Lab-04

Negar

2/10/2022

```
dictionary <- readr::read_csv(here::here("data", "green_dictionary.csv"))
```

```
## Rows: 36 Columns: 4
## -- Column specification -------------------------------------------------------
## Delimiter: ","
## chr (3): Item, Content, Options
## dbl (1): Keying
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
green_data <- readr::read_csv(here::here("data", "green_data.csv"))
```

```
## Rows: 373 Columns: 37
## -- Column specification -------------------------------------------------------
## Delimiter: ","
## chr  (1): id
## dbl (36): green1, green2, green3, green4, green5, comp1, comp2, comp3, comp4...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

Question 1

```
dictionary |>
  summarize(dictionary)
```

```
## # A tibble: 36 x 4
##    Item   Content                                         Options        Keying
##    <chr>  <chr>                                           <chr>           <dbl>
##  1 green1 Overall, I am regarded as an environmentally c~ 1-5; SD, D, NA~      1
##  2 green2 I am not particularly known for protecting the~ 1-5; SD, D, NA~     -1
##  3 green3 My friends know I enjoy nature.                 1-5; SD, D, NA~      1
##  4 green4 I have a reputation for living in harmony with~ 1-5; SD, D, NA~      1
##  5 green5 Among people I know, I seem to be the one most~ 1-5; SD, D, NA~      1
##  6 comp1  I feel others' emotions.                        1-5; SD, D, NA~      1
##  7 comp2  I inquire about others' well-being.             1-5; SD, D, NA~      1
##  8 comp3  I sympathize with others' feelings.             1-5; SD, D, NA~      1
##  9 comp4  I take an interest in other people's lives.     1-5; SD, D, NA~      1
## 10 comp5  I like to do things for others.                 1-5; SD, D, NA~      1
## # ... with 26 more rows
```

```
green_data_changed <- green_data |>
             na_if("-99")
```

```r
reversed <- dictionary |>
  filter(Keying == -1) |>
  pull(Item)
green_data_reverse <- green_data_changed |>
  mutate(across(all_of(reversed),
                ~ recode(.x, "1"= 5, "2" = 4, "3" = 3, "4" = 2, "5" = 1),
                .names = "{.col}r"))
green_data_changed$student <- recode_factor(green_data_changed$student, '1' = "Non-student", '2' = "Stud
```

Question 2

```r
green_data_changed |>
  rowwise() |>
  mutate(
    Green_total = mean(c_across(green1:green5), na.rm = TRUE),
    Comp_total = mean(c_across(comp1:comp10), na.rm = TRUE),
    Intel_total = mean(c_across(intel1:intel10), na.rm = TRUE),
    Open_total = mean(c_across(open1:open10), na.rm = TRUE)
  ) |>
  ungroup()
```

```
## # A tibble: 373 x 41
##    id     green1 green2 green3 green4 green5 comp1 comp2 comp3 comp4 comp5 comp6
##    <chr>   <dbl>  <dbl>  <dbl>  <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
##  1 9099        4      2      5      4      3     2     5     4     4     2     2
##  2 6275        3      5      3      3      1     4     5     4     5     5     1
##  3 8116        4      2      5      4      3     4     5     5     4     5     1
##  4 8586        4      3      5      4      3     4     5     5     4     5     1
##  5 0406        3      2      4      4      4     4     4     4     4     4     2
##  6 5645        4      3      4      4      3     5     5     5     4     4     2
##  7 3788        1      1      1      1      1     3     3     3     3     3     3
##  8 8424        4      3      2      2      2     1     1     1     2     2     4
##  9 8450        4      2      3      5      2     4     4     4     4     4     1
## 10 0512        4      2      5      5      4     5     2     5     4     2     2
## # ... with 363 more rows, and 29 more variables: comp7 <dbl>, comp8 <dbl>,
## #   comp9 <dbl>, comp10 <dbl>, intel1 <dbl>, intel2 <dbl>, intel3 <dbl>,
## #   intel4 <dbl>, intel5 <dbl>, intel6 <dbl>, intel7 <dbl>, intel8 <dbl>,
## #   intel9 <dbl>, intel10 <dbl>, open1 <dbl>, open2 <dbl>, open3 <dbl>,
## #   open4 <dbl>, open5 <dbl>, open6 <dbl>, open7 <dbl>, open8 <dbl>,
## #   open9 <dbl>, open10 <dbl>, student <fct>, Green_total <dbl>,
## #   Comp_total <dbl>, Intel_total <dbl>, Open_total <dbl>
```

Question 3

```r
green_data_pomp <- green_data_changed |>
  mutate(across(c(green1:open10),
         ~ recode(.x, "1"= 0,"2" = 25, "3" = 50, "4" = 75,"5"= 100),
                .names = "{.col}_pomp"))
print(green_data_pomp)
```

