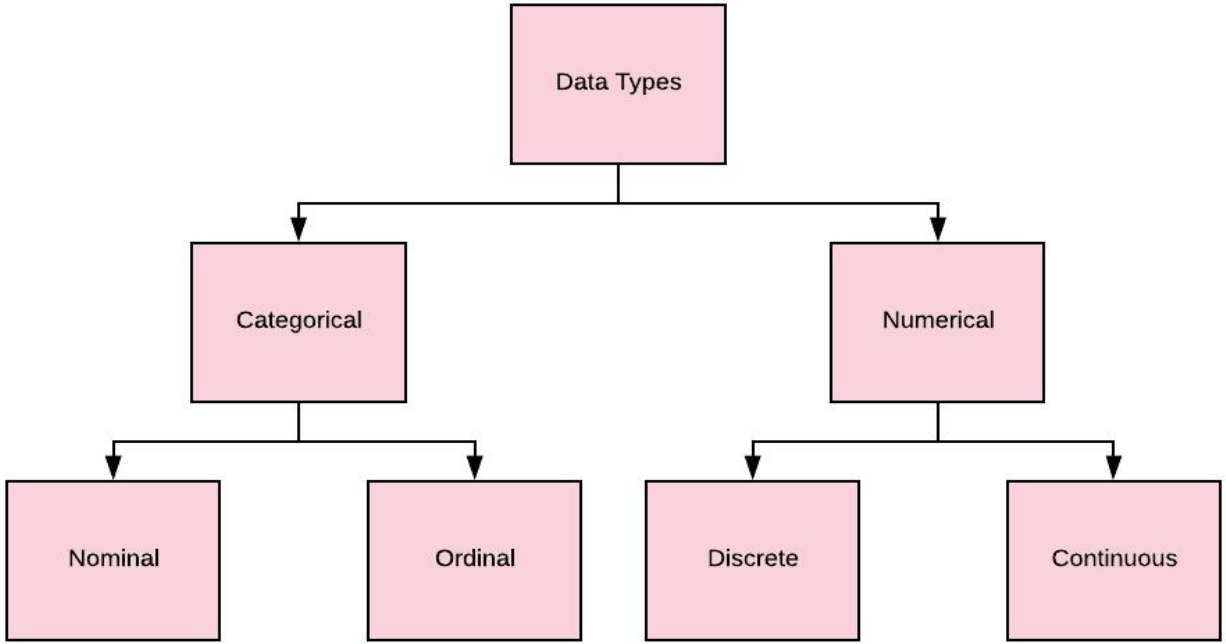
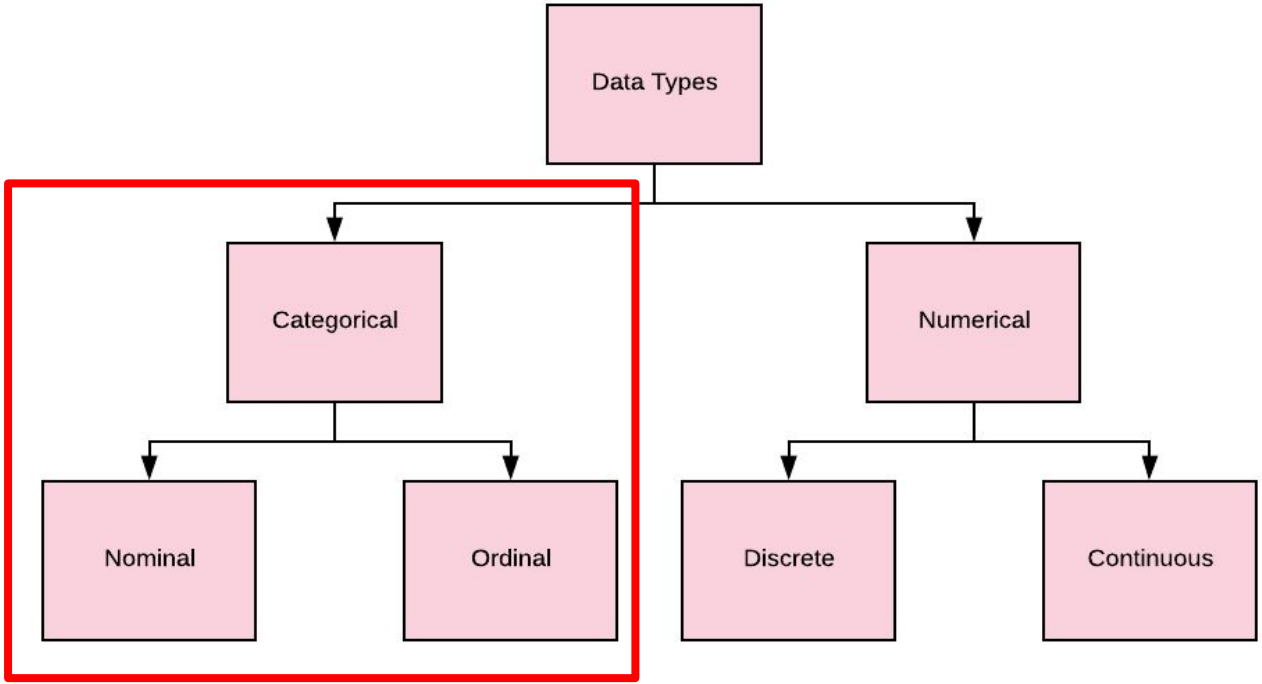
Variables Types

