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.

```
library(tidyverse)
library(openintro)
library(GGally)
library("gridExtra")
```

You will be using the `ggpairs` function from this package later in the lab.

Creating a reproducible lab report

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)
```

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.

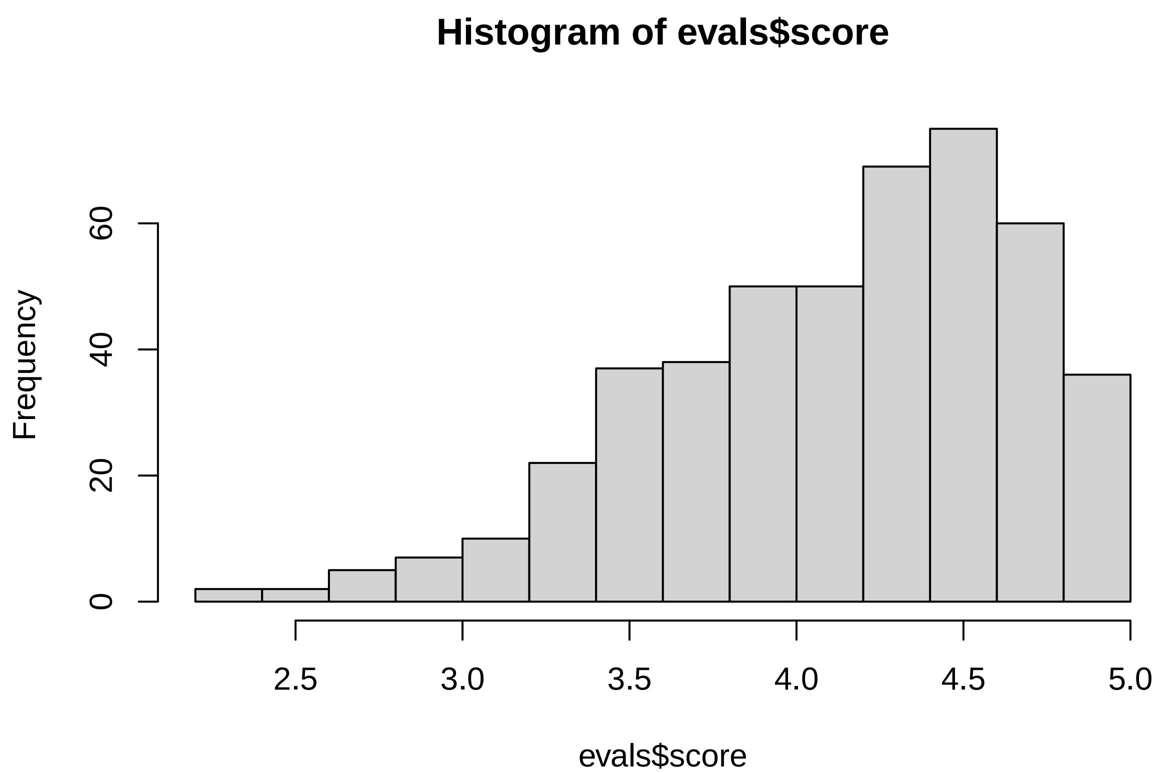
The study is observational.

It is a difficult question to answer since each evaluator will have a different perception of the teachers appearance.

I would rephrase it as "Are student evaluations unduly biased by the physical appearance of the teacher?"

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)
```



Its left skewed. Most of the values are on the high end.

I would think most students appreciate their teachers and give them good grades.

3. Excluding `score`, select two other variables and describe their relationship with each other using an appropriate visualization.

Im curious how the beauty scores were distributed by age and gender.

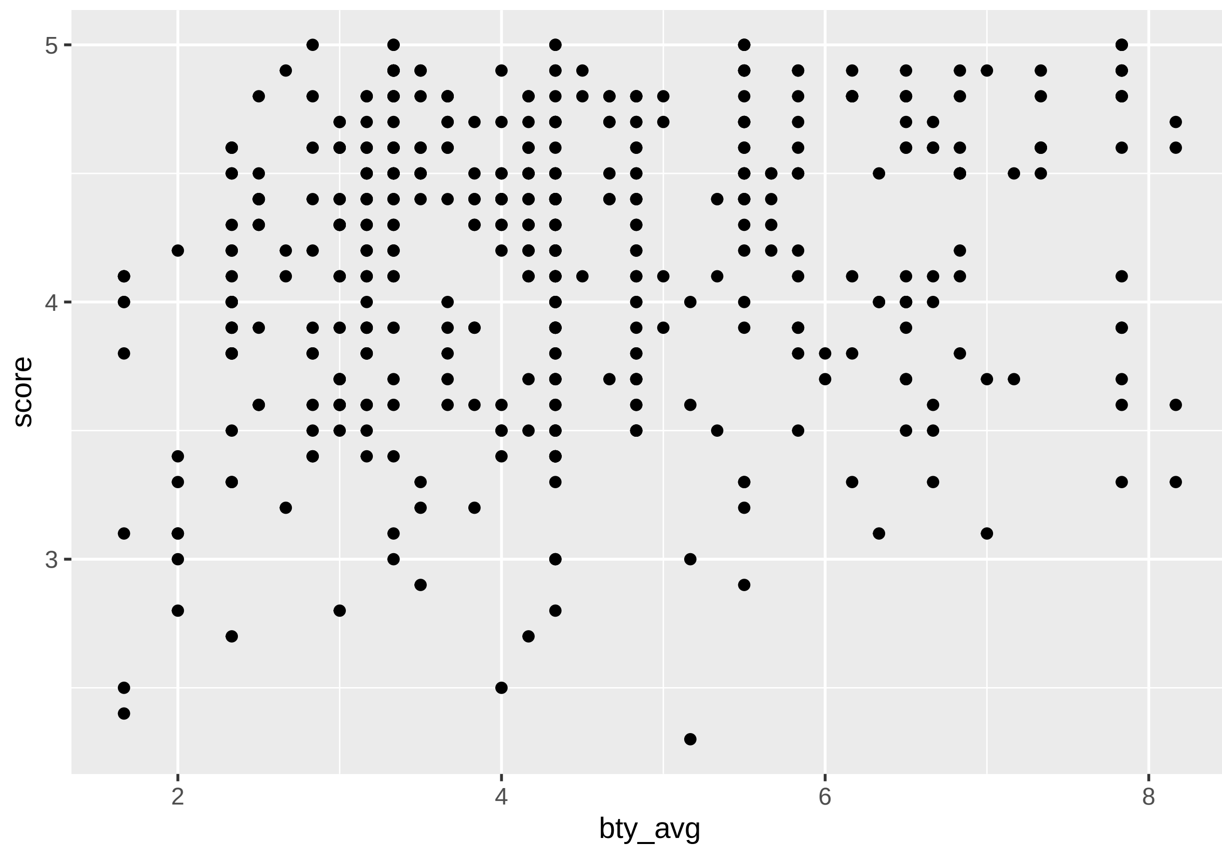
```
ggplot(data = evals, aes(x = age, y = bty_avg, fill = gender)) +
  geom_bar(stat = "identity", position = position_dodge(), alpha = 0.75)  +
  ylim(0,6) +
  labs(x = "\n Age", y = "Score\n", title = "\n Scores by Age and Gender \n") +
  theme(plot.title = element_text(hjust = 0.5),      # center the title
        axis.title.x = element_text(face="bold", colour="red", size = 12),
        axis.title.y = element_text(face="bold", colour="red", size = 12),
        legend.title = element_text(face="bold", colour="blue", size = 10))
```

