

# ECON30520: Final Project on Beauty dataset

Nikita Fernes - 18205420

18 January, 2022

## 1 Introduction

This paper will study the impact of an instructor's appearance on their course evaluation. Our main dependent variable of interest is course evaluation *eval* which is on a scale of 1 to 5 and our main explanatory variable is *beauty* that is averaged across a panel of 6 students. Overall, we have 9 explanatory variables 4 of which are numeric and the rest categorical. First we have some descriptive statistics, then we introduce some relationships between the variables and finally we perform a multiple regression using OLS to estimate the effect of beauty of course evaluation.

## 2 Literature Review

A study by Ponzio and Scoppa (2012) done in an Italian university with a fairly large data set shows that beauty does in fact have a significant positive correlation with course evaluation even in an intellectual profession such as teaching. They also state that beauty in this profession may be more important for females than for males.

However, the magnitude of this correlation is much lower in a German study by Süssmuth (2006) where the impact of overall beauty in the instructor's course evaluation is positive but only significant at a 10% level of significance in one group of younger students. There was no significant relationship with the other groups of students. An interesting point to note is that Süssmuth (2006) found that courses taught in English had better instructional

ratings showcasing that language can have a great impact on course evaluation. They also found that there was some indication of gender discrimination by advanced students.

## 3 Descriptive Statistics

### 3.1 Numerical Variables

To start with we describe the numerical data in the table given below. Here, we see that the average teaching evaluation score is 4 out of 5 which seems to be quite good considering it is above average. The beauty rating that is averaged across the panelists was shifted to have a mean of 0 as seen in the table. Average age of an instructor is 48 and the mean number of students in a course is 55.

	study (N = 463)
<b>Overall teaching evaluation score</b>	
min	2.1
max	5
mean (sd)	$4.00 \pm 0.55$
<b>Beauty rating of instructor by 6 students</b>	
min	-1.450494
max	1.970023
mean (sd)	$0.00 \pm 0.79$
<b>Age of instructor</b>	
min	29
max	73
mean (sd)	$48.37 \pm 9.80$
<b>Number of students in course</b>	
min	8
max	581
mean (sd)	$55.18 \pm 75.07$

## 3.2 Categorical Variables

Most of our categorical variables are binary variables so I have combined them with one another and some numerical variables to get a more in depth and stylised analysis of the data. This part of the paper deals with the features of the instructor.

We have a pie chart that shows us the number of minority professors, separated by gender. We see that in a total of 195 females, 36 belong to a minority background and 159 do not i.e., they are caucasian. Again, in a total of 268 males only 28 are from minority backgrounds. This shows that there is a majority of males as well as caucasian instructors.

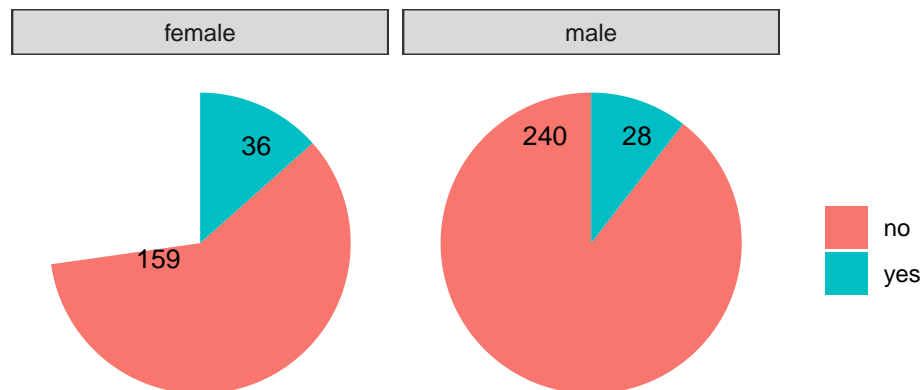


Figure 1: Pie chart of minority professors by gender

Next, we have a simple histogram paired with a density plot that smoothes out the the distribution of age using a kernel density estimate. We see as before the age distribution is largely concentrated in the 48-50 range same as the mean previously described. ‘we also see some concentration around the 58-60 age group followed by 35-37 year olds.

