Physical Appearance Affects Instructors Ratings More Than We Think*

Marcin Jaczynski

20 April 2023

At the end of each semester of school, students are able to rate their professors and teachers through course evaluations. There are two schools of thought for course evaluations: they are repersentative of how well the teacher was able to teach a course, and if the students liked them or they are meaningless as only students who received a bad grade fill out the course evaluations and they are not representative of the teachers ability to teach. However, there are many factors that influece a teacher's rating. In this paper I observe how the beauty has an impact on their overall evaluations and see how other factors also play a role in the rating outcome.

1 Introduction

Teacher ratings have a direct impact on the validity of a teacher or professor. While there are a lot of factors that play on how students will perceive a teacher and what kind of rating they will give them at the end of the semester, one that is not taken into consideration many times is how beautiful they are. As such, I have chosen to try and answer "How does an instructors beauty, according to their students, affect the instructors course evaluation?" This paper seeks to examine what the trends in instructor evaluations are from the students at the University of Texas during the years 2000-2002. Among beauty, things such as age, gender, department, tenure, and how many students they were teaching are going to be taken into account and evaluated. Based on the observations from the 463 instructors, there is a correlation between the beauty of the instructor and the evaluation they have received.

*Code and data are available at: **repo link HERE**_

2 Data

2.1 Data Source and Methodology

The data collected for this paper, was sourced form an Applied Econmetrics with R (AER) library. The AER is a collection of data sets in applied econometrics research, and is widely used in the academic fields of empirical analysis of microeconomic data. The package is free to download from the Comprehensive R Archive Network (CRAN), or can be installed using the RConsole or RStudio.

2.2 Data Collection

The data set used in this paper contains data pertaining to course, instructor and student evaluation data of 463 instructors. The data was collected from the 2000-2002 academic years at the University of Texas at Austin. The data was provided by Prof. Hamermesh, and includes the students ratings of their instructors as well as their beauty rating (average from six independent judges), as well as other charasteristics pertaining to the instructor. The details for the data collection methodology of the TeachingRatings data set are unclear. However, assumptions can be made that the data was collected through anonymous teacher ratings websites such as rateMyProfessor.com or course evaluations that were done at the school. However, the exact methodology is not explicitly mentioned in the documentation of the data set or the academic literature which uses the following data set.

The data set has been previously cleaned, and so it does not require any additional cleaning, however, minor adjustments have been made, such as omitting the columns "…1" and "prof" as both of them only contained the entry number, which was already present in data frame. Additionally, as a safety check, a function was ran in order to remove any N/A values, however, the result showed that no entries were removed as a result of said function.

Each entry in the following data set contains the following details:

- Minority - factor. Does the instructor belong to a minority (non-Caucasian)?
- Age - the professor's age.
- Gender - factor indicating instructor's gender.
- Credits - factor. Is the course a single-credit elective (e.g., yoga, aerobics, dance)?
- Beauty - rating of the instructor's physical appearance by a panel of six students, averaged across the six * panelists, shifted to have a mean of zero. On a scale of -2.0 to 2.0.
- Eval - course overall teaching evaluation score, on a scale of 1 (very unsatisfactory) to 5 (excellent).
- Division - factor. Is the course an upper or lower division course? (Lower division courses are mainly large freshman and sophomore courses)?

- Native- factor. Is the instructor a native English speaker?
- Tenure - factor. Is the instructor on tenure track?
- Students - number of students that participated in the evaluation.
- Allstudents - number of students enrolled in the course.
- Prof - factor indicating instructor identifier.

2.3 Data Analysis

The data set from AER was gathered by installing the AER package [source AER] in the RStudio [source R studio] console, and then the package contents were extracted to only include the TeachingRatings data set in the "download_data" file under "/scripts". The data was then analyzed using the R programming language (**rcoreteam2022?**). Additonally, the programming packages dplyr [source dplyr] as well as ggplot2 [source ggplot2], were used to analyze the data, and create the graphs.

2.4 Advantages and Disadvantages

The following data set has both some advantages and disadvantages. In terms of advantages, the data set is really easy to obtain and use. It is available online for free, and requires 3 steps in order to gather the main TeachingRatings data set. The data set that is then obtained has also already been cleaned, as such it does not require any additional work other than picking which columns will be used. Additionally, the documentation which is provided is very thoroughly written, with each column explained and examples of how to install the package and obtain the data. However, in terms of the disadvantages, firstly the documentation that is provided, does not provide any background context on the data gathering methodology, and if there was any incentive to complete the survey. The data set also contains 463 entries, which over a two year span, does not seem like a large number of entries. Additionally, the study was performed in the year 2002, which would not qualify it as a recent study, and would likely need to be redone to be applicable to more recent years, especially after a global pandemic in the years 2020-2022.

