A SURVEY ON THE EFFECTS OF COVID-19 ON THE EDUCATION, SOCIAL LIFE AND MENTAL HEALTH OF STUDENTS

Reza Fathurahman Sihab

Data Background

- The emergence of Corona Virus disease (COVID-19) has led the world to an unprecedented public health crisis.
- The closure of educational institutions during the COVID-19 pandemic has prompted a rapid transition from physical to digital learning. Online learning has emerged as a crucial alternative to conventional learning, providing students with access to education despite the restrictions on public movement.

Analysis Overview

- This research aims to provide an understanding of the learning situation during the COVID-19 pandemic, particularly in in Delhi National Capital Region (NCR) and other areas.
- Statistical data and visualization will be presented to provide an overview of the learning situation and student satisfaction with online classes during the pandemic.

Data Understanding

A cross-sectional survey is conducted 19 coloumns with a sample size of 1182 students of different age groups from different educational institutions in Delhi National Capital Region (NCR) and outside Delhi NCR.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1182 entries, 0 to 1181
Data columns (total 19 columns):
 #   Column                                                                  Non-Null Count  Dtype  
---  ------                                                                  --------------  -----  
 0   ID                                                                      1182 non-null   object 
 1   Region of residence                                                     1182 non-null   object 
 2   Age of Subject                                                          1182 non-null   int64  
 3   Time spent on Online Class                                              1182 non-null   float64
 4   Rating of Online Class experience                                       1158 non-null   object 
 5   Medium for online class                                                 1131 non-null   object 
 6   Time spent on self study                                                1182 non-null   float64
 7   Time spent on fitness                                                   1182 non-null   float64
 8   Time spent on sleep                                                     1182 non-null   float64
 9   Time spent on social media                                              1182 non-null   float64
 10  Prefered social media platform                                          1182 non-null   object 
 11  Time spent on TV                                                        1182 non-null   object 
 12  Number of meals per day                                                 1182 non-null   int64  
