1. Why did you choose your prediction methodology?

I am most comfortable applying statistical analysis to recorded data. In addition, I believe that curve fitting is less resource-intensive than most ML methodologies, while potentially providing an easy-to-understand modeling curve for future analysis.

1. What are the strengths and weakness of your approach?

For statistical behaviors that can be reduced to a simple curve (e.g. a Gaussian, exponential growth curve, etc.), an easily reproducible formula can be fitted and applied for extrapolation. However, more complex behaviors are more difficult to model, and extrapolation beyond the available data has the potential to give wildly inaccurate results.

1. Important caveats or assumptions?

* Big Assumption: The data provided will be reasonably fitted by the chosen curves (square root for % screening visits/total visits and 6th-degree polynomial for total visits over time).
* Assumption: In the interests of time, weekend data can be omitted without negatively affecting the fit (ideally, some “weekend factor” could be calculated to allow those data points to contribute to calculating a smooth curve, if a statistical method is still used).

1. What did you provide as your likelihood metric and why?

I did not have enough time to construct an adequate likelihood metric. Given more time, I would aggregate the errors from the fitted curves to generate a band of maximum and minimum screenings over the month of July, then calculate the odds of each given day reaching the given threshold value.

1. How could you improve the quality of your prediction?

If I had visiting data from previous years, I could create a more accurate curve of the expected clinical visits over the course of the year, leading to a more insightful formula for fitting the data & extrapolating onward. Data on screening rates during other health emergencies in the U.S. (or other countries) could have a similar use in devising a more accurate fitting curve with which to model our projections.

1. What questions would you ask the stakeholders if you were tasked with this problem?

What is your risk appetite for “Probability that COVID-19 screenings exceed the screening threshold”? E.g. “If likelihood is 51% or higher, report that we expect to cross the threshold.”

By “any given day” do you mean “return a likelihood for each day” or “return a likelihood that at least 1 day in the month will be over the threshold”?

1. Given the output of your prediction, what recommendation would you make to the stakeholders?

With the work currently presented, I do not currently possess enough information to make a detailed recommendation to the stakeholders. What I would recommend that the stakeholders look to increase their screening threshold for most weekdays in July. The number of screenings is trending upward, albeit at a less steep rate than between April and May, and weekends are normally low-volume. There could be a reduction in cases by the end of the month, so I would recommend that most extra resources be allocated earlier in the month.