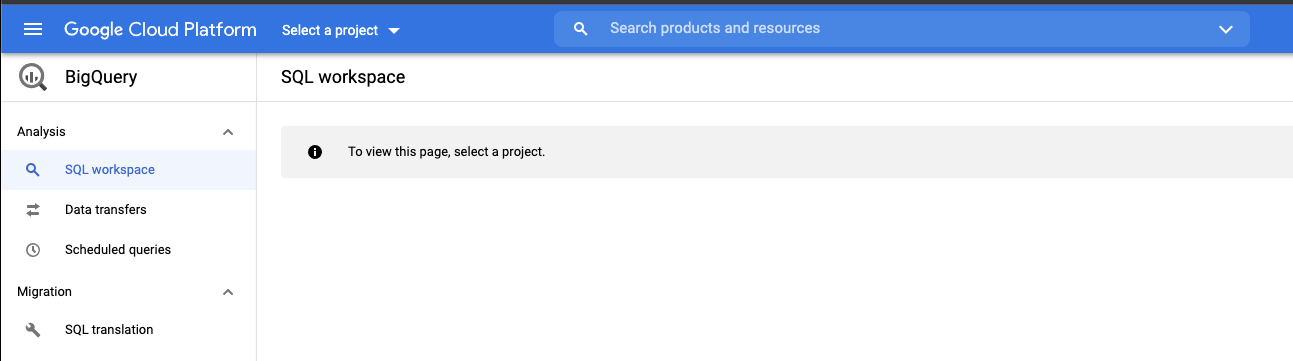
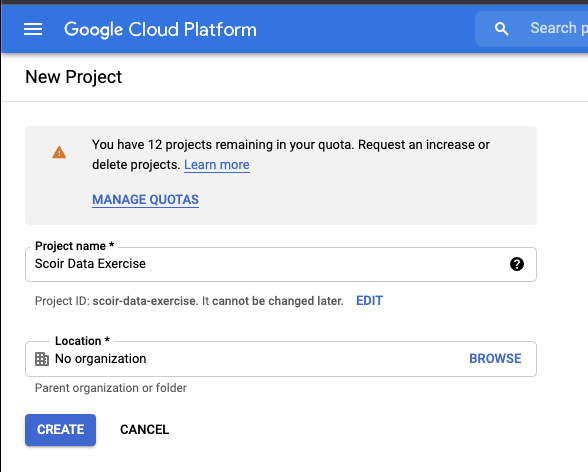
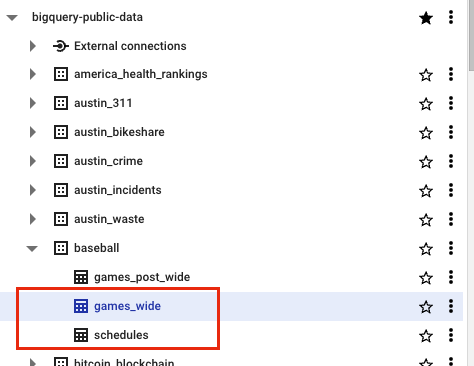
**Scoir Data Analyst Assessment**

1. Open [Google BigQuery](https://console.cloud.google.com/bigquery)
2. Create a new project called “Scoir Data Exercise” by clicking “Select a Project” on the top left of the screen and then select “New Project”

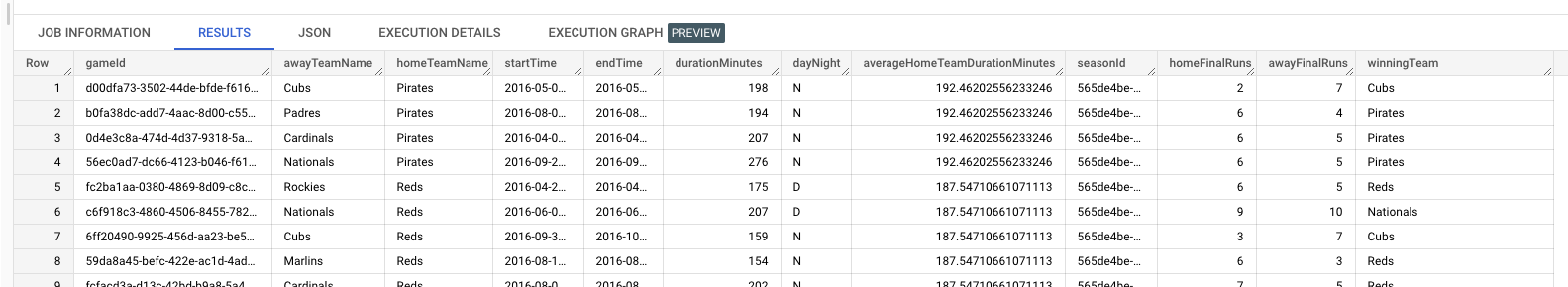


1. After creating the new project, you should see “bigquery-public-data” in the “Explorer” section of BigQuery. Open games\_wide and schedules under baseball
   1. Schedules provides some high level data on the games (one row per game)
   2. Games\_wide provides some high level data on the games as well as more detailed data on pitches (many rows for one game)



1. Write a query joining the 2 tables to create one table with gameId, awayTeamName, homeTeamName, startTime, duration\_minutes as durationMinutes, dayNight, seasonId, homeFinalRuns, awayFinalRuns,

and the following additional calculated columns. Please use at least one CTE. You can copy and paste your code into this document OR (bonus) check your sql file into a personal git repo and share the repo here. Header rows for the desired result are below:



1. endTime
2. winningTeam - calculated from homeFinalRuns and awayFinalRuns, null if there is no definitive winner
3. Home team average duration in minutes - averageHomeTeamDurationMinutes

5. Which team won the most games?

1. Identify the gameId for their longest game
2. Write a summary paragraph with at least 3 details about this game using games\_wide (pitcher details/stats/, day/time/location, anything interesting about the outcomeDescription frequency
   1. Helpful info:
      1. You can assume that each row is a pitch with details about the pitcher, pitch speed, batter, and pitch outcome. To make this true, exclude rows where inningEventType = 'LINEUP'
      2. The home team is pitching and the away team is at bat when inningHalf = 'TOP'. The away team is pitching and the home team is at bat when inningHalf = 'BOT'

Note: I don’t know much about baseball so I do not expect you to either. I am only interested in the information you can pull out of the data and how you deliver it.