Coinbase Data Science Intern Take Home Exercise

Background

Before they are able to trade, new Coinbase users need to go through a multi-step onboarding process which includes email, phone, and ID verification. The Global Onboarding team at Coinbase owns the end-to-end onboarding experience with a goal of driving new traders while maintaining guardrails such as user fraud.

As a part of planning, the team has asked you to perform a deep-dive analysis into the current state of onboarding and help identify opportunities for improvement in the flow. The findings from this analysis will act as a key input into the team's roadmap and prioritization over the next year.

Instructions

Use the attached data to perform a basic analysis into the onboarding flow and share your analysis as you would with another Data Scientist in an Rmarkdown or Jupyter notebook. Be sure to include any data exploration, processing, and visualizations you performed as a part of this analysis. At the end of this exercise you should be equipped to make a few recommendations to business stakeholders on next steps.

Your submission will be evaluated on both **technical acumen** and **communication quality**. You have full freedom to use whatever analytical tools and techniques you feel is most appropriate for this exercise.

Some guiding questions to get you started:

- Where in the onboarding process are users currently churning from the onboarding flow?
- What are the biggest reasons for users to churn?
- Does this vary over time and for different user segments?
- What product recommendations would you make based on your analysis?

Please submit notebook output in PDF form. Do not spend more than 5-6 hours on this assignment.

Data dictionary

users.csv
- **user_id:** unique id for Coinbase users
- **country:** country code where users first signed up
- **platform:** indicator for whether user used web, ios, or android to signup
- **signed_up_at:** timestamp in UTC when the user signed up
- **phone_verified_at:** timestamp in UTC when the user verified their phone number
- **email_verified_at:** timestamp in UTC when the user verified their email address

id_verification.csv
- **id:** unique id for the id verification attempt
- **user_id:** unique id for Coinbase users
- **created_at:** timestamp in UTC when the user attempted id verification
- **status:** status of this id verification attempt
- **failure_reason:** reason why the id verification attempt failed

trades.csv
- **id:** unique identifier for initiated trade attempt
- **user_id:** unique id for Coinbase user
- **created_at:** timestamp in UTC when the trade was initiated
- **volume_usd:** estimated trade volume in USD