 13  Change in your weight                                                   1182 non-null   object 
 14  Health issue during lockdown                                            1182 non-null   object 
 15  Stress busters                                                          1182 non-null   object 
 16  Time utilized                                                           1182 non-null   object 
 17  Do you find yourself more connected with your family, close friends , relatives  ?  1182 non-null   object 
 18  What you miss the most                                                  1182 non-null   object 
dtypes: float64(5), int64(2), object(12)
memory usage: 175.6+ KB
```

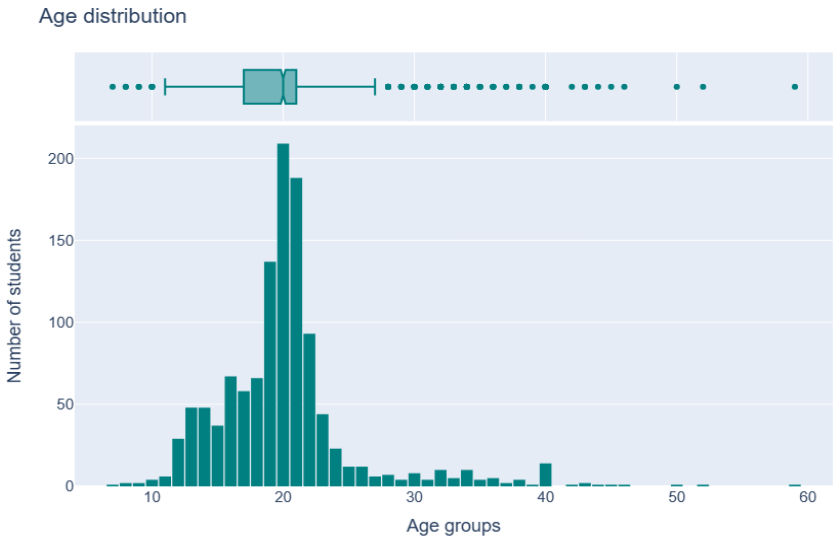
Data Preprocessing

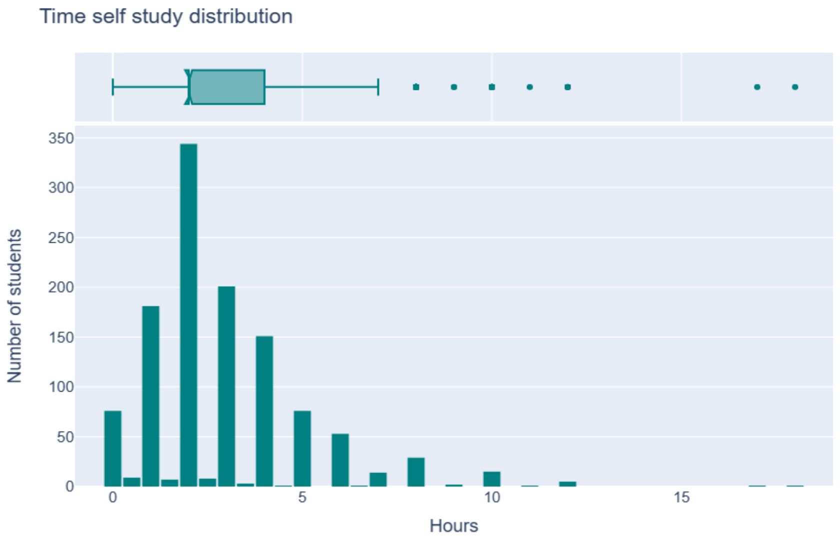
1. Handling missing values
2. Cleaning data
3. Handling duplicated values
4. Feature Engineering
5. Handling outliers

Missing Value

	feature	missing_value	percentage
0	medium	51	4.315
1	rating_online_class	24	2.030

Handling Outliers

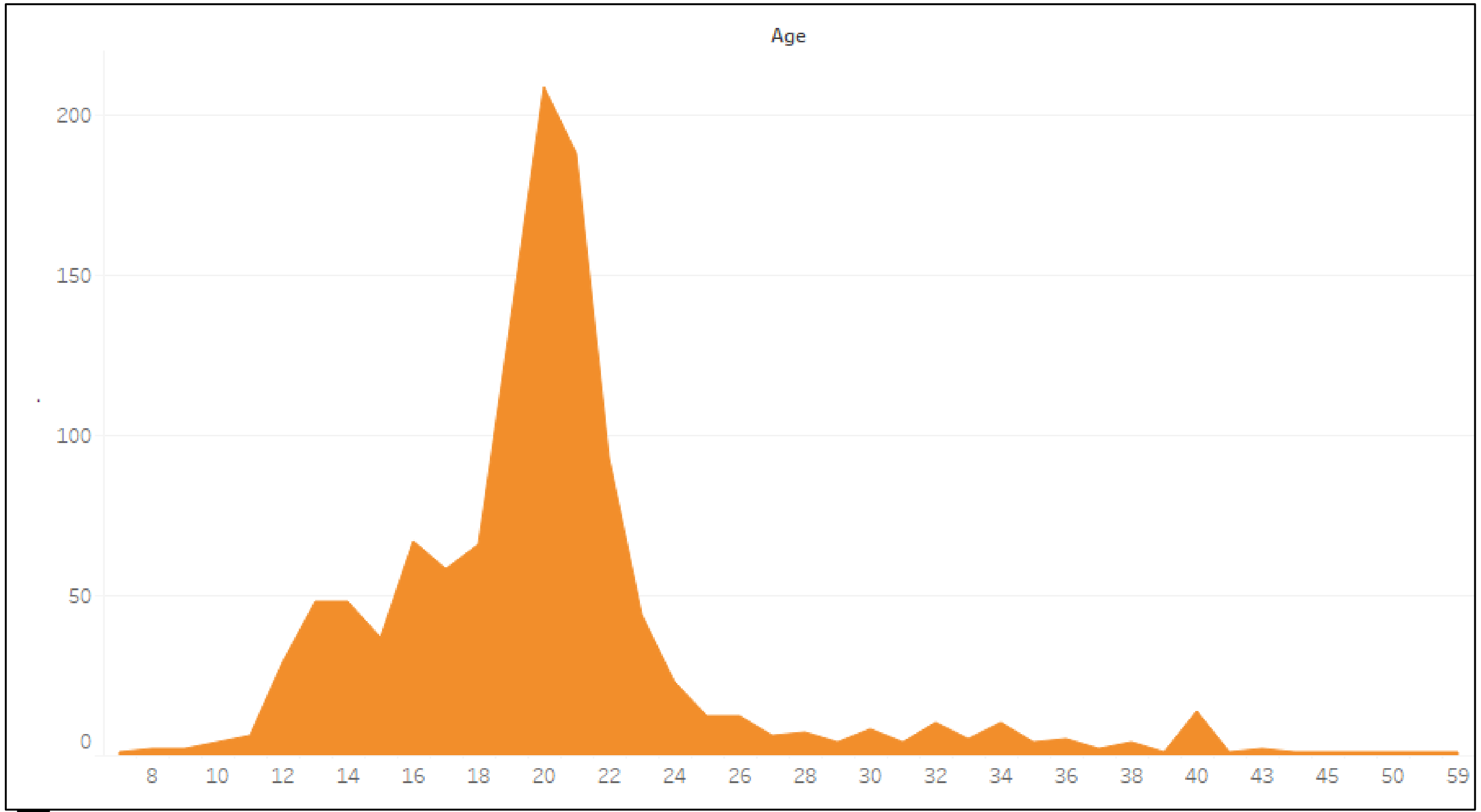


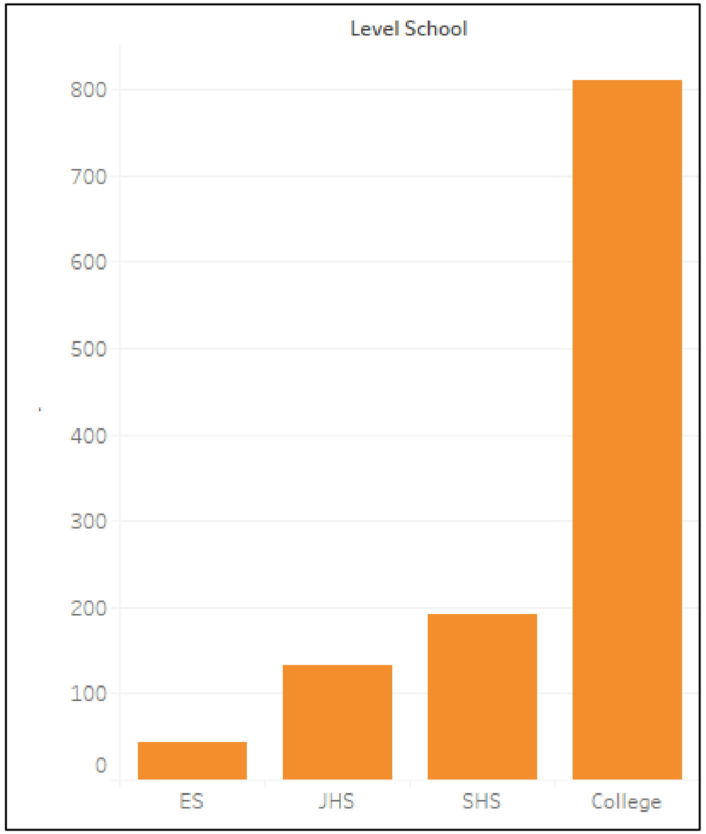


Explanatory Data Analysis and Visualizations

Distribution of Age

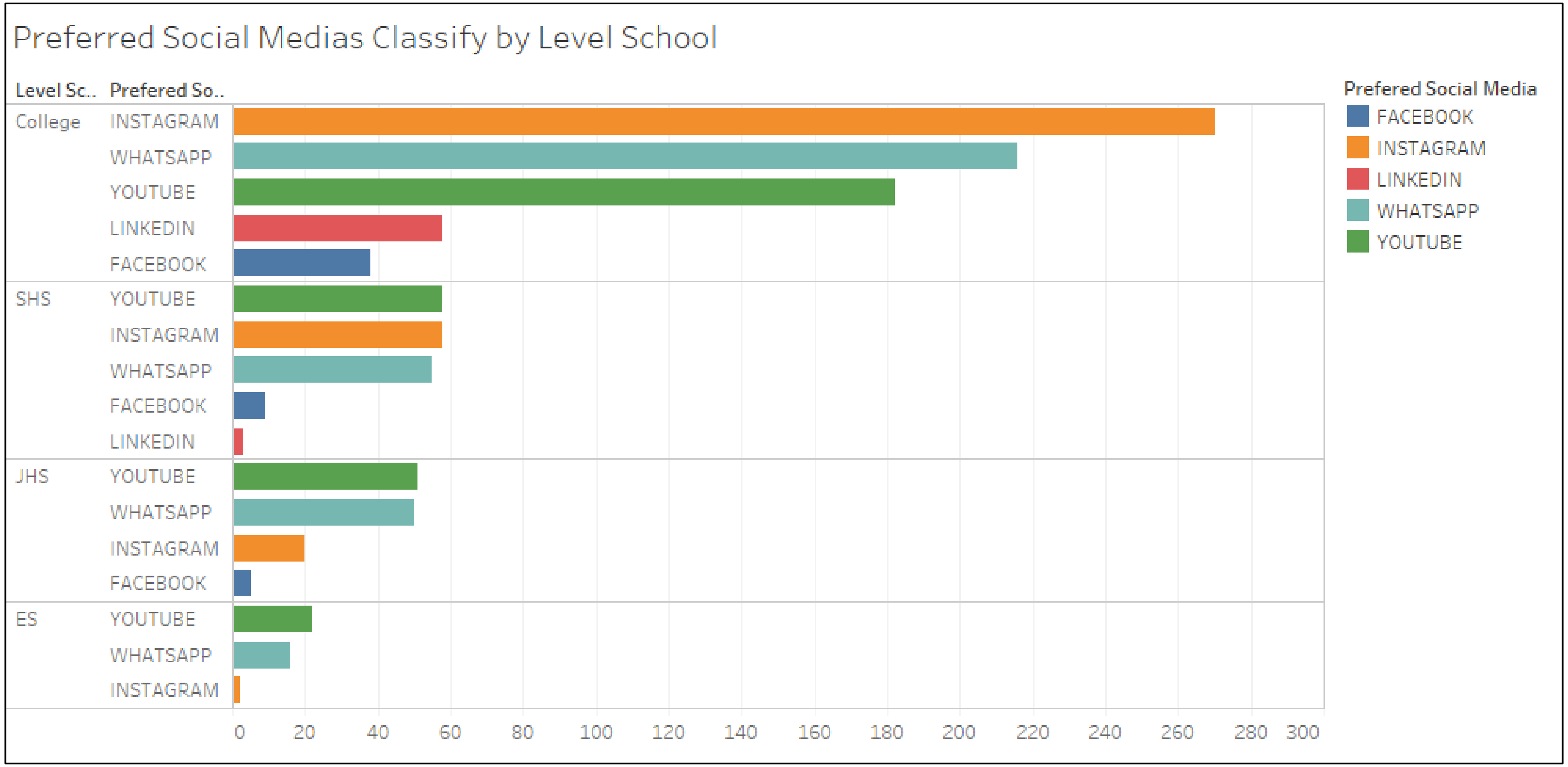
Performing feature engineering on the age column by classifying it into 4 categories, and it turns out that the samples are dominated by students with an age range of 18 years and above, approximately 811 out of 1180 samples.





