Personality and Worldview

your name here

October 7, 2021

Measuring Personality

Carney et al. (2009) use several strategies to assess personality in their investigation of the link between personality and political ideology. This week's assignment will help you understand aspects of these methodologies and associated challenges.

The survey you took included two different short versions of the Big Five Inventory. You will have the chance to compare these two inventories and to get some experience with the Personal Living Space Cue Inventory (PLSCI) (Gosling et al. 2005), which was also used in Carney et al. (2009).

Data Details:

- File Name: `Oct7ClassData.csv`
- Source: These data are from the survey you took in class. You took the BFI-10, a short version of the Big Five Inventory (Rammstedt and John 2007) and the Ten-Item Personality Inventory (TIPI), which is a different instrument designed to quickly meausre the Big Five (Gosling et al. 2003). You then answered the same questions used by Carney et al. (2009) to assess political attitudes in some of their studies.

Variable Name	Variable Description
`Overall`	Self-reported overall ideology on a 1-5 scale with 1 being extremely liberal and 5 being extremely conservative
`Social`	Self-reported social ideology on a 1-5 scale with 1 being extremely liberal and 5 being extremely conservative
`Economic`	Self-reported economic ideology on a 1-5 scale with 1 being extremely liberal and 5 being extremely conservative
`Random ID`	A randomly generated respondent identifier
`BFI_extraversion`	The average of the two items on the BFI-10 associated with extraversion
`BFI_agreeableness`	The average of the two items on the BFI-10 associated with agreeableness
`BFI_conscientiousness`	The average of the two items on the BFI-10 associated with conscientiousness
`BFI_emot_stability`	The average of the two items on the BFI-10 associated with emotional stability (also known as neuroticism, although low emotional stability is the same as high neuroticism)
`BFI_openness`	The average of the two items on the BFI-10 associated with openness

Variable Name	Variable Description
`TIPI_extraversion`	The average of the two items on the TIPI-10 associated with extraversion
`TIPI_agreeableness`	The average of the two items on the TIPI-10 associated with agreeableness
`TIPI_conscientiousness`	The average of the two items on the TIPI-10 associated with conscientiousness
`TIPI_emot_stability`	The average of the two items on the TIPI-10 associated with emotional stability(also known as neuroticism)
`TIPI_openness`	The average of the two items on the TIPI-10 associated with openness

The last 20 variables in the data are the actual items from the BFI-10 and the TIPI. If you want to refresh your memory about the questions and learn which questions were meant to go together, you can find more about the BFI-10 here (scroll down to the "Is there a shorter version of the BFI available?" question) and the TIPI here.

Don't forget to load the data.

Question 1

Check the Google Drive for your photo assignments and download the assigned photos of living space (you should be assigned to two photos). **Use the worksheet on the Google Drive to do an assessment based on the PLSCI. Based on the PLSCI and your own general impressions, fill out the new class survey with your guess about each person's openness, conscientiousness, and overall political ideology without looking at the actual data. Make sure to write down your responses to the survey since you won't have access to them after you submit. (Page 391 of Gosling et al. 2002 might help you figure out which items on the inventory will be most helpful here.) Use the random ID to check your guesses. How did you do?** (If you are finding the PLSCI difficult and frustrating, peek ahead to the next few questions.)

Question 2

Each photo should have at least two people assigned to code it. Confer with each of the people who were also assigned to your photo and see what their ratings were on the PLSCI, as well as their guesses about openness, conscientiousness, and ideology. **What does this tell you about the PLSCI? Read a bit about intercoder (aka inter-rater) reliability and reflect on its importance in research like this.**

Vivian: mostly same answers: 4, 3, 2 Kayla: 5, 4, 1

Question 3

What challenges did you encounter administering the PLSCI? What do you notice that might complicate applying this inventory today?

Question 4

How would you redesign the PLSCI to make it more useful/current?

