

# **Optimizing SQL for Analytics in the Cloud Kick-off**

# Welcome and Introductions

- Instructor Introduction
- Student Introductions
  - Name
  - Bootcamp and Cohort
  - What do you like to do outside of computers/coding/data?

# Course Logistics - Class Schedule and Attendance Policy

- Days/Time: Tuesday and Thursday, 6:00 - 9:00 PM
- Dates: 10/7/2025 - 12/4/2025
  - Note: We will take Thanksgiving week off (11/24 - 11/28)
  - There will be optional office hours Tuesday, 11/25, during normal class time (attendance on this date is optional and not recorded)
- You must attend at least 80% of class sessions
  - (Maximum 3 absences)
  - Notify instructor via Slack of any absences, or if you'll be late to class

# Class Expectations & Philosophy

- The role of AI in this class
  - Do not depend on AI for your answers.
  - The goal is to become a better analyst by gaining comfort and fluency writing advanced SQL. Relying on AI cheats you of this opportunity.
  - Rule of Thumb: Use AI to *support* your learning, not *replace* it
- Collaborate, don't compete
  - You will work with other analysts and data experts in your job; collaboration is expected.
  - Work together with you fellow students. Teaching a concept is the best way to master it yourself.

# Class Expectations & Philosophy

- Participation
  - You are expected to participate by walking through answers to exercise questions, asking questions, etc
  - At work, it feels better (especially early on) to decide when you're going to talk about what you've been working on vs being called on unexpectedly; practice choosing your moment in class

# Understanding the Data

- Dataset: The CMS Synthetic Public Use Files
  - Synthetic dataset based on real-world medical claims data from the Centers for Medicare & Medicaid Services (CMS)
- Key resources:
  - [CMS page about the data](#): for background on the data source
  - [OMOP CDM GitHub.io](#): information about table schema in the OMOP common data model
- Short “About” Document: pinned in the Slack channel
- Note: our database has some slight differences from the standard schema to support some learning concepts that will come up over the coming weeks
  - You can ignore the drug\_exposures\_summary table and columns with the suffix “no\_idx” until directed