There are two questions in this assessment, and you are expected to answer both of them. The first question is on estimation, where you will be required to estimate a number of something (i.e. number of stars in the galaxy). The second question is to build a classification model. Good luck!

Q1: This is the estimation question. Please write down all the assumptions you are making, or any data sources that you use as reference. The main goal of this question is to understand your thinking process, and not the accuracy of the answer. The better you help me understand your thoughts, the better.

QUESTION: Please estimate the number of users that transact via TnG eWallet in any given month.

Q2: There are two csv files attached in this folder. These files contain the following:

ds_programming_transaction_table.csv

txn_id	account_id	use_case	pay_amt	paid_time
1000000001	A000007	apple_store	5000	11/10/2020
1000000002	A000010	retailer	2853	19/11/2020
1000000003	A000003	toll	95	12/12/2020
1000000004	A000010	retailer	1522	1/1/2021
1000000005	A000011	apple_store	5000	25/2/2021
1000000006	A000007	prepaid_mobile	2000	9/3/2021
1000000007	A000002	postpaid_mobile	4690	12/3/2021
1000000008	A000009	ecommerce	15800	9/4/2021
1000000009	A000008	ecommerce	1000	29/5/2021
1000000010	A000001	retailer	2375	3/8/2021
1000000011	A000001	retailer	3500	24/8/2021
1000000012	A000003	retailer	1520	27/8/2021
1000000013	A000012	toll	500	12/10/2021
1000000014	A000010	retailer	2807	18/10/2021
1000000015	A000008	ecommerce	16600	28/10/2021
1000000016	A000006	prepaid_mobile	1020	17/11/2021
1000000017	A000010	apple_store	2000	21/11/2021
1000000018	A000003	retailer	3070	22/12/2021

ds_programming_customer_target_table.csv

account_id	nric	latitude	longitude	is_active
A000001	780805415511	3.131604296	101.7516121	1
A000002	941114105797	3.143431913	101.7416554	0
A000003	561101025240	3.157315399	101.7510118	1
A000006	790623085224	3.136489721	101.7471487	0
A000007	410313238221	3.159029946	101.7132453	1
A000008	440404440000	3.177541534	101.7508412	1
A000009	900622210013	3.177972392	101.7426017	0
A000010	311222100067	3.197682918	101.7438038	1
A000011	310222100288	3.146901176	101.7369396	0
A000012	660622210002	3.150557	101.708989	0

Your task is to predict whether a user is **active**, based on the target column *is_active*. You are welcome to use any language of choice.

You can derive customer features from the transactional and customer tables attached (hint: # of transactions, gender, age).

Our expected output is a table containing 2 columns, `account_id` and `is_active` based on your model's prediction. Please submit your final code, comments, and results in a Microsoft Word file, along with any analysis or findings you come across during the development process.