3 Results

3.1 Male vs. Female

The first area I explored, while trying to find what affected an instructors evaluation is the gender of the instructor. According to my initial hypothesis, if one gender scored a higher in beauty metrics , then they would also score higher in terms of evaluation. As such, I decided to create two seperate graphs in order to show the correlation between beauty and evaluation

for both genders. Figure 1 shows the beauty and evaluation scores for all male instructors, and, the results show that for male instructs for every additional point of beauty rating, we would expect the evaluation to increase by 0.20027.

```
Call:
lm(formula = eval ~ beauty, data = males)

Residuals:
     Min       1Q   Median       3Q      Max
-1.83820 -0.37318  0.05899  0.39397  1.06764

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.08595    0.03298 123.882  < 2e-16 ***
beauty       0.20027    0.04337   4.617 6.06e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5366 on 266 degrees of freedom
Multiple R-squared:  0.0742,    Adjusted R-squared:  0.07072
F-statistic: 21.32 on 1 and 266 DF,  p-value: 6.056e-06
```

The next figure Figure 2 explores the correlation between the beauty and evaluation of female instructors. The figure conforms to my hypothesis, as for every additional point in beauty ratings the instructor should gain an extra 0.08762 evaluation points.

```
Call:
lm(formula = eval ~ beauty, data = females)

Residuals:
     Min       1Q   Median       3Q      Max
-1.62661 -0.37392  0.01511  0.41156  1.01511

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.89085    0.03873 100.468   <2e-16 ***
beauty       0.08762    0.04700   1.864   0.0638 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

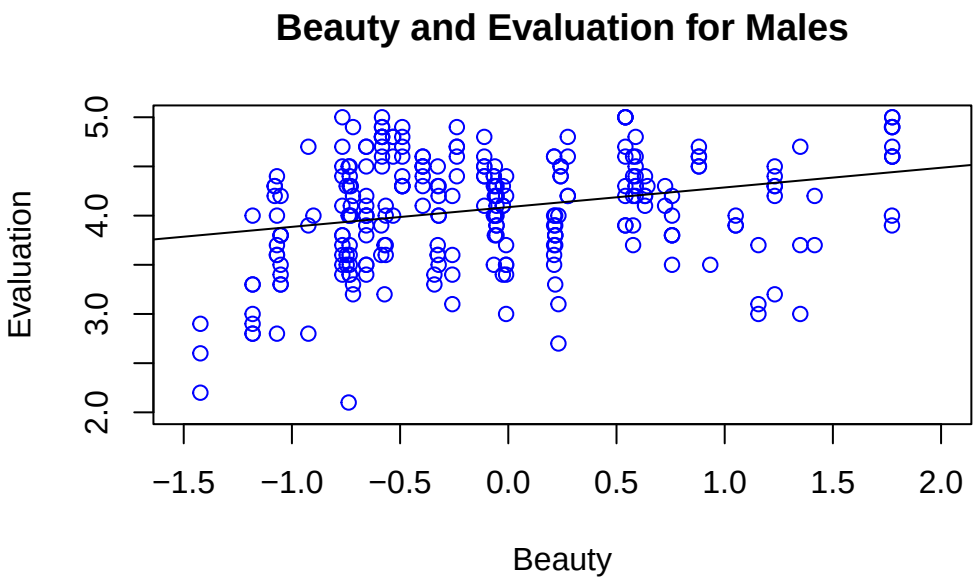


Figure 1: DON'T FORGET..

