

ALL_DATA(범주형_시간추가_결측치0) (1).CSV

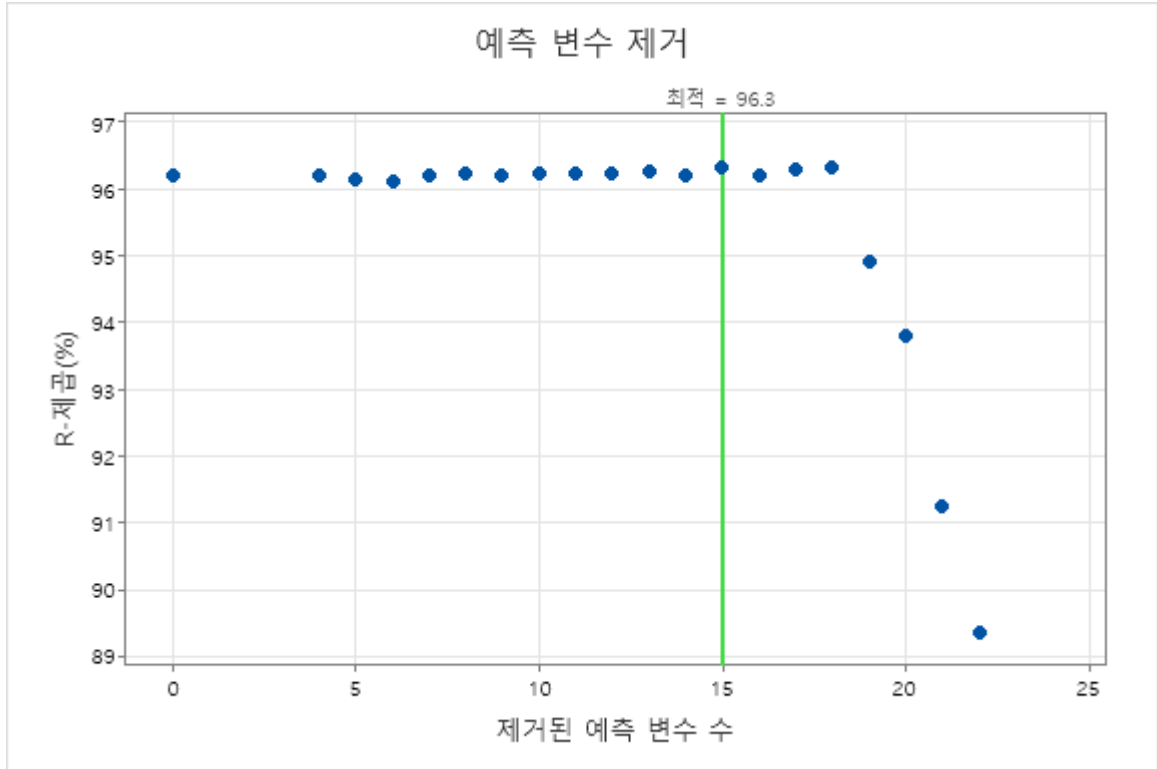
TreeNet® 회귀 - 주요 예측 변수 검색: 전력거래량(MWh) vs
기온(°C), 풍속(m/s), 풍향(16방위), 습도(%), 증기압(hPa),
이슬점온도(°C), 현지기압(hPa), 해면기압(hPa), 전운량(10분위),
중하층운량(10분위), 최저운고(100m), 시정(10m), 지면온도(°C),
5cm 지중온도(°C), 10cm 지중온도(°C), 20cm 지중온도(°C), 30cm
지중온도(°C), 지역, 년도, 월, 일, 요일, 시간

방법

손실 함수	제곱 오차
최적 트리 수 선택 기준	최대 R-제곱
모형 검증	70/30% 교육/검정 집합
학습률	0.1
하위 표본 부분	0.5
트리당 최대 터미널 노드 수	6
최소 단말 노드 크기	3
노드 분할을 위해 선택된 예측 변수 수	총 예측 변수 수 = 23
사용된 행	978816

