Background:
Nations Info Corp is an online subscription provider of real estate and credit monitoring services. Our customers pay a monthly membership fee, after a usually $1 trial for several days. One of the challenges of a subscription business is to forecast the subscription revenue we collect from each customer (also known as LTV or CLV, lifetime value). On one hand, it's necessary to know the LTV because we need that value to know if the business is profitable. For instance, if the LTV is $100 but we are paying $120 to acquire each customer, that is clearly not a profitable business to pursue. On the other hand, it is very difficult to know what the LTV is when you first acquire the customer, since it may take many months for the revenue to come in, as we charge the customer each month they remain a member subscribed to our product.

Data:
We are using a cohort approach for data analysis. For this exercise, we will define the cohort as the group of customers that sign up in a given month (e.g. January 2020).

We collect cohort data from 2020 to 2022 for 2 different verticals of business (rto: rent-to-own and credco: credit monitor). We wanted to use various short term metrics (i.e. how they performed within 1 month from signup) to predict long term LTV. Data is provided in the "historical data.csv" file. Here is the data dictionary:

Numbers #:
M0#: number of signups we get. Serve as the denominator of all following metrics.

Dollar Amounts $:
LTV 0-15: average money we collect from customers in 15 days from signup.
LTV 0-30: average money we collect from customers in 30 days from signup.
LTV 0-360: average money we collect from customers in 360 days from signup.

Percentages %:
C0%: cancels on the same day of signup.
C1%: cancels during the trial period, before the first monthly bill.
D0%: failed to pay for signup due to card declines.
D1%: succeeded to stay the trial period but failed to pay for the first monthly bill due to card declines.
M1%: succeeded to stay the trial period and pay for the first monthly bill.
MOBILE%: signups using mobile device.
PREPAID%: signups using prepaid card.
LOGIN RATE%: login to account after initial singup
SEARCH RATE%: search for properties after initial signup (not available for credit business)
PDP VIEW RATE%: view the property details after initial signup (not available for credit business)

Task 1:
Our team is tasked with forecasting the LTV 0-360 using the given metrics and any external data. Marketing and product teams are interested in how data analysis and predictive modeling could help with their business decisions.

Please compile your Python / R codes and results in a .html file. Also feel free to use any business intelligence tools to present insights.

Task 2:
We are constantly testing new features on the sites and want to assess performance of the changes and optimize profitability of the overall business. We tested different price points of subscription fees on “rto” business recently, where our old version (Variant A) was put head to head in a test against the new version (Variant B). The visitor traffic is supposed to split evenly between A and B.

Analyze the test results and present findings, giving a recommendation about what we should do for the traffic that is being tested.

Variant A: $49 monthly subscription fee after 7 days trial.
Variant B: $39 monthly subscription fee after 7 days trial.

Data is provided in the “test result.csv” file. Here is the dictionary for additional metrics than the historical dataset:

VISITORS#: number of unique people who visit our website, before signup.
CPA$: cost we pay to partners on each signup.

Task 3 (SQL Question):
Table1: “orders” - the information of each order being placed

Column Name	**Type**
order_id	number
order_status	varchar
signup_type	varchar
order_datetime	timestamp
jluvr	varchar

Table 2: “activities” - the users activities on our website, whether before or after the order

Column Name	Type
id	number
action_type	varchar
user_data	varchar
created_at	timestamp
jluvr	varchar

Prompt:
“jluvr” is a key to define each individual user visiting our website, and can be used to link “orders” and “activities” tables. Note jluvr is not unique in either orders or activities table (i.e. same user can place multiple orders and can have multiple activities).
We wanted to find the last activity of users before each order being placed. The result will show each unique order, and the matched activity if it exists (if not exists, shows null and keeps the order info). Feel free to state proper assumptions if any are not clarified above.
Please submit the SQL codes in a plain text file / doc.

Sample Output:

order_id	jluvr	order_datetime	created_at	action_type	user_data
1001	abc-123-def	2023-12-01 08:00:00	2023-12-01 07:58:00	form_submit	rent
1002	abc-123-def	2023-12-01 09:00:00	2023-12-01 08:02:00	form_submit	own
1003	abc-123-efg	2023-12-01 10:00:00	2023-12-01 09:55:00	form_view	own
1004	abc-123-fgh	2023-12-01 10:01:00	(null)	(null)	(null)
…	…	…	…	…	…