Age didnt matter that much. This helps me to understand what is meant by beauty.

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:

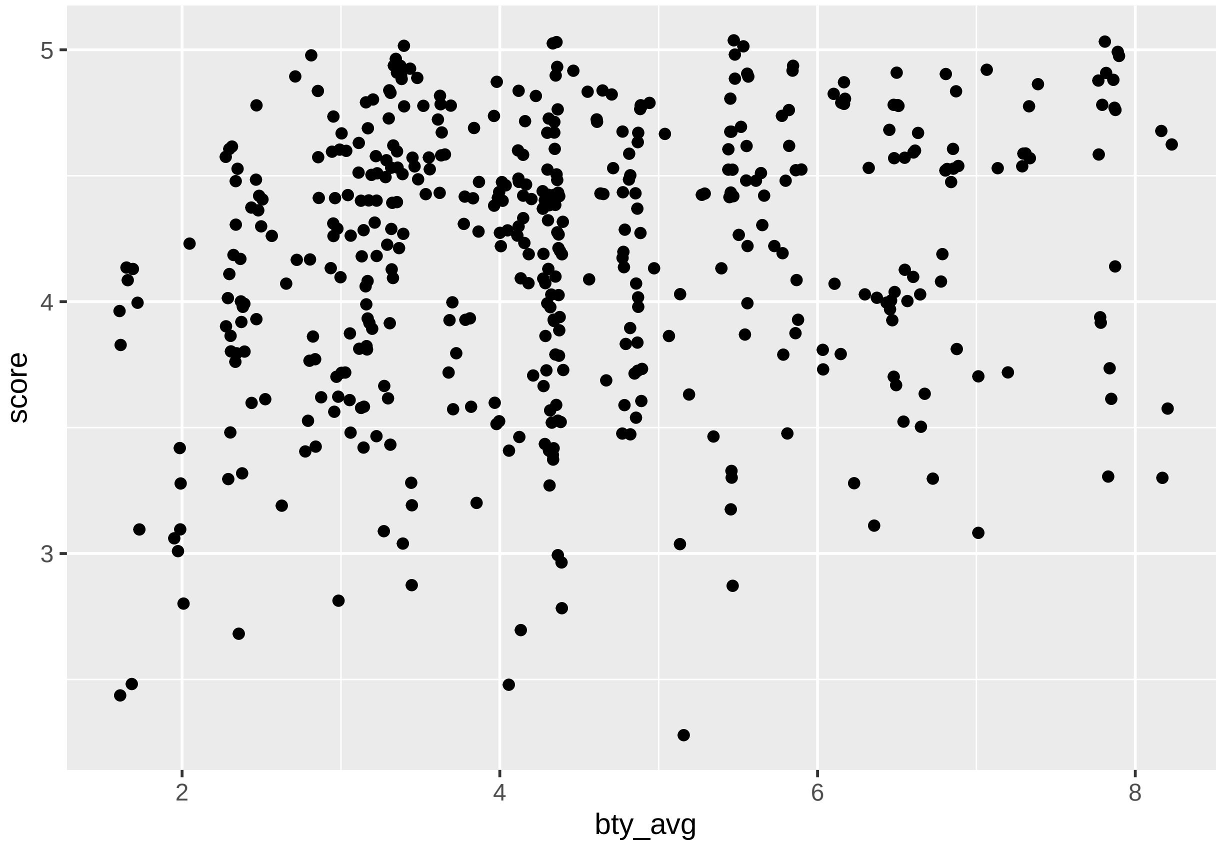
```
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_point()
```



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?

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()
```



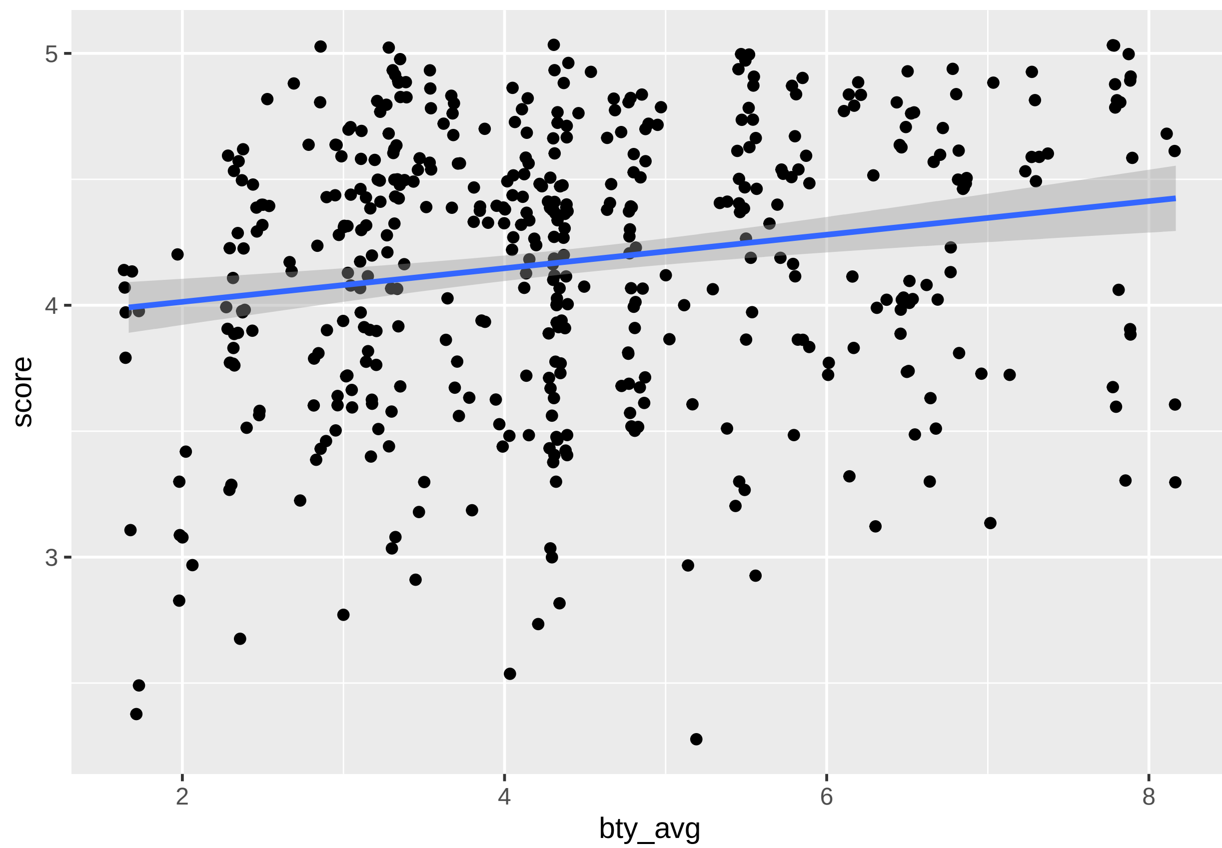
geom jitter enables 3 observations that fall on the same dot to appear as 3 dots close together.

It sacrifices precision for the sake of a better overall view.

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:

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