반응 정보

데이터 집합	N	N의 비율(%)	평균	표준 편차	최소값	Q1	중위수	Q3	최대값
교육	685180	70.00	3784.30	3952.80	0	330	2771	6203	18838
검정	293636	30.00	3782.21	3951.06	0	330	2768	6205	18880



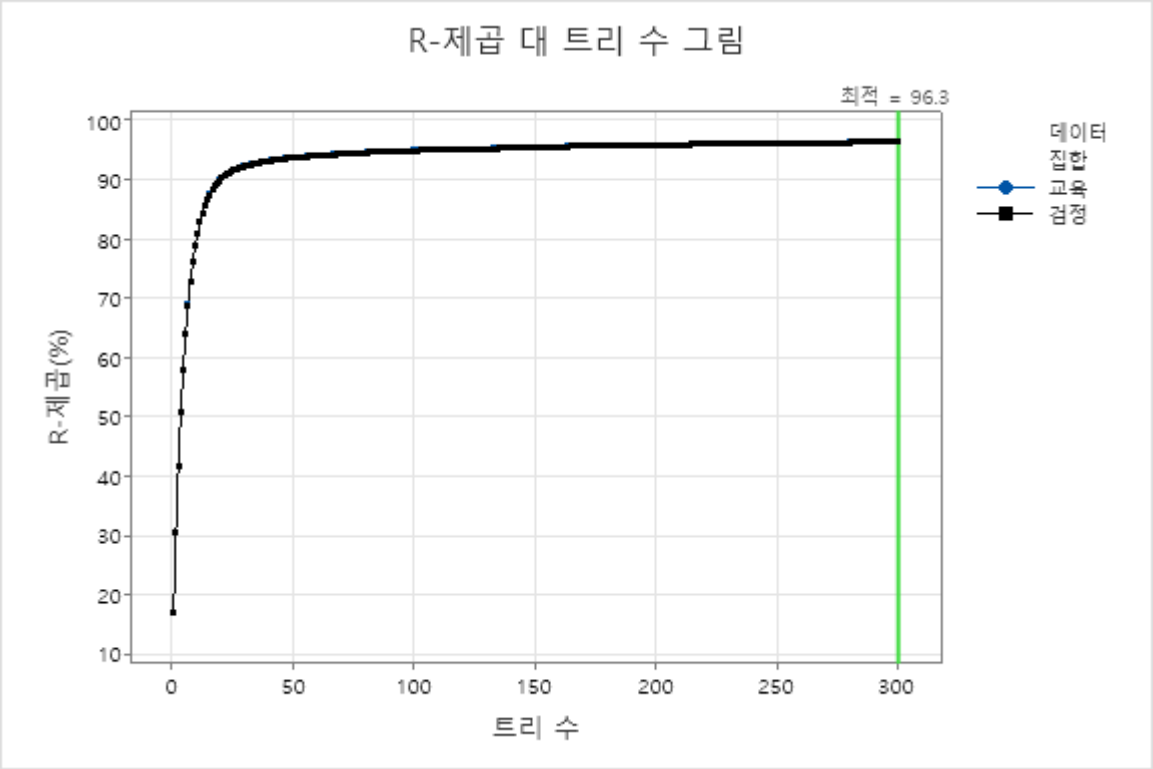
중요하지 않은 예측 변수를 제거하여 모형 선택

검정

최적의 트리		예측 변수	
모형	수	R-제곱(%)	제거된 예측 변수
1	300	96.19	23 없음
2	300	96.19	19 해면기압(hPa), 풍향(16방위), 풍속(m/s), 중하층운량(10분위)
3	300	96.14	18 최저운고(100m)
4	300	96.11	17 시정(10m)
5	300	96.18	16 습도(%)
6	300	96.23	15 전운량(10분위)
7	300	96.20	14 이슬점온도(°C)
8	300	96.22	13 현지기압(hPa)
9	300	96.24	12 20cm 지중온도(°C)
10	300	96.22	11 10cm 지중온도(°C)
11	300	96.27	10 지면온도(°C)
12	300	96.20	9 5cm 지중온도(°C)
13*	300	96.30	8 일
14	300	96.20	7 증기압(hPa)
15	300	96.27	6 30cm 지중온도(°C)
16	300	96.30	5 기온(°C)
17	300	94.92	4 요일
18	300	93.79	3 시간
19	298	91.24	2 년도
20	68	89.37	1 월