Question 5

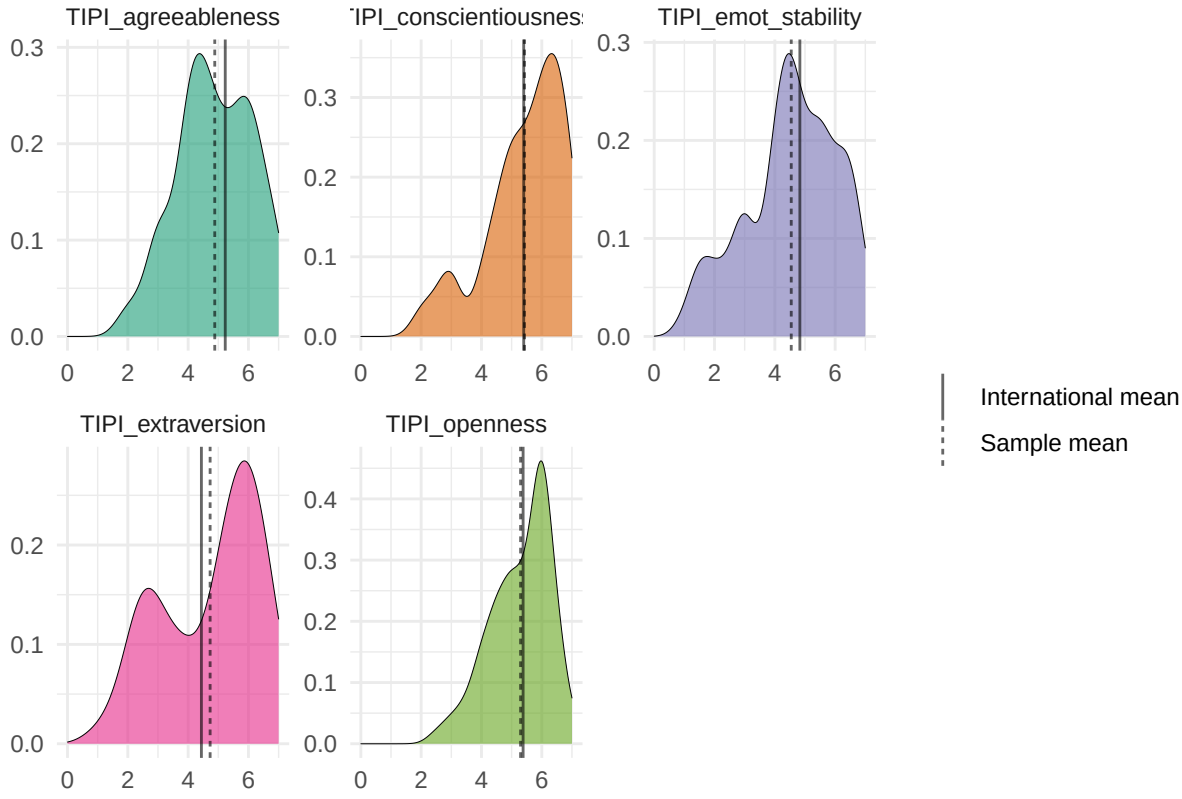
Let's see how the class compares to a large dataset of people who have taken one of these personality inventories. Gosling et al. (2003) report findings from administering the TIPI to about 1800 undergraduates. **Compare the class results to the published norms (look at the table on page 526, all ethnicities, Whole sample; i.e. the upper left corner of the table). Make a table or a plot of the class results and then write your thoughts about how these are similar or different to international norms and why that might be the case.**

```
# Create data frame for international averages/standard deviations for
# personality traits.

inter_norms <- tibble(trait = c("TIPI_extraversion", "TIPI_agreeableness",
                                "TIPI_conscientiousness",
         "TIPI_emot_stability", "TIPI_openness"),
         inter_mean = c(4.44, 5.23, 5.40, 4.83, 5.38),
         inter_sd = c(1.45, 1.11, 1.32, 1.42, 1.07))

# Pivot class data so rows correspond to every individual trait measured for
# individuals. group_by trait and summarize for mean and standard deviation.

TIPI_dist_comp <- personality_data %>%
  select(TIPI_extraversion, TIPI_agreeableness, TIPI_conscientiousness,
         TIPI_emot_stability, TIPI_openness) %>%
  pivot_longer(cols = TIPI_extraversion:TIPI_openness, names_to = "trait",
               values_to = "score") %>%
  group_by(trait) %>%
  mutate(mean = mean(score)) %>%
  ungroup() %>%

  # Join international data

  left_join(inter_norms, by = "trait") %>%
  select(-inter_sd) %>%

  # Pivot means such that they can be plotted and distinguished by aes mapping
  # in geom_vline

  pivot_longer(cols = c("mean", "inter_mean"), values_to = "means", names_to = "mean_type") %>%

  # Plot histograms and vlines for means

  ggplot(aes(x = score, fill = trait)) +
    geom_density(alpha = 0.6, size = 0, show.legend = FALSE) +
    facet_wrap("trait", scales = "free") +
    geom_vline(aes(xintercept = means, linetype = mean_type), alpha = 0.6) +
    theme_minimal() +
    scale_fill_brewer(palette = "Dark2") +
    scale_x_continuous(limits = c(0, 7)) +