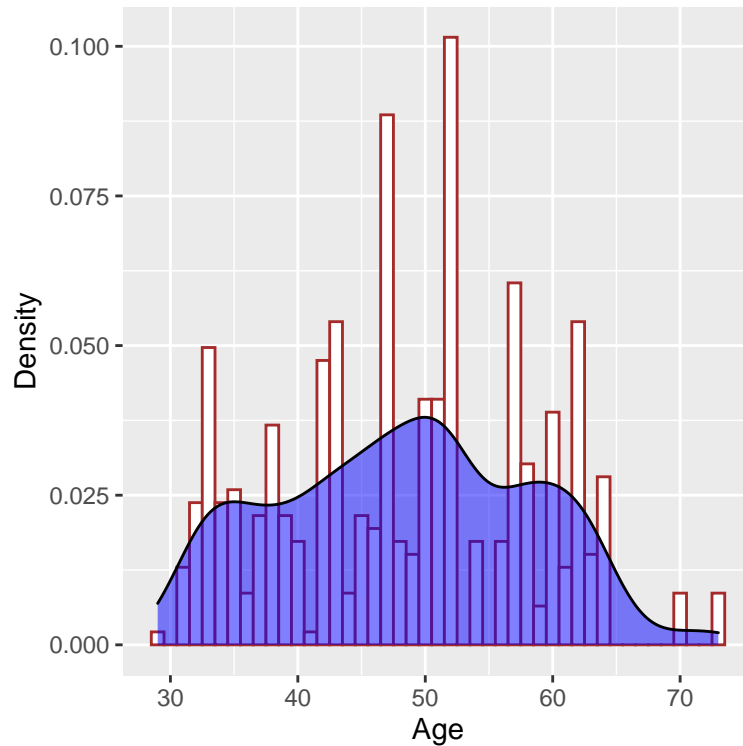


Figure 2: Histogram by age

Here, I wanted to see the relationship between the beauty rating with gender and whether the instructor was a native english speaker. We already know that the mean of beauty is 0 that is easily noticeable in this violin plot. However, we see that native speakers overall have a higher beauty rating, more for women than for men. Important to note from the graph that females that are non-native english speakers have more negative averaged beauty ratings than males that are non-native english speakers. However, overall females have a higher beauty rating as well showing that perhaps females as instructors are more criticized for their looks.

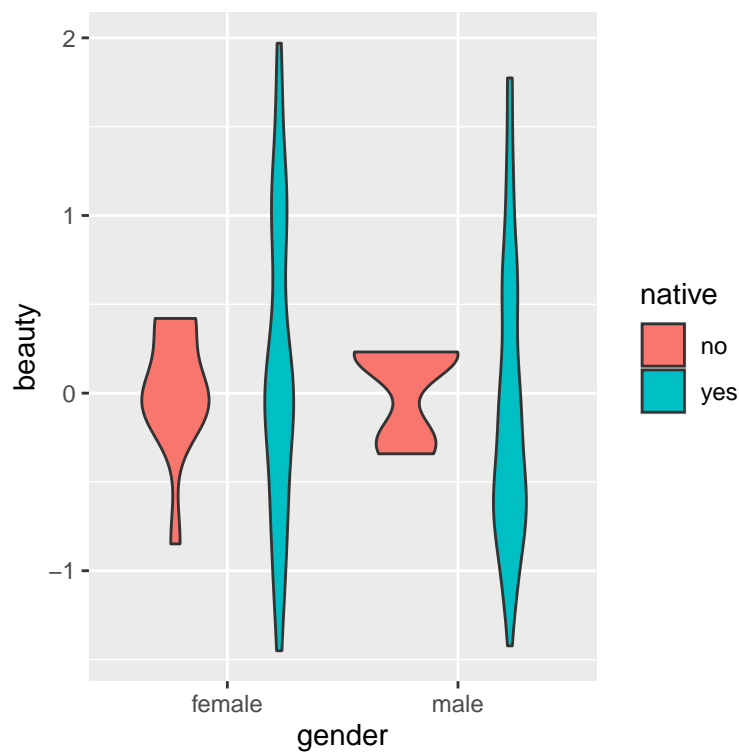


Figure 3: Violin plot of age by Gender, separated by native

## 4 Relationships with evaluation

Just like the relationship seen before with beauty and gender we will be looking at some relationships with evaluation which is the main dependent variable. In the boxplot below in figure 4 we see that native english speakers have a higher average in course evaluation than non-native speakers. The whiskers that represent the confidence intervals also differ in both groups.

