



DATA SCIENCE WITH PYTHON : USAGE OF GET_DUMMIES TO CREATE DUMMY VARIABLES WITH PANDAS #3176

PRESENTED BY: DEEPTHI M

BATCH NUMBER: 05

SERIAL NUMBER: 172

Why to create Dummy Variables?

Conversion of categorical data to numeric data.

Many of the model in data mining treats or computes the dataset when all the datatype is in numeric.

Methods that are used to convert:

Label encoder: Label Encoding refers to converting the labels into a numeric form so as to convert them into the machine-readable form.

One- hot encoding: A one hot encoding is a representation of categorical variables as binary vectors. This first requires that the categorical values be mapped to integer values.

Water Temperature	
A	Hot
B	Cold
C	Warm
D	Cold



Dummy Variables

Water Temperature		var_hot	var_warm	var_cold
A	Hot	1	0	0
B	Cold	0	0	1
C	Warm	0	1	0
D	Cold	1	0	0