

The Dropout Crisis in India: A Survival Analysis

Vanshika Wah

Topic Statement:

Analysis and Prediction of the Student Dropout Rate in India

Abstract

“Education is the basic requirement for human development.”

Retention of students in schools all over the country has been a concern for many decades now, as it is a key indicator of social and economic well-being. Student attrition results in consequences which affect the functioning of the overall institution. The early identification of ‘at-risk’ students is a crucial task that needs to be addressed precisely. In this paper, we use the method of Survival Analysis - a popular technique in the field of Data Analytics - to predict the student dropout rate in the country and to analyse the reasons for the same, along with providing possible solutions to prevent student attrition.

Introduction

In this study, a sample of 1335 students (class1- class12) from Indian schools has been analysed. The data has been collected from the National Family Health Survey (NFHS-3). In the NFHS, specific questions are asked to household respondents about schooling for children aged 5-17 years. It is observed that the dropout rate in India is higher as compared to many other countries (figure-1). Our aim is to predict, when students are likely to drop out of school and to find out the factors which play an important role behind this. This paper will also discuss some probable solutions to improve student retention in schools. To find out when students are most likely to drop out, we proceed with the help of Survival Analysis using the Python Program.

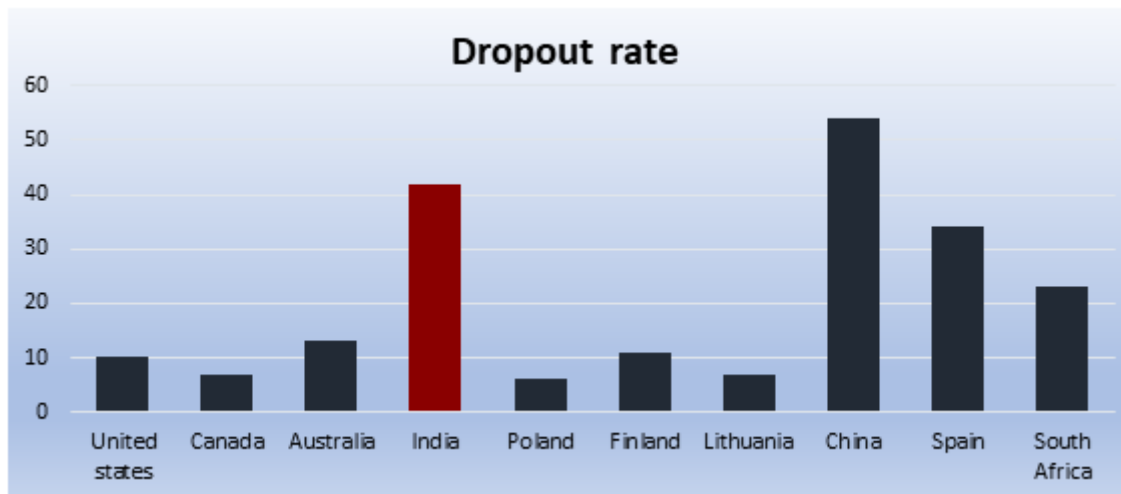


Figure 6: Comparson of Dropout rates between different countires

What Is Survival Analysis?

Survival analysis is a technique that is used to study the time until some event of interest, often referred to as ‘death’, occurs. The survival function is:

$$S(t)=P(T>t)$$

where,

T= subject’s survival time

t= specific number of years for which we want to check the probability for

In our case, we check whether T years (actual years spent in school by a student) is greater than t years (t=1-12 depending on the number of years spent in school which we want to check probability for).

We use the Kaplan-Meier estimator and curve. The Kaplan- Meier curve is a step function that is used to estimate the survival function. The estimator is given by-

$$\hat{S}(t) = \prod_{i: t_i \leq t} \left(1 - \frac{d_i}{n_i}\right),$$

Where,

t_i = time when at least one event happened

d_i = the number of events

n_i = the individuals known to survive (censored cases) Censored cases refer to the events of interest which did not occur in the study time

period. In this paper, the event of interest is the student dropping out and the censored case refers to students who have not yet dropped out. Using the Kaplan Meier Fitter function on our sample data, we obtain the curve in [figure-2](#) using Python Programming.

