- This case study will go through a BLIND EVALUATION by the Director of Advanced Analytics and senior Data Scientists. **DO NOT PUT YOUR NAME ON ANY DOCUMENTS**. You will be assigned a number by the HR representative that you must put on your submission(s).
- For this case study you will be given a business question and a dataset and are asked to perform all steps in the analytical process necessary to answer the business question.
- Your submission should include 2 formats:
    1. A Jupyter Notebook or other format that showcases your code and analysis, that a data scientist will evaluate. The notebook should include markdowns and documentation that explain your decision-making.
    2. An executive style presentation (i.e., a presentation to executive level who understand the business but are limited in advanced analytics knowledge). This presentation should demonstrate your ability to use visualization to tell a story while limiting the need for extensive text.
- The customer data you will be receiving is real, but only a sample of the full list of customers. Customer names have been replaced with other identifiers.
- You should use Python and its associated libraries to manipulate the data.
- You are encouraged to go beyond what is explicitly required if it aligns with the overall purpose of the analysis.
- Your analysis and presentation should include:
    1. Highlight any irregularities in data and how you addressed them
    2. Outlines the objectives of the study in your own words
    3. Includes the results of any analysis you performed to answer the business question
    4. Optional: Include any next steps or recommendations you wish to offer

Submissions must be returned by the due date assigned to you by the HR representative. Email your submission directly to the HR Rep and they will ensure it is clear of any identifiers regarding the author before sending the documents to the Director.

Business Question:

Canada Post began a marketing campaign on September 1st, 2022 to increase customer parcel volumes. The campaign is ongoing. Only a subset of customers was targeted for the marketing campaign. The marketing campaign was only instituted in one city.

Your task is to determine if the marketing campaign has been effective thus far in increasing parcel volumes above expected volumes for the targeted customers in the targeted city.

Information related to the dataset:

The marketing campaign only took place in City #1. The marketing campaign is not expected to influence volumes in City #2.

Only customers #1-7, inclusive, were targeted by the marketing campaign.

In your Jupyter Notebook file you are expected to demonstrate:

Exploratory analysis and data cleaning, including any research steps that you have taken to determine a solution.

The use of 2 different time-series forecast algorithms (e.g., ARIMA, fbprophet, etc) to assess the overall effectiveness of the marketing campaign.

An assessment of the two algorithms to determine which better answers the business question.

Key Dates for your Consideration:

Initial COVID Impacts: April 1st, 2020 to June 30th, 2020

Additional COVID wave: April 1st, 2021 to June 15th, 2021

REMEMBER:

THIS IS A BLIND EVALUATION. DO NOT PUT YOUR NAME ON ANY DOCUMENTS.