알고리즘은 각 단계에서 하나의 예측 변수와 중요도가 0인 임의의 예측 변수를 제거했습니다.

* 선택한 모형에는 최대 R-제곱이 있습니다. 선택한 모형의 출력에 뒤에 나옵니다.

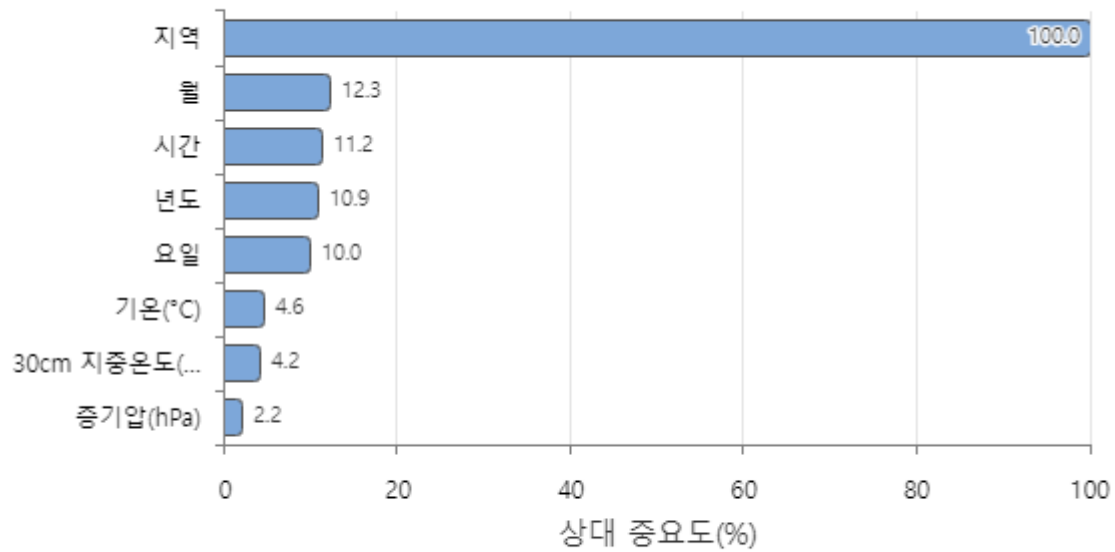


모형 요약

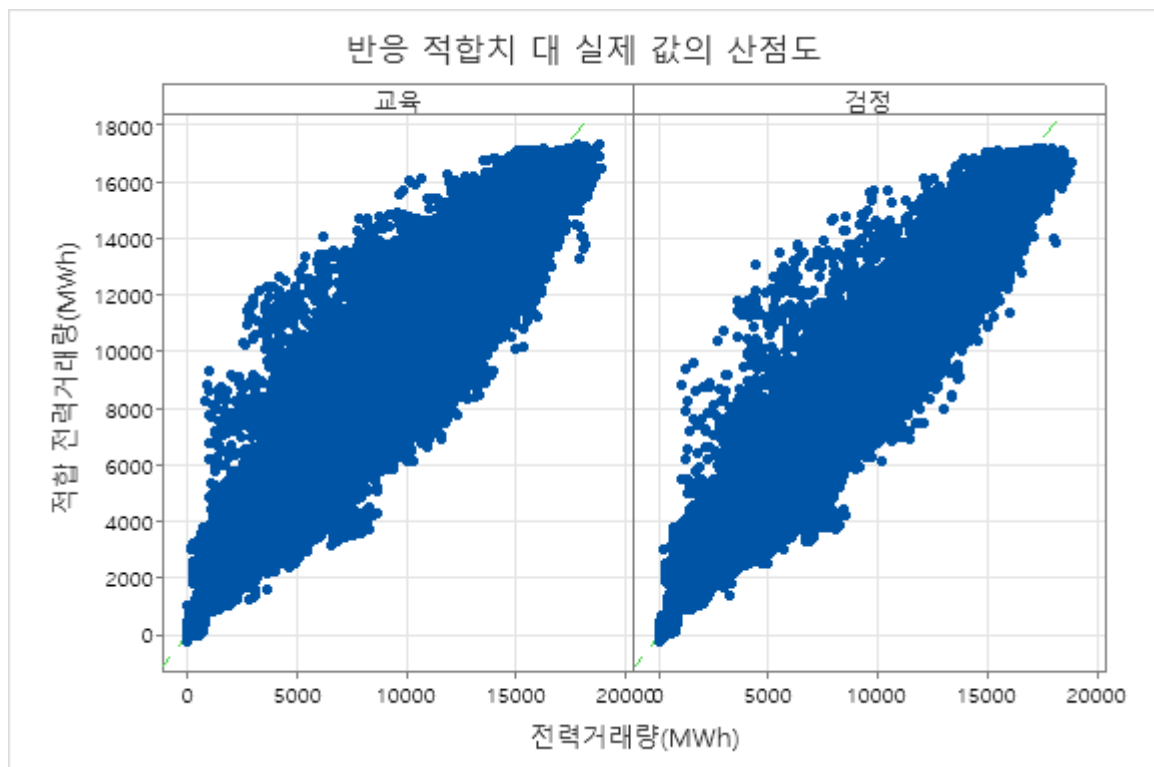
전체 예측 변수 8
 중요 예측 변수 8
 성장한 트리 수 300
 최적의 트리 수 300

통계량	교육	검정
R-제공	96.31%	96.30%
루트 평균 제곱 오차(RMSE)	759.3885	759.6609
평균 제곱 오차(MSE)	576670.8975	577084.6157
평균 절대 편차(MAD)	480.1979	480.3744
평균 절대 백분율 오차(MAPE)	1.6879	1.6723

