1. 데이터 셋 확인

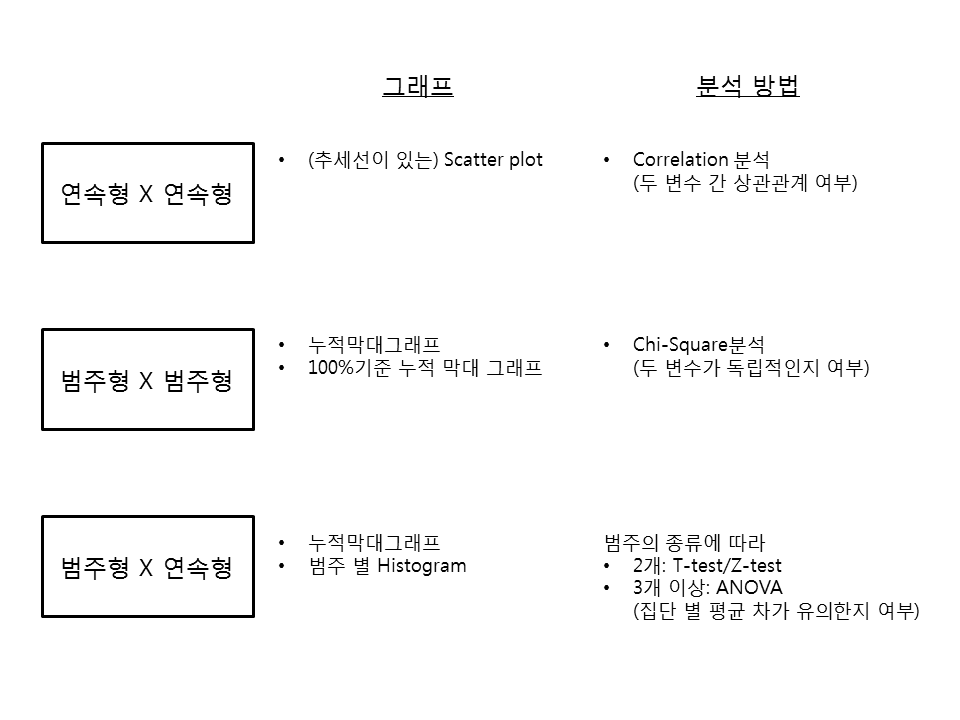
A. 변수 확인

B. RAW 데이터 확인

B-1 단변수 분석

Histogram, Boxplot

B-2 이변수 분석



B-3 셋 이상의 변수

범주형 변수가 하나이상 포함되어 있는 경우 변수를 범주에 따라 쪼갠 후에 위 분석 방법에 따라 분석

예를 들어 남성-여성의 정보가 있고 연소득, 학력, 키의 정보가 있다고 할때 성별로 구분해서 연소득과 학력이 독립적인지 T-test로 확인해 볼 수 있음. 학력으로 구분해서 연소득과 키의 상관관계를 확인해도 됨

기온별(cold, cool, mild, hot)으로

Useage 와 average temperature 독립적인지 아닌지 알 수 있음 …?

Useage 와 subway 독립적인지 아닌지 알 수 있음 …?

기간별(Q1, Q2, Q3, Q4)로

Useage 와 average temperature 독립적인지 아닌지 알 수 있음.

T-test

등분산성 var.test

독립표본 t-test: 두개의 그룹 간 평균의 차이가 유의미 한지 여부를 판단

대응표본 t-test: 동일한 집단의 전-후 차이를 비교하기 위해 사용

단일표본 t-test : 하나의 집단의 평균이 특정 기준보다 유의미하게 다른지 혹은 큰지/작은지를 알아보는 분석 방법