Don't Panic, Be Happy - Introspecting the Happiness of University Students.

Pooja Priyadarshini[1] and Subhankar Mishra[2]: 0000-0002-9910-7291

[1] Utkal University, Bhubaneswar-751004, India. pooja.wit1@gmail.com
[2] National Institute of Science Education and Research, Bhubaneswar-752050, India.
Homi Bhabha National Institute,
Anushaktinagar, Mumbai - 400094, India. {smishra@niser.ac.in

Abstract. Happiness underlines the intuitive constructs of a specified population based on the positive psychological outcomes. It is the cornerstone of the cognitive skills covering a major research area in Economics and Psychology. Exploring university student's happiness and its determinants has been the essence of the researchers lately. In this study, we have analyzed the university student's happiness and its facets using statistical distribution charts; designing research questions. Furthermore, regression analysis, machine learning and clustering algorithms were applied on world happiness dataset and university students dataset for training and testing respectively. Philosophy was the happiest department while Sociology the saddest; average happiness score being 2.8 and 2.44 respectively. Pearson coefficient of correlation was 0.74 for Health. Predicted happiness score=5.2, MSE=0.26, RMSE=0.51, goodness of model fit=51%, train and test error being 0.52, 0.47 respectively. On a Confidence Interval(CI) of 5% p-value was least for Campus Enviroment(CE) and University Reputation(UR) and maximum for Extracurricular Activities(ECA) and Work Balance(WB) (i.e. 0.184 and 0.228 respectively). RF+Clustering got the highest accuracy(89%) and Fscore(0.98) and the least error(17.91%), hence turned out to be best for our study.

Keywords: Happiness· Student's Happiness· Machine Learning· clustering.

1 Introduction

"Happiness is not a state to arrive at, but a manner of travelling" - Margaret Lee Runbeck.

Happiness is a floating or a fluctuant concept that has been explicated by the researchers(Economics and Psycology) throughout the ages. The present-day psychology - Affirmative Psychology has pinpointed the psychological health and well being since late 1990's[1]. Ed Diener[2], being one of the pioneer in the area of psychological research has set forth that human's well being is closely associated with happiness in context with modern psychology. Happiness is tagged

with manifold outcomes namely, broadened attention, efficiency, innovations, social factors, health[5]. During the past decades, researchers have surveyed several facets of happiness namely, income, economic growth, escalation, human development index, institutions, expenditure, unemployment and globalization[6]. Ahn et al.[7] work threw light on the national well being of people on the basis of GDP and highlighted other parameters for attaining happiness specifically, physical and mental health, good governance, crime, family and social relationships, economy, corruption etc. Using such factors we can precisely answer a question, "What is the happiness score of a country?" or "Which is the happiest country based on happiness ranks?" or "How is the country performing?"

The problem statement of the study can be picked out as follows. Firstly, despite the exponential growth in research on happiness and the relationships amongst the factors, scholars are finding it difficult to define the term happiness exactly[8]. In context with expressiveness, appraisals and lifestyles happiness still remains a fuzzy terminology with existing debates amongst researchers. The vocabularies namely, happiness, psychological sentiments, well being and inter-disciplinary satisfaction are closely inter-related and prejudiced in nature, thereby complicating the results of empirical analysis[9]. Secondly, the ample happiness dataset is available for research, manual calculation is extremely troublesome. So, machine learning algorithms are used along with statistics. Thirdly, in this fast-paced and technology-driven world, the inter-personal relationships has deteriorated drastically. People lack time to ask and answer questions like "How was your vacation?" or "How happy are you with the college life?" or "How satisfied are you with your work life?" or "Do you like the current government?"[10]

One of the motivation to conduct this study was to overcome the research gap between general public and the students[11]. Student's happiness is directly proportional to university rankings as well. University student's happiness can be enumerated by the criterion namely, age, gender, university reputation, time management, work balance etc[12]. We get to see the world happiness rankings every year through the Gallup World polls highlighting the overall world population. However, education linked with happiness has caught the attention of researchers lately.

