Data Scientist Technical Assessment

Please complete the following tasks using Files 1-7.

1) Combine the seven files into a single file that can be uploaded as a SQL table.

For the next section, write SQL queries for each question that will return the requested results for the full time period covered in Files 1-7.
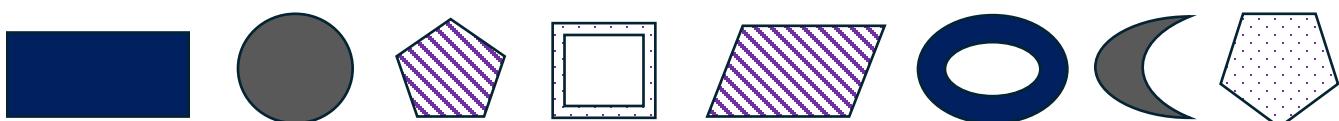
2) How many unique users are there?
3) Who are the top 10 users who took calls? How many calls did each take?
4) What is the average total handle time of a call?
5) Who are the top 10 users with the highest ratio of disconnect type 'peer' to total calls taken? What is the ratio for each of these users? How does this differ from disconnects 'client' and 'peer'? What can you infer from this?
6) How many inbound queue calls were received per month?
7) What is the percentage difference of inbound queue calls taken in 2023 vs 2024?
8) What is the month and year with the most inbound queue calls?

For the next section, utilize python to model the forecasted calls per month in 2025 then answer each question.

9) How many calls were taken per month during 2023 and 2024?
10) What is the forecasted call count per month in 2025 and 2026?
11) What model was used to build the forecast?
12) Why did you choose this model?
13) Paste the visualization of the actual data and forecast.

Please answer the following questions.

14) Is there a correlation between hold time and handle time?
15) How is talk time affected by hold time?
16) Is there a correlation between survey scores and disconnect type?
17) Is there a correlation between hold time and disconnect type? If so, please explain it.
18) Group the shapes below based on their characteristics. List all possible groupings.