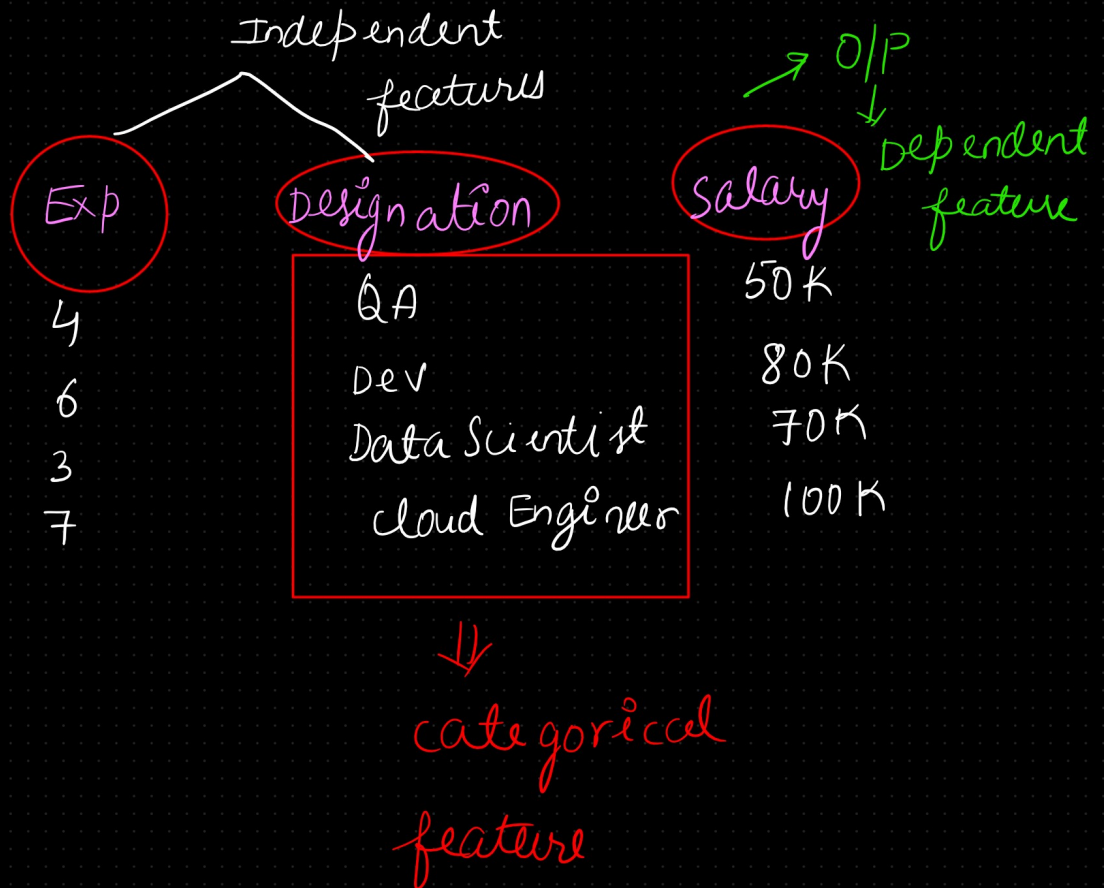


Feature Engineering

Day - 03

DATA ENCODING

Data Encoding is a process of converting data from one form to another.



Data Encoding → categorical features
↓
Numerical



Types of Data Encoding

- ① Nominal / One - Hot Encoding
- ② Label Encoding
- ③ Ordinal Encoding
- ④ Target Guided Ordinal Encoding

categorical \longrightarrow Numerical

Disadvantages

① Sparse Matrix

[Overfitting]

② 1000 features [1000 categories]

Nominal / OHE

categorical \rightarrow numeric

$\swarrow \searrow$
 $0, 1 \rightarrow$ binary
vector

Eg \Rightarrow House price

No. of Rooms House size Location Price

—

—

Delhi

Noida

Pune

Delhi

Noida

Nominal
Encoding \Rightarrow

	Delhi	Noida	Pune
	1	0	0
	0	1	0
	0	1	0
	0	0	1

