Multiple linear regression

Grading the professor

Many college courses conclude by giving students the opportunity to evaluate the course and the instructor anonymously. However, the use of these student evaluations as an indicator of course quality and teaching effectiveness is often criticized because these measures may reflect the influence of non-teaching related characteristics, such as the physical appearance of the instructor. The article titled, "Beauty in the classroom: instructors' pulchritude and putative pedagogical productivity" by Hamermesh and Parker found that instructors who are viewed to be better looking receive higher instructional ratings.

Here, you will analyze the data from this study in order to learn what goes into a positive professor evaluation.

Getting Started

Load packages

In this lab, you will explore and visualize the data using the **tidyverse** suite of packages. The data can be found in the companion package for OpenIntro resources, **openintro**.

Let's load the packages.

```
library(tidyverse)
library(openintro)
library(GGally)
```

This is the first time we're using the `GGally` package. You will be using the `ggpairs` function from this package later in the lab.

The data

The data were gathered from end of semester student evaluations for a large sample of professors from the University of Texas at Austin. In addition, six students rated the professors' physical appearance. The result is a data frame where each row contains a different course and columns represent variables about the courses and professors. It's called `evals`.

```
glimpse(evals)
```

```
## Rows: 463
## Columns: 23
## $ course_id    <int> 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 1~
## $ prof_id      <int> 1, 1, 1, 1, 2, 2, 2, 3, 3, 4, 4, 4, 4, 4, 4, 4, 4, 5, 5,~
## $ score        <dbl> 4.7, 4.1, 3.9, 4.8, 4.6, 4.3, 2.8, 4.1, 3.4, 4.5, 3.8, 4~
## $ rank         <fct> tenure track, tenure track, tenure track, tenure track, ~
## $ ethnicity    <fct> minority, minority, minority, minority, not minority, no~
## $ gender       <fct> female, female, female, female, male, male, male, male, ~
```

```
## $ language    <fct> english, english, english, english, english, english, en~
## $ age         <int> 36, 36, 36, 36, 59, 59, 59, 51, 51, 40, 40, 40, 40, 40, ~
## $ cls_perc_eval <dbl> 55.81395, 68.80000, 60.80000, 62.60163, 85.00000, 87.500~
## $ cls_did_eval <int> 24, 86, 76, 77, 17, 35, 39, 55, 111, 40, 24, 24, 17, 14,~
## $ cls_students <int> 43, 125, 125, 123, 20, 40, 44, 55, 195, 46, 27, 25, 20, ~
## $ cls_level   <fct> upper, upper, upper, upper, upper, upper, upper, upper, ~
## $ cls_profs   <fct> single, single, single, single, multiple, multiple, mult~
## $ cls_credits <fct> multi credit, multi credit, multi credit, multi credit, ~
## $ bty_f1lower <int> 5, 5, 5, 5, 4, 4, 4, 5, 5, 2, 2, 2, 2, 2, 2, 2, 2, 7, 7,~
## $ bty_f1upper <int> 7, 7, 7, 7, 4, 4, 4, 2, 2, 5, 5, 5, 5, 5, 5, 5, 5, 9, 9,~
## $ bty_f2upper <int> 6, 6, 6, 6, 2, 2, 2, 5, 5, 4, 4, 4, 4, 4, 4, 4, 4, 9, 9,~
## $ bty_m1lower <int> 2, 2, 2, 2, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 3, 3, 3, 7, 7,~
## $ bty_m1upper <int> 4, 4, 4, 4, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 3, 6, 6,~
## $ bty_m2upper <int> 6, 6, 6, 6, 3, 3, 3, 3, 3, 2, 2, 2, 2, 2, 2, 2, 2, 6, 6,~
## $ bty_avg     <dbl> 5.000, 5.000, 5.000, 5.000, 3.000, 3.000, 3.000, 3.333, ~
## $ pic_outfit  <fct> not formal, not formal, not formal, not formal, not form~
## $ pic_color   <fct> color, color, color, color, color, color, color, color, ~
```

We have observations on 21 different variables, some categorical and some numerical. The meaning of each variable can be found by bringing up the help file:

```
?evals
```

Exploring the data

1. Is this an observational study or an experiment? The original research question posed in the paper is whether beauty leads directly to the differences in course evaluations. Given the study design, is it possible to answer this question as it is phrased? If not, rephrase the question.

This is an observational study, there is no control group.

Beauty is in the eye of the beholder (highly subjective) and there is no control group against which to compare results. In addition, only six students rated the professors on their physical appearance so the sample size is small.

Therefore, it is not possible to answer the question as posed because there may be other factors that contribute to teacher evaluations apart from a teacher's perceived attractiveness. Perhaps a better phrasing might be 'whether there is a relationship between a teacher's perceived attractiveness and course evaluations'.

2. Describe the distribution of `score`. Is the distribution skewed? What does that tell you about how students rate courses? Is this what you expected to see? Why, or why not?

```
hist(evals$score, xlab = 'Score')
```

Histogram of evals$score
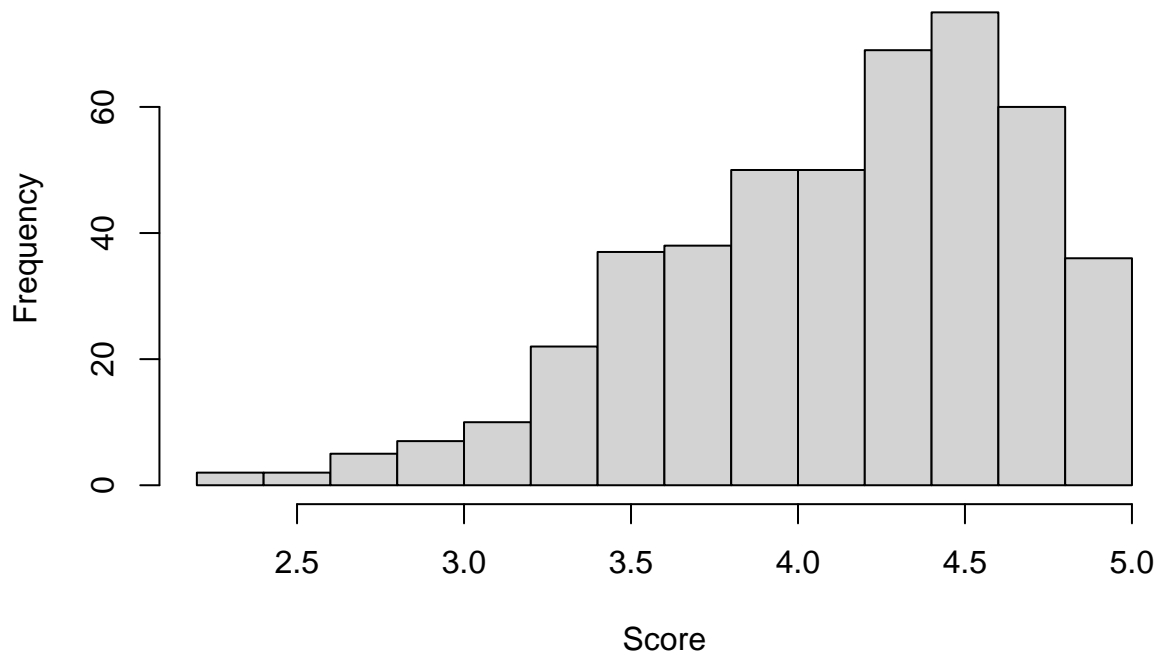
The distribution for `score` is left-skewed, meaning more students rated their professors more positively. This is somewhat surprising as one would expect the distribution to be normal. The skew might be due to other factors such as the selected departments, course levels, etc.