```
## # A tibble: 373 x 72
##    id     green1 green2 green3 green4 green5 comp1 comp2 comp3 comp4 comp5 comp6
##    <chr>   <dbl>  <dbl>  <dbl>  <dbl>  <dbl> <dbl> <dbl> <dbl> <dbl> <dbl> <dbl>
```

```
##  1 9099         4     2     5     4     3     2     5     4     4     2     2
##  2 6275         3     5     3     3     1     4     5     4     5     5     1
##  3 8116         4     2     5     4     3     4     5     5     4     5     1
##  4 8586         4     3     5     4     3     4     5     5     4     5     1
##  5 0406         3     2     4     4     4     4     4     4     4     4     2
##  6 5645         4     3     4     4     3     5     5     5     4     4     2
##  7 3788         1     1     1     1     1     3     3     3     3     3     3
##  8 8424         4     3     2     2     2     1     1     1     2     2     4
##  9 8450         4     2     3     5     2     4     4     4     4     4     1
## 10 0512         4     2     5     5     4     5     2     5     4     2     2
## # ... with 363 more rows, and 60 more variables: comp7 <dbl>, comp8 <dbl>,
## #   comp9 <dbl>, comp10 <dbl>, intel1 <dbl>, intel2 <dbl>, intel3 <dbl>,
## #   intel4 <dbl>, intel5 <dbl>, intel6 <dbl>, intel7 <dbl>, intel8 <dbl>,
## #   intel9 <dbl>, intel10 <dbl>, open1 <dbl>, open2 <dbl>, open3 <dbl>,
## #   open4 <dbl>, open5 <dbl>, open6 <dbl>, open7 <dbl>, open8 <dbl>,
## #   open9 <dbl>, open10 <dbl>, student <fct>, green1_pomp <dbl>,
## #   green2_pomp <dbl>, green3_pomp <dbl>, green4_pomp <dbl>, ...
```

Question 4

```
Avg_green_data_pomp <- green_data_pomp |>
  rowwise() |>
  mutate(Green_pomp = mean(c_across(green1:green5), na.rm = TRUE),
         Comp_pomp = mean(c_across(comp1:comp10), na.rm = TRUE),
         Intel_pomp = mean(c_across(intel1:intel10), na.rm = TRUE),
         Open_pomp = mean(c_across(open1:open10), na.rm = TRUE))
Avg_green_data_pomp$student <- recode_factor(Avg_green_data_pomp$student, '1' = "Non-student", '2' = "S

