UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIEROS INFORMÁTICOS

Máster Universitario en Ciencia de Datos

Homework 1.1: Course Evaluation Dataset Analysis

Xiao Luo
Nabil Aziz
Jiayi Lin
Widad El Abbassi

Madrid, October 20^{th}, 2019

1 Homework 1.1: Course Evaluations

1.1 Data description

The dataset chosen for this data analysis report is **Course Evaluations[1]**, one of many offered on Moodle Platform.

The dataset is made up of 463 observations (rows) of 22 variables. Each observation contains a different course and variety of information related to both the course and the professor who was teaching it.

The course evaluation system is often critizised because of possible influence of non-teaching related characteristics, such as the physical appearance, gender, ethnicity, etc. With this dataset we want to analyse these possible influence on the final evaluation.

1.2 Research questions

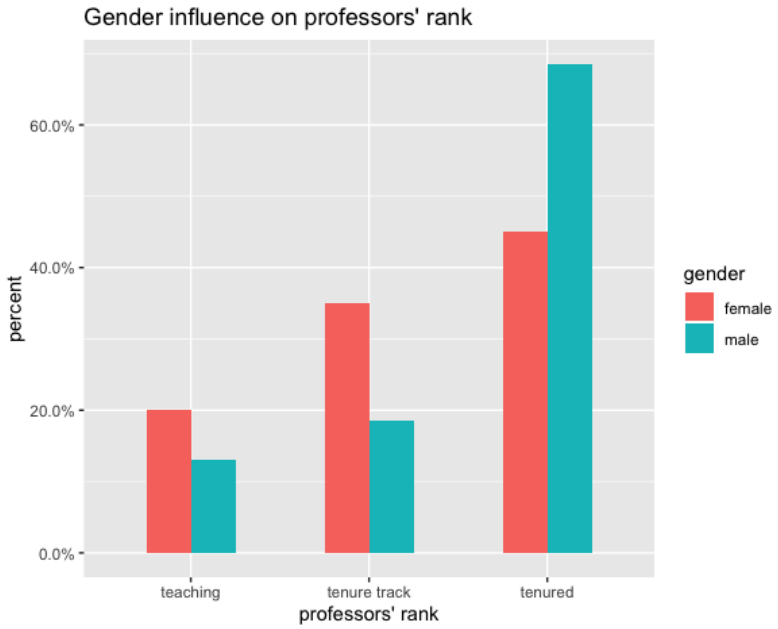
1. Based on these data, is there evidence that professor rank is independent of gender?
2. There are many categorical variables that can have an influence on course evaluations. Focus on two of them.
3. Explore the relationship between course evaluations and a **beauty score** for each professor. Is there evidence of a different relationship depending on gender?
4. Explore the relationship between course evaluations and professors' rank.

1.3 Data preparation and manipulation

- **Data cleaning:** The data cleaning process concluded with 0 changes. The dataset is clean, which means there's no incomplete, incorrect, inconsistent or duplicate data.
- **Aggregating:** Taking into account the dataset structure, for the first research question it is needed to aggregate the dataset by unique professor, since evaluations from different courses of a same professor (stored in different observations) are irrelevant, which means they are duplicates for this specific research question.

1.4 Plots and findings

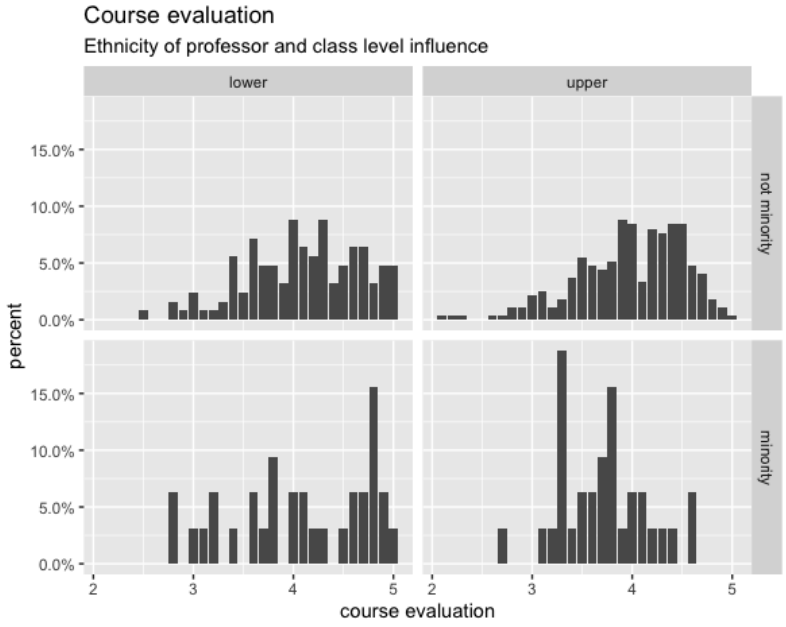
1. Based on these data, is there evidence that professor rank is independent of gender?