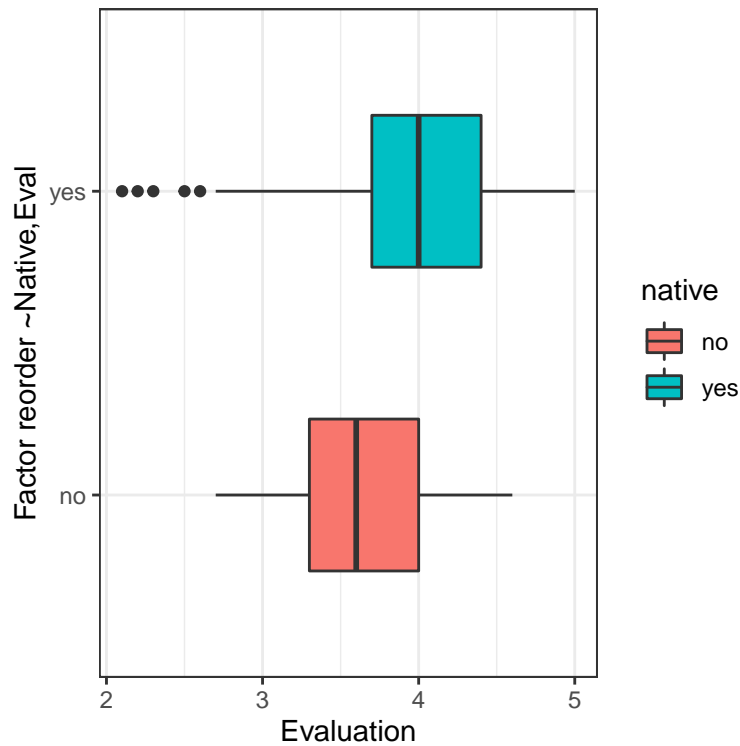


Figure 4: Boxplot showing evaluation if native or not

In the relationship below in figure 5 we regress evaluation on beauty however we further distinguish the two by the type of course i.e., whether it is a single-credit elective or not. The graph below shows us that when the course is not a single credit elective, beauty and evaluation have a positive correlation. In contrast, when the course is a single credit elective beauty and evaluation have a negative correlation.