```
m_bty<-lm(score~bty_avg, data=evals)

sum_m_bty<-summary(m_bty)

sprintf("The r squared is %.3f which is pretty low so its not a great linear relationship.",
        sum_m_bty$r.squared)
```

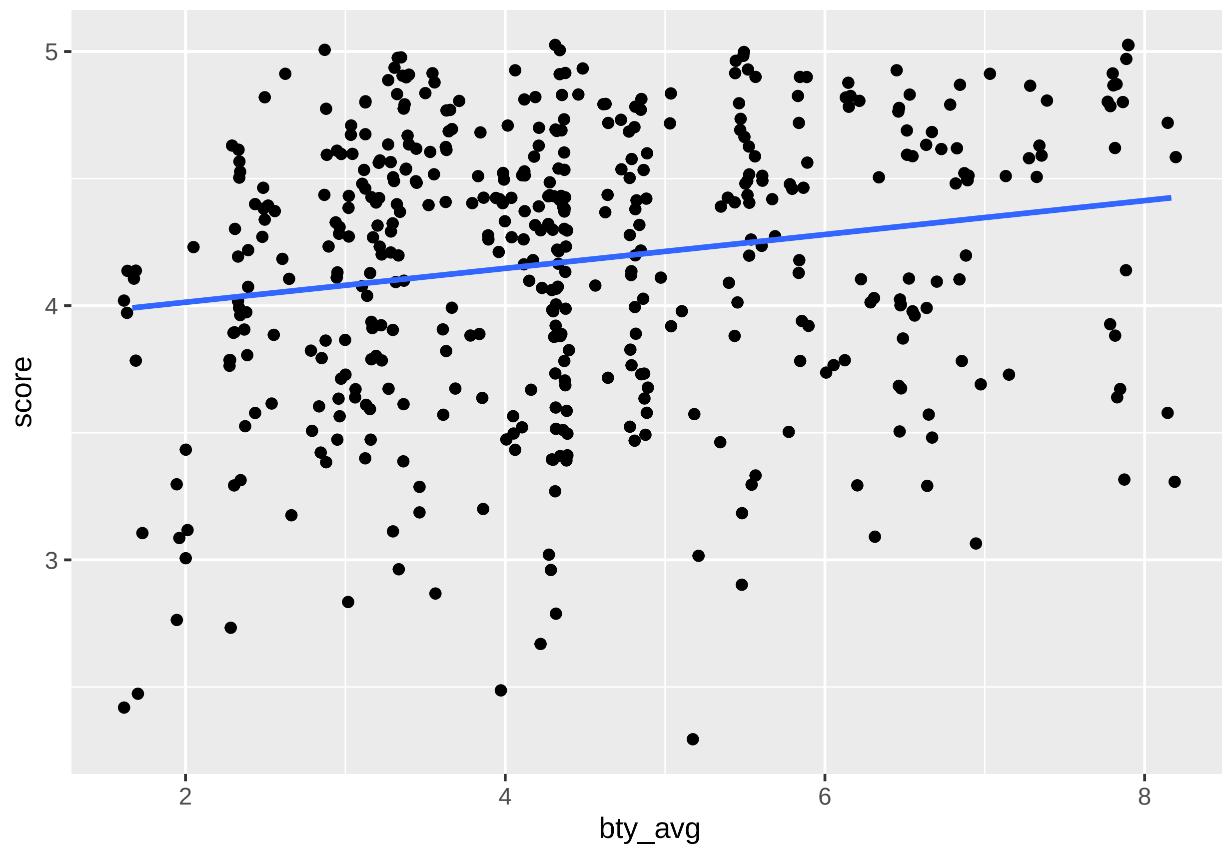
```
## [1] "The r squared is 0.035 which is pretty low so its not a great linear relationship."
```

```
sprintf("The residual standard error is %.3f which is pretty high, again showing for any given bty_avg
        ,sum_m_bty$sigma)
```