```
Residual standard error: 0.5354 on 193 degrees of freedom
Multiple R-squared:  0.01769,   Adjusted R-squared:  0.0126
F-statistic: 3.475 on 1 and 193 DF,  p-value: 0.06383
```

Figure Figure 3 shows the graph of the beauty and evaluation correlation for female instructors with the line of best fit for the male instructors. According to the models, for men the beauty affects their evaluations more than the female instructors.

Figure (**table-avg-male-and-female-beauty-and-eval?**) shows the average beauty and evaluation rating for both males and females. After seeing the results of the male instructors having their evaluation be affected more by their beauty, I wanted to see what the average rating for both males and females were. According to the data, male instructors have a lower average beauty rating than female instructors, yet males are on average more highly rated. After seeing the following results, I decided to take on another route to figure out what impacted the instrucots evaluation the most.

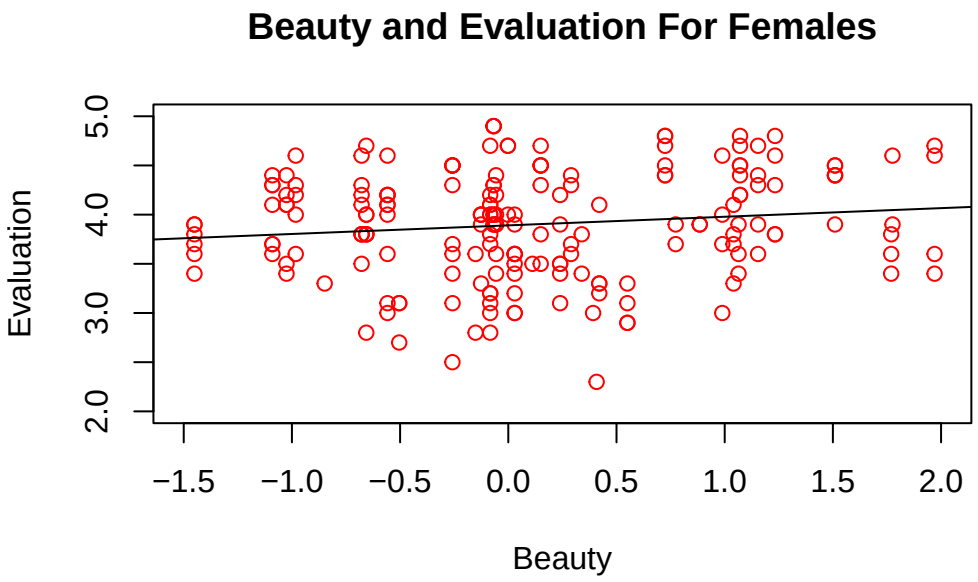


Figure 2: DON'T FORGET..

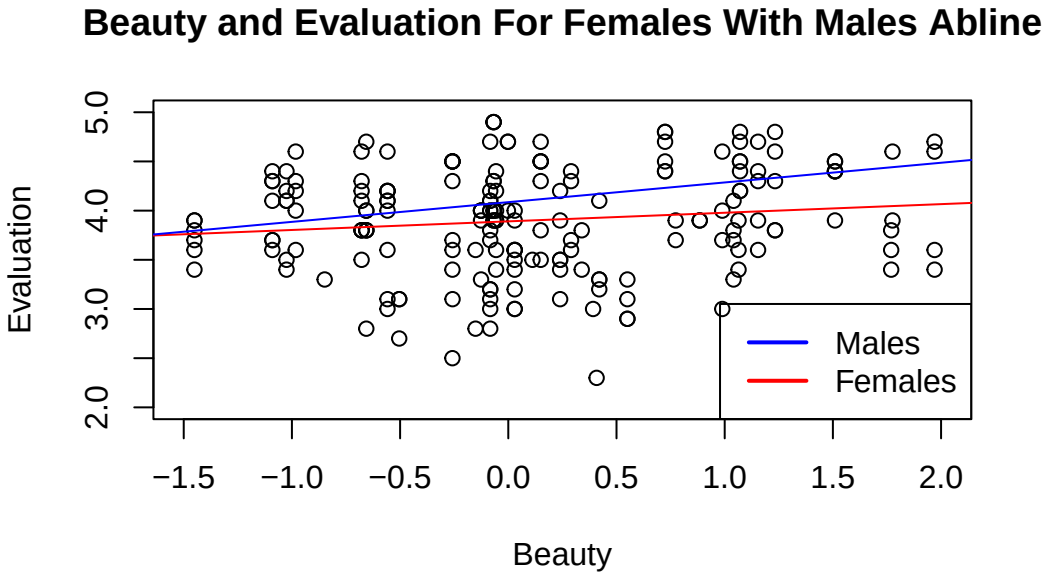


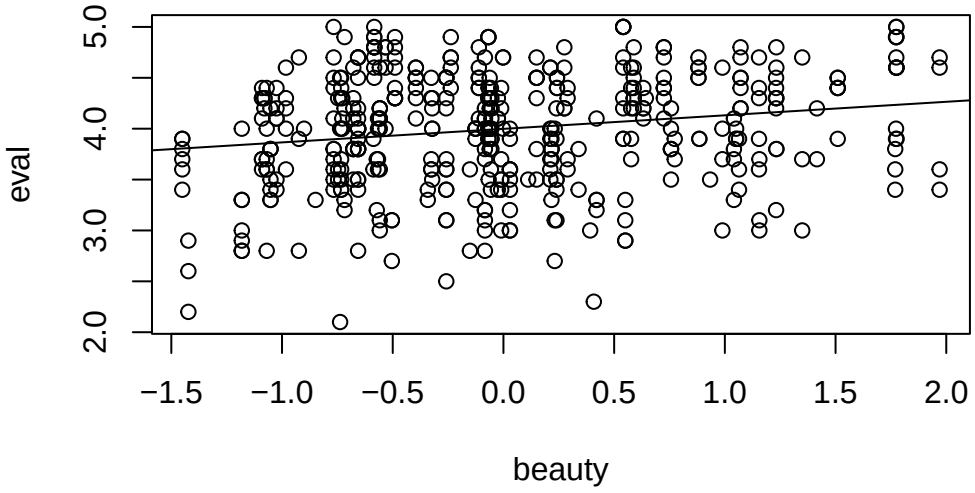
Figure 3: DON'T FORGET..

Table 1: DON'T FORGET

Gender	Beauty	Eval
Male	-0.08448224	4.069030
Female	0.11610907	3.901026

DON'T FORGET..

According to (**table-overall-beauty-and-eval?**) there is a correlation between the beauty of the instructor and their evaluation. However, what had the biggest impact? The slope of the line for both male and female instructors is 0.133, meaning that every increase in beauty should yield an evaluation higher by 0.133. Additionally the r^2 value is 0.03364 or 3% , indicating that evaluation can be explained by the variability in beauty, which is the variability of y that is explained by the x. The r^2 indicates how much it affects the evaluation.




