Cluster 3

  `SMAPE : 2.92323`

▼ 테스트 : Pycaret - compare model

- Cluster 0 : et

  `SMAPE : 4.05499`

- Cluster 1 : et

  `SMAPE : 4.69945`

- Cluster 2 : CatBoost

  `SMAPE : 2.27453`

- Cluster 3 : et

  `SMAPE : 2.92448`