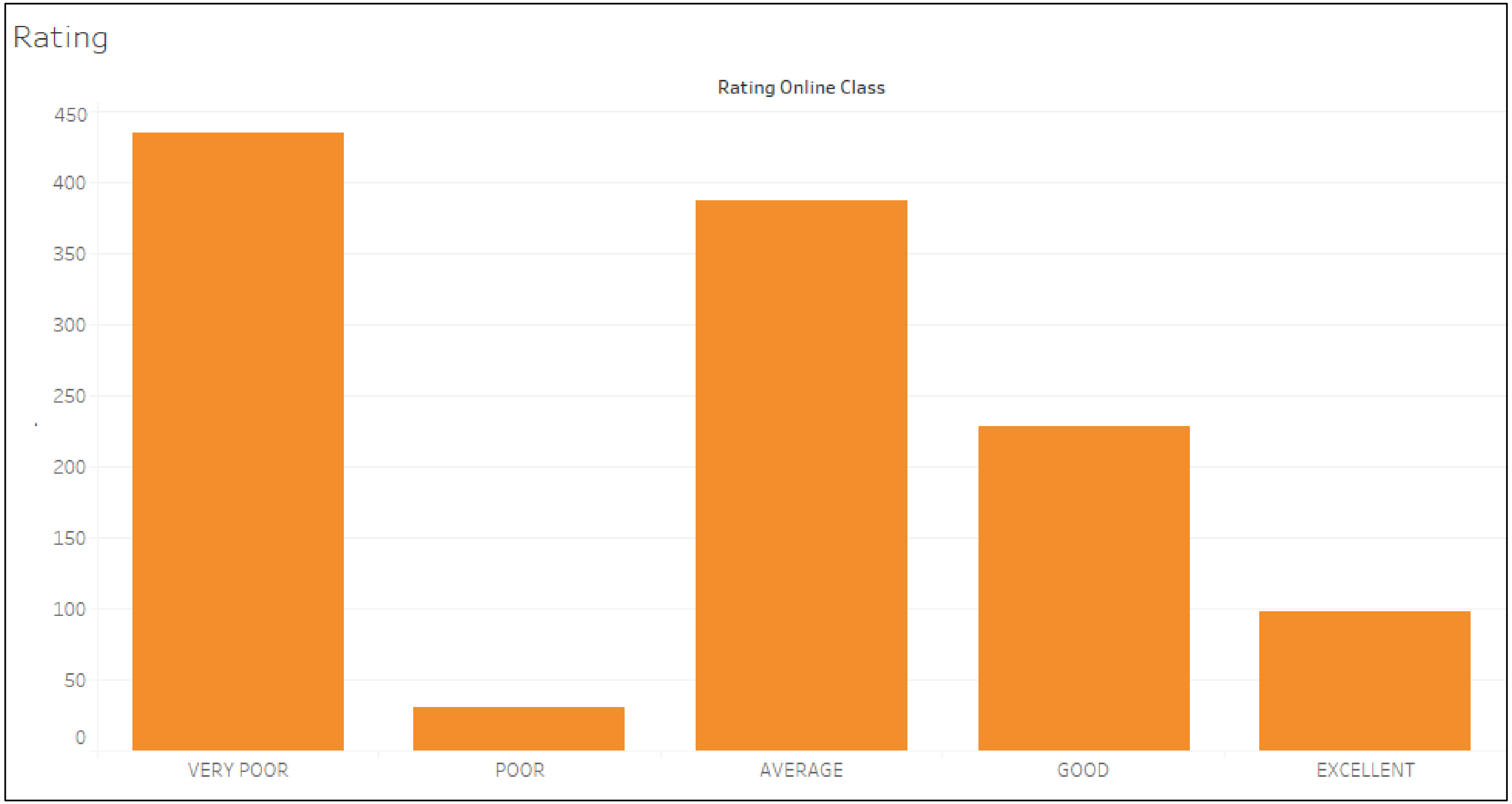
Preferred Social Media

Sorting the preferred social media graph based on Level school, it was found that the most visited social media are Instagram to Facebook. After analyzing based on the classification of educational levels, it is apparent that the prominent total users are on LinkedIn and Facebook.

