

# ggplot – boxplots, more density

Professor Halterman

Michigan State University

PLS 397 Analyzing and Visualizing Data  
Fall 2023

# Check-in and logistics

Today's quick check-in:

**<https://forms.gle/narLSm6uPxWG9Uen6>**

This should take about 5 minutes.

# This week's reading

- ▶ Wednesday: 5.3, 5.4, 5.5 (labeling text in a figure)
- ▶ Next Monday: ggplot cheat sheet: <https://posit.co/wp-content/uploads/2022/10/data-visualization-1.pdf>

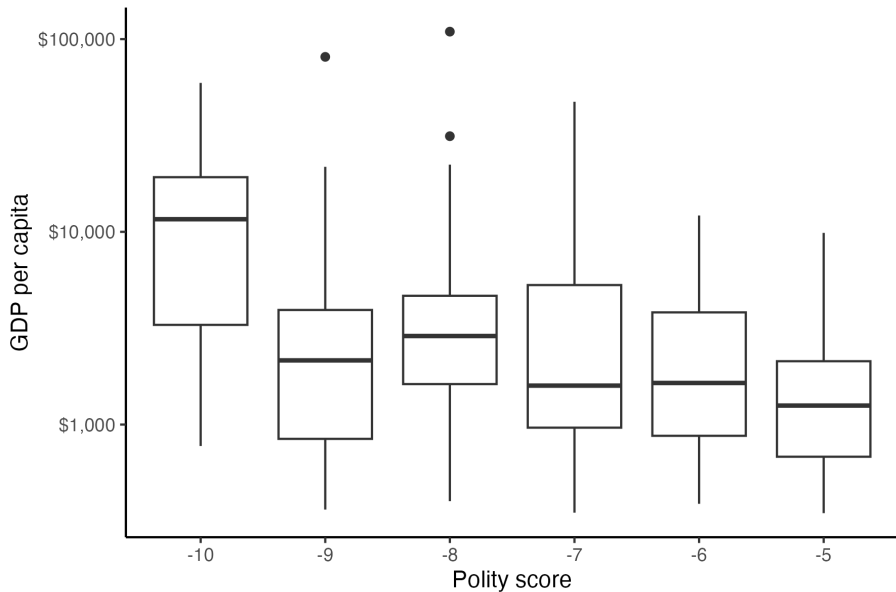
# Reordering plots

- ▶ By default, ggplot will order categorical variables alphabetically.
- ▶ But this usually isn't the most useful order.
- ▶ We can also arrange by the value of a different column.
- ▶ For example, to sort countries by their GDP, we can reorder them within our aes call:
- ▶ `aes(x = reorder(country, gdpPercap, na.rm=TRUE))`

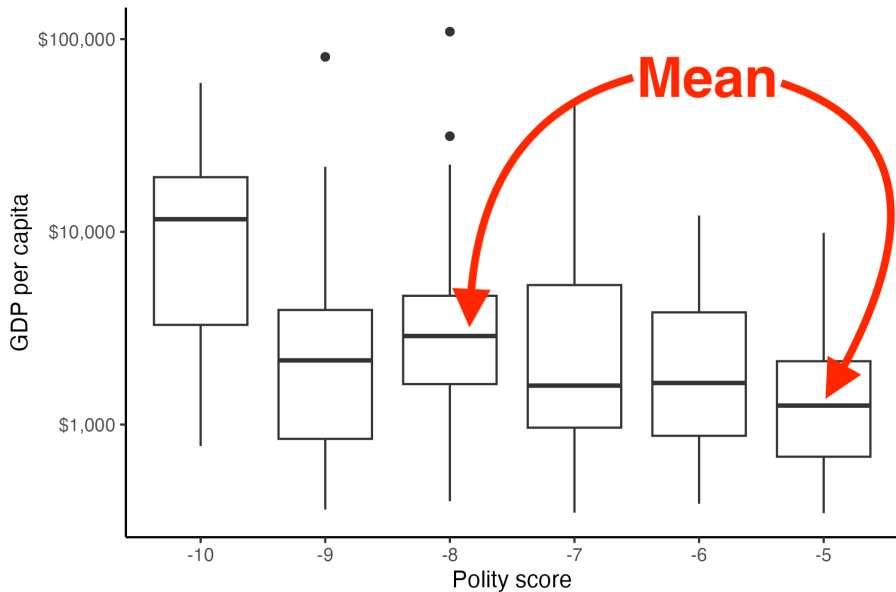
# Box and whisker plots

- ▶ Box and whisker plots (“boxplots”) show a useful set of summary statistics:
  - ▶ Mean
  - ▶ interquartile range
  - ▶ outliers

