

Final Case Study 2017

SEDS 625

due December 8, 2017 11:59PM

Problem Statement

An older female applicant for a job as a teacher (40 years or older is an acceptable division between “older” and “younger” according to the legal system for cases like this) believes she was not hired because of her age. She hires you to help support her lawsuit; her lawyers provide this data set, collected from school files. She accuses the school of preferring to hire younger applicants (which would be illegal). The school denies the charges, saying that any apparent differences, if present, might be explained by other characteristics of the applicants.

You should write a short report, written for the legal team (who typically had one introductory statistics course 20 years ago). Is there evidence to support this case? If so, provide a clear, concise, and coherent explanation to help the legal team (who will be very grateful that you supported the case). If not, provide a clear, concise, and coherent explanation to inform the legal team (who of course will be disappointed).

The Data

```
x <- read.csv("TeacherHires.csv", as.is=TRUE)
head(x)
```

##	interviewed	hired	appdate	age	sex	residence	GPA.u	GPA.g	MA
## 1	yes	no	4/10/04	46	Female	1	N/A	4.03	yes
## 2	yes	no	5/4/01	46	Female	2	3.19	3.55	yes
## 3	yes	no	5/24/04	48	Female	2	2.1	3.3	yes
## 4	yes	no	4/23/03	57	Female	2	3.65	3.97	yes
## 5	yes	no	2/22/02	43	Male	1	2.36	3.56	no
## 6	yes	no	3/21/03	42	Female	2	N/A	3.36	yes
##	substitute	teaching	experience	workkids	volunteer	agegroup			
## 1	no	yes	19	yes	yes	older			
## 2	no	yes	23	yes	yes	older			
## 3	no	no	24	yes	no	older			
## 4	no	yes	13	yes	no	older			
## 5	yes	yes	22	yes	yes	older			
## 6	yes	yes	19	yes	no	older			

Note: the variable **residence** is coded as 1 = “lives in the same town”, 2 = “lives in the state but not in this town”, 3 = “lives out of state”, and 4 = “lives out of country”. The two GPA variables are for undergraduate and graduate grade point averages; **substitute** and **teaching** are indicating whether the applicant has prior experience as a substitute teacher, or as a full-time teacher. Other variables should be evident from the names.

Logistics

You must work alone. If you have any questions, you should email me (Susan) directly; do not use Piazza. I'll answer as a member of the legal team for this applicant (not as a statistician).

Your report should be professionally organized in sections, using our usual Abstract, Data, Analysis, and Conclusion sections. While there is no strict page limit, you should do your best to keep the report concise.

Please: Keep this problem off the web and out of circulation because it provides a really nice example for a course like Case Studies. Please do not share with other students. Please do not post to the web. Please do not share with students who might take my classes in the future. The process of discovery is very important! The “answer” (if there is one) is far less important.

DUE Friday 12/8, 11:59 PM

Use either R Markdown to produce a polished document. All your work must be done in R and be completely reproducible. You must upload your script `.Rmd` ***and*** the resulting PDF to your Dropbox.