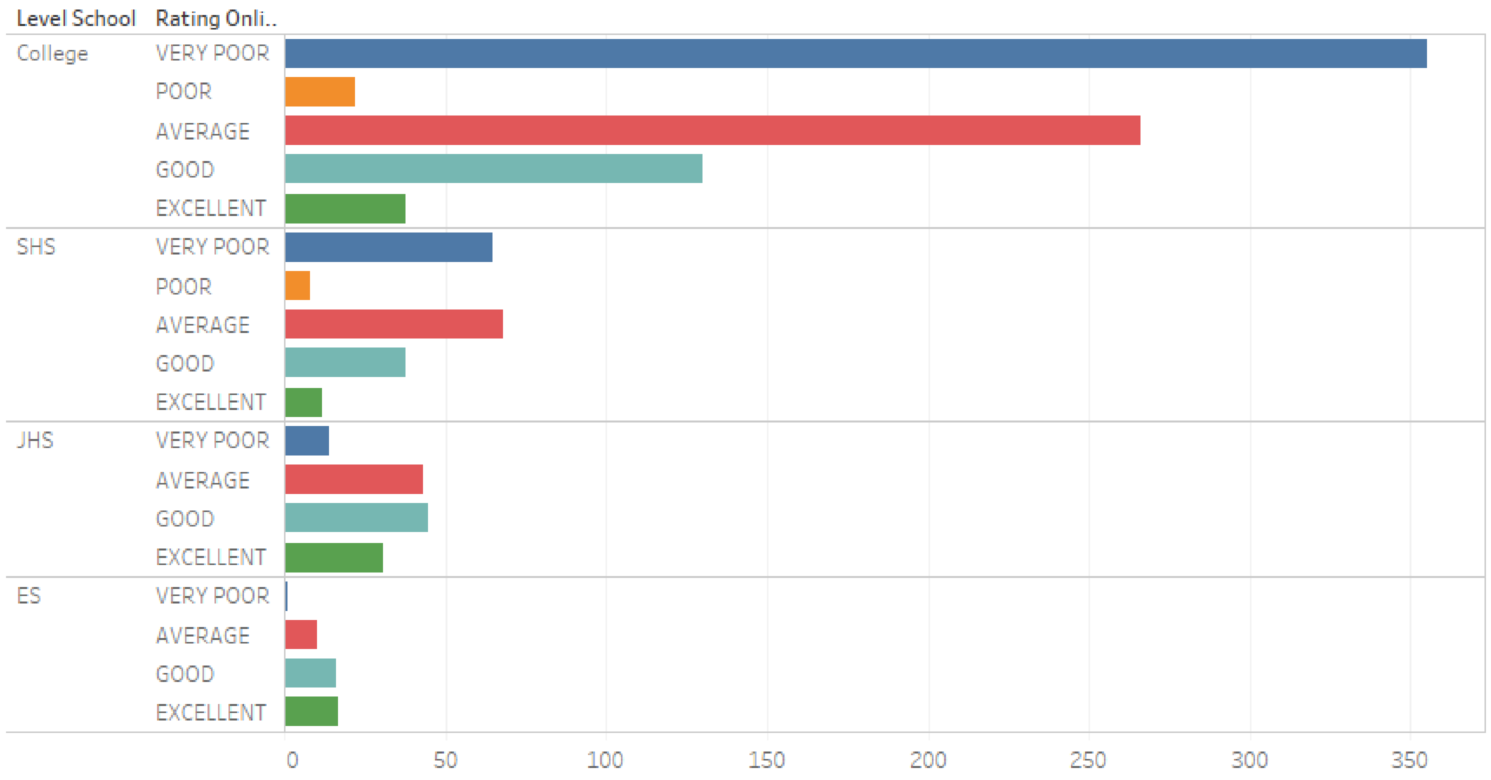


Rating Online Class

The majority of respondents gave a low rating for online classes. Meanwhile, the Very Poor and Poor ratings have a significant number compared to other ratings.



How do students rate their classes?
Does the level of school influence their ratings?



The higher the level of education, the more difficult it becomes for students to adapt to online learning. This may be due to several factors, such as the increasing complexity of the material, the growing need for physical interaction, and the inability to fully maximize the available technology for online learning.

Rating Online Class