Data Types



Analytics Vidhya
Learn everything about analytics

Data Types

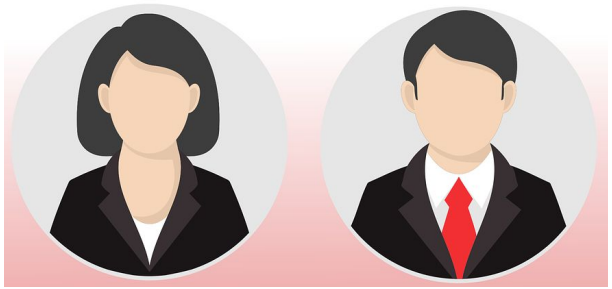


Analytics Vidhya
Learn everything about analytics

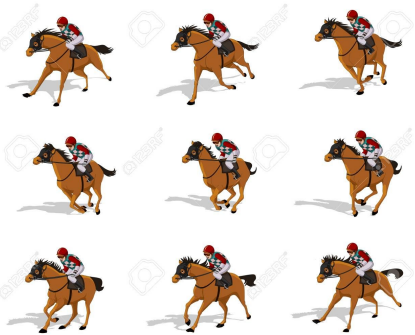
Categorical Data

- Also known as Qualitative Data
- Represent Characteristics or Qualities

Ex. Your Gender ?