```
## [1] "The residual standard error is 0.535 which is pretty high, again showing for any given bty_avg
```

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`.

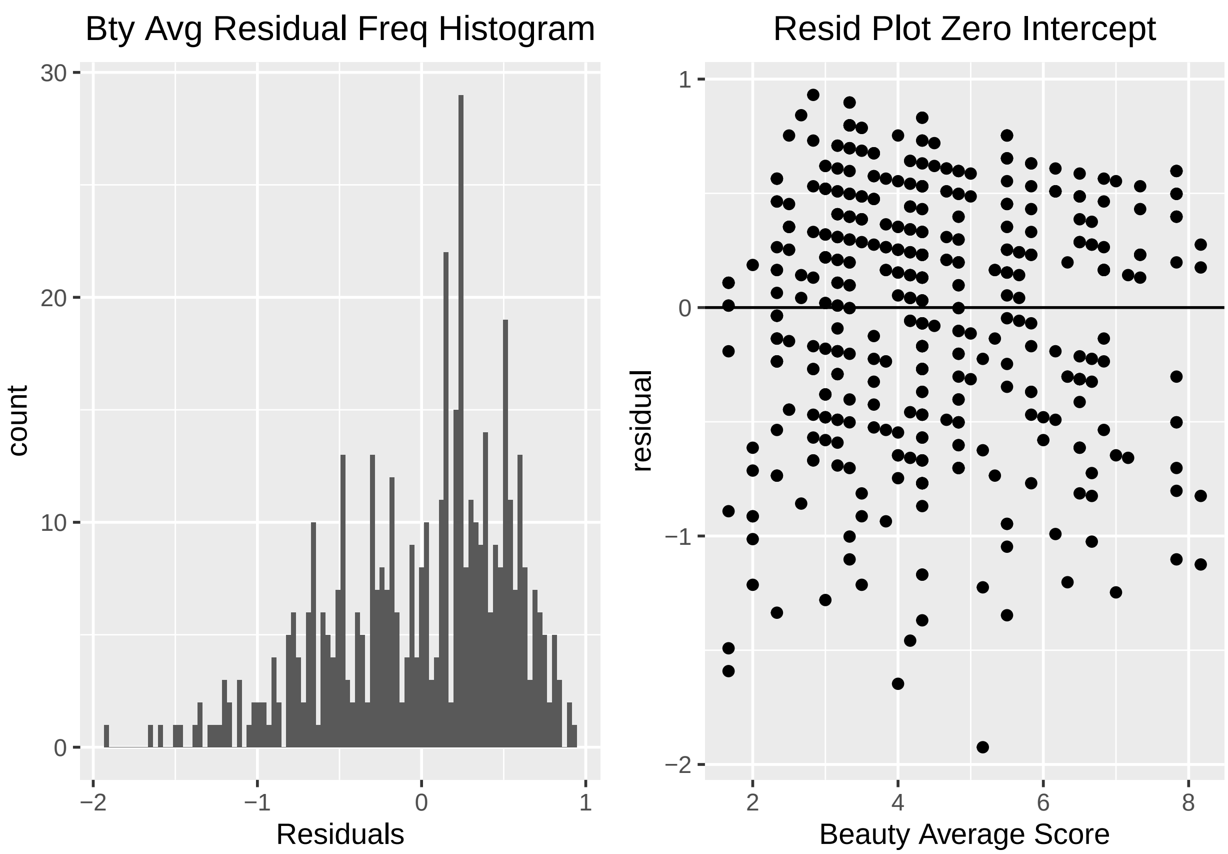
```
ggplot(data = evals, aes(x = bty_avg, y = score)) +
  geom_jitter() +
  geom_smooth(method = "lm", se = FALSE)
```



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).

The plot on the left does a better job showing that about .3 over the mean is very common. The plot on the right highlights the dispersion of the below average beauty averages

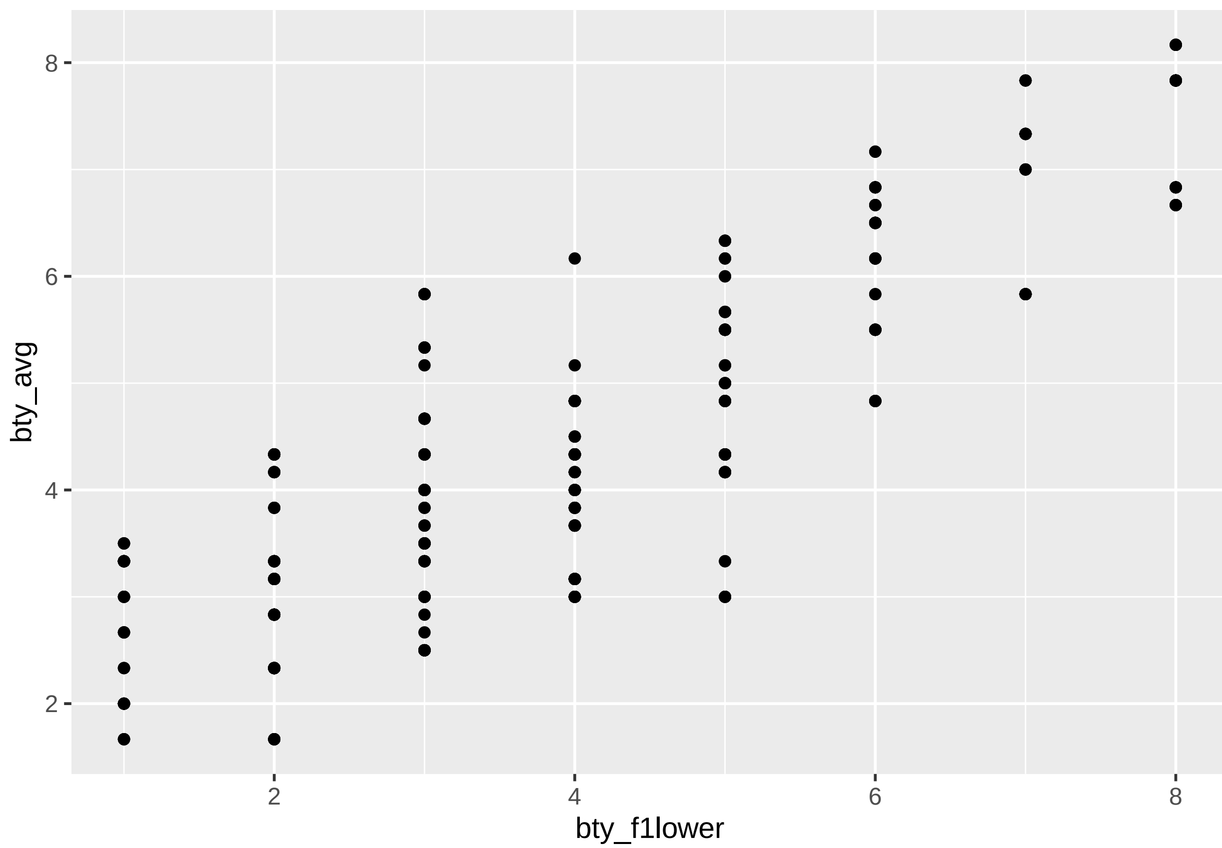
```
evals$residual <- resid(m_bty)



plot1<-ggplot(data = evals, aes(x = residual)) +
  geom_histogram(binwidth = .03) +
   ggtitle("Bty Avg Residual Freq Histogram") +
  xlab("Residuals") +
  theme(plot.title = element_text(hjust = 0.5))


plot2<-ggplot(evals, aes(x = bty_avg, y = residual)) +
  geom_hline(yintercept = 0) +
  geom_point() +
  ggtitle("Resid Plot Zero Intercept") +
  xlab("Beauty Average Score") +
  theme(plot.title = element_text(hjust = 0.5))

grid.arrange(plot1, plot2, ncol = 2)
```



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()
```