```
Call:
lm(formula = eval ~ beauty, data = TeachingRatings)

Residuals:
```

```
     Min       1Q   Median       3Q      Max
-1.80015 -0.36304  0.07254  0.40207  1.10373

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.99827    0.02535 157.727  < 2e-16 ***
beauty       0.13300    0.03218   4.133 4.25e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5455 on 461 degrees of freedom
Multiple R-squared:  0.03574,   Adjusted R-squared:  0.03364
F-statistic: 17.08 on 1 and 461 DF,  p-value: 4.247e-05
```

The table (**coefficients-affecting-eval?**) shows off all of the coefficients and how they respectively affect the evaluation of the instructor. The p-value is the indicator of how likely something is due to pure randomness, and so a low p-value means that something is very unlikely to be due to pure chance. For all of the variables that affect the instructors evaluation, the p-value is 7.822e-15, which is really small. This indicates that the variables present do in fact affect the evaluation of an instructor, and there is a small chance that it is due to randomness. The estimate shows how much each variable is able to affect the evaluation of an instructor as it increases. The p-value for each again indicates how likely it is that each variable could be affecting the evaluation due to pure chance. Based on the data I use the 95% rule to say that there is less than 5% that the strong affecting variables are not due to chance and they have a great impact, these include: minority, gender, credits, beauty, native, students, allstudents.

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
(Intercept)	13.1386716936	0.3096867649
minorityyes	-1.6043289751	0.4520520991
age	-0.1020625493	0.0418871547
genderfemale	0.6301714363	0.3274755674
creditssingle	0.3926076356	0.6132490428

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
beauty	-1.3043190049	6.662269728
divisionlower	-0.0584321797	0.073322770
nativeno	0.0466903767	7.188884419
tenureyes	-3.7843180138	14.492788250
students	-0.0003342839	0.002452138
allstudents	-0.0029406024	0.001431006
prof2	-0.6494872450	1.842894885
prof3	-1.0498918905	4.064545000
prof4	-2.6839369475	9.390106155
prof5	-0.3245685038	1.462649369
prof6	3.0135233139	9.304811369
prof7	-2.8180458240	8.645891942
prof8	-0.9993020564	2.000353865
prof9	-2.0263746032	6.809810085
prof10	-2.9915172086	12.814936043
prof11	-0.5674735835	4.090744188

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
prof12	-21.043375955	5.2362882925
prof13	-2.575776160	0.8893127547
prof14	-5.995218546	1.0592935627
prof15	-2.479872762	2.3315373856
prof16	-3.181547403	0.3477871083
prof17	-19.207291381	5.0550548785
prof18	-16.361288768	3.9270390608
prof19	-2.825926848	0.9322029351
prof20	-12.049900488	4.3660729643
prof21	-12.746012003	4.5256696850
prof22	-13.806587188	3.1912761616
prof23	0.3225346547	7.628250794
prof24	1.6081037052	2.472469925
prof25	0.1573405242	2.464787433
prof26	0.4967518100	0.623250549
prof27	1.8327860028	2.947853504
prof28	1.1144267614	2.209644613
prof29	0.9001249972	1.866007713

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
prof30	-2.9289497737	1.790283121
prof31	-2.9770805785	9.780532419
prof32	-1.9190174706	8.080780666
prof33	-0.0535542018	1.446482756
prof34	0.0492405402	5.016055025
prof35	-0.9337593884	6.189451443
prof36	1.3231659198	1.258219750
prof37	-0.6304171823	4.434034968
prof38	0.1984292659	9.612422527
prof39	0.7934674462	2.939091610
prof40	-0.8055888495	0.434149061
prof41	-0.6729397571	5.602840465
prof42	-0.7982813104	5.126298203
prof43	-3.0353337497	6.679813949
prof44	1.1035560455	0.928010597
prof45	1.7699294815	0.555892256
prof46	-0.3622876285	1.239930864

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
prof47	-3.439298619	0.3107110065
prof48	-0.316872098	1.0727251233
prof49	-19.557437650	5.6826891686
prof50	-4.203355656	0.6246915724
prof51	-2.992747165	1.0313551386
prof52	1.741164912	5.6486981340
prof53	0.071913256	1.4455067503
prof54	-17.054024139	4.4452412871
prof55	-0.431898387	0.9945627463
prof56	-10.376511130	2.0520824308
prof57	-1.889015612	0.3872227745
prof58	-7.344833387	2.0562549322
prof59	-12.415657527	2.7999304530
prof60	-1.032359392	1.2816454674
prof61	-10.017887160	1.9744361650
prof62	-1.070389580	0.1425869053

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
prof63	0.4275225385	1.322962510
prof64	0.5248497977	7.251376471
prof65	-5.6579335838	21.600435664
prof66	-0.4879983255	1.956584077
prof67	1.5875309118	8.774077401
prof68	-2.4585655898	5.966855861
prof69	-1.0555422626	2.389081180
prof70	-1.1839153300	6.882657864
prof71	-2.6727492968	14.555725034
prof72	0.9727224945	0.223849077
prof73	3.4017002516	6.652151198
prof74	-2.0374717657	9.478659233
prof75	-0.9431379285	0.457967117
prof76	-0.4763092969	0.359222674
prof77	-7.0507263057	27.045084128
prof78	-0.8982174115	3.778859863
prof79	-0.4938072390	2.725697900

Table 2: DON'T FORGET

Coefficient	Estimate	Standard Error
prof80	0.4676497206	0.3331758207
prof81	-4.8609457426	0.818294432
prof82	-0.3883115352	2.869961749
prof83	-3.4104781146	8.770241964
prof84	3.1205716861	13.703558236
prof86	1.6174389467	6.792994874
prof87	-0.3530190754	2.981398513
prof88	-1.0902594625	3.536027639
prof89	1.4308337819	12.330021353