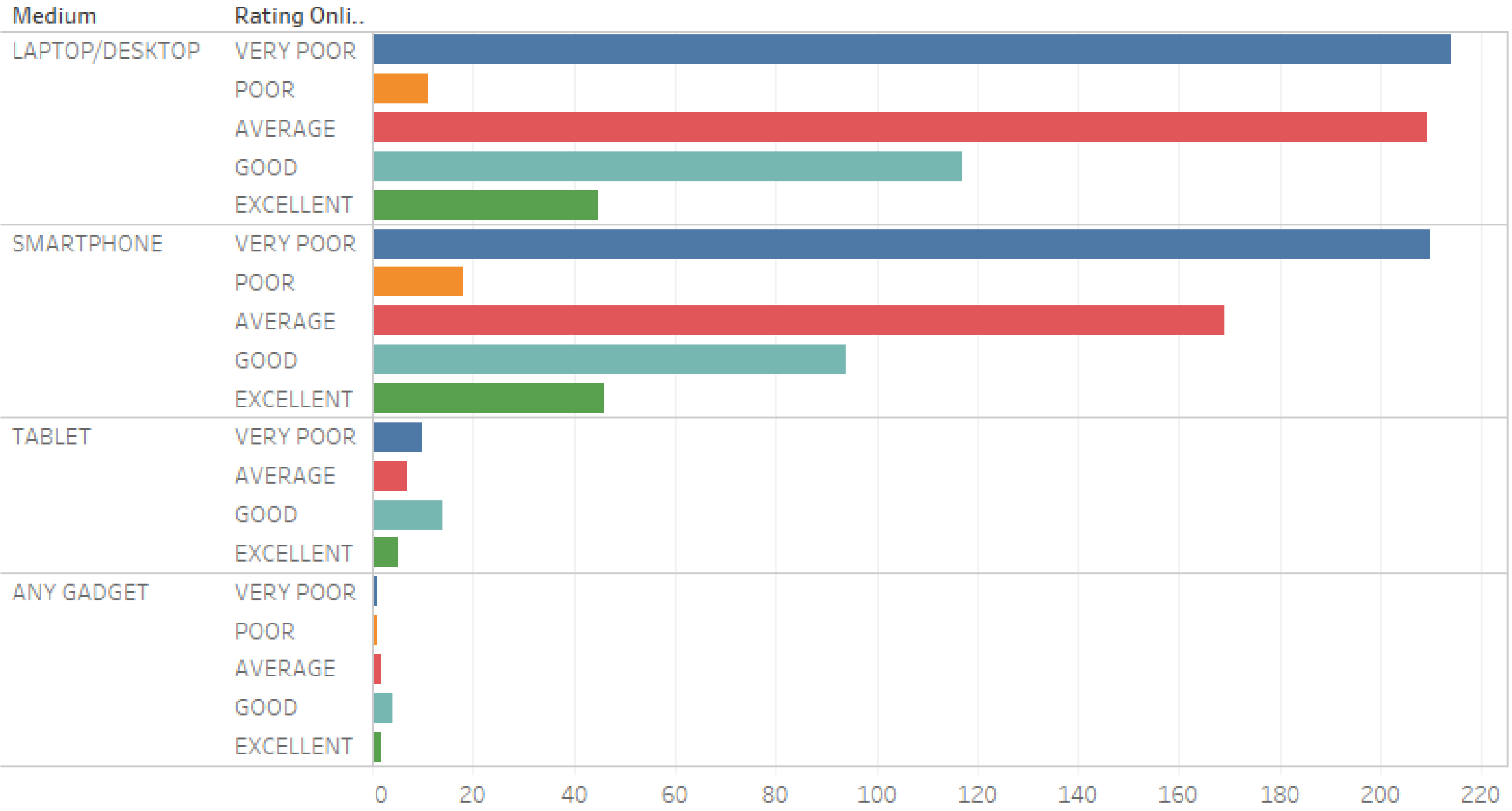
Level School	VERY POOR	POOR	AVERAGE	GOOD	EXCELLE..
College	355	22	266	130	38
SHS	65	8	68	38	12
JHS	14		43	45	31
ES	1		10	16	17

How do students rate their classes?
Does the medium influence their ratings?

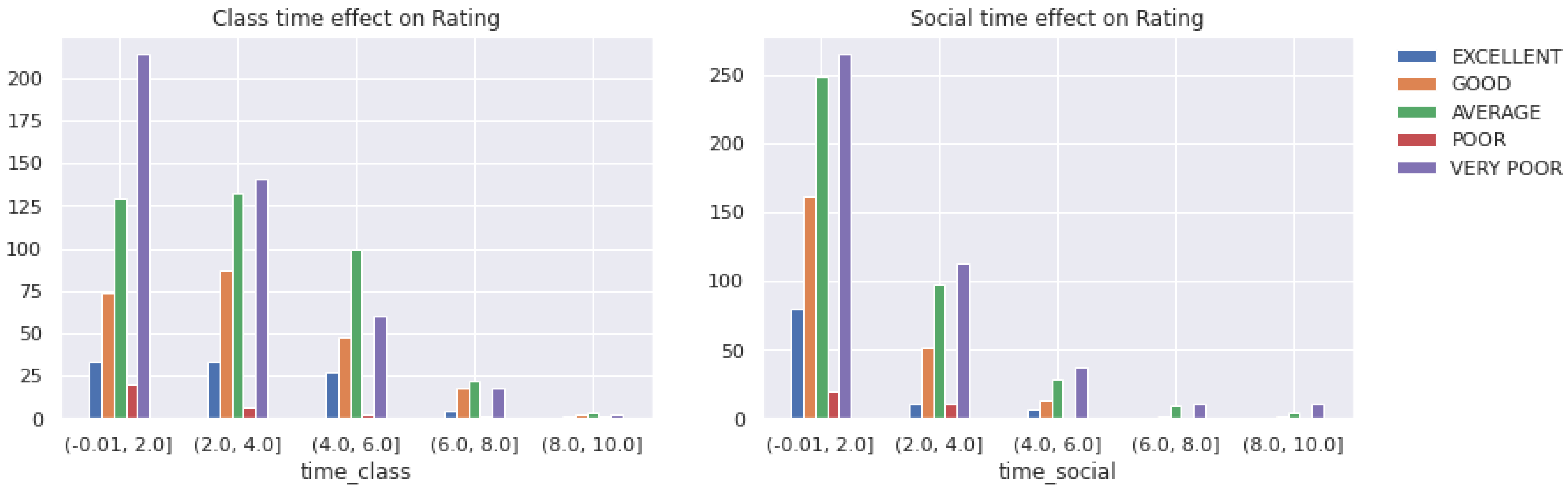
The percentage of online class ratings based on the gadget or medium used by students shows that those who have more classes and use desktop computers tend to give better ratings, possibly because it is easier to focus, take notes, and sit for longer periods while using a desktop, which leads to a better experience compared to using a smartphone or other devices.