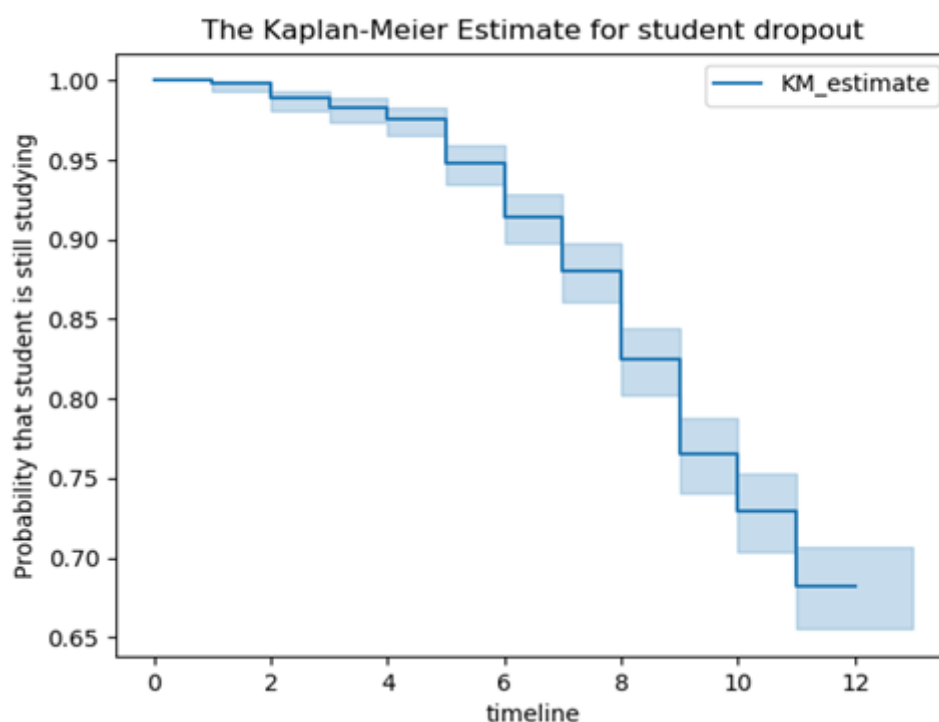


Figure 7: The Kaplan Meier Estimate for student dropout

We observe that the probability of a student still studying reduces as the years in school pass by. The curve begins with the probability of 1, where every student has just enrolled into school; hence, there is no possibility of leaving school at that time, as it is the time of joining (we have taken class 1 as the year of enrolment in our study). The probability of the students still studying keeps reducing till it comes down to 0.68 after 11 years of education and remains constant after that, till the end of year 12. The curve stops at year 12 because after 12 years of education, i.e. after class 12, a student graduates and hence cannot dropout. The curve moves in a step function, as it shows the probability of the students still studying, at every consecutive passing year, with no period in between any two classes as it is assumed that students leave before the next academic year begins.

To check the probability of student retention at any year, we use the survival function, where t can be any number from 0-12, depending on the number of years spent in school that we want to check the probability for.

The major drops in retention (probability that student is still studying) can be seen at the following values in Figure-2 or by putting t = 'number of years' in the survival function-

- year 5 ($t=5$): probability drops from 0.97 in year 4 to 0.94 in year 5
- year 8 ($t=8$): probability drops from 0.88 in year 7 to 0.82 in year 8
- year 9 ($t=9$): probability drops from 0.82 in year 8 to 0.76 in year 9
- year 11 ($t=11$): probability drops from 0.72 in year 10 to 0.68 in year 11

These are the years of education at which the dropout rates are the highest.

Since year 5 is the end of primary school and year 8 is the end of middle school, we can say that the dropout rates are highest at the time of transitioning to a higher level of education, i.e. from primary school to middle school and from middle school to high school.

Factors affecting the dropout rate

The factors that affect the dropout rate are stated in [table-1](#).

Factor	Percentage of dropouts
Lack of interest	29.8
No female teacher	0.2
Transportation problem	1
Repeated failures	6
Required for household work	11.9
High cost of education	18.8
Safety issues	1.4
Others	5.1

Percentage of dropouts based on different factors

For a simpler analysis, we create a column chart (figure-3) from the data given in Table-1.

