

Categorical Variables

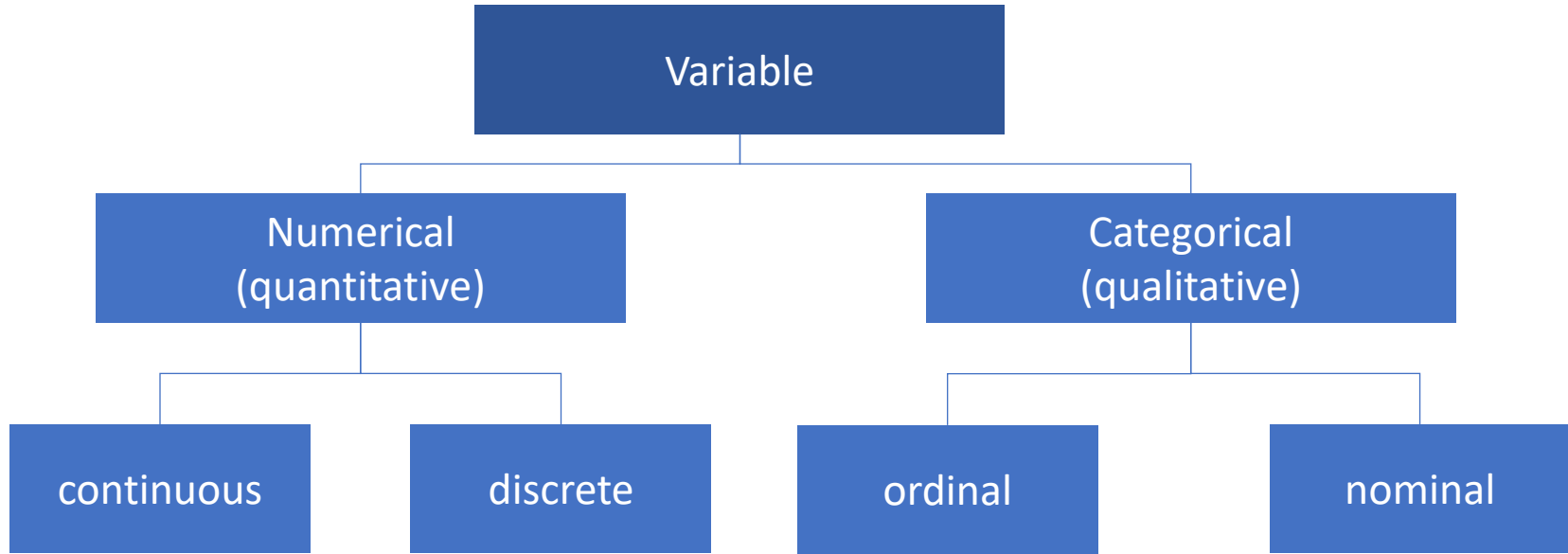
Categorical Variables

Introduction

- S

Categorical Variables

Variable Types



Property: any value in range

integer values

ordered,
mutually exclusive

unordered,
mutually exclusive

Example: height in cm

count

ranking in competition

driving / walking

Categorical Variables

Ordinal Feature Conversion

- Ordinal features are converted to integers

Pain
no pain
mild
moderate
intense
severe



Pain
0
1
2
3
4

Categorical Variables

Nominal Feature Conversion

- nominal features are unordered, mutually exclusive
- new columns created for binary (dummy) variables
- one dummy variable created for each level of column
- technique is called *One Hot Encoding*
- use pandas `get_dummies()`

Color		Color_green	color_yellow	color_red
green	→	1	0	0
yellow		0	1	0
red		0	0	1