HUMAN AND SOCIAL DATA SCIENCE MASTER DISSERTATION

What potential associations exist between education related factors and the mental health of adolescents between the ages of 11 – 17?

Supervisors
Dr Adam Barrett
Dr Darya Gaysina

Author
Megan Jennefer Robertson

DEPARTMENT OF MATHMATICS AND PHYSICAL SCIENCE
UNIVERSITY OF SUSSEX

WORD COUNT: 11,913

Content Page

1.0. Introduction

In 2021, the NHS conducted a study looking into the mental health of adolescents between the ages of 5 and 22 (NHS Digital, 2021). The study found that 1 in 6 children aged 6 to 16 had a probable mental disorder, a significant increase when compared to the 1 in 9 children from the 2017 study. This study is not alone in its findings, with numerous other research papers spanning the last 2 decades finding similar growing trends in adolescent mental ill health across the world (Wiens et al, 2020; Patalay & Gage, 2019; Griffin et al, 2018; Mojtabai et al, 2016).

This increase in the prevalence of mental health issues has led to a large amount of research being conducted, looking into diagnosis, treatment, and, most importantly for this dissertation, potential risk factors and impacts of mental ill health. In general, risk factors like child abuse, social media, socioeconomic difficulty and social isolation, to name a few, are frequently referenced as having a negative impact on adolescent mental health (WHO, 2022). Additionally, mental health is often linked to negative impacts such as increased risk of substance use, increased aggressive behaviour, poor peer relationships and poor school performance (Youth, 2022). While the body of research into the impacts of adolescent mental health is significant, there is less research looking at how education related factors might impact pupils' mental health. This is unfortunate as children between the ages of 5 and 18 spend a large percentage of their time a week in an academic setting, and it is likely to have an impact on the development of their mental health.

Viewing both sides of the current research into these 2 fields, there are several key concepts which emerge for both. When reviewing how adolescent mental health impacts a child's education, the general consensus is that there will be a negative impact of grade attainment, school attendance and drop-out rates (Smith et al, 2021; Lereya & Deighton, 2019; Schulte-Korne, 2016; Riglin et al, 2012). Alternatively, when examining how education impacts adolescent mental health, there are associations between poor academic performance, academic stress, and academic expectations (Crenna-Jennings, 2021; Cohen et al, 2020; Mboya et al, 2020; Pascoe et al, 2019). These concepts will be further explored in the literature review for this project. Although, they provide a valuable starting point, many of the studies tend to use small sample sizes, focusing on a limited age range and typically researching one or two specific aspects of the field. This project will be taking on a more top-down macro-approach, using the studies within the literature review as a guide when attempting to discover new associations between education related factors and mental health. However, all educational features will be tested for association with mental health, with no particular pre-assumptions in mind which would potentially result in some associations not being found. This will be important, as although this project does not have the time or resources to prove causal relationships between any of the associations discovered, it will provide more grounds for further study into less well researched relationships between education and mental health

The overall goal of this dissertation is to analyse the data of over 8000 pupils from ages 11 to 17, and report any associations discovered between education related factors and poor mental health. There are 2 specific research questions which have been asked to achieve this goal. The primary question this dissertation seeks to answer is "What associations can be found between education related factors and the adolescent mental health of cohort members aged 11 to 17?". The second question is "What is the gender difference between any associations discovered?". This is because a number of studies have found distinct gender differences but there is no clear consensus on how gender fully impacts different associations.

In order to answer the 2 research questions outlined above, the first step is to create a list of education related factors and a mental health score, before performing a series of statistical test to determine significance and strength of association. Finally, the final features which demonstrate a significant association will be used to train a classifier model, which will determine the features' importance and impact on its final predictions. Through these tests, this project will be able to determine significance, strength and direction of any associations found.

The following chapters of this project are laid out as such. The second chapter will address in more detail the current literature surrounding these 2 fields of research, while the third chapter outlines the specific methodologies used in this project and results found. Finally, the fourth chapter will conclude the project by discussing the most prominent associations discovered and answering the 2 research questions proposed.

2.0. Literature Review

In 2021, a study by NHS Digital (2021) found that 17.6% of secondary school aged children had a probable mental disorder, an increase from 12.6% in 2017. Additionally, the World Health Organisation (2022) estimates that globally 14% of 10–19-year-olds have experienced a mental disorder, with anxiety (8.2%), conduct disorder (6%), ADHD (5.5%), and depression (3.9%) being the leading causes. This increase in prevalence of adolescent mental ill health is concerning as the ramifications will continue to affect them into their adulthood. A study by Kessler et al (2005) found that around 50% of all lifetime mental illnesses presented by the age of 14 and 75% presented by the participants mid-20s. Additionally, around 75% of adults suffering from mental ill health started experiencing difficulties before they were 18 years old (NHS England, 2015, Kim-Cohen et al, 2003). These studies demonstrate how important a timeframe adolescence is for mental health.

Alongside this increase in global prevalence of mental ill health, there has also been a steady rise in mental health awareness over the last 70 years (Venters, 2019). In the 1950s mental health was viewed as a taboo topic which was often either ignored or patients were removed from the public eye and sent to psychiatric asylums. Now, in 2022 there has been a significant shift in society's attitude towards mental health, with people being more comfortable to talk about mental health and provide support to struggling individuals (Relative Insight, 2021). Additionally, governments are starting to place more spending into the research, diagnosis, and treatment of different mental health conditions, with the UK's mental health budget being just under £15 billion (NHS England, 2022).

Alongside the growing mental health prevalence and awareness, is research which ranges from diagnosis to new treatment. Specifically for this project, the relevant studies are those which have examined the impact adolescent mental health and education have on each other.

2.1. The Impact of Adolescent Mental Health on Education

There has been substantial research examining how education is impacted by mental illness in adolescents. The large body of this work can be broken down into 2 categories of interest.

Starting with the first category of academic attainment, due to the large body of research into this field, the general consensus is that mental health difficulties in adolescences is associated with lower academic attainment (Agnafors et al, 2021; Lereya et al, 2019; Seller et al, 2019; Cornaglia et al, 2012; Mcleod & Kaiser, 2004). However, this association does not equally apply to both sexes. There are several key pieces of literature which examine this difference in association in further detail. The first is a study by Smith et al (2021) which analysed longitudinal associations between mental difficulties at age 11-14 and academic attainment at age 16 for 1100 pupils. The study found that, when accounting for potential compounding variables, pupils with mental health difficulties were twice as likely to not achieve 5 GCSE grades A* to C, with the effect being greater for male attainment. Another study by Lopez-Lopez et al (2021) found similar gendered results by examining the association between self-reported depressive symptoms and academic attainment for 3800 pupils aged 11 – 18. The study discovered an overall negative association, which was stronger in male participants. The last study which also show similar trends to the previous two was conducted by Riglin et al (2013). Riglin et al (2013) found that certain mental difficulties resulted in a decline in academic attainment, with depressive symptoms and conduct~~s~~ problems being the primary culprits. Additionally, this study found that the associations discovered were stronger for male participants (Riglin et al, 2013). All these studies demonstrate the negative effect mental ill health has on academic achievement, with the impact being greater for male participants. Finally, it should be noted that not all studies agree with the association between academic performance and mental health difficulties. Some studies have shown no effect between internalizing problems, like depression and anxiety, and academic performance (Van Der Ende et al, 2016; Miech et al, 1999).

The second category regards attendance and dropout rates, which is less well research then the previous category. However, there are still several studies which have looked into this relationship and found a negative association (Lereya & Deighton, 2019; Schulte-Korne, 2016). Specifically, a study by Cornaglia et al (2011) found an association between poor mental health measured at ages 13 & 14 and the probability of drop-out, with the association being stronger for girls but boys still showing a similar pattern when using later measures of mental health. Another study by Hjorth et al (2016) found that poor mental health was associated with dropout rates among students in both vocational and higher education, except that men were 5x at risk of dropout, while the association was not present for woman. These studies offer slightly conflicting information, as both agree on an association between mental health and dropout rates but disagree on the gendered aspect of the association.

Overall, the studies discussed for both categories are extensive and provide some interesting, if somewhat conflicting, insight, regarding the impact mental health has on adolescent education.

2.2. The Impact of Education of Adolescent Mental Health

Unlike the previous field, the research into how education related factors impact adolescent mental health has been limited. There are however a few studies have shown that school does have a small, but significant, influence on adolescent mental health (Hale et al, 2014; Roeger et al, 2001). Specifically, a large study by Ford et al (2021), which used multi-level linear regression models to analysed data from over 26,500 first/second year secondary school pupils. The study found a small but significant impact of school level factors on pupil's mental health. It is however important to note that this study only looked at participants aged between 11 and 14 years old and although this age range needs to be accounted for, this study fails to account for the later stages of adolescence where the risk and prevalence of mental illness increases (Jones, 2013).

Looking beyond studies which examine the general influence education had on adolescent mental health, there are several studies which research education related risk factors. These studies provide 6 potential influences which can negatively impact adolescent mental health.

The first of these risk factors identified is high academic stress, with secondary school students frequently self-reporting ongoing stress related to their education (UNESCO, 2012). A study by the Organisation for Economic Co-operation and Development (2017) which surveyed 540,000 students aged 15-16, found that 66% reported feeling stressed about grades while 59% felt anxiety around examination. Additionally, girls consistently reported having greater anxiety surrounding schoolwork compared to boys. This study demonstrates how prevalent academic stress is at secondary school age. Academic stress can often be broken down to include concerns surrounding grade performance, examination, long study hours, rigid busy schedules, high parental expectations, and vastness of curriculum (Sani et al, 2012; Sreeramareddy et al, 2007). One particular study by Pascoe et al (2019), which performed a literature review of 13 studies, found that students experiencing high levels of ongoing academic stress were more likely to suffer a serious mental illness. This is not surprising as the relationship between stress and mental disorders like depression and anxiety is well documented (Moylan et al, 2013; Dantzer, 2012; Dantzer et al, 2011).

The second risk factor is poor academic performance, with multiple studies suggesting that poor academic performance can lead to an increased risk of suffering from depressive symptoms or the anxiety condition (Sorberg Wallin et al, 2019; Miech et al, 1999). There are also several key studies which explore this association in more detail, starting with a study by Crenna-Jennings (2021) which found an association between poor wellbeing, self-esteem, and mental distress and secondary school aged pupils being placed in the bottom stream at school. Additionally, another study by Mboya et al (2020) which looked at students in higher education, found an association between heightened mental distress and poor grade performance. Finally, a study mentioned previously by Lopez-Lopez et al (2021) found that girls with lower academic attainment would be more vulnerable to depressive symptoms in the future. The limitations with these studies is that the associations discovered where by chance and not investigated in further detail. The association between mental health and poor academic performance is significant, as seen in the previous section yet despite this, it is rare that poor academic performance is researched as a potential risk factor for adolescent mental ill health. Finally, it should be noted that a study by Agnafors et al (2021) found no association between academic performance at age 15 and 19 and an increased risk for mental health problems at age 20. This again indicates that the association between mental health and academic performance is a contentious field, regardless of the direction of the relationship.

Similar to the previous factor, the third risk factor is academic expectation. A study by Cohen et al (2020), which examined how levels of educational attainment had affected participants long term mental health, found that individuals with lower expectations regarding the level of education they would attain were at a higher risk of depression later in life. However, low personal expectations are not the only potential risk factor. Another concern is overbearingly high parental expectations which can also have a negative impact on adolescent mental health. Several studies examining Asian-American's mental health specifically, found a negative impact on mental health due to heightened stress from attempting to meet high parental expectation and feelings of shame and embarrassment should they fail (Chung, 2017; Chung, 2016; Lee & Zhou, 2015). Although another study by Warikoo et al (2020) suggests that this negative impact on mental health is instead caused by poor relationships with parents, rather than high academic expectations. This indicates that parental relationships and involvement might be another potential risk factor that should be examined in this project.

The fourth risk factor looks at forced attendance for low achievers. Avendano et al (2017) reviewed the 1972 school reforms act which increased the minimum school leaving age, finding that there was an increased prevalence of mental health conditions in adulthood after the reform. Avendano et al (2017) theorised this was caused by compelling low-achievers, who may benefit little from additional schooling, to remain in education and incur further psychological and emotional costs for little benefit (Echstein & Wolpin, 1999).

The fifth risk factor is poor attitudes towards school, which includes low enjoyment, interest and engagement in the educational material being provided. Deighton et al (2021) which analysed the Millennium Cohort Study found that 59% of children who experience some period of mental health difficulties had also suffered from low enjoyment of school. This indicates that other attitude-based factors, such as interest, engagement and boredom might also impact adolescent mental health, which this project will be examining in greater detail.

The final risk factor identified is school connectedness, which represents the extent to which students feel accepted, respected, included and supported by others in the school environment (Cavioni, 2021; Patalay et al, 2020; Goodenow, 1993). Several studies found that lower feelings of school connectedness were associated with emotional distress, worry, loneliness, and suicidal ideations (Ford et al, 2021; Joyce & Early, 2014; Pretty et al, 1994). Although one study by Patalay & Fitzsimons (2018) found that low school connectedness was only linked to worse self-reported mental health in females.

Overall, the aforementioned associations will be used in this project as a starting point when analysing education and mental health in the Millennium Cohort Study (MCS). This project will hope to build upon the previous research, discovering and testing further associations between education and mental health.

3.0. Research & Results

In order to answer the research questions proposed in this dissertation, the Millennium Cohort Study will be used as the primary data source for analysis. The mental health condition of adolescent participants will be the dependent variable while a variety of education related features will be the independent variables. The methodology for this project can be broken down into 4 major stages. The first stage is preparing the primary dataset into workable data and breaking down the dependent and independent variables which will be further examined during this analysis. The second stage comprises of several frequency distribution tests to allow for a better understanding of any patterns within the data. The third step is performing a series of statistical inference tests, to check for associations between educational and mental health factors. The final step will be developing a classifier model, designed to predict the mental health conditions of cohort members. Once the model is built, features will then be reviewed for their importance in the model's decision, as well as the specific impact on the final prediction.

3.1. The Millennium Cohort Study

In order to answer the research question proposed in this dissertation, the Millennium Cohort Study (MCS) was used as the primary data source for analysis. The MCS is one of numerous longitudinal studies performed by the University College London, designed to "provide multiple measures of cohort members' physical, socio-emotional, cognitive and behavioural development over time, as well as detailed information on their daily life, behaviour and experiences" (UCL, 2022). Specifically, the MCS follows the lives of around 19,000 cohort members born between 2000-2002 across the UK. Although this study has measured numerous aspects of its cohort members lives, this dissertation will be focusing on the educational and mental health data collected when participants ranged from 11 to 17 years old. Due to this specific focus and a need to ensure consistency of participants across all 3 sweeps at ages 11, 14 and 17, the relevant sample size has reduced to just over 8000 participants.

3.2. Data Set Up

The MCS provided several data files for each sweep, ranging from cohort member to parental interviews. The files specifically chosen were all cohort member interviews and the parental interviews whose questions related to the cohort member at each sweep, the teacher survey from sweep 5, age 11, and the cohort member qualification survey from sweep 7, age 17. This resulted in 8 separate files which needed to be reduced, cleaned, and combined into one dataset which could be analysed. First, the files were manually reduced, with only educational and mental health variables being kept. Then, in order to merge the 8 files into one cohesive dataset, a Cohort ID was created to ensure continuity across all 3 sweeps. The Cohort ID was created by using the anonymised identifiers used in the MCS, specifically the MCSID which is the household identifier, and the CNUM which identifies the specific cohort member in that household.

