PORTCHAIN

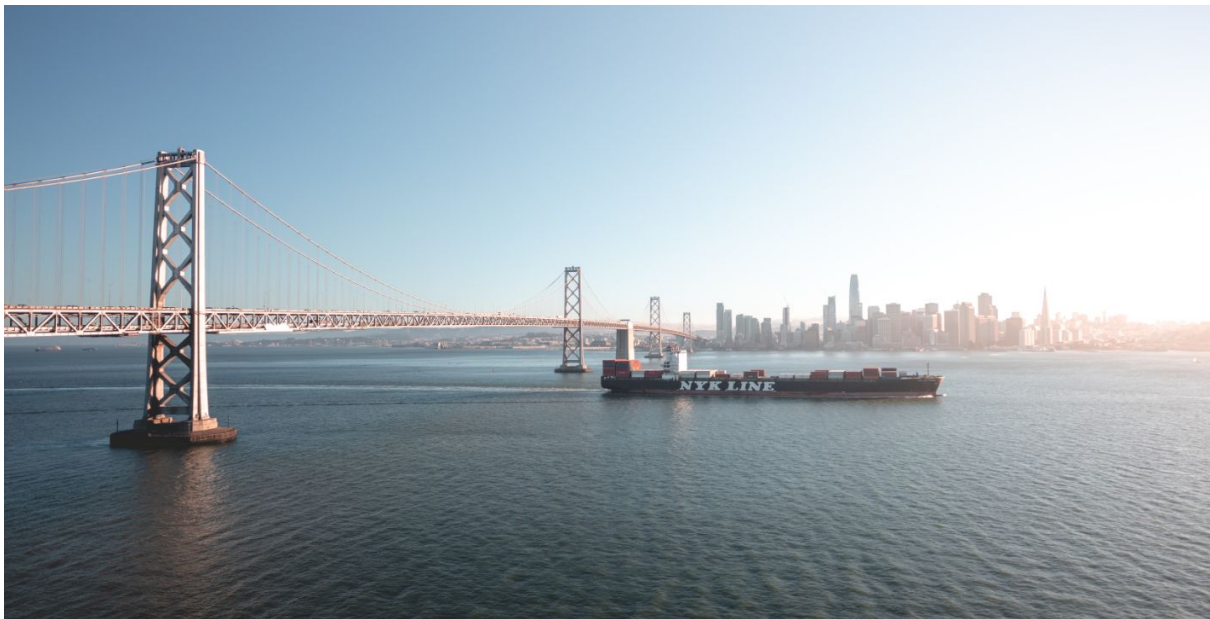**Coding challenge:**

Import and analyse container vessel schedules

Context about the shipping industry

90% of all manufactured goods in the world - worth $15 trillion - are moved through a container. The backbone of this global trade is made up of two main types of companies: Carriers and Terminal Operators.

Carriers are also known as shipping lines, and they operate the 6,000 container vessels sailing around the world at any given time between the different terminals. The vessels come in all sizes, with the biggest ones being 400m long, 60m wide and the capacity to carry more than 23,000 containers.

Terminal Operators operate the 1,200 terminals that exist globally. A terminal is a part of a port and they operate large cranes that move the containers between a vessel and a

PORTCHAIN

parking space for the containers behind the cranes (called a Yard). Loaded and unloaded containers are continuously delivered and picked up by trucks, rail or other vessels, and transported either inland or to another port.

The vessels run on Schedules, which is the equivalent of a bus having a route with planned stops along the way, and when it will arrive at each stop. A stop for a vessel is called a Port Call, and globally there are 500,000 port calls annually.

Context about the data

- A port call is a vessel berthing at a container terminal to load and unload containers.
- A Port may have more than one container terminal. For simplicity in this exercise, the port calls only reference the port and not the actual container terminal.
- Sometimes a port call is present in a schedule but is temporarily skipped as a one-off event. This is called an "omitted" port call and is reflected through the `portCall.isOmitted` boolean field.
- The data in this challenge is real data that was gathered between Jan 1st 2019 and May 31st 2019.
- Although each vessel schedule is planned months in advance, delays and schedule changes happen frequently as the vessel approaches its different ports due to a range of reasons..
- A port call duration is the time during which the vessel stays berthed to load and unload its containers. This can be calculated by subtracting the arrival to the departure time:
  `portCallDuration = portCall.departure - portCall.arrival`

Your task

Your task is to build a tool to import vessel schedules from an external data source and display interesting statistics (defined below) about these schedules.