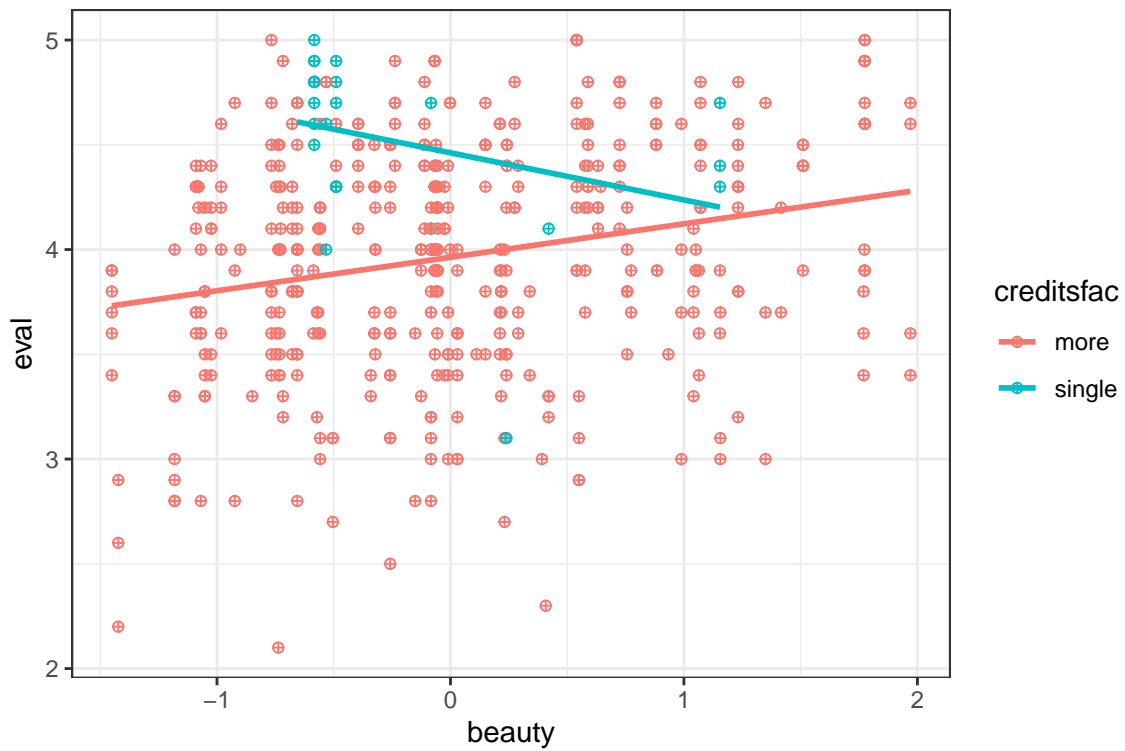


Figure 5: Scatter plot of evaluation on beauty separated by credits

Here in figure 6 we see how course evaluation on beauty differs in a lower division versus in a upper division course, wherein lower division consist mainly of freshmen and sophomores whereas upper division consists largely of seniors. Using the following graph we can see that in the lower division the correlation between beauty and evaluation begins as positive however after a few fluctuations it shows a negative correlation when beauty is rated above 0.5. Upper division also shows one fluctuation in the correlation between beauty and evaluation however overall it has a positive relationship as with an increase in beauty there is an increase in course evaluation.

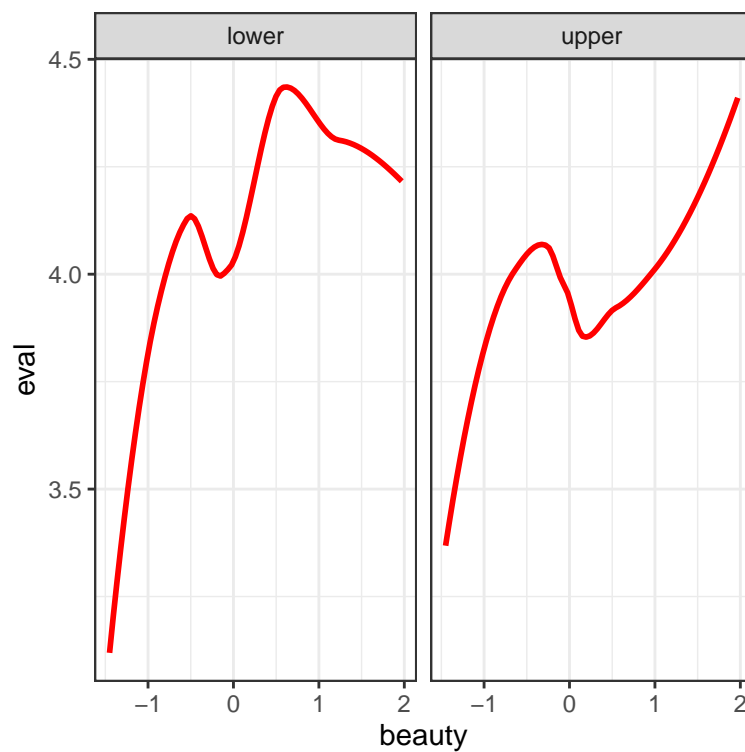


Figure 6: Relationship of evaluation on beauty separated by division that is ordered by age



This relationship below in figure 7 shows us how tenure affects the relationship between course evaluation and beauty. We see a positive relationship between beauty and evaluation for instructors that have tenure as well as those that don't. At a glance it seems that they run parallel to one another and may have the similar if not the same slopes. For those without tenure, the evaluation seems to be higher at every level of beauty rating.

