### **[COPY]** [**EXTERNAL LINK**](https://docs.google.com/document/d/13NhC0F5F9IK83QKJfVzGXHyKM9xoTq19fa_Ql4DyZZA/edit#heading=h.8be9prerbvsc) **- please duplicate any updates there**

### **Product Operations > Business Technology Analyst Case Study** [Weekly Presentations, Python Notebook, Sql Analysis]

### **Goal of the Case Study** [**Important]**

*[To qualify for the next technical round i.e code presentation please remember 3 points]*

(i) Understand the basics of container shipping with the steps laid out in “[Learn the Basics](#_9pdzaonx1ngc)”

(ii) DELIVERABLES: Add all of the below in a Gdrive folder shared with us on and [recruit@portcast.io](mailto:recruit@portcast.io) who have been dealing with Customers Accounts

1. **Ppt - Powerpoint presentation** - summarize the questions, graphs and data you want to explain
2. **Google Colab - Python Notebook** equivalent available in your Gdrive for FREE use your gmail account to sign-in and create a file in your Gdrive using <https://colab.research.google.com/notebooks/intro.ipynb>?
3. **CSV exports of Anomalies** of specific containers in the data dump that you downloaded
4. **Data Dump Downloaded:** The converted JSON to CSV downloaded from S3

(iii) In ML/AI product the backbone is data, you should be able to demonstrate data challenges you encountered and how did you handle them in your Ipython notebook analysis and the presentation in your next round

### **Learn the Basics:**

**Tracking Assets: A.** Containers# **B.** Bill of Lading# (invoice holding multiple Containers#)

1. Hit the url <https://www.msc.com/track-a-shipment> [Use the search bar for values given]
2. Use **Container #’s** [Each Container is shipped by a Carrier]
   1. SEGU9765052 **(On Vessel, ETA)**
   2. CAIU7626180 **(On Vessel, Transhipment, ETA available)**
   3. MSCU4847068 **(No ETA, already Arrived)**
   4. FDCU0449590[[1]](#footnote-0) **(Invalid Container Number**)
3. Use **Bill of Lading #’s** [Each of bill of lading is a collection of Container #’s]
   1. MEDUC6599558 **(Active)**
   2. MEDUAQ786421 **(Old, but Valid i.e expired)**
   3. WWLLWYTN0101088 **(Invalid BL**)

**CheatSheet:** Study the entries of each of the data points in search results and try to understand the meaning of each especially

* POL = Port of Loading and POD = Port of Discharge [Start & Finish of the ocean journey], and
* Arrival Date which is ETA (Expected Time of Arrival) as per Schedule: please note this is not a prediction just the time table schedule of a given carrier for a container

### **Case Study:**

**Customer Profile**

Changi Express Logistic (CEL) is an international freight forwarder with a monthly container volume of up to 50K containers globally via ocean, air and land. They pride themselves as one of the leading freight forwarders that deliver On-Time and In-Full delivery to their customers. More recently, CEL observed a sharp decline in carrier reliability schedules across all major trade lanes which in turn affected their operations globally. The data that the CEL operators receive from the carriers directly is often missing, is inaccurate or is only updated very late into the journey/just a day before the arrival. That does not offer enough time for the CEL teams to communicate the delay back to their own customers.

Thus, CEL’s Innovations Team has sourced a handful of vendors (including Portcast) for testing on how to improve existing Ocean Supply Chain Visibility. Portcast is now 4 weeks into the POC for the predictive ETA solution. Nikana-San leads Innovations Projects and is our key stakeholder within CEL. Our weekly 1-hour check-ins are typically used to clarify on predictions being generated from our backend APIs/Emailed Reports for the cargo containers uploaded by the customer on a daily basis at a predefined schedule. E.g. 5.00am daily, each batch of approx. 200 new cargo containers (Bill of Lading No.s or Booking No.s or Container No.s are used to make a POST request on our API system)

With 80% of the container volume allocated for the trial having arrived at the POD, you are now working closely with the Account Manager to prepare for the next mid-trial review of the performance so far.