Once the different datafiles were then merged together using the Cohort ID, the final dataset was checked for missing values. The data used several negative values to represent missing entries, which did not add any specific additional information. These values were replaced with a simple NaN entry, to ensure they were not counted during the statistical test phase. However, these missing values would be addressed in more depth during the modelling phase. Additionally, any potential duplicate rows were removed from the dataset to ensure they would not throw off the analysis.

Finally, the data was divided into two datasets based on the cohort members sex, of which there were around 300 more female participants to their male counterparts. Additionally, the proportion of participants classified with an 'Abnormal' mental health condition was different when broken down by sex. 23.7% of woman likely suffered a mental illness compared to only 12.7% of men. This indicates that women will be overrepresented in the mental health categories when compared to men.

Once the data had been completely cleaned and prepared for analysis, the dependent and independent variables were selected.

3.3. Dependent & Independent Variables

The dependent variable for this study was the cohort members mental health, which was drawn from the age 17 sweep, as it was the first-time participants had answered questions in regards to their own mental health. The dependent variable was created by scoring different mental health data points, in order to create a binary banding of either 'Normal' or 'Abnormal' mental health conditions. The 4 features which created this final banding were 2 mental health questionnaires, a mental health diagnosis and finally whether the participant had made any suicide attempts. Specifically, the 2 mental health questionnaires used were the Strengths and Difficulties Questionnaire (SDQ) and the Kessler 6 scale (K6). The SDQ is a commonly used behavioural screening test which consists of 25 questions assessing both positive and negative attributes, including the areas of conduct problems, hyperactivity, emotional symptoms, peer problems and prosocial behaviour (Goodman et al, 1998). The final scoring for this test results in 3 bandings of Normal, Borderline and Abnormal mental conditions. Meanwhile the other questionnaire, the K6, measures nonspecific psychological distress over a 30-day period, touching on areas of whether participants felt sad, nervous, restless, hopeless, worthless, and the degree of effort it felt like to perform tasks (Kessler et al, 2003). The final scoring of this test breaks down into 2 bandings of Normal and Abnormal mental conditions. Alongside these bandings given by these 2 tests, participants who had a mental health diagnosis and participants who had attempted suicide were also given an 'Abnormal' banding. These 3 bands of 'Normal', 'Borderline' and 'Abnormal' were given the scores of 0, 0.5 and 1. Once the scores were totalled together, any participants with a score of 1 or greater were classified as having an 'Abnormal' mental health condition, while the rest were classified with a 'Normal' mental health condition. These 2 bandings of 'Abnormal' and 'Normal' will be used as the dependent variable for this project.

Meanwhile, the independent variables came from a wide range of data collected regarding the participants education across all 3 sweeps. The data was drawn from answers provided by cohort members, parents, teachers and even qualification records. These different perspectives provided a huge range of data and resulted in 92 independent variables, which were broken down into 8 categories for easier analysis. These categories are as follows:

1) Academic Performance: *This category includes self-reported subject performance, as well as subject performance reports from teachers and final grades from GCSE and BTEC qualifications.*

2) Attitudes towards school*: This category includes both cohort members perceived attitudes and self-reported attitudes towards school and subjects, ranging from enjoyment and interest to concentration and engagement.*

3) Misbehaviour & Punishment: *This category includes the absenteeism of cohort members and their misbehaviour at school, as well as the more extreme punishments like suspension and expulsion.*

4) Homework: *This category includes parental and school expectations surrounding homework, as well as how often cohort members received additional help from family with homework.*

5) Academic Expectations: *This category includes parental and teachers' academic expectations of the cohort member, from likelihood of attending higher education to what kind of role they will be performing at the age of 16.*

6) Parental Involvement: *This category includes how involved parents both were perceived and self-reported to be with the cohort members education.*

7) School Environment: *This category looked at environmental factors such as class size, gender make up and faith-based schools. Whether cohort members had received additional support from the school is also included in this category.*

8) Post 16: *This category included what kinds of activities participants were involved in post-secondary education.*

There was little alteration needed for the independent variables with one exception, GCSE and BTEC grades. In order to gain a single entry which summarised the cohort members overall grade performance, the grades were provided a numerical score and then totalled together, with a higher overall score indicating a better overall grade performance.

After finalising both the dependent and independent variables, the first statistical tests could begin.

3.4. Statistical Tests

The first statistical test conducted were a series of frequency distribution tests which were performed on the different independent variables. However, due to a large percentage of variables all displaying a tendency towards one or two 'common' answers, the frequencies were changed to percentages instead of tallies. This allowed for easier observation of how the 'Normal' and 'Abnormal' mental health conditions were represented in the less common answers. Specifically, the total sample proportion for the 'Normal' and 'Abnormal' bandings were 81.6% and 18.4% respectively. If any of the frequency percentages were significantly higher or lower than this sample proportion, it would indicate a possible association.

Once a better understanding of the patterns and potential associations within the data had been achieved, the second statistical test performed was a Chi-Square Test of Independence which is designed to check for association between 2 variables (McHugh, 2013). Additionally, this particular test is non-parametric in nature which made it an appropriate choice as all the data being analysed was either nominal or ordinal. A P-value of 0.05 would be used to assess the significance of the association found between the independent and dependent variable.

However, alone the Chi-Square Test of Independence is unable to determine the strength of any association found. For this reason, the final statistical test performed was the Cramer's V Test which would provide an effect size, which is a measure of the power of association between 2 categorical variables (Kearney, 2017). Depending on the degrees of freedom, the threshold for the effect size will change. This can be seen in Figure 1, however all the tests had a degrees of freedom size of 1 which means only the first row is relevant when examining the effect size threshold.

Degrees of Freedom	Small Effect	Medium Effect	Large Effect
1	0.10	0.30	0.50
2	0.07	0.21	0.35
3	0.06	0.17	0.29
4	0.05	0.15	0.25
5	0.04	0.13	0.22

Figure 1: Cramer's V Test Degrees of Freedom effect size threshold table.

Combined, both the Chi-Square Test and the Cramer's V Test can determine whether 2 variables are associated and to what degree of strength their association is. The following section will outline the results from both tests and the final features selected to be run during the modelling phase of this project.

3.5. Statistical Test Results

The Chi-Square Test of Independence and the Cramer's V Test were performed on 3 different datasets, with the first dataset accounting for all cohort members, while the other 2 datasets divided the cohort members by sex, resulting in a Female and Male group.

3.5.1. Chi-Square Test of Independence Results

Starting with the Chi-Square Test, the goal was to ascertain whether certain features were associated with the dependent variable and reduce the original 92 independent variables. Reviewing the 3 different datasets, the combined dataset removed 30 independent variables which lacked any significant association, while the female dataset removed 23 variables and the male dataset removed 44 variables. The following Figure 2 includes all the features removed from the respective datasets, as well as including which survey provided the feature in brackets.

	All Participants	Female Participants	Male Participants
1	English performance (Teacher)	English performance (Teacher)	English performance (Teacher)
2	English streamed (Teacher)	Science streamed (Teacher)	Maths performance (Teacher)
3	Science streamed (Teacher)	Good at science (Sweep 5)	Science performance (Teacher)
4	BTEC subject grades	BTEC subject grades	English streamed (Teacher)
5	Bored at school (Teacher)	Expelled from school (Sweep 5 parents)	Maths streamed (Teacher)
6	Tries at school (Teacher)	Expelled from School (Sweep 6 parents)	Science streamed (Teacher)
7	Hands homework in late (Teacher)	Missed school (Teacher)	Good at science (Sweep 6)
8	Misbehaves at school (Teacher)	Missed school (Sweep 5)	GCSE subject grades
9	Suspended from school (Teacher)	Frequency of missed school (Sweep 6)	BTEC qualifications
10	Expelled from school (Sweep 5 parents)	Staying in school post age 16 (Sweep 5)	BTEC subject grades
11	Missed school (Sweep 5)	Support group (Teacher)	Bored at school (Teacher)
12	Frequency of missed school (Sweep 6)	Receiving support (Sweep 6 parent)	Tries at school (Teacher)
13	Support from individual teacher (Teacher)	Support from teaching assistant (Teacher)	Frequency of missed school (Sweep 6 parents)
14	Support from teaching assistant (Teacher)	Expectations post age 16 (Sweep 5 parents)	Support from teaching assistant (Teacher)
15	Support group (Teacher)	Class size (Teacher)	Misbehaves at school (Teacher)
16	Receiving support (Sweep 6 parent)	Faith school (Sweep 5 parents)	Suspended from school (Teacher)
17	Faith secondary school (Sweep 5 parents)	Faith secondary school (Sweep 5 parents)	Expelled from School (Sweep 5 parents)
18	Faith school (Sweep 5 parents)	Faith school (Sweep 6 parents)	Missed school (Teacher)
19	Expectations post age 16 (Sweep 6 parents)	Same & mixed sex school (Sweep 5 parents)	Missed school (Sweep 5 parents)
20	Expectations post age 16 (Sweep 5 parents)	Attending parents evening (Sweep 5 parents)	Hands in Homework late (Teacher)
21	Staying in school post age 16 (Sweep 5)	Time spent on Homework (Teacher)	Missed school (Teacher)
22	Faith school (Sweep 6 parents)	Time spent on homework (Sweep 5 parent)	Support from teaching assistant (Teacher)
23	Same & mixed sex school (Sweep 5 parents)	Check homework completed (Sweep 5 parents)	Support group (Teacher)
24	Class size (Teacher)		Receiving support (Sweep 6 Parent)
25	Time spent on homework (Teacher)		Likes science (Sweep 5)
26	Time spent on homework (Sweep 5 parent)		Staying in school post age 16 (Sweep 5)
27	Check homework completed (Sweep 5 parents)		Expectations post age 16 (Sweep 6 parents)
28	Apprenticeships (sweep 7)		Likelihood university (sweep 6)
29	Traineeship (sweep 7)		Interested mother (teacher)
30	University (sweep 7)		Interested father (teacher)
31			Attending parents evening (Sweep 5 Parents)
32			Faith school (Sweep 5 Parents)
33			Faith secondary school (Sweep 5 Parents)
34			Faith school (Sweep 6 Parents)
35			Same & mixed sex school (Sweep 5 Parents)
36			Same & mixed sex school (Sweep 6 Parents)
37			Class size (Teacher)
38			Time spent on homework (Teacher)
39			Time spent on homework (Sweep 5 Parent)
40			Help with homework (Sweep 5 Parents)
41			Full or part time education (Sweep 7)
42			Apprenticeship (Sweep 7)
43			Traineeship (Sweep 7)
44			University (Sweep 7)
Removed	**30**	**23**	**44**

Figure 2: Chi-Square Test of Independence results table

3.5.2. Cramer's V Test Results

Although the Chi-Square Test significantly reduced the original 92 features for all 3 datasets, when considering the V scores from the Cramer's V Test, a lot of the associations found were of negligible effect and so were discounted. After performing the Cramer V Test on the associated features for the 3 different datasets, only 12 features remained for the combined dataset, while 10 remained for the male dataset and 21 remained for the female data.

All Participants	P value	V value	Female Participant	P value	V value	Male Participants	P value	V value
				Academic Performance				
GCSE subject grades	0.0074	0.1458	GCSE subject grades	0.0002	0.2127			
Good at maths (Sweep 6)	0.0	0.111	Good at maths (Sweep 6)	0.0	0.1304			
			Good at maths (Sweep 5)	0.0	0.1004			
				Attitude towards School				
Unhappy at school (Sweep 5)	0.0	0.1208	Unhappy at school (Sweep 5)	0.0	0.1447	Unhappy at school (Sweep 5)	0.0	0.1218
Unhappy at school (Sweep 6)	0.0	0.2578	Unhappy at school (Sweep 6)	0.0	0.2869	Unhappy at school (Sweep 6)	0.0	0.1892
Tired at school (Sweep 6)	0.0	0.1727	Tired at school (Sweep 6)	0.0	0.188	Tired at school (Sweep 6)	0.0	0.115
Hard to focus on school (Sweep 6)	0.0	0.1913	Hard to focus on school (Sweep 6)	0.0	0.2257	Hard to focus on school (Sweep 6)	0.0	0.139
Happy with school (Sweep 6)	0.0	0.1801	Happy with school (Sweep 6)	0.0	0.2042	Happy with school (Sweep 6)	0.0	0.142
Happy with school work (Sweep 6)	0.0	0.1832	Happy with school work (Sweep 6)	0.0	0.199	Happy with school work (Sweep 6)	0.0	0.1563
School is a waste (Sweep 6)	0.0	0.1078	School is a waste (Sweep 6)	0.0	0.143	School is a waste (Sweep 5)	0.0	0.1069
			Enjoys school (Teacher)	0.0	0.1174			
			Enjoys school (Sweep 5 parent)	0.0	0.135			
			Finds school Interesting (Sweep 6)	0.0	0.1133			
			Tired at school (Sweep 5)	0.0	0.1039			
				Academic Expectations				
Likelihood university (Sweep 7)	0.0	0.1657	Likelihood university (Sweep 7)	0.0	0.2041	Likelihood university (Sweep 7)	0.0185	0.2015
Likelihood university (Sweep 6)	0.0	0.1375	Likelihood university (Sweep 6)	0.0002	0.1927	Move to secondary school (Sweep 5 parents)	0.0	0.1298
			Likelihood university (Teacher)	0.0	0.1009			
			Likelihood university (Sweep 7 parents)	0.0	0.1236			
			Prepared for secondary school (Teacher)	0.0	0.1256			
				Misbehaviour & Punishment				
Missed school (Sweep 6)	0.0	0.1032	Missed school (Sweep 6)	0.0	0.1032	Suspension (Sweep 6 parents)	0.0	0.1142
				Post 16 Activities				
			Where studying post school (Sweep 7)	0.0	0.1204			
12			**21**			**10**		

Figure 3: Cramer's V Test results table

After completing both the Chi-Square Test for Independence and the Cramer's V Test, the remaining variables for each of the 3 groups will be used to train a classifier model, designed to assess their importance and impact on the model's final prediction regarding cohort members' mental health.

3.6. Modelling

The goal for creating a classifier model was that once a list of education related features had been selected, the model would be trained to predict the mental health conditions of participants based on the features provided. Then, once the model had been trained, the importance and impact of certain features as predictors would be determined using SHAP values. However, before the final model could be trained and run, there were 2 major limitations of the data that needed to be overcome.

Firstly, there was a significant amount of missing data from a number of the selected features for the model. This had not been an issue earlier when performing the series of statistical tests, as all the ones used were able to ignore any NaN values. However, all the classifier models being considered did not have this capability and so these missing values had to be addressed. The solution to this issue was to replace the missing values, for which 2 different methods of data imputation were tested. Those 2 methods were mode imputation, which replaces all missing values with the most common value, and KNN imputation, which replaces the instances of a missing value with the nearest other instance (Ismiguzel, 2022). These 2 methods were tested on all 5 classifiers, to determine the effects on the different models. When comparing the 2 methods, there was little significant difference in performance, except in regard to the Decision Tree Classifier and the Random Forest Classifier, where the KNNImputer performed slightly better. For this reason, the KNNImputer was chosen as the data imputation methodology to use going forward.

The second issue which was encountered was the significant class imbalance between the 'Normal' and 'Abnormal' classifications, with only 18.4% of participants having the latter classification. This meant, that without any preprocessing, all models would simply predict the majority 'Normal' class in order to achieve a false 'high accuracy' of 81.6%. In order to rectify this issue, the solution was to perform a resampling to ensure that the majority and minority classifications were equal. Due to the small sample size of 8000, the only option was to perform an oversampling, where artificial entries are created to increase the instances of the minority class to be equal to the majority class. The main disadvantage with oversampling is that by creating exact copies of existing instances, there likelihood of overfitting increases (Weiss, 2007). For this reason, instead of a random duplication method, the synthetic minority oversampling technique (SMOTE) was used instead. SMOTE uses the k-nearest neighbour model to produce new artificial samples rather than replicating pre-existing samples, which should mitigate the degree of overfitting in the model. After performing the oversampling on the data, the performance across all classifier models significantly improved.