Although this seems to be true regardless of `rank`, `teaching` professors have a shorter left-tail (teaching professors were, in general, rated higher than tenured and tenure track professors).

```
ggplot(data = evals, aes(x = score, fill = rank)) + geom_histogram()
```
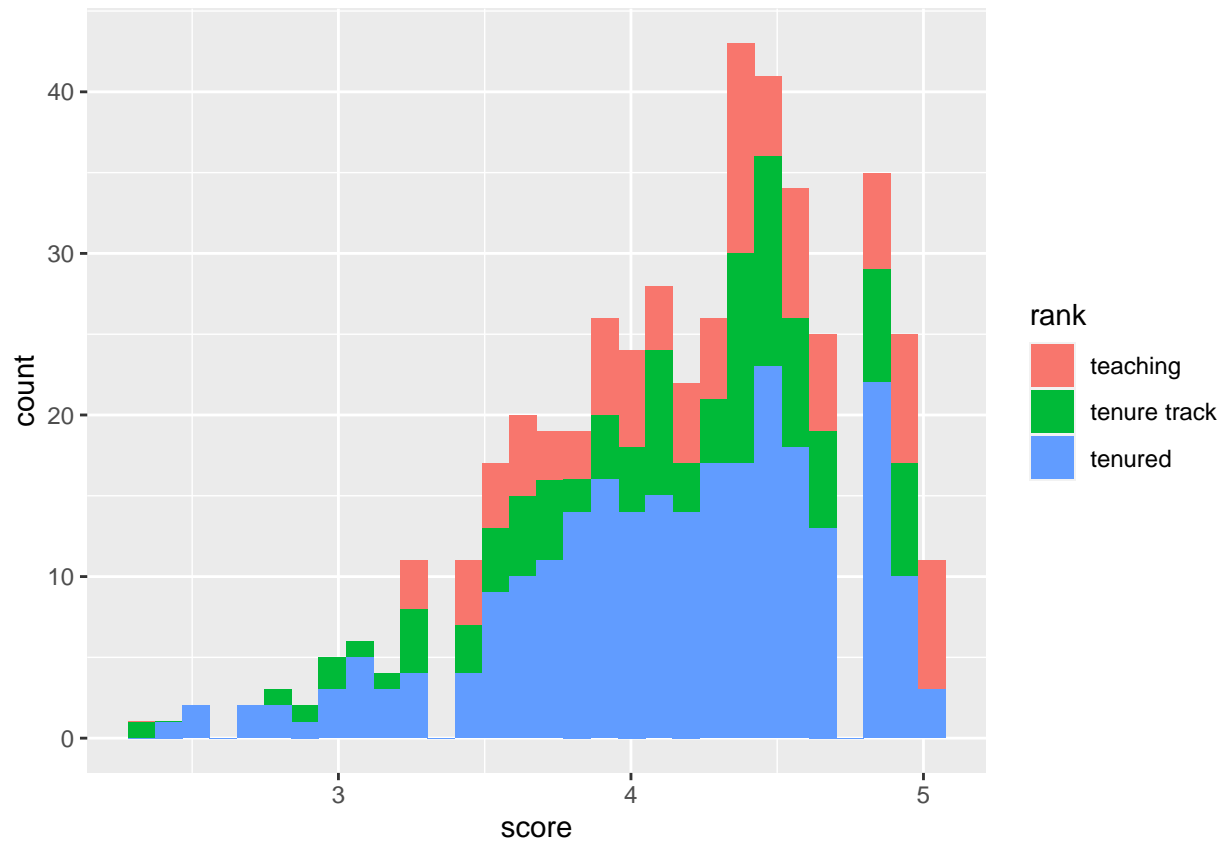
3. Excluding `score`, select two other variables and describe their relationship with each other using an appropriate visualization.

```
boxplot(evals$bty_avg ~ evals$pic_color, main = "Boxplot of Beauty Average Score by Pic Color", ylab = "
```

Boxplot of Beauty Average Score by Pic Color
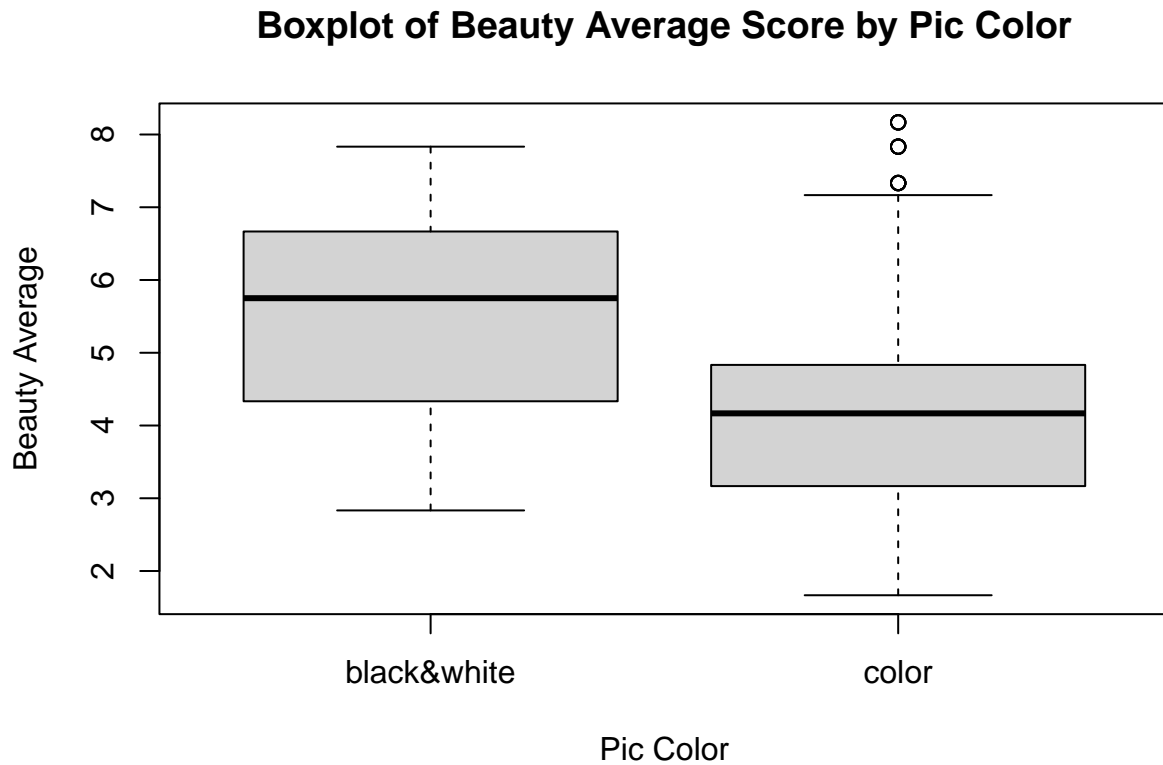
Teachers with Black&White photos received higher Beauty Average scores (3-8) than those with Color photos (1.5-5).

Simple linear regression

The fundamental phenomenon suggested by the study is that better looking teachers are evaluated more favorably. Let's create a scatterplot to see if this appears to be the case:

```
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_point()
```
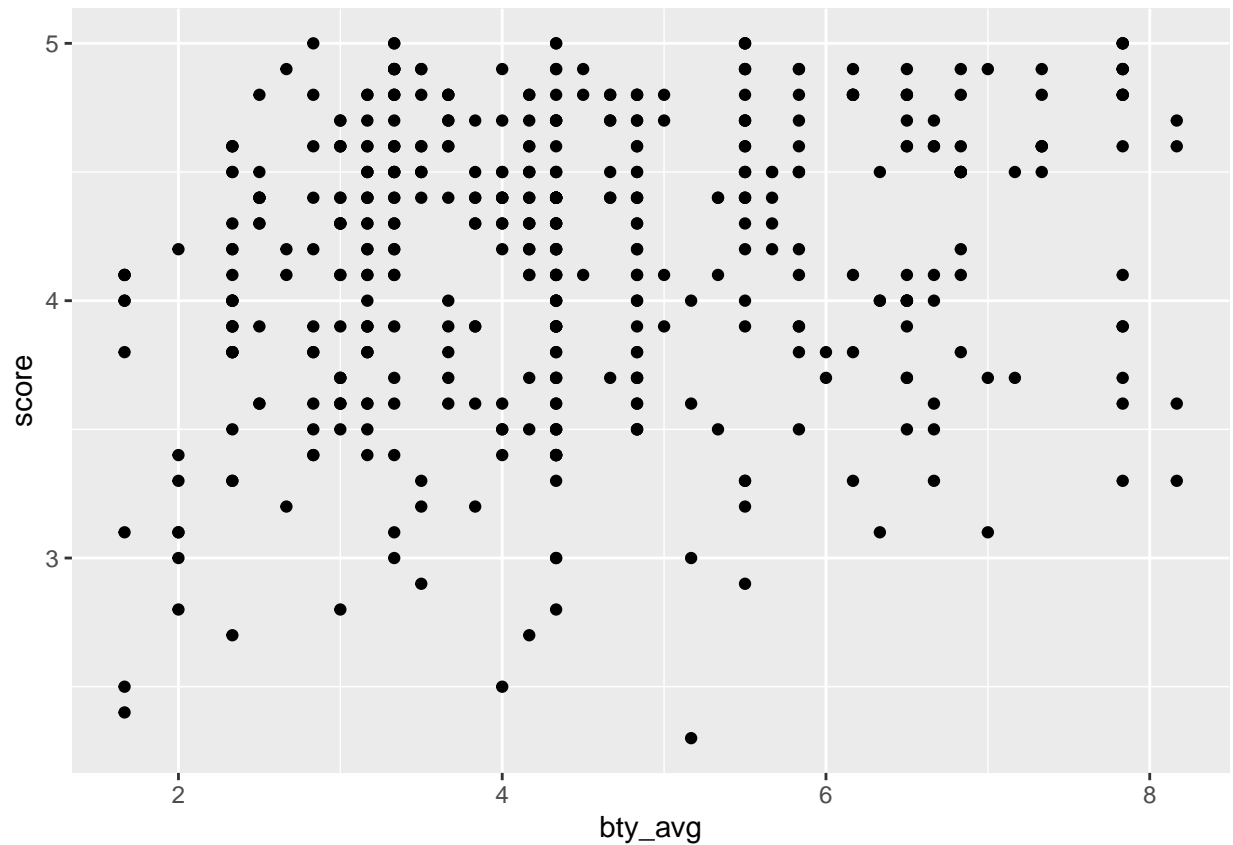
Before you draw conclusions about the trend, compare the number of observations in the data frame with the approximate number of points on the scatterplot. Is anything awry?

There are 463 observations and the scatterplot shows far fewer.