    scale_linetype_discrete(name = "", labels = c("International mean", "Sample mean")) +
    theme(panel.spacing = unit(2, "lines")) %>%
    labs(x = "", y = "")

# Pivot class data so rows correspond to every individual trait measured for
```

```r
# individuals. group_by trait and summarize for mean and standard deviation.

class_inter <- personality_data %>%
  select(TIPI_extraversion, TIPI_agreeableness, TIPI_conscientiousness,
         TIPI_emot_stability, TIPI_openness) %>%
  pivot_longer(cols = TIPI_extraversion:TIPI_openness, names_to = "trait",
               values_to = "score") %>%
  group_by(trait) %>%
  summarize(mean = mean(score),
            sd = sd(score),
            .groups = "drop") %>%

  # inner_join combines inter_norms with class data using trait.

  inner_join(inter_norms, by = "trait")

TIPI_mean_comp <- class_inter %>%
  pivot_longer(cols = c(mean:inter_sd), names_to = "stat", values_to = "values") %>%
  filter(str_detect(stat, "mean") == TRUE) %>%
  ggplot(aes(x = trait, y = values, fill = stat)) +
    geom_col(position = "dodge") +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 30, vjust = 0.9, hjust = 1)) +
    labs(x = "", y = "", title = "Comparing class means to international means")

TIPI_sd_comp <- class_inter %>%
  pivot_longer(cols = c(mean:inter_sd), names_to = "stat", values_to = "values") %>%
  filter(str_detect(stat, "sd") == TRUE) %>%
  ggplot(aes(x = trait, y = values, fill = stat)) +
    geom_col(position = "dodge") +
    theme_minimal() +
    theme(axis.text.x = element_text(angle = 30, vjust = 0.9, hjust = 1)) +
    labs(x = "", y = "", title = "Comparing class sd to international sd")

class_inter
```

```
## # A tibble: 5 x 5
##   trait                  mean    sd inter_mean inter_sd
##   <chr>                 <dbl> <dbl>      <dbl>    <dbl>
## 1 TIPI_agreeableness     4.88 1.25        5.23     1.11
## 2 TIPI_conscientiousness 5.43 1.29        5.4      1.32
## 3 TIPI_emot_stability    4.54 1.51        4.83     1.42
## 4 TIPI_extraversion      4.73 1.62        4.44     1.45
## 5 TIPI_openness          5.30 0.990       5.38     1.07
```

```r
TIPI_dist_comp
```