Rating Online Class

Medium	VERY POOR	POOR	AVERAGE	GOOD	EXCELLE..
LAPTOP/DESKTOP	214	11	209	117	45
SMARTPHONE	210	18	169	94	46
TABLET	10		7	14	5
ANY GADGET	1	1	2	4	2



What affects the rating, whether it's students who have better class time or students who are more active on social media?



People who are less active on social media give mixed reviews, but generally, they also provide sufficiently good reviews. Therefore, most students who enjoy the class and give average or good reviews spend less time on social media, and the majority of votes for ratings come from them. They tend to give higher ratings. On the other hand, students who are more involved in learning tend to give better reviews. Additionally, it should be noted that people who spend more time on social media or less time in learning are fewer than those who spend less time on both sides.

Classification Model for Mental Health

Filltering Features

Out of the 19 columns in the dataset, only 11 features are considered influential towards student mental health. These features are:

- age
- time_online_class
- time_self_study
- time_fitness
- time_sleep
- time_social_media
- time_tv
- num_meals_per_day
- delta_weight
- delta_weight
- health_issue_in_lockdown

Classification Model for Mental Health

Model Performance Baseline

Model Performance Baseline

	Model	Recall	AUC	F1 Score	precision	accuracy
0	RandomForestClassifier	7.89%	52.68%	12.24%	27.27%	87.85%
1	DecisionTreeClassifier	34.21%	58.40%	24.53%	19.12%	77.40%
2	LogisticRegression	0.00%	50.00%	0.00%	0.00%	89.27%
3	XGBClassifier	15.79%	54.41%	18.18%	21.43%	84.75%
4	GradientBoostingClassifier	2.63%	49.89%	4.17%	10.00%	87.01%
5	LGBMClassifier	10.53%	51.94%	12.70%	16.00%	84.46%
6	ExtraTreesClassifier	7.89%	52.52%	12.00%	25.00%	87.57%
7	HistGradientBoostingClassifier	15.79%	54.57%	18.46%	22.22%	85.03%

Classification Model for Mental Health

Model Performance Oversampling

Model Performance Oversampling

	Model	Recall	AUC	F1 Score	precision	accuracy
0	RandomForestClassifier	18.42%	53.99%	17.95%	17.50%	81.92%
1	DecisionTreeClassifier	18.42%	49.72%	13.33%	10.45%	74.29%
2	LogisticRegression	55.26%	59.12%	23.86%	15.22%	62.15%
3	XGBClassifier	23.68%	56.15%	21.69%	20.00%	81.64%
4	GradientBoostingClassifier	42.11%	56.81%	22.22%	15.09%	68.36%
5	LGBMClassifier	21.05%	53.09%	17.20%	14.55%	78.25%
6	ExtraTreesClassifier	5.26%	51.05%	8.00%	16.67%	87.01%
7	HistGradientBoostingClassifier	23.68%	54.72%	19.57%	16.67%	79.10%

Classification Model for Mental Health

