

Out-Role Individual Assignment

Ian Brandenburg (2304791)

Introduction

This course gave me the opportunity to explore new ways to use the causal analysis and interpretation of data in a more practical sense. Before enrolling in this course, I sought to better understand the “so what?” of data analysis, and how regression coefficients and interpretations can be applied to different disciplines and use cases. This court case that was conducted developed a better understanding for how to utilize causal interpretations to take action against issues in discrimination. Furthermore, discrimination can be difficult to quantify, especially in legal cases, but this course gave a very clear approach to arguing for discrimination cases by making use of data analysis procedures.

Point 1: Critically evaluate statistical analysis based on subject matter expertise and general human judgement

The first point that I took away from this course was how to critically evaluate statistical analysis based on the subject matter, expertise, and general human judgement. The project we conducted taught active collaboration in solving problems with data, but also formulating arguments using regression coefficients and p-values. The subject matter of the wage gap was discussed quite frequently, but considering potential confounder variables is very important in determining what the unexplained gender wage gap may be.

The consideration of experts in the court case is important as well. Expert data analysts identify potential biases, such as the selection bias. This is something that can be presented by either the plaintiffs or defendants. It is common because either side will select a dataset or regression models that will represent their case, without acknowledging issues in the data.

Finally, the general human judgement component of data analysis is crucial to consider. Not every problem can be solved or explained by numbers alone, and just because we identified a significant correlation between billable hours and other variables, does not mean that these variables explain the entirety of the wage gap. An unexplained portion of the wage gap was identified, and the court charged the defendants with paying the plaintiff the unexplained amount. Companies may not have deceitful intentions, or perhaps the way the “optimize” their wage setting has systematically been discriminating against employees because of considering factors that would be inherently discriminatory; however, companies may also do this intentionally, so understanding their true motives is difficult. This is why it is important to recognize these issues and make adequate judgements to improve intentional or systematic discrimination in pay so that women can be paid fairly in comparison to men.

Point 2: Use causal graphs to illustrate how three or more factors are related, and how they affect one another directly or indirectly

The court case we constructed also gave a clear use case for causal graphs to illustrate how three or more factors are related. This was seen through how billable hours were influenced by

potential confounder variables, such as the department, children, and employees having faced discrimination. The causal graphs provided us with clear visual representation, which in turn allowed for our team to advocate for wage discrimination in gender. These causal graphs and tables gave insight into direct relationships, such as the associations between billable hours and different departments, indirect relationships, and confounder variables. The court cases we worked on contained several potential confounder variables but could have included more that were not present. This reminded me of the importance of considering confounder variables when conducting regression models, as they can provide thorough explanations of the variance in the target dependent variable.

The causal graphs and tables effectively illustrate indirect relationships between variables by employing regressions across various dependent variables to analyze the associations between the explanatory variables. This can help understand the associations between the explanatory variables and the target dependent variable.

Conclusion

The course provided a clear overview of methods for utilizing data with the purpose of mitigating discrimination in this workplace. The same concept can be applied to many different fields and can potentially have a positive impact on society. Before taking this course, I was having difficulties establishing a strong “so what?” component. These methods can also be applied from within a company, to ensure the employees are being treated fairly. In large scale companies, such as Google, internal discrimination can easily happen. It is important to monitor wages across the company to maintain fair pay and pursue a better future.