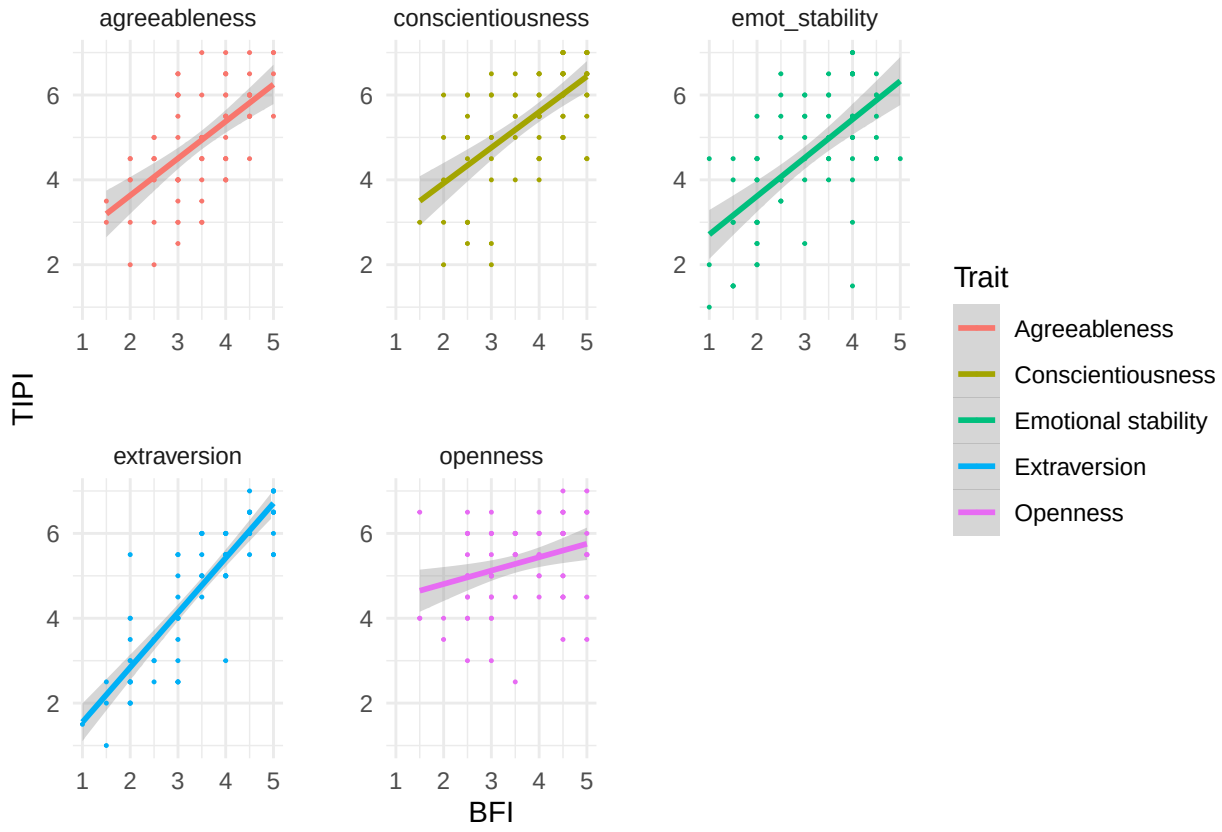


Question 6

The BFI-10 and the TIPI are supposed to measure the same five personality traits. To what degree do they seem to be measuring the same constructs in our class sample? **Compare each BFI-10 index to its counterpart TIPI index. You can do this numerically by calculating the correlation coefficient, graphically with a plot, or both. Comment on what you find.**

```
cor_clean <- personality_data %>%
  select(BFI_extraversion:TIPI_openness) %>%
  mutate(id = row_number()) %>%

  # Add an id column before pivoting such that id number reflects observations
  # from each respondent

  pivot_longer(cols = BFI_extraversion:TIPI_openness, names_to = "scale_trait",
               values_to = "values") %>%
  separate(col = "scale_trait", into = c("scale", "trait"), sep = "_",
           extra = "merge") %>%
  pivot_wider(names_from = scale, values_from = values) %>%
  drop_na()

TIPI_BFI_comp <- cor_clean %>%
  ggplot(aes(x = BFI, y = TIPI, color = trait)) +
    geom_smooth(method = "lm", formula = y ~ x) +
    geom_point(size = 0.2) +
```

```
    facet_wrap("trait", scales = "free") +
    theme_minimal() +
    theme(panel.spacing = unit(2, "lines")) +
    scale_x_continuous(limits = c(1, 5)) +
    scale_y_continuous(limits = c(1, 7)) +
    scale_color_discrete(name = "Trait", labels = c("Agreeableness", "Conscientiousness",
                                                    "Emotional stability", "Extraversion",
                                                    "Openness"))

TIPI_BFI_comp
```




```
cor_clean %>%
  group_by(trait) %>%
  summarize(meanBFI = mean(BFI),
            meanTIPI = mean(TIPI),
            cor = cor(x = BFI, y = TIPI),
            .groups = "drop")
```

```
## # A tibble: 5 x 4
##   trait             meanBFI meanTIPI   cor
##   <chr>               <dbl>    <dbl> <dbl>
## 1 agreeableness        3.43     4.88 0.601
## 2 conscientiousness    3.8      5.43 0.629
## 3 emot_stability       3.02     4.54 0.624
## 4 extraversion         3.46     4.73 0.869
```

```
## 5 openness             3.55     5.30 0.307
```

Question 7: Data Science Question

