Stat 230 Group HW 2 - solution

This homework solution is for the sole benefit of students taking Stat 230 from Prof. St. Clair during Fall term 2021. Dissemination of this solution to people who are not registered for this course is not permitted and will be considered grounds for Academic Dishonesty for the all individuals involved in the giving and receiving of the solution.

Due: 3pm Friday, Sept 24

Grading due: noon Sunday, Sept 26

Use this Markdown template to answer the questions below. Knit to a Word or pdf doc and upload this doc to the group homework forum. Grader comments can be added as comments to this forum post.

Problem: Teacher evaluations

This problem looks at end of semester student evaluations for 463 courses taught by a sample of 94 professors from the University of Texas at Austin. We want to investigate a possible relationship between a teacher's physical appearance and their teaching evaluation, both evaluated by students. The variables we will consider in this HW are:

- `score`: Average professor evaluation score, from (1) very unsatisfactory - (5) excellent
- `bty_avg`: Average beauty rating of professor, from (1) lowest - (10) highest

```
evals <- read.csv("https://math.carleton.edu/kstclair/data/teacher_evals.csv")
str(evals)
## 'data.frame':    463 obs. of  4 variables:
##  $ prof_id: int  1 1 1 1 2 2 2 3 3 4 ...
##  $ score  : num  4.7 4.1 3.9 4.8 4.6 4.3 2.8 4.1 3.4 4.5 ...
##  $ bty_avg: num  5 5 5 5 3 ...
##  $ gender : chr  "female" "female" "female" "female" ...
```

(a) Load the `skimr` package and run `skim_without_charts(evals)` to get quick univariate EDA summaries for this data set. Report the mean values of `score` or `bty_avg` and determine if there are any missing values for these variables. (Note: the `skim` function includes mini-histogram charts but these cause an error when knitting to a pdf format.) *answer:* The mean eval score is 4.17 and the mean beauty score is 4.42. There are no missing cases for either variable (`n_missing` column entries are 0).

```
library(skimr)
skim_without_charts(evals)
```

Table 1: Data summary

Name	evals
Number of rows	463

Number of columns	4
Column type frequency:	
character	1
numeric	3
Group variables	None

Variable type: character

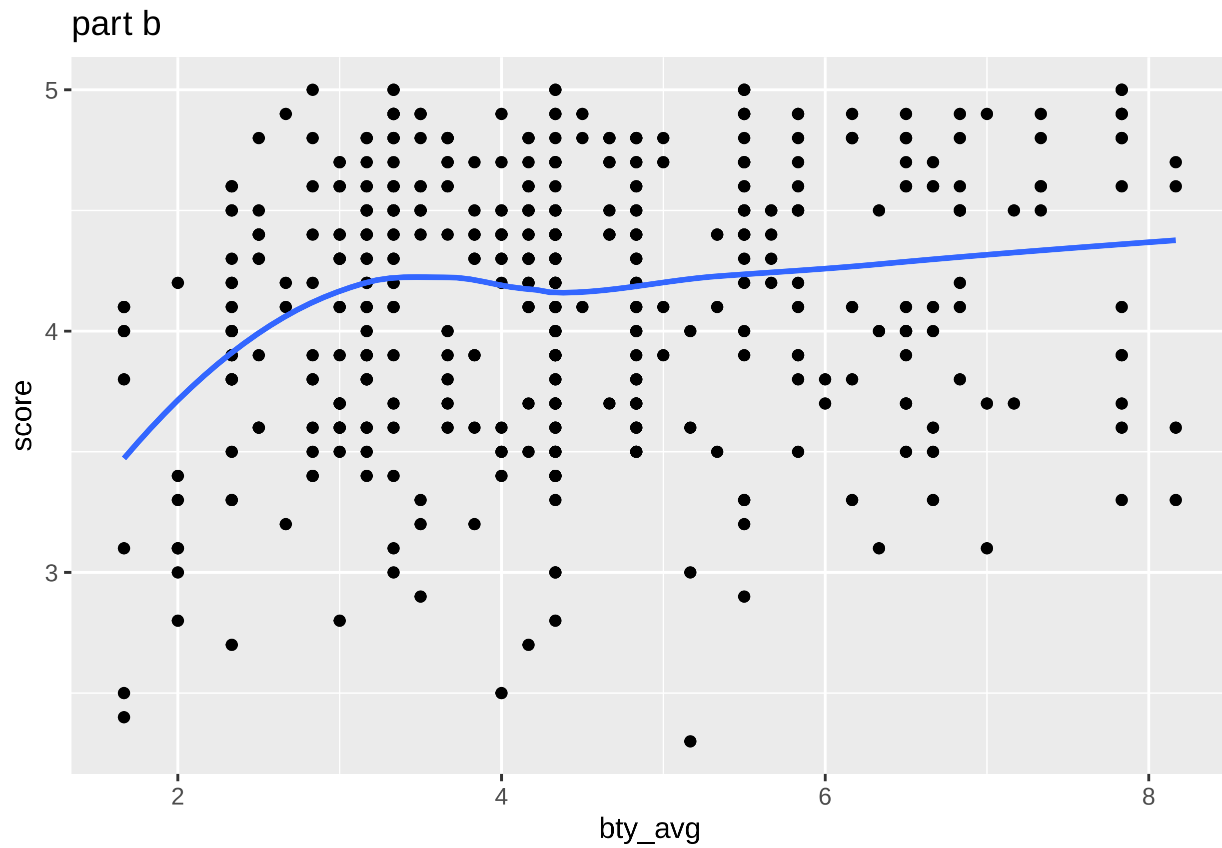
skim_variable	n_missing	complete_rate	min	max	empty	n_unique	whitespace
gender	0	1	4	6	0	2	0

Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100
prof_id	0	1	45.15	27.55	1.00	20.00	43.00	70.5	94.00
score	0	1	4.17	0.54	2.30	3.80	4.30	4.6	5.00
bty_avg	0	1	4.42	1.53	1.67	3.17	4.33	5.5	8.17

(b) Use the `ggplot2` package to create a scatterplot of `score` against `bty_avg` with a `geom_smooth` smoother line ("loess") added without error shading around the line. Comment on the form, direction and strength of the relationship between eval score and beauty score. *answer* - form: shows some curvature but mostly linear - direction: positive - strength: fairly weak, lots of variation in `score` around the smoother line

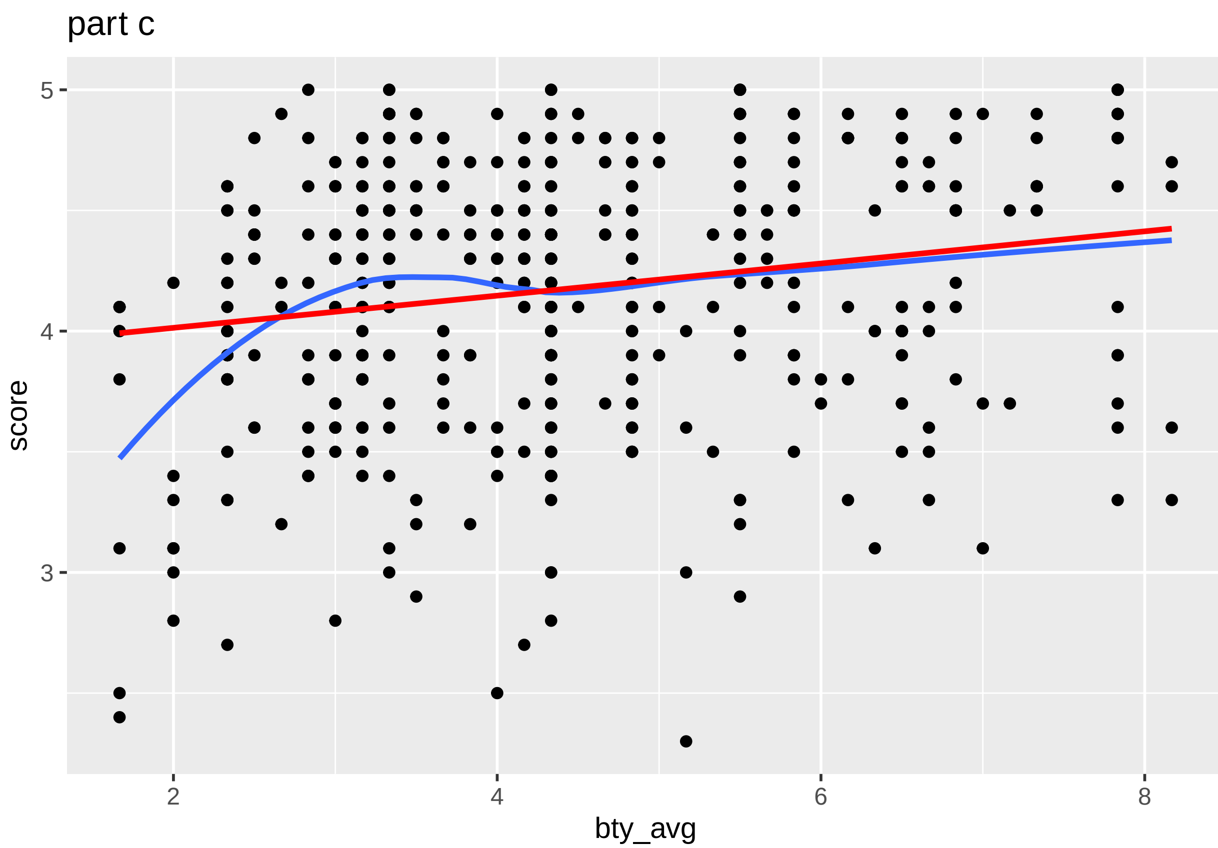
```
library(ggplot2)
ggplot(evals, aes(x = bty_avg, y = score)) +
  geom_point() +
  geom_smooth(se = FALSE)+
  labs(title = "part b")
```



