* 在Excel文档中有2个Sheet：behavior和user。这两个表可以用列ID\_1和ID\_2来合并。
* 你的任务是预测‘behavior’表中的‘target’列，可以利用两个表中的其他信息来预测（除了ID\_1和ID\_2）。
  + 请编写代码来进行EDA（探索性数据分析），更好的了解数据情况。
  + 基于你对数据的理解，你认为应该如何建立一个准确、可靠、可理解的模型？请写出你的大纲，包含主要的建模步骤。你可以用中文、英文或psedo code来表述。
  + 请把你的代码和结果整理到易读的格式（如利用juypter notebook）。
* 如果有疑问、或需要更长时间来完成项目，请及时与我们联系。
* In the excel file you received, there are two working sheets: behavior and user. The two tables can be merged based on the columns ID\_1 and ID\_2 together.
* Imagine you are asked to build a process to predict the ‘target’ column in the sheet ‘behavior’ based on the other information in the two tables (excluding ID\_1 and ID\_2).
  + Please write code to conduct EDA (exploratory data analysis) to better understand the data.
  + Based on your understanding of the data, please outline the important steps you would take in building an accurate, robust, and interpretable model. Please feel free to choose Chinese, English, or pseudo code in describing your thoughts.
  + Please organize your code, result and writing in an easily readable fashion (jupyter notebook is preferred).
* Please do not hesitate to reach out if you have any questions or if you need to more time to finish the project.