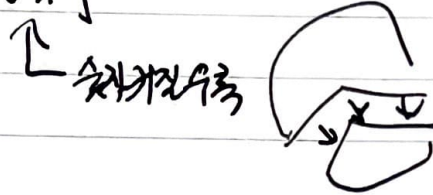


1인

- Value_counts 는 Null값 반환 x
- explode(0.0.1]



→ explode(0.0.1]

sns.set(font_scale=2)

autoplot = '%.1.2f%.1.2f' 소수점자리.

ax = ax[0]

↳ subplots(1,2, ~) 는 1열 2행 즉, 그래프 2개를 뜻함.

ax[0]은 첫번째 그래프를 뜻함.

Countplot(x=)

! x값이 여러개 있으면 나눠서 그리기.

Categorical features 는 순서가 없다. *

Ordinal Features 는 순서가 있다.

가위 종류 Tall, medium, short 분류된다.

Ordinal Features 에는 data bloting이 가능하다.

A → 1

B → 2

C → 3

이렇게 computer bias 안함

but, Categorical features 에 order를 붙여 bias가 생긴다.

feature를 따라올라 순서가 있느냐가 중요하다.

순서 ~~없음~~ Label encoding 후

있으면

(one hot encoding
frequency encoding
mean encoding)

→ 3중 하나

있으면 4중 label encoding.

Label encoding을 해야 순서라는 정보가 살아난다.
안하면 정보가 사라진다.

* 데이터타입에 따라 분석하는 방법이 달라진다.

예. category - barplot

Continuous - histogram, scatter.

성공이냐? 실패이냐? ← target. 카테고리임.

카테고리는 scatter그림 안돼.

· 시각화를 하는 이유는 data points를 쉽게 보기 위함이다.
데이터를 가지고 그림을 그려서 거기에서 트렌드와 해석. 로직을 만들면서
그 로직을 통해서 . .

우리가 할 일은

· Countplot은 실제 '수'

Shape 함수 ⇒ DataFrame 또는 배열의 구조 및 차원에
대한 정보를 얻는 데 유용하다. 인덱싱으로 가능 가능.