



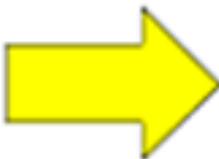
# Ordered ordinal encoding

# Ordered ordinal encoding: definition

- Categories are replaced by integers from 1 to  $k$ , where  $k$  is the number of distinct categories in the variable, but this numbering is informed by the mean of the target for each category.

# Ordered ordinal encoding: example

Color	Red	Red	Yellow	Green	Yellow
-------	-----	-----	--------	-------	--------



Target	0	1	1	0	1
--------	---	---	---	---	---

Color	2	2	1	3	1
-------	---	---	---	---	---

The diagram illustrates the process of ordered ordinal encoding. It shows a single input table with a 'Color' column being transformed into two separate output tables. The first output table, labeled 'Target', assigns numerical values to the colors based on their order of appearance: Red (0), Yellow (1), and Green (0). The second output table, labeled 'Color', assigns numerical values based on a unique mapping: Red (2), Yellow (1), and Green (3). A large yellow arrow points from the input table to the output tables, indicating the transformation process.



# Ordered ordinal encoding: Advantages

- Straightforward to implement
- Does not expand the feature space
- Creates monotonic relationship between categories and target

# Ordered ordinal encoding: Limitations

- May lead to over-fitting
- Difficult to implement together with cross-validation with current libraries

# THANK YOU

[www.trainindata.com](http://www.trainindata.com)