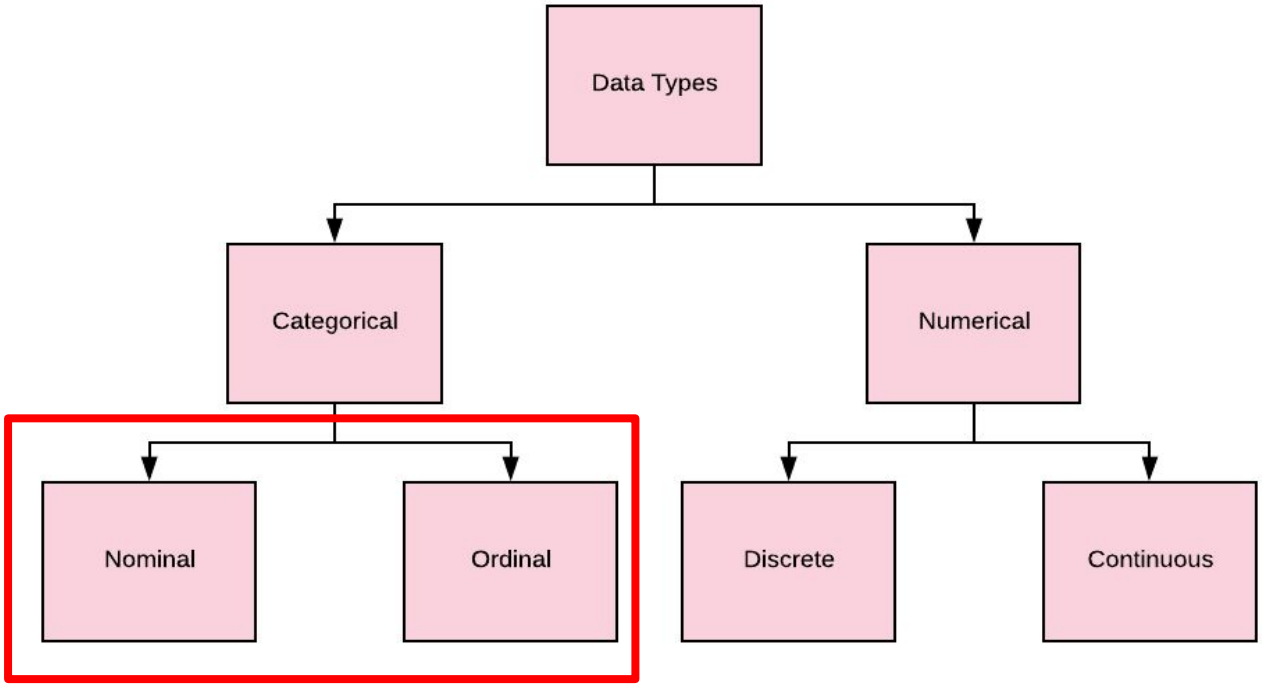


Horse racing?



Analytics Vidhya
Learn everything about analytics

Data Types

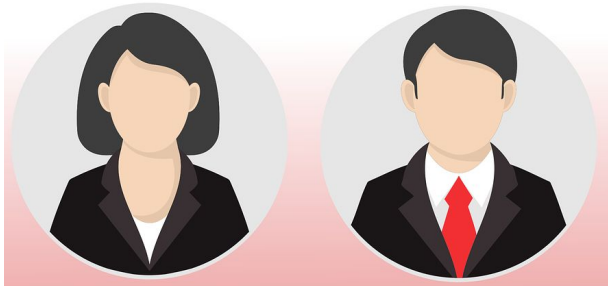


Analytics Vidhya
Learn everything about analytics

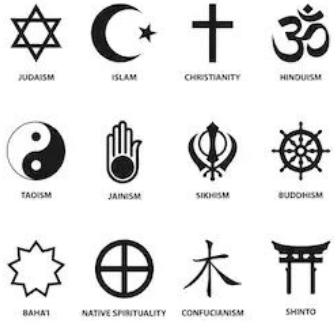
Nominal Data

- Discrete
- No Specific Order
- Changing Label doesn't impact Meaning

Ex. Your Gender ?



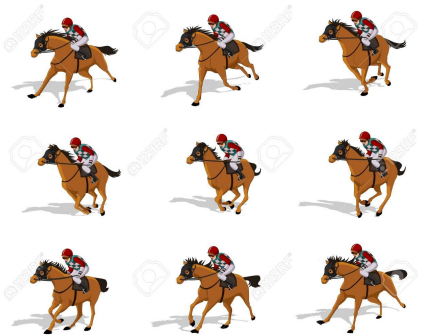
Your religious Preference?



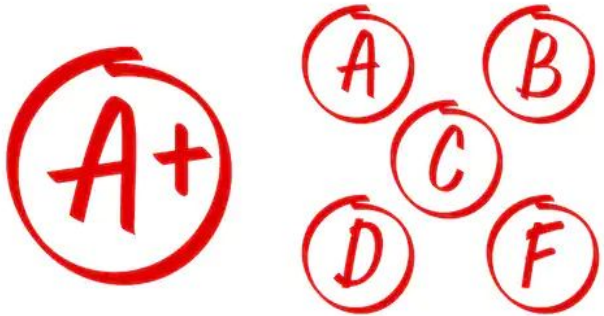
Ordinal Data

- Discrete
- Specific Order
- Labels represent Meaning

Horse racing?