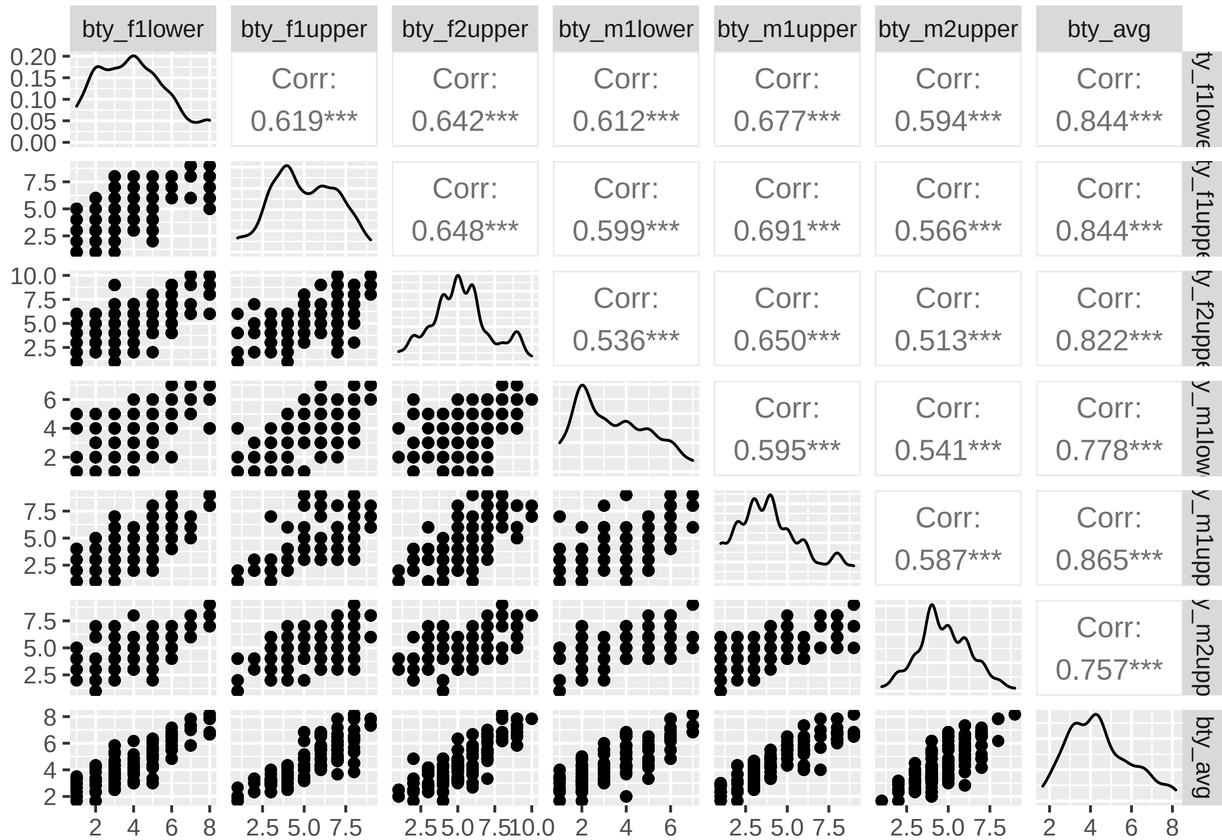



```
evals %>%
  summarise(cor(bty_avg, bty_f1lower))
```

```
## # 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)
sum_m_bty_gen<-summary(m_bty_gen)
sum_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.

The conditions for least square regression are :

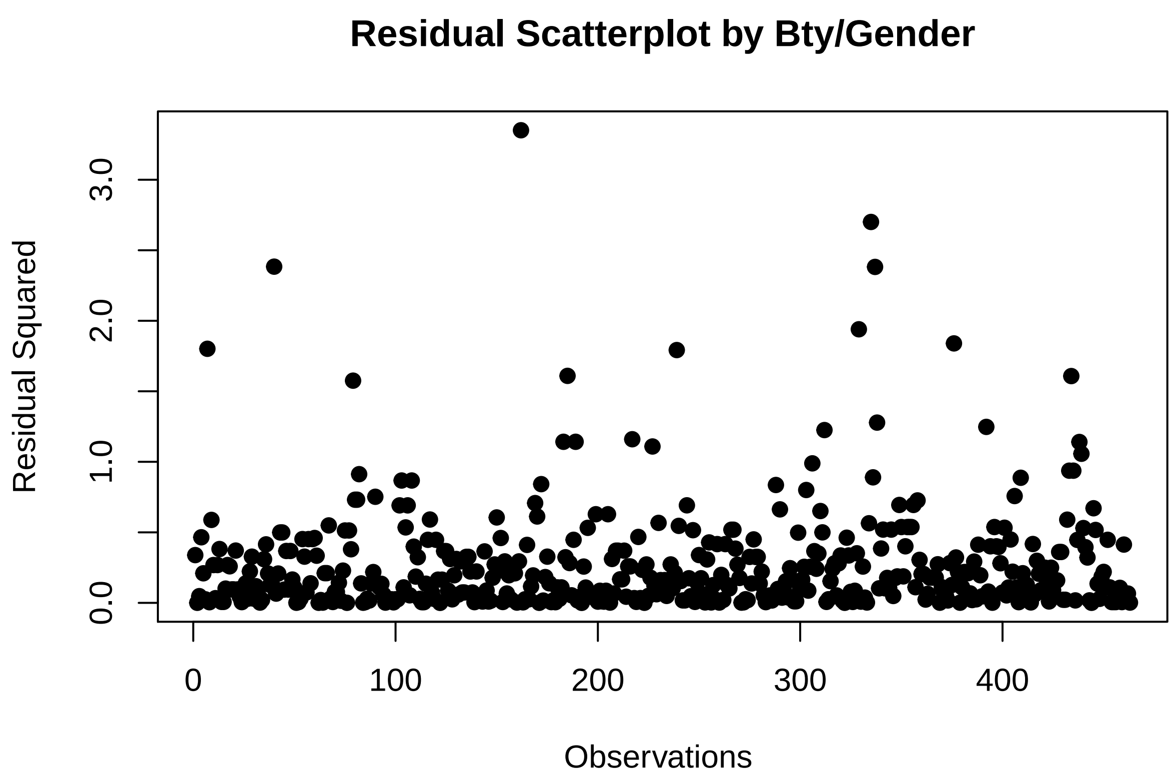
```
1)  Linearity
2)  Normality of the Residuals
3)  Constant Variability
4)  Independence
```

Here is a scatterplot of residual variability (each residual squared).

The most significant outlier was professer 30.

That professor had a bueaty of 5.1 yet scored only 2.3