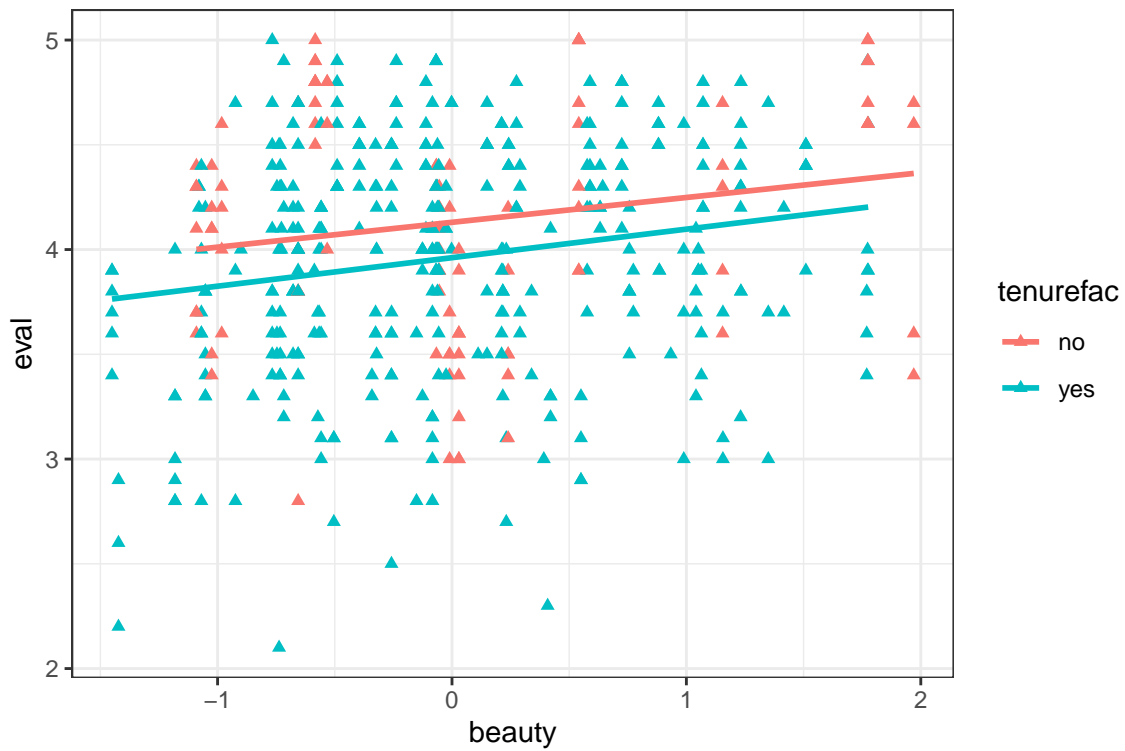


Figure 7: Course evaluation on beauty separated by Tenure

This final graph 8 showcases the distribution of age and gender as we regress our dependent variable = evaluation on our main explanatory variable = beauty. We see an overall positive correlation between beauty and course evaluation, the relationship of which we will be looking in depth in the next segment. This graph makes it easy for us to see that female instructors are given a higher beauty rating than men as we observe on the right side of the graph. Another point to note is that when beauty  $> 0$ , the age groups seem to be of a younger cohort of instructors.