상대 변수 중요도

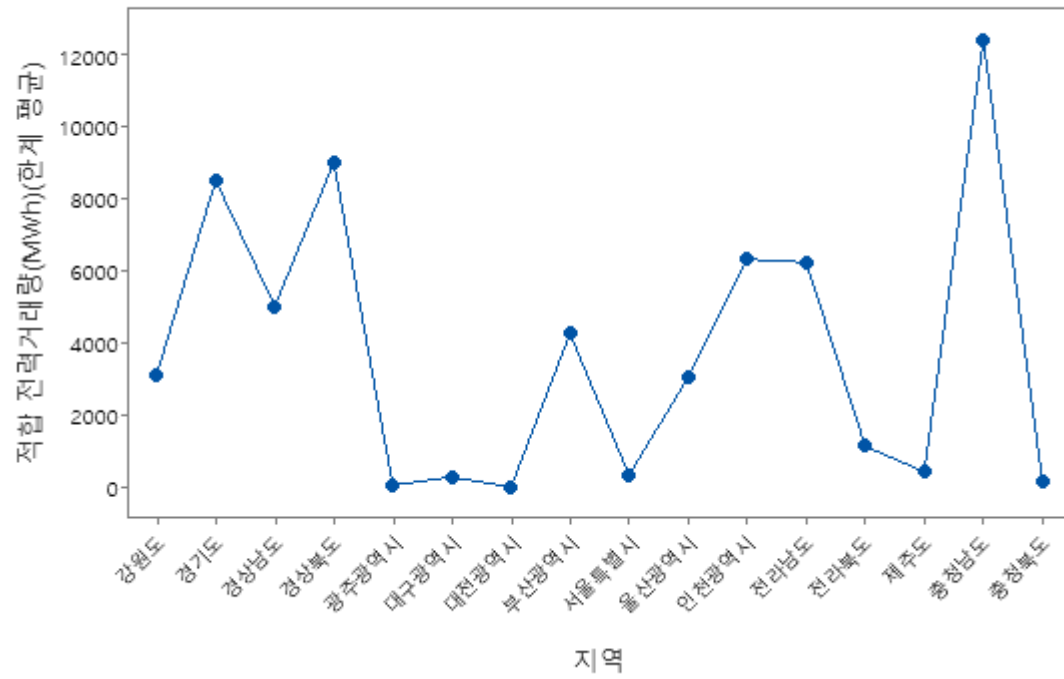


변수 중요도는 예측 변수에 분할을 적용할 때 모형 개선을 측정합니다. 상대 중요도는 상위 예측 변수를 기준으로 한 % 개선으로 정의됩니다.

