

Supply Chain Analyst

Candidate assessment

Guidelines

Your task is to exhibit creativity and resourcefulness in data gathering to support your hypotheses. Focus on demonstrating clear thought processes and applying fundamental principles to deconstruct the problem statement. Present your findings in a format of your choice, emphasising the reasoning behind your problem-solving approach. Please also include access to your code, either in Github or attached to your submission.

This assessment aims to evaluate your rate of learning and ability to iterate effectively. Given the complexity of the task, it's understood that you might not have sufficient time to complete every aspect. Therefore, you will have a timeframe of 3 to 4 days to work on this. This constraint is intended to assess how you prioritise and manage your time under a tight deadline.

This assessment is challenging, and we place greater emphasis on your approach to tackling the challenge rather than the final outcome. The key criteria for evaluation include:

- Problem-solving approach and clarity of thought
- Effective communication of your methodology
- Assumptions and decision-making rationale
- Quality of insights derived from the data

Objective

Exploratory Data Analysis

Analyse the data from 'allocation_algorithm_logs_20240105.csv' to gain insights into the fulfilment network and the allocation algorithm. Pay special attention to metrics that reflect the quality of allocations, including:

- Distribution of orders across facilities
- Allocation costs
- Stock availability
- Geographical aspects (provinces of destination and fulfilment centres)
- Other relevant or interesting metrics you identify

Optimal Fulfilment Centre Allocation

Develop a script/function to determine the most cost-effective allocation of order items to fulfilment centres, based on an allocation request and the associated available inventory (available in the allocation_algorithm_logs_20240105 file). Your script should aim to minimise allocation costs.

bash

Present your script's output and compare your results with the existing allocation algorithm. Include any shortcomings of your approach and the next steps you would take to improve your results.

Feel free to reach out with any queries about the data or to validate your assumptions during the assessment.

Data description

`courier_costs.csv`

This file contains the Rand value associated with shipping a parcel with each courier depending on the route classification ("Local", "Regional", "Main") and order type ("DoorToDoor", "StorePickup")

`courier_post_code_mapping.csv`

This file contains the mapping of postcodes to their nearest hubs for each courier, as well as the classification of the postcode relative to its nearest hub (LMR = "Local"/"Main"/"Regional")

`bash_fcm_facilities.csv`

A mapping of facility_id to facility postal code

`dc_branch_numbers.csv`

The branch numbers of all DC facilities. If a branch number does not appear in this file, then the facility is a "store".

`allocation_algorithm_logs_20240105.csv`

This file contains the logs of the current allocation algorithm running in production. Each row contains JSON objects containing:

- request: The allocation request for an order
- available_inventory: The stock available at any facility which can fulfil at least one item in the order
- plan: The output of the allocation algorithm. A list of order items and which facilities they were allocated to

bash

request object key fields

- `destination.branch_number`: The branch number of the selected collection point if the delivery type is “StorePickup”
- `destination.delivery_type`: Whether the order is for home delivery (DoorToDoor) or a collection from a designated store (StorePickup)
- `items`: A list of all items in the order
- `item.order_item_id`: the id of the order item
- `item.quantity`: the number of units of the item that were ordered
- `item.sku`: the SKU number of the item
- `item.trading_company_number`: There are distinct trading companies under the general TFG umbrella. This value represents the trading company the SKU belongs to.

available_inventory object key fields

- Each item in the list represents the quantity of a specific SKU available at a specific facility
- `FacilityCourierCActive`: whether `courier_c` is eligible to fulfil parcels from this store
- `FacilityCoordinates`: The Well-Known Binary (WKB) representation of the facility coordinates
- `FacilityType`: Whether the facility is a store or distribution centre
- `NormalisedSku`: This is the SKU number appended to the `TradingCompanyNumber`. It is an important field, as SKU number is not necessarily unique across trading companies
- `OnHandQuantity`: The number of units of the SKU in that facility
- `QtyAvailable`: The number of units of the SKU that are available to fulfil from that facility

plan object key fields

- `external_order_reference`: The order number associated with the allocation
- `reservations`: a per item list of the allocations created by the algo
- `reservation.courier`: The suggested courier to be used for fulfilment
- `reservation.sku`: as above
- `reservation.trading_company_number`: as above

DC Branch Numbers

In TFG, each brand that carries stock in a distribution centre has its own “virtual branch” within the DC. Multibrand orders allocated to these DCs can still be packed into a single parcel, so the allocation algorithm should consider all “DC virtual branches” as a single facility.

The `dc_branch_numbers.csv` file can be used to identify which branches are in fact virtual branches within the same dc.

bash

Route guide logic

Couriers bill based on the hub that the collection and drop off postcodes fall in, and the classification of the route into either “Local”, “Main” or “Regional”.

For courier_a and courier_c, the overall route classification can be derived as follows:

- If either the collection postcode or the delivery postcode is “Regional” to its closest hub, the overall route is classified as “Regional”.
- If the postcode is either “Local” or “Main” relative to the hub, and the parcel needs to be collected from one hub and delivered to a different hub, the overall route is classified as “Main”.
- If the postcode is either “Local” or “Main” relative to its closest hub, and the parcel needs to be collected and delivered in the same hub, the overall route is classified as “Local”

For courier_b, the mapping is simpler as all courier_b parcels are collected from the same facility. As such, the location of the delivery post code defines the overall route classification.

Courier availability

- courier_b is only able to fulfil parcels from dc_b branches (identified in the dc_branch_numbers file).
- courier_c is only able to fulfil parcels from facilities with FacilityCourierCActive = True, and to customer addresses that fall within 20km of any Courier C Active facilities.
- courier_a is able to deliver from all facilities except for dc_b branches.

General assumptions

- Assume that all items allocated to a single facility can be packed into a single parcel, regardless of the number of items.