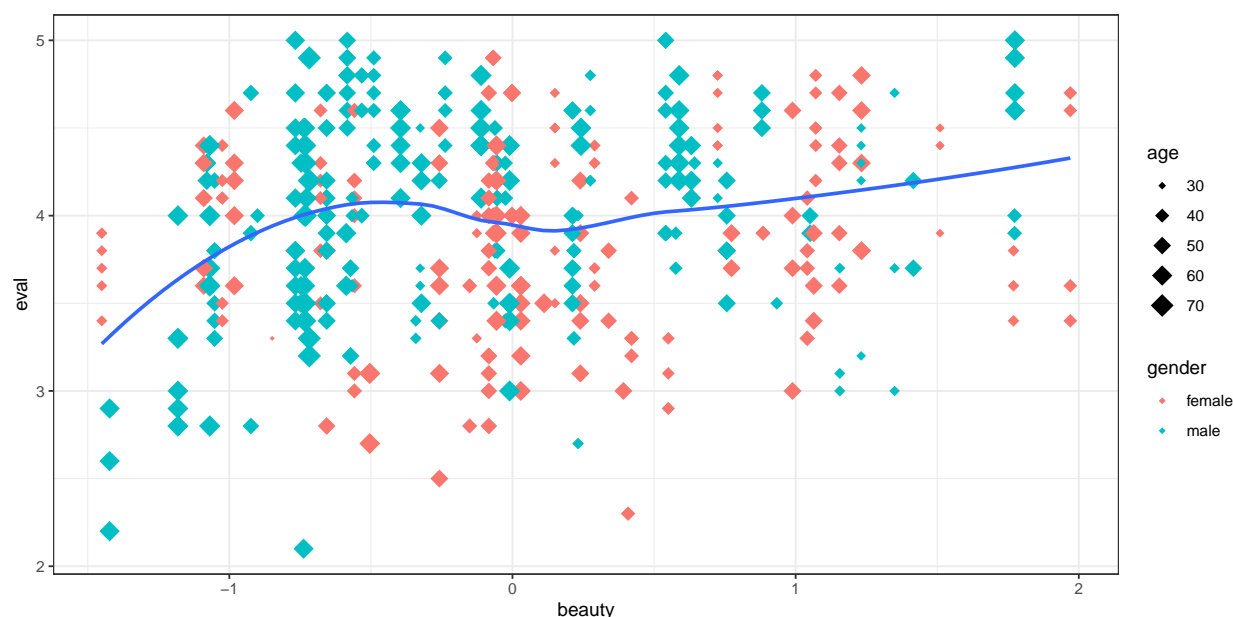


Figure 8: Scatter plot with age and gender, regressing beauty on evaluation

Looking into gender further, we see in figure 9 that beauty has a higher impact on evaluation of male instructors than females which is contrary to the findings of Ponzio and Scoppa (2012) where beauty for females has a greater impact.