Avg_green_data_pomp |>
ggplot() +
aes(x = Green_pomp) +
  geom_histogram(fill = "darkgreen") +
  geom_density(fill="black") +
  ylab("Count")
```
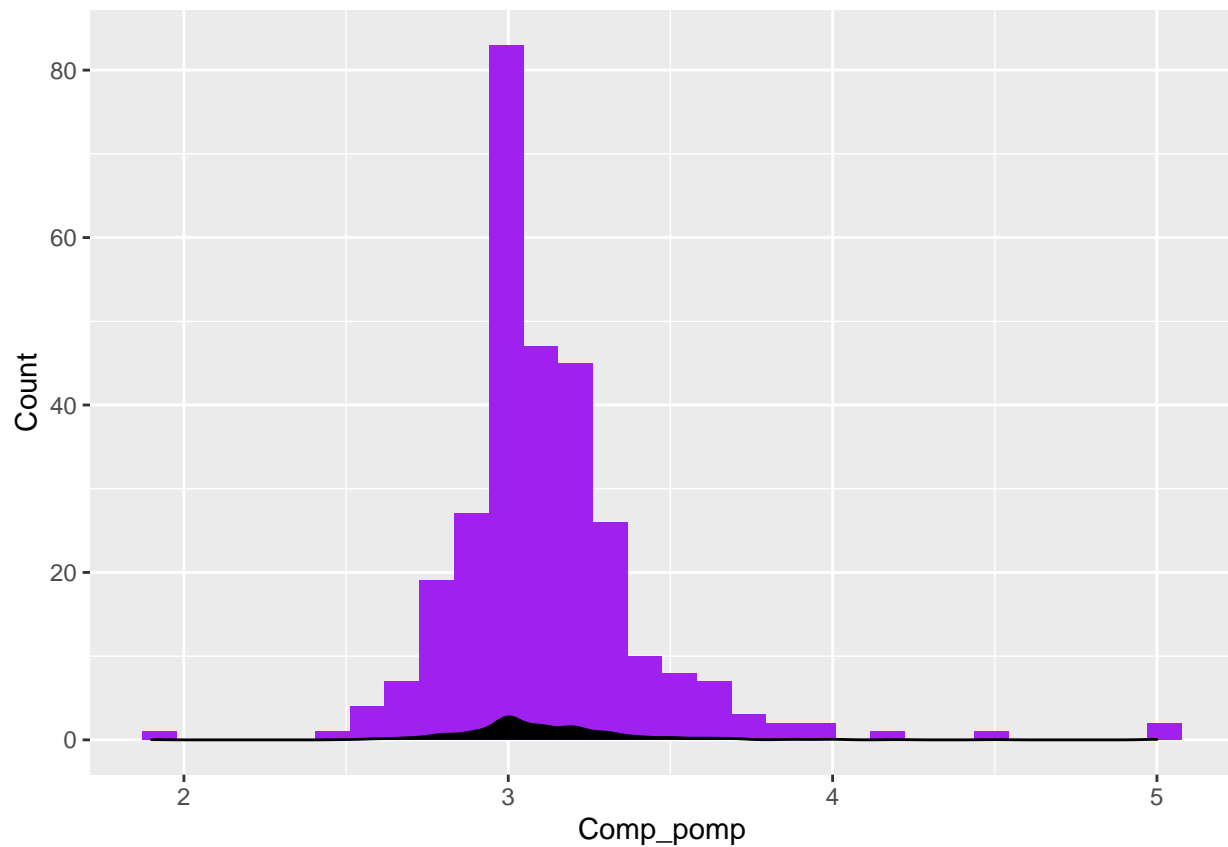
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 58 rows containing non-finite values (stat_bin).
```

```
## Warning: Removed 58 rows containing non-finite values (stat_density).
```
