# Case Study for Logistics Product Analytics

## — Context —

At Delivery Hero, we aim to delight our customers with **seamless on-demand delivery** for any choice of products. Sometimes, there are issues caused by a variety of factors: higher demand at vendors, availability of riders who make delivery or other region-specific causes. We'd like to understand these causes and their frequency, in order to help improve customer experience.

## — Instructions —

Logistics management team wants to understand the causes of poor order experience, and would like data-driven insights to guide and improve the share of seamless orders.

An order is considered non-seamless if it has one or more of these "issues": ●
Cancellations – which can be done by customer, rider, vendor, or app failures ●
Delays – deliveries being delayed too much compared to originally promised time ●
Customer Service Contacts – customers reaching out for help resolving issues

Some orders may have multiple issues. If so, for classification purposes, Cancellations take precedence over Delays, which take precedence over Customer Service Contacts.

**Tasks**

A. Based on the available data, build the dataset(s) you need for your analysis using ***SQL, Python, R***, or another data programming language.
B. Present your insights and ***top 3 recommendations*** to improve the share of seamless orders, in a slide deck (share as ***PDF***).
C. Include any assumptions you have made and whether additional data could have helped. Call out any data issues you observed.
● Besides the specific tasks listed above, feel free to include any additional insights you uncover from the data. For data visualization, please use your tool of choice other than Excel/Sheets eg. Tableau, PowerBI, Looker, or anything else.
● Please limit your analysis to 4-6 slides.

We estimate the analysis to take maximum 8 hours to complete. Please limit your analysis approach based on this guidance. Although you could spend more time on this, we don't want you to :)

**Submission**

Separately include all code, dashboards and other files you built as part of the analysis; send everything in a zip file, or upload content via cloud file sharing.

## — Data —

Use your choice of language to analyze the data available here (access via the google account

linked to your email address).

*Data-set:* `orders`
*Description:* Basic order metadata
*Fields:*

- `country_code`: ISO country code
- `enc_order_id`: Encrypted Order ID; uniquely identifies an order with country_code. - `enc_customer_id`: Encrypted Customer ID
- `vertical`: Groceries, Restaurants, or Local shops (all others)
- `order_placed_at`: Timestamp of order placement, in UTC
- `status`: ACCEPTED, CANCELLED, DELIVERED, PICKED_UP
- `order_value`: Order amount, in local currency

*Data-set:* `cancellations`
*Description:* Cancellation attributes for cancelled orders
*Fields:*

- `country_code`: ISO country code
- `enc_order_id`: Encrypted Order ID; uniquely identifies an order with country_code. - `cancelled_at`: Timestamp of order cancellation, in UTC
- `cancellation_owner`: Party that cancelled the order: RIDER, VENDOR (i.e. restaurants or shops), CUSTOMER (self-cancellation via the app), PLATFORM (app errors)

*Data-set:* `deliveries`
*Description:* Estimated & actual delivery times of orders that had a rider involvement. Some orders may not have values if they were cancelled prior to assignment of a rider. *Fields:*

- `country_code`: ISO country code
- `enc_order_id`: Encrypted Order ID; uniquely identifies an order with country_code. - `promised_delivery_time`: Original delivery estimate provided to customer at time of placing order, in seconds
- `actual_delivery_duration_secs`: Actual time of delivery since order placement, in seconds
- `delivery_status`: Rider-side status of a delivery: completed, or cancelled

*Data-set:* `contacts`
*Description:* Contacts (chats only) related to an order (if available). An order may have more than one contact. Assume chat is the only mode of contacting customer service. *Fields:*

- `country_code`: ISO country code

- `enc_order_id`: Encrypted Order ID; uniquely identifies an order with country_code. - `creation_timestamp`: Timestamp of contact, in UTC

- `served_chat_ind`: 1 or 0, indicating whether the chat was answered - `missed_chat_ind`: 1 or 0, indicating whether the chat was missed. Mutually exclusive with `served_chat_ind`.

- `contact_reason`: customer's reason for contacting support