4. Replot the scatterplot, but this time use `geom_jitter` as your layer. What was misleading about the initial scatterplot?

```
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_jitter()
```
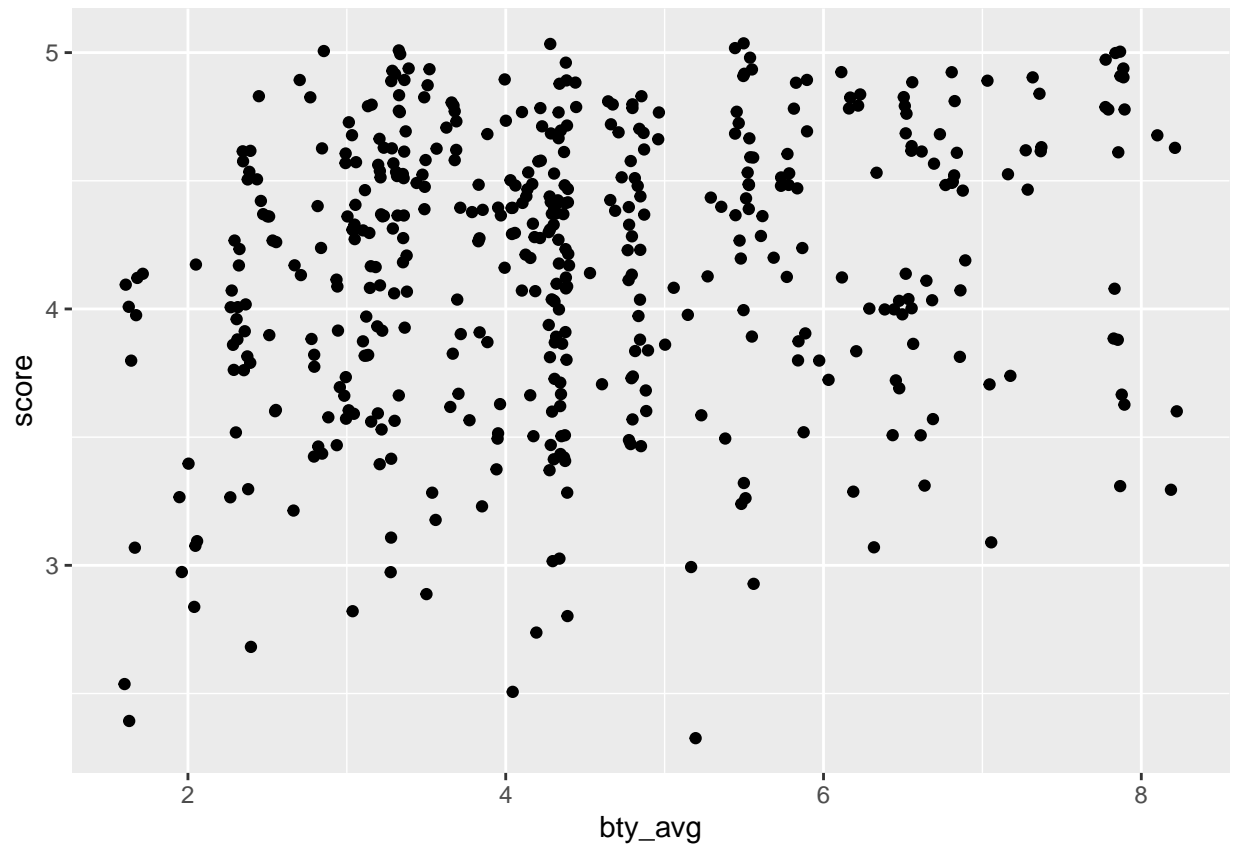
There was overplotting in the initial scatterplot - i.e. multiple points mapped to same coordinates.

5. Let's see if the apparent trend in the plot is something more than natural variation. Fit a linear model called `m_bty` to predict average professor score by average beauty rating. Write out the equation for the linear model and interpret the slope. Is average beauty score a statistically significant predictor? Does it appear to be a practically significant predictor?

Add the line of the bet fit model to your plot using the following:

```
lm_bty <- lm(evals$score ~ evals$bty_avg)
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_jitter() +
  geom_smooth(method = "lm")
```
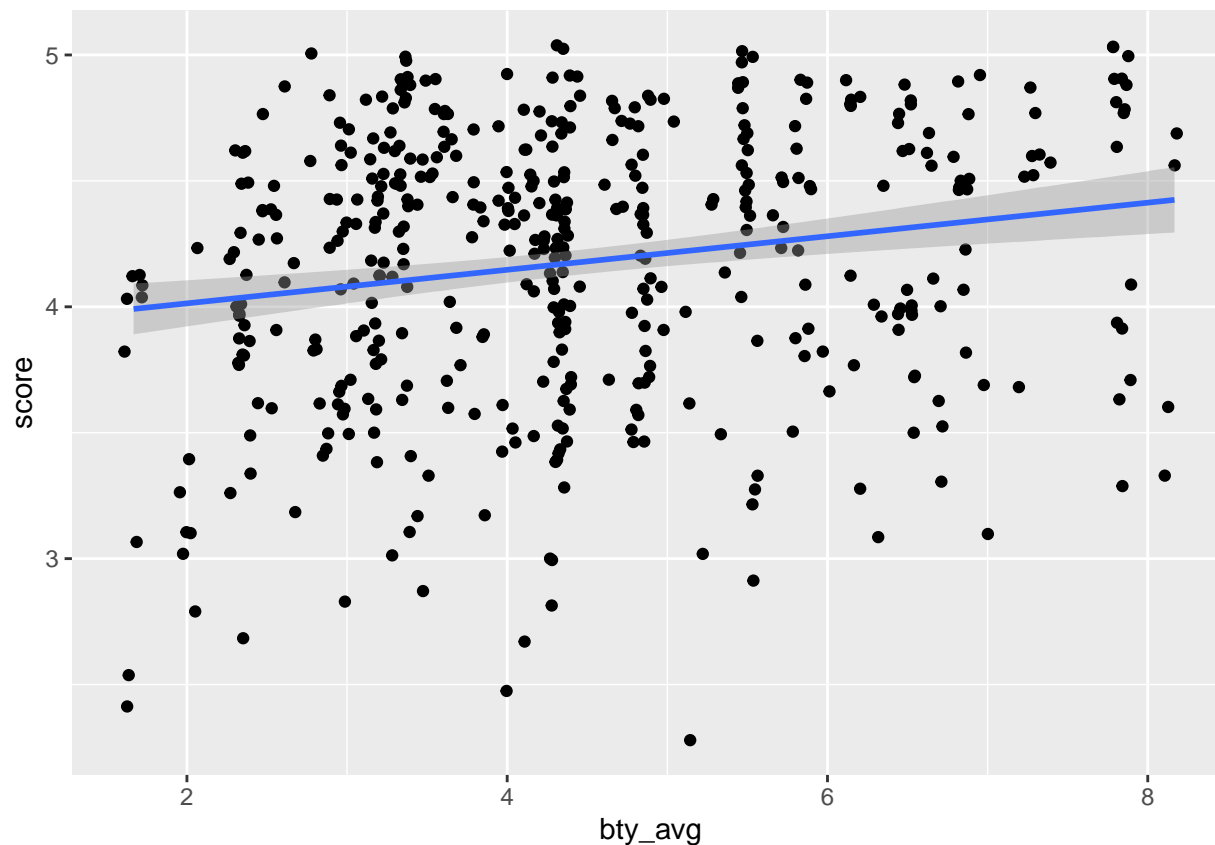
```
summary(lm_bty)
```

```
##
## Call:
## lm(formula = evals$score ~ evals$bty_avg)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.9246 -0.3690  0.1420  0.3977  0.9309
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.88034    0.07614   50.96  < 2e-16 ***
## evals$bty_avg  0.06664    0.01629    4.09 5.08e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5348 on 461 degrees of freedom
## Multiple R-squared:  0.03502,    Adjusted R-squared:  0.03293
## F-statistic: 16.73 on 1 and 461 DF,  p-value: 5.083e-05
```

The blue line is the model. The shaded gray area around the line tells you about the variability you might expect in your predictions. To turn that off, use `se = FALSE`.