```
Call:
lm(formula = eval ~ ., data = df_predictions)

Residuals:
     Min       1Q   Median       3Q      Max
-1.46212 -0.19511  0.00898  0.18983  1.00008

Coefficients: (6 not defined because of singularities)
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   13.1386717 36.0968676   0.364 0.716081
minorityyes   -1.6043290  4.5205210  -0.355 0.722870
age           -0.1020625  0.4188715  -0.244 0.807630
genderfemale   0.6301714  0.2747557   2.294 0.022383 *
creditssingle  0.3926076  0.1324904   2.963 0.003243 **
beauty        -1.3043190  6.6622697  -0.196 0.844894
divisionlower -0.0584322  0.0733228  -0.797 0.426017
nativeno       0.0466904  3.1888844   0.015 0.988326
```

```
tenureyes     -3.7843180 14.4927882  -0.261 0.794149
students      -0.0003343  0.0024521  -0.136 0.891641
allstudents   -0.0029406  0.0014310  -2.055 0.040598 *
prof2         -0.6494872  1.8428949  -0.352 0.724721
prof3         -1.0498919  4.0645450  -0.258 0.796317
prof4         -2.6839369  9.3901062  -0.286 0.775173
prof5         -0.3245685  1.4626494  -0.222 0.824512
prof6          3.0135233  8.3048114   0.363 0.716916
prof7         -2.8180458  8.6458919  -0.326 0.744656
prof8         -0.9993021  2.0003539  -0.500 0.617684
prof9         -2.0263746  6.8098101  -0.298 0.766203
prof10        -2.9915172 12.8149360  -0.233 0.815551
prof11        -0.5674736  4.0907442  -0.139 0.889747
prof12        -5.2362883 21.0433760  -0.249 0.803630
prof13        -0.8893128  2.5757762  -0.345 0.730098
prof14        -1.0592936  5.9952185  -0.177 0.859850
prof15        -2.3315374  2.4798728  -0.940 0.347745
prof16        -0.3477871  3.1815474  -0.109 0.913014
prof17        -5.0550549 19.2072914  -0.263 0.792557
prof18        -3.9270391 16.3612888  -0.240 0.810449
prof19        -0.9322029  2.8259268  -0.330 0.741684
prof20        -4.3660730 12.0499005  -0.362 0.717313
prof21        -4.5256697 12.7460120  -0.355 0.722746
prof22        -3.1912762 13.8065872  -0.231 0.817334
prof23         0.3225347  0.6282508   0.513 0.607993
prof24         1.6081037  4.4724699   0.360 0.719387
prof25         0.1573405  4.4647874   0.035 0.971907
prof26         0.4967518  0.6232505   0.797 0.425949
prof27         1.8327860  2.9478535   0.622 0.534504
prof28         1.1144268  4.2096446   0.265 0.791366
prof29         0.9001250  2.8660077   0.314 0.753648
prof30        -2.9289498  1.7902831  -1.636 0.102696
prof31        -2.9770806  9.7805324  -0.304 0.761005
prof32        -1.9190175  8.0807807  -0.237 0.812418
prof33        -0.0535542  1.4464828  -0.037 0.970486
prof34         0.0492405  2.0160550   0.024 0.980528
prof35        -0.9337594  6.1894514  -0.151 0.880167
prof36         1.3231659  8.2582198   0.160 0.872793
prof37        -0.6304172  4.4340350  -0.142 0.887019
prof38         0.1984293  0.6124225   0.324 0.746118
prof39         0.7934674  0.9390916   0.845 0.398703
prof40        -0.8055888  0.4341491  -1.856 0.064322 .
prof41        -0.6729398  5.6028405  -0.120 0.904465
```