We are interested in whether personality is associated with political ideology. Multiple regression is one approach to simultaneously testing associations between several independent variables and a single dependent variable of interest. **Pick one of the personality inventories and use all of the trait indices in a regression model with at least one of the political ideology questions as the dependent variable. Interpret the results. (Note that you can use OLS for this question, even though it might not be the most appropriate model. As a bonus, you can explain why OLS might not be the best model and suggest an alternative. As another bonus, take a look back at all of the regressions you have just run. Why might we be skeptical of any individual p-value associated with one of these regression coefficients?)** Non-data science students should consider tackling part of this question but only using bivariate regression (one political ideology dependent variable and one personality independent variable.)

```
q7_lm1 <-  lm(formula = Overall ~ TIPI_extraversion + TIPI_agreeableness +
       TIPI_conscientiousness + TIPI_emot_stability + TIPI_openness, data = personality_data)

q7_lm2 <-  lm(formula = Overall ~ TIPI_agreeableness + TIPI_conscientiousness, data = personality_data)

summary(q7_lm1)
```

```
##
## Call:
## lm(formula = Overall ~ TIPI_extraversion + TIPI_agreeableness +
##     TIPI_conscientiousness + TIPI_emot_stability + TIPI_openness,
##     data = personality_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.6430 -0.3910 -0.1012  0.4936  1.6204
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.44749    0.59294   4.128  9.4e-05 ***
## TIPI_extraversion        0.02569    0.05854   0.439  0.66207
## TIPI_agreeableness      -0.18782    0.07088  -2.650  0.00982 **
## TIPI_conscientiousness   0.18741    0.06872   2.727  0.00795 **
## TIPI_emot_stability     -0.01645    0.06022  -0.273  0.78555
## TIPI_openness           -0.07487    0.09730  -0.769  0.44404
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7425 on 75 degrees of freedom
## Multiple R-squared:  0.1535, Adjusted R-squared:  0.09706
## F-statistic:  2.72 on 5 and 75 DF,  p-value: 0.02591
```

```
summary(q7_lm2)
```

```
##
## Call:
```

