**Some information on specific columns:**

CLNDR\_DT = revenue transaction date

CLNDR\_DT\_MONTHSEQ = sequential month number that corresponds with CLNDR\_DT

Auth\_Date = The date from which a project is considered new. Because we do 12 month forecast, we look at this date +12 months as the date frame from where a project is new. For example if we are doing a forecast for January 2023- December 2023, if a project has an auth date of January 2023, we would not have historical data from <December 2022 so any revenue generated by this project would be considered ‘new’ for 2023.

Auth\_Date\_Monthseq = sequential month number that corresponds with Auth\_Date

**Question:**

How would you approach predicting what NEW PROJECT revenue will be for the next 12 months at the CATEGORY level. As noted above there will be no history for a new project. However, there are histories of projects that were considered ‘new’ from a particular point in time. The attached dataset was raw export with the data not filtered or cleaned up in anyway. Please review the data to be able to provide some specific details in answering the questions below.

* How would you think creating a ‘new project’ model? What type of approaches would you try?
* How would you transform this dataset so that it could be used for the approach you chose above? How would you filter data, would you create any new variables, etc? What would your test and training datasets look like? -🡪 **Please see the Jupyter Notebook**
* Please provide examples of what these inputs into the model would be for (variables: dependent, independent, static, dynamic):
  + Predicting from the beginning of 2023 (forecast for all new project revenue for full year (months 1-12).
  + Would your datasets look any different for predicting from July 2024 to July 2025?