

The first five minutes of the Zoom interview will consist of you presenting data visualizations to address TWO of the following questions.

The dataset you have been provided is a simulated dataset based a pattern of attendance typical of a large program.

The main purpose of this part of the interview is to hear you talk about data. It is more about how you explain your exploration based on the questions than it is about the visualizations themselves. You should use whatever tools you are most comfortable with (Python, R, Excel, Tableau, etc.) to visualize the data.

You will be sharing your screen for this part of the interview, so you can use any method that can be shared that way.

If you have questions on any of this, please let me know.

- Greg

Exploratory Questions:

1. Is there anything interesting in the overall pattern of attendance throughout the year for the agency as a whole? Do any patterns jump out?
2. What does the distribution of "attend\_description" look like? Does this vary from location (agency-site or agency-site-classroom level aggregations) to location?
3. Are there any interesting patterns in the reasons that children are absent? Do these vary by location?
4. Head Start is very concerned with making sure that children attend on a regular basis. What is the rate absence rate (to keep it simple, assume you are comparing "absent" to all other values of attend\_description in calculating this) over time for this agency? Does that rate vary by location?