Happiness denotes the transient period of pragmatic responses that can be boosted by watching funny episodes, receiving BRICS award twice, completion of a degree or turning COVID negative. Theoretically, in context with the literature happiness has covered an encyclopedic domain. Diener et al.[3] describes happiness as a controversial term. For instance, Sonja Lyumbomirsky[13] have classified happiness as an immeasurable domain that includes the knowledge of cheerfulness, delight, ecstasy, fortune and prosperity tagged with a purposeful life. Veenhoven[14] proposed the impressionistic definition of happiness i.e., the satisfaction with one's own life. However, Kahneman and Riis[15], referred happiness as a narrow-mined euphoria or a short-lived episode that a person experiences. Veenhoven[14], contemplates happiness as the gratification of one's own life on a whole. Furthermore, Deci and Ryan[9] have italicized the self- in-

dulgence principle of happiness so as to solicit the pleasing experiences and keep off the distressing ones. Kahneman and Jason[16] have proposed the dual angle to compute happiness. Firstly, the "impersonal happiness" can be answered via questionnaires like "How happy are you at this place?" or "How was your interview?" Secondly, questionnaires designed to rate happiness "How was your college life ?" or "What is the feedback regarding a professor?". This assess the intuitiveness of the respondents. The Satisfaction with Life Scale(SWLS)[17] and Subjective Happiness Scale(SHS)[18] are the vital tools needed for the same.

In our study, we had collected 280 university students' reviews in context with happiness with 11 fields for conducting empirical analysis from the students via local survey, department wise in Utkal university, India . We incorporate a step-by-step approach to perceive the contribution of our paper :

- Several Research Questions(RQ) encircling happiness were designed; evaluated on the university students dataset; answered through visualization using statistical charts and make intuitive judgment.
- Using basics of statistics to find the mean, Standard Errors(SE), coefficients, Mean Squared Error(MSE) and Root Mean Squared Error(RMSE), Standard Deviation(SD), F-statistics and p-value of our survey data from Utkal University, India and comparing them with outcomes of study of Finland[19] and Australia[20].
- Linear regression and multiple regression was applied on the 2019 world happiness corpus and using students happiness dataset for testing. Also happiness facets namely, Freedom, Social and Health from world dataset were used to predict university students' happiness score. Machine learning(Linear Regression, KNN, Random Forest(RF), Support Vector Machines(SVM))and clustering algorithms were used for empirical analysis and compute performance(Fscore and accuracy). Hence, comparison graphs were picturized.

This paper has been drafted into five sections. The study begins with a brief introduction presented in Section 1. Section 2 underlines the background of happiness. Section 3 organizes the techniques utilized in the paper. Section 4 tells about the empirical analysis and results obtained. Finally, Section 5 pinpoints the conclusions of the work and upcoming developments.

2 Related Works

1. In context with happiness of university students, Frey and Stuzer[10] argued contribution of age on happiness was negligible however, Esa[19] found age to be an important explanatory variable. Their studies revealed older Finns are less happy as compared to youngsters. The results seemed more interesting when an age group of 18 and 28 year old respondents were taken under consideration. In an another instance of age group, Blanchflower and Oswald's[12] research studies revealed that youngsters of US and Europe were dissatisfied and seemed unjustified due to growth in education. Their statistics explained that the European students perusing higher education after 18 years and below 30 years were rewarded at max in lieu of happiness.