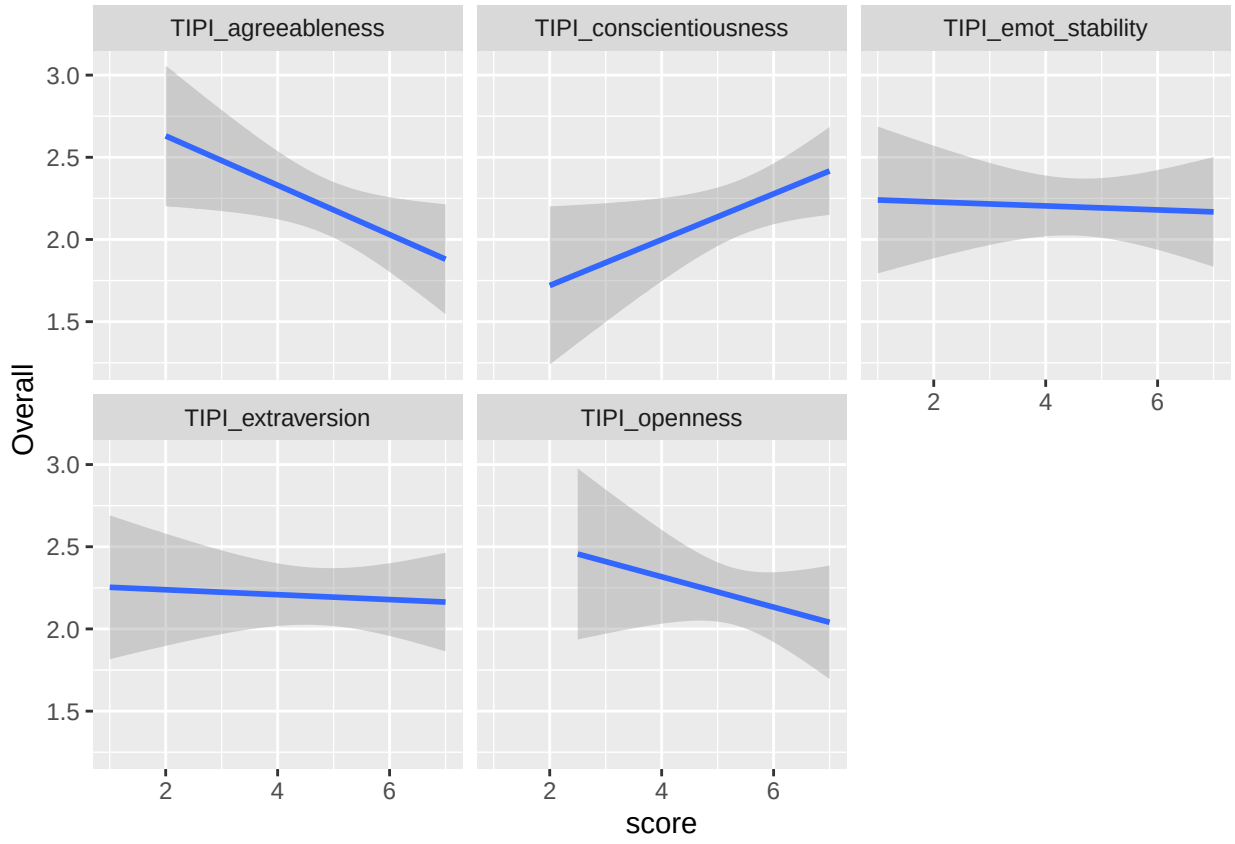
```
## lm(formula = Overall ~ TIPI_agreeableness + TIPI_conscientiousness,
##     data = personality_data)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.57784 -0.44327 -0.08519  0.53093  1.63401
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              2.15347    0.43050   5.002 3.42e-06 ***
## TIPI_agreeableness      -0.19479    0.06758  -2.882  0.00510 **
## TIPI_conscientiousness  0.18341    0.06503   2.820  0.00608 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7319 on 78 degrees of freedom
## Multiple R-squared:  0.1444, Adjusted R-squared:  0.1225
## F-statistic: 6.584 on 2 and 78 DF,  p-value: 0.002279
```

```
stargazer(q7_lm1, q7_lm2, type = "html")
```

