Lab 4 Activity

Let's look at the `attitude` dataset, which we used in the activity for Lab 2. This dataset is always loaded into R.

1. Run `help("attitude")` to get some info on the data and the meaning of each variable. You can also run `View(attitude)` to open the data in the data viewer window. Then, create a dataset called `dat` that only includes the `rating`, `privileges`, and `complaints` variables.

2 Use the `ggpairs()` function from `GGally` to visualize the relation and distribution of the 3 variables in the the `dat` object. Do these variables seem linearly related? If so, do they seem positively or negatively related?

4. Run two individual regressions, with one regression having the variable `privileges` predicting `rating` and the other having the `complaints` variable predicting `rating`. What are the values of the two regression slopes? Do they match your expectations? Are the values of the slopes significant?

5. Run a multiple linear regression where both `privileges` and `complaints` predict `rating`. What are the values of the slopes now? Is there anything that you find surprising?

6. Visualize your regression model with an interactive 3D plot that includes a regression plane. Note that this plot will not be visible in a PDF once you knit (why? PDFs are not interactive). If you are wondering why the slope of `privileges` is now negative, looking at the plot from a specific angle may give you some insight.

- What are some of the points with the highest residuals? Try finding them in the interactive plot by hovering over the dots (HINT: if the residuals are large, they should be far away from the plane, and the sign indicates whether they will be above or below the plane!)