8

```
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_jitter() +
  geom_smooth(method = "lm", se = FALSE)
```
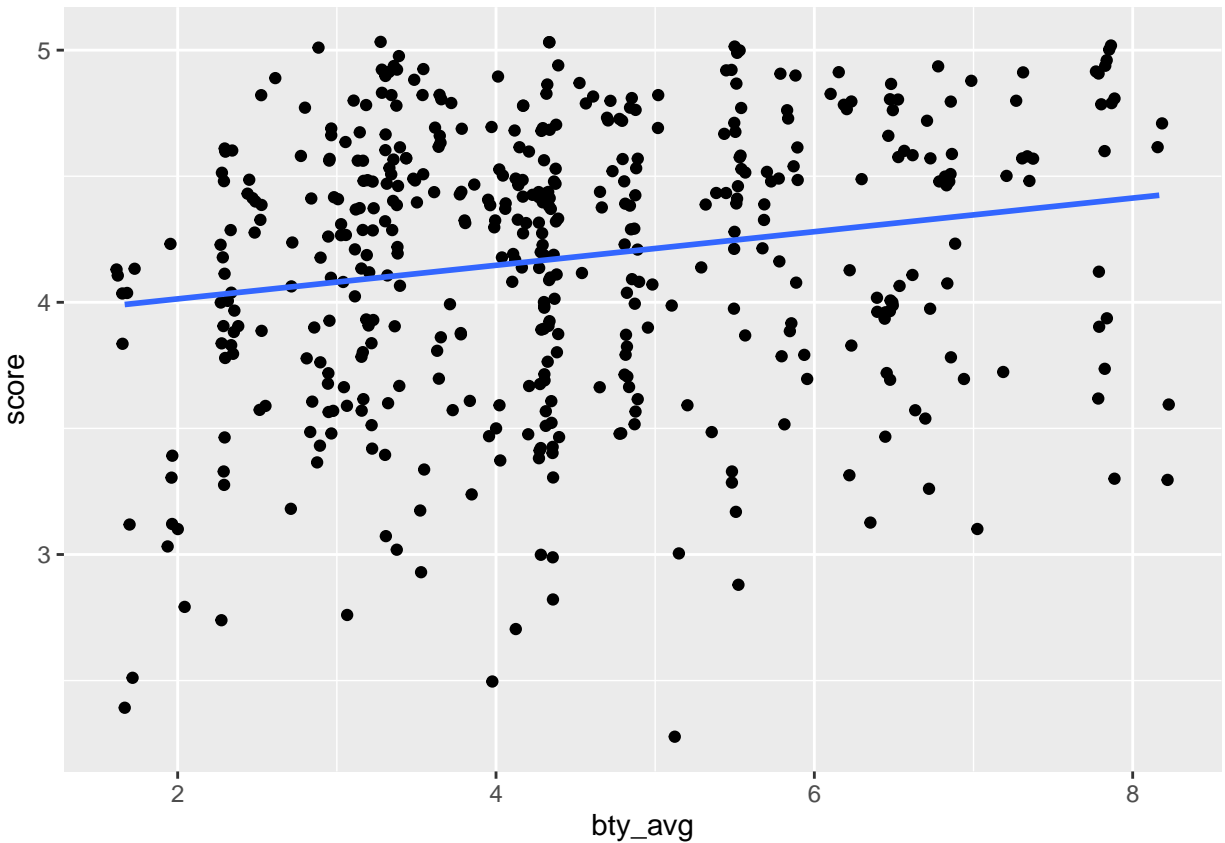

6. Use residual plots to evaluate whether the conditions of least squares regression are reasonable. Provide plots and comments for each one (see the Simple Regression Lab for a reminder of how to make these).

```
score_bty_avg = lm(evals$score ~ evals$bty_avg)
ggplot(data = score_bty_avg, aes(x = .fitted, y = .resid)) +
  geom_point() +
  geom_hline(yintercept = 0, linetype = "dashed") +
  xlab("Fitted values") +
  ylab("Residuals")
```
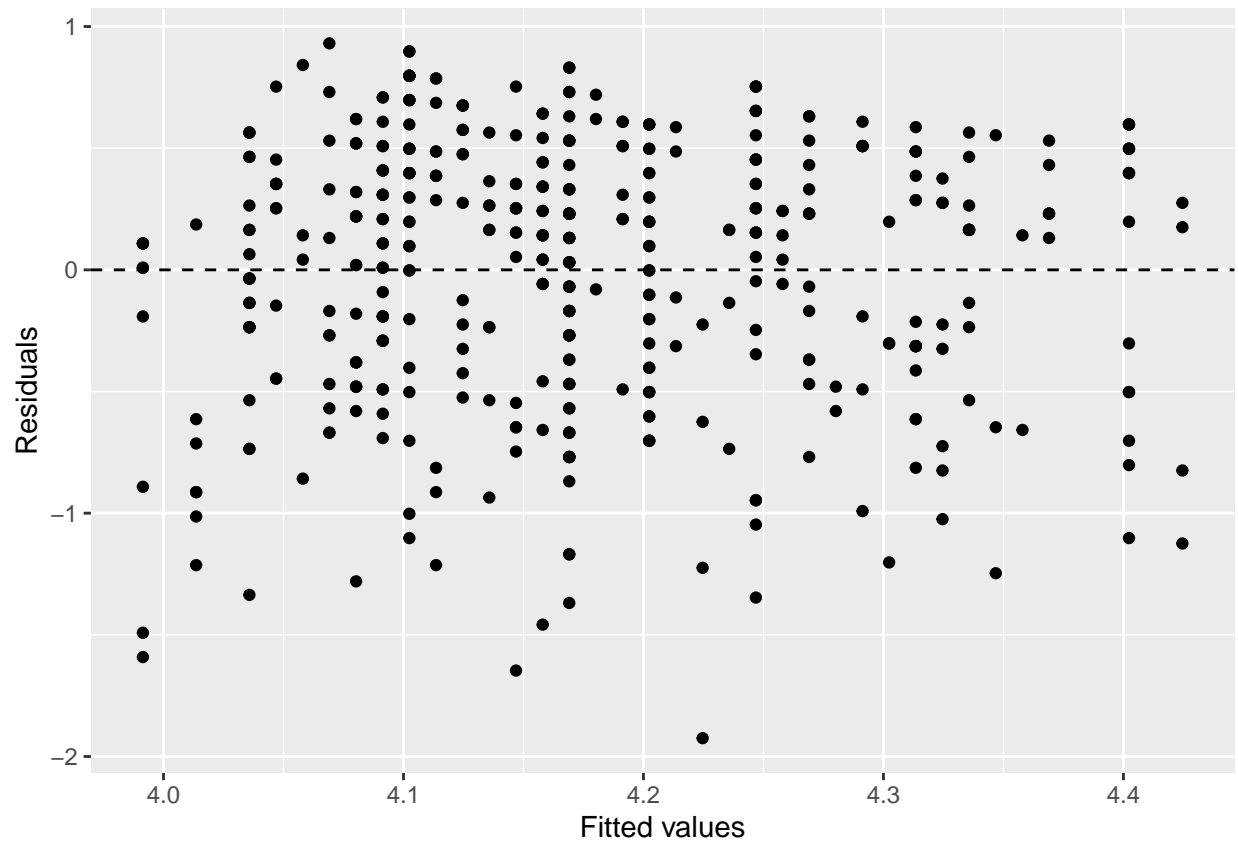
Fitted values are evenly distributed above and below the zero line. This shows a **liner relationship** between the two variables.

```
ggplot(data = score_bty_avg, aes(sample = .resid)) +
  stat_qq()
```
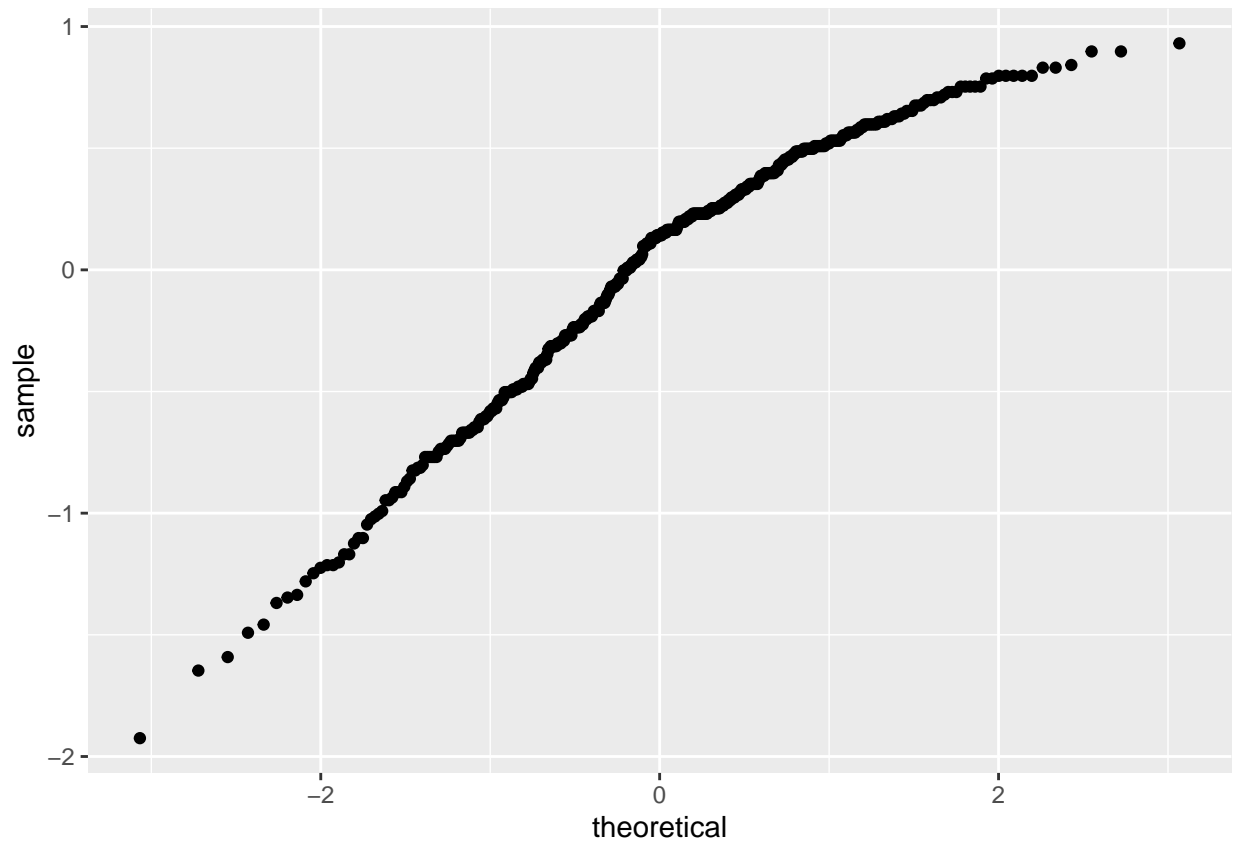
Residuals appear to be **nearly normal** with a left-skew.