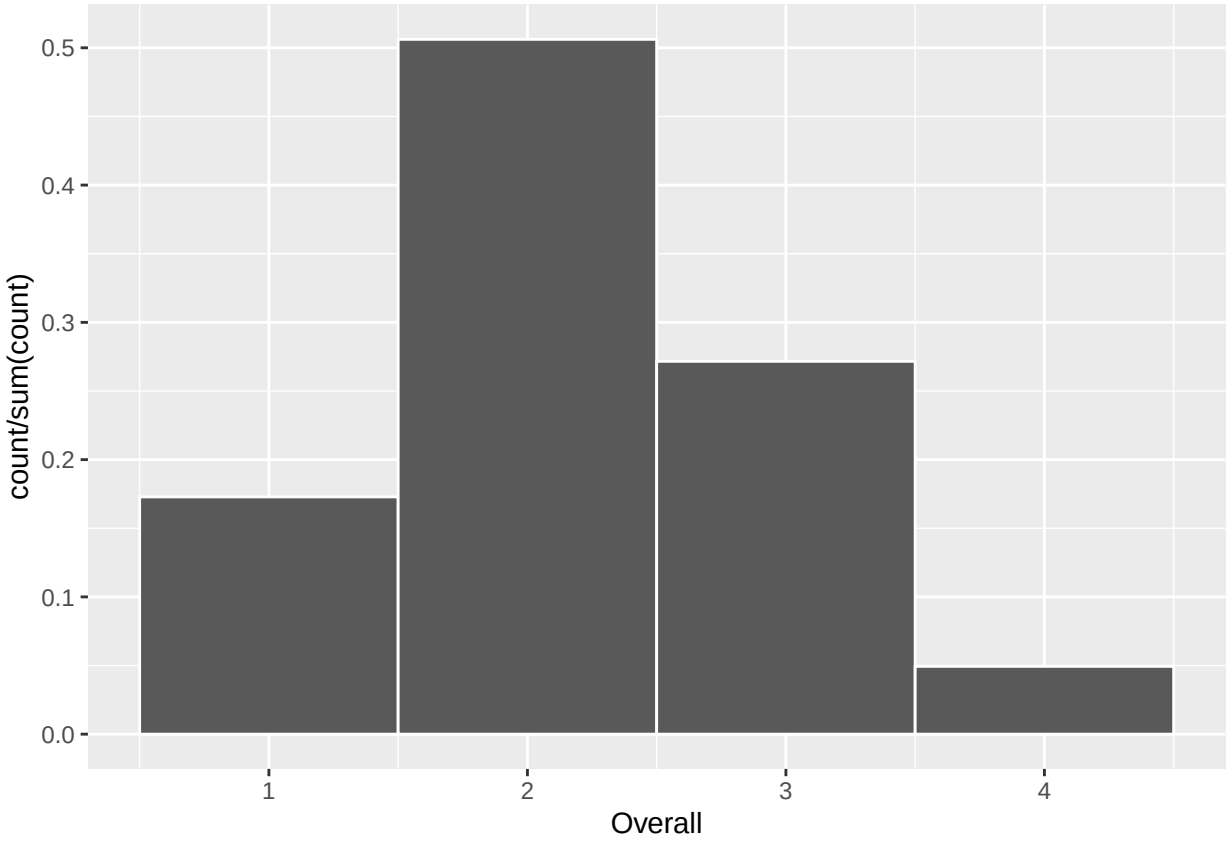
```
##
## <table style="text-align:center"><tr><td colspan="3" style="border-bottom: 1px solid black"></td></t
## <tr><td></td><td colspan="2" style="border-bottom: 1px solid black"></td></tr>
## <tr><td style="text-align:left"></td><td colspan="2">Overall</td></tr>
## <tr><td style="text-align:left"></td><td>(1)</td><td>(2)</td></tr>
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## <tr><td style="text-align:left"></td><td>(0.059)</td><td></td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td style="text-align:left">TIPI_agreeableness</td><td>-0.188<sup>***</sup></td><td>-0.195<sup>**
## <tr><td style="text-align:left"></td><td>(0.071)</td><td>(0.068)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td style="text-align:left">TIPI_conscientiousness</td><td>0.187<sup>***</sup></td><td>0.183<sup
## <tr><td style="text-align:left"></td><td>(0.069)</td><td>(0.065)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td style="text-align:left">TIPI_emot_stability</td><td>-0.016</td><td></td></tr>
## <tr><td style="text-align:left"></td><td>(0.060)</td><td></td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td style="text-align:left">TIPI_openness</td><td>-0.075</td><td></td></tr>
## <tr><td style="text-align:left"></td><td>(0.097)</td><td></td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td style="text-align:left">Constant</td><td>2.447<sup>***</sup></td><td>2.153<sup>***</sup></td
## <tr><td style="text-align:left"></td><td>(0.593)</td><td>(0.430)</td></tr>
## <tr><td style="text-align:left"></td><td></td><td></td></tr>
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## <tr><td style="text-align:left">R<sup>2</sup></td><td>0.153</td><td>0.144</td></tr>
## <tr><td style="text-align:left">Adjusted R<sup>2</sup></td><td>0.097</td><td>0.123</td></tr>
## <tr><td style="text-align:left">Residual Std. Error</td><td>0.742 (df = 75)</td><td>0.732 (df = 78)</
## <tr><td style="text-align:left">F Statistic</td><td>2.720<sup>**</sup> (df = 5; 75)</td><td>6.584<su
## <tr><td colspan="3" style="border-bottom: 1px solid black"></td></tr><tr><td style="text-align:left"
## </table>
```

```r
personality_data %>%
  select(Overall, TIPI_extraversion, TIPI_agreeableness, TIPI_conscientiousness,
         TIPI_emot_stability, TIPI_openness) %>%
  pivot_longer(cols = TIPI_extraversion:TIPI_openness, names_to = "trait",
               values_to = "score") %>%
  ggplot(aes(x = score, y = Overall)) +
    geom_smooth(method = "lm") +
    facet_wrap("trait")
```

```
## `geom_smooth()` using formula 'y ~ x'
```




```r
personality_data %>%
  select(Overall) %>%
  ggplot(aes(x = Overall, y = after_stat(count/sum(count)))) +
    geom_histogram(binwidth = 1, color = "white")
```



