

CheckPoint2

Rodolfo Viana

01-06-2015

Vários cursos pedem que os alunos avaliem o professor no final de um período/curso. Porém, algumas pessoas acreditam que algumas características não relacionadas à docência podem influenciar no resultado. No artigo intitulado “Beauty in the classroom: instructors’ pulchritude and putative pedagogical productivity” (Hamermesh & Parker, 2005) o autor reforça a ideia que professores vistos como tendo uma melhor aparência recebem avaliações mais altas.

Utilizando os dados das avaliações dos discentes da Universidade do Texas em Austin gostaríamos de saber se existem indícios de que a percepção da beleza de um professor impacta na avaliação de sua docência

```
library(dplyr)
```

```
evals <- read.csv("~/Projetos/DataAnalysis/Assignment4/evals.csv", head=TRUE)
```

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

```
##
## Call:
## lm(formula = score ~ bty_avg, data = evals)
##
## 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 ***
## 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
```

Com isso temos indícios de que a beleza influencia na avaliação da docência do professor (pois o $\Pr(>|t|)$ foi igual à $5.08e-05$). Porém acreditamos que existem diversos outros fatores que impactam na avaliação da docência do professor, para se ter uma ideia, a beleza para essa nossa amostra e modelo, influencia apenas 3% na avaliação da docência do professor (o R-squared foi igual à 0.03293).

Em busca de encontrar mais fatores que influenciam na avaliação da docência do professor criamos novos modelos. O próximo modelo que vamos avaliar é lavando em consideração a percepção da beleza e a idade.

```
mod2 <- lm(score ~ bty_avg + age, data = evals)
summary(mod2)
```

```
##
## Call:
## lm(formula = score ~ bty_avg + age, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9427 -0.3474  0.1293  0.3957  0.9478
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.054732   0.169865  23.870  < 2e-16 ***
## bty_avg      0.060656   0.017098   3.548 0.000429 ***
## age         -0.003059   0.002664  -1.148 0.251396
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5347 on 460 degrees of freedom
## Multiple R-squared:  0.03778,    Adjusted R-squared:  0.0336
## F-statistic: 9.031 on 2 and 460 DF,  p-value: 0.0001422
```

Com esse novo modelo foi observado que a percepção da beleza continua com indícios de que influência na avaliação da docência do professor, porém não podemos dizer o mesmo da idade do professor (pois o $\text{Pr}(>|t|)$ foi igual à 0.251396)

O próximo modelo que vamos avaliar é levando em consideração a percepção da beleza e o nível do professor (horista, assistente, titular).

```
mod3 <- lm(score ~ bty_avg + rank, data = evals)
summary(mod3)
```

```
##
## 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 ***
## ranktenured    -0.12623    0.06266  -2.014  0.0445 *
## ranktenure track -0.16070    0.07395  -2.173  0.0303 *
## ---
## 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
```

Com esse modelo foi observado que a percepção da beleza continua com indícios de que influência na avaliação da docência do professor, além disso existe indícios de que o nível do professor influencia na avaliação do professor. Quando o professor tem o nível de assistente ou titular temos um impacto negativo na avaliação do professor (-0.12623 e -0.16070 respectivamente). Dentre os níveis do professor horista foi mais bem avaliado.

Com esse modelo temos uma influência de apenas 4% na avaliação do professor, por esse motivo criamos um novo modelo, dessa vez levando em consideração a percepção da beleza, nível do professor e a lingua da universidade frequentada pelo professor (inglês ou não inglês)

```
mod4 <- lm(score ~ bty_avg + rank + language, data = evals)
summary(mod4)
```

