431 Deliverable A Sketch

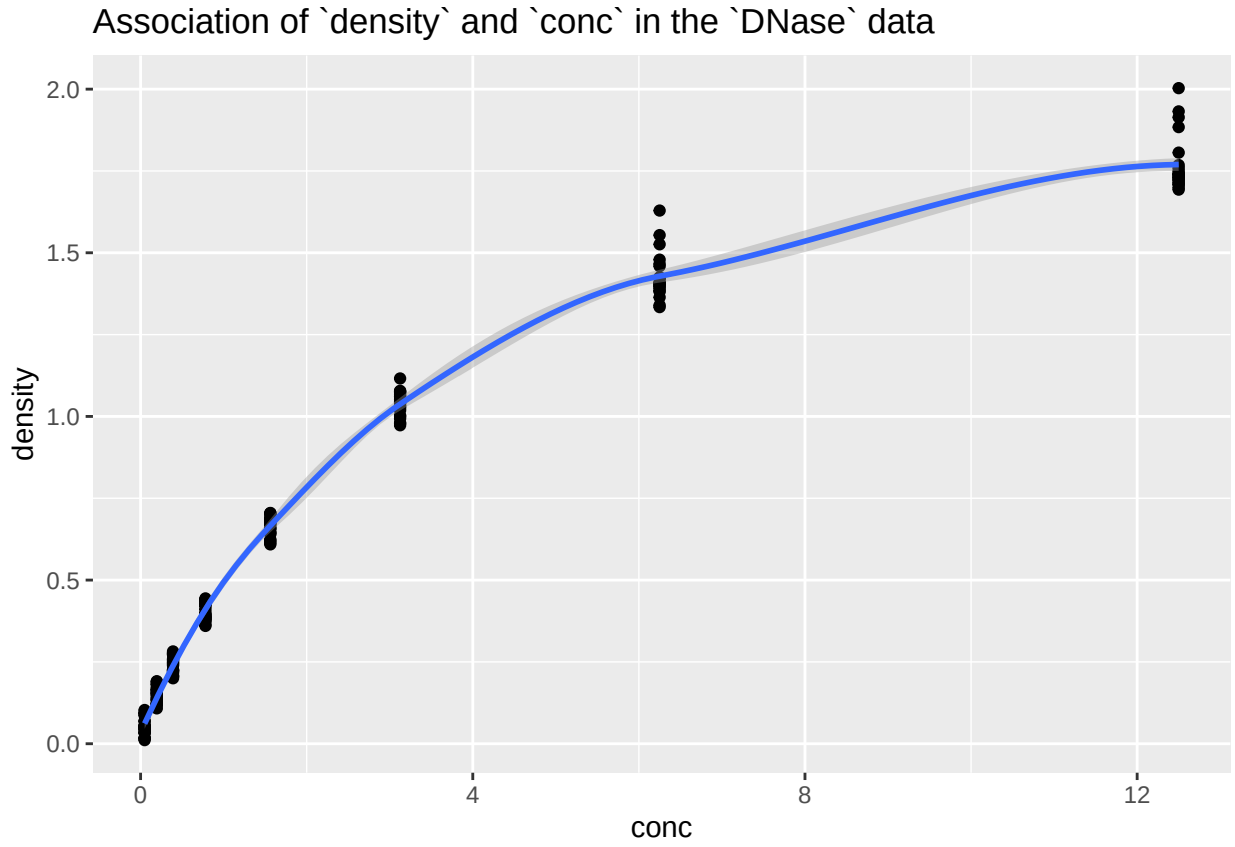
*Due **2018-08-30** at 2 PM. Last Edited 2019-08-29 21:55:41*

Activity 1. Interpreting a Visualization Built in R

Professor Love used R and the `tidyverse` to build the plot below using the `DNase` data set from the `datasets` package automatically loaded by R. Here's the plot again, and the code I used to build it.

```
ggplot(DNase, aes(x = conc, y = density)) +
    geom_point() +
    geom_smooth() +
    labs(title = "Association of `density` and `conc` in the `DNase` data")
```

```
`geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Use the Help window in R to learn about the `DNase` data set, and in particular, about the two variables displayed in the plot and their scientific context. Then write a paragraph (no more than 100 words) which explains what the plot indicates about the relationship between the two variables, and (more generally) what you have learned about the data (or science) from the plot.

The `DNase` data set: Help file

- *Description*: The `DNase` data frame has 176 rows and 3 columns of data obtained during development of an ELISA assay for the recombinant protein DNase in rat serum.

- The variable `conc` is a numeric vector giving the known concentration of the protein.
- The variable `density` is a numeric vector giving the measured optical density (dimensionless) in the assay. Duplicate optical density measurements were obtained.

What Were We Looking For?

We don't write answer sketches for essay questions, and that's sort of what this is. We'll likely share some nice materials written by students in the class (anonymously) in class soon, but we can tell you what we were hoping to see.

1. We want you to write in complete, grammatically correct English sentences. We want you to make your points as clearly as possible, in your own words, not, for instance, just copy-and-pasting what's in the help file.
2. We want you to accurately describe what the graph indicates about the relationship between `conc` and `density` as shown by the data, specifically that higher concentrations of the DNase protein are associated with higher values of measured optical density in the assay.
3. We wanted you also to describe the shape of the relationship, specifically that indicate that the relationship appeared somewhat non-linear. It appears that the impact of changing the `conc` level is a bit more substantial on `density` at lower `conc` levels than at higher levels.
4. We wanted you to remark on the nature of the experiment, that several `density` measures were taken at each `conc` level, and perhaps to suggest that the blue smooth curve follows fairly closely to the average of those `density` measures at each observed `conc` level. This explains why the points in the plot fall in vertical lines at certain `conc` levels, and do not appear at other levels - that is the design of this study.

The TAs will provide a few comments, centered around these ideas, in reaction to your paragraph. We hope this is helpful to you, as you think through future work.

Activity 2. Completing a Survey - Google Form

No real need here for an answer sketch. We're not looking for particular answers, just trying to understand where your attitudes are at the start of the class.