Finally, the last step before running the model was determining which classifier performed the best on the data provided. The 5 most common classifier models were chosen to be tested and included the Logistic Regression model, the Naïve Bayes model, the K-Nearest Neighbour model, the Decision Tree Classifier and the Random Forest Classifier. These models were trained on all the final selected features and assessed using k-fold cross validation on the 5 metrics of accuracy, precision, recall, F1 score and ROC AUC score, with the model's performance visualised through a confusion matrix. When examining the performance of each model, the Logistic Regression performed the weakest, followed by the Naïve Bayes model, while the Random Forest Classifier performed the best overall.

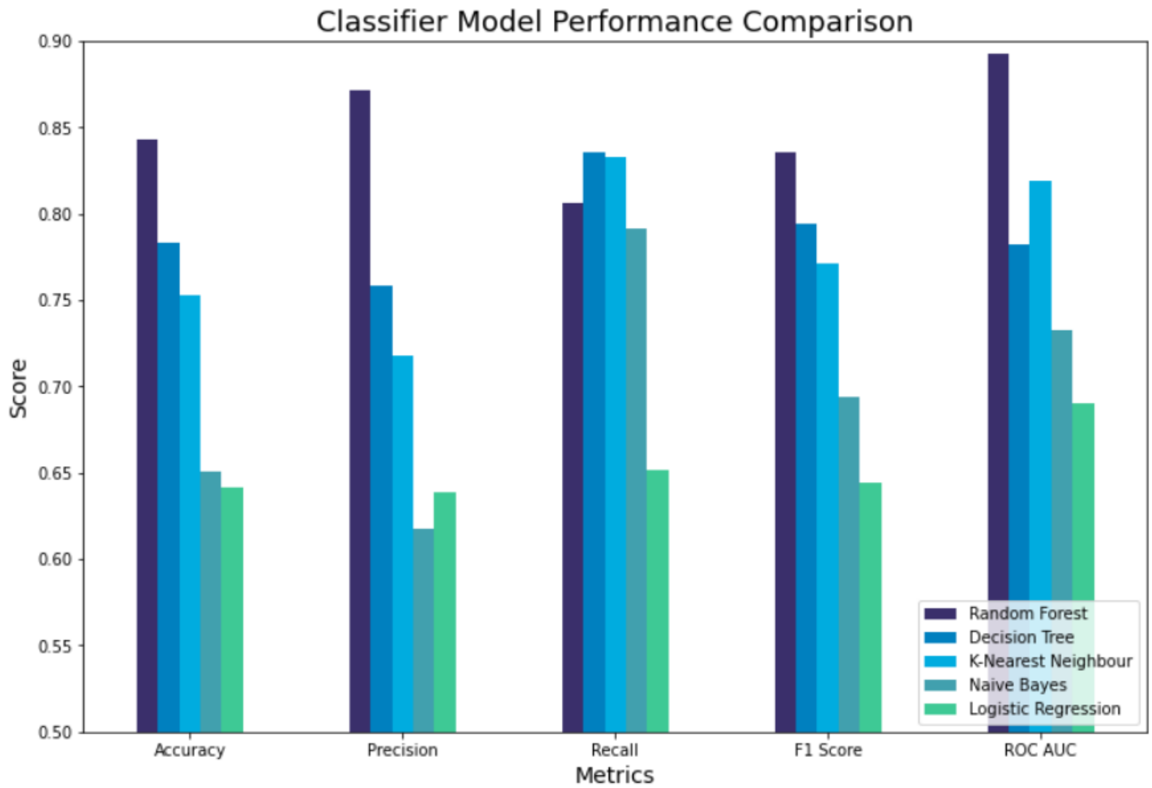


Figure 4: Classifier model performance comparison bar graph

Once the final classifier model had been chosen, the model was run over the final feature selection and the feature importance and impact was analysed using SHAP values. SHAP, otherwise known as Shapley Additive Explanations, is a method to explain individual predictions and is based on Shapley values from game theory (Lundberg & Lee, 2017). In essence the Shap value is the contribution of each feature to the prediction, with negative Shap values indicating a contribution towards the 0 classification, while a positive Shap value indicates a contribution towards the 1 classification (Lundberg & Lee, 2017). These Shap values helped demonstrate whether the association between the education related factors and mental health were positive or negative.

The following results section will outline in detail the performance of the different models, as well as the overall importance and impact certain features had on the model's final predictions regarding cohort members' mental health.

3.7. The Model Results

After constructing and resolving issues that arose during the modelling stage, the final model chosen was a random forest classifier. This classifier was then trained on the 3 different final feature sets chosen for the combined, female and male groups. Additionally, while the models were trained on the oversampled train set of data, the SHAP values were acquired using the test dataset which had not undergone any oversampling processing. This was to ensure that the SHAP values were derived from the original class proportions.

3.7.1. Final Model Performance

Similar to the metrics used to evaluate the different classifier models during the model building phase, the 3 random forest models were each evaluated again using k-fold cross validation with the 5 metrics of accuracy, precision, recall, f1 score and ROC AUC score. The 3 models where trained using the final features selected for each dataset. This meant that the model with the fewest features was the male group at 10, followed by the combined group at 12 and then finally the female group at 21 features. Figure 5 shows the different metric scores achieved for each of the 3 random forest models run.

	Accuracy	Precision	Recall	F1 Score	ROC AUC Score
Combined Dataset	0.8439	0.8719	0.8062	0.8377	0.8951
Female	0.8749	0.9136	0.8282	0.8687	0.9357
Male	0.8330	0.8391	0.8240	0.8314	0.8895

Figure 5: Final 3 models performance metrics table

Overall, the female model performed the best across all metrics, followed by the combined model and then the male model. Although even the male model which performed worst still has an accuracy of 83% and an F1 Score of 83%, which is consistent with 70-90% industry standard for a good model metric score (Barkved, 2022). The strong performance from all 3 models adds validity to the SHAP values used to determine feature importance and impact.

Reviewing the feature importance and impact using the SHAP values, each model ranked the features differently due to being trained on different feature sets.

3.7.2. Combined Model Results

Starting with the combined model, Figure 6 looks at feature importance and shows that feature **GCSE Subject Grades** was ranked the most important, followed by **Happy with School (Sweep 6)** and **Likelihood of University (Sweep 7)**. Meanwhile the feature **Missed School (Sweep 6)** was the least important feature, followed by **Unhappy at School (Sweep 5)** and **School is a waste (Sweep 6)**.

Additionally, Figure 7, which also orders the features in terms of importance, highlights how different values from the feature impact the final prediction. From this graph, we are able to see that lower values for the top 4 features are predictive of class 1 ('Abnormal'), while higher values are predictive of the class 0 ('Normal'). However, due to the ordinal nature of the majority of the final features selected, low and high values do not inherently indicate negative and positive connotations, but rather will vary.

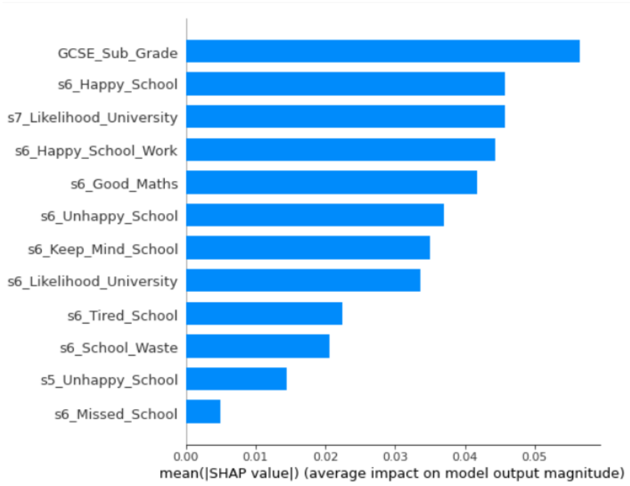


Figure 6: Combined model feature importance bar graph

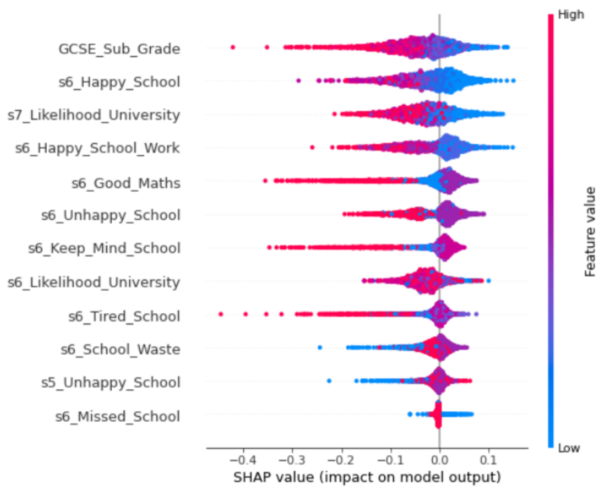


Figure 7: Combined model SHAP values summary graph

Figure 8 is a summary table created to outline all the features from the combined model, including information regarding who provided the information, their scoring and what type of relationship was highlighted through the SHAP values.

	Feature	Survey	Scoring	Association (0 – Normal | 1 – Abnormal)
1	GCSE Subject Grades	Qualifications	Score of overall grade performance, with a higher score indicating a better performance	Low values predict class 1 and high values predict class 0
2	Happy with School	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = '*Not at all happy'*	Low values predict class 1 and high values predict class 0
3	Likelihood of University	Sweep 7	0 – 100 scale indicating percentage chance of cohort member attending university	Low values predict class 1 and high values predict class 0
4	Happy with School Work	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = '*Not at all happy'*	Low values predict class 1 and high values predict class 0
5	Good at Maths	Sweep 6	0 – 3 Scale with 0 = *'Strongly Agree'* and 3 = '*Strongly Disagree'*	High values predict class 0 while no values strongly predict class 1
6	Unhappy at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The highest value (3) predicts class 0 while middle values predict class 1
7	Hard to focus on School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	High values predict class 0 while no values strongly predict class 1
8	Likelihood of University	Sweep 6	0 – 100 scale indicating percentage chance of cohort member attending university	No clear association pattern
9	Tired at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The highest value (3) predicts class 0 while no values strongly predict class 1
10	School is a Waste	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	Low values predict class 0 and the second highest value (2) predicts class 1
11	Unhappy at School	Sweep 5	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	No clear association pattern
12	Missed School	Sweep 6	Binary values with 0 = '*Yes'* and 1 = *'No'*	Value 0 (Yes) slightly predicts class 1.

Figure 8: Combined model feature impact summary table

Some of the more interesting feature association are reviewed in more detail using dependency graphs.

Looking at the top 4 features, the most important feature, **GCSE Subject Grade**, can be seen in Figure 9 to have a clear negative association between improving grade scores and predicting the 'Abnormal' class. A similar negative association can be seen with the feature **Likelihood of University (Sweep 7)**, although it is less distinct, with lower values predicting the 'Abnormal' class. This indicates that individuals who are more likely suffering from an 'Abnormal' mental health condition regard their likelihood of attending university as lower.

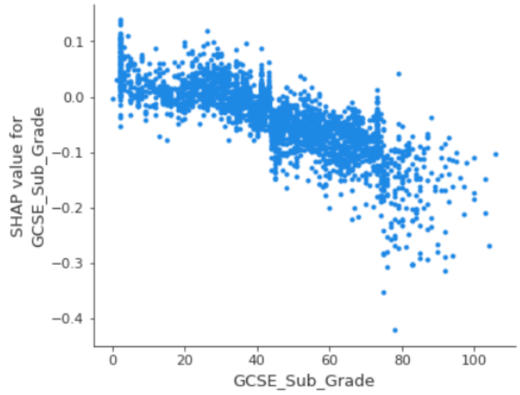


Figure 9: Combined Model – GCSE Subject Grade

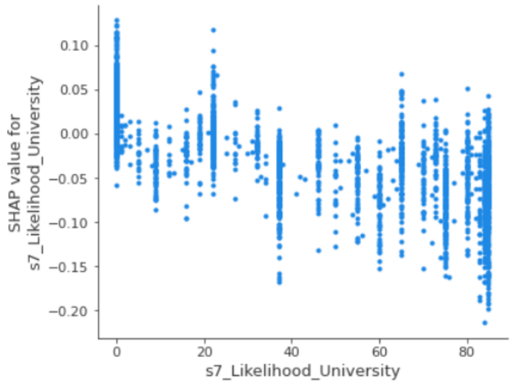


Figure 10: Combined Model – Likelihood of University (sweep 7)

Looking at the other 2 top features, **Happy with School (Sweep 6)** and **Happy with School Work (Sweep 6)**, which also both demonstrate a similar negative association with lower values predicting the 1 class. However, unlike the previous 2 features, these low values represent positive connotations. These features suggest an association between happiness with school and schoolwork and the 'Abnormal' mental health condition, which is unexpected.

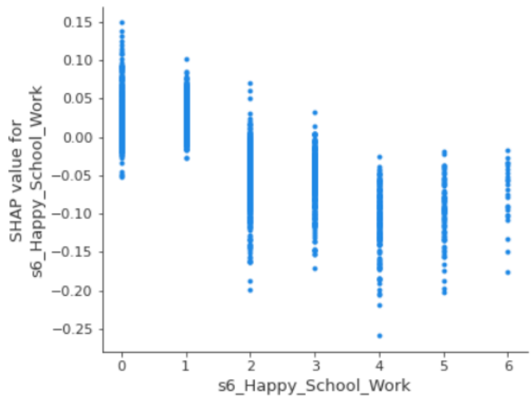


Figure 11: Combined Model – Happy with School Work (sweep 6)

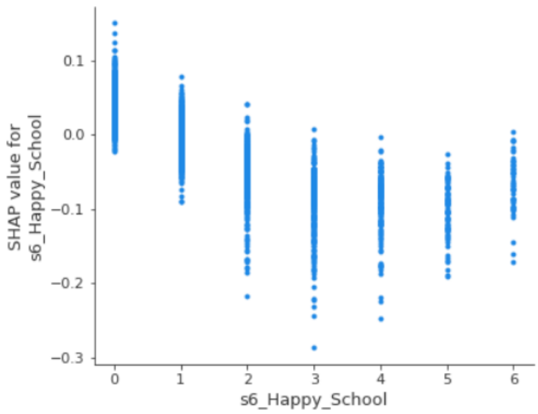


Figure 12: Combined Model – Happy with School (sweep 6)

Besides the 4 most important features, other interesting features included **Good at Maths (Sweep 6)**, **Likelihood of University (Sweep 6)** and **School is a Waste (Sweep 6)**. Starting with the feature, **Good at Maths (Sweep 6)**, which showed that high values were a good predictor for the 0 class while low values had little impact. This is interesting as it indicates that perceiving oneself to be bad at maths is not inherently associated with the 'Abnormal' class. Another interesting feature is the **Likelihood of University (Sweep 6)** which shows no clear pattern of association, unlike its corresponding feature at sweep 7. This change could potentially be due to the 3-year difference between the 2 sweeps. Finally, the last feature **School is a Waste (Sweep 6)** found that low values predicted the 0 class, while the value 2 predicted the 1 class. This was unusual as higher values represented a positive connotation, implying that not viewing school as a waste of time was associated with the abnormal class.