2. In the study of Esa[19] showed significant results when happiness results of the countries namely, Finland, Sweden, Norway, Spain, Germany, Mexico were combined; but the results were found out to be negative. Clark and Oswald[21] reached to similar inferences. People with higher education were dissatisfied with work. They were to explain whether the university students are happy or not?
3. In the research of Diener et al.and Argyle[4] there was a minor association between education and happiness on the basis of top level of qualification attained. The pleasure from education ensues from work, earnings, occupations. Hartog and Oosterbeek[22] revealed schooling period was happy and satisfying as compared to happiness at the university.
4. Diener et al.[2] devised the SWLS to rate critical facets of prosperity and happiness. For instance, reviews like "I like my job culture and I am satisfied with the income as well." are answered on a 7-point Likert scale(1 denotes strongly disagree; 7 denotes strongly agree). This scale has caught attention amongst all age groups and circumstances and proved to be justifiable and authentic estimation.[17] Lyumbomirsky and Lepper proposed another scale of ratings namely, SHS[18]. It denotes a 4 item scale to identify global delight.
5. Young and McIntyre[23] had proposed a comparison of happiness scores of 'Piers-Harris Children' Self concept Scale' amongst the gifted students studying in 8 grade and its distinguished fellow mates in US. A "yes" or "no" response was captured by the respondents and mean(6.68), SD(2.11), t(78)=0.98, p=0.33, d-0.22, was computed on 95% confidence interval. In the study of Ash and Huebner conducted amongst the US students studying in grades 6-8, included 61 gifted and 61 distinguished students[24].

The exemplary results of Esa [19] and Chan et al[20] was one of the strongest reason to follow them for our study. The idea was to analyze happiness amongst the students of Jyvaskyla University, Finland and Australia respectively. The survey data so obtained comprised of components influencing happiness and delight. They inferred student's level of joy was highly influenced by WB, TM and RF and raised issues on Satisfaction. They suggested to extend the study to other countries and introspect whether the factors influencing the happiness of Australian and Finnish students holds goods for data crawled from across the globe. Moshe[25] had several unanswered research questions encircling happiness of gifted students to be addressed in future. Table-1, Table-2 shows the comparative study of existing literature.

3 Methodology

3.1 Data Sources:

Fig.1 explains the architectural design of our study.

In our study, we used two datasets. Firstly, the corpus, World happiness-2019, was crawled in .csv format with 200 records. It contained 9 columns namely, Overall rank, Country or Region, Score, GDP, Freedom, Corruption, Healthy

Table 1: Basic descriptive statistics of Respondents

Parameters	Chan[20]	Esa[19]	Our Study
Sample size	749	246	280
Female% of total	50.6	53.9	63
Directly from high school% of total	74.8	27.2	-
Part-time employment% of total	48.1	35.4	-
Average working hours(of above),h	11.2	10.5	-

Table 2: Students satisfaction with university life: responses to statement 'Overall, I am happy with my university life'

Parameters	Chan[20]	Esa[19]	Our Study
Strongly agree(or Very Happy)	13.9	22.8	47
Agree(or Happy)	54.6	63.8	34.7
Neutral	23.1	8.5	12.2
Disagree(or Sad)	7.2	4.9	6.1
Strongly disagree(or Very Sad)	1.2	0.0	-

life, Social Support and Generosity. Secondly, our line of action was to highlight the happiness of Utkal university(India) students and rank the Department wise happiness based on the survey data. Hence, we piled up real-time dataset from the university students department-wise(namely, Analytical and Applied Economics, Geography, Odia, Sociology etc.). The survey was conducted for a month and responses were captured on Google form. It comprised of 10 questionnaires in context with happiness in order to understand the student's sentiments. They were asked to assign their responses on the designed 4-point Likert scale (1-very happy, 2-Happy, 3-Neutral, 4-Sad). We collected an adequate sample size i.e.,300 from amongst the university students for a comparative research study. The survey conducted was entirely non-compulsory in nature and no baits were offered to gather data. The fields of the survey were Department, UR, CoE, CE, GSR, RF, TM, WB, ECA, Gender, Age Group. Data cleaning and other preprocessing schemes(eliminating lost data, discarding NA values) were initiated on our dataset to enhance the overall efficiency of our model. Post cleaning, we obtained 157 training and 280 test records in the data set.

3.2 Quantitative Measures

Likert Scale Researchers employ this in a survey to trap the responses of respondents in a close ended fashion to answer the questionnaires and generate useful insights. Respondents gives their level of pleasure or displeasure on a symmetric scale(from 1-5 or 1-7 (1-very happy, 2-happy, 3-neutral, 4-sad, 5-very sad)) of agreement or disagreement.

