

# Python Task: Calculate the efficiency ( $\eta$ ) of a factory line.

## 1 Scenario

You need to report on performance of a factory line. You have access to the production data of the line and have decided to calculate the efficiency for the line.

## 2 Efficiency calculation

Efficiency of a factory line for each batch of an article  $A$  is given by the equation :

$$\eta_A = \frac{Q_A}{S_A \times P_A \times T_A}$$

where,

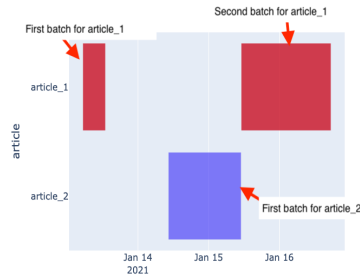
- $Q_A$  : Total number of sacks produced for the article  $A$
- $S_A$  : Target speed [meters/min] of the factory line for the article  $A$
- $P_A$  : Quantity of sacks of Article  $A$  that can be produced per meter of material
- $T_A$  : Duration of the batch of article  $A$  in minutes

**Duration of a batch**  $T_A$  can be calculated as the time duration between the timestamp for first sack of the article produced to the timestamp for the last sack of the article.

$$T_A = last\_timestamp\_for\_A - first\_timestamp\_for\_A$$

A **batch** for an article is defined by a single continuous production of that article on a factory line.

Figure 1: Following figure shows different batches for two articles. For article\_1, there are two different batches identified, as in between the batches there is another article being produced.



## 2.1 Data description

You have received two files:

**Product data (product\_data.xls):** This file contains timestamps for each sack of article produced by the factory line.

**Article Specification data (article\_specifications.csv):** This file contains the important specification for the articles - Target speed and quantity of sacks per meter of material produced by the line.

## 3 Task

Please identify the production batches from the **Product data** and calculate the efficiency of the batches.

Please present your solution in a **jupyter notebook**. You can use any package in python for the task. We would use **pandas** and **numpy**. Also if you can, visualising the batches as seen in the Fig. 1.