```
evals$residual_gen <- resid(m_bty_gen)

evals$residual_gen_var <- (evals$residual_gen)^2

plot(x=evals$residual_gen_var, main="Residual Scatterplot by Bty/Gender",
     xlab="Observations ", ylab="Residual Squared", pch=19, frame=T)
```



Condition 2 looks ok as we saw earlier, its a bit left skewed. 3 and 4 look very good.

The relationship between beauty and assessment score is weak.

It seems to me using linearity as a condition for least squared regression is somewhat circular logic since the metrics are in part designed to assess linearity.

The R squared improved by adding gender.

In both models the RSE was high, so neither relationship is very linear.

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`?

Reviewing the estimates, it seems adding gender adjusts the slop up a bit.

```
# lets put them in a table

est.df<-data.frame(predictor="Intercept",
                   bty=sum_m_bty$coefficients[1][1],
                   btygender=sum_m_bty_gen$coefficients[1][1]
                   )

est.df<-rbind(est.df, data.frame(predictor=labels(sum_m_bty_gen$terms)[1],
                                 bty=sum_m_bty$coefficients[2][1],
                                 btygender=sum_m_bty_gen$coefficients[2][1]
                                 ))

est.df<-rbind(est.df, data.frame(predictor=labels(sum_m_bty_gen$terms)[2],
                                 bty=NA,
                                 btygender=sum_m_bty_gen$coefficients[3][1]
                                 ))



knitr::kable(est.df, caption='Estimate Comparison')
```

Table 1: Estimate Comparison

predictor	bty	btygender
Intercept	3.880338	3.7473382
bty_avg	0.066637	0.0741554
gender	NA	0.1723895

Reviewing the output metrics, it seems adding gender improves our correlation.

Reviewing the output metrics, r squared is higher, that infers a better linearity, better correlation.

The Fstatistic, however, decreased which infers the overall residual increased compared to the total sum of squares.

```
out.df<-data.frame( metric="r.squared",
                   bty=sum_m_bty$r.squared,
                   btygender=sum_m_bty_gen$r.square
```

```
                    )

out.df<-rbind(out.df, data.frame(metric="adj.r.squared",
                                  bty = sum_m_bty$adj.r.squared,
                                 btygender = sum_m_bty_gen$adj.r.squared
                                 ))

out.df<-rbind(out.df, data.frame(metric="rse",
                                  bty = sum_m_bty$sigma,
                                 btygender = sum_m_bty_gen$sigma
                                 ))



out.df<-rbind(out.df, data.frame(metric="fstat",
                                  bty = sum_m_bty$fstatistic[1],
                                 btygender = sum_m_bty_gen$fstatistic[1]
                                 ))

knitr::kable(out.df, caption='Model Output Comparison')
```

Table 2: Model Output Comparison

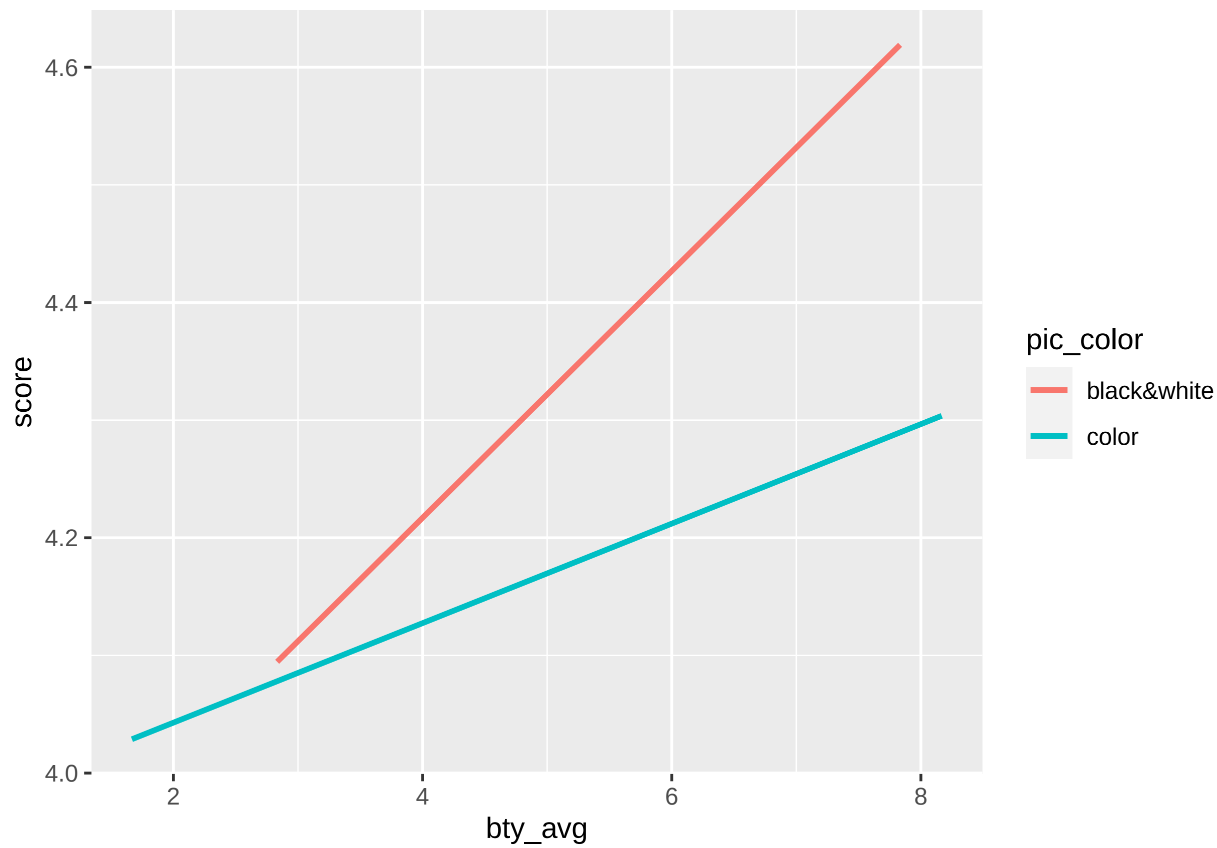
	metric	bty	btygender
1	r.squared	0.0350223	0.0591228
2	adj.r.squared	0.0329290	0.0550320
3	rse	0.5348351	0.5286878
value	fstat	16.7312277	14.4527273

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.

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

```
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?