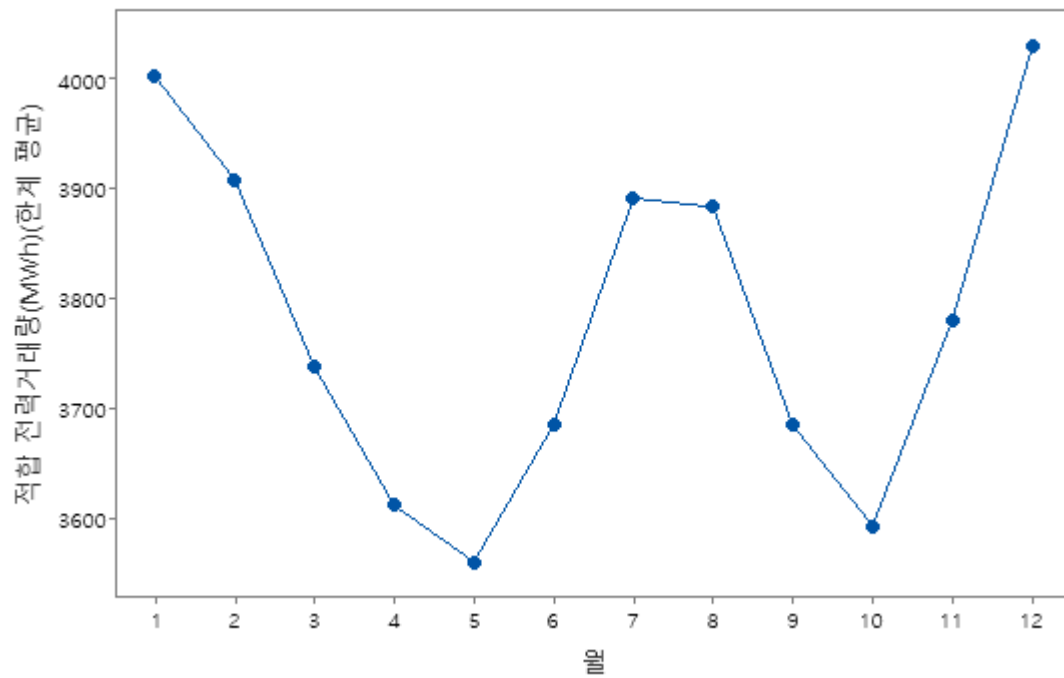


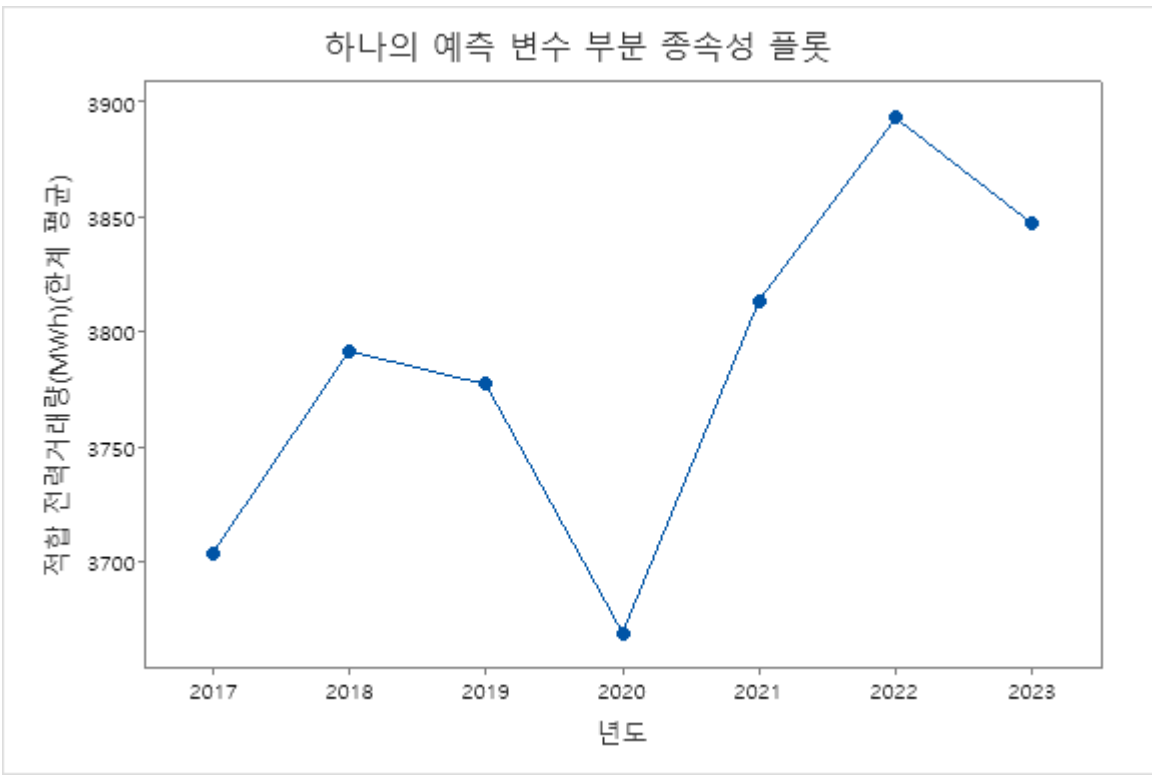