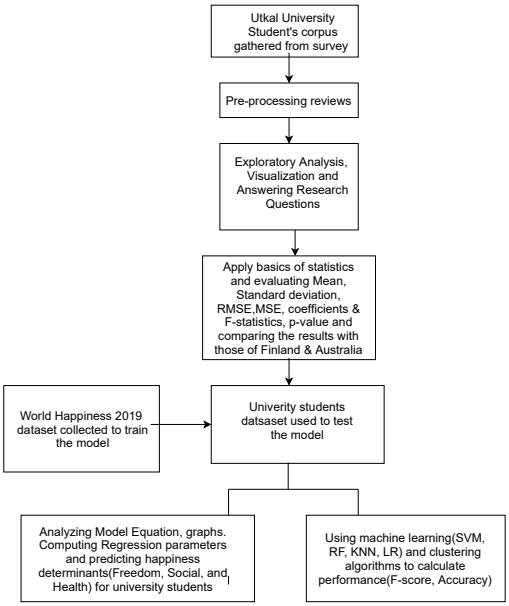


Fig. 1: Architectural Design of Student's Happiness.

Performance Metrics

$$F - score = \frac{2 * precision * recall}{precision + recall} \tag{1}$$

$$accuracy = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative} \tag{2}$$

3.3 Algorithms

Regression Analysis is the key highlight of our study and highly emphasized, and is denoted by:

$$Y_i = f(X_i, \beta) + e_i \tag{3}$$

$$f(X_i, \beta) = \beta_0 + \beta_i . X_i. \tag{4}$$

In order to evaluate the best-fit line for every explanatory variable, Multiple linear regression follows the given steps: 1.The regression coefficients for the model is computed by finding out the least model error. 2.The t-statistic for the entire model. 3.The corresponding p-value for the model. 4.The t-statistic and p-value for every regression coefficient is taken into account for the model.

Machine learning algorithms are used to make predictions and classifications from the abundant dataset available. For instance: **SVM** plots the data to an N(no. of features) dimensional space and a hyperplane identifies such points during regression and classification. **RF** utilizes decision trees for training the dataset based on certain attributes. **KNN** classifies the dataset based on the no. of votes obtained from k nearest neighbors. Aggregation of values from k such neighbours is used as output. **Clustering**, an unsupervised learning technique, deals with large, complex, unlabelled dataset. K-means and agglomerative methods were used for our study.

4 Experiments and Results

The three-fold experimental demonstrations so performed included - The exploratory analysis of the survey data and answering the RQ's through statistical distribution graphs. The visualization aspect,listed in Table-3, showcased the cognitive skills of students incorporated with fundamentals of statistics. Secondly, a statistical analysis was done on the survey dataset using STATA. The statistical outcomes(RMSE, MSE, Test and Train Error, Mean, SD, Coefficient, F-statistics, p-value etc.) so obtained were compared with those from the study of work of Chan[20] and Esa[19] from university of Australia and Finland. Thirdly, Linear Regression and Multiple Regression was applied on the world happiness-2019 dataset for training the model. For testing we used our survey data of university students. Particularly, we predicted the happiness score for the happiness factors(Freedom, Social, Health) for our test data. Model equations were designed along with computation of regression parameters. Ranking amongst the department was highlighted based on happiness of students. We used scatterplots and modeling plots for making inferences. Then the performance of model was tested using classifiers(LR, KNN, SVM, RF) and clustering algorithms. The Python code for experiment is available at "https://github.com/smlab-niser/2020happiness".

