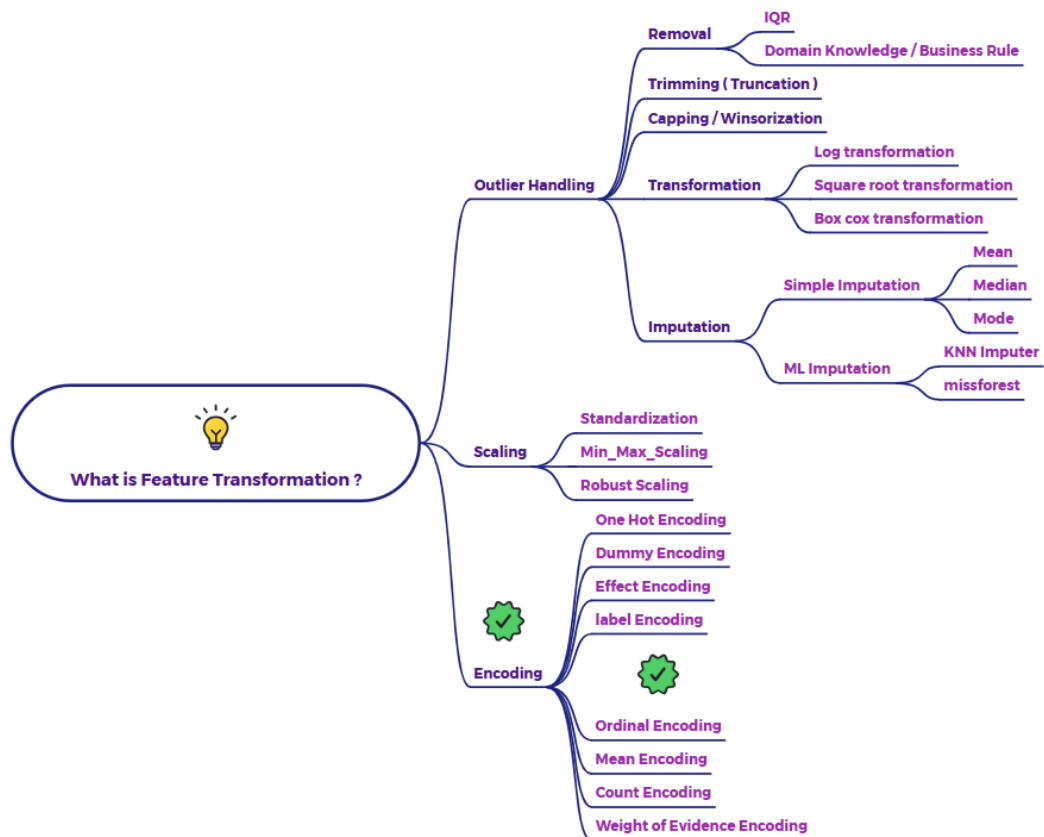


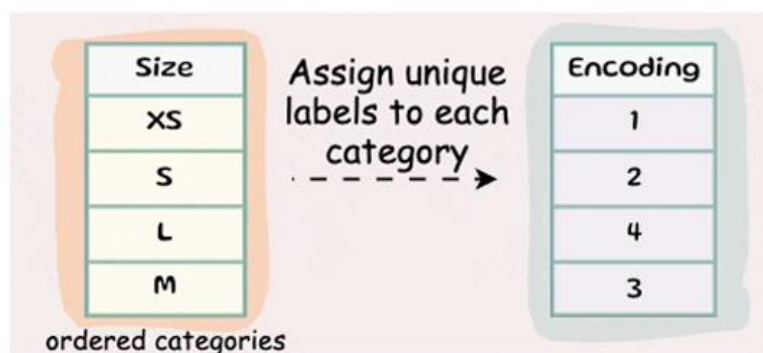
Explain Ordinal Encoding with an example



1. Explanation of Ordinal Encoding

Ordinal encoding is a technique used to convert categorical data into a numerical format, specifically for variables where the categories have a meaningful order or ranking. Each unique category is assigned an integer that reflects this order.

Ordinal Encoding



2. How to Calculate Ordinal Encoding

Here's a step-by-step explanation with an example:

Example:

Suppose we have a dataset with a "Education Level" column:

Education Level
High School
Bachelor's
Master's
High School
PhD

1. Identify Unique Categories and Their Order: The unique categories are "High School," "Bachelor's," "Master's," and "PhD," and they have a clear order:

* High School < Bachelor's < Master's < PhD

2. Assign Integers Based on Order: Assign integers to each category, reflecting their order:

* High School -> 1

* Bachelor's -> 2

* Master's -> 3

* PhD -> 4

3. Replace Categories with Integers: Replace the original categories with their corresponding integers.

The resulting ordinal-encoded data looks like this:

Education Level
1
2
3
1
4

3. When to Use Ordinal Encoding

- When dealing with categorical variables where the categories have a natural order or ranking. Examples include:
 - Education level (e.g., High School, Bachelor's, Master's, PhD)

- Size (e.g., Small, Medium, Large)
- Customer satisfaction (e.g., Very Unsatisfied, Unsatisfied, Neutral, Satisfied, Very Satisfied)

4. Strengths and Weaknesses of Ordinal Encoding

- **Strengths:**

- Simple to implement.
- Preserves the ordinal relationship between categories.
- Does not increase the dimensionality of the data.
- Suitable for machine learning algorithms that can handle ordered numerical input.

- **Weaknesses:**

- Only applicable when there is a clear and well-defined order among the categories.
- If the order is not correctly defined, it can lead to incorrect model interpretations.