Multiple linear regression

The data set contains several variables on the beauty score of the professor: individual ratings from each of the six students who were asked to score the physical appearance of the professors and the average of these six scores. Let's take a look at the relationship between one of these scores and the average beauty score.

```
ggplot(data = evals, aes(x = bty_f1lower, y = bty_avg)) +
  geom_point()
```
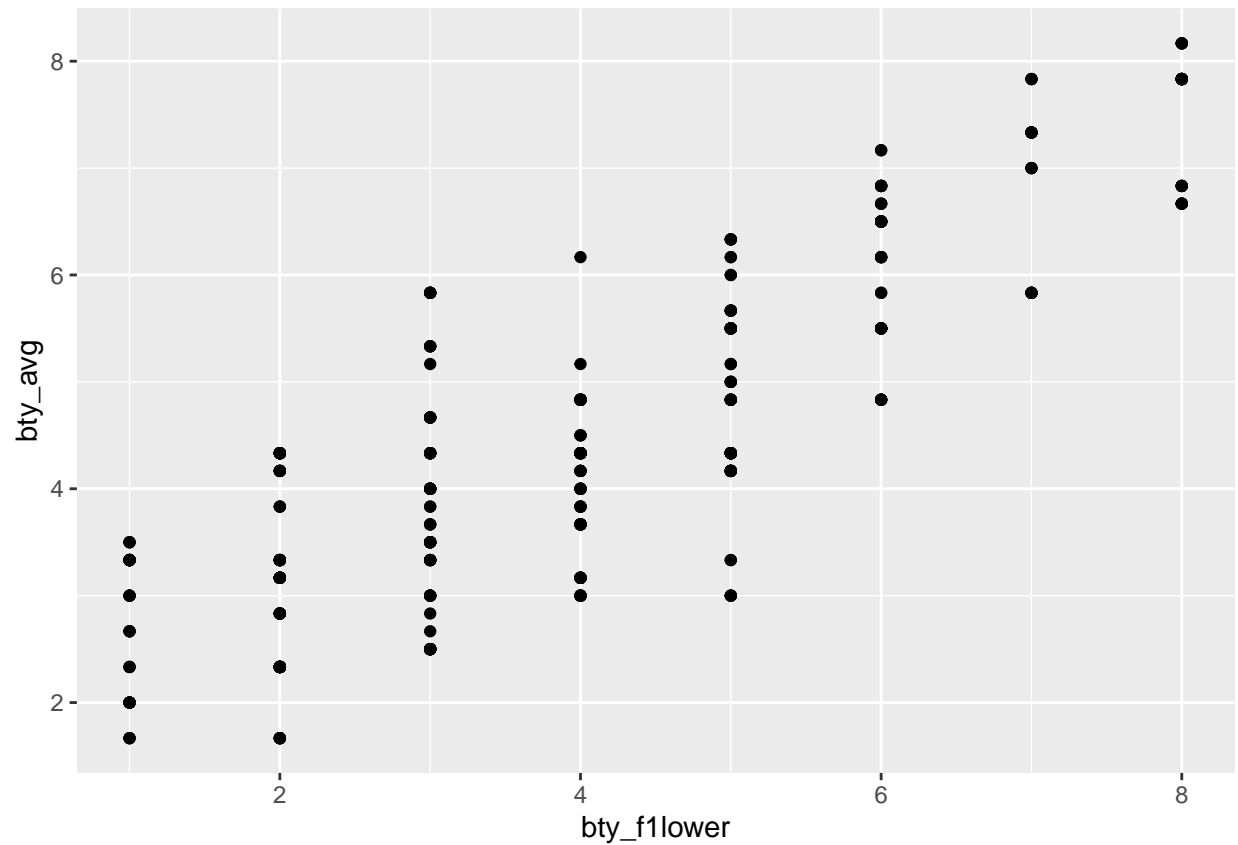
```
evals %>%
  summarise(cor(bty_avg, bty_f1lower))
```

```
## # A tibble: 1 x 1
##   `cor(bty_avg, bty_f1lower)`
##                         <dbl>
## 1                       0.844
```

As expected, the relationship is quite strong—after all, the average score is calculated using the individual scores. You can actually look at the relationships between all beauty variables (columns 13 through 19) using the following command:

```
evals %>%
  select(contains("bty")) %>%
  ggpairs()
```
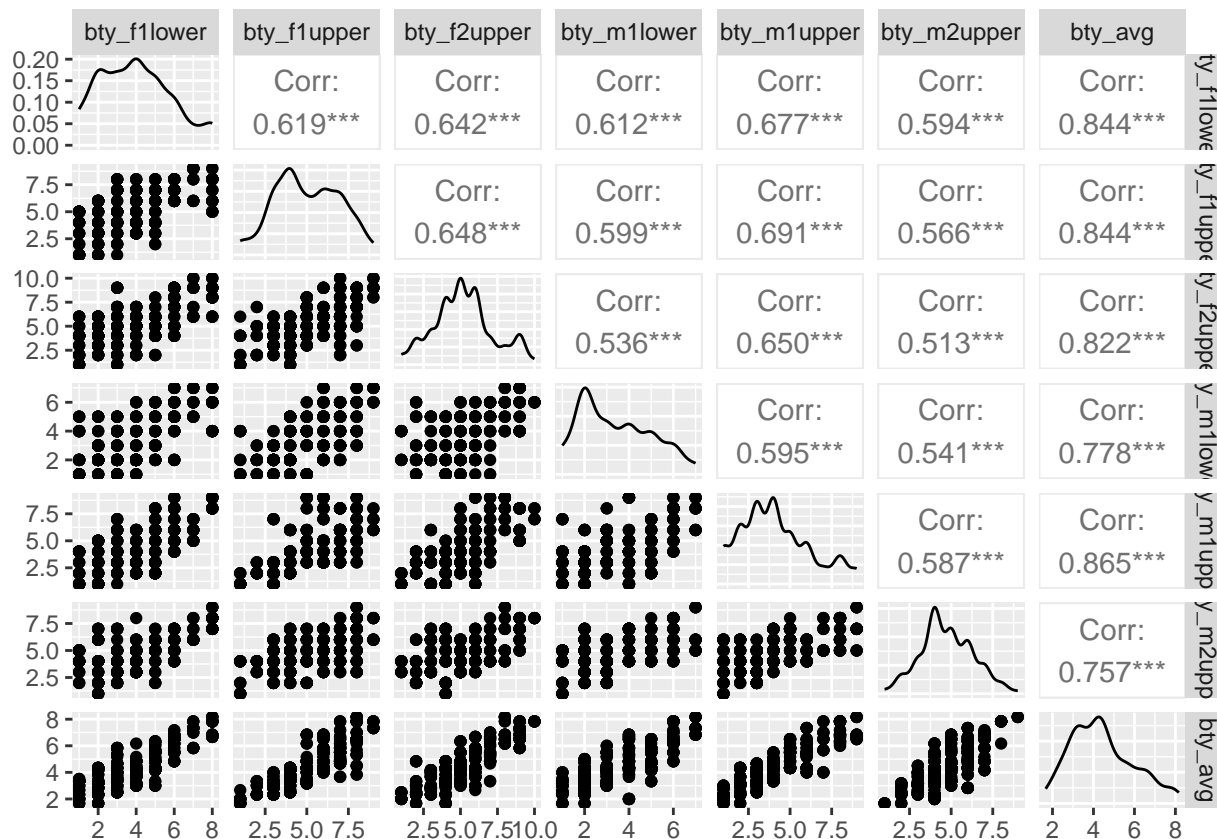
These variables are collinear (correlated), and adding more than one of these variables to the model would not add much value to the model. In this application and with these highly-correlated predictors, it is reasonable to use the average beauty score as the single representative of these variables.

