Correlation between Catholic Literacy Rates and University Applications*

Ishaan Bansal

2022-02-06

Abstract

Literacy rates are a good indicator of number of university applications. We usually assume that a higher literacy rate is indicative of increased university applications. This paper will work to prove this relationship using data sourced from Open Data Toronto

1 Introduction

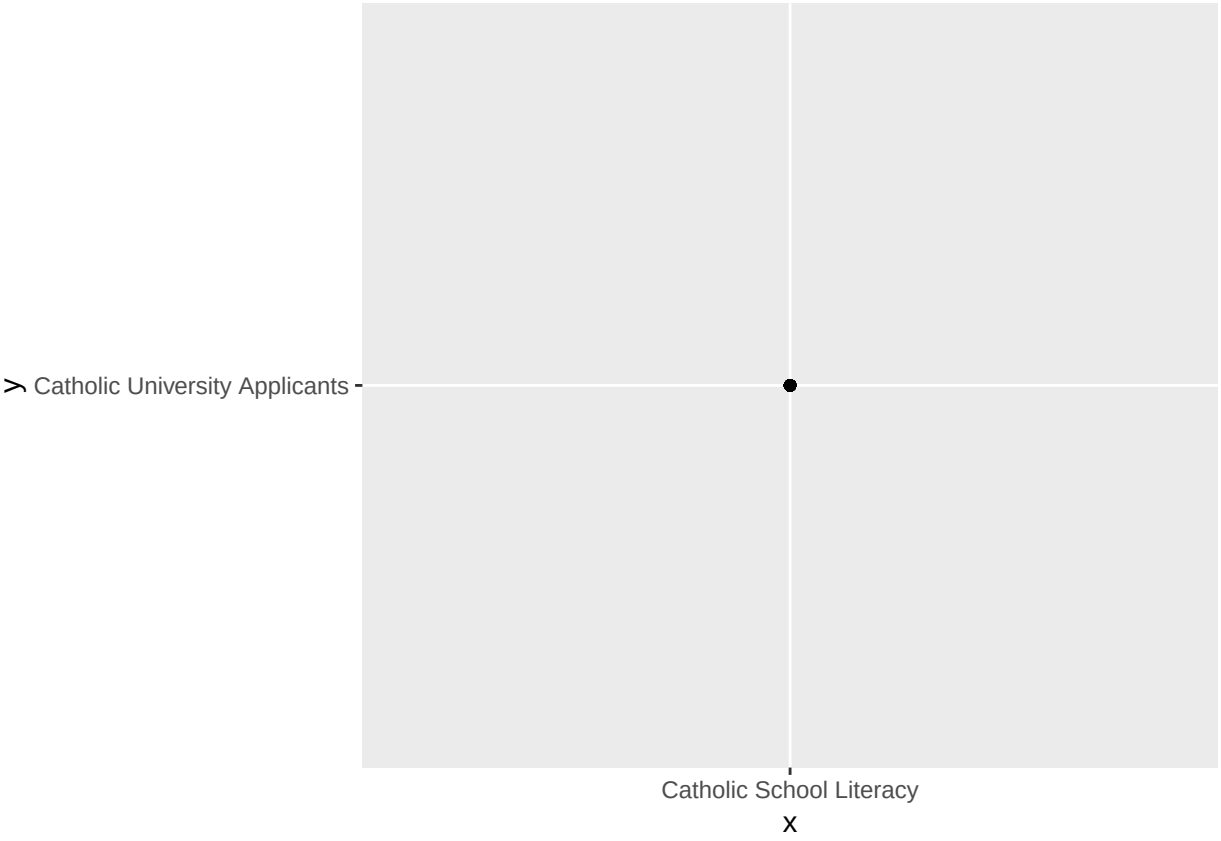
Data is taken from different regions of Canada. Shows raw data for Catholic literacy rates and University applications. Using this data this paper will attempt to prove a positive correlation between Literacy rates and University Applications.

2 Data

The data is sourced from Open Data Toronto (Gelfand 2020). By filtering the data I was able to select only the columns I require for this analysis - Catholic Literacy rates and Catholic University applications using dplyr (Wickham et al. 2021). I will then create a plot using tidyverse (Wickham et al. 2019) and ggplot (Wickham 2016)

```
### Creating Plot ###
date1 = read.csv('Raw_data.csv')
ggplot(date1, aes(x='Catholic School Literacy', y= 'Catholic University Applicants')) +
    geom_point()
```

*Code and data are available at:https://github.com/ishaan-bans/CatholicSchoolLiteracy.git



3 Conclusion

As the plot shows, there is significant presence of a strong positive relationship between Catholic Literacy rates and Catholic University Applications. This proves our initial hypothesis.

4 References

(R Core Team 2020) (Xie 2020)

Gelfand, Sharla. 2020. *Opendatatoronto: Access the City of Toronto Open Data Portal*. https://CRAN.R-project.org/package=opendatatoronto.

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy DAgostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2021. *Dplyr: A Grammar of Data Manipulation*. https://CRAN.R-project.org/package=dplyr.

Xie, Yihui. 2020. *Bookdown: Authoring Books and Technical Documents with r Markdown*. https://github.com/rstudio/bookdown.