SWIFT Assignment - Choked Hubs Prioritization

You are working for a logistics company which works with courier partners to ensure that customers who place orders on online websites get their orders delivered to them in a timely and efficient manner. The dataset provided contains information about orders (or shipments) journeying from a brand's warehouse to the customer's address through a series of courier warehouses. The journey starts by the parcel to be delivered to the customer getting picked up from the seller/brand's warehouse and making it's way through a road network within large vehicles or trucks and stopping at certain courier warehouses during the journey, where it is removed from the current vehicle, sorted based on it's destination and moved into another vehicle destined for the next warehouse on the way to the customer's location. Using the given dataset, your task is to find out which Courier Warehouses are choked and which should be prioritized by the on-ground operations team for clearing / avoiding usage.

Dataset Fields:

1. **shipment_id** : A unique identifier for a parcel
2. **latest_status:** The latest state of the parcel
3. **latest_location:** The latest location of the parcel
4. **deduped_track_details:** Array (or list) containing the entire journey details of the parcel, i.e. the time at which a parcel was last seen **(ctime)** in a particular courier warehouse **(location)**

Notes:

- The analysis should be run in a SQL Workspace. The dataset is in a json newline-delimited file. Most SQL workspaces support loading newline-delimited json files into tables. If your chosen workspace does not support json (or this format of json), we leave the dataset file in your able & innovative hands dear candidate and hope you figure out a method to load the data into a SQL workspace, or else, this entire exercise shall be in vain.
- Take the current date to be 07-October-2023 and work accordingly.
- You might find blank/missing data in the dataset. Dealing with it on your own is a part of the assignment.
- A choked warehouse is one where shipments have entered but are not exiting within an optimal period of time due to some operational issue, causing delays in delivery of the

shipment to the customer. The mathematical definition of a Choked Warehouse needs to be generated by you, the candidate and presented in your solution.

- The objective of this assignment is to present a logical approach to the solution generated by you. The solution generated should tell us which Courier Warehouses to “Prioritize for Clearing” or to “Ignore” because performance is within expectations.
- The approach should be presented in a document format (.doc or .pdf file) along with a .sql file containing the code used to generate the result. The resulting solution should contain a list of Courier Warehouses with a category tag (“Prioritize for Clearing” or “Ignore”) and be in a .csv file.

  Hence, we need 3 files from you as part of completing the assignment:

  1. DOCX or PDF containing your approach
  2. SQL/SQLX file containing the code used to generate the result
  3. CSV file containing a list of warehouses along with the priority category tag.
- Feel free to add any new insights into the DOCX/PDF file that you can extract out from the data provided to you.

We prefer you send the solution within 24 hours of receiving the assignment regardless of the official deadline, unless of course you want to take more time and surprise us with analytical wizardry.

Dataset Link (if not accessible, check with recruiter):
https://drive.google.com/file/d/1UPLKnrLrzTMBLR8IEbuAnKwVp0bDX1vA/view?usp=drive_link