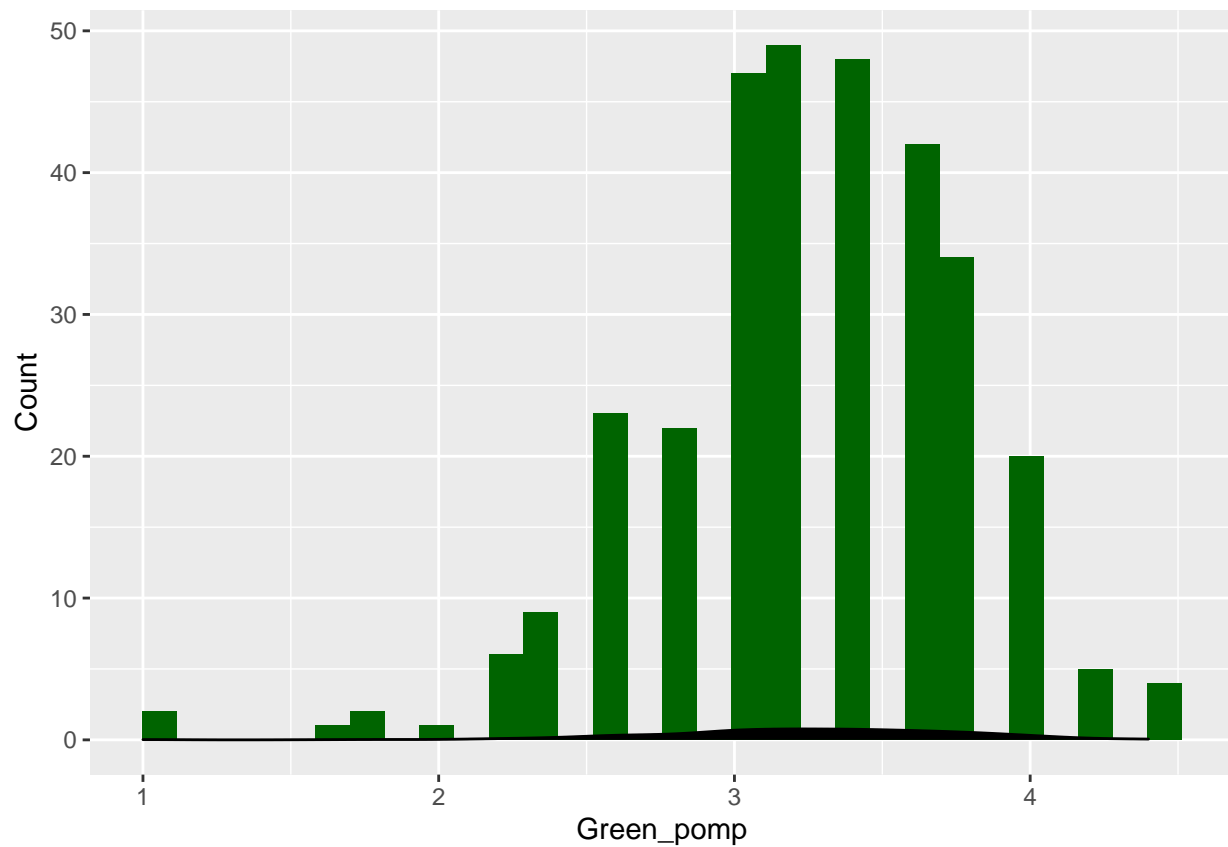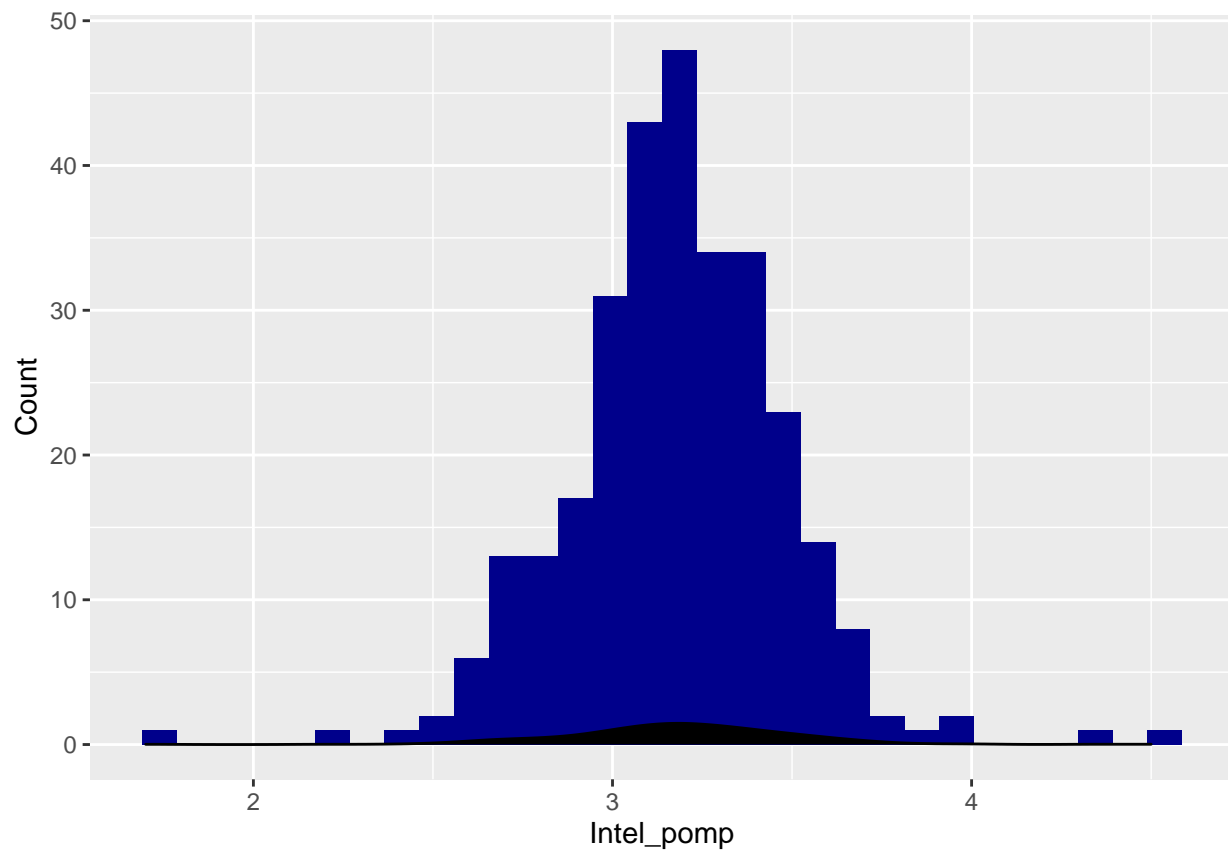
```
Avg_green_data_pomp |>
ggplot() +
aes(x = Comp_pomp) +
  geom_histogram(fill = "purple") +
  geom_density(fill="black") +
  ylab("Count")
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Warning: Removed 77 rows containing non-finite values (stat_bin).

## Warning: Removed 77 rows containing non-finite values (stat_density).