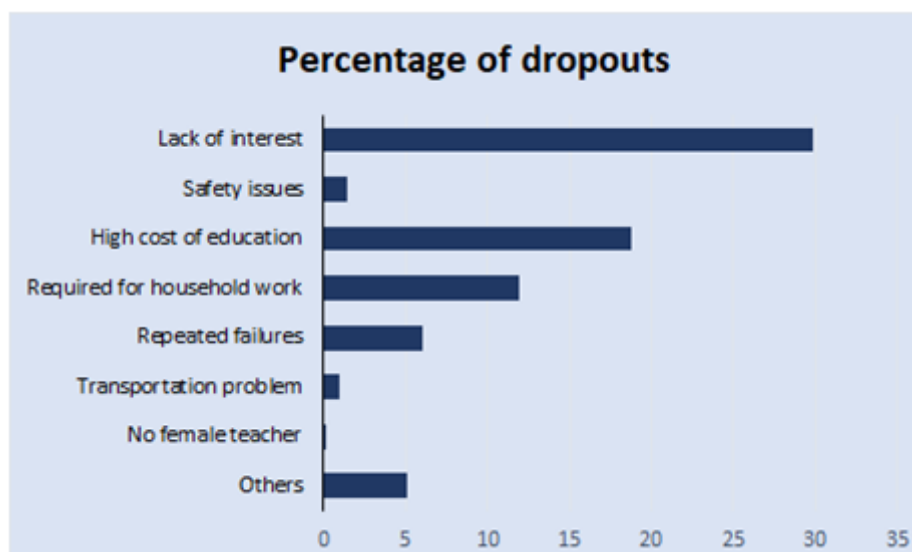


Figure 8: Percentage of dropouts based on different factors

We can see that the factors which play the most important role are ‘Lack of interest’, followed by ‘High cost of education’ and then ‘Child required for household work’. Since these three factors have a significant effect on the dropout rate, our analysis of prediction for future dropout rates will be based on these three factors.

Early prediction of dropout rates

It can therefore be said that the majority of students who tend to dropout are those who lack interest in studies. These students see little value in the content of the course and feel that they can utilize their time in better fields to start earning early. This problem exists in both urban as well as rural areas, whereas the problem of high costs mainly affects students from rural areas who come to urban cities for education. Hence, students from rural areas tend to leave school due to the disability of their families to bear high costs. The other factor, which is requirement of children for household work, mainly affects students from rural areas due to the traditional mindset of rural families. Girls specifically are stopped from attending school in order to be taught household chores or for being married off. We can see the percentage difference between boys

and girls who drop out in figure-4. We notice that there is a 60:40 ratio which indicates the lack of social and economic well-being. We also see in figure-5 that the majority of students who dropout are from backward classes.