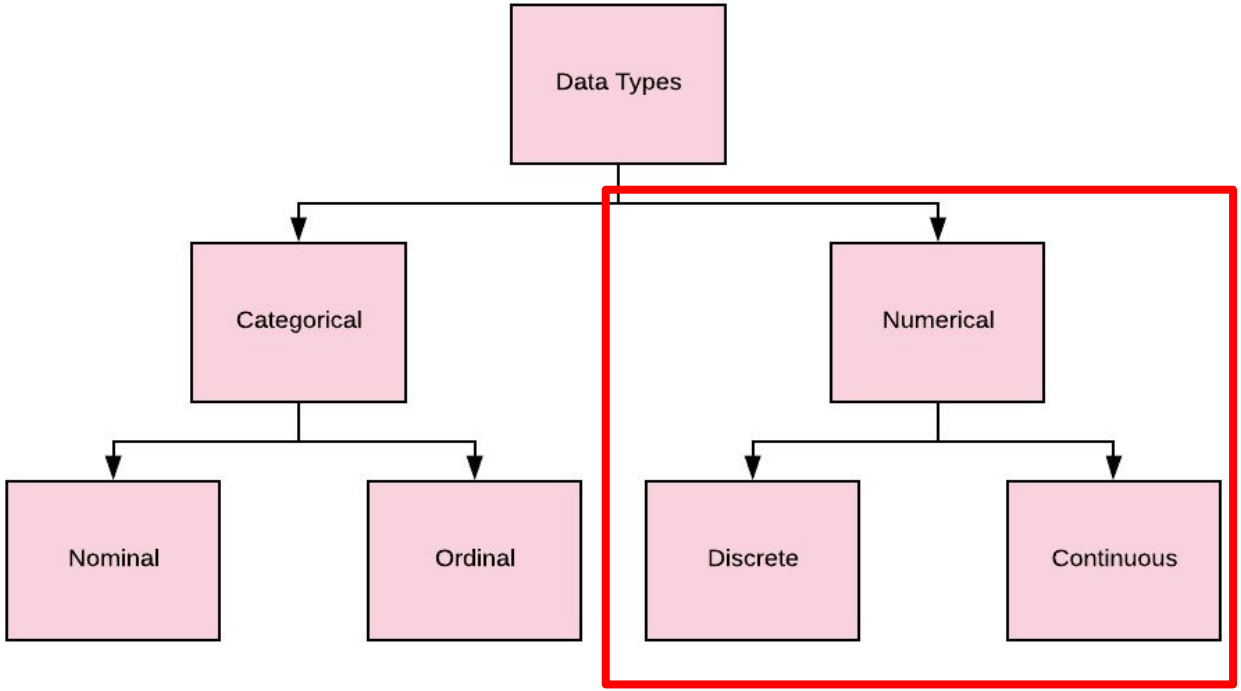
Your Grades?



Nominal vs Ordinal

- Name vs Order
- Quantitative Meaning
- Information type

Data Types



Analytics Vidhya
Learn everything about analytics

Numerical Data

- Quantitative
- Numerical significance

Ex. Your sleeping hours ?

Ex. When did you sleep last night?

Do both these questions represent similar values ?

Discrete Data

- Takes only certain values
- Can be counted

  Ex. Your sleeping hours ?

Continuous Data

- Can take any value
- Can not be counted but measured

  Ex. When did you sleep last night?

Thank You!

How Data is Stored in Pandas

Pandas dtype	Python type	NumPy type	Usage
object	str or mixed	string_, unicode_, mixed types	Text or mixed numeric and non-numeric values
int64	int	int_, int8, int16, int32, int64, uint8, uint16, uint32, uint64	Integer numbers
float64	float	float_, float16, float32, float64	Floating point numbers
bool	bool	bool_	True/False values
datetime64	NA	datetime64[ns]	Date and time values
timedelta[ns]	NA	NA	Differences between two datetimes
category	NA	NA	Finite list of text values

Data Type: Integers and Float

- Integer and Float are kinds of numerical data
  - Integer is number without decimal point (example - 5, 902, 12)
  - Float is number with decimal point (example - 0.5, 22.7)

- Numerical operations

Example: Age of Customer (int)

Current Balance (float)

Analytics Vidhya
Learn everything about analytics

Data Type: Boolean

- Boolean variable has only two possible values:- True/False
- Comparison

  Ex. ((7>3) & (5>6))
  False

  Ex. ((7>3) | (5>6))
  True

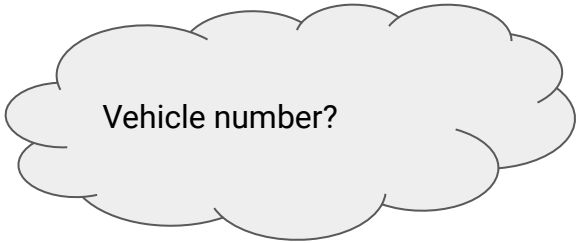
- Conditional Statements

Identify all such 40 years old customers who are prone to churn?