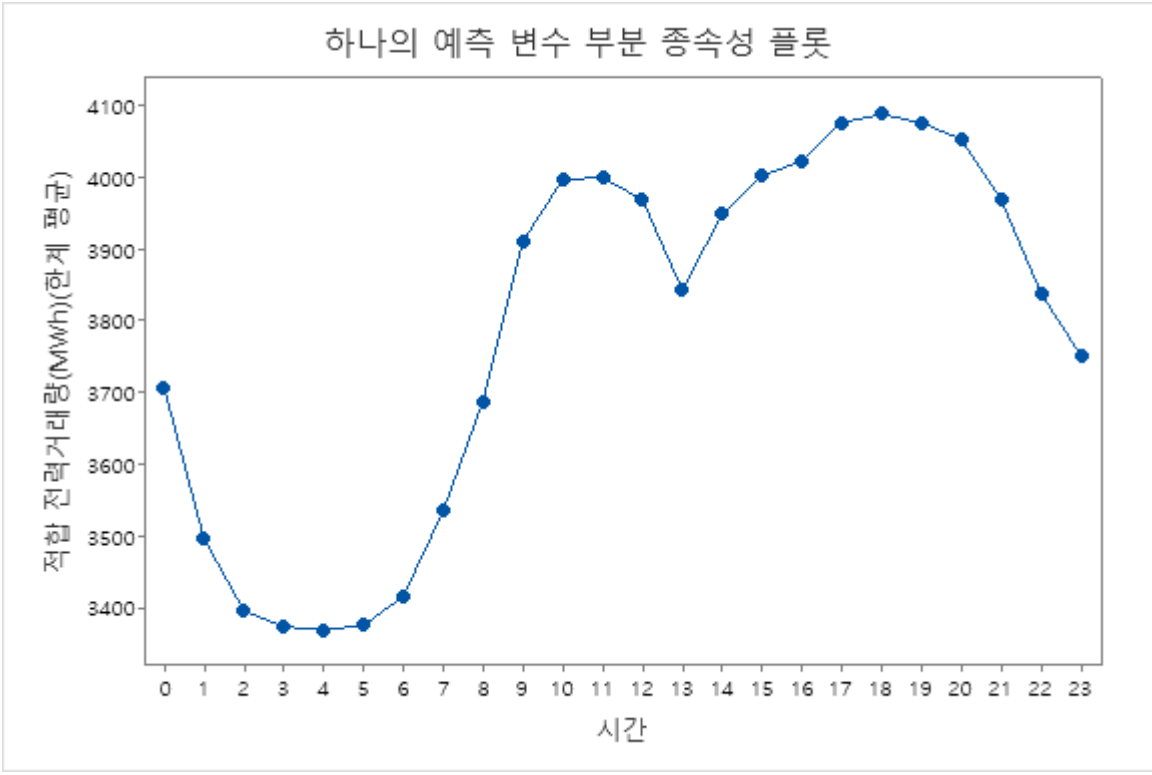
하나의 예측 변수 부분 종속성 플롯

