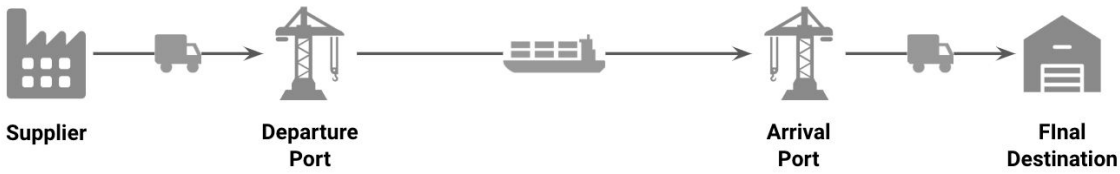
Flexport Data Analytics Interview Case Assignment

Context: As a freight forwarder, Flexport manages the complexity of shipment processes and provides a seamless transportation experience for our clients.

The profile of these global supply chains are innumerably complex and varied - one of the primary reasons we're succeeding by using structured data to upgrade the process - but for the purpose of this exercise we'll present a simplified view.



A **Consignee** purchases goods from a **Supplier** to then sell them to another party. An example consignee might be a Californian retailer of teddy bears that partners with a Taiwanese supplier to manufacture and deliver those teddy bears. The retailer might then sell those teddy bears directly to consumers via their network of retail stores.

Consignees place a **Purchase Order (PO)** with their Supplier. Today Flexport's consignees can choose to do this via our platform ("online") or through their own established channel ("offline"). Suppliers then confirm a **Booking** against the Purchase Order(s). Similarly, Flexport's suppliers can choose to do this online or offline. Bookings to POs can be M:M.
After a booking has been made, the **Shipment** is created in Flexport's system and freight services are procured for its transit. A booking to shipment is 1:1.
Therefore, there are 3 distinct ways that a shipment can originate in Flexport's system:

1) Consignee and supplier are fully using the system
   Purchase Order -> Booking -> Shipment
2) Consignee and supplier are not using our PO service, but still using our Booking service
   Booking -> Shipment
3) Consignees and suppliers are interacting offline and sending us finalized shipment notice
   Shipment

The two primary modes of freight are Air and Ocean. For Ocean shipments, the **Products** are transported within **Containers**. Air shipments do not use containers.

Assignment: Leveraging the data that you received, we want you to extract and understand key insights and opportunities inherent in our data. Think of this as "data in; insights out". Keep in mind that Flexport aims to provide our clients with fast, reliable, on-time deliveries at low costs.

Please consider the following:

- Feel free to ask questions at any time!
- You may employ any analytical techniques or tools in your solution, ranging from data visualization to machine learning. Note, we find that the strongest submissions have some element of metric design, dashboard design, and *actionable* exploratory analysis.
- Please share your detailed analysis (SQL queries, Python/R scripts, notebooks, and any other underlying work). **Accepted formats are PDFs, any text file, .ipynb, word/powerpoint, and public web links**.
- If you need to take any reasonable assumptions on missing or undefined data, feel free to do so and clearly state them in your work.
- For information that is not included and you think would be interesting or helpful to have, describe the additional dataset(s) that would be beneficial.
- If invited onsite for an interview, you will be asked to present your analyses in front of several interviewers (10-15 mins) with an additional 10-15 minutes of Q&A. You can present your analysis in any format.

Attachments included

Data Dictionary
Datasets*:

- Suppliers
- Purchase Orders and Bookings
- Consignees
- Shipments
- Containers

*These tables have been randomized/anonymized for obfuscation, but are based on actual datasets. As with real datasets, you may see errors and irregularities. In general, feel free to make any assumptions and exclusions you think are appropriate and state them in your work.