**Objective:**

In this competition you will work with a challenging time-series dataset consisting of daily events of multiple companies. The ask is to analyze, handle data in CSV, an make predictions on events (which event will occur when) in next 12 months. This is a real and live case study and the data will be provided.

Scope Items

• Identify potential repeating events over the historical pattern  
• Be able to provide a percentage of accuracy for each event identified as repeating  
• Potential to predict “when it will occur”. This could be a day of week and week of month pattern or specific dates  
• Potential to predict the hotel it will book. (Do it if possible)  
• Forecast the potential size of the event (Do it if possible)

Data fields:

* Market
* Hotel Name
* Hotel Address (this would be for number 4 above)
* Event ID
* Event Dates
* Company/Account Name
* Event Name
* Largest Meeting Room used (Square Footage)
* Total Square Footage Used
* Estimated Attendees
* Chain Scale
* Brand Group
* Brand

Take care of following points:

1. Going forward, events with the Company/Account as “unknown” should be excluded from the analysis.
   1. These are events which are classified in our data as “SMERF – Social” by nature, they are generally unpredictable like a birthday party, unlikely to repeat versus those events associated with a defined “company”
2. Any dates that are tightly clustered and/or sequential are likely the same event (or sub events of a group) even though the event names may differ. Likely need to build logic around these events so we keep them grouped together and not considered multiple repeats in the same month.
3. Remove the Events if occurs only once.
4. Account / Company that has started a monthly meeting and has repeated for 15 months should be identified as a monthly repeating event even though the history is thin.
5. In expectation, companies which have returned to booking events would have the highest likelihood to be repeat business as they have resumed their “normal” booking pattern
6. Companies that have a strong booking pattern, but have not started to book events since March 2020 will be a lower likelihood to repeat, however, they may begin booking in the future as restrictions fade, so should not be omitted entirely
7. Any Company that does not have sufficient history in a given period should be omitted from being flagged as a “repeat”

AI-ML Model Output:

* + - 1. Number of events occurring in upcoming individual months.
      2. Name of the events along with month in which they occur. Also try to predict in week of month or particular date event will happen.
      3. Accuracy of the repeating events.
      4. Accuracy, Precision Score, F1 score of the model.
      5. Potential to predict the hotel it will book and Forecast the potential size of the event.