```
Avg_green_data_pomp |>
ggplot() +
aes(x = Intel_pomp) +
  geom_histogram(fill = "darkblue") +
  geom_density(fill="black") +
  ylab("Count")
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Warning: Removed 77 rows containing non-finite values (stat_bin).
Removed 77 rows containing non-finite values (stat_density).

```
Avg_green_data_pomp |>
ggplot() +
aes(x = Open_pomp) +
  geom_histogram(fill = "darkred") +
  geom_density(fill="black") +
  ylab("Count")
```

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.

Warning: Removed 77 rows containing non-finite values (stat_bin).
Removed 77 rows containing non-finite values (stat_density).
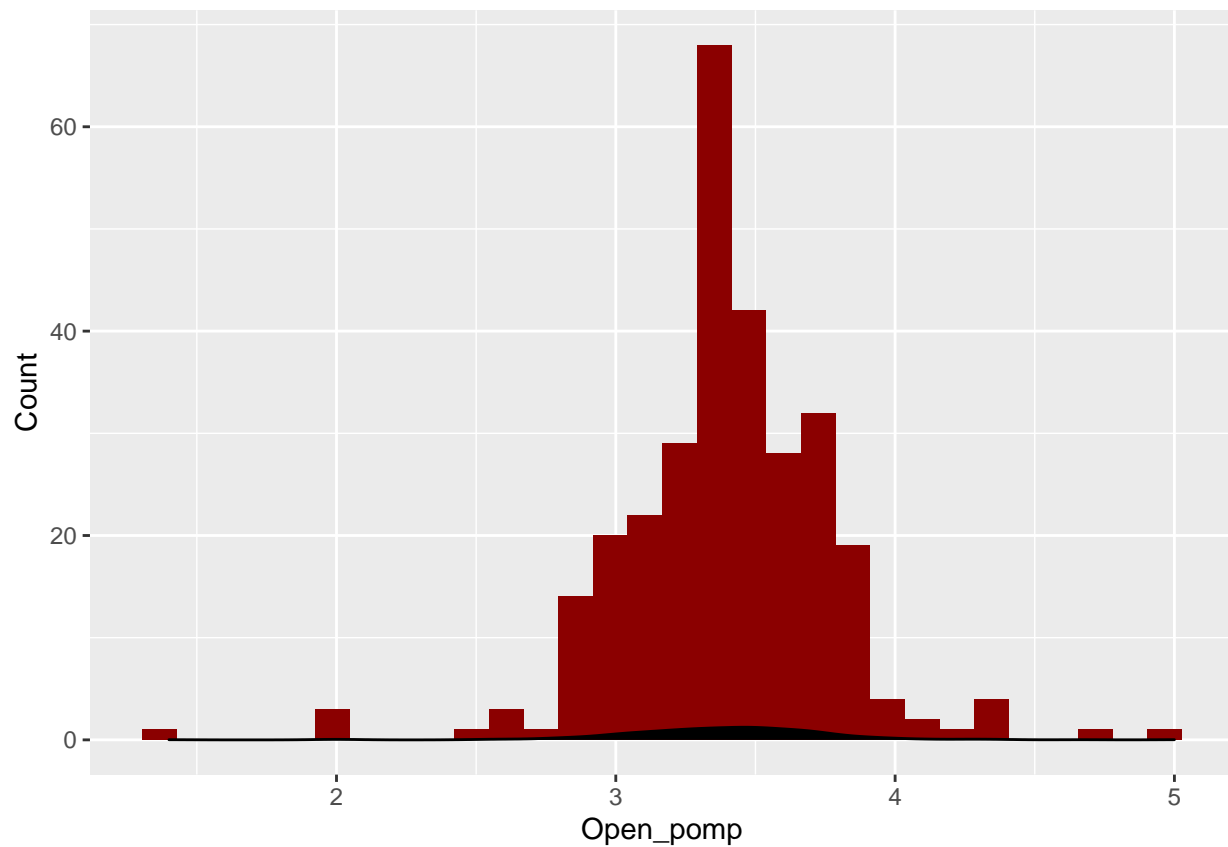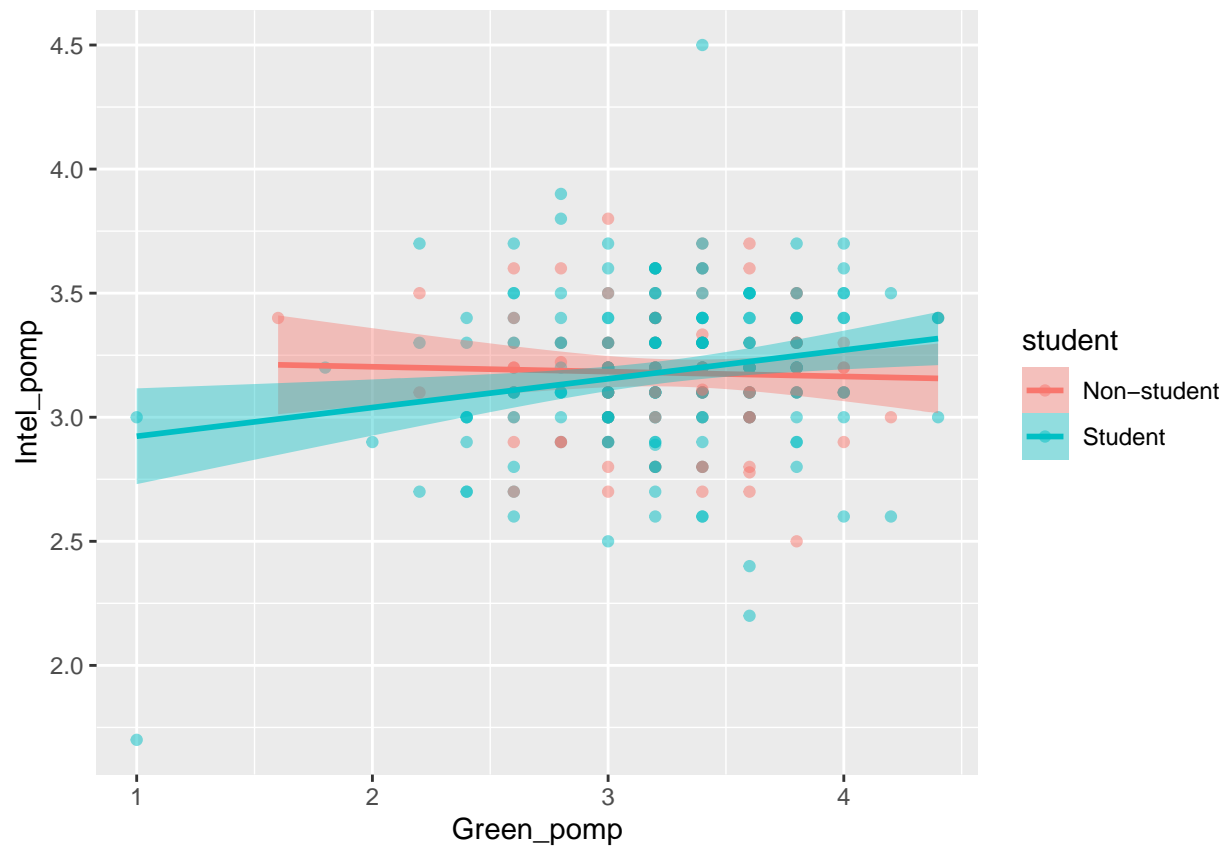
Question 5

```
Avg_green_data_pomp |>
  filter(!is.na(student)) |>
  ggplot() +
  aes(x = Green_pomp, y = Comp_pomp,
      group = student,
      color = student,
      fill = student) +
  geom_point(alpha = .5) +
  geom_smooth(method = "lm")
```

```
## `geom_smooth()` using formula 'y ~ x'
```
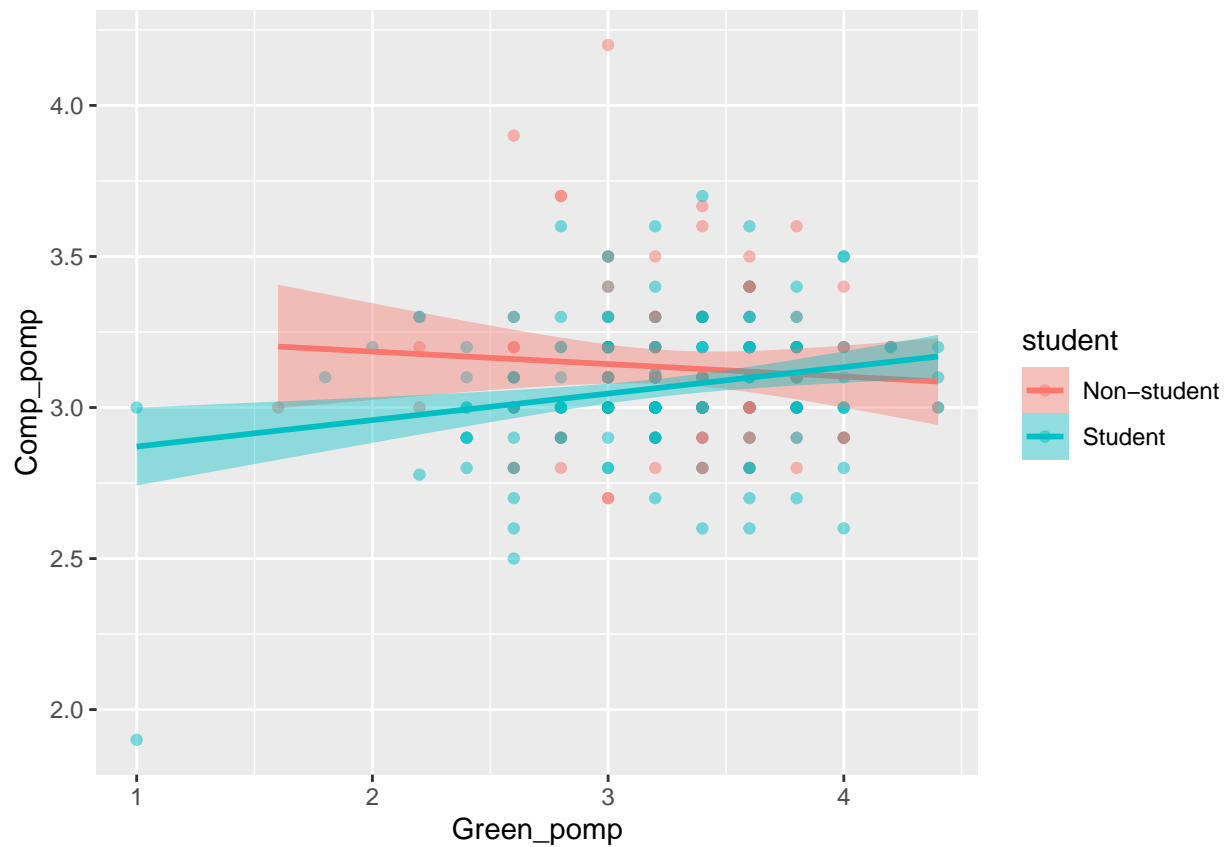
```
Avg_green_data_pomp |>
  filter(!is.na(student)) |>
  ggplot() +
  aes(x = Green_pomp, y = Intel_pomp,
      group = student,
      color = student,
      fill = student) +
  geom_point(alpha = .5) +
  geom_smooth(method = "lm")
```

