

05

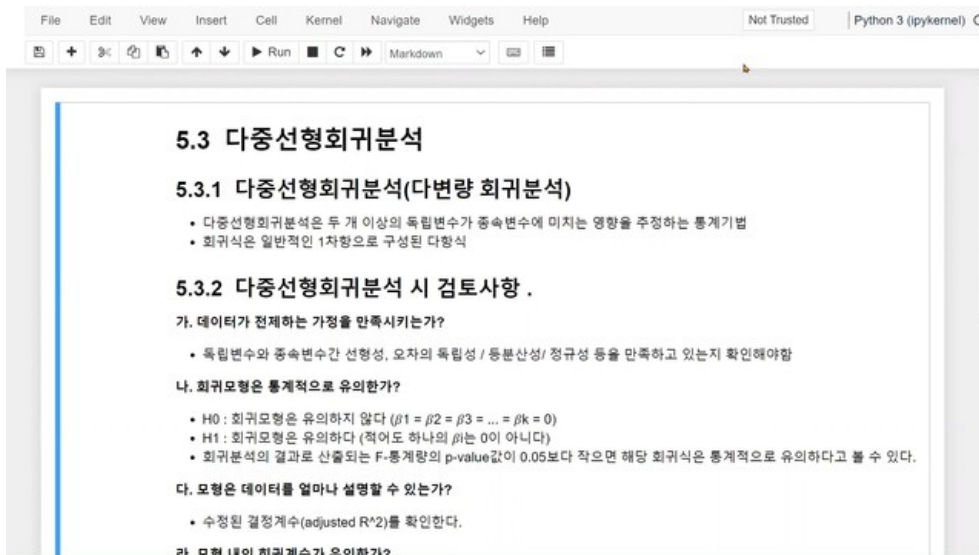
## 다중 선형 회귀분석

[다중선형회귀분석]

# 변수가 추가된 것 외에는 단순선형회귀분석과 같습니다.

- 변수가 추가된다면 무엇이 달라질까요?
- 독립변수가 많아지면 다중공선성이 생기므로, 변수선택법을 통한 최적의 회귀모델을 선택해야 하는 과정이 추가됩니다.

00:00



[주의해야 할 점]

03:16

# 회귀분석에서 독립변수에 범주형 데이터가 있다면 더미변수화를 시켜주어야 합니다.

File Edit View Insert Cell Kernel Navigate Widgets Help Not Trusted Python 3 (ipykernel) O

• 회귀분석의 결과로 산출되는 F-통계량의 p-value값이 0.05보다 작으면 해당 회귀식은 통계적으로 유의하다고 볼 수 있다.

**다. 모형은 데이터를 얼마나 설명할 수 있는가?**

- 수정된 결정계수(adjusted R<sup>2</sup>)를 확인한다.

**라. 모형 내의 회귀계수가 유의한가?**

- 단변량 회귀분석에서 회귀계수의 유의성 검토와 마찬가지로 회귀계수에 대한 t통계량의 p-value값이 0.05보다 작으면 해당 회귀 계수가 통계적으로 유의하다고 볼 수 있다. 단, 다중회귀분석을 할 때는 모든 회귀계수가 유의한지를 검정한 후 해당 회귀식을 해석해야 함

**마. 모형이 데이터를 잘 적합하고 있는가?**

- 모형의 잔차와 증속변수에 대한 산점도를 그리고, 회귀진단을 수행하여 판단한다.

---

**5.3.2.1 더미변수화**

**가. 범주형 변수 변환**

- 회귀분석은 연속형 변수를 다루는 기법이므로 범주형 데이터의 경우 형태를 변환해주어야 회귀분석을 수행할 수 있다.
- 더미변수란 0 or 1값만 가지며 어떤 특징에 해당 하는지의 여부를 표현하는 변수이다.

In [ ]: 1