



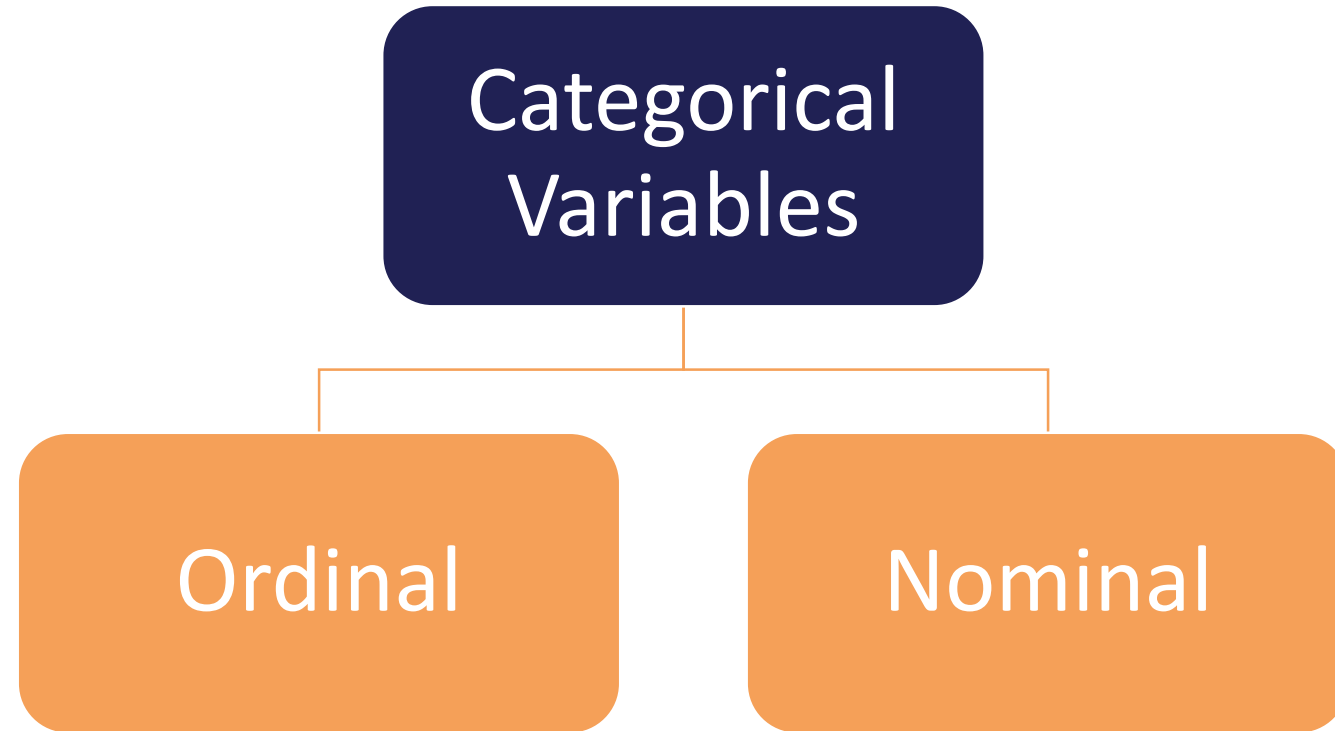
Categorical Variables

Categorical Variables

The values of a categorical variable are selected from a group of **categories**, also called **labels**. Examples:

- Marital status (married, single, ...)
- Intended use of loan (debt-consolidation, car purchase, ...)
- Mobile network provider (Vodafone, Orange, ...)
- Gender (male, female)

Variable Types



Ordinal Variables

Categorical variables in which categories can be meaningfully ordered are called ordinal. Examples:

- Student's grade in an exam (A, B, C or Fail)
- Days of the week (Monday = 1 and Sunday = 7)
- Educational level, with the categories: Elementary school, High school, College graduate and PhD ranked from 1 to 4

Nominal Variables

Show no intrinsic order of the labels. Examples:

- Country of birth (Argentina, England, Germany)
- Postcode
- Vehicle make (Citroen, Peugeot, ...)



Special Cases

Special cases

- Categorical variables where categories are encoded as numbers (e.g. gender may be coded as 0 for males and 1 for females)
- Id variables: number that uniquely identifies an observation