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

```
Avg_green_data_pomp |>
  filter(!is.na(student)) |>
  ggplot() +
  aes(x = Green_pomp, y = Open_pomp,
      group = student,
      color = student,
      fill = student) +
  geom_point(alpha = .5) +
  geom_smooth(method = "lm")
```

`geom_smooth()` using formula 'y ~ x'
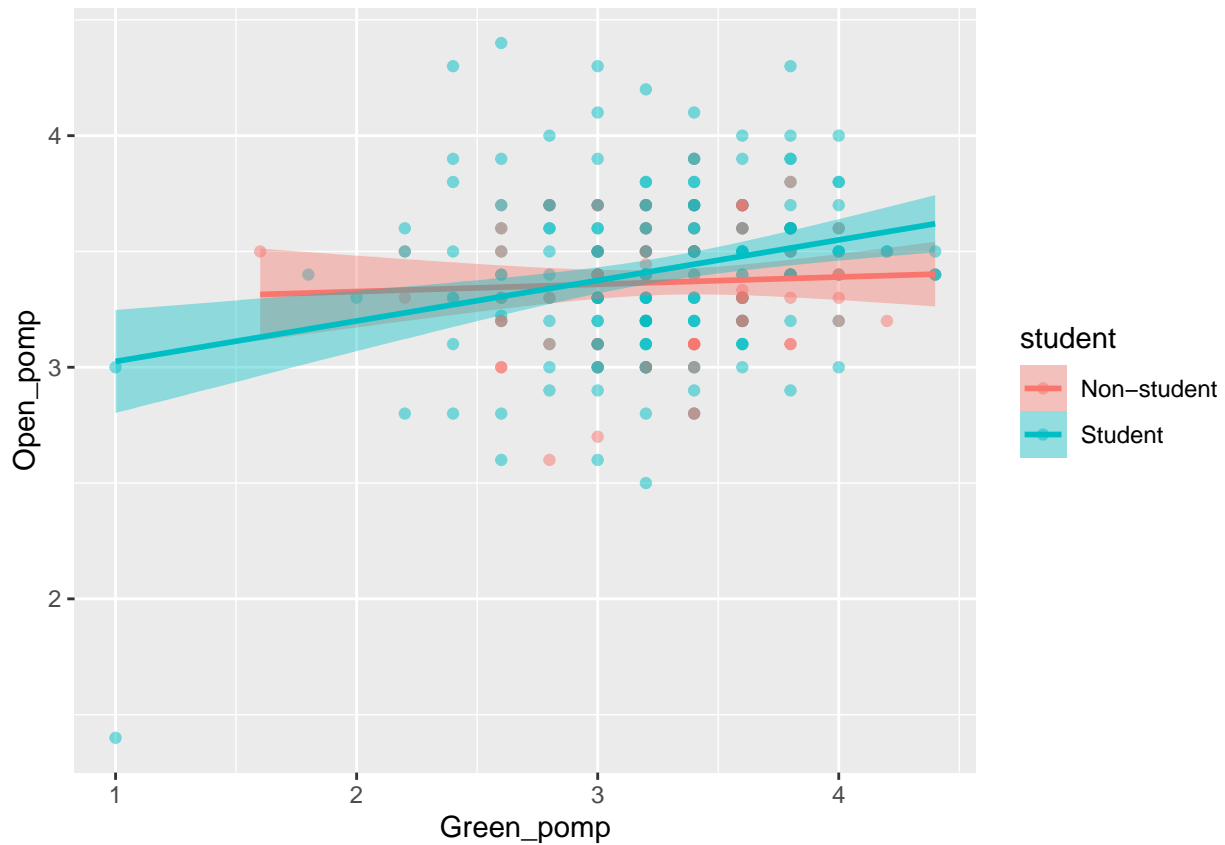
In all three plots, the non-student group showed higher scores at first, and then the scores for the student group went higher than the non-student group.

