ClearMetal Data Science Analysis Challenge

Instructions:
For this challenge, we are interested in understanding your organization, data analysis, and communication skills. You have free rein to take this analysis in any direction you may choose.

We do not expect you to spend more than 3 hours on this challenge. You may share your thoughts on future work if you don't finish within the timeframe of the analysis.

About the data set:
The dataset accompanying this document is an anonymized sample of predictions for a subset of shipping containers. The column definitions are as follows:

- Container_trip_id: An integer that uniquely identifies a single container shipment and all the milestones associated with that shipment
- Milestone_type: A string that identifies key events in the shipment life cycle
- Milestone_time: A datetime that indicates the date and time at which the above milestone occurred
- Milestone_location: A string that indicates the name of the location at which the above milestone occurred
- Milestone_unlocode: A 5-character string that is used as a standardized location code (one-to-one with milestone_location)
- Date_received: A datetime that indicates when the milestone update was received (distinct from when the milestone actually occurred)
- Ocean_carrier_code: A 4-character string that uniquely identifies the ocean carrier that is responsible for the shipment
- Port_call_id: An integer that uniquely identifies a vessel stop at a port
- Vessel_id: An integer that uniquely identifies the vessel associated with the current milestone (although sometimes the carrier attaches the incorrect vessel or attaches a future vessel to the current milestone)
- Vessel_name: A string that indicates the name of the vessel associated with the current milestone
- Status_code: A raw code associated with the current milestone
- Eta_type: Either 'future' or 'past' - indicating whether the discharge_eta (see below) is for a milestone that has already occurred or has yet to occur
- Planned_discharge_location: A string that indicates the name of the location at which the container is scheduled to be discharged (aka unloaded from the vessel at the final port in the container's journey)
- Planned_discharge_unlocode: A 5-character string that is used as a standardized location code (one-to-one with planned_discharge_location)
- Discharge_eta: A datetime that indicates when the carrier thinks the container will be discharged at the final port in its journey

Expected output:
This is mostly up to you. You can choose to present your results as a Jupyter notebook (or something similar), a slide deck, a PDF, or whatever you feel is most appropriate. Keep in mind that we might choose to use this dataset and your output in subsequent conversations to ask clarifying questions and/or dig deeper into the data and your analysis.