Pictures that are black and white have a coefficient of 0.

Whatever effect black and white photos have is built in to the other coefficients.

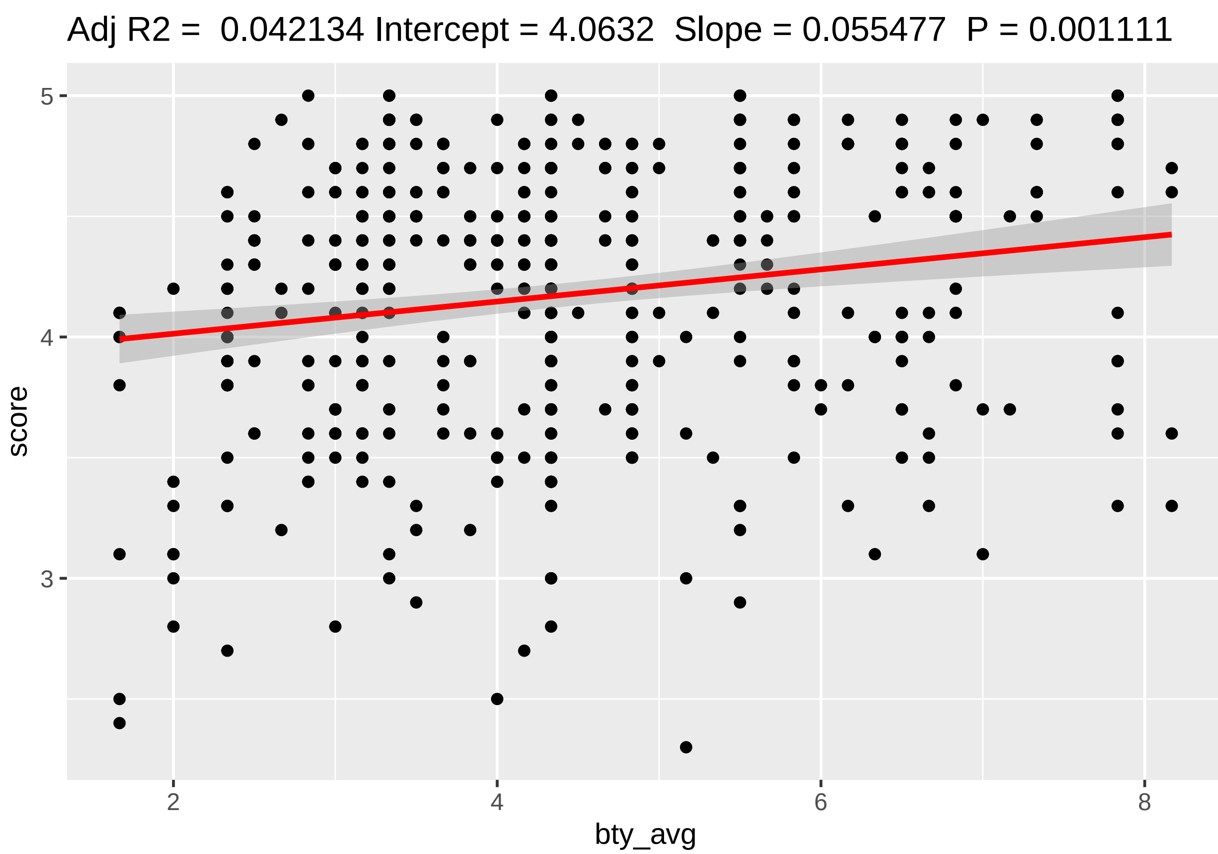
```
fit<-lm(score ~ bty_avg + pic_color, data = evals)
beta_0<-fit$coefficients[1][1]
beta_1<-fit$coefficients[2][1]
beta_2<-fit$coefficients[3][1]

sprintf("beta_0 = %.2f   beta_1 = %.2f    beta_2 = %.2f ", beta_0,beta_1,beta_2)
```

```
## [1] "beta_0 = 4.06   beta_1 = 0.06    beta_2 = -0.16 "
```

$$score = \begin{cases} pic = color & \text{-0.16059pic + btyavg * 0.05548 + 4.06318} \\ pic = blackwhite & \text{0 + btyavg * 0.05548 + 4.06318} \end{cases}$$

```
ggplot(fit$model, aes_string(x = names(fit$model)[2], y = names(fit$model)[1])) +
  geom_point() +
  stat_smooth(method = "lm", col = "red") +
```

```
labs(title = paste("Adj R2 = ",signif(summary(fit)$adj.r.squared, 5),
                   "Intercept =",signif(fit$coef[[1]],5 ),
                   " Slope =",signif(fit$coef[[2]], 5),
                   " P =",signif(summary(fit)$coef[2,4], 5)))
```



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
```

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.

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.

The gendermale pvalue is very low, I guess it really is important to the score

Below I added the number of students coefficients to the previous table.

```r
m_bty_std <- lm(score ~ bty_avg + cls_students, data = evals)
sum_m_std<-summary(m_bty_std)

# intercept
est_new_col.df<-data.frame(btystudents=sum_m_std$coefficients[1][1]  )

est_new_col.df<-rbind(est_new_col.df, data.frame(
                                 btystudents=sum_m_std$coefficients[2][1]
                                 ))

est_new_col.df<-rbind(est_new_col.df, data.frame(
                                 btystudents=NA
                                 ))

est_new.df<-cbind(est.df,est_new_col.df)



est_new_row.df<-data.frame(predictor="n_students",
                  bty=NA,
                  btygender=NA,
                   btystudents=sum_m_std$coefficients[3][1]
                   )

est_new.df<-rbind(est_new.df, est_new_row.df)

knitr::kable(est_new.df, caption='Estimate Comparison')
```

Table 3: Estimate Comparison

predictor	bty	btygender	btystudents
Intercept	3.880338	3.7473382	3.8785042
bty_avg	0.066637	0.0741554	0.0663721
gender	NA	0.1723895	NA
n_students	NA	NA	0.0000544

13. Interpret the coefficient associated with the ethnicity variable.

I repeated the exercise. In this case ethnicity (not minority) improved the score a little less than gender (male)

```
m_bty_std <- lm(score ~ bty_avg + ethnicity, data = evals)
sum_m_std<-summary(m_bty_std)

