# Data Science (Product) Take Home Exercise

At Clover Health, we pride ourselves on considering candidates with non-traditional resumes. To make sure that you can showcase your full range, we'd like you to complete this take home exercise. We know your time is valuable, but we've given you this exercise for the following reasons:

**It gives you the opportunity to learn about us.** Did we choose a thoughtful and challenging exercise? Did we give you the opportunity to get excited about ways you could improve a complicated system with analytics? Did you feel like you were able to show yourself off as "more than a resume"? If so, you'll be more excited about working for us. (If not, we may not be right for you - and we don't want to waste your time.)

**It gives us the opportunity to learn about you.** Were you able to clearly communicate your ideas? Were you able to find a good balance between completeness and time to ship? Can what you did be leveraged by the rest of your team? This is your time to shine, so show us what we can't see on your resume.

**Our expectations:** We don't expect you to spend all weekend on this (an afternoon, maximum). We don't expect you to deliver production-ready models. We know you have limited information; it's okay to make some assumptions and base your work off those. (But if you have questions, don't hesitate to ask!) We look forward to getting a glimpse into your thinking process and creativity!

Complete the exercise below in whatever format you prefer. For some loose requirements, imagine yourself as a member of our team. This means that:

- Other members of the team will need to read your code.

- They will expect to be able to reproduce your results.

- They may want to reuse some of your analyses.

- They (or you) might need to do something similar in the near future.

- They will certainly want to make use of the same data.

- They don't have a license for paid statistical software like SAS, SPSS or Stata. Please use something free like Python, R or SQL.

The intention of this assignment is to learn more about candidates as part of Clover Health's interview process, not to acquire confidential information. Contact Otis at otis.anderson@cloverhealth.com if you have any questions.

## Assignment

You work for an insurance company that is contemplating adding several new doctors to their coverage network. Before the company reaches out to the doctors, they'd like to have some idea of the doctors' clinical effectiveness.

You have obtained a sample of data that includes a particular treatment for a chronic condition that is common in your membership population. You have also tracked down whether or not the patient had a particular type of health failure (recorded by a follow-up visit).

1) If you were to only consider the provider's effectiveness with regard to this particular treatment, would you recommend reaching out to any of these doctors in particular? If so, which ones, and why?

2) What other data would be helpful in understanding which doctors to reach out to? What other data would help you evaluate the overall clinical effectiveness of the doctor? How would you use this data?

Please email us your answers, write up, and code attachments. Feel free to include any visualizations you think may be helpful.

Data Dictionary

| | Column | Description |
|---|---|---|
| 1 | event_id | unique id for each treatment |
| 2 | servicing_provider_id | unique id for physician administering treatment |
| 3 | servicing_provider_name | name of physician administering treatment |
| 4 | treatment_date | date of treatment |
| 5 | member_id | unique id for each member of plan |
| 6 | member_age | age of member |
| 7 | member_sex | member sex; 0 = female ,1 = male |
| 8 | health_risk_assesment | a variable scaled 1 to 10 that quantifies overall member health risk, using demographic (including age and sex) and health history surveys |
| 9 | outcome | failure = treatment failed to control condition as of follow up visit |