36-315: Statistical Graphics & Visualization

- course objectives
  - learn useful principles for making appropriate statistical graphics.
  - critique existing graphs and remake better ones.
  - visualize statistical analyses to facilitate communication.
  - pinpoint the statistical claims you can/cannot make from graphics.
  - write and speak publicly about statistical graphics.
  - practice tidy data manipulation in R using the tidyverse
  - practice reproducible workflows with Quarto
- grammar of graphics defined and used in ggplot2
  - see `01lec.pdf`
- goal of data visualization: show data, communicate a story
  - induce viewer to think about substance, not graphical methodology
  - make large, complex datasets more coherent
  - encourage comparison of different pieces of data
  - describe, explore, and identify relationships
  - avoid data distortion and data decoration
  - use consistent graph design
  - avoid graphs that lead to misleading conclusions!
- data types
  - quantitative
    - discrete
    - continuous
  - categorical (`factor`)
    - nominal
      - no order
      - e.g. race, species
    - ordinal
      - ordered!
      - ranking
      - DEFAULT IN R! manually define `factor` levels, or alpha. default
- area plots
  - pie chart (BAD!!!)
  - bar chart
  - stacked bar/spine chart (for variable comparison)
  - waffle charts????
  - rose diagrams (temporal or directional context can justify usage)