**Goal of this Case Study:**

The required outcome of the following exercise is to convert CELL from a Trial Account/POC-Proof of Concept to a Paid account. You may want to consider the following Business consideration when preparing the slide deck:

0) Defining 3 key metrics to determine success of the trial

1) Showcasing existing visibility supply chain performance of the customer

2) Performance improvement from using Portcast predictions

3) Improvement plan for the anomalies, including

* [Internal]: communicated internally within Portcast to deliver better service levels to CELL and other customers
* [External] communicated externally with CELL

### **Assignment:**

**Please complete the following:**

In this task, you will create data summarization for weekly check-in with the customer after downloading the data-dump from the AWS S3 link provided. *(Below the words dataset and data dump are interchangeably used)*

1. Create an Ipython Notebook in Google Colab using your gmail account for FREE (**Hint:** [Welcome To Colaboratory - Colaboratory](https://colab.research.google.com/notebooks/intro.ipynb)? )
2. Download the JSON data-set and build a python client to read data in Google Colab. (attached in the email along with this assignment)
3. You are expected to transform this file from JSON to CSV and upload in your GDRIVE folder to make this usable for yourself both to manually explore using Google Sheets and programmatically using Ipython Notebook in Google Colab (**Hint:**  you will have to colocate both the dataset and Google Colab to load the data in Google Colab)
4. Define 3 most Imp Questions based by sampling the data manually/programmatically to evaluate what 3 questions your customer might ask in the next call. (**Hint:** Here the above case study is going to be super handy.)
5. Use Google Colab: Ipython Notebookto explore these by using pandas (**Hint:** <https://github.com/pandas-dev/pandas>) for data exploration in the above dataset. (discovery of anomalies in the dataset)
6. After identifying anomalies draw distribution graphs (**Hint:** <https://github.com/plotly/plotly.py> library to draw your charts) to demonstrate data challenges you have encountered
7. Export CSV in the same folder where you have Google Colab. (**Hint:**Only anomalies CSV(s) from the above Google Colab notebook)
8. Check against carrier website Anomalies (refer to carrier specific track & trace page e.g. Use Carrier name to Google the website for Track and Trace to witness what is wrong with these Anomalies(**Hint:** Using the Carrier Name, also called Carrier number or scac Google the track and trace website)
9. Create note what is wrong in these anomalies csv exported in a separate column in the CSV sheet to highlight your findings by cross checking against carrier websites that you just googled in the previous step. (**Hint:** Create an extra column for your findings as notes only for anomalies csv exports)
10. Write an email to us with a deliverable link Gdrive having all the above files as mentioned at the start of the assignment to address [oleg.kurochka@portcast.io](mailto:oleg.kurochka@portcast.io) and [ghassen.chniti@portcast.io](mailto:ghassen.chniti@portcast.io) **with summary email treating them as Nikana-San from Changi Express Logistic (CEL)**. Note email skills are important in how you craft your findings briefly nudging them to open your attachments. Remember customers in our industry get tons of CSVs in emails hence crafting email becomes important on a day to day basis.

### **Guidance Timeline to Complete:**

Submission **Tuesday 11:59pm Midnight, 20th March 2022(End of Day Singapore Time).**

**Don’t forget to set a new submission deadline if the above mentioned [20th March 2022] is not possible in a reply to the email when you receive the assignment. Customer communication of deadlines and sticking to them is key expectation on the role.**

## **Business Background**

### **What does Portcast Do?**

Portcast helps logistic players with google maps like predictions for estimated arrival time of cargo goods transported via ships, airlines, trucks and trains. ETA predictions is one of the product lines at Portcast and is the team you have applied for.

### **Why is ETA prediction important ?**

"Think of your anxiety to correctly predict and communicate to your friend your estimated time of arrival(ETA) when you are looking at the google maps when stuck on a red light"

### **Impact & Cost Savings?**

Now think of the entire world moving cargo through multi-modal transport, where every delay in one mode of transport has knock-on effects on other parts of the supply chain thereby creating billion dollar inefficiencies and lost man-hours in the system.

### **Why ML/AI use case?**

Humanly impossible to generate rules, patterns around delay reasons(think the live impact Google Maps brings vs human patterns of old taxi drivers remembering congestion free routes). Ground reality and the always on nature of cargo movement demands a superior AI based approach to this problem at hand

## **Visual Analogy: Mental Model**

[Download high resolution image at <https://drive.google.com/file/d/1BClxQMRnNqpCY9tZidXCH6GrMDEO2F87/view>]

### **Frontend Screenshot**

[Download high resolution design image at <https://drive.google.com/file/d/1Zb9v7MFoUkMJyFFlMwSGxHhIc84yZf2-/view?usp=sharing>]

## 

### **Contact: For any Queries/Clarifications [Please Only Whatsapp]**

Sagun Garg

+65 83000962 [Whatsapp]

1. Try <https://www.track-trace.com/container> instead to find out the right carrier [↑](#footnote-ref-0)