```
prof42       -0.7982813  5.1262982  -0.156 0.876338
prof43       -3.0353337  6.6798139  -0.454 0.649808
prof44        1.1035560  5.9280106   0.186 0.852423
prof45        1.7699295  5.5558923   0.319 0.750236
prof46       -0.3622876  1.2399309  -0.292 0.770312
prof47       -0.3107110  3.4392986  -0.090 0.928065
prof48       -1.0727251  0.3168721  -3.385 0.000788 ***
prof49       -5.6826892 19.5574376  -0.291 0.771550
prof50       -0.6246916  4.2033557  -0.149 0.881938
prof51       -1.0313551  2.9927472  -0.345 0.730580
prof52        1.7411649  6.4869813   0.268 0.788536
prof53        0.0719133  1.4550675   0.049 0.960610
prof54       -4.4452413 17.0540241  -0.261 0.794504
prof55       -0.9945627  0.4318984  -2.303 0.021854 *
prof56       -2.0520824 10.3765111  -0.198 0.843341
prof57       -0.3872228  1.8890156  -0.205 0.837697
prof58       -2.0562549  7.3448334  -0.280 0.779667
prof59       -2.7999305 12.4156575  -0.226 0.821704
prof60       -1.2816455  1.0323594  -1.241 0.215229
prof61       -1.9744362 10.0178872  -0.197 0.843866
prof62       -0.1425869  1.0703896  -0.133 0.894100
prof63        0.4275225  5.3229625   0.080 0.936029
prof64        0.5248498  7.2513765   0.072 0.942340
prof65       -5.6579336 21.6004357  -0.262 0.793518
prof66       -0.4879983  1.9565841  -0.249 0.803181
prof67        1.5875309  8.7740774   0.181 0.856520
prof68       -2.4585656  5.9668559  -0.412 0.680554
prof69       -1.0555423  2.3890812  -0.442 0.658882
prof70       -1.1839153  6.8826579  -0.172 0.863522
prof71       -2.6727493 14.5557250  -0.184 0.854412
prof72        0.9727225  5.2238491   0.186 0.852385
prof73        3.4017003  5.6521512   0.602 0.547653
prof74       -2.0374718  9.4786592  -0.215 0.829924
prof75       -0.9431379  0.4579671  -2.059 0.040165 *
prof76       -0.4763093  0.3592227  -1.326 0.185687
prof77       -7.0507263 27.0450841  -0.261 0.794469
prof78       -0.8982174  3.7788599  -0.238 0.812251
prof79       -0.4938072  2.7256979  -0.181 0.856337
prof80        0.4676497  6.3317582   0.074 0.941164
prof81       -4.8609457 20.8182944  -0.233 0.815509
prof82       -0.3883115  2.8699617  -0.135 0.892448
prof83       -3.4104781  8.7702420  -0.389 0.697600
prof84        3.1205717 13.7035582   0.228 0.819992
```

```
prof85                NA         NA      NA       NA
prof86         1.6174389  7.7929949   0.208 0.835696
prof87        -0.3530191  2.9813985  -0.118 0.905810
prof88        -1.0902595  3.5360276  -0.308 0.758008
prof89         1.4308338  9.3300214   0.153 0.878201
prof90                NA         NA      NA       NA
prof91                NA         NA      NA       NA
prof92                NA         NA      NA       NA
prof93                NA         NA      NA       NA
prof94                NA         NA      NA       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3908 on 365 degrees of freedom
Multiple R-squared:  0.6082,    Adjusted R-squared:  0.504
F-statistic:  5.84 on 97 and 365 DF,  p-value: < 2.2e-16
```

Table (**strong-coefficients-affecting-eval?**) shows all of the variables that affect an instructor's evaluation the most. The estimate is how much the evaluation is how much the evaluation will either increase or decrease based on the change in the variable, and the standard error is how likely it is to be due to randomness.