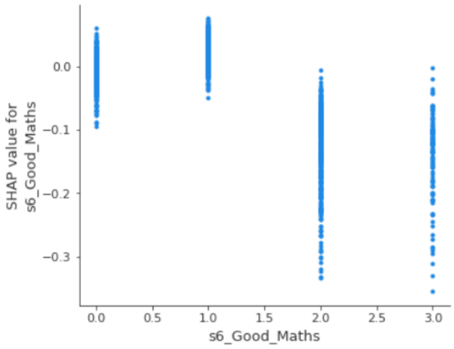


Figure 13: Combined Model – Good at Maths (sweep 6)

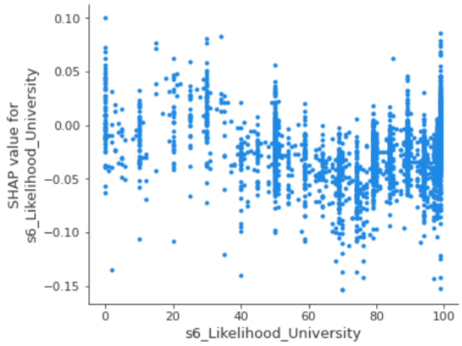


Figure 14: Combined Model – Likelihood of University (sweep 6)

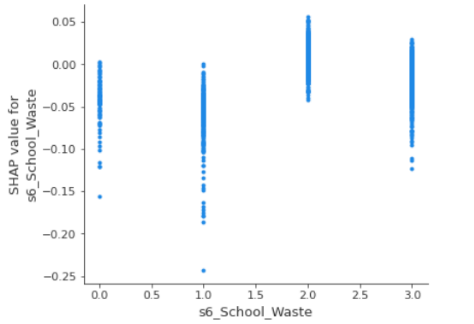


Figure 15: Combined Model – School is a Waste (sweep 6)

3.7.3. Male Model Results

Turning to look at the male model next, when examining Figure 16, the most important feature was **Likelihood of University (Sweep 7)**, followed by **Happy with School (Sweep 6)**. Meanwhile the least important features were **School is a Waste (Sweep 5)** and **Suspensions (Sweep 6)**.

When accounting for how the different values of certain features impacted the models' predictions, there was some variation. For the features **Likelihood of University (Sweep 7)** and **Unhappy at School (Sweep 6)**, higher values were good predictors for the 0 class. Meanwhile lower values were predictors for the 1 class for the features **Happy with School (Sweep 6)** and **Happy with School Work (Sweep 6)**. It is worth noting, just because one set of values predicts a certain class, this does not mean the opposite values will predict the other class as seen across all 3 models.

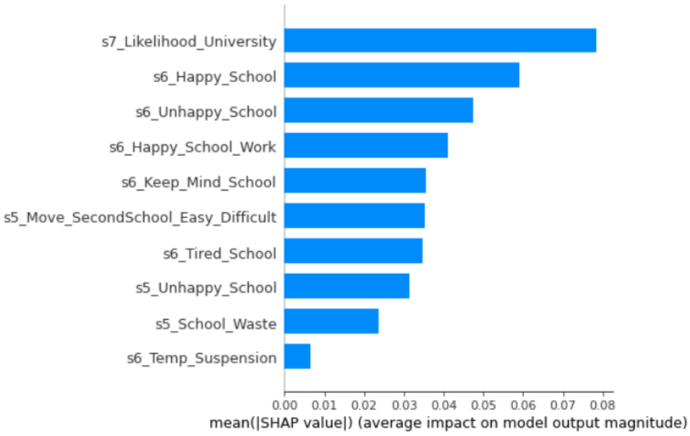


Figure 16: Male Model feature importance bar graph

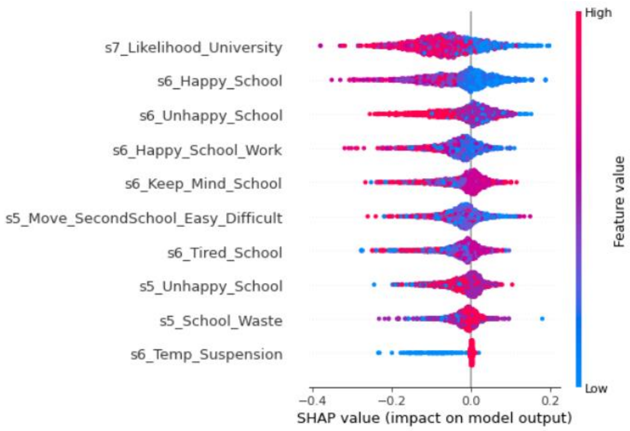


Figure 17: Male Model feature impact summary graph

Like previously, Figure 18 summarises the different features, their scoring system and the associations found by the SHAP values for the male model.

	Feature	Survey	Scoring	Association (0 – Normal | 1 Abnormal)
1	Likelihood University	Sweep 7	0 – 100 scale indicating percentage chance of cohort member attending university	Lower values slightly predict class 1 and higher values slightly predict class 0.
2	Happy with School	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = *'Not at all happy'*	Higher values predict class 0 while no values clearly predict class 1
3	Unhappy at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The highest value (3) predicts class 0, while middle values predict class 1
4	Happy with School Work	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = *'Not at all happy'*	Higher values predict class 0 while no values clearly predict class 1
5	Focused on School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	No clear association
6	Move to Secondary School	Sweep 5 Parents	0 – 4 Scale with 0 = *'Very Easy'* and 4 = *'Very Difficult'*	The highest value predicts class 0 while no other values have any strong association.
7	Unhappy at School	Sweep 5	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	No clear association but all values tended toward class 0
8	Tired at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	No clear association but all values tended toward class 0
9	School is a Waste	Sweep 5	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	Lower values slightly predict class 0, while the other values have no association.
10	Suspension	Sweep 6 Parents	Binary values with 0 = *'Yes'* and 1 = *'No'*	The lowest value (Yes) slightly predicts class 0.

Figure 18 Male Model feature summary table

Unlike the combined model, a lot of the features from the male model either tend towards predicting the Normal class or show no clear associations. This might be due to the lower percentage of cohort members with the 'Abnormal' classification, only 12.7%, when compared to the other datasets. Still, 3 features stood out as interesting in the male model which were further explored with dependency graphs, which were the top 3 features, **Likelihood of University (Sweep 7)**, **Happy with School (Sweep 6)** and **Unhappy at School (Sweep 6)**.

Starting with **Likelihood of University (Sweep 7)**, a negative association can be seen as lower values slightly predict class 1. However, this association is not very clear, with a lot of the other values showing a range of impact for both class 0 and 1. Then looking at the second feature, **Happy with School (Sweep 6)**, higher values clearly predict class 0, while lower values are shown to have no clear association. This feature tells us that there is an association between being unhappy with school and the Normal class, but the opposite association between being happy with school and the Abnormal class, is not true. Both these 2 features show similar trends to the combined model, although to a less clear degree. Finally, the last feature, **Unhappy at School (Sweep 6)**, shows that middle values predict class 1 prediction while higher values more strongly predict class 0. This is somewhat expected as it shows an association with frequent unhappiness at school and an Abnormal mental health condition, although it is surprising that the lowest score does not have a clear association with either class.

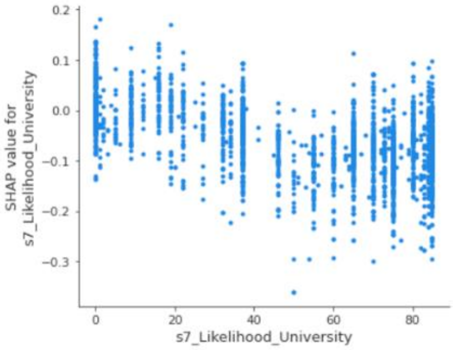


Figure 19: Male Model – Likelihood of University (sweep 7)

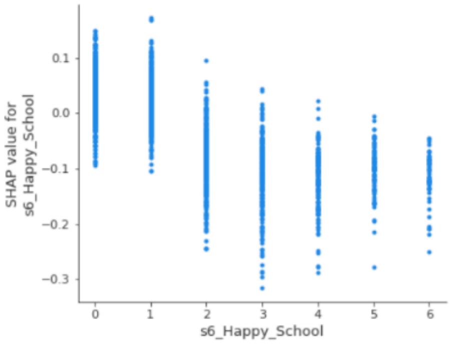


Figure 20: Male Model – Happy with School (sweep 6)

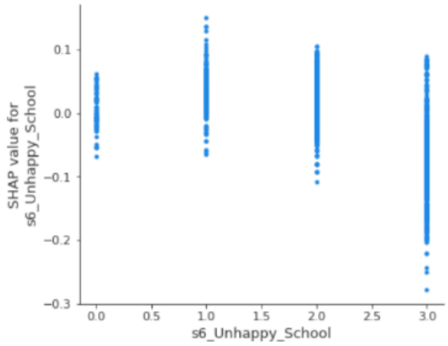


Figure 21: Male Model – Unhappy at School (Sweep 6)

3.7.4. Female Model Results

Finally, when reviewing the female model, Figure 22 shows the most important feature by a large margin is **GCSE Subject Grades** followed by both **Unhappy at School** features from sweep 5 and 6. While the least important feature was **Enjoys School (Sweep 5 Parents)**, followed by **School is Interesting (Sweep 6)** and **Where Studying or Training (Sweep 7)**.

Figure 23 then highlights how the 6 most important features, demonstrate a general pattern of lower values predicting the 'Abnormal' class, while the higher values are predictors for the 'Normal' class.

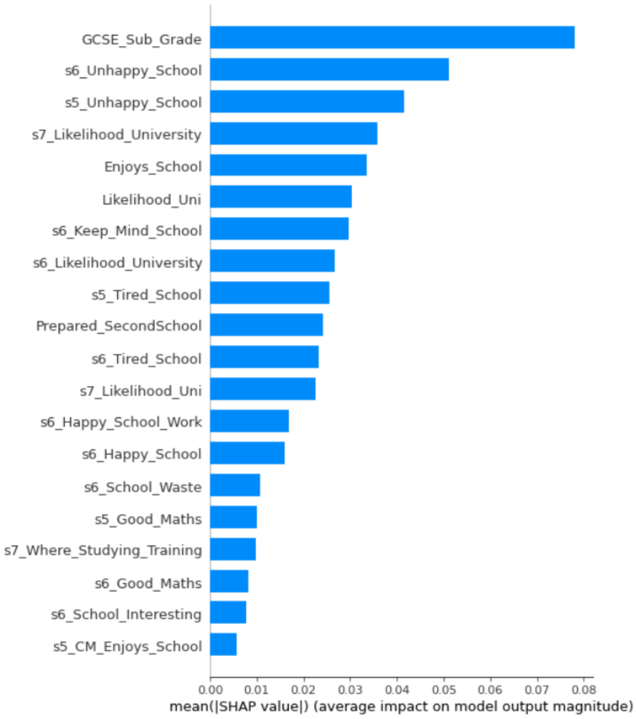


Figure 22: Female Model feature importance bar graph

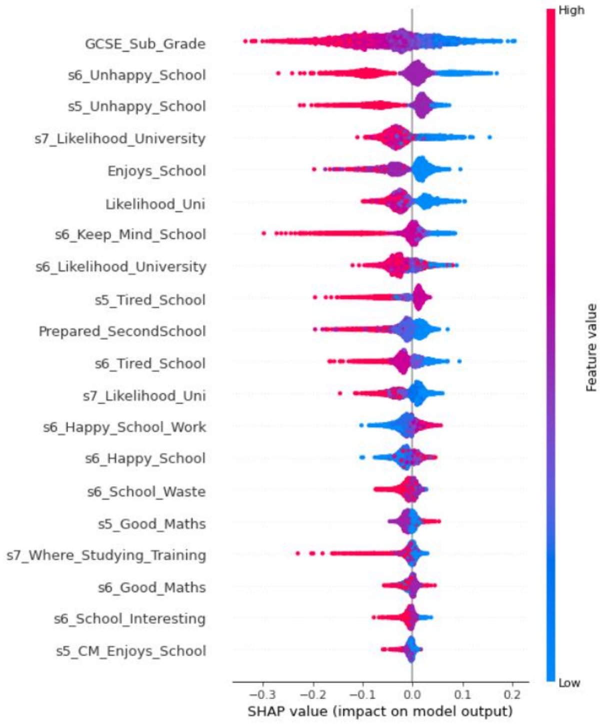


Figure 23: Female Model feature impact summary graph

The following Figure 24 is a summary of the patterns discovered about each features impact on the model's prediction of the 2 classes.

	Feature	Survey	Scoring	Association
1	GCSE Subject Grades	Qualification	Score of overall grade performance, with a higher score indicating a better performance	Low values predict class 1 while high values predict class 0.
2	Unhappy at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	Low values predict class 1 while high values predict class 0.
3	Unhappy at School	Sweep 5	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	All values predict class 1 except the highest value (3) which predicts class 0.
4	Likelihood University	Sweep 7	0 – 100 scale indicating percentage chance of cohort member attending university	Low values predict class 1 while high values predict class 0.
5	Enjoys School	Teacher	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The lowest value (0) predicts class 1 while all other values predict class 0
6	Likelihood Uni	Teacher	0 – 3 Scale with 0 = *'Very Likely'* and 3 = *'Very Unlikely'*	The lowest value (0) predicts class 1 while all other values predict class 0
7	Hard to Focus on School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The highest value (3) predicts class 0 while lower values predict class 1
8	Likelihood University	Sweep 6	0 – 100 scale indicating percentage chance of cohort member attending university	Lower values predict class 1 while higher values predict class 0. Except highest value (100) which shows no clear association.
9	Tired at School	Sweep 5	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	All values predict class 0, except for the second highest value (2) which predicts class 1
10	Prepared for Secondary School	Teacher	0 – 3 Scale with 0 = *'Very Prepared'* and 3 = *'Not Prepared'* Additional group 4 = *'Not going to Secondary School'*	The lowest value (0) predicts class 1, while all other values predict class 0.
11	Tired at School	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	Lower values predict class 1 while higher values predict class 0
12	Likelihood University	Sweep 7 Parents	0 – 3 Scale with 0 = *'Very Likely'* and 3 = *'Very Unlikely'* Additional group 4 = *'Do not Know' and 6 = 'No answer'*.	Lower values predict class 1 while higher values predict class 0.
13	Happy with School Work	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = *'Not at all happy'*	Low values predict class 0 while higher values predict class 1
14	Happy with School	Sweep 6	0 – 6 Scale with 0 = *'completely happy'* and 6 = *'Not at all happy'*	Lower values predict class 0, while other values don't have clear association
15	School is a Waste	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	No clear association except the highest value (3) which predicts class 0
16	Good at Maths	Sweep 5	0 – 3 Scale with 0 = *'Strongly Agree'* and 3 = *'Strongly Disagree'*	The second lowest value (1) predicts class 0, while the higher values slightly predict class 1.
17	Where studying post school	Sweep 7	6 categories: 1 = *'6th Form School'*, 2 = *'6th Form College'*, 3 = *'FE College'*, 4 = *'Other'*, 5 = *'Training Provider'*, and 6 = *'Specialist College'*	Lower values have no clear association, while the higher values predict class 0.
18	Good at Maths	Sweep 6	0 – 3 Scale with 0 = *'Strongly Agree'* and 3 = *'Strongly Disagree'*	No clear associations, but the highest value (3) slightly predicts class 1
19	Finds School Interesting	Sweep 6	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	The highest value (3) slightly predicts class 0, all other values have no clear association
20	Enjoys School	Sweep 5 Parents	0 – 3 Scale with 0 = *'Always'* and 3 = *'Never'*	Middle values slightly predict class 0, no other clear association.

