

## SUBMISSION INSTRUCTIONS

Please submit through the greenhouse link.

## SQL Exercise

For a planned CRM campaign, FanDuel Sportsbook is going to send an email to existing players letting them know the available markets and new site features for the upcoming NFL season. In addition there will be targeted offers to certain players based on their transactional history. As an analyst your job is to help define the segmentation and then use the SQL database to generate lists of users. The data tables you have at your disposal are below:

**USER TABLE** – This is where we store information about user’s account (one row per customer)

FIELD	DESCRIPTION
playerid	Unique identifier for customer
alias	Username of customer on site
email	Email address of customer
registration_date	Date of customer’s registration
activation_date	Date of customer’s first activity with their own money (i.e. after a deposit)
account_balance	Snapshot of latest customer balance

**ACTIVITY TABLE** – This is where we store information about user’s activity

FIELD	DESCRIPTION
playerid	Unique identifier for customer
state	“NJ” or “PA”
betplaceddate	Date of bet placement
betsettledate	Date of bet settlement
betid	Unique identifier of bet
legid	Unique identifier of leg (might be more than one per bet)
bettype	Bet type: can be Straight (1 leg) or Parlay (2+ legs)
stake	Amount wagered per leg
winning	Amount won per leg
ggr	Gross Revenue. Formula: Stake – Winning (+ if customer lost, - if customer won)
sportname	Sport selected per leg (Football, Basketball, etc.)

**DEPOSIT TABLE** – This is where all the information about every deposit

FIELD	DESCRIPTION
playerid	Unique identifier for customer
cartid	Unique identifier for deposit
payment_status	Status of payment: “a” if approved, “d” if declined
payment_amount	Amount deposited
cartdateid	Date of deposit
sb_flag	Flag to identify Sportsbook deposits: 1 if SB, 0 if non-SB

- 1) Based upon the information you have in the database what variables would you consider most when designing a segmented offer for users in the new football season and why do you think that they are important? What sort of groups would you like to create and how might you vary the incentives/offer to each of those groups?
- 2) The CRM Manager has an idea for a test, but wants to know if there are enough users available to make the results significant, so asks you to work out the count of users who activated in 2019 (note: a customer is considered activated when he goes through at least one successful deposit and places a bet). How would you do this with SQL code? Are there any other constraints you might consider adding to give a more accurate estimate?
- 3) The CRM Manager is happy with the base user size so it's time to generate the list of users to be sent the email. With the same constraints as above you need to generate a list with the following fields, so that the CRM Manager can start to build his custom segments:
  - i. playerid
  - ii. alias
  - iii. email
  - iv. total stake generated in 2019
  - v. margin in 2019 (i.e. gross revenue/stake)
  - vi. last bet placed date
  - vii. total amount deposited on Sportsbook (approved deposit only)

How would you create this list in SQL?

- 4) The Marketing VP is concerned that there has been a decrease in Parlay bets (i.e. bets with more than 1 leg) as a % of total bet count. Specifically he wants to know in each month from January 2019 to April 2019 what percentage of Parlay bets we had. Can you provide the SQL code required to answer his questions?
- 5) The Acquisition Manager is concerned that there has been a decrease in Registration to Activation rate (i.e. many customers registered but they never made a deposit). Specifically he wants to know what Reg-to-Act ratio we have had in each month from January 2019 to April 2019. Can you provide the SQL code required to answer his questions?