With the barplot it's very clear that professor's rank is highly dependent of gender. Female professors' occupation at mid-low rank is by far superior to male professors, but the latter has overwhelming advantage at high rank. This fact could be explained in two different ways:

- Despite of the big number of female professors at tenure track, the amount who gets to tenured rank is quite low.
- Because of the gender equality tendency, the university has already reacted. But due to the extremely slow promotion process, female professors are still on their halfway.

2. There are many categorical variables that can have an influence on course evaluations. Focus on two of them.

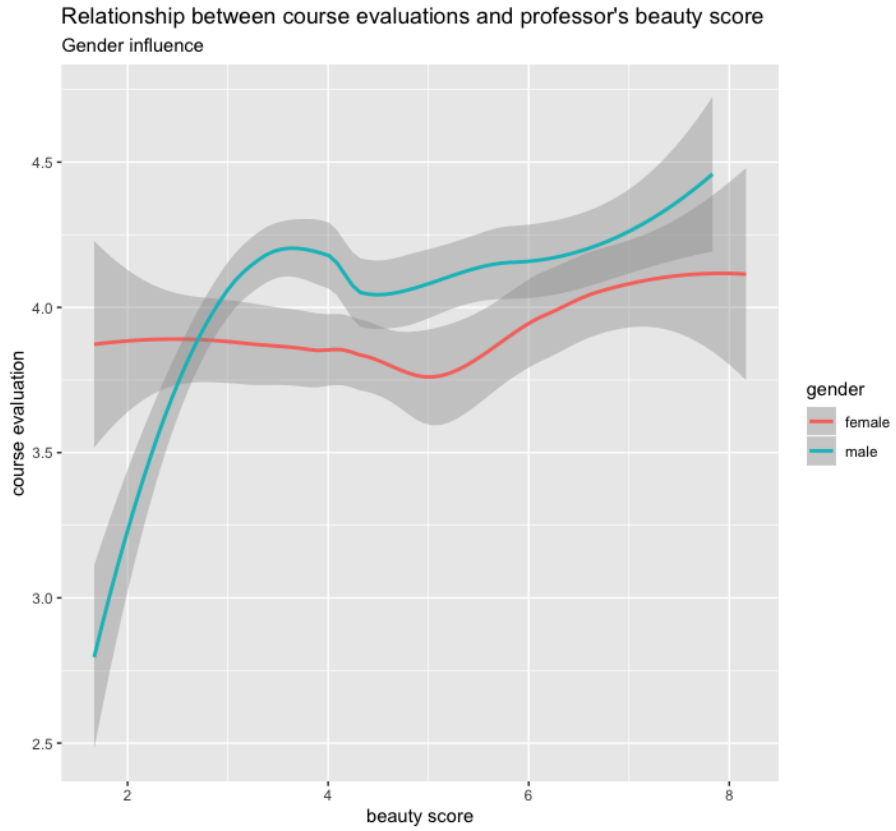


With this plot we want to compare the influences of professor's ethnicity, classified as minority or not minority, and class level, as lower or upper, on course evaluations.

We can observe that there is a huge difference between evaluations toward courses holded by not minority professors and minority ones. The former's distribution experiences a quite regular change, but the latter's is clearly very irregular.