```
Call:
lm(formula = eval ~ minority + gender + credits + beauty + native +
    students + allstudents, data = df_predictions)

Residuals:
     Min       1Q   Median       3Q      Max
-1.93274 -0.32062  0.03453  0.36360  1.08001

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.062923   0.040380 100.616  < 2e-16 ***
minorityyes   -0.162284   0.075036  -2.163 0.031080 *
genderfemale  -0.190863   0.049250  -3.875 0.000122 ***
creditssingle  0.622270   0.105846   5.879 7.99e-09 ***
beauty         0.155475   0.030818   5.045 6.57e-07 ***
nativeno      -0.259451   0.104557  -2.481 0.013446 *
students       0.007198   0.002267   3.175 0.001601 **
allstudents   -0.004459   0.001360  -3.279 0.001121 **
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5081 on 455 degrees of freedom
Multiple R-squared:  0.1741,    Adjusted R-squared:  0.1614
F-statistic:  13.7 on 7 and 455 DF,  p-value: 4.034e-16
```

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
(Intercept)	13.1386716936	0.8096867649
minorityyes	-1.6043289751	0.4520520991
age	-0.1020625493	0.0418871547
genderfemale	1.6301714363	0.3274755674
creditssingle	0.3926076356	0.6132490428
beauty	-1.3043190049	0.6622697280
divisionlower	-0.0584321797	0.0733227700
nativeno	0.4669037637	0.7188884419
tenureyes	-3.7843180138	0.4927882500
students	-0.0003342839	0.0024521380
allstudents	-0.0029406024	0.0014310060
prof2	-0.6494872450	1.8428948850
prof3	-1.0498918905	4.0645450000
prof4	-2.6839369475	9.3901061550

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
prof5	-0.3245685038	1.462649369
prof6	3.0135233139	9.304811369
prof7	-2.8180458240	8.645891942
prof8	-0.9993020564	2.000353865
prof9	-2.0263746032	6.809810085
prof10	-2.9915172086	12.814936043
prof11	-0.5674735835	4.090744188
prof12	-5.2362882925	21.043375955
prof13	-0.8893127547	2.575776160
prof14	-1.0592935627	5.995218546
prof15	-2.3315373856	2.479872762
prof16	-0.3477871083	3.181547403
prof17	-5.0550548785	19.207291381
prof18	-3.9270390608	16.361288768
prof19	-0.9322029351	2.825926848

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
prof20	-4.3660729643	12.049900488
prof21	-4.5256696850	12.746012003
prof22	-3.1912761616	13.806587188
prof23	0.3225346547	7.628250794
prof24	1.6081037054	2.472469925
prof25	0.1573405244	2.464787433
prof26	0.4967518100	0.623250549
prof27	1.8327860028	8.947853504
prof28	1.1144267614	2.209644613
prof29	0.9001249972	1.866007713
prof30	-2.9289497737	1.790283121
prof31	-2.9770805785	9.780532419
prof32	-1.9190174706	8.080780666
prof33	-0.0535542018	1.446482756
prof34	0.0492405402	2.016055025
prof35	-0.9337593884	6.189451443
prof36	1.3231659198	1.258219750
prof37	-0.6304171823	4.434034968
prof38	0.1984292659	9.612422527

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
prof39	0.7934674462	2939091610
prof40	-0.8055888495	0.434149061
prof41	-0.6729397571	5.602840465
prof42	-0.7982813104	5.126298203
prof43	-3.0353337497	6.679813949
prof44	1.1035560450	0928010597
prof45	1.7699294815	0555892256
prof46	-0.3622876285	1.239930864
prof47	-0.3107110065	3.439298619
prof48	-1.0727251233	0.316872098
prof49	-5.6826891686	19.557437650
prof50	-0.6246915724	4.203355656
prof51	-1.0313551386	2.992747165
prof52	1.7411649126	6486981340
prof53	0.0719132561	4455067503
prof54	-4.4452412871	17.054024139
prof55	-0.9945627463	0.431898387

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
prof56	-10.3765111302.0520824308	
prof57	-1.8890156120.3872227745	
prof58	-7.3448333872.0562549322	
prof59	-12.4156575272.7999304530	
prof60	-1.0323593921.2816454674	
prof61	-10.0178871601.9744361650	
prof62	-1.0703895800.1425869053	
prof63	0.427522538	1.322962510
prof64	0.524849797	7.251376471
prof65	-21.6004356645.6579335838	
prof66	-1.9565840770.4879983255	
prof67	1.587530911	8.774077401
prof68	-5.9668558612.4585655898	
prof69	-2.3890811801.0555422626	
prof70	-6.8826578641.1839153300	
prof71	-14.5557250342.6727492968	
prof72	0.972722494	0.223849077

Table 3: DON'T FORGET

Coefficient	Estimate	Standard Error
prof73	3.4017002516	6652151198
prof74	-2.0374717657	9.478659233
prof75	-0.9431379285	0.457967117
prof76	-0.4763092969	0.359222674
prof77	-7.0507263057	27.045084128
prof78	-0.8982174115	3.778859863
prof79	-0.4938072390	2.725697900
prof80	0.4676497206	3331758207
prof81	-4.8609457426	20.818294432
prof82	-0.3883115352	2.869961749
prof83	-3.4104781146	8.770241964
prof84	3.1205716861	13.703558236
prof86	1.6174389466	6792994874
prof87	-0.3530190754	2.981398513
prof88	-1.0902594625	3.536027639
prof89	1.4308337812	2330021353

4 Discussion

#4.1 Beauty and Evaluation After reviewing all of the data, my initial hypothesis that beauty does influence an instructor's evaluation stands true, as the data shows that for each unit increase in beauty should result in a corresponding increase in evaluation score of 0.16. However, other factors also play a significant factor to the instructors evaluation.