DROPOUT RATE- SEX BASED

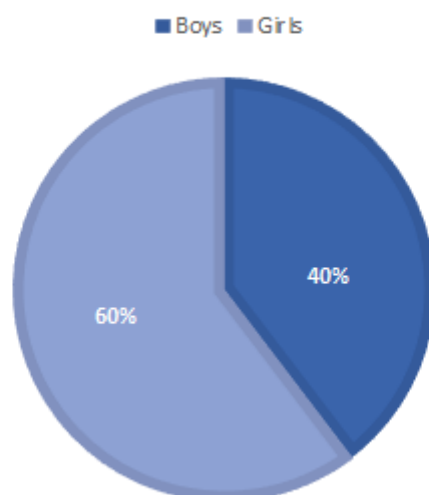


Figure 9: Percentage difference between boys and girls who dropout

DROPOUT RATE- CASTE BASED

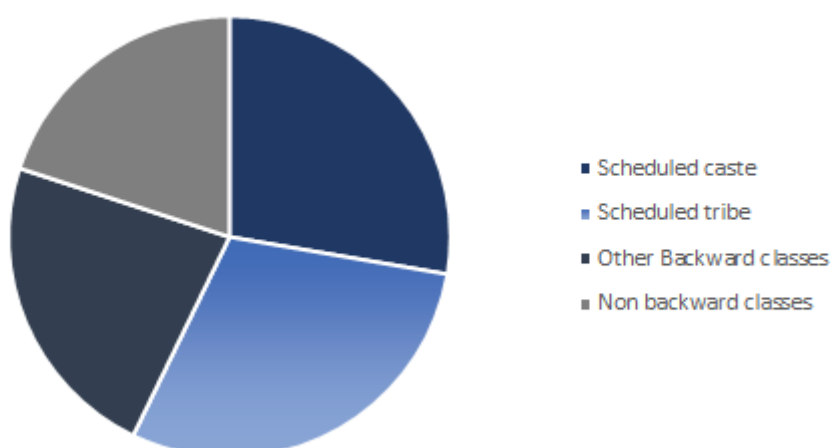


Figure 10: Percentage difference between dropouts based on caste

What can be done to prevent student attrition?

“Knowledge is power.”

The major problem lies in the education system of the country which is mostly about rote learning in order to score higher marks, rather than understanding and gaining real knowledge. If students could clearly see how coursework connects to their goals, interests and concerns, they would be more likely to value it, and thus more motivated to invest time and effort in it. There is a huge scope for improvement in India's education system. The improvisation should begin by making the coursework more practical and application based. Reduction in syllabus will also be a relief for students and will create more room for creativity in their minds.

The problem of rising inflation in the economy and hence rising costs has led students from poor families to drop out of schools. The issue of black money collected by wrong means by private schools needs to be investigated. Also, there needs to be an increase in the number of government schools like Kendriya Vidyalaya in India. The government should be encouraged to begin self-financing schooling under these bodies.

According to NFHS, almost 90 million girls, i.e. approximately 29% of the total student population, are stopped from attending school in India. India should adopt some approaches similar to those adopted in countries like Pakistan, Senegal and Nepal where non-formal education with flexible timetables is provided to girls. Schools with only women as teachers and supervisors should be opened to address the issue of safety for girls.

Conclusion

In this paper, we attempted to understand the differentials and factors associated with school dropouts in India along with an analysis of the time when these students are most likely to leave, based on the data from NFHS-3. The major reasons cited for dropping out were ‘lack of interest’, ‘high costs’ and ‘children required for household work’. The most concentrated time interval of dropouts was the time of transitioning to a higher level of education, i.e. from primary school to middle school and

from middle school to high school. This analysis can help schools all over the country to be prepared for future dropouts as the dropout rate affects the overall functioning of an institution. Unless and until there is considerable improvement in the economic status of households and change in the social attitudes of the citizens of India, achieving the goal of universalisation of school education will remain a major challenge for the country.

References

1. Marie Lall, '*The Challenges for India's Education System*', Chatham House, New Delhi, 2005
2. Fan Yang, '*A competing risks survival analysis of high school dropout and graduation: a two-stage model specification approach*', University of Iowa, 2017
3. Tarannum Siddiqui, '*Influencing Factors of Dropout Students in Elementary Education: a Study of two Districts in India*', IOSR Journal Of Humanities And Social Science, 2017
4. Sattar Ameri, '*Survival Analysis Approach for Early Prediction of Student Dropout*', Wayne State University, 2015
5. Jomon Mathew, '*Why do children dropout of the schools?*'
6. Patricio Valdivieso, '*Survival Analysis: Exploring the dropout motives in a panel of Peruvian children, using the "Young Lives" program dataset*', The University of Edinburgh, 2015
7. Patricia Lone, '*Keeping Girls in School*', UNICEF
8. Allie Knofczynski, '*Why Global Dropout rates aren't Improving*', The Borgen Project, 2017,
9. NFHS-3 Fact Sheets for Key Indicators based on Final Data, <http://rchiips.org/NFHS/pdf/India.pdf>
10. Rasheeda Khatoon, '*A study of student absenteeism in the elementary schools of Eastern UP and its relation with socio-economic and familial factors*'
11. Graduation Rate in Different Countries,

[https://data.oecd.org/eduatt/
graduation-rate.htm](https://data.oecd.org/eduatt/graduation-rate.htm)

12. Sunita Chugh, '*Dropout in Secondary Education: A Study of Children Living in Slums of Delhi*', National University of Educational Planning and Administration, 2011