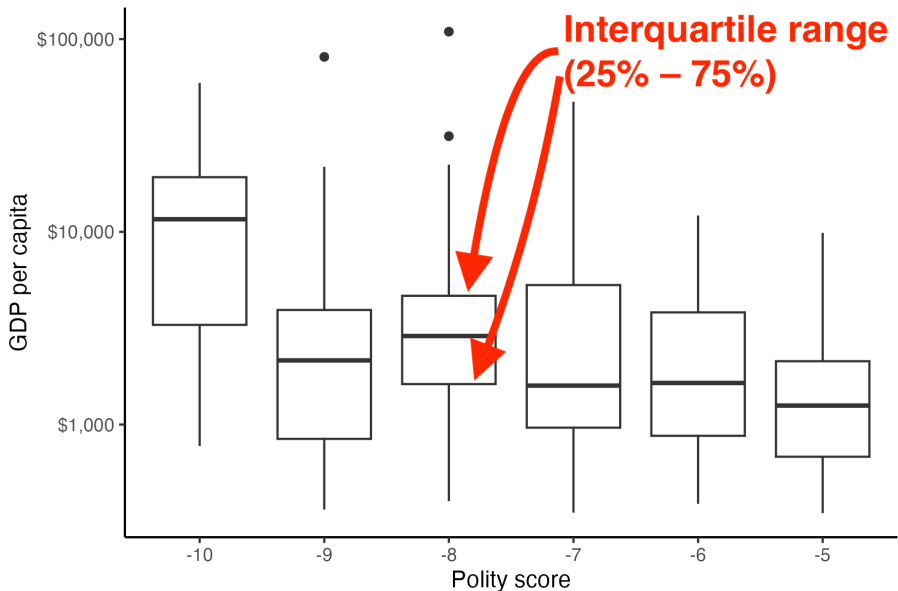
# Boxplot example



# Boxplot example

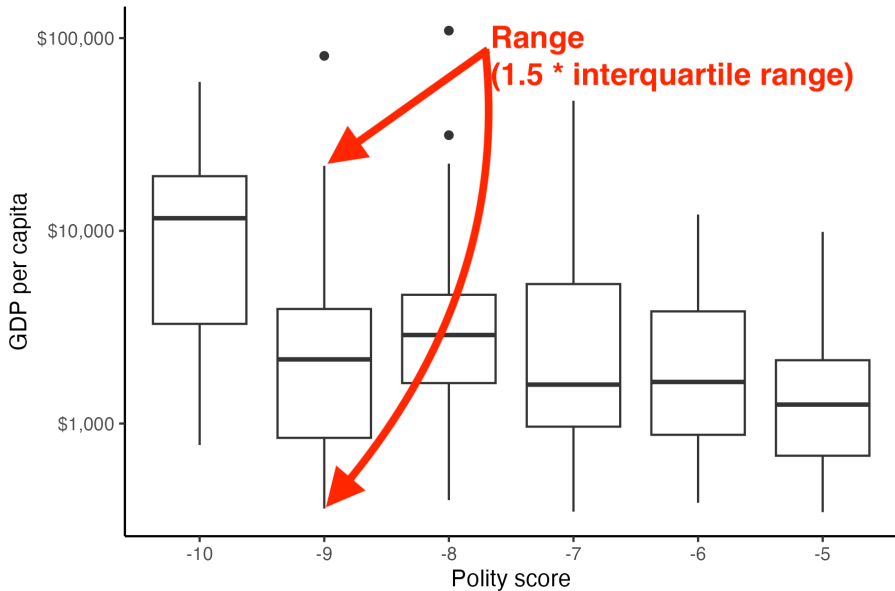


# Boxplot example

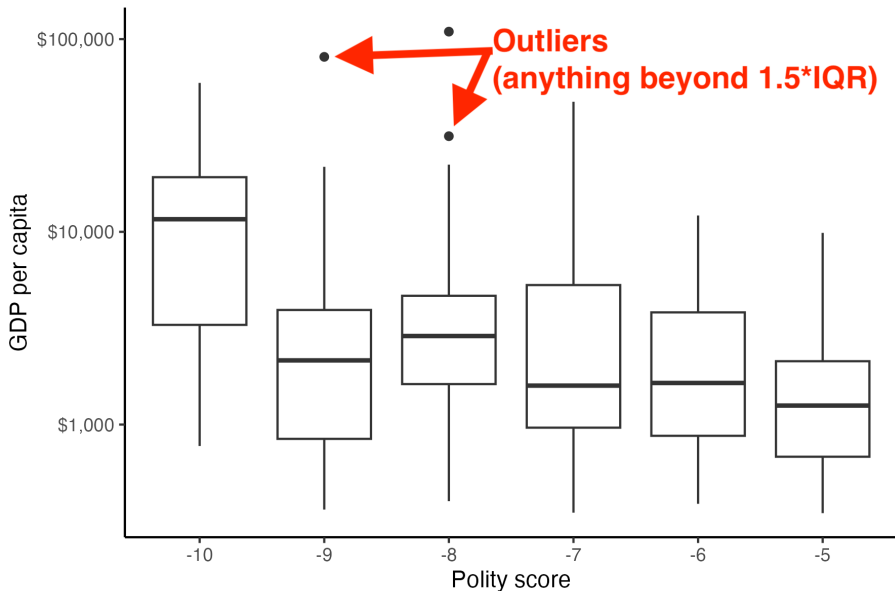




# Boxplot example



# Boxplot example



- ▶ Sometimes we want to show multiple density plots (across a categorical variable)
- ▶ We know how to do this with a facet
- ▶ But sometimes the facets take up too much space.
- ▶ Alternative: `library(ggridges)`
- ▶ + `geom_density_ridges()`, with the `y` variable set to our categorical variable