(c) Add a SLR `lm` line to the scatterplot in (b) with an added `color = "red"` argument and exclude the error shading. Your graph should now have 2 lines, one a red SLR line and one a blue loess smoother line. Where does the relationship between beauty and eval score look nonlinear? Is there a range of beauty scores where the relationship looks approximately linear?
answer approximately linear for beauty ratings between 3 and 9, nonlinear below ratings of 3.

Extra comment: this curvature at the "start" of the smoother could be due to the sparseness of low beauty rating data.

```
ggplot(evals, aes(x = bty_avg, y = score)) +
  geom_point() +
  geom_smooth(se = FALSE) +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = "part c")
```

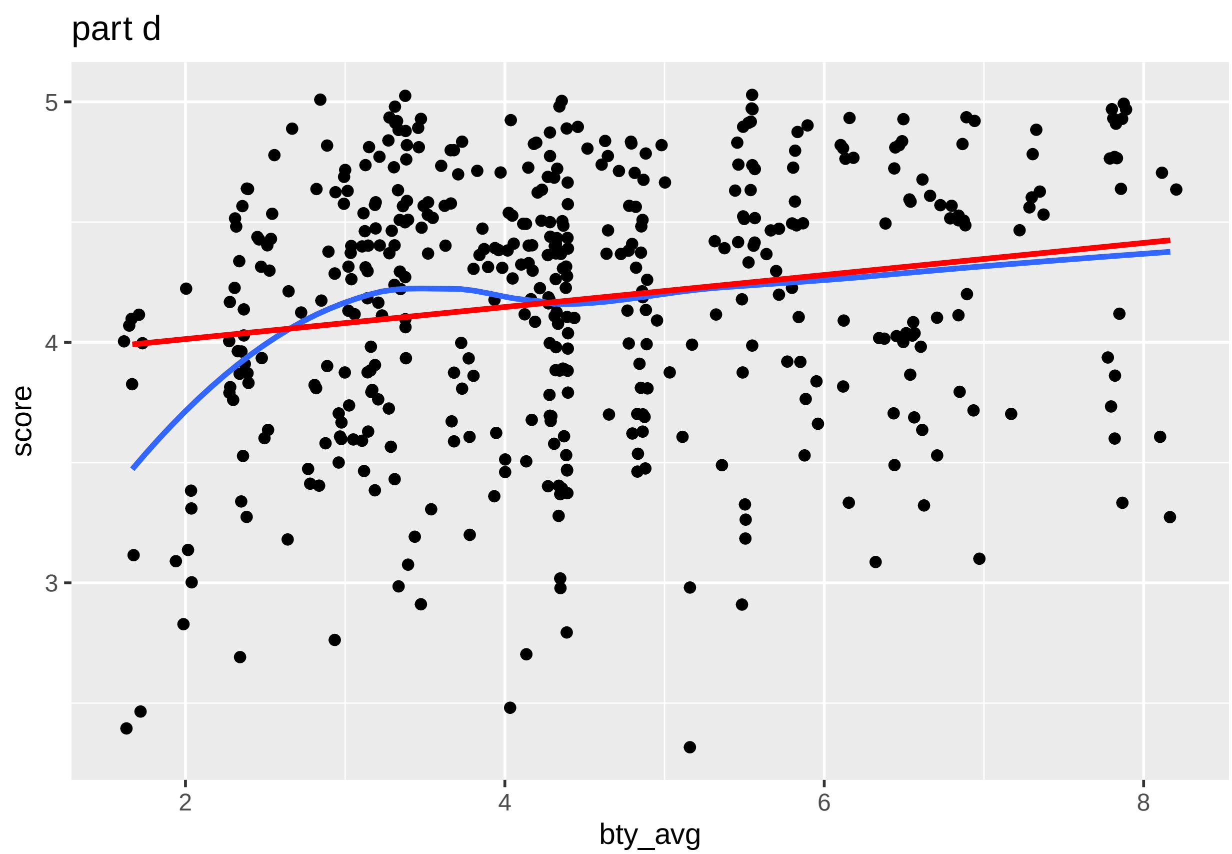


(d) Modify your graph in part (c) to use `geom_jitter` instead of `geom_point`. Explain the difference in plots between (c) and (d) and explain why using `geom_jitter` reveals about the data that `geom_point` does not. *answer* The jittered plot adds a small amount of random error to each x/y value being plotted. This reveals the fact that there are cases in the data that have the same (exact) combinations of x/y (`bty_avg` and `score`) values. For example, there are 3 cases with a beauty average of 5.5 and eval of 5:

```
library(dplyr)
evals %>% filter(bty_avg == 5.5, score == 5)
##   prof_id score bty_avg gender
## 1      10     5     5.5   male
## 2      10     5     5.5   male
## 3      10     5     5.5   male
```

These cases are plotted over one another in `geom_point` which hides the fact that more than one case is represented by one plotting dot. The jittered plot reveals these overplotted cases.

```
ggplot(evals, aes(x = bty_avg, y = score)) +
  geom_jitter() +
  geom_smooth(se = FALSE) +
  geom_smooth(method = "lm", se = FALSE, color = "red") +
  labs(title = "part d")
```



