

Simulation Modelling Exercise

Objective: Business Understanding: Extract information from customer briefing.

Tools: Combine multiple data sources

Mathematical: Calculate descriptive statistics

Technical: Build simulation model based on extracted data.

A Muesli distribution company has approached you to optimize their delivery schedules. They want to offer Service Level Agreements (SLA) to customers that promises 95% of their deliveries will arrive within the specified lead times. As their product has a long shelf, the length of the lead time is not as important as ensuring orders do not arrive late. Currently the variability of their order process has made it difficult for them to calculate the maximum number of days they should offer as delivery lead time in order to meet a 95% service level.

The warehouse manager described the workflow as follows:

Order received (Day 1) - order processed in warehouse and made ready to ship (normally 2 days) - order leaves warehouse in truck following day - order delivered to customer (handled by logistics company)

The company has provided a list of their transactions over the course of the past years. They have full data on Order Date and the 'On Truck Scan' date but have limited visibility of what happens in between. They have on occasion sent some interns into the warehouse to record the 'Ready to Ship' date for as many orders as they could. The warehouse manager says they have not changed their processes much in the past year so they think it should be a good estimate.

Customers can send orders every day but the warehouse only works Monday to Friday so any orders received on the weekends wait until Monday to be actioned.

Trucks Leave the warehouse on Mondays, Wednesdays and Fridays

Orders leave on trucks the day after the are made ready for shipping (or two days later if there is no truck)

Customers can pay for Express Processing that means the orders leave on the truck the day the order is ready for shipping.

The logistics company has said they have on average 3 day delivery times to all locations. They transport goods on weekends but only deliver to customers from their local distribution centers on weekdays. The Muesli company managed to get their hands on some data about exact delivery dates for a number of shipments via promotions they ran where customers scanned a QR code on the package in order to register for a prize. (We assume customers always scanned the code on the day of arrival).

Deliverables:

1. Make an exploratory data analysis to understand the time taken in each step in the order process. Show how long a 'normal' order takes as well as how much variation exists for each stage. Do the data align with the explanations of the processes given by the company?

2. Based on existing processes and the historical sample data, create a simulation model to determine the lead time that the company should include in the SLA to ensure they are able to meet their 95% on-time commitment for both standard and express shipping options.
3. Could the company offer different lead times for orders received on different days of the week (and for express and standard)? Build a visualisation of the expected shipping times per day of week and shipping method.
4. The company would of course like to improve the lead time, but don't know the best way. The logistics provider has said that they are able to offer daily pick up (Monday to Friday) for an additional cost. Alternatively the warehouse could hire more people in order to be able to handle all shipments as 'Express'. Assuming the costs are equivalent, use the simulation model to determine which option would lead to better improvement in the 95% delivery promise in the SLA.
5. Use a Jupyter notebook to organise the text, code and visualisation that can be delivered to the customer.
6. Present the results and your method to the class.

Hints:

Google sheets can be downloaded as excel files.

Pandas has a `read_xls()` method that can be used for reading excel files.

To use this an additional package called *openpyxl* must be installed using conda.

Alternatively there are packages that allow python to read directly from google sheets, feel free to explore these.

```
df_total["processing_time"] = df_total["processing_time"].apply(lambda x:
np.random.choice(df_total["processing_time"].dropna().values) if np.isnan(x) else
x)
```