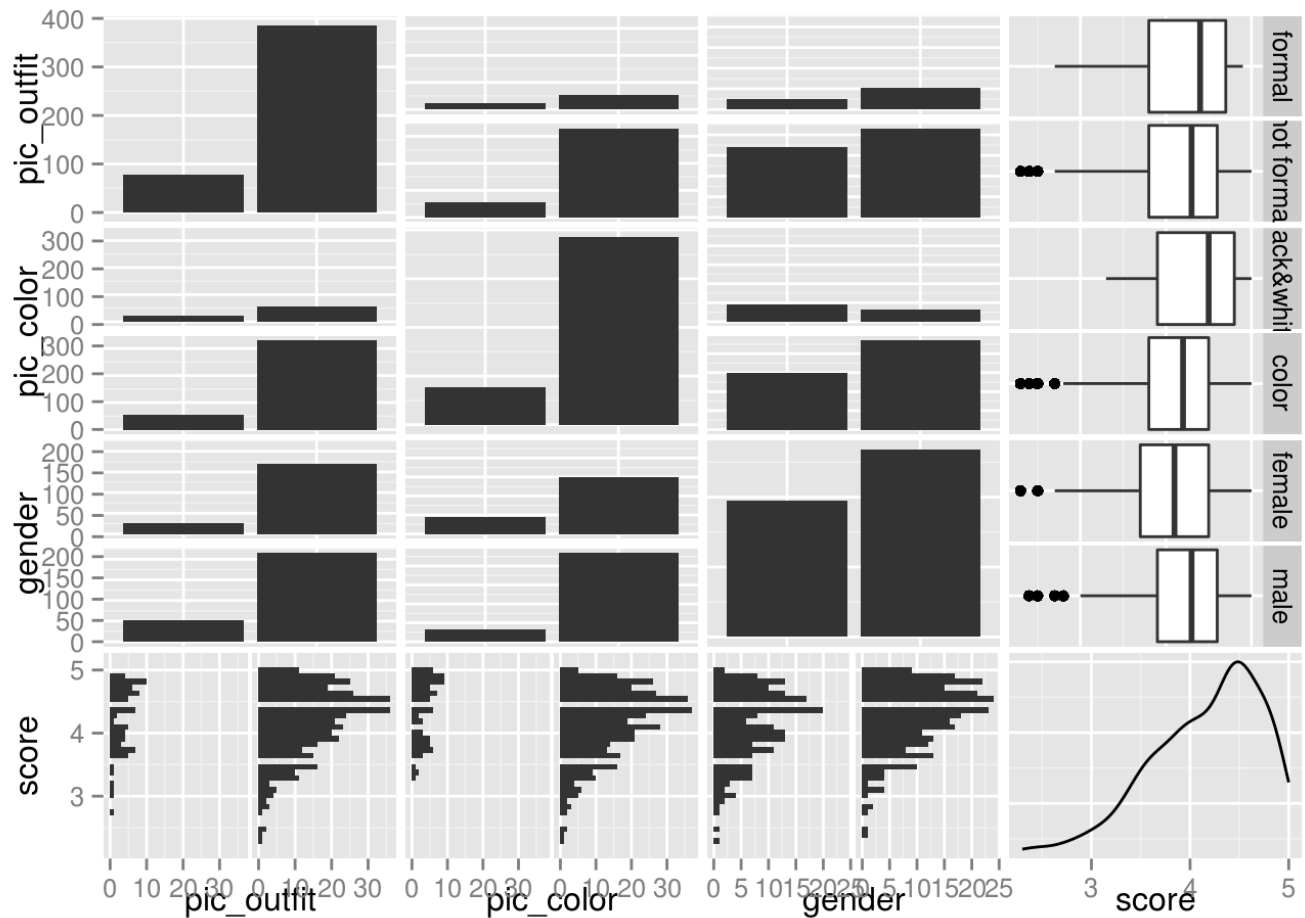
```
##
## Call:
## lm(formula = score ~ bty_avg + rank + language, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9061 -0.3392  0.1412  0.4077  0.9418
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      3.98492    0.09046  44.050 < 2e-16 ***
## bty_avg           0.06707    0.01649   4.066 5.62e-05 ***
## ranktenured      -0.11677    0.06260  -1.865  0.0628 .
## ranktenure track -0.12537    0.07560  -1.658  0.0980 .
## languagenon-english -0.22227    0.10653  -2.087  0.0375 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5309 on 458 degrees of freedom
## Multiple R-squared:  0.0555, Adjusted R-squared:  0.04725
## F-statistic: 6.728 on 4 and 458 DF,  p-value: 2.874e-05
```

Com esse modelo temos indícios de que a percepção da beleza influencia na avaliação da docência do professor, além disso para os docentes que frequentaram uma universidade não inglesa temos indícios de que influencia negativamente na avaliação da docência do professor. Além disso agora temos indícios de que o nível do professor não influencia na avaliação da docência do professor ($\text{Pr}(>|t|)$ igual à 0.0628 e 0.0980).

Nesse modelo temos o R-squared igual à 4%. O que significa que as variáveis escolhidas explicam em 4% a variável resposta.

Antes de escolher uma nova variável para o nosso modelos, foi feita uma análise com 3 variáveis. Para que dessa forma talvez aumente significativamente o valor de R-squared. As variáveis escolhidas foram `pic_outfit` (roupa do professor na foto avaliada: informal, formal), `pic_color` (cor da foto avaliada: colorida, preto e branco) e por último `gender` (sexo do professor: feminino, masculino).

```
library(GGally)
ggpairs(select(evals, pic_outfit, pic_color, gender, score))
```



O gráfico acima possui diversas informações, para essa parte do problemas focamos apenas na busca de alguma variável que tenha alguma impacto na variável score. Por essa razão escolhemos a variável pic_color por ser a varável que possui maior diferença no valor do 1,2,3,4 quartil.

```
mod5 <- lm(score ~ bty_avg + language + pic_color, data = evals)
summary(mod5)
```

```
##
## Call:
## lm(formula = score ~ bty_avg + language + pic_color, data = evals)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9024 -0.3468  0.1432  0.3897  0.8980
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    4.09942    0.10916  37.556 < 2e-16 ***
## bty_avg         0.05477    0.01679   3.262  0.00119 **
## languagenon-english -0.28122    0.10364  -2.713  0.00691 **
## pic_colorcolor  -0.17998    0.06882  -2.615  0.00921 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5286 on 459 degrees of freedom
## Multiple R-squared:  0.06134,    Adjusted R-squared:  0.0552
## F-statistic: 9.998 on 3 and 459 DF,  p-value: 2.148e-06
```

Com esse modelo temos que é alta a significância de que é pouco provável o fato de que não exista nenhuma relação entre as variáveis língua da universidade frequentada pelo professor, cor da foto avaliada e percepção da beleza com a avaliação da docência do professor.

Nesse modelo temos o R-squared igual à 5%. O que significa que as variáveis escolhidas explicam em 5% a variável resposta. O que é considerado um valor baixo e ao mesmo tempo esperado, pois é considerado que a avaliação da docência do professor é algo bem complexo, que naturalmente não vai ser explicado completamente apenas com as variáveis disponíveis no nosso conjunto de dados.