Figure 24: Female Model feature summary table

Overall, when comparing the associations between all 3 models, the female model seems to display the clearest associations as can be further seen in the dependency graphs for certain features. This is potentially due to the fact that the female model had the highest amount of cohort members with the 'Abnormal' classification, 23.7%.

Starting with the top feature **GCSE Subject Grades** there is a very clear association between low values predicting the 'Abnormal' class while the higher values predict the 'Normal' Class. This shows a very clear relationship that there is a negative relationship between mental health conditions and overall grade attainment.

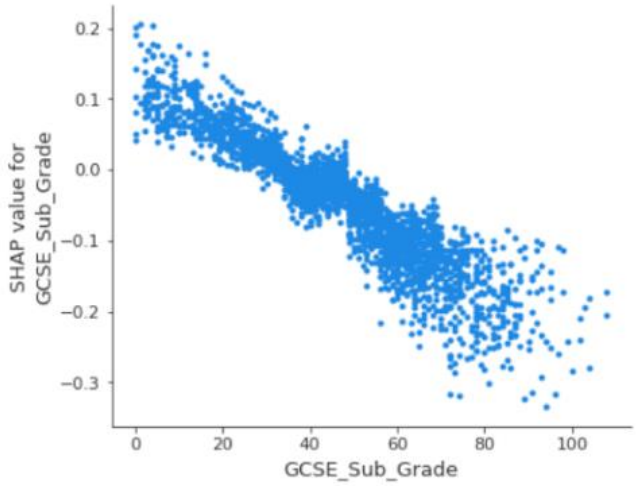


Figure 25: Female Model – GCSE Subject Grades

Next there are the 4 features which examine the cohort members likelihood of attending university as rated by different participants. Starting with the cohort members from sweep 6 and 7, there is a very similar pattern. Specifically, lower values tend to predict the 'Abnormal' class while higher values predict the 'Normal' class. The exception to this occurs in sweep 6, where the highest score of 100 seems to slightly predict the Abnormal class. This is unique as it implies there is a category of cohort members with the 'Abnormal' classification who hold high personal academic expectations which breaks the trend.

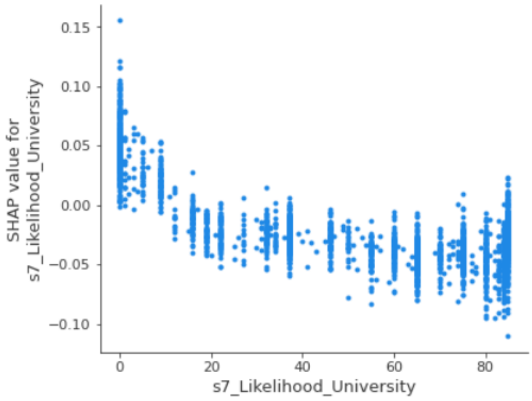


Figure 26: Female Model – Likelihood of University (sweep 7)

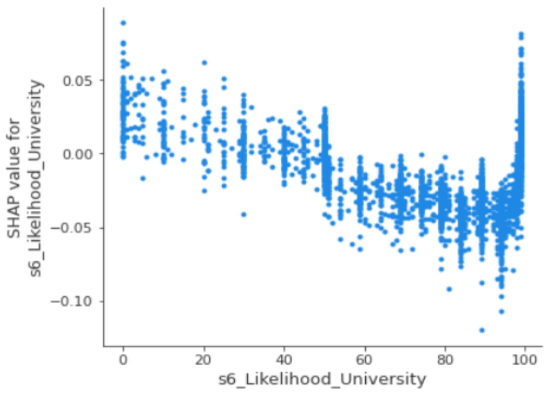


Figure 27: Female Model – Likelihood of University (sweep 6)

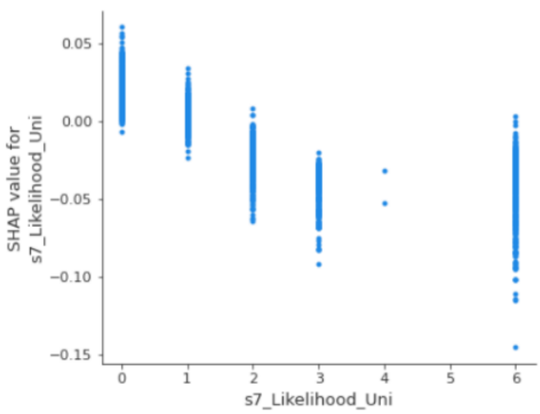


Figure 28: Female Model – Likelihood of University (sweep 7 parents)

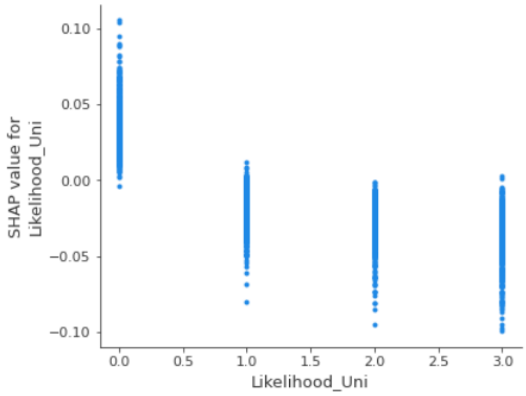


Figure 29: Female Model – Likelihood of University (teacher)

This pattern seen in the sweep 6 feature for likelihood of attending university can be seen to be mirrored ~~in~~ when looking at parents from sweep 7 and teachers from sweep 5. Specifically, it can be seen that lower values, which have a positive connotation in this case, predict the Abnormal class, while lower values predict the Normal class. This implies that external academic expectation is higher for female cohort members with the 'Abnormal' classification.

The next four features all look at cohort members happiness and enjoyment surrounding school. Starting with the feature **Unhappy at School** from both sweep 5 and 6, they both share a similar trend. In general, the lower values predict the 'Abnormal' class, while the highest value of 3 tends to strongly predict the Normal class. This indicates an association between being unhappy at school and having an Abnormal mental health condition.

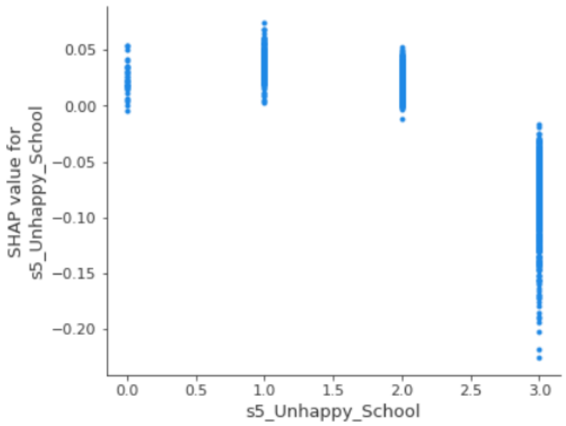


Figure 30: Female Model – Unhappy at School (sweep 5)

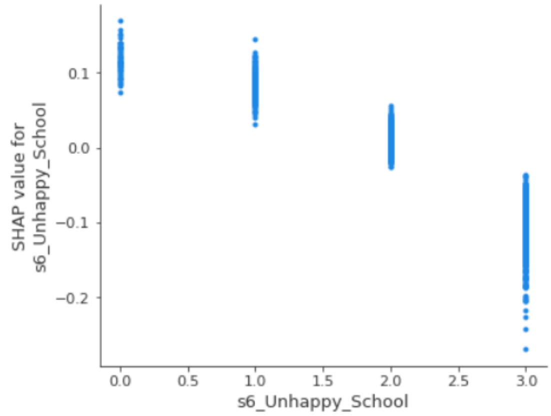


Figure 31: Female Model – Unhappy at School (sweep 6)

The trend seen above is also seen in the next feature of **Happy with School Work (Sweep 6)**, which shows that lower values predict class 0 and higher values predict class 1, indicating again that there is an association between being unhappy with school work and the Abnormal class. Looking at the final feature in this grouping, **Enjoys School (Teacher)**, this was interesting as it bucked the trend seen in the other features from this group. Low values tended to predict the Abnormal class, while middle and higher values predicted the normal class. This indicates that teacher's perceived female cohort members with the abnormal classification as enjoying school more than their peers.

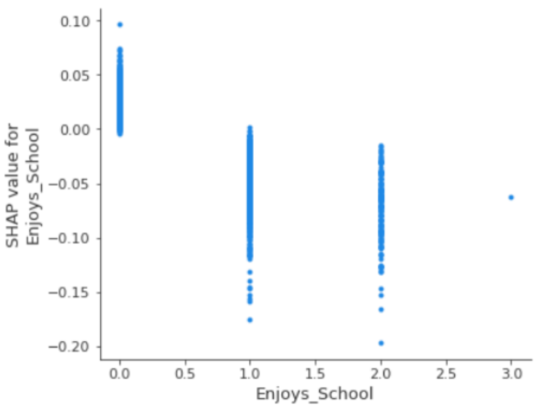


Figure 32: Female Model – Enjoys School (teacher)

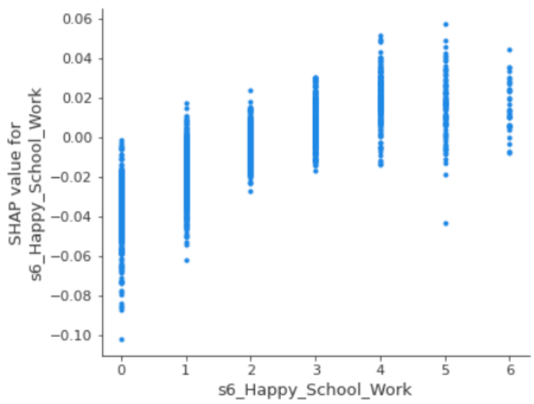


Figure 33: Female Model – Happy with School Work (sweep 6)

Finally, the last 2 features to be examined will be **Difficult to Focus on School (Sweep 6)** and **Tired at School (Sweep 6)**. Both these features have a similar trend with lower values predicting the Abnormal class and higher values predicting the Normal class. Both these features indicate an association between struggling to engage in school, either due to tiredness or lack of focus, and the cohort members mental health condition.

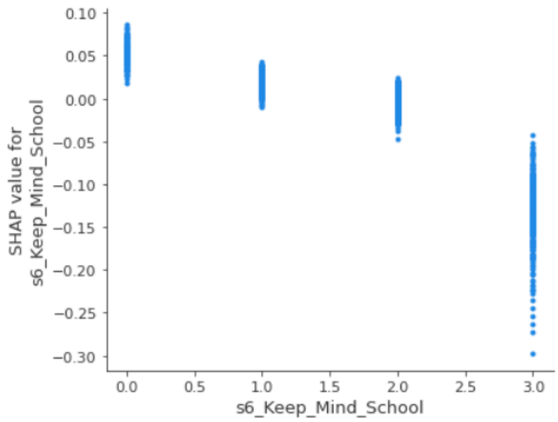


Figure 34: Female Model – Difficult to Focus on School (sweep 6)

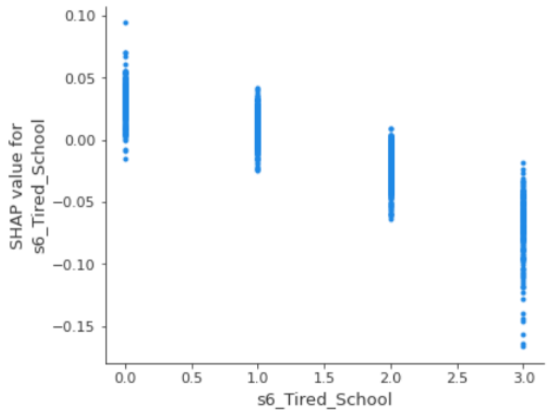


Figure 35: Female Model – Tired at School (sweep 6)

4.0. Discussion

4.1. School Performance

Starting with all features from the School Performance category, the male dataset found no associated features, while the combine and female dataset found three.

The **Good at Maths** feature from both sweep 5 and 6, which were often low ranked in terms of importance for both models due to a lack of a clear predictive pattern. The female model's **Good at Maths** features SHAP values were more definitive than the combined model. The highest feature value (3) was shown to typically predict the Abnormal class, implying that perceiving oneself as being bad at maths is associated with female pupils' poor mental health.

The other feature from this category, **GCSE Subject Grades**, was ranked as the most important feature for both the combined and female models. For both models, this feature's SHAP values show that lower values predict the Abnormal class while higher values predict the Normal class, although again this predictive pattern is clearer in the female model. Additionally, when looking at the Cramer's V scores for this feature, the female V score was 0.2127, while the combined score was 0.1458, indicating a stronger association in the female dataset. This is potentially the result of the male dataset not having any association with the **GCSE Subject Grades** feature.

Overall, there is a strong association between female cohort members academic performance and their mental health, while this association does not exist for male cohort members. This is unique from previous research addressed, which predominantly found that the academic performance of male participants was more heavily impacted by their mental health than their female participants (Smith et al, 2021; Lopez-Lopez et a, 2021; Riglin et al, 2013).

4.2. Attitude towards school

The category which examined the cohort members attitudes towards school contributed the most associated features for all 3 datasets.

Unsurprisingly, of those features the most significant in both the statistical tests and the model were ones that revolved around the cohort members happiness at school. Specifically, the following 4 features, **Happy with School (Sweep 6)**, **Happy with School Work (Sweep 6)**, and **Unhappy at School** from sweep 5 and 6, were present in every dataset.

Starting with the features, **Happy with School (Sweep 6)** and **Happy with School Work (Sweep 6)**, which ranked very high in terms of importance for the male and combined model. Overall, the general predictive pattern was lower values predicted the Abnormal class. This was unusual as for these features, lower values represented happiness, implying that being happier with school and school work was associated with Abnormal mental health for the male and combined dataset. Although, for the female model the predictive pattern reverses for these 2 features, with higher values predicting the Abnormal class instead.

Turning to the feature **Unhappy at School** from sweep 5 and 6, across all models the feature measured at sweep 6 was viewed as more important. For the male and combined model, the SHAP values were unable to show a clear predictive pattern, although in general the highest value (3) tended to predict the Normal class. Meanwhile with the female model, there is a clear pattern of lower values predicting the Abnormal class, while higher values predicted the Normal class.

Additionally, there are 2 further features which did not revolve around cohort members happiness. When examining both **Tired at School (Sweep 6)** and **Difficult to Focus on School (Sweep 6)** the features followed very similar trends. The female model SHAP values provided a clear predictive pattern with lower values predicting the Abnormal class, while the male and combined model had no clear association.

Overall, after examining the features in this category, the female model has had stronger predictive patterns when compared to the combined and male model. These findings imply that male participants mental health is less connected to the academic environment and instead are having greater impacts in other areas of their life. This is also supported by the fact that the male model has the least number of final features selected and why the SHAP values predictive patterns all tend towards a Normal prediction, even when there is no distinct pattern.

4.3. Misbehaviour & Punishment

All features selected from this category tended to be ranked the lowest in terms of feature importance across all 3 models. This is possibly due to the binary nature of the features and the imbalance in answers, with the Yes value having significantly fewer instances then the No value.

Looking at the first feature, **Missed School (Sweep 6)**, which was present in the combined and female models, it was ranked as the least important feature for both models. Although, the SHAP values found that the Yes value did slightly predict the Abnormal class for both models, which was consistent with previous research (Lereya & Deighton, 2019; Schulte-Korne, 2016).

The second feature is **Suspension (Sweep 6 Parents)** which only had a significant association in the male dataset but was ranked as the least important feature in the male model. The SHAP Values found that the Yes value slightly predicted the Normal class.