(e) Fit the regression of `score` on `bty_avg` and write down the equation for the estimated mean score regression line and give the estimated model SD. (Note: the limitations of this model will be assessed in the next HW set)

answer:

$$\hat{\mu}_{score|btyavg} = 3.88 + 0.06664(btyavg) \quad \hat{\sigma} = 0.5348$$

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

(f) Use your model summary from (e) to give the p-value and conclusion for the test of $H_O : \beta_1 = 0$ vs. $H_A : \beta_1 \neq 0$ answer: The p-value is 0.00005 which means we have evidence that the *estimated* effect of beauty score on mean evaluation is *statistically significant*.

(g) Math check: show how the test statistic and p-value shown in the model summary for part (e) were computed. answer: test stat:

$$t = \frac{0.06664 - 0}{0.01629} = 4.09 \quad p - value = 2P(T > 4.09) = 2(0.000025) = 0.00005$$

```
.06664/.01629  # test stat
## [1] 4.090853
pt(4.09, df = 463 - 2, lower.tail = FALSE)  # P(T > 4.09)
## [1] 2.545411e-05
1-pt(4.09, df = 463 - 2)  # another way to get P(T > 4.09)
## [1] 2.545411e-05
2*pt(4.09, df = 463 - 2, lower.tail = FALSE)
## [1] 5.090821e-05
```

(h) Use `confint` or `broom::tidy` to get a 95% CI for β_1 and interpret the CI in context with a directional statement. answer: We are 95% confident the a one point increase in beauty score is associated with an increase in mean course evaluation score of 0.035 to 0.099 points.

```
confint(eval_lm)
##                  2.5 %     97.5 %
## (Intercept) 3.73070764 4.02996827
## bty_avg     0.03462292 0.09865116
library(broom)
tidy(eval_lm, conf.int = TRUE)
## # A tibble: 2 x 7
##   term        estimate std.error statistic   p.value conf.low conf.high
##   <chr>          <dbl>     <dbl>     <dbl>     <dbl>    <dbl>     <dbl>
## 1 (Intercept)   3.88      0.0761     51.0  1.56e-191   3.73      4.03
## 2 bty_avg       0.0666    0.0163      4.09 5.08e-  5   0.0346    0.0987
```