Step 1: Capture vessel schedules from the provided APIs

All available vessels can be listed here:
https://import-coding-challenge-api.portchain.com/api/v2/vessels

For each vessel, you can call the following API to get the full list of port calls and how they evolved with time (i.e., if they are omitted, delayed etc):
https://import-coding-challenge-api.portchain.com/api/v2/schedule/<vesselIMO>

For example, the following link requests the vessel schedule between January and May 2019 for the container ship MILANO BRIDGE:
https://import-coding-challenge-api.portchain.com/api/v2/schedule/9757187

Step 2: Analyse the data

From the data you have retrieved in Step 1, please now display the following information:

- The top 5 ports with the most arrivals, and the corresponding number of total port calls for each port.
- The top 5 ports that have the fewest port calls, and the number of total port calls for each port.
- For each port, the percentiles of port call durations: 5th, 20th, 50th, 75th and 90th percentiles.
- For each vessel, calculate the 5th, 50th and 80th percentiles for the port call delay when the vessel is 14, 7 and 2 days from arrival.
  Cf. Annex 1 below.

You can choose to display the data through a terminal/console or a web user interface, depending on your own preference.

Guidelines

- Provide a solution to the steps above that you are proud of, and that you think reflect your skill sets. Pay special attention to the code organization and its readability.
- There is no need to store the data in a database or build a full webapp.
- Write tests for your code, including anything that may cover edge cases.
- Make the code easy to run (include a README file).
- Do not reinvent the wheel. Where possible, use existing tools and libraries.
- Don't hesitate to reach out if you have any questions along the way

Good luck!

PORTCHAIN

Annex 1

A **delay** is the difference between a port call's 'forecasted' time of arrival vs. its 'actual' time of arrival .
- The 'forecasted' time of arrival is the time of arrival 'as we know it' from the port call logEntries, at any given point in time.
- The 'actual' time of arrival is fixed and is the latest known arrival time of the port call. The API returns this actual value as the `portCalls[N].arrival` value.
- Therefore delays are functions of time and depend on the forecast point. These forecast points are found in the `logEntries[N].createdDate` field.
- All delays are positive: always take the absolute value

Example

We're given a port call for which the vessel arrived at 13:00 on March 15. Let's say we only have 3 log entries:

  - On March 15 at 13:00, the vessel is arriving at this exact time. This update is made within 2 days of arrival: ignore this data point because it was created less than 2 days from arrival.

  - On March 13 at 12:59, the vessel arrival time was forecasted for 10:00 on March 15. This update is the latest regarding the vessel arrival time that is more than 2 days from the actual arrival time. In that case, the 2 day delay is 3 hours: `abs(forecastedTime - actualTime) = 3 hours`.

  - On March 1 at 00:00, the vessel arrival time was forecasted to arrive at 18:00 on March 15th. This update is the latest we have, both 7 and 14 days before the actual arrival time; therefore both the 7 and 14 day delays are in this case both 5 hours.

The *simplified* data for this example is below:

```
{
 "vessel": {...},
 "portCalls": [{
     "arrival": "2019-03-15T13:00:00+00:00",
     "port": { "id": "DJPOD", "name": "Djibouti" },
     "logEntries": [{
       "updatedField": "arrival",
       "arrival": "2019-03-15T18:00:00+00:00",
       "createdDate": "2019-03-01T00:00:00+00:00"
     }, {
       "updatedField": "arrival",
       "arrival": "2019-03-15T10:00:00+00:00",
       "createdDate": "2019-03-13T12:59:00+00:00"
     }, {
       "updatedField": "arrival",
       "arrival": "2019-03-15T13:00:00+00:00",
       "createdDate": "2019-03-15T13:00:00+00:00"
     }]
 }]
}
```