In order to see if beauty is still a significant predictor of professor score after you've accounted for the professor's gender, you can add the gender term into the model.

```
m_bty_gen <- lm(score ~ bty_avg + gender, data = evals)
summary(m_bty_gen)
```

```
##
## Call:
## lm(formula = score ~ bty_avg + gender, data = evals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8305 -0.3625  0.1055  0.4213  0.9314
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.74734    0.08466  44.266  < 2e-16 ***
## bty_avg      0.07416    0.01625   4.563 6.48e-06 ***
## gendermale   0.17239    0.05022   3.433 0.000652 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.5287 on 460 degrees of freedom
## Multiple R-squared:  0.05912,    Adjusted R-squared:  0.05503
## F-statistic: 14.45 on 2 and 460 DF,  p-value: 8.177e-07
```

7. P-values and parameter estimates should only be trusted if the conditions for the regression are reasonable. Verify that the conditions for this model are reasonable using diagnostic plots.

```
ggplot(data = m_bty_gen, aes(sample = .resid)) +
  stat_qq() + stat_qq_line()
```
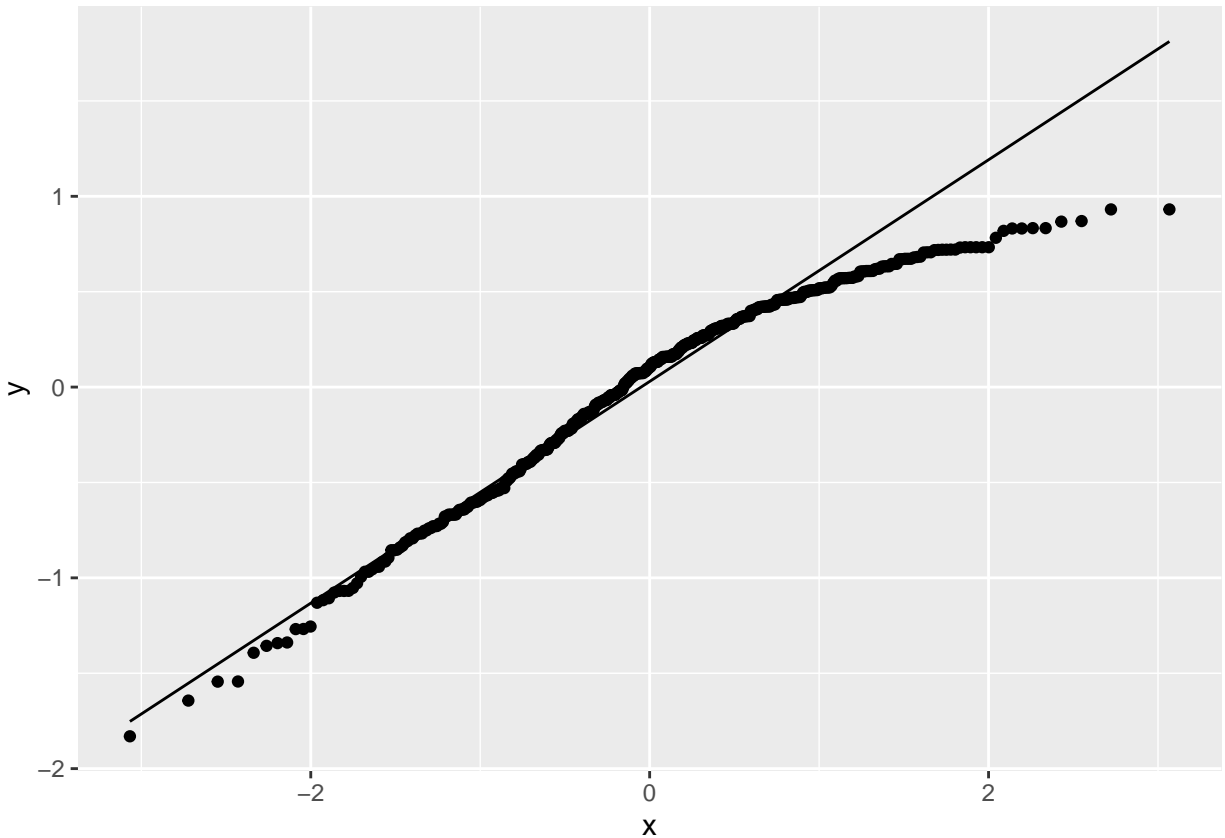

Although the points drift off below the normal line towards the upper end, most of the data falls along the normal line.

```
plot(m_bty_gen$residuals)
abline(h = 0, lty = 3)
```

The residuals are spread evenly along the zero line. Conditions for this model have been met.

8. Is `bty_avg` still a significant predictor of `score`? Has the addition of `gender` to the model changed the parameter estimate for `bty_avg`?

```
summary(m_bty_gen)
```

```
##
## Call:
## lm(formula = score ~ bty_avg + gender, data = evals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8305 -0.3625  0.1055  0.4213  0.9314
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  3.74734    0.08466  44.266  < 2e-16 ***
## bty_avg      0.07416    0.01625   4.563 6.48e-06 ***
## gendermale   0.17239    0.05022   3.433 0.000652 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5287 on 460 degrees of freedom
## Multiple R-squared:  0.05912,    Adjusted R-squared:  0.05503
## F-statistic: 14.45 on 2 and 460 DF,  p-value: 8.177e-07
```
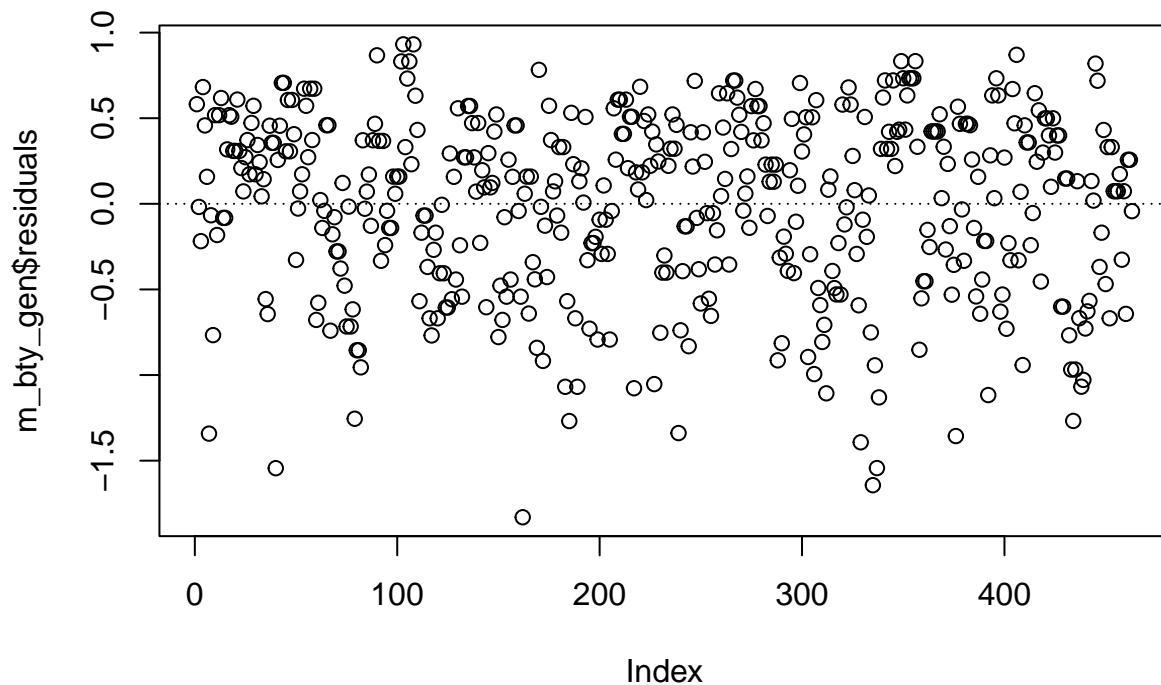
`bty_avg` is still a significant predictor of score and adding the gender variable to the model has increased the poarameter estimate from 0.06664 to 0.07416. However, the R-square is very low for this model.

Note that the estimate for `gender` is now called `gendermale`. You'll see this name change whenever you introduce a categorical variable. The reason is that R recodes `gender` from having the values of `male` and `female` to being an indicator variable called `gendermale` that takes a value of 0 for female professors and a value of 1 for male professors. (Such variables are often referred to as "dummy" variables.)

As a result, for female professors, the parameter estimate is multiplied by zero, leaving the intercept and slope form familiar from simple regression.

