Train In Data

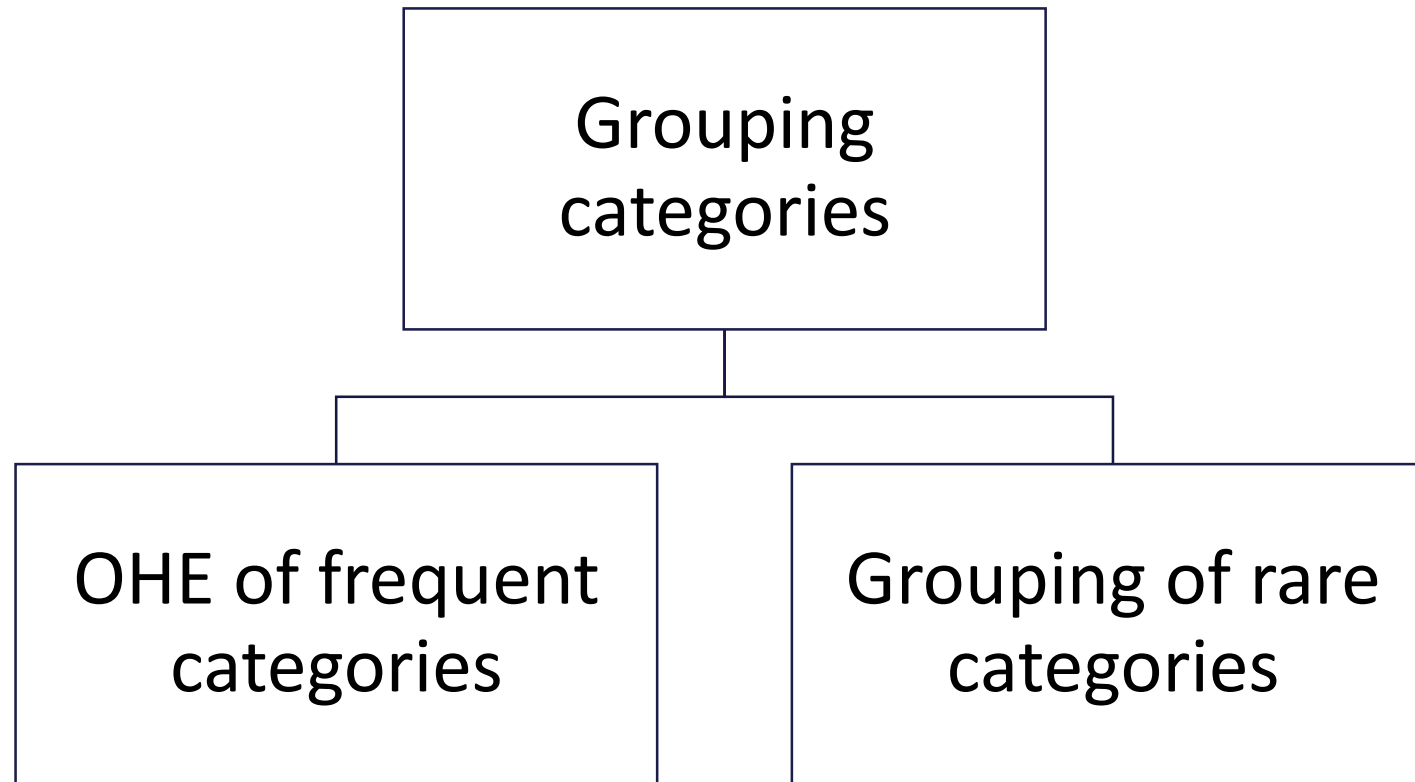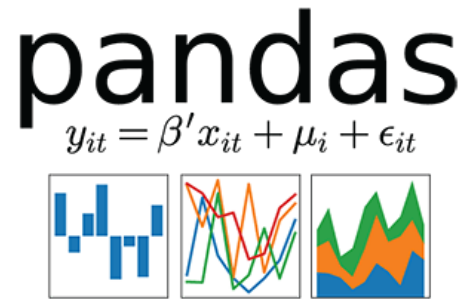# Grouping categories

Wrap up

Grouping rare categories

Grouping categories in Python

Unseen categories

Unseen categories will be treated as infrequent categories

➢ Zero in all dummies (Feature-engine)

➢ Additional dummy for infrequent (sklearn)

➢ Grouped with infrequent categories

Types of target

Binary

Multi-class

Continuous

These encoding methods are independent of the target.

Content

- Accompanying Jupyter notebook

  - Implementation in **pandas**

  - Implementation in **Scikit-learn**

  - Implementation in **Feature-engine**

THANK YOU

www.trainindata.com