1) 'What' information would you derive from it, and 'how' will you derive it (you can give SQL queries, pseudo code, ... whatever you're comfortable with)

- **Step1: Create dummy data sets : User table** and **Purchase table** by leveraging Python faker package and random function, for simulation purpose:
    - **User table (name of table: df_user)**
        - **id:** primary id of user, assuming total 997 unique users
        - **name**: full name of user (created from faker package, not real names)
        - **referring_user_id:** id of user who refer the user to the platform
            - Assume 300 of the 997 users have signed up the platform through the referral program posted by 100 other users
        - **Example of first 10 records of User table:**

id	name	referring_user_id
0	Gloria Lee	891
1	Karen Christensen	<NA>
2	Nicole Lopez	<NA>
3	Rodney Mccarthy	<NA>
4	Heather Brooks	<NA>
5	Andrew Sandoval	<NA>
6	Margaret Crane	<NA>
7	Heather Solis	<NA>
8	William Warner	975
9	Terry Armstrong	720

    - **Purchase table (name of table: df_purchase)**
        - **id:** primary id of transactions
        - **user_id:** id of each user, forigen id
        - **date :** date when the translation occurred, assuming all in month of Nov'20
        - **total:** total $ amount of the transaction (range from $0.0 ~ $50.0)
        - **Example of first 10 records of Purchase table:**

id	user_id	date	total
0	325	2020-11-01	17.319093
1	372	2020-11-01	16.678905
2	973	2020-11-01	26.341131
3	689	2020-11-01	25.594784
4	809	2020-11-01	14.450100
5	370	2020-11-01	4.158096
6	297	2020-11-01	40.096005
7	881	2020-11-01	17.395389
8	429	2020-11-01	16.761624
9	713	2020-11-01	7.224745

- **Step2: Derive a new table Purchase with Discount** by adding a discount column, following the logics below:
    - **If** a user signed up the platform through the referral program **AND** the user has posted the referral group program, the user receives $20 + $10 = $30 discount
    - **If** a user signed up the platform through the referral program **AND** the user has not posted the referral group program, the user receives $20 + $0 = $20 discount
    - **If** a user did not sign up the platform through the referral program **AND** the user has posted the referral group program, the user receives $0+ $10 = $10 discount
    - **If** a user **Neither** signed up the platform through the referral program **NOR** posted the referral group program, the user receives $0+ $0= $0 discount
    - **SQL query to create the Derived Table**

```
query = """
SELECT
df_purchase.id AS id,
user_id,
referring_user_id,
date,
total,
CASE
    WHEN
        ( referring_user_id IS NOT NULL ) AND ( df_user.id IN (
            SELECT DISTINCT referring_user_id
            FROM df_user
            WHERE referring_user_id IS NOT NULL
        )) THEN 30

    WHEN
        ( referring_user_id IS NOT NULL ) AND ( df_user.id NOT IN (
            SELECT DISTINCT referring_user_id
            FROM df_user
            WHERE referring_user_id IS NOT NULL
        )) THEN 20
    WHEN
        ( referring_user_id IS NULL ) AND ( df_user.id IN (
            SELECT DISTINCT referring_user_id
            FROM df_user
            WHERE referring_user_id IS NOT NULL
        )) THEN 10
    ELSE 0
END AS discount
FROM df_purchase LEFT JOIN df_user ON df_user.id = df_purchase.user_id

"""
```

    - **Example of first 10 records of Purchase with Discount table:**

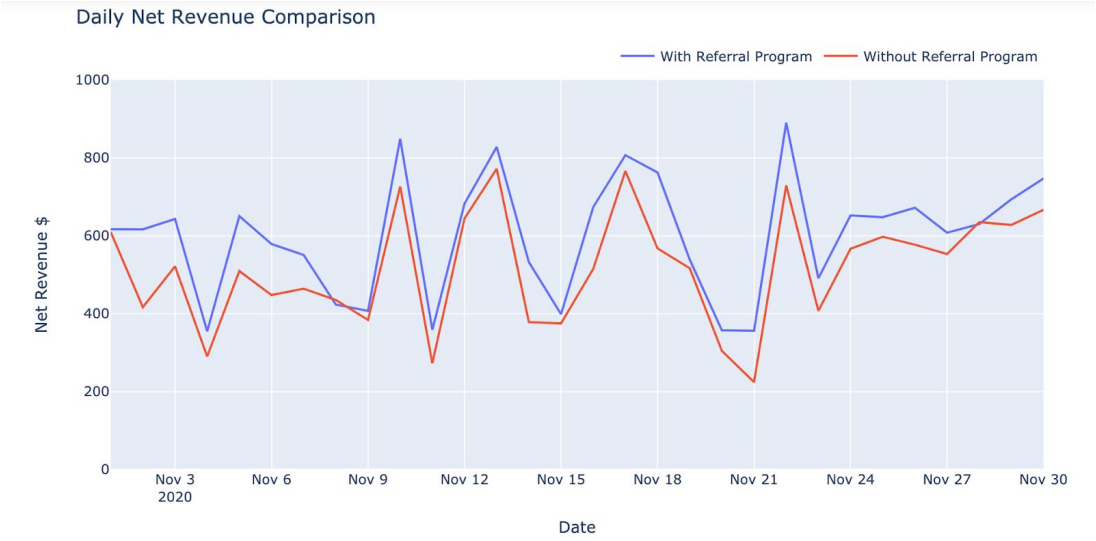
id	user_id	referring_user_id	date	total	discount
0	325	<NA>	2020-11-01	17.32	0
1	372	<NA>	2020-11-01	16.68	0
2	973	<NA>	2020-11-01	26.34	0
3	689	<NA>	2020-11-01	25.59	0
4	809	<NA>	2020-11-01	14.45	0
5	370	<NA>	2020-11-01	4.16	10
6	297	<NA>	2020-11-01	40.10	0
7	881	544	2020-11-01	17.40	20
8	429	57	2020-11-01	16.76	20
9	713	259	2020-11-01	7.22	30

2) Using the information from Step 1), how would you make a recommendation on whether this rewards program should be continued or discontinued

First of all, we look at **Net $ Revenue** for the **two scenarios:**

- **With the referral group**; using the complete set of data
- **Without the referral group**; using the sunset of the data which only includes the users that did not participate the referral program; in another word, neither posted the referral program, nor got referred to the platform by other users.

As shown in "**Daily Net Revenue Comparison**", the daily net revenue with the referral program are general high than the net revenue without the referral program



The conclusion is more clear by plotting the both lines in **Daily Running Total.** If we look at the daily running total of net revenue with referral program vs. without referral program, we can clearly see that overall the net revenue with referral program is increasing faster than the scenario without program.

By end of Nov'20, the net revenue with the referral program is $18K vs. $15K without the program. So the recommendation is to continue the referral program.