Overall, the primary conclusion that can be drawn from this category is that female participants suffering from a mental health disorder are more likely to be absent from school.

4.4. Academic Expectation

When looking at the category of academic expectations, the only feature which all groups had as being significantly associated was **Likelihood of University (Sweep 7)**. Across all the models, this feature was ranked very highly, with it being the most important feature for the male model. The SHAP values found that lower values predicted the Abnormal class for this feature while higher values tended to predict the Normal class. Although, like with previous features, the predictive pattern was clearest with the female model.

The only other shared feature for this category was **Likelihood of University (Sweep 6)**, which was present in the female and combined dataset and scored middling in terms of feature importance for both. The SHAP values for both models showed a predictive pattern with lower scores predicting the Abnormal class and higher scores predicting the Normal class, similar to the feature at sweep 7, although the pattern was more distinct for the female model. However, for the female model the highest value (100) was shown to slightly predict the Abnormal class. This is a change from the overall trend and implies female cohort members with very high personal expectations are also likely to suffer from an Abnormal mental health condition, possibly due to higher levels of stress.

The last 2 features which examine **Likelihood of University** are from the teacher and sweep 7 parental surveys, which only showed up in the female dataset as significantly associated. Both these features SHAP values showed that lower values predicted the Abnormal class. However, these lower values had positive connotations, implying a relationship between high expectations of female cohort members attending university and the Abnormal classification, which is a similar predictive pattern to the feature measured at sweep 6. This indicates that either, female cohort members are more likely to experience high external expectations in regard to higher education, or that female cohort members have higher personal expectations regarding their education which is perceived and then mirrored by others in their expectations.

The final feature from the female model worth mentioning is **Prepared for Secondary School (Teacher)** which was of middling importance to the model. This feature's SHAP values showed the lowest value, which represented 'very prepared', predicted the Abnormal class, while all other values predicted the Normal class. This feature continues to tie into the idea that there are higher expectations on female cohort members and their capabilities within an academic setting.

Overall, academic expectation was more associated with mental health for the female dataset when compared to the other 2 datasets. Additionally, although the overall predictive pattern was that lower personal academic expectation predicted the Abnormal class, there was also a clear outlier to this pattern in the female model, with higher personal and external academic expectations also predicting Abnormal class. This could be explained by a couple possible theories. For one, female students self-report higher levels of effort in their schoolwork and place more importance on hard work at school (Kessel & Heyder, 2017; Lam et al, 2012; McCrea et al, 2008). Additionally, female students tend to outperform male students at the secondary school level (Workman & Heyder, 2020; DiPrete & Jennings, 2012; Lam et al, 2012). Both these factors might result in higher personal and external academic expectations, which will come with higher levels of stress, a known risk factor for mental health.

4.5. Limitations

There are several limitations of this project that should be acknowledge when considering the results presented.

Firstly, the MCS has a number of linked administration datasets which include data on health and academic performance. However, these datasets contain personal information which could be used to identify participants. For this reason, the datasets are not available for use without extensive training and an ethics review. This makes it more difficult to attain hard data on the cohort's academic achievements, but this was mitigated by using self-reported and perceived subject performance data, as well as self-reported qualification performance data. These data points can act as a suitable substitute for the linked education administration dataset, although there is a risk that this data is not 100% accurate and objective, meaning any findings relating to academic performance is less reliable.

Secondly, although this dissertation would like to establish causal relationships between educational factors and mental health, this is not possible. Unfortunately, in order to establish causality, there are 3 conditions which must be met (Antonakis et al, 2010). Firstly, the cause needs to have occurred before the effect. Secondly, the relationship between the cause and the effect cannot have occurred by chance. Thirdly, there cannot be other intervening or unaccounted variables that could be responsible for the causal relationship. Unfortunately, only the second condition is easily met through statistical tests with the available data and the current time and resource constraints on this project. The first condition is difficult to meet due to the inconsistent measurement of self-reported mental health, making it hard to establish a timeline for each participant. The third condition can't be met due to the complexity of both education and mental health, and the many different variables which impact those fields, making it almost impossible to account for every confounding variable which could be impacting that relationship. Due to the design of this project, there is no way to truly establish a causal relationship between education related factors and mental health. Normally causality is confirmed through either randomized controlled experiments or the use of multiple data sources that contain 2 homogenous groups to compare (Antonakis et al, 2014). However, despite this, this project can still provide support for associations found between education and mental health which can be further researched in projects with a greater time and resource threshold.

Finally, this research unfortunately does not account for confounding variables like race, socioeconomic backgrounds, parental and peer relationships, health, and living conditions, to name a few. These are some of the known factors which could explain the potential associations found. These were not accounted for due to the large scope of this dissertation, the number of confounding variables which needed to be addressed was significant, with many not being available in the sample used. Despite this, these associations can still provide potential avenues for further investigation in studies with more of a focused question, that can account for the numerous confounding variables to prove a causal relationship.

4.6. Conclusion

To conclude, the 2 main research questions were answered for this project with a number of significant associations found between education and mental health. Additionally, the gender difference in these association explored, with female cohort members showing stronger associations then their male counterparts. There were 4 key discoveries' though the course of answering the 2 research questions which can be summarised as follows.

Firstly, the feature **GCSE Subject Grades** was found to have a significant association with female mental health, but this association was absent from the male dataset. Additionally, this feature ranked very highly in the female model, with a worse overall grade performance indicating poor mental health. This is unique from the previous research done into this field, with a number of studies finding a stronger association in male participants. However, this finding adds further complexity to the debate surrounding the gender difference in academic performance and mental health.

Secondly, the male dataset had the fewest associations between education related features and mental health, as well as the SHAP values for a majority of their features showing a preference for the Normal class, even when there was no clear predictive pattern. This suggests that boys' mental health is potentially not as connected to education, instead being impacted by other areas of their lives.

Thirdly, the feature **Missed School (Sweep 6)** showed that missing school had a significant association with 'Abnormal' mental health, but only for the female dataset. This further supports previous research which examined the connection between absenteeism and mental health.

Fourthly, the feature **Likelihood of University** at sweep 7 was found to be associated with both the female and male dataset, with lower expectations being associated with mental ill health. Although, it was also found that higher personal and external academic expectations were also associated with poor mental health for the female dataset. These higher expectations of female cohort members could be potentially explained by a number of factors, as previously discussed, which then leads to higher levels of stress that can result in a worse mental health condition.

Finally, it should be noted that there was a female bias in this study, with female cohort members being overrepresented in the final sample by 300 instances. Additionally, female cohort members had a higher Abnormal or Normal percentage ratio of 23.7%, compared to the 12.7% of male participants. This could possibly explain why the gender difference leans heavily in the female favour.

Overall, these 4 key discoveries both build upon old discussions as well as indicating potential new avenues of research. This dissertation has achieved the goals it set out with, and provide ground proof for further research into these 2 fields.

5.0. References

Agnafors, S., Barmark, M. & Sydsjo, G. (2021). Mental Health and academic performance: A study on selection and causation effects from childhood to early adulthood. *Social Psychiatry and Psychiatric Epidemiology, 56:857-866.*

Antonakis, J., Bendahan, S., Jacquart, P. & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly, 21(6):1086-1120.*

Antonakis, J., Bendahan, S., Jacquart, P. & Lalive, R. (2014). "*Causality and Endogeneity: Problems and Solutions.*" In D. V. Day (Ed.), The Oxford Handbook of Leadership and Organisation. New York: Oxford University Press.

Avendano, M., De Coulon, A. & Nafilyan, V. (2017). *Does more education always improve mental health? Evidence from a British Compulsory schooling reform.* Health, Econometrics and Data Group, Department of Economics, University of York.

Barkved, K. (2022). How to know if your machine learning model has good performance. *Obviously AI, 9th March 2022.* https://www.obviously.ai/post/machine-learning-model-performance (Accessed 14th August 2022)

Cavioni, V., Grazzani, I., Ornaghi, V., Agliati, A. & Pepe, A. (2021). Adolescents' Mental Health at School: The Mediating Role of Life Satisfaction. *Frontiers in Psychology, 12.*

Chung, A. Y. (2016). *Saving Face: The Emotional Costs of the Asian Immigrant Family Myth.* New Brunswick: Rutgers University Press.

Chung, A. Y. (2017). Behind the Myth of the Matriarch and the Flagbearer: How Korean and Chinese American Sons and Daughters Negotiate Gender, Family and Emotions. *Sociological Forum, 32(1):28-49.*

Cohen, A. K., Nussbaum, J., Weintraub, M. L. R., Nichols, C. R. & Yen, I. H. (2020). Association of Adult Depression with Educational Attainment, Aspirations, and Expectations. *Preventing Chronic Disease, 17.*

Cornaglia, F., Crivellaro, E. and McNally, S. (2012). Mental Health and Education Decisions. *Discussion Paper No. 6452.*

Crenna-Jennings, W. (2021). Young people's mental and emotional health: Trajectories and drivers in childhood and adolescence. *Education Policy Institution.*

Dantzer, R. (2012). Depression and inflammation: An intricate relationship. *Biological Psychiatry, 711:4-5.*

Dantzer, R., O'Connor, J. C., Lawson, M. A. & Kelley, K. W. (2011). Inflammation-associated depression: From serotonin to kynurenine. *Psychoneuroendocrinology, 363:426-436.*

Deighton, J., Lereya, S. T. & Wolpert, M. (2021). Enduring Mental Health in Childhood and Adolescence: Learning from the Millennium Cohort Study. *Journal of the American Academy of Child and Adolescent Psychiatry, 60(8):1030-1039.*

DiPrete, T. A. & Jennings, J. L. (2012). Social and behavioral skills and the gender gap in early educational achievement. *Social Science Research, 41(1):1–15.*

Echstein, Z. & Wolpin, K. I. (1999). Why Youths Drop out of High School: The Impact of Preferences, Opportunities and Abilities. *Econometrica, 67(6):1295-1339.*

Ford, T., Esposti, M. D., Crane, C., Taylor, L., Montero-Marin, J., Blakemore, S., Bowes, L., Byford, S., Dalgleish, T., Greenberg, M. T., Nuthall, E., Phillips, A., Raja, A., Ukoumunne, C., Viner, R. M., Williams, M. G., Allwood, M., Aukland, L., Casey, T., De Wilde, K., Farley, E., Kappelmann, N., Lord, L., Medlicott, E., Palmer, L., Petit, A., Pryor-Nitsch, I., Radley, L., Warriner, L. & Sonley, A. (2021). The Role of Schools in Early Adolescents' Mental Health: Findings from the MYRIAD Study. *Journal of the American Academy of Child & Adolescent Psychiatry, 60(12):1467-1478.*

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlation. *Psychology in the Schools, 30:79-90.*

Goodman, R., Meltzer, H. & Bailey, V. (1998). The Strengths and Difficulties Questionnaire: A pilot study on the validity of the self-report version. *European Child and Adolescent Psychiatry, 7:125-130.*

Hale, D. R., Patalay, P., Fitzgerald-Yau, N., Hargreaves, D. S., Bond, L., Gözig, A., Wolpert, M., Stansfeld, S. A. & Viner, R. M. (2014). School-level variation in health outcomes in adolescence: Analysis of three longitudinal studies in England. *Prevention Science, 15(4):600-610.*

Hjorth, C. F., Bilgrav, L., Frandsen, L. S., Overgaard, C., Torp-Pedersen, C., Nielsen, B. & Boggild, H. (2016). Mental health and school dropout across educational levels and gender: a 4.8-year follow-up study. *BMC Public Health, 16.*

Ismiguzel, I. (2022). Imputing Missing Data with Simple and Advanced Techniques. *Towards Data Science, 12th May 2022.* https://towardsdatascience.com/imputing-missing-data-with-simple-and-advanced-techniques-f5c7b157fb87 (Accessed 15th August 2022)

Jones, P. B. (2013). Adult Mental Disorders and their Age at Onset. *The British Journal of Psychiatry, 202:5-10.*

Joyce, H. D. & Early, T. J. (2014). The impact of school connectedness and teacher support on depressive symptoms in adolescents: A multilevel analysis. *Children & Youth Services Review, 39:101-107.*

Kearney, M. W. (2017). *"Cramer's V".* Sage Encyclopedia of Communication Research Methods, M. R. Allen (eds). United States, California: Sage Publications.

Kessels, U. & Heyder, A. (2017). Personal value placed on academic effort as a mediator of gender differences in grades: A domain-specific analysis. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 49:86–97.*

Kessler, R. C., Barker, P. R., Colpe, L. J., Epstein, J. F., Gfroerer, J. C., Hiripi, E., Howes, M. J., Normand, S. L., Manderscheid, R. W., Walters, E. E. & Zaslavasky, A. M. (2003). Screening for serious mental illness in the general population. *Archives of General Psychiatry, 60(2):184-189.*

Kessler, R., Berglund, P., Demler, O., Jin, R., Merikangas, R. & Walters, E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry, 62:593-602.*

Kim-Cohen, J., Caspi, A., Moffitt, T. E., Harrington, H., Milne, B. J. & Poulton, R. (2003). Prior juvenile diagnoses in adults with mental disorder: development follow-back of a prospective-longitudinal cohort. *Archives of General Psychiatry, 60(7):709-717.*

Lam, S., Jimerson, D., Kikas, E., Cefari, C., Veiga, F. H., Nelson, B. Hatzichristou, C., Polychroni, F., Basnett, J., Duck, R. & Farrell, P. (2012) Do girls and boys perceive themselves as equally engaged in school? The results of an international study from 12 countries. *Journal of school psychology, 50(1):77-94.*

Lee, J. C. & Zhou, M. (2015). *The Asian American Achievement Paradox.* New York: Russell Sage Foundation.

Lereya, S. T., Patel, M., Dos Santos, J. P. G. A. & Deighton, J. (2019). Mental Health Difficulties, Attainment and Attendance: A cross-sectional study. *European Child & Adolescent Psychiatry, 28:1147-1152.*

Lereya, T. & Deighton, J. (2019). Learning from HeadStart: the relationship between mental health and school attainment, attendance and exclusions in young people aged 11 to 14. London: EBPU.

Lopez-Lopez, J. A., Kwong, A. S. F., Washbrook, L., Tilling, K., Fazel, M. S. & Pearson R. M. (2021). Depressive symptoms and academic achievement in UK adolescents: A cross-lagged analysis with genetic covariates. *Journal of Affective Disorders, 284:104-113.*

Lundberg, S. M. & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information and processing systems, 30.*

Mboya, I. B., John, B., Kibopile, E. S., Mhando, L., George, J. & Ngocho, J. S. (2020). Factors associated with mental distress among undergraduate students in northern Tanzania. *BMC Psychiatry, 20(28):1-7.*

McCrea, S. M., Hirt, E. R. & Milner, B. J. (2008). She works hard for the money: Valuing effort underlies gender differences in behavioral self-handicapping. *Journal of Experimental Social Psychology, 44(2):292–311.*

McHugh, M. L. (2013). The chi-square test of independence. *Biochemia Medica, 23(2):143-149.*

Mcleod, J. D. & Kaiser, K. (2004). Childhood emotional and behavioural problems and educational attainment. *American Sociological Review, 69(5):636-658.*

Miech R. A., Caspi, A., Moffitt, T. E., Wright, B. R. E. & Silva, P. A. (1999). Low socioeconomic status and mental disorders: A longitudinal study of selection and causation during young adulthood. *American Journal of Sociology, 104(4):1096-1131.*

Moylan, S., Maes, M., Wray, N. R. & Berk, M. (2013). The neuroprogressive nature of major depressive disorder: Pathways to disease evolution and resistance, and therapeutic implications. *Molecular Psychiatry, 185:595-606.*

NHS Digital. (2021). Mental Health of Children and Young People in England 2021 – Wave 2 follow up to the 2017 survey. *NHS Digital, 30th September 2021.* https://digital.nhs.uk/data-and-information/publications/statistical/mental-health-of-children-and-young-people-in-england/2021-follow-up-to-the-2017-survey/copyright (Accessed 8th June 2022)

NHS England. (2015). Future in Mind – Promoting, protecting, and improving our children and young people's mental health and wellbeing. *Department of Health.*

NHS England. (2022). NHS Mental Health Dashboard, *NHS England.* https://www.england.nhs.uk/mental-health/taskforce/imp/mh-dashboard/#:~:text=When%20spend%20on%20specialised%20commissioning,14.98%20billion%20in%202021%2F22. (Accessed 13th August 2022)

OECD. (2017). *Students' Well-being.* PISA 2015 Results (Volume III). Paris, France.

