(3) Ordinal Encoding :-
" Categories are replaced with numbers based on their rank/order.
Eg:- Size = { Small, Medium, Large }
 — Small → 1, Medium → 2, Large → 3

→ Pro :- Maintains order of categories.
→ Cons:- Should only be used when data has a natural order.

* One-Hot Encoding :- (Day - 27)

" Each Category is converted into a new column with binary values (0 or 1)".

⇒ Dummy Variable Trap :-

→ Here we drop one column (any one) from one of the splited columns.
→ In one-hot encoding we only keep ~~to~~ (n-1) columns.
bcz

Eg:-

We remove this column

	Yellow	Blue	Red	
Yellow	1	0	0	→ Means Yellow
Blue	0	1	0	→ Means Blue
Blue	0	1	0	
Red	0	0	1	→ Means Red

→ ~~If be~~ for blue & Red it is clearly visible bcz there is 1 for them but we can identify that if both are zero than it would be yellow.

→ So by this we can reduce the dimensions and also we can know the value.

→ This new created columns are known as dummy variables.

→ If we don't remove column than there would be ~~the multicolinearity problem.~~ multicolinearity problem.

⇒ OHE using most frequent variables :-

→ We have to do this bcz agar ada pase ek column named car_brand than ema to bau badhi brands hase. so enu apde OHE karye to min. 40 jevi columns add thai. So e dimensions bau vadhari dese so processing bau vadhi jase. So enu solution che keh apde jeh frequently used brands che ena alag alag column banaie and jeh brand amuk amuk vaar j aai che ene apde ek column (others) ma nakhi daisu so jeh apdi 40 columns banvani hoi e 10 j banse have to apda dimensions pan ocha thaya and apdu kaam pan thai gayu.