Lab/HW 6: ggplot2; factors

Question 1

[5 pt]

The data for the exercises is available in the "EconomistData.csv" file on Moodle. Read it in with

dat <- read.csv("EconomistData.csv")

These data consist of Human Development Index and Corruption Perception Index scores for several countries.

- Create a scatter plot with CPI on the x axis and HDI on the y axis.
- Color the points in the plot blue.
- Color the points in the plot according to Region.
- Make the points bigger by setting the size to 2.
- Make the points bigger by setting the size to HDI rank.
- Create boxplots of CPI by Region.
- Overlay jittered points on top of the box plots.

Question 2

[5 pt]

Re-create a scatter plot with CPI on the x axis and HDI on the y axis (as you did in the previous exercise).

- Overlay a smoothing line on top of the scatter plot using the default method.
- Overlay a smoothing line on top of the scatter plot using the lm method (based on linear regression). Hint: see ?stat_smooth.
- Overlay a smoothing line on top of the scatter plot using the loess (locally weighted smoothing) method, but make it less smooth. Hint: see ?loess.

Question 3

[up to 10 pt]

Create a single plot (or multi-faceted plot) that highlights an interesting relationship between 4 variables from a dataset you provide that has at least 10 variables. Explain in words how the plot you created highlights this relationship and why you chose the aesthetic mappings / faceting method you ended up using.

Question 4

[5 pt] With the `gss_cat` data from lecture 7 part I, find how to manipulate the data and create a visualization that demonstrates how the proportions of people identifying as Democrat, Republican, and Independent changed over time.