Candidate Exercise
Senior Actuarial Analyst, Actuarial Data

Background:

- A partner team has reached out to us for support as they consider a reinsurance agreement. The proposal under consideration is for Uber to purchase reinsurance for the 250k - 1.5M layer of a certain segment of the Rides book. This segment comprises multiple policies:
    - Policy 123: covers Period 3 losses up to a 1M limit per occurrence, subject to a deductible
    - Policy 456: covers Period 2 & 3 losses 1M xs 1M per occurrence
- Our partner team needs us to calculate incurred loss costs and claim frequency for the ceded layer at various maturities.
- Period definitions:
    - P0 = Offline or the Driver app is off
    - P1 = Available or waiting for a ride request
    - P2 = En route to pick up riders / deliveries
    - P3 = On trip (ie. pickup is complete and en route to dropoff)

Questions

- You are given a dataset containing historical claims data for Policy 123 and a separate dataset containing historical claims data for Policy 456. Please take a few minutes to familiarize yourself with that data. What clarifying questions would you ask the team after your review?
- You are given a dataset containing trip information and driver information. This table will serve as your exposure data. Are there any clarifying questions you need to ask about the exposure data? What will you consider as you select an appropriate exposure base?
- Provide an example of a SQL query or Python script that will fulfill our partner's request.
- Describe what you would do with the output of your query in order to answer your partner's question.
- Feel free to make assumptions as necessary, just make sure you state them clearly!

Data

- Dataset 1: loss run containing data for policies 123
- Dataset 2: loss run containing data for policy 456
- Dataset 3: exposure data