하나의 예측 변수 부분 종속성 플롯



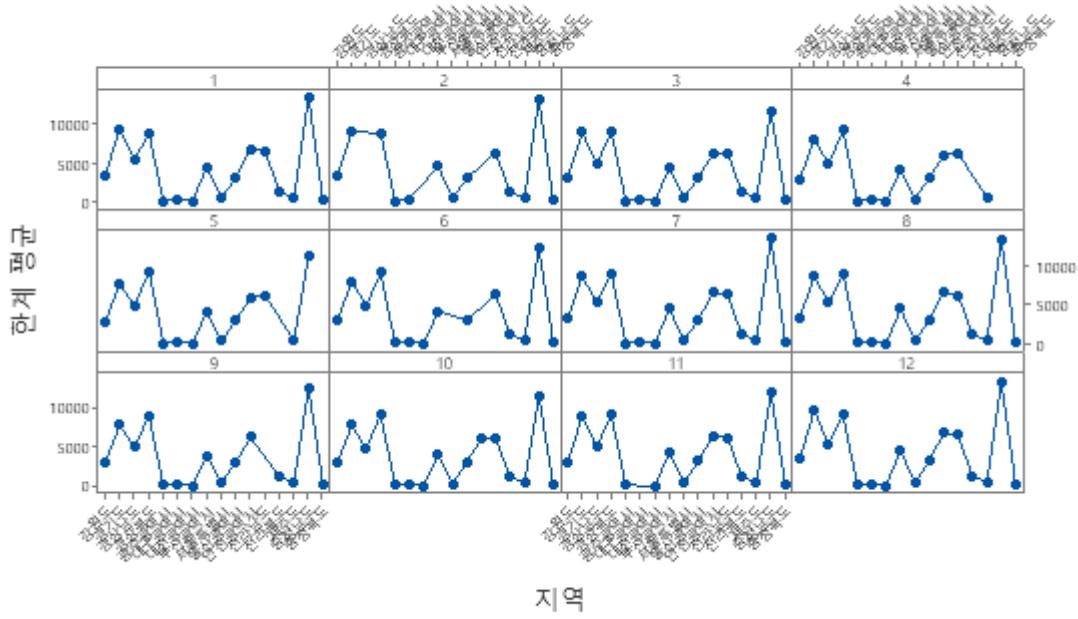
하나의 예측 변수 부분 종속성 플롯





두 개의 예측 변수 부분 종속성 플롯

장착된 분산형 플롯 전력거래량(MWh)



패널 변수: 월