Question 6

```
ggplot(na.omit(Avg_green_data_pomp)) +
  geom_jitter(aes(x = Green_pomp, y = student,
      fill = student,
      color = student), height = .15) +
  geom_point(aes(x = Green_pomp, y = student,
      fill = student,
      color = student)) +
  geom_bar(aes(x = Green_pomp,
      fill = student,
      color = student), alpha = 0.5) +
  geom_density(aes(x = Green_pomp,
      fill = student,
      color = student), alpha = 0.5)
```
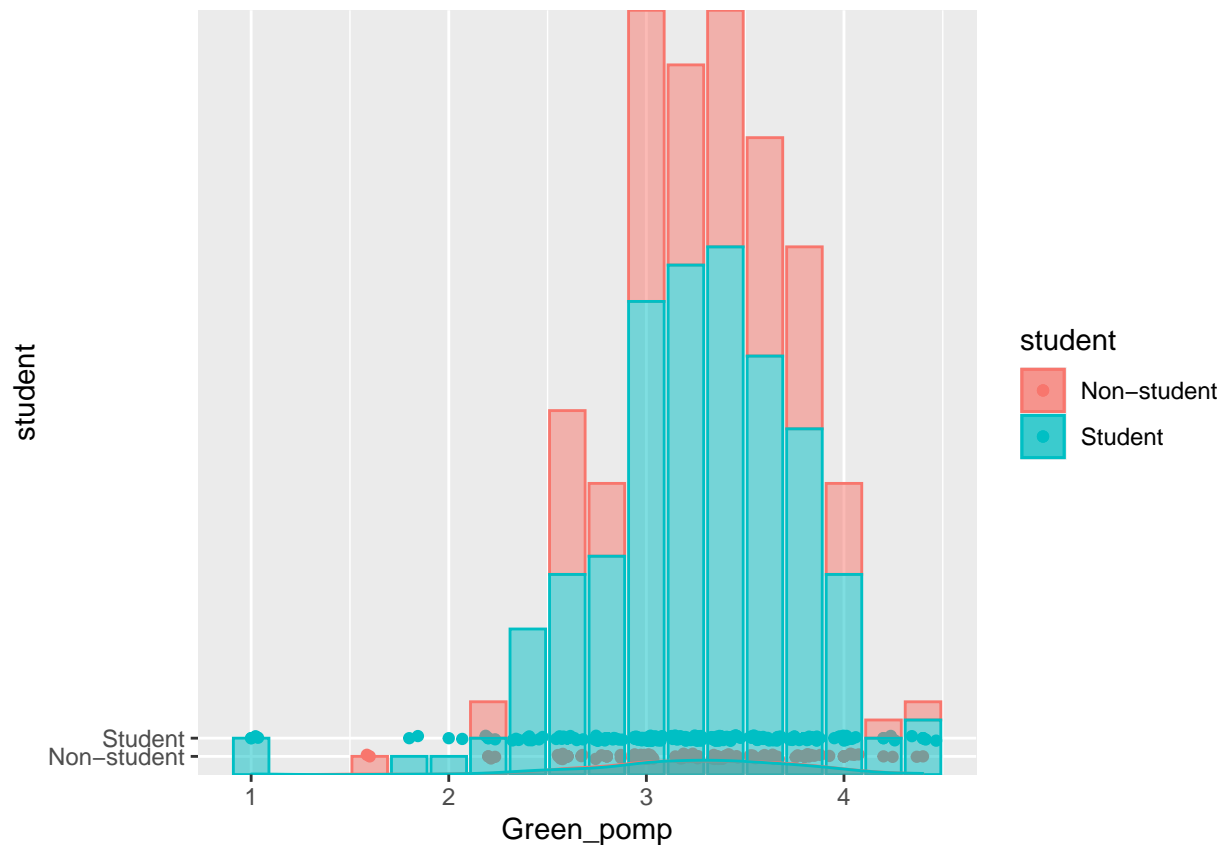
Overall, non-student group had higher environmental awareness scores compared to student group.

Question 7

```
table <- Avg_green_data_pomp |>
  na.omit() |>
  group_by(student) |>
  summarize(across(c(Green_pomp, Comp_pomp, Intel_pomp, Open_pomp),
          list(
            Mean = ~ mean(.x, na.rm = T),
            SDs = ~ sd(.x, na.rm = T),
            Medians = ~ median(.x, na.rm = T),
            Minima = ~ min(.x, na.rm = T),
            Maxima = ~ max(.x, na.rm = T)
          )))
table
```

```
## # A tibble: 2 x 21
##   student      Green_pomp_Mean Green_pomp_SDs Green_pomp_Medians Green_pomp_Mini~
##   <fct>                  <dbl>          <dbl>              <dbl>            <dbl>
## 1 Non-student             3.27          0.492                3.2              1.6
## 2 Student                 3.25          0.541                3.2              1
## # ... with 16 more variables: Green_pomp_Maxima <dbl>, Comp_pomp_Mean <dbl>,
## #   Comp_pomp_SDs <dbl>, Comp_pomp_Medians <dbl>, Comp_pomp_Minima <dbl>,
## #   Comp_pomp_Maxima <dbl>, Intel_pomp_Mean <dbl>, Intel_pomp_SDs <dbl>,
## #   Intel_pomp_Medians <dbl>, Intel_pomp_Minima <dbl>, Intel_pomp_Maxima <dbl>,
## #   Open_pomp_Mean <dbl>, Open_pomp_SDs <dbl>, Open_pomp_Medians <dbl>,
```

```
## #   Open_pomp_Minima <dbl>, Open_pomp_Maxima <dbl>
```