# intercept
est_new_col.df<-data.frame(btyethnicity=sum_m_std$coefficients[1][1]  )

est_new_col.df<-rbind(est_new_col.df, data.frame(
                                btyethnicity=sum_m_std$coefficients[2][1]
                                ))

est_new_col.df<-rbind(est_new_col.df, data.frame(
                                btyethnicity=NA
                                ))

est_new.df<-cbind(est.df,est_new_col.df)



est_new_row.df<-data.frame(predictor="ethnicity",
                  bty=NA,
                  btygender=NA,
                   btyethnicity=sum_m_std$coefficients[3][1]
                   )

est_new.df<-rbind(est_new.df, est_new_row.df)

knitr::kable(est_new.df, caption='Estimate Comparison')
```

Table 4: Estimate Comparison

predictor	bty	btygender	btyethnicity
Intercept	3.880338	3.7473382	3.7613598
bty_avg	0.066637	0.0741554	0.0678685
gender	NA	0.1723895	NA
ethnicity	NA	NA	0.1317495

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?

\textcolor{blue}{The highest pvalue belonged the cls_profs when category level=multiple}

```
m_almost_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_almost_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
```

It didnt make a significant difference. The original coefficient was quite small.

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.

Ill prioritize adjusted r squared but also look at rse.

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

sprintf("New Model Adj R Squared = %.3f.   RSE = %.3f",
        summary(m_backwards)$adj.r.squared, summary(m_backwards)$sigma)
```

```
## [1] "New Model Adj R Squared = 0.163.   RSE = 0.498"
```

```r
sprintf("Full Less Gender Model Adj R Squared = %.3f.   RSE = %.3f",
        summary(m_almost_full)$adj.r.squared, summary(m_almost_full)$sigma)
```

```
## [1] "Full Less Gender Model Adj R Squared = 0.163.   RSE = 0.497"
```

```r
sprintf("Full Model Adj R Squared = %.3f.   RSE = %.3f",
        summary(m_full)$adj.r.squared, summary(m_full)$sigma)
```

```
## [1] "Full Model Adj R Squared = 0.162.   RSE = 0.498"
```

\textcolor{blue}{Seems like each time i remove something r-squared decreases and rse increases. I removed cls_profs and cls_level which helped a little}

I tried to do this programatically but it was taking too much time so i used notepad to format this...

score = \ 4.0856255 + \ ranktenure track * -0.142 \ ranktenured * -0.089 \ gendermale * 0.203 \ ethnicitynot minority * 0.142 age * -0.008 languagenon-english * -0.209 cls_perc_eval * 0.005 cls_students * 0.0003 bty_avg * 0.041 cls_creditsone credit * 0.473 pic_colorcolor * -0.197 pic_outfitnot formal * -0.117

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

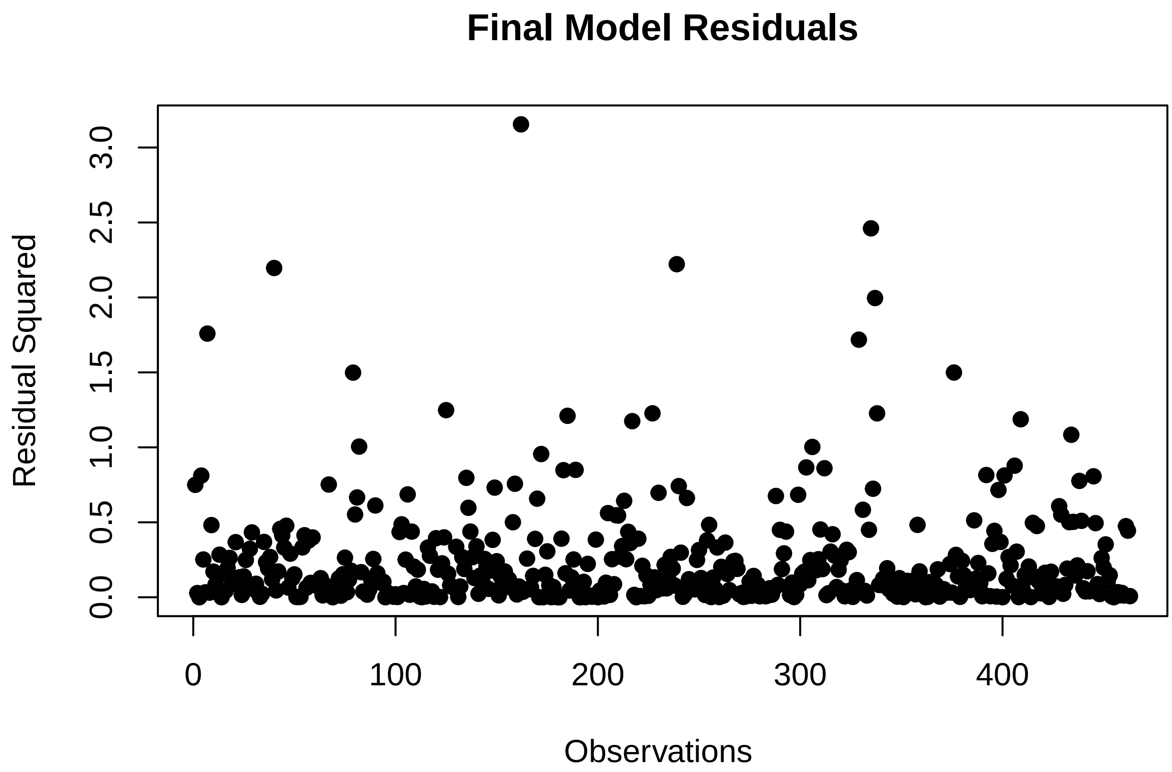
The conditions for least square regression are :

```
1)  Linearity
2)  Normality of the Residuals
3)  Constant Variability
4)  Independence
```

Here is a scatterplot of residual variability (each residual squared).

My comments are the same as from question 7 which used a different equation.

```r
evals$residual_b_gen <- resid(m_backwards)

evals$residual_gen_var <- (evals$residual_b_gen)^2

plot(x=evals$residual_gen_var, main="Final Model Residuals", xlab="Observations ", ylab="Residual Square
```



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?

Sure some class are relative low stress where no one gets upset.

Other classes stressful on the students so maybe they are more critical in the evaluations.

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.

It does appear the men received better grades (by .2)

The model has any class with "some credits" receiving an extra .47 so I dont understand that.

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

\textcolor{blue}{When it comes the bty_avg, I think the study reveals more about what the beauty assessors are thinking than anything else.}

The other attributes are more interesting. It is possible that age and gender influence students in general.