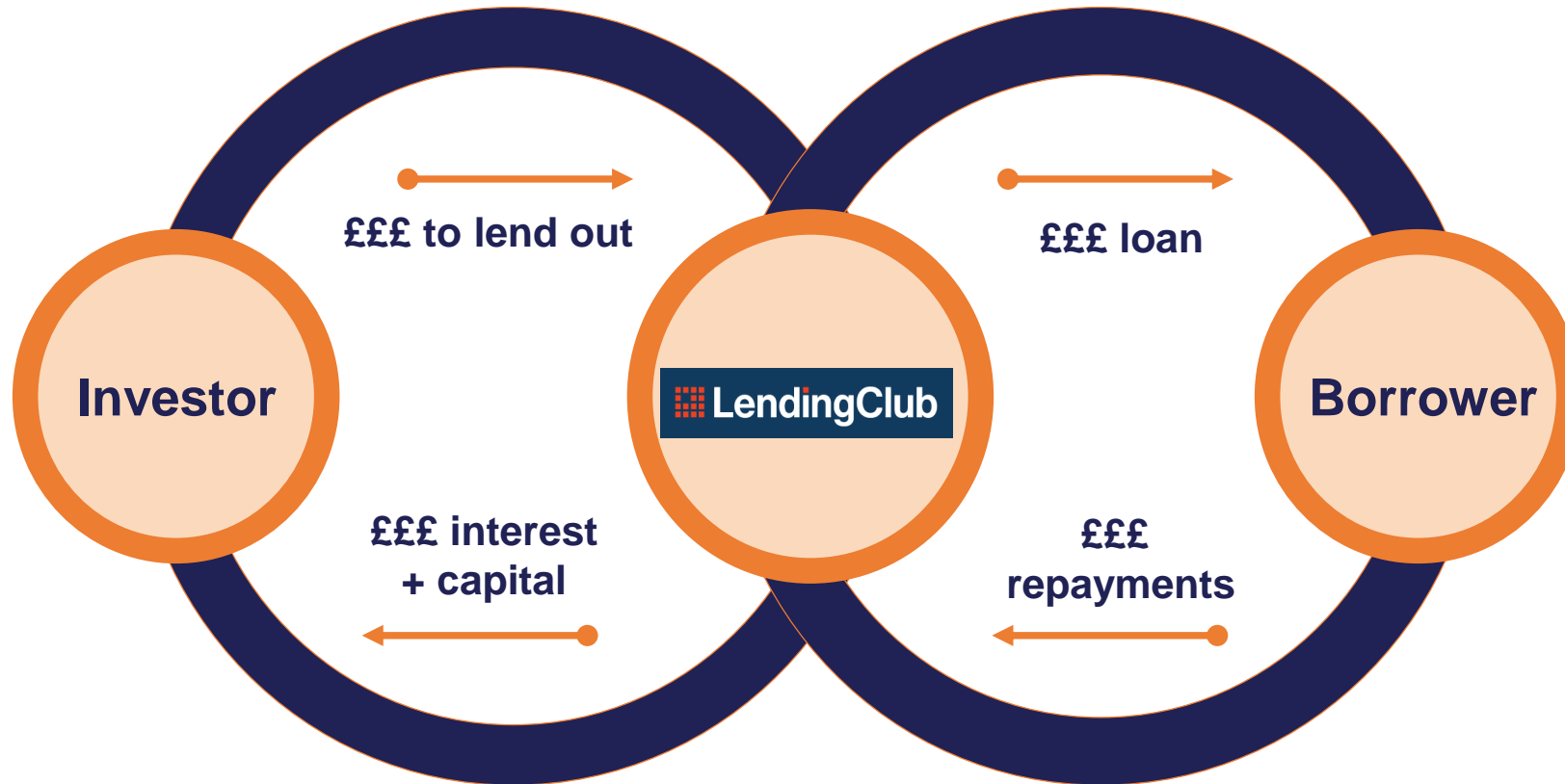


Categorical Variables

Notebook demo



Peer to Peer Finance

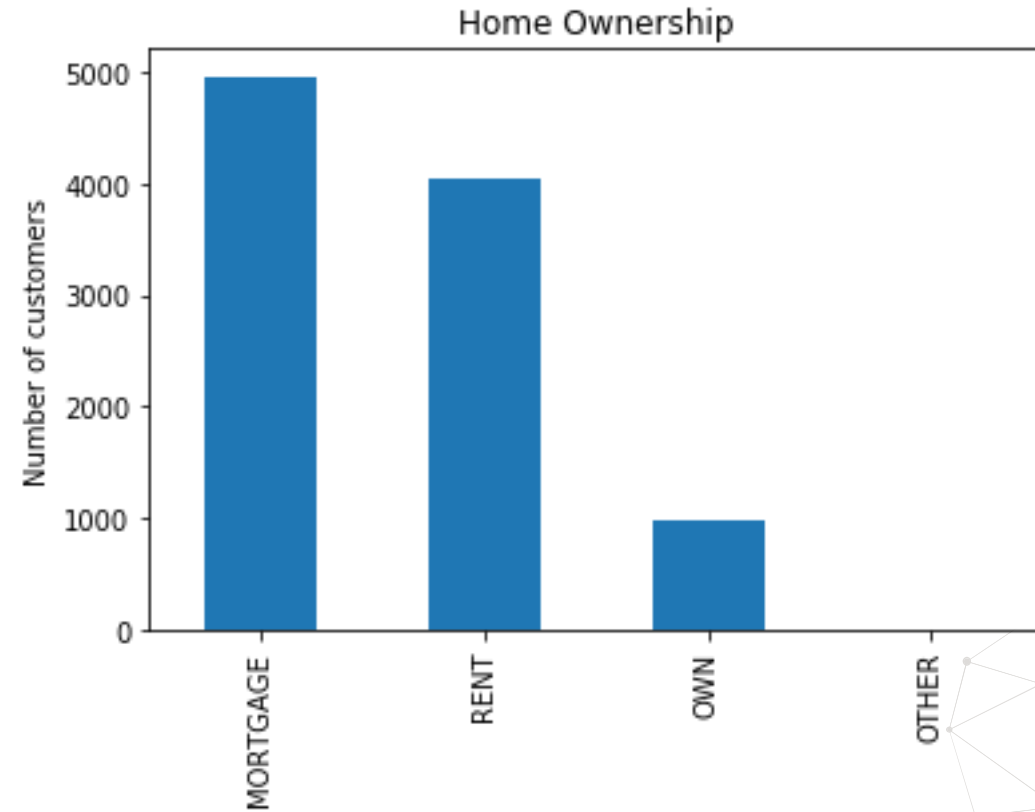


Categorical Variable Examples

- Intended use of loan - Nominal
- Home ownership - Nominal
- Loan status - Nominal

