



Encoding categorical variables with Python.

Introduction

Categorical data is data that has more than one category. When working with that type of data we have two types, nominal and ordinal. *Nominal data* is data that has no particular order or hierarchy to it, and *ordinal data* is categorical data where the categories have order, but the differences between the categories are not important or unclear.

We will be working with a dataset of used cars for this article to truly understand and demonstrate how to work with categorical data. Let's explore it and see what type of data we are working with.

```
import pandas as pd

# import data
cars = pd.read_csv('cars.csv')

# check variable types
print(cars.dtypes)
## OUTPUT
# year          int64
# make          object
# model         object
# trim          object
# body          object
# transmission  object
# vin           object
```