$$\widehat{score} = \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg + \hat{\beta}_2 \times (0)$$
$$= \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg$$

```
ggplot(data = evals, aes(x = bty_avg, y = score, color = pic_color)) +
 geom_smooth(method = "lm", formula = y ~ x, se = FALSE)
```
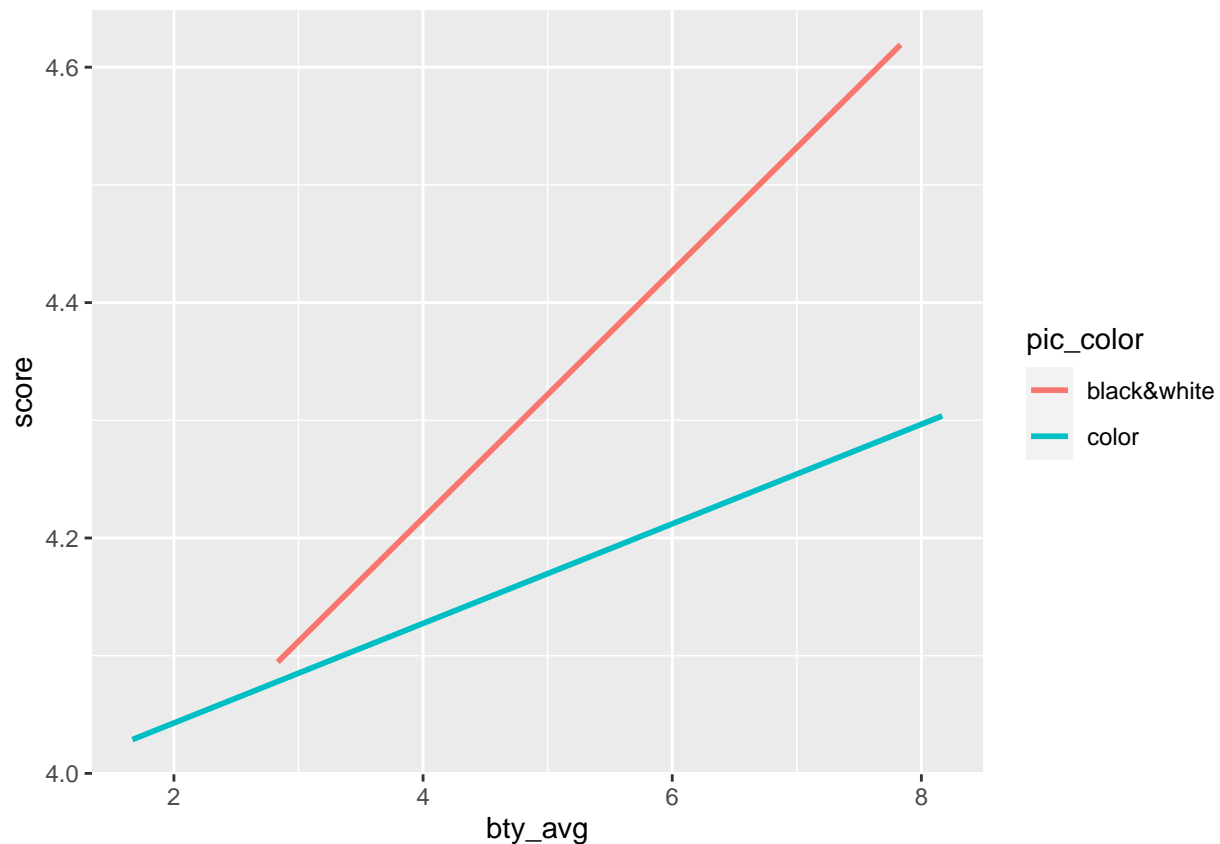

9. What is the equation of the line corresponding to those with color pictures? (*Hint:* For those with color pictures, the parameter estimate is multiplied by 1.) For two professors who received the same beauty rating, which color picture tends to have the higher course evaluation score?

$$\widehat{score} = \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg + \hat{\beta}_2 \times (1)$$
$$= \hat{\beta}_0 + \hat{\beta}_1 \times bty\_avg + \hat{\beta}_2$$

The decision to call the indicator variable `gendermale` instead of `genderfemale` has no deeper meaning. R simply codes the category that comes first alphabetically as a 0. (You can change the reference level of a categorical variable, which is the level that is coded as a 0, using the`relevel()` function. Use `?relevel` to learn more.)

10. Create a new model called `m_bty_rank` with `gender` removed and `rank` added in. How does R appear to handle categorical variables that have more than two levels? Note that the rank variable has three levels: `teaching`, `tenure track`, `tenured`.

```
m_bty_rank = lm(score ~ bty_avg + rank, data = evals)
summary(m_bty_rank)
```

```
##
## Call:
## lm(formula = score ~ bty_avg + rank, data = evals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.8713 -0.3642  0.1489  0.4103  0.9525
##
## Coefficients:
##                   Estimate Std. Error t value Pr(>|t|)
## (Intercept)        3.98155    0.09078  43.860  < 2e-16 ***
## bty_avg            0.06783    0.01655   4.098 4.92e-05 ***
## ranktenure track  -0.16070    0.07395  -2.173   0.0303 *
## ranktenured       -0.12623    0.06266  -2.014   0.0445 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5328 on 459 degrees of freedom
## Multiple R-squared:  0.04652,    Adjusted R-squared:  0.04029
## F-statistic: 7.465 on 3 and 459 DF,  p-value: 6.88e-05
```

Since rank variable has three levels (teaching, tenure track and tenured), R has added another line into the regression summary to account for it. R leaves out one level but mentions the rest as variables.

The interpretation of the coefficients in multiple regression is slightly different from that of simple regression. The estimate for `bty_avg` reflects how much higher a group of professors is expected to score if they have a beauty rating that is one point higher *while holding all other variables constant.* In this case, that translates into considering only professors of the same rank with `bty_avg` scores that are one point apart.

The search for the best model

We will start with a full model that predicts professor score based on rank, gender, ethnicity, language of the university where they got their degree, age, proportion of students that filled out evaluations, class size, course level, number of professors, number of credits, average beauty rating, outfit, and picture color.

11. Which variable would you expect to have the highest p-value in this model? Why? *Hint:* Think about which variable would you expect to not have any association with the professor score.

I would expect the number of credits to not have any association with the professor score.

Let's run the model...

```
m_full <- lm(score ~ rank + gender + ethnicity + language + age + cls_perc_eval
             + cls_students + cls_level + cls_profs + cls_credits + bty_avg
             + pic_outfit + pic_color, data = evals)
summary(m_full)
```

```
##
## Call:
## lm(formula = score ~ rank + gender + ethnicity + language + age +
##     cls_perc_eval + cls_students + cls_level + cls_profs + cls_credits +
##     bty_avg + pic_outfit + pic_color, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.77397 -0.32432  0.09067  0.35183  0.95036
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             4.0952141  0.2905277  14.096   < 2e-16 ***
## ranktenure track       -0.1475932  0.0820671  -1.798  0.07278 .
## ranktenured            -0.0973378  0.0663296  -1.467  0.14295
## gendermale              0.2109481  0.0518230   4.071 5.54e-05 ***
## ethnicitynot minority   0.1234929  0.0786273   1.571  0.11698
## languagenon-english    -0.2298112  0.1113754  -2.063  0.03965 *
## age                    -0.0090072  0.0031359  -2.872  0.00427 **
## cls_perc_eval           0.0053272  0.0015393   3.461  0.00059 ***
## cls_students            0.0004546  0.0003774   1.205  0.22896
## cls_levelupper          0.0605140  0.0575617   1.051  0.29369
## cls_profssingle        -0.0146619  0.0519885  -0.282  0.77806
## cls_creditsone credit   0.5020432  0.1159388   4.330 1.84e-05 ***
## bty_avg                 0.0400333  0.0175064   2.287  0.02267 *
## pic_outfitnot formal   -0.1126817  0.0738800  -1.525  0.12792
## pic_colorcolor         -0.2172630  0.0715021  -3.039  0.00252 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.498 on 448 degrees of freedom
## Multiple R-squared:  0.1871, Adjusted R-squared:  0.1617
## F-statistic: 7.366 on 14 and 448 DF,  p-value: 6.552e-14
```

12. Check your suspicions from the previous exercise. Include the model output in your response.

My guess was off the mark. In fact the variable `cls_profsingle` (number of professors teaching the course: single, multiple) has the highest p-value in this model.

13. Interpret the coefficient associated with the ethnicity variable.

The `ethnicity` variable increases the `score` by 0.1234929, all other variables being held constant.

14. Drop the variable with the highest p-value and re-fit the model. Did the coefficients and significance of the other explanatory variables change? (One of the things that makes multiple regression interesting is that coefficient estimates depend on the other variables that are included in the model.) If not, what does this say about whether or not the dropped variable was collinear with the other explanatory variables?

```
m_full <- lm(score ~ rank + gender + ethnicity + language + age + cls_perc_eval
             + cls_students + cls_level +  cls_credits + bty_avg
             + pic_outfit + pic_color, data = evals)
summary(m_full)
```

```
##
## Call:
## lm(formula = score ~ rank + gender + ethnicity + language + age +
##     cls_perc_eval + cls_students + cls_level + cls_credits +
##     bty_avg + pic_outfit + pic_color, data = evals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7836 -0.3257  0.0859  0.3513  0.9551
##
## Coefficients:
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)             4.0872523  0.2888562  14.150  < 2e-16 ***
## ranktenure track       -0.1476746  0.0819824  -1.801 0.072327 .
## ranktenured            -0.0973829  0.0662614  -1.470 0.142349
## gendermale              0.2101231  0.0516873   4.065 5.66e-05 ***
## ethnicitynot minority   0.1274458  0.0772887   1.649 0.099856 .
## languagenon-english    -0.2282894  0.1111305  -2.054 0.040530 *
## age                    -0.0089992  0.0031326  -2.873 0.004262 **
## cls_perc_eval           0.0052888  0.0015317   3.453 0.000607 ***
## cls_students            0.0004687  0.0003737   1.254 0.210384
## cls_levelupper          0.0606374  0.0575010   1.055 0.292200
## cls_creditsone credit   0.5061196  0.1149163   4.404 1.33e-05 ***
## bty_avg                 0.0398629  0.0174780   2.281 0.023032 *
## pic_outfitnot formal   -0.1083227  0.0721711  -1.501 0.134080
## pic_colorcolor         -0.2190527  0.0711469  -3.079 0.002205 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4974 on 449 degrees of freedom
## Multiple R-squared:  0.187,  Adjusted R-squared:  0.1634
## F-statistic: 7.943 on 13 and 449 DF,  p-value: 2.336e-14
```

After dropping `cls_profsingle` the coefficients and significance of the other explanatory variables has changed very slightly. My guess is that the `cls_profsingle` was collinear with the other explanatory variables, and thus dropping it did not change the model significantly.

