

WEEKLY STATUS REPORT

REPORT DATE	TRACK NAME	PREPARED BY
Sep 12th, 2021	IntegrityM (Data cleaning and Data wrangling)	Yihang Zhao, Yuwen Luo, SeungHeon Han

STATUS SUMMARY

During the 1st week, we mainly read and understood the paper and documents clients provided us. Furthermore, we followed the instruction in the paper to clean the datasets and wrangle with data (group & aggregate). Then, we labeled each datasets using fraudulent labels within the LEIE dataset. We also explored the constructed data and prepared them for modeling. Clients will provide more documents in the following week, so we expect to accomplish more further.

PROJECT OVERVIEW

MAIN TASK	SUB TASK	% DONE	DUE DATE	ASSIGNED TO	NOTES
Load in and Clean datasets	Part B	100%	09/12/2021	SeungHeon Han	
	Part D	100%	09/12/2021	Yihang Zhao	
	DMEPOS	100%	09/12/2021	Yuwen Luo	
	LEIE			All of us	
Data Exploration	Part B	80%	09/12/2021	SeungHeon Han	
	Part D	80%	09/12/2021	Yihang Zhao	
	DMEPOS	80%	09/12/2021	Yuwen Luo	

RISK AND ISSUE HISTORY

ISSUE	ASSIGNED TO	DATE
-------	-------------	------

Imbalanced datasets	All of us	09/08/2021

CONCLUSIONS & OUTLOOK

We have cleaned and wrangled datasets this week and we are waiting for more documents from clients. After these steps, we realized the datasets are very imbalanced, so we need to solve this problem afterwards.