Pascoe, M. C., Hetrick, S. E. & Parker, A. G. (2019). The impact of stress on students in secondary school and higher education. *International Journal of Adolescence and Youth, 25(1):104-112.*

Patalay, P. & Fitzsimons, E. (2018). Development and predictors of mental ill-health and wellbeing from childhood to adolescence. *Social Psychiatry and Psychiatric Epidemiology, 53(12):1311-1323.*

Patalay, P., O'Neill, E., Deighton, J. & Fink, E. (2020). School characteristics and children's mental health: A linked survey-administrative data study. *Preventive Medicine, 141.*

Pretty, G. M., Andrewes, L. & Collett, C. (1994). Exploring adolescents' sense of community and its relationship to loneliness. *Journal of Community Psychology, 22:346-358.*

Relative Insight. (2021). Mental Health Awareness: How has the media conversation changed? *Relative Insight, 14th May 2021.* https://relativeinsight.com/mental-health-awareness-how-has-the-media-conversation-changed/ (Accessed 13th August 2022)

Riglin, L., Fredrickson, N., Shelton, K. H. & Rice, F. (2013). A longitudinal study of psychological functioning and academic attainment at the transition to secondary school. *Journal of Adolescence, 36(3): 507-517.*

Roeger, L., Allison, S., Martin, G., Dadds, V. & Keeves, J. (2001). Adolescent depressive symptomatology: Improve schools or help students? *Australian Journal of Psychology, 53:134-139.*

Sani, M., Mahfouz, M., Bani, I., Alsomily, A., Alagi, D., Alsomily, N., Madkhaly, F., Madkhali, R., Hakami, A. & Hakami, A. (2012). Prevalence of stress among medical students in Jizan University, Kingdom of Saudi Arabia. *Gulf Medical Journal, 1(1):19-25.*

Schulte-Korne, G. (2016). Mental Health Problems in a School Setting in Children and Adolescents. *Deutsches Arzteblatt International, 113(11):183-190.*

Sellers, R., Warne, N., Pickles, A., Maughan, B., Thaper, A. & Collishaw, S. (2019). Cross-Cohort change in adolescent outcomes for children with mental health problems. *Journal of Child Psychology and Psychiatry, 60(7):813-821.*

Smith, N. R., Marshall, L., Albakri, M., Smuk, M., Hagell, A. & Stanfeld, S. (2021). Adolescent mental health difficulties and educational attainment: Findings from the UK household longitudinal study. *BMJ Open, 11:1-10.*

Sorberg Wallin, A., Koupil, I., Gustafsson, J. E., Zammit, S., Allebeck, P. & Falkstedt, D. (2019). Academic Performance, externalizing disorders and depression: 26,000 adolescents followed into adulthood. *Social psychiatry and psychiatric epidemiology, 54(8):977-986.*

Sreeramareddy, C. T., Shankar, P. R., Binu, V. S., Mukhopadhyay, C., Ray. B. & Menezes, R. G. (2007). Psychological morbidity, sources of stress and coping strategies among undergraduate medical students of Nepal. *BMC Medical Education, 7(26)*

UCL. (2022). Millennium Cohort Study. *University College London, Centre for Longitudinal Studies.* https://cls.ucl.ac.uk/cls-studies/millennium-cohort-study/ (Accessed 1st July 2022)

UNESCO. (2012). *International standard classification of education (ISCED) 2011*. United Nations Educational, Scientific and Cultural Organisation: Institute for Statistics. Montreal, Quebec.

Van Der Ende, J., Verhulst, F. C. & Tiemeier, H. (2016). The Bidirectional pathways between internalizing and externalizing problems and academic performance from 6 to 18 years. *Development and Psychopathology, 28(3):855-867.*

Venters, N. (2018). The past, present and future of innovation in mental health. *NHS Digital, 10th July 2018.* https://digital.nhs.uk/blog/transformation-blog/2018/the-past-present-and-future-of-innovation-in-mental-health#:~:text=One%20major%20change%20has%20been,health%20professionals%20and%20seek%20treatment (Accessed 13th August 2022)

Warikoo, N., Chin, M., Zillmer, N. & Luthar, S. (2020). The Influence of Parent Expectations and Parent-Child Relationships on Mental Health in Asian America and White American Families. *Sociological Forum, 35(2).*

Weiss, G. M., McCarthy, K. & Zabar, B. (2007). Cost-sensitive learning vs. sampling: Which is best for handling unbalanced classes with unequal error costs? *Dmin, 7(35-41):24-31.*

WHO. (2022). Adolescent Mental Health. *World Health Organisation.* https://www.who.int/news-room/fact-sheets/detail/adolescent-mental-health (Accessed 8th June 2022).

Workman, J. & Heyder, A. (2020). Gender Achievement gaps: the role of social costs to trying hard in high school. *Social Psychology of Education, 23:1407-1427.*

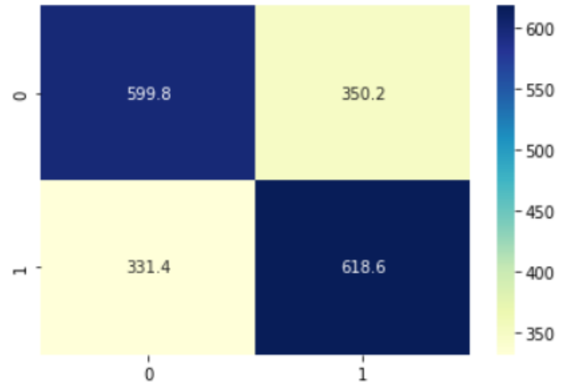
Youth. (2022). How mental health disorders effect youth. *Youth.gov.* https://youth.gov/youth-topics/youth-mental-health/how-mental-health-disorders-affect-youth (Accessed 15th August 2022)

6.0. Appendix

6.1. Model Comparison – Confusion Matrix

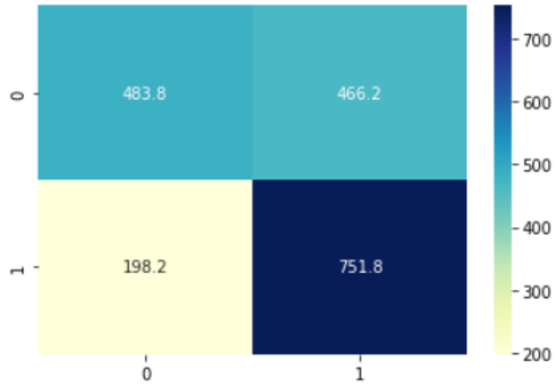
Logistic Regression Model

```
Accuracy: 0.6412631578947369
Precision: 0.6384079157697671
Recall: 0.6511237519620913
F1: 0.6445680304200692
Roc AUC Score: 0.6899686530782339
```



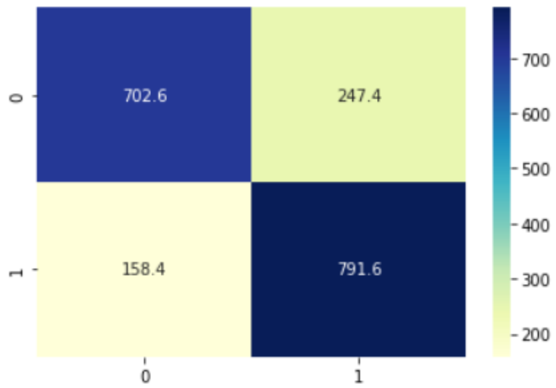
Naïve Bays Model

```
Accuracy: 0.6503157894736843
Precision: 0.6172917296755173
Recall: 0.7912566164464421
F1: 0.6934509225408527
Roc AUC Score: 0.7327031303469838
```



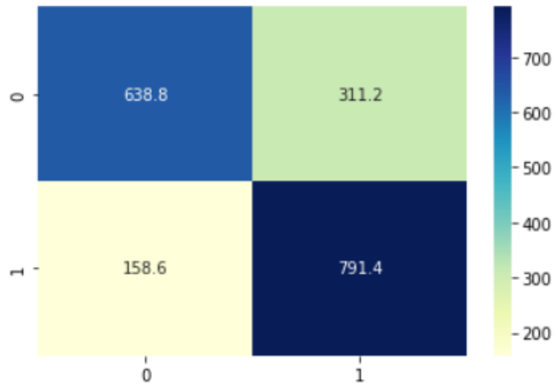
Decision Tree Classifier

```
Accuracy: 0.7856842105263159
Precision: 0.759805483928489
Recall: 0.8355791076463129
F1: 0.795370374867616
Roc AUC Score: 0.784934504248326
```



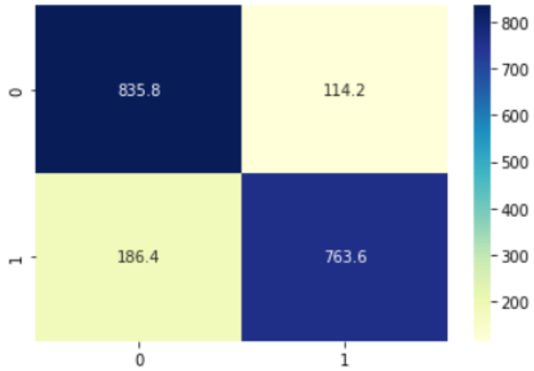
K-Nearest Neighbour Model

```
Accuracy: 0.7527368421052631
Precision: 0.7178098566760677
Recall: 0.8330274267466141
F1: 0.7710840494579847
Roc AUC Score: 0.8189214712776176
```



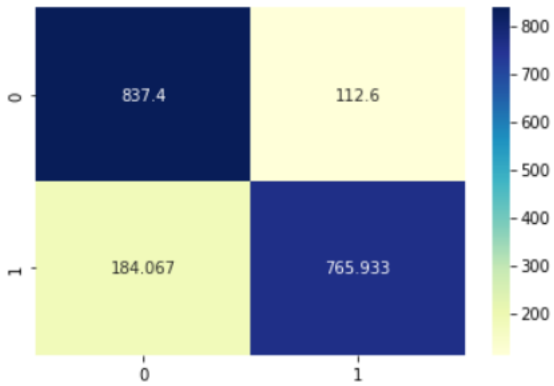
Random Forest Model

```
Accuracy: 0.842736842105263
Precision: 0.8695848315368556
Recall: 0.8045571411702381
F1: 0.8369357555252405
Roc AUC Score: 0.8931531012702809
```

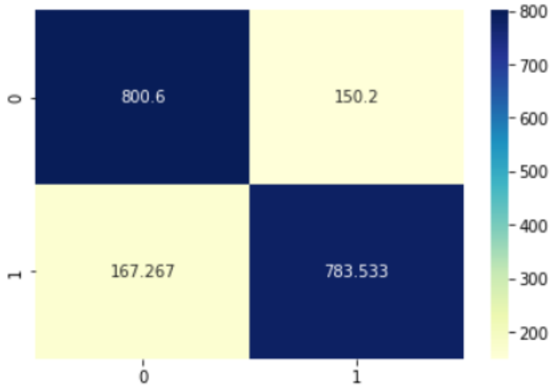


6.2. Final Models – Confusion Matrix

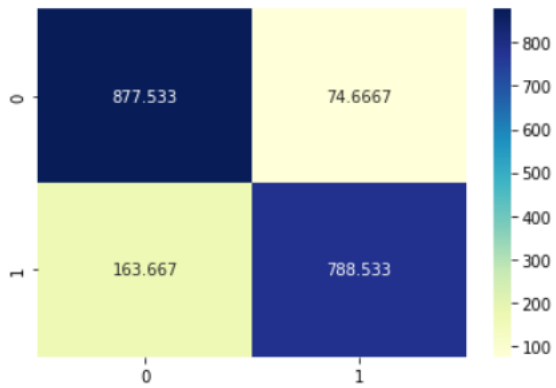
Combined Model



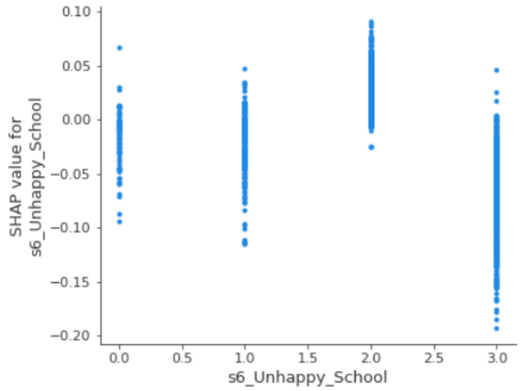
Male Model

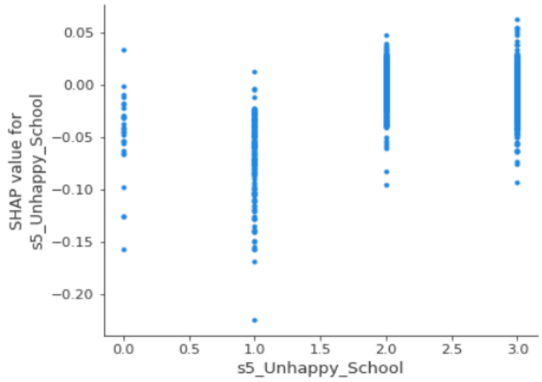


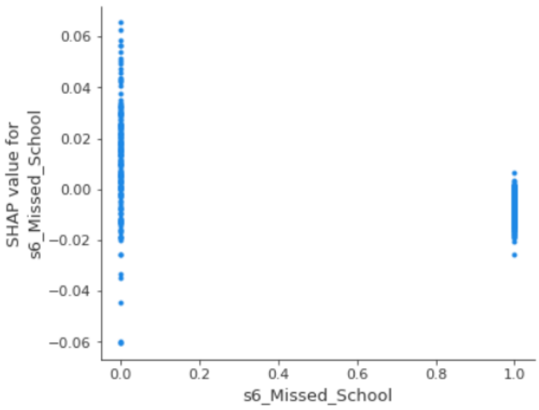
Female Model

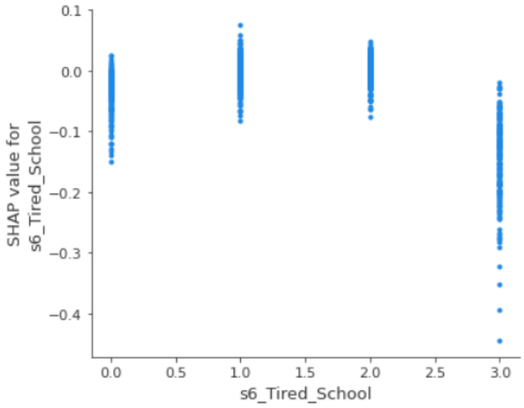


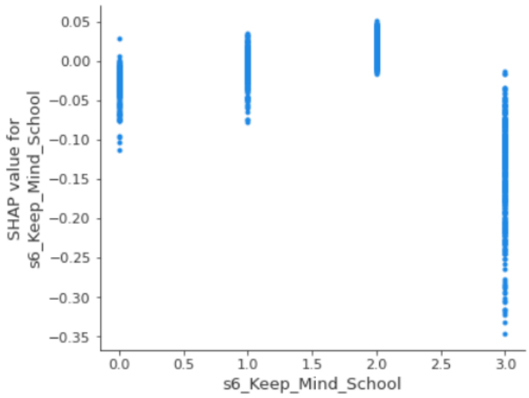
6.3. Combined Model – Other feature dependency graphs