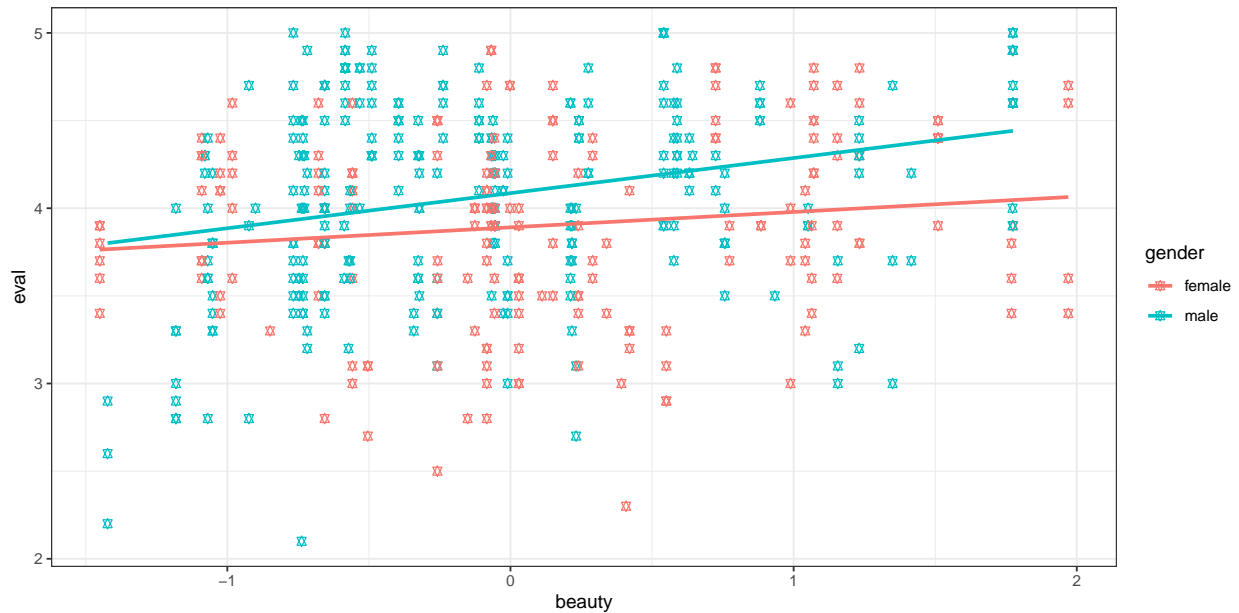


Figure 9: Evaluation on beauty, by gender

## 5 Regressions

### 5.1 Ordinary Least Squares

Using all the information above we create a regression to check if there exists a significant correlation between the explanatory and dependent variable - course evaluation *eval*. I have done 1 simple regression between evaluation and beauty which was previously plotted in figure 8 as well as 2 multiple regressions.

```
lm_main= lm(eval~beauty, data=study)

lm_professor = lm(study$eval ~ study$beauty + study$age +
                  study$gender + study$minority +study$native, study)

lm_mult = lm(study$eval ~ study$beauty + study$age +
              study$gender + study$minority+study$native+
```

```
study$division + study$tenure + study$credits+study$allstudents)

models<-list(lm_main,lm_professor,lm_mult,lm_mult)
```

$$eval = \beta_0 + \beta_1 beauty + \beta_2 age + \beta_3 gender + \beta_4 minority + \beta_5 native \\ \beta_6 division + \beta_7 tenure + \beta_8 credits + \beta_9 allstudents$$

In the **Main** regression, we see that as beauty increases by 1, eval increases by 0.13. As the p value is 0 we see that it is quite a strong significant relationship at all the significant levels - 1%, 5% & 10%. The standard error is 0.032. The 95% confidence interval is  $[b-2(0.032), b+2(0.032)]$  and it has a very low R square.

In the second **Instructor** model that takes into account the attributes of the instructor we see that beauty, gender and native speaker are highly significant. The rest are statistically insignificant and therefore I do not feel the need to discuss them. Although, it is interesting to note how age has a negative correlation with evaluation though insignificant. Holding every other variable constant in each interpretation: - As beauty increases by 1, evaluation increases by 0.14 - Being male increases the evaluation score by 0.206 - Being a native speaker, increases the evaluation score by 0.313 Overall R square has also increased to 0.89.

In the **Final** model, that includes all the given explanatory variables shows us that beauty, gender and credits are statistically significant at all levels of significance. Minority that was statistically insignificant before is now significant at the 5% and 10% levels of significance. In contrast, native speaker that was statistically significant at all levels before is now significant only at 5% and 10% levels. Upper-Division and Tenure are weakly significant at the 10% level of significance. Age is still statistically insignificant. Holding others constant in each interpretation: - As beauty increases by 1 unit eval increases by 0.159 - Being male increases the evaluation score by 0.195 - Increase in elective by 1 unit increases evaluation by 0.587 - Being a native speaker, increases the evaluation score by 0.238 - Being from a minority background decreases the evaluation score by 0.164 - Being an instructor in an upper-division

Table 1: Regression Table

	Main	Instructor	Overall	Robust
beauty	0.133 (0.032) [p = 0.000]			
Beauty		0.141 (0.033) [p = 0.000]	0.159 (0.032) [p = 0.000]	0.159 (0.031) [p = 0.000]
Age		-0.003 (0.003) [p = 0.325]	-0.002 (0.003) [p = 0.381]	-0.002 (0.003) [p = 0.381]
Male		0.207 (0.053) [p = 0.000]	0.195 (0.052) [p = 0.000]	0.195 (0.054) [p = 0.000]
Minority		-0.044 (0.076) [p = 0.558]	-0.164 (0.077) [p = 0.034]	-0.164 (0.070) [p = 0.034]
Native speaker		0.313 (0.109) [p = 0.004]	0.238 (0.108) [p = 0.027]	0.238 (0.101) [p = 0.027]
Upper-Division			-0.015 (0.057) [p = 0.789]	-0.015 (0.059) [p = 0.789]
Tenure			-0.058 (0.063) [p = 0.361]	-0.058 (0.060) [p = 0.361]
Credits			0.587 (0.118) [p = 0.000]	0.587 (0.115) [p = 0.000]
No. of Students			0.000 (0.000) [p = 0.458]	0.000 (0.000) [p = 0.458]
Num.Obs.	463	463	463	463
R2	0.036	0.089	0.159	0.159
Std.Errors	Classical	Classical	Classical	Robust

P-values in square brackets

In final column, robust standard errors in parentheses

Standard errors in parentheses

course decreases the evaluation score by 0.015 - Being an instructor on tenure can decrease the evaluation score by 0.58

R square increases to 0.159 showing that this final model has a better goodness-of-fit than the other models. There is also a slight but notable difference in standard errors between the robust and classical model.