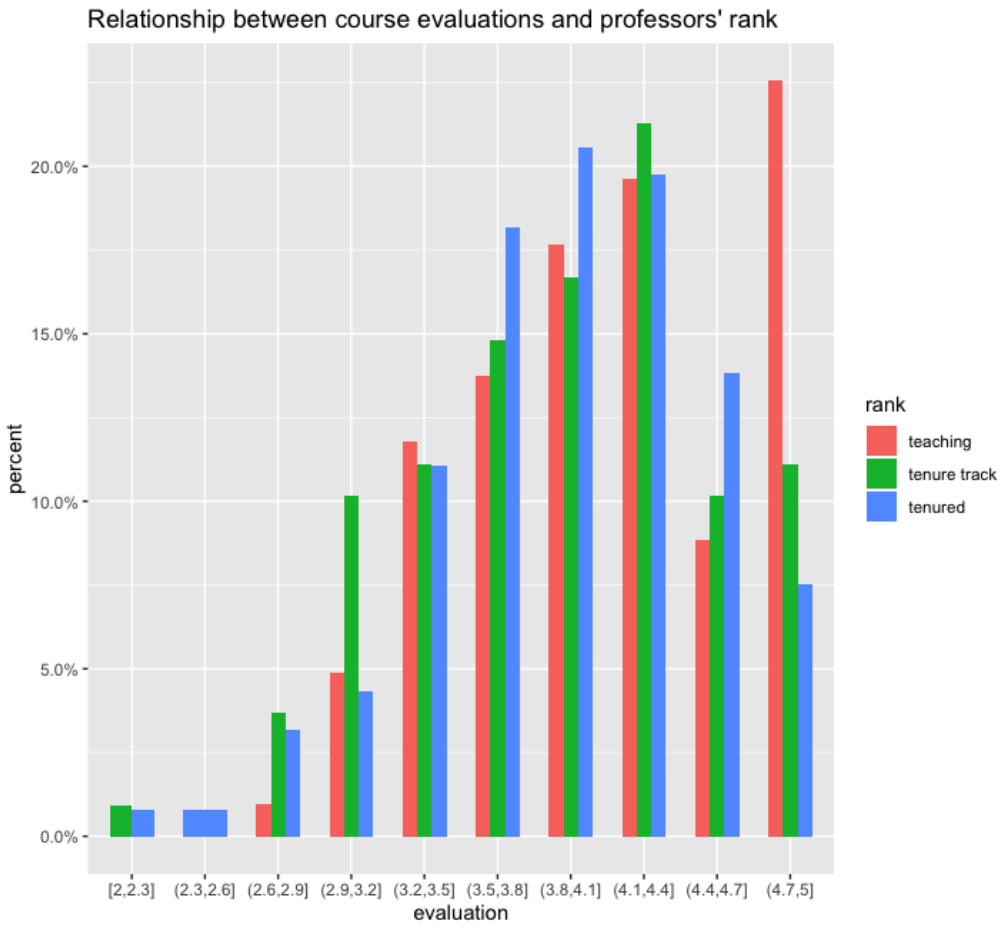
Focus on class level POV, students from upper classes are much more strict evaluating the course. In the contrary, students from lower classes tend to give high evaluation score.

3. Explore the relationship between course evaluations and a beauty score for each professor. Is there evidence of a different relationship depending on gender?



The plot shows solid evidence of different relationship depending on gender. Even the confidence interval for female professors is wider than for male professors, the overall mean value varies poorly for the former. Male professors receive a better course evaluation in general way, but it drops drastically when beauty score falls down of 3 ("ugly"). It should be noted that in both male and female curves present inflection points at beauty score of 4.3 and 5, respectively.

4. Explore the relationship between course evaluations and professors' rank.



All three ranks' (of professors) distribution experience an monotonically increasing tendency on mid-low evaluation scores. At the hightest score interval, the teaching rank professors has huge advantage to others two ranks. At the second highest score interval, teaching rank professors' occupation falls below others two ranks, bringing an interesting conclusion: tenure track and tenured rank professors are more "stable" than teaching rank ones, but latters are those who can hold the most interesting and entertaining courses.

1.5 Data Analysis plan

blablabla

1.6 Conclusions and limitations of findings

References

[1] Moodle. Course evaluation. https://moodle.upm.es/titulaciones/oficiales/pluginfile.php/1947512/mod_resource/content/4/_site/evals.html.