Problem Statement:

In a supply chain network, A shipment journey can have several sequential events before successful delivery to destination. Delay in an event can cause delay to subsequent events. Also, a shipment may need multiple vessels to complete the journey.

E.g. A shipment journey from Mumbai to Chandigarh might have 2 sub journeys. First, Mumbai to Delhi in a large Truck and then Delhi to Chandigarh in a small truck. Each sub journeys have events like loading and unloading from trucks and some events like inspection. On the other hand, the truck from Mumbai to Delhi might carry several other shipments of different customers of Delhi, Chandigarh, Kanpur, Shimla and many more locations.

The shipments can also get delayed due to several reasons like natural calamities, geo-political unrests, accidents, human strike etc. A delay in shipment can cause a potential disruption at receiver's end, and hence business wants to build a forecasting model to forecast delays at a destination so that receiver can plan to mitigate the risk of disruptions.

Data Description:

We have 2 data sources.

1. 2 years temporal data of vessel movement across the globe.
2. Geographical data

Due to congestion in Port sometimes vessels need to wait near the proximity of the ports for several days (known as stationary hours) leading to delays in shipments. So, it is very important to forecast delays so that subsequent the ground operations at destination side can be better planned.

You are free to use any **external data sources** that you might think useful.

Expectations from this Challenge:

We expect you to work on this challenge as a take home assignment. This should not take more than 8 to 10 active hours. You should spend time on Weekend and send us back the solution by early next week so that our team can go through the solution before You present a summary (of any form) on this challenge focusing on the below key aspects on Wednesday.

1. Your realisation on the problem based on Exploratory Data Analysis.
2. High level solution approach.
3. Data preparation and Feature engineering to achieve the solution describe in 4.
4. Choice of forecasting models with justification. We expect you to capture long term dependencies in the model for accurate forecasting. Ideally you should start with a simple statistical model, evaluate the results and identify the gaps in the model. Suggest an advance model (preferably a using Deep Neural Networks) to overcome the challenges that you faced in the simple model.
5. Evalution of your forecasting model. Your realisation after Evalution.
6. The assumptions you have made to complete the analysis with justification.

7. Sample results – 1. Forecast of stationary hours for a port (or at a Terminal as you think as suitable), 2. Explanation of your results – Reasons behind your prediction. These are just 2 examples; you are free to add any results that you might think useful for business decisions
8. Scope of improvements.

Note: In case you got any questions on the data and problem statement, please write to abhishek.maity@maersk.com for better understanding. You can also reach to Abhishek on +91 – 9740240822 for quick response.

Please find the data descriptions below.

Vessel journey data:

'call_id': ID variable to identify a Vessel at a Port.

'arrival_date': arrival date of a vessel at a port.

'departure_date': departure date of a vessel at a port.

'arrival_time': arrival timestamp of a vessel at a port.

'departure_time': departure timestamp of a vessel at a port.

'imo': Vessel identifier

'vessel_name': Name of the vessel

'vessel_type': Type of the vessel

'Country': Name of the country

'CountryCode': Identifier of the Country

'Port': Name of The Port

'Terminal': Name of the Terminal

'Duration': Duration of vessel stay in a Port

'Service_name': Name of the service

'length': vessel length

'vesteu': capacity of a vessel

'prev_departure_time': departure time of a vessel from the previous port.

'prev_port': Previous port visited by the vessel

'prev_arrival_time': arrival time of a vessel at the previous port.

'prev_duration': Duration of vessel stay in previous Port

'prev_leg_duration': Time taken to travel from previous port to current port

'prev_leg_distance_nm': Distance travelled from previous port to current port

'prev_leg_stationary_hours': waiting time at the vicinity of Port to get a place to unload the cargo

'prev_call_id': ID variable to identify a Vessel at its previous Port.

'next_port': Next port visited by the vessel

'next_arrival_time': arrival time of a vessel at the next port.

'next_duration': Duration of vessel stay in next Port.

'next_leg_duration': Time taken to travel from current port to next port

'next_leg_distance_nm': Distance travelled from current port to next port

'next_leg_stationary_hours': waiting time at the vicinity of next Port to get a place to unload the cargo.

'next_call_id': ID variable to identify a Vessel at its next Port.

Geographical Data:

'CountryCode': Country Code

'Country': Country Name

'PORT_UNLOCODE': Port Code

'PORT': Port name

'LATITUDE': Latitude

'LONGITUDE': Longitude