## 5.2 Coefficient Plots

The coefficient plot shows us the 95% confidence intervals of all the coefficients. We see that at and number of students are 0 with no confidence intervals.

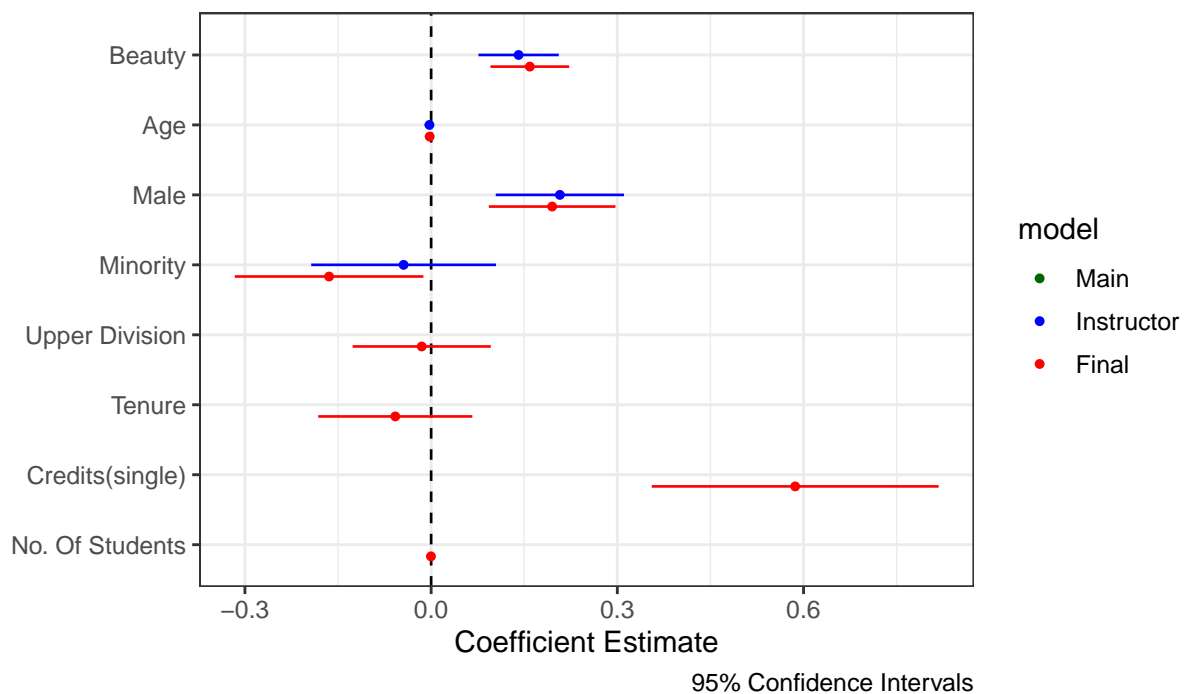


Figure 10: Coefficient Plot of OLS models

## 6 Conclusion

Overall we see that beauty is statistically significant in each model and therefore has a positive correlation with the evaluation of the course even while controlling for all other factors. This matches the study done in Italy by Ponzio and Scoppa (2012) as previously mentioned. The characteristics of the instructor that is out of his/her control that have an effect on this relationship is gender, being a native speaker and coming from a minority background. This is similar to what we see in the study in Germany by Süssmuth (2006) that shows us languages and gender can have an effect on evaluation. In terms of the course attributes that have nothing to do with beauty, being on tenure, teaching a single/multiple credit course and whether they teach an upper-division or lower-division course also have an impact on course evaluation.

## References

- Ponzo, Michela, and Vincenzo Scoppa. 2012. “The Good, the Bad, and the Ugly: Teaching Evaluations, Beauty and Abilities.” <https://ideas.repec.org/p/clb/wpaper/201204.html>.
- Süssmuth, Bernd. 2006. “Beauty in the Classroom: Are German Students Less Blinded? Putative Pedagogical Productivity Due to Professors’ Pulchritude: Peculiar or Pervasive?” *Applied Economics* 38: 231–38.