

# Documentation - Term Project - 5250H Data Analytics with R

Jishant Prashant Acharya

## Introduction

Academia has evolved a lot, there have been a lot of new enhancements in the subjects the level of studies. Research output has been on the rise as people come up with innovations and new theories to understand the world we live in. Psychology is a scientific study that involves the human mind and behavior. It involves the conscious and the unconscious phenomena that includes feelings and thoughts. Whenever a new implementation is done in the realm of academics the psychological impact of the that change has looked at for the best implementation. This helps in understanding if there is pressure coming on the students with the new change. Academia and Psychology go hand in hand and are dependent on one another in many ways.

A student that is undergoing academic stress will have psychological signs that point towards him/her having stress<sup>6</sup>. This a best example that can be used as a testimonial to put academic performance and mental or social factors into context where one effects the other. We will use previous research conclusions to set the background of the dashboard created to show the effects of the social and economic factors on the performance of students.

The dashboard created can help understand the data set in a graphic manner which will make it easier to get the understanding of the effects of different factors on one another using the chi square tests and regression between all the continuous variables available to show the correlation strength and relation between one another.

## Background

We find provenance of this dashboard in many researches that have concluded that various social and economic factors affect a student's performance. There are a lot of variables that we can use to conclude the same using various hypothesis testing methods.

Distance Learning has become routine in the recent times when we all are forced to stay away from each other due to COVID-19. When this idea was just a new thought there was very little data to play with and understand a student's performance<sup>1</sup>. When the students are not in the same room as the professor, getting to know the students and their performance is very difficult. The only resort that the professors have for having a better reach to the class

is to understand if the topic has been understood by most of the students. Assignments are one way of looking into things alongside with the demographic from where they come. The demographic being an important part of the equation along side with the marks that they score in the assignments. The research<sup>1</sup> works on a supervised learning algorithm that uses the data and predicts a set of students that will or can perform poorly. This will help the professor to get in touch with them proactively and make them understand by pushing them over the limit for that extra performance. Here as we see demographic information of the student is an important variable that helps the model predict the student performance. The success of this model making approach is supported by the accuracy of 62% initially which increases to a whopping 82% before the final examinations. This goes to show that factors other than marks affect a student's performance and can be pushed to achieve greater marks by supporting them in things that are not directly related to academia.

The data used has been used in the research<sup>2</sup> to predict the best setup for the school to look into the life of the students to improve their marks. The research is based in Portugal where the school leaving percentage is nearly 2.5x the European Average of school leaving in the age group of 18-24. The model in the research was used to predict better student conditions and a great amount of accuracy was reached when the record had one of the both other period marks other than the target period marks. This model takes into consideration that all the factors that are not directly related to a student's score also influence the performance to a considerable extent. The school management can look at the research that quantifies a variable's effect on the student performance and can improve in those areas to get to a better scoring student population that doesn't flunk out or leave school preemptively.

Age is an imminent fact of life, we are all growing towards being old and mature. The research upon careful inspection concluded that with increase in the age the ability to score and comprehend decreases in vigor. The research studies the *Relative Age Effect (RAE)*. This research has the socioeconomic status and the type of school in context to check the effects of all these combined on a population of eighth graders. The article<sup>3</sup> works on structural equation modelling (SEM) to get the latent variables and work on the measured or indicator variables to get the relationships between each other. The method involves using a pool of students that are enrolled in the same year of school to keep the *RAE* at the minimum. The pool of students is further divided into 5 groups of G1, G2, G3, G4, and G5 where G1 is a group of people who are oldest and G5 are the youngest. Study conducted on 15,234 students supported that *RAE* is a real phenomenon which was seen when the eldest group didn't stand out in performance and the G2 group scored the most marks which then dwindled down until the last group of G5. This shows the effect of *RAE* that proves that when a student is enrolled in the schools earlier are at a greater advantage to score more but also shows how being a little mature for the class can help with a better progressed neurological development.

Student's attendance in the class can play a very important role when it comes to performance directly related to the score in the application under the regression tab. As we can clearly see that when the absences are put against any period marks the correlation comes out to be negative. We can conclude that as the absences increase the performance starts to dip. However the sample space is too small to conclude that generally but the correlation

stands negative in the small space too. Researchers tried to get the non-academic reasons for their poor academic results. They found a significant relationship between the attendance of the student and the results. Furthermore, apart from the absences they could also find Other significant findings which included the relationship between attendance and the level of study; work shift; ethnicity; whether participants had any dependents; and how studies were funded<sup>4</sup>. There is a constant willingness and want to push the country's education and research output for the better. Which cannot be done without keeping in mind the latent factors of the social standing and the economic status of the student. To improve the output an all around development with respect to each facet of life of a student will have to ameliorate so as to give the best result possible.

Gender can also play an important role in understanding the results of the student. When put into context the socio-economic factors, this could support all the previous researches in this realm. The research also deals with a bias from the selector's point of view. Nonetheless, gender's effect was seen during the study where most of the girls were found to be overachievers and very less of them were found to be in the underachiever group. The study also found the association of parents with the school to be a factor that added to be a significant factor in making an impact on a student's performance<sup>6</sup>.

## **Instructions to Run**

## **Section Description**

## **References**

1. S. Kotsiantis , C. Pierrakeas & P. Pintelas ; Predicting Students' Performance In Distance Learning Using Machine Learning Techniques
2. Paulo Cortez and Alice Silva ; Using Data Mining to Predict Secondary School Student Performance
3. Juan-José Navarro1, Javier García-Rubio, Pedro R. Olivares ; The Relative Age Effect and Its Influence on Academic Performance
4. Kawtar Tani, Elizabeth Dalzell, Nathan Ehambaranathan, Sheela Murugasu and Anne Steele ; Eval- uation of factors affecting students' performance in tertiary education
5. Thomas R. Ford ; Social Factors Affecting Academic Performance: Further Evidence
6. Yaoran Li , Jeff Allen , Alex Casillas ; Relating psychological and social factors to academic perfor- mance: A longitudinal investigation of high-poverty middle school students