

변수 선택

Type2 ANOVA사용 변수 선택

```
In [ ]: mod_full.compare_f_test(mod_non_humi)
        #mod_full 은 모든 변수가 포함된 모델의 잔차제곱합
        #mod_non_humi 는 습도만 제거한 모델의 잔차제곱합
```

출력값 중 p값이 0.578 으로 나옴
→ 습도는 매상에 유의미한 영향을 끼치지 않음.

AIC이용 변수 선택

```
In [ ]: print("모든 변수를 포함한 모델: ", mod_full.aic.round(3))
        print("습도를 제외한 모델      : ", mod_non_humi.aic.round(3))
```

습도를 제외한 모델의 AIC가 더 작아짐
→ 습도는 매상 예측 모델에서 제외해야함.

모든 독립변수가 포함된 모델을 구축해서 잔차 제곱합을 구한 후
습도만 제거한 모델의 잔차제곱합을 구해서 잔차제곱합의 차를 구해서 P값을 구하면 0.578이 나온다. 이를 통해 습도는 매상에 대해
유의미한 영향을 끼친다고 볼 수 없다는 결과를 얻을 수 있다.

습도를 뺀 모델에서 다시 type2분산분석을 하여 다른 유의미하지 않은 변수가 있는지도 확인해주어야 한다.

계수 등의 결과에 관한 해석은 변수 선택 후의 모델을 이용해서 수행한다.
AIC를 이용해서도 변수를 선택할 수 있다.

전체 모든 변수를 포함한 모델인 mod_full과 습도를 제외한 모델인 mod_non_humi의 AIC를 비교하여 습도를 제거한 모델의 AIC
가 더 작아진 것을 확인하고 습도를 매상 예측 모델에서 제외한다.

AIC를 이용한 모델 선택은 기계적으로 수행할 수 있고, 검정의 다중성과 검정의 비대칭성 문제도 없어서 사용하기 편리하다.