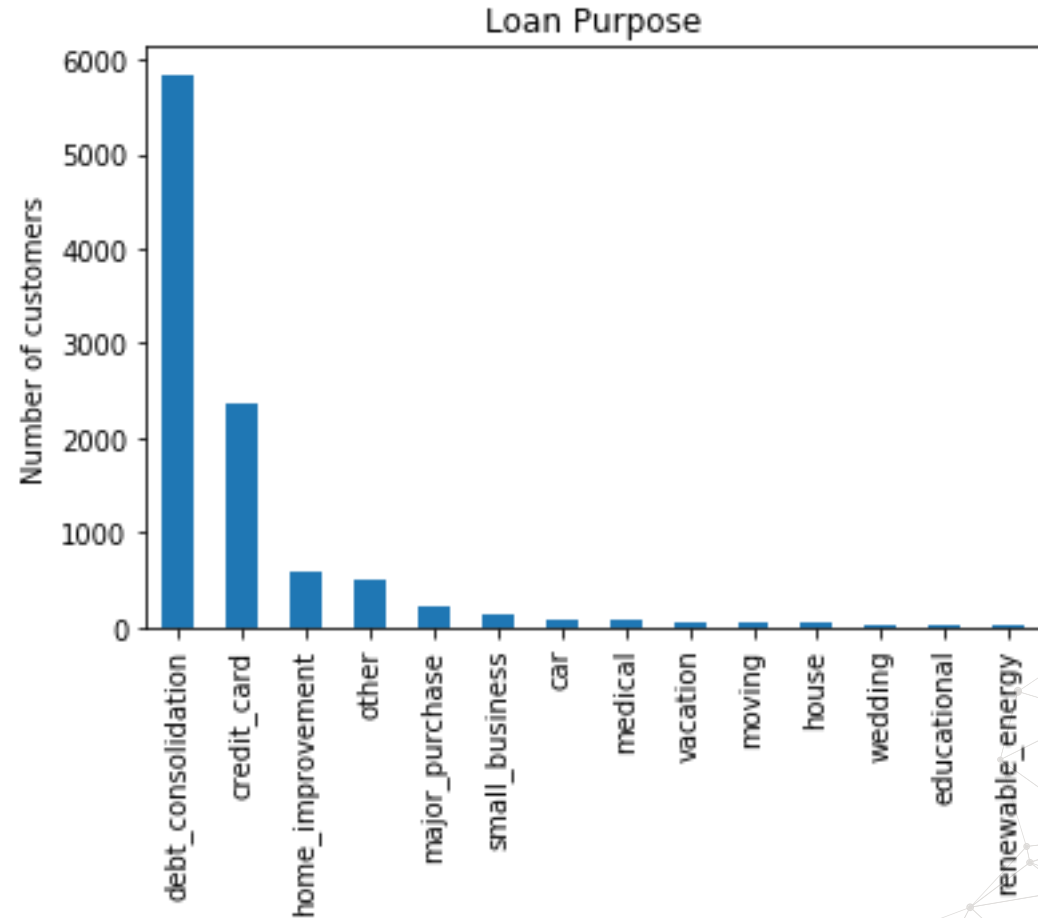
Home Ownership

- Example values: [Mortgage , Rent, Own, Other]



Loan Purpose

- Example values: [Debt consolidation, car, credit car, moving, etc.]



Accompanying Jupyter Notebook



- Read the accompanying Jupyter Notebook

THANK YOU

www.trainindata.com