Analytics Vidhya
Learn everything about analytics

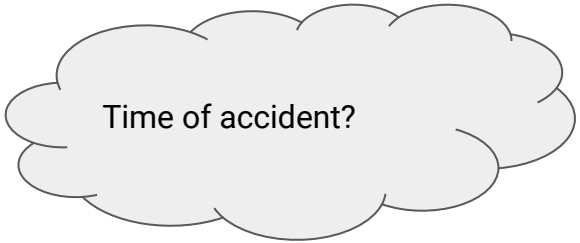
Data Type: Object

- Object can be single or mixed type of variables
  - Text (Movie Script)
  - Mixed numeric (Passport ID)
  - Non numeric values (Currency Symbol)



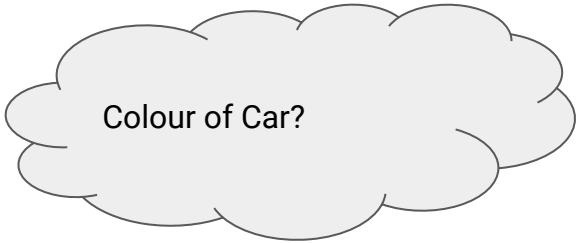
Data Type: Date Time

- Date time variable for date and time values
- Timedelta variable for difference between two datetimes

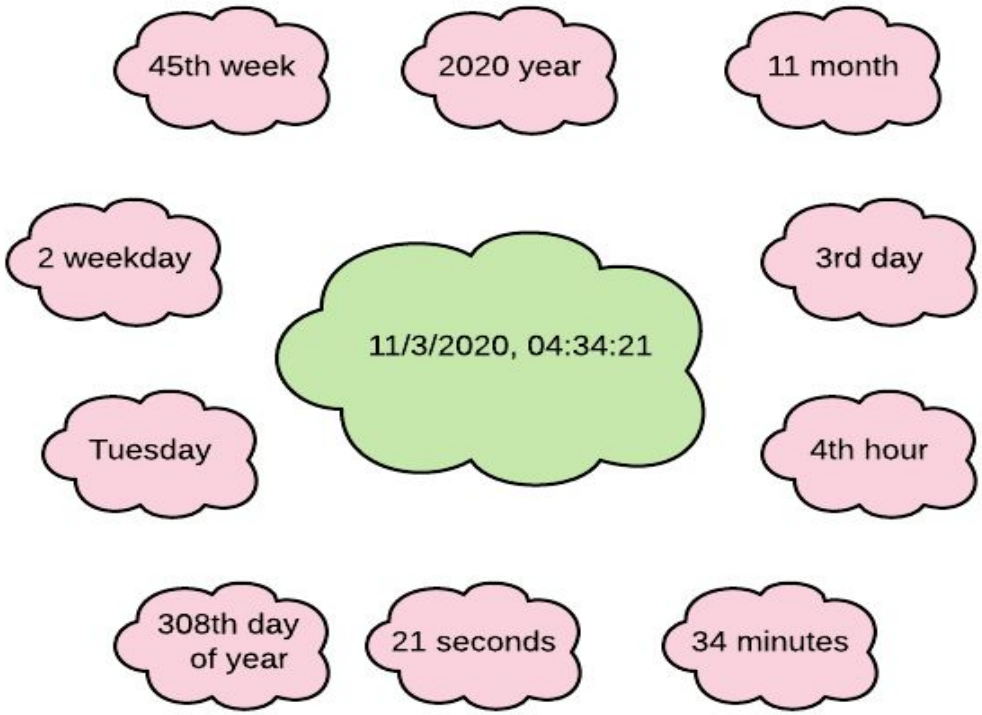


Data Type: Category

- Categorical variable is used to represent the categorical data.
- Can Specifying an order.



Date Time



Analytics Vidhya
Learn everything about analytics