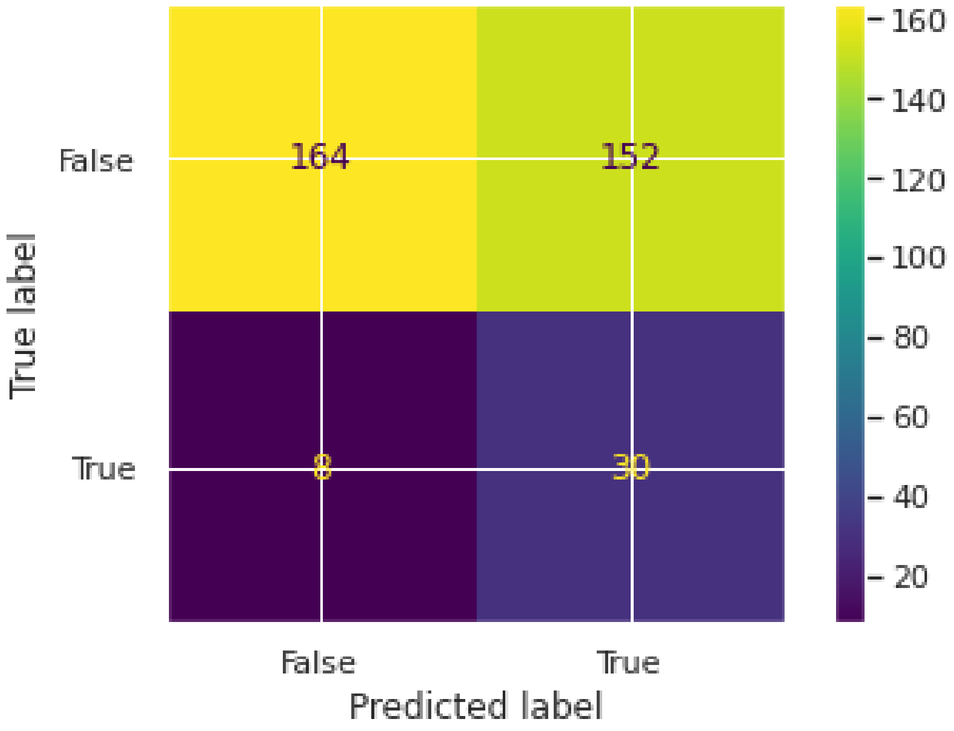
Model Performance Undersampling

Model Performance Undersampling

	Model	Recall	AUC	F1 Score	precision	accuracy
0	RandomForestClassifier	78.95%	65.42%	27.27%	16.48%	54.80%
1	DecisionTreeClassifier	65.79%	59.32%	23.58%	14.37%	54.24%
2	LogisticRegression	57.89%	57.27%	22.34%	13.84%	56.78%
3	XGBClassifier	65.79%	60.27%	24.27%	14.88%	55.93%
4	GradientBoostingClassifier	63.16%	58.95%	23.41%	14.37%	55.65%
5	LGBMClassifier	60.53%	57.00%	22.12%	13.53%	54.24%
6	ExtraTreesClassifier	76.32%	64.42%	26.73%	16.20%	55.08%
7	HistGradientBoostingClassifier	73.68%	63.74%	26.42%	16.09%	55.93%

Classification Model for Mental Health

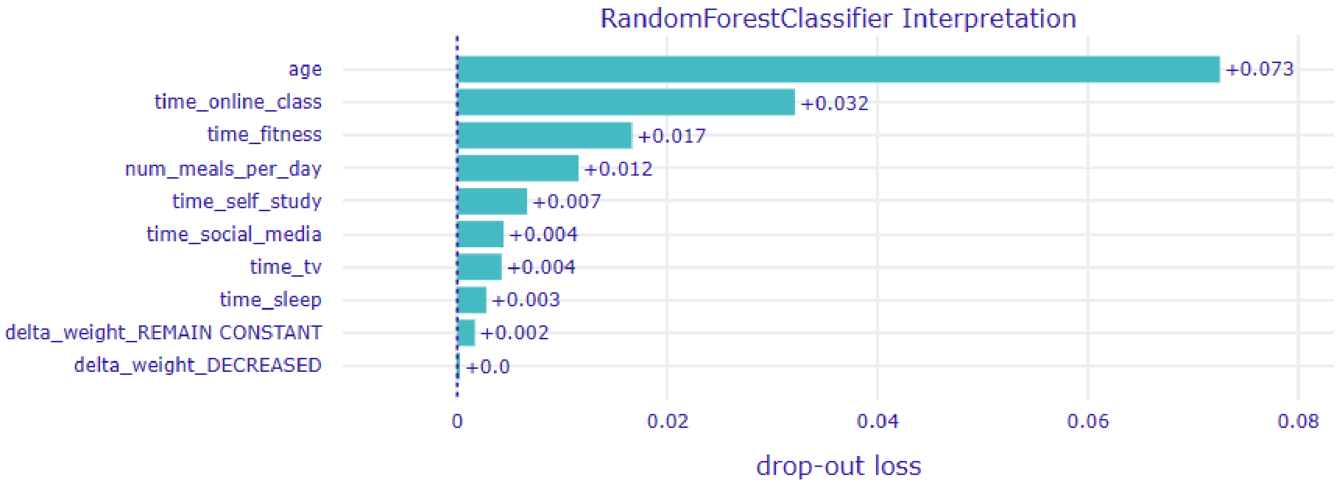
Best Model



Explanation:

- True Negative (TN) = 164
- False Positive (FP) = 152
- False Negative (FN) = 3
- True Positive (TP) = 30

Variable Importance



Model Performance Undersampling

Model	Recall	AUC	F1 Score	precision	accuracy
RandomForestClassifier	78.95%	65.42%	27.27%	16.48%	54.80%

Conclusions

1. The time spent by students for online classes is not in line with what they are supposed to do.
2. Limited class interaction and inefficient schedules significantly affect the satisfaction level of students.
3. Based on the analysis, peer impact in the school environment motivates individuals to work hard and learn social skills.
4. The biggest challenge for online learning is the requirement for efficient digital infrastructure and digital skills for both students and teachers.

Recommendations

Based on the results of data analysis and visualization, it is important for schools, colleges and governments to continue working together to improve online learning systems and methods in order to provide a better learning experience for students by:

1. Optimizing the use of technology and online learning platforms
2. Providing training and support to teachers and students in dealing with online learning
3. Creating a more interactive and participatory learning environment
4. Accommodating students' learning needs at home
5. Due to the COVID-19 situation, many students are likely to suffer from stress, anxiety, and depression, so emotional support should be provided to students.

Thank You

Credits:

- Kaggle
- Slidesgo
- Freepik

Rezafsihab98@gmail.com

https://www.linkedin.com/in/rezafsihab/

Source:

- Dataset: https://www.kaggle.com/datasets/kunal28chaturvedi/covid19-and-its-impact-on-students
- Journal: https://www.sciencedirect.com/science/article/pii/S019074092032288X