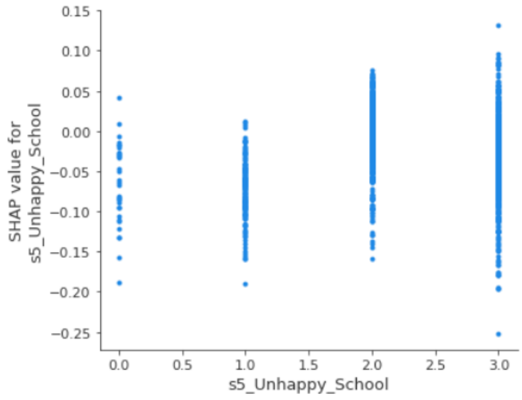


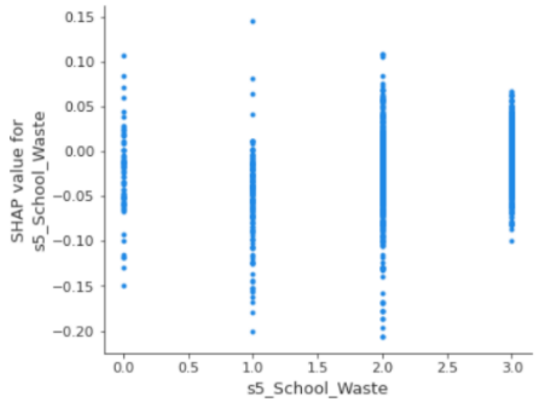


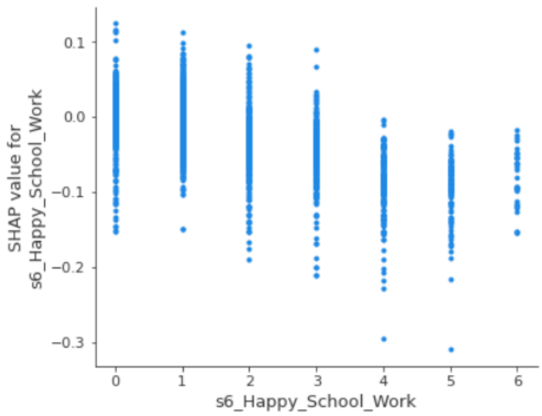


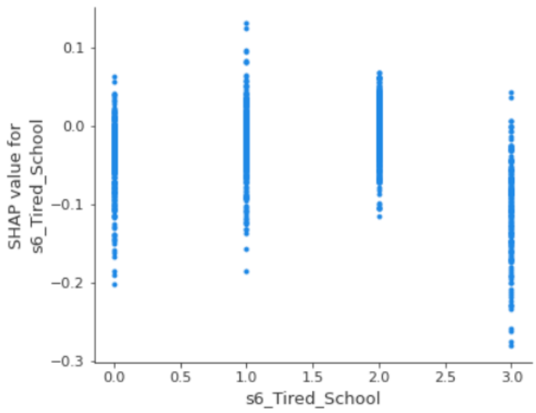


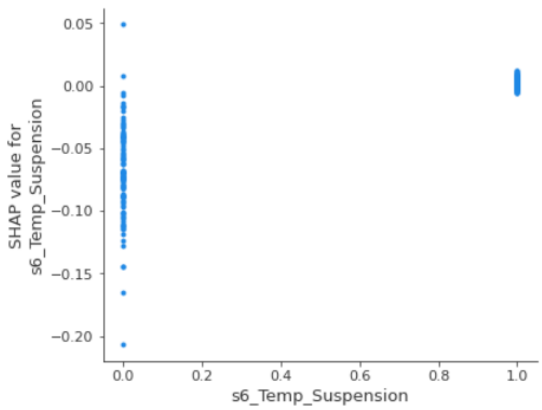
6.4. Male Model – Other feature dependency graphs

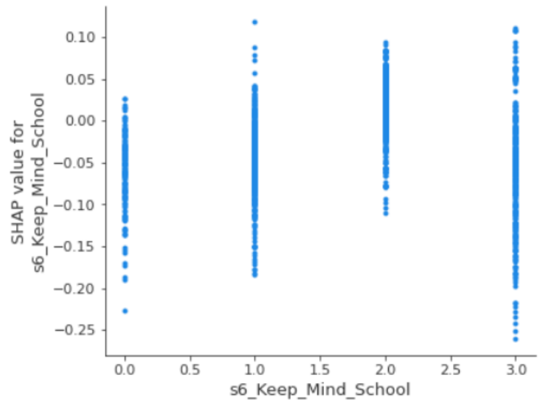


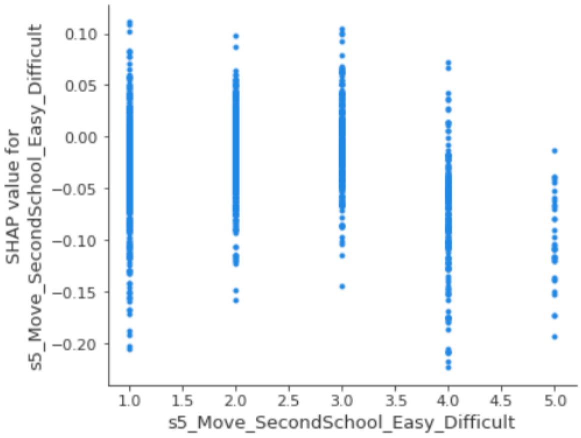




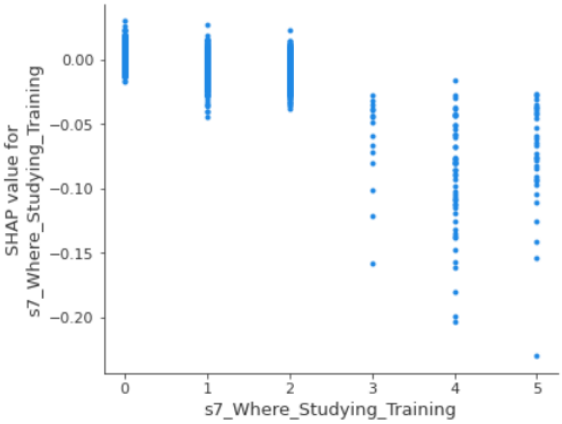


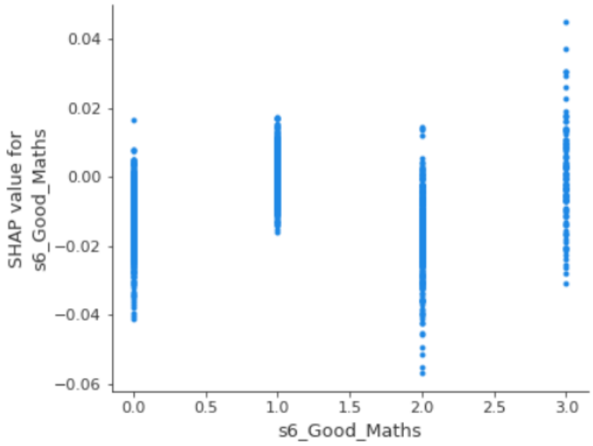


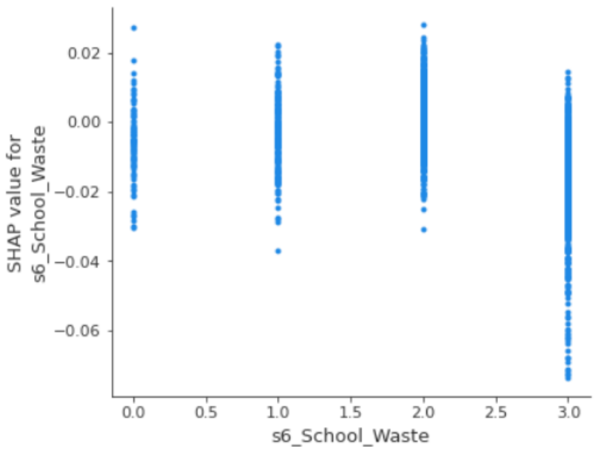


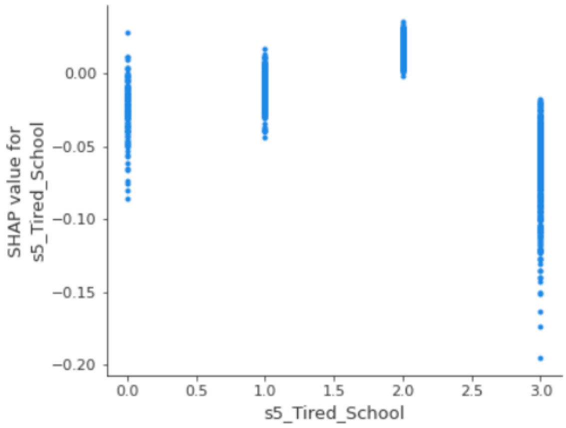


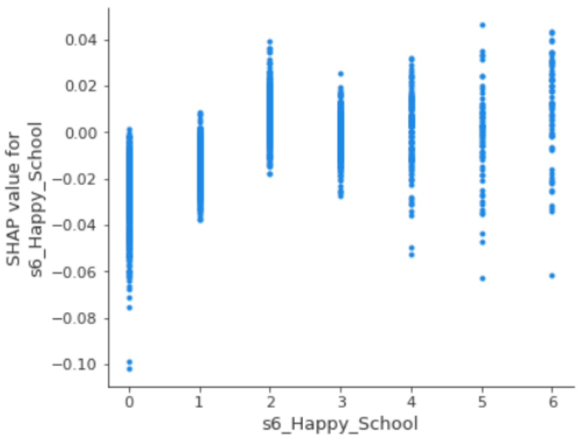
6.5. Female Model – Other feature dependency graphs

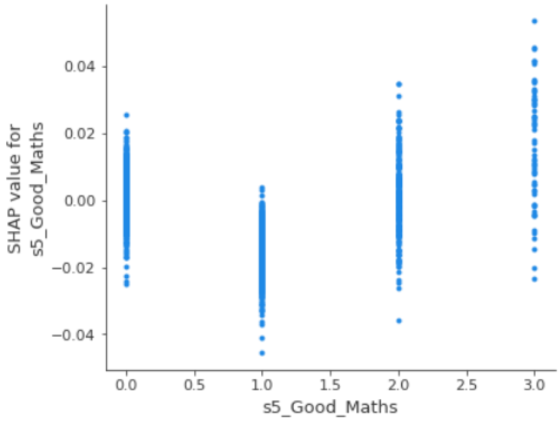


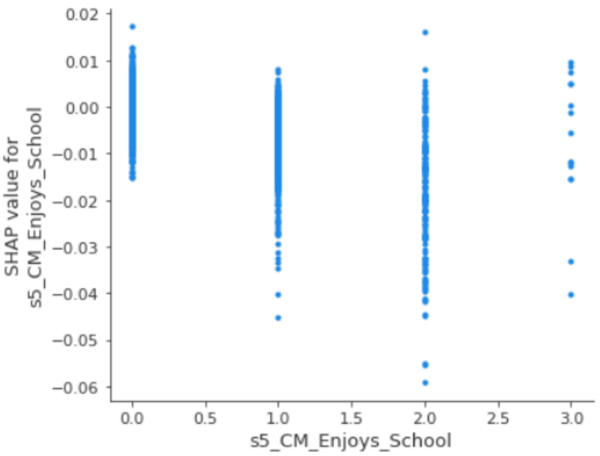


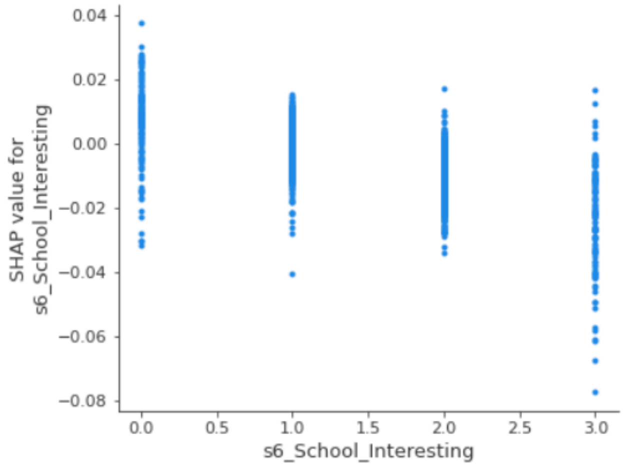


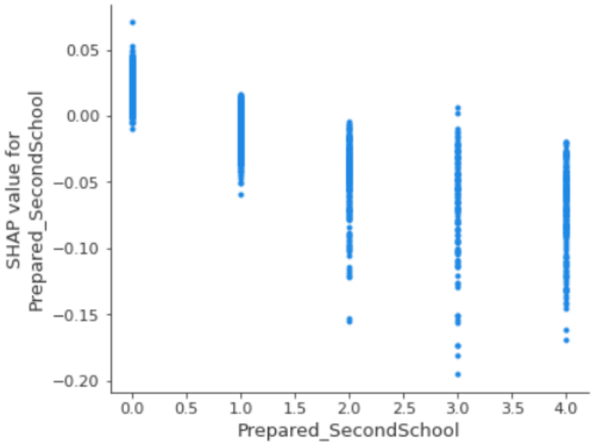


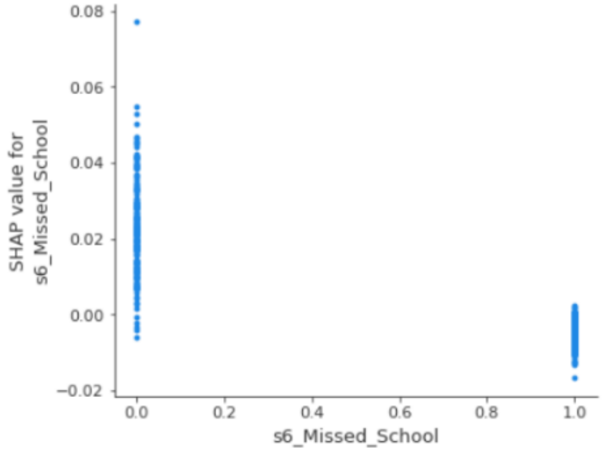





SHAP value for
s5_CM_Enjoys_School
0.02
0.01
0.00
−0.01
−0.02
−0.03
−0.04
−0.05
−0.06
0.0
0.5
1.0
1.5
2.0
2.5
3.0
s5_CM_Enjoys_School


SHAP value for
s6_School_Interesting
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.08
0.0
0.5
1.0
1.5
2.0
2.5
3.0
s6_School_Interesting


SHAP value for
Prepared_SecondSchool
0.05
0.00
−0.05
−0.10
−0.15
−0.20
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
Prepared_SecondSchool


SHAP value for
s6_Missed_School
0.08
0.06
0.04
0.02
0.00
−0.02
0.0
0.2
0.4
0.6
0.8
1.0
s6_Missed_School