Data Challenge Senior Data Analyst

About OptioPay:

We are a Payout enabler – when a company wants to send money to their customer, we step in and we receive requests to pay out money in the form of *Payment Jobs*. These users then go to our platform where they have the opportunities to convert these Payment Jobs into one or more *Transfer Jobs*, each of which can have one and only one *Transfer Method*, which can be either a Payout Method (Bank Transfer), or a *Campaign* (Voucher). Our goal is to convert as many of those Payment Jobs into Vouchers as we can, and to make the customer happy in the process, by giving them an *Upside*.

Entities: Payment Jobs can spawn one or more Transfer Jobs when the status changes to Closed. Each Transfer Job has a Transfer Method, either a Campaign or Payout Method, depending on whether the person opted for a Voucher or Payout. A successful Voucher delivery means that the Status is Closed and a Payout Method ID does not exist. The Catalog table shows the journey a user has taken through the platform, from the first click to the last. It shows whether the user looked at Payout Methods or Transfer Methods, and which ones.

NB: If a Payment Job is not acted upon and it reaches a certain date, the status automatically changes to Fallback and it triggers a Transfer Job for sending the money to the Recipient's Bank Account.

State Transitions:

Payment Job: Start -> New -> Open -> Closed (or Cancelled or Fallback)
Transfer Job: NEW -> OPEN -> CLOSED (or CANCELED)

Dataset:

Connect to this Postgres instance:
Hostname: http://optiopay-data-challenge.cawbir4qpzsc.eu-central-1.rds.amazonaws.com
Port: 5432
Database: optiopay
User: *sent separately*

1. Walk us through a Cohort analysis of the data from the included tables. How would you identify a cohort and what would you look for?

2. The Product Team is brainstorming ways to increase the Conversion Rate, I.e. the number of Payment Jobs that convert to at least one Voucher. Analyze the data you are given and find some insights and suggestions, and present those to a Product Owner, and explain how you would convince them of any course of action that you can think of.

3. There are conflicting opinions on the matter of whether a long or a short user journey leads to better conversion rate. How would you analyze the data to find out whether it has a positive or negative effect?

Please send the results as reproducible SQL/Python/R scripts or Jupyter notebook.