Linear regression shows that only two of the personality traits (agreeableness, conscientiousness) are significantly correlated with the independent variable (overall ideology). The corresponding p values for those predictors are comparatively low, meaning they are statistically significant. But most of the traits are not significantly correlated with overall ideology. Both the Multiple R-squared and Adjusted R-squared values are low, meaning little of the variance in observations is actually explained by the model.

OLS might not be the best linear regression model to fit the data because the variables might not be linearly related. Breaking down the relationships between each independent variable with the dependent variable shows that there aren't many strong relationships between the personality predictors and social ideology – only agreeableness and conscientiousness show a strong independently linear relationship with overall ideology.

Additionally, even when p-value coefficients indicate significance, other model outputs can be misleading. For example, when you run the regrression with all 5 personality predictors, the multiple R-squared value is

Question 8: Data Science Question

Is the Big Five really best characterized as five factors? If we ask ten questions on a personality inventory, we might think that these questions actually reflect only five underlying (or latent) variables. In fact, this is the supposition of the BFI-10 and the TIPI. Factor analysis is one way to examine data and investigate if the dimension of the data can be reduced from many variables to fewer underlying factors. **Conduct a factor analysis on either the BFI-10 or the TIPI questions (note, these are the numbered variables, not the named variables) You may want to use the `fa.parallel` function from the `psych` package, which you can read more about here. How many factors does your analysis suggest best explain the class data? Optional bonus: run the code several times. Does your answer change? If so, why?**

Question 9

If you still have your random ID, take a look at your own BFI-10 and TIPI index scores for each factor. **How do your scores compare to one another across instruments? How do your ratings on each of the Big Five factors compare to your own self image? (Remember that the BFI-10 uses a five point scale while the TIPI uses a seven point scale.)**