15. Using backward-selection and p-value as the selection criterion, determine the best model. You do not need to show all steps in your answer, just the output for the final model. Also, write out the linear model for predicting score based on the final model you settle on.

```
lm_backwards = lm(score ~ rank + ethnicity + gender + language + age + cls_perc_eval + cls_credits + bty
summary(lm_backwards)
```

```
##
## Call:
## lm(formula = score ~ rank + ethnicity + gender + language + age +
```

```
##      cls_perc_eval + cls_credits + bty_avg + pic_color, data = evals)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1.7984 -0.3158  0.0832  0.3699  0.9173
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)             3.952938   0.262949  15.033  < 2e-16 ***
## ranktenure track       -0.118607   0.081205  -1.461  0.14482
## ranktenured            -0.052912   0.063940  -0.828  0.40837
## ethnicitynot minority   0.153074   0.076062   2.012  0.04476 *
## gendermale              0.209150   0.051321   4.075 5.43e-05 ***
## languagenon-english    -0.175573   0.106605  -1.647  0.10026
## age                    -0.007946   0.003060  -2.597  0.00972 **
## cls_perc_eval           0.004536   0.001438   3.155  0.00171 **
## cls_creditsone credit   0.476406   0.110992   4.292 2.17e-05 ***
## bty_avg                 0.049660   0.017036   2.915  0.00373 **
## pic_colorcolor         -0.197815   0.067797  -2.918  0.00370 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.4991 on 452 degrees of freedom
## Multiple R-squared:  0.1761, Adjusted R-squared:  0.1578
## F-statistic: 9.658 on 10 and 452 DF,  p-value: 1.168e-14
```
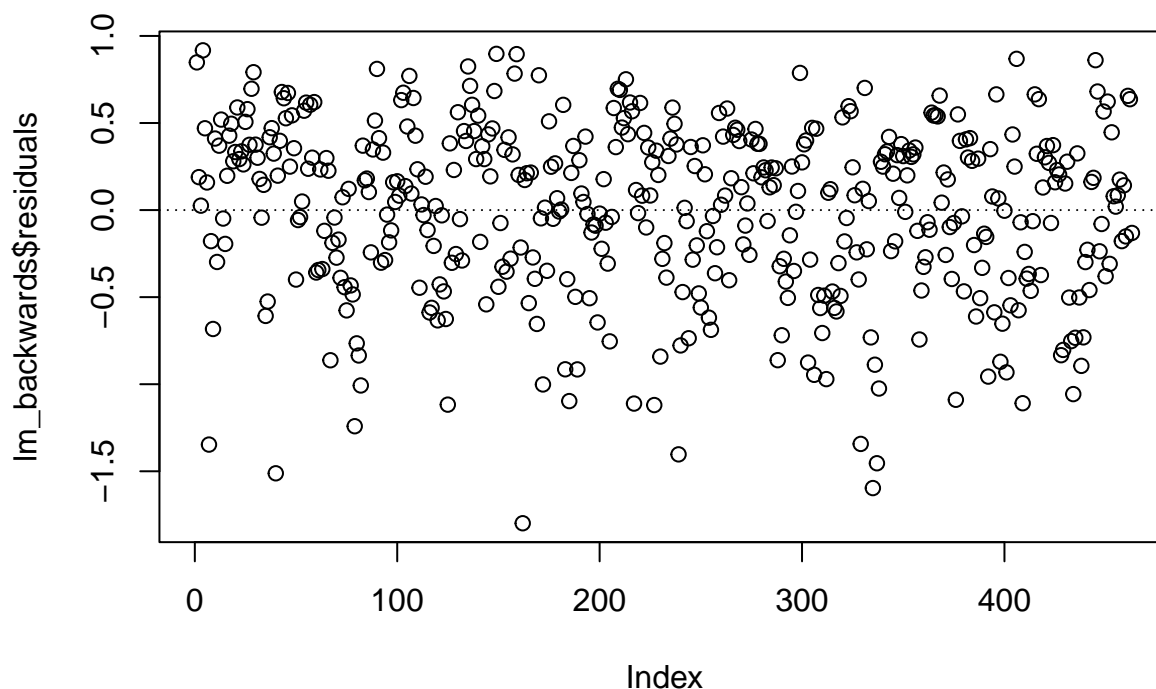
$$\widehat{score} = \hat{\beta}_0 + \hat{\beta}_1 \times rank + \hat{\beta}_2 \times ethnicity + \hat{\beta}_3 \times gender + \hat{\beta}_4 \times language \times \hat{\beta}_5 \times age + \hat{\beta}_6 \times cls\_perc\_eval + \hat{\beta}_7 \times cls\_credi$$

16. Verify that the conditions for this model are reasonable using diagnostic plots.

```
ggplot(data = lm_backwards, aes(sample = .resid)) +
  stat_qq() + stat_qq_line()
```
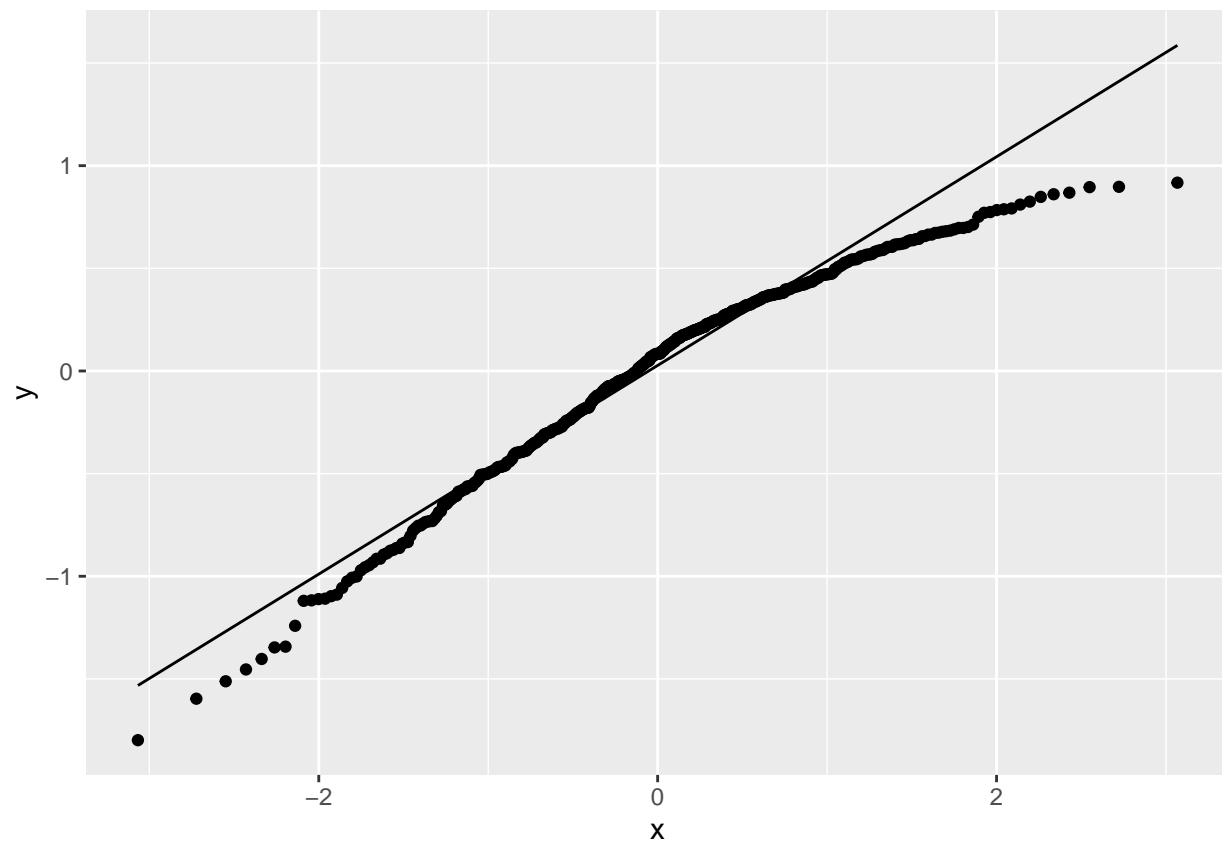
```
plot(lm_backwards$residuals)
abline(h = 0, lty = 3)
```

17. The original paper describes how these data were gathered by taking a sample of professors from the University of Texas at Austin and including all courses that they have taught. Considering that each row represents a course, could this new information have an impact on any of the conditions of linear regression?

Clearly the independence assumption is not longer valid.

18. Based on your final model, describe the characteristics of a professor and course at University of Texas at Austin that would be associated with a high evaluation score.

A high evaluation scoure at UT-Austin would be associated with a professor with a *teaching* rank, non-minority, male, English educated, with a Black&White profile photograph.

19. Would you be comfortable generalizing your conclusions to apply to professors generally (at any university)? Why or why not?

I would not be comfortable generalizing my conclusions to professors generally. In order to make a generalized conclusion I think it would be necessary to take a much larger sample of students from across a broader set of universities (public/private, 2/4 year colleges, etc) reflecting a large cross-section of student population.