

## Problem

It can be difficult to navigate the logistics when it comes to buying art. These include, but are not limited to, the following:

- Effective collection management
- Shipping the paintings, antiques, sculptures, and other collectibles to their respective destinations after purchase

Though many companies have made shipping consumer goods a relatively quick and painless procedure, the same rules do not always apply while shipping paintings or transporting antiques and collectibles.

## Task

You work for a company that sells sculptures that are acquired from various artists around the world. Your task is to predict the cost required to ship these sculptures to customers based on the information provided in the dataset.

## Data description

The dataset folder contains the following files:

- `train.csv`: 6500 x 20
- `test.csv`: 3500 x 19
- `sample_submission.csv`: 5 x 2

The columns provided in the dataset are as follows:

Column name	Description
Customer Id	Represents the unique identification number of the customers
Artist Name	Represents the name of the artist
Artist Reputation	Represents the reputation of an artist in the market (the greater the reputation value, the higher the reputation of the artist in the market)

Height	Represents the height of the sculpture
Width	Represents the width of the sculpture
Weight	Represents the weight of the sculpture
Material	Represents the material that the sculpture is made of
Price Of Sculpture	Represents the price of the sculpture
Base Shipping Price	Represents the base price for shipping a sculpture
International	Represents whether the shipping is international
Express Shipment	Represents whether the shipping was in the express (fast) mode
Installation Included	Represents whether the order had installation included in the purchase of the sculpture
Transport	Represents the mode of transport of the order
Fragile	Represents whether the order is fragile
Customer Information	Represents details about a customer
Remote Location	Represents whether the customer resides in a remote location
Scheduled Date	Represents the date when the order was placed
Delivery Date	Represents the date of delivery of the order
Customer Location	Represents the location of the customer
Cost	Represents the cost of the order

## Evaluation metric

```
score = 100*max(0, 1-metrics.mean_squared_log_error(actual, predicted))
```

# Result submission guidelines

- The index is **Customer Id** and the target is the **Cost** column.
- The result file must be submitted in **.csv** format only.
- The size of this result file must be 3500 x 2.

**Note:** Ensure that your submission file contains the following:

- Correct index values as per the test file
- Correct names of columns as provided in the **sample\_submission.csv** file