**11회 대회 에너지 데이터 설명**

**[Files]**

1. train.csv

국내(인천지역) 특정 지역의 모 아파트 및 모 상가의 전력사용량. (1300호)

2016년 7월 26일 11시 ~ 2018년 6월 30일 24시까지 시간 당 전력사용량

1. test.csv

국내(인천지역) 특정 지역의 모 아파트 및 모 상가의 전력사용량. (200호)

2017년 7월 1일 00시 ~ 2018년 6월 30일 24시까지 시간 당 전력사용량

* train의 세대와 다른 세대

1. submission.csv

test셋에 제시된 세대(상가)에 대한 예측값들을 제출하는 포맷

**[Data fields]**

1. train.csv

2016년 7월 26일 11시 ~ 2018년 6월 30일 24시까지 시간 당 전력사용량

|  |  |
| --- | --- |
| Time | 특정 날짜, 특정 시간대  ex) 2016.7.26 11:00 => 2016년 7월 26일 11~12시 |
| X1, X2 ... X1300 | 각 세대의 미터 ID (**실제 열 이름은 다름**) |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Time | X1 | X2 | X3 | **···** | X1299 | X1300 |
| 2016.7.26 11:00 |  |  |  |  |  |  |
| 2016.7.26 12:00 |  |  | 1302 |  |  |  |
| **:** |  |  |  |  |  |  |
| 2018.6.30 23:00 |  |  |  |  |  |  |

예) 2016년 7월 26일 12:00 ~ 13:00의 X3 세대의 전력사용량은 1302이다.

1. test.csv

2017년 7월 1일 00시 ~ 2018년 6월 30일 24시까지 시간 당 전력사용량

|  |  |
| --- | --- |
| Time | 특정 날짜, 특정 시간대 |
| X1301, X1302 ... X1500 | 각 세대의 미터 ID (**실제 열 이름은 다름**) |

국내(인천지역) 특정 지역의 모 아파트 및 모 상가의 전력사용량.

train.csv와 포맷은 동일함. 그러나 예측해야 하는 세대의 아이디가 중복되지 않음

1. **submission\_1002**.csv

test셋에 제시된 세대(상가)에 대한 예측값들을 제출하는 포맷

미터ID(세대ID), 시간별전력사용량(24개), 10일간전력사용량(10개), 5개월간전력사용량(5개)

즉 각 열은 다음과 같은 형태임.

|  |  |
| --- | --- |
| meter\_id | 각 세대의 미터 ID |
| X2018\_7\_1\_1h | 2018년 7월 1일 00:00 ~ 01:00의 전력사용량 |
| X2018\_7\_1\_2h | 2018년 7월 1일 01:00 ~ 02:00의 전력사용량 |
| **:** | **:** |
| X2018\_7\_1\_23h | 2018년 7월 1일 22:00 ~ 23:00의 전력사용량 |
| X2018\_7\_1\_24h | 2018년 7월 1일 23:00 ~ 24:00의 전력사용량 |
| X2018\_7\_1\_d | 2018년 7월 1일 일일 전력사용량 |
| **:** | **:** |
| X2018\_7\_10\_d | 2018년 7월 10일 일일 전력사용량 |
| X2018\_7\_m | 2018년 7월 한달 전력사용량 |
| **:** | **:** |
| X2018\_11\_m | 2018년 11월 한달 전력사용량 |

+ 10월2일 부로 meter\_id가 오름차순으로 sort되었습니다. 리더보드 제출에는 문제가 없으니 이에 양해 바랍니다.