

## Handling Mixed Variables:

Sometimes there are some features which contains code. They are not fully numerical or not fully categorical.

Suppose a feature "Cabin" →

Cabin:

B5

C23

D32

E15

B → class name → categorical data

5 → Number of boggies → Numerical

The solution is to make two new features from it and delete the original.

<u>Cabin</u>	→	<u>Cabin class</u>	→	<u>Cabin Number</u>
B5		B		5
C23		C		23
D32		D		32
E15		E		15

But if the values are like this below?

7  
C  
1  
A  
C  
4  
D

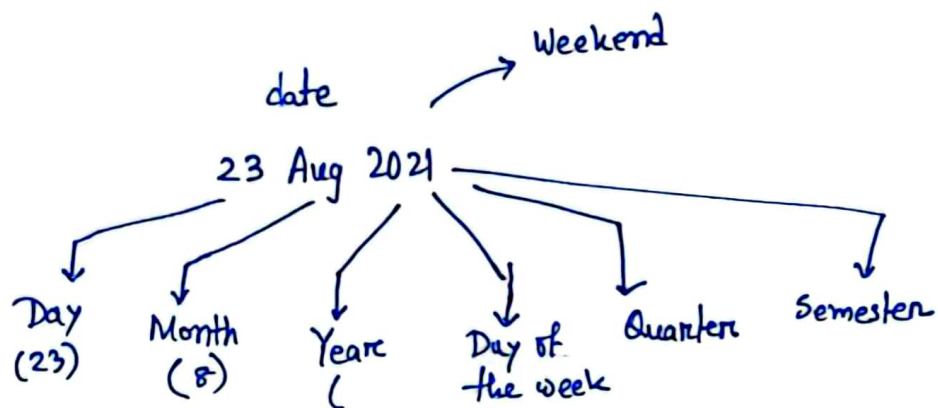
[mix of categorical and numerical]

Then we have to separate like this

<u>x</u>		<u>x-cat</u>	<u>x-num</u>
7	→	NaN	7
C	→	<del>NaN</del> C	NaN
1	→	NaN	1
A	→	A	NaN
C	→	C	NaN
4	→	NaN	4
D	→	D	NaN

### Handling Date and Time variables:

From a date column, many information can be derived to make other new features.



Detailed code example has been uploaded to github.