4.2 Male and Female Instructors

From the findings, it has been shown that female instructors receive lower evaluations than male instructors, while holding all other variables constant. The average female evaluation was 3.90, while the average male evaluation was 4.07, despite females scoring higher in terms of beauty than males. The difference is small, however statistically significant, as it contradicts my hypothesis of higher beauty leading to a higher evaluation. However, more data and research would need to be done in order to explore the possible reasons, and see if the trend continues.

4.3 Class Size

Another factor which affected the evaluation of instructors, was the class size. The data shows that instructors who teach larger classes receive higher evaluations, than those who teach smaller ones. However, the following trend is seen to take less of an effect as class sizes continue to increase. As such, there is a positive correlation between the class size and evaluation, but this effect is limited. # 4.4 Minorities Instructors who are minorities are seen to receive lower evaluations than those who are not a minority. The data shows that minority instructors received evaluations that are 0.16 units lower on average than those instructors who are not a minority.

4.5 Native

Based on the data, instructors who are native receive lower evaluations than those who are not. The coefficient estimate for "nativeyes" is 0.259451, meaning that being a native instructor would increase the evaluation score by 0.259451 units, as compared to those who are non native. This effect has a significant effect on the evaluation of the instructors with a p-value of 0.013446. However, according to the data, non-native instructors actually receive higher evaluations than their native counterparts. This seems counter intuitive that while the estimate is positive, they receive lower evaluations, however, the estimate is holding all other variables constant which means that the estimate is the difference in the evaluation is everything else about two instructors remains identical. As such, though the estimate is positive, the actual

data shows that the native instructors actually receive lower evaluations that the non-native instructors.

4.1 Limitations

There are several limitations to consider from the following study. Firstly the study was conducted at one university, and therefore cannot be a generalization of other universities or students. Additionally, the study is relatively, old given by the fact that it was conducted in the years 2000-20002, and cannot reflect the changes in student demographics, teaching methodologies, and evaluation practices since then. Additionally, the method of data collection is unknown, therefore the reliability of the data is unknown. Moreover, the grading scale which was used in the study ranges from -2.0 to 2.0 for measuring beauty, which could potentially be unfamiliar to some and cause invalid ratings. The measure of beauty in the study is also based on personal preference of the students which were assessing the instructors beauty, which cannot be a unified measurement, and would vary depending on which students were chosen to evaluate the instructors beauty. This could potentially introduce variability and bias in the study.

Going back to the results of the native instructors versus the non-native ones. While the "native-yes" estimate is positive, indicating that native instructors should receive a higher evaluation, the actual data shows they receive a lower one. This means there could be some hidden variables which have not been accounted for such as the instructors language proficiency, that may also affect the students evaluation of their instructors.

Lastly, the study lacks information about the students themselves which carried out the evaluations. Cross referencing instructor evaluations to student data such as their age, gender, academic performance could have significant impact on the instructor. For example female students could find a male instructor highly attractive, but if only 6 male students were chosen to evaluate their beauty, then the data could be skewed. Additionally, students which do well in a course are more likely to review their instructor positively, while those who performed poorly tend to leave more negative evaluations. Leaving an abnormally negative score, could lead to the results being inaccurate, and without said information it is hard to understand the factors which influence an instructor's evaluation.

5 Appendix