# EDD estimation for an e-commerce company

## Problem statement

Logistics is a booming sector with millions of orders getting shipped on a daily basis. Being able to predict the exact time an order can be delivered can be as tricky as it is important.

In this problem you are required to predict the estimated delivery date (EDD) for each order of an ecommerce enterprise. EDD is defined as the number of days between shipment date and order delivery date.

The training dataset comprises daily shipment data of an ecommerce enterprise from June 2022 to August 2022.

You’re challenged to predict the target column for each shipment for the 3 weeks’ data provided in the test dataset. The following target variable is to be predicted:

1. predicted\_exact\_sla

There are the following deliverables for this project:

1. submission.csv with predicted values corresponding to column ‘id’
2. A short description for the model used and the variable importance / trends
3. Data visualization showing important / useful patterns in the data
4. Code used for generating the predictions
5. A brief engineering doc with suggestions on how this can be integrated with the Clickpost system where EDD can be consumed on an order or bulk level

Your submission will be evaluated on the following parameters:

1. Accuracy of predicted values for the target values
2. RMSE between the predicted values and the observed values for the target column for shipments that were delivered within the promised SLA
3. Any other metric that you would like to add or evaluate on

## Dataset

The dataset comprises of three files:

1. **train.csv** - the dataset on which you need to train your model
2. **test.csv** - the dataset on which you need to make your predictions
3. **pincodes.csv** - list of pincodes with corresponding geographic details

Following table provides more details about the fields of the training dataset.

| **Field name** | **Description** |
| --- | --- |
| id | Unique identifier for a shipment |
| order\_placed\_date | Date when the order was placed |
| order\_delivered\_date | Date when the order was delivered |
| courier\_partner\_id | Courier partner identification code |
| order\_shipped\_date | Date when the order was shipped |
| account\_type\_id | Identification code for courier partner account type |
| drop\_pin\_code | Drol location pincode |
| pickup\_pin\_code | Pickup location pincode |
| quantity | Number of items in the shipment |
| account\_mode | Mode used to ship the package |
| order\_delivery\_sla | Number of days taken to ship the item |