4.1 Discussions

– Philosophy was the happiest department and Sociology the saddest amongst all other and departments. Their average happiness scores resulting in 2.85 and 2.44 respectively. Sanskrit had a negligible impact on student's happiness with a minor variation. School of Women Studies, Odia and Law had a wide fluctuation in happiness score. The gender distribution of our study was balanced with male and female participation accounting to 37% and 63%. Age groups was divided in 2 slots of 25-30 and 30-35 years respectively in both the genders. SD resulting in 2.82. Mean happiness for two slots turned up to 20 and 22.97. Respondents were basically young(25-30 years) studying in first or second year post-graduation department. 27% were part time workers. This hints at they need to have a balanced work and university life leaving meagre time for ECA. Females should a higher interest in survey

as compared to male counterparts. Furthermore, part-time employment and their working hours were excluded from our study.

- In our study, Department of Biotechnology and Philosophy(33% of the total) captured the maximum respondents; while Chemistry the least(15% of the total). Overall based on all departments, happiness factors, UR had an extreme variation in participation percentage(29.6%). RF captured the maximum ratio of participation(31.6%). COE and GSR is also catching up . Respondents had least variation for WB and CE with a similarity in mindset in all departments(25%). The Finnish students had maximum respondents from School of Business and Economics followed by Information Technology and Humanities.
- Table 5 captures the satisfaction in university of the respondents in context with the statement-" Overall, I'm happy with my university life". The happiness percentage of all the respondents were captured on a 4-point Likert scale(4-Very happy, 3-Happy, 2-Neutral and 1-Sad) based on the study of Chan and Mangeloja and Hirvonen. 86.6% Finnish students were found to be more content compared to Australian counterparts. The statistics of Indian respondents were as follows- Very Happy(UR=15%; TM, GSR, COE= 13% each, CE+RF+WB+ECA=46%). Happy(TM=13.5%; CE=13%; WB=13%; RF=16%; UR+GSR+COE+ECA=44.5%). Neutral(RF+WB=13.5%, ECA=12% , COE=16%, TM+UR+CE+GSR+WB=45%). Sad(CE,GSR,TM=13%; ECA=14.5%; WB+COE+UR+RF=46.5%).
- The model equation so obtained after applying regression analysis is - $Happinessscore = 0.0001289+1.000005Social+0.999869Health+0.999912Freedom.$ Average score of happiness is calculated to be 5.4. Predicted happiness score is depicted to be maximum for Social and Health factor. Pearson's coefficient of correlation is found out to be 0.65 for Social and 0.74 for Health. Correlation coefficient of Australian students was 0.66. Table 4 and Table 6 depicts the outcomes of empirical analysis upon applying regression. R-square was found to be 0.986. It was found out that there were a few missing answers given by respondents denoting their uncertainty and inability to comprehend. Lyard[13] underlined this as the soundness of the questionnaire in a survey.
- The noteworthy findings by considering 5% CI are- 82% of the respondents pertaining to departments other than Odia, Philosophy, Biotechnology, Public Administration were happy with UR. 83% of the respondents pertaining to departments other than Geography, Zoology Business Administration and Commerce were happy with COE. - 77% of the respondents pertaining to departments other than Philosophy, Law, Computer Science, Geography were happy with CE. - 80% of the respondents pertaining to departments other than Philosophy, English, Botany, Statistics were happy with GSR.
  However, by considering 10% CI - 69% of the respondents pertaining to departments other than Biotechnology, Analytical and Applied Economics, History and Archaeology, Commerce were happy with RF. 80% of the respondents pertaining to departments other than Philosophy, Anthropology, Chemistry, Geology were happy with TM 75% of the respondents pertaining

to departments other than Physics, Personal Management, Psychology, Business Administration were happy with WB. 89% of the respondents pertaining to departments other than Computer Science, Commerce, Biotechnology, Chemistry were happy with ECA.
- Predicted value of score from the explanatory variable is computed and listed in Table-7. The explanatory variables namely, ECA and WB seems to be significant(p-value= 0.854 and 0.759); while CE and UR seems to be quite insignificant for our survey. GSR is also one of the predominant criterion for student's satisfaction, similar to Finland study. Goodness of model fit was found out at 51% at 10% CI.
- Fig 6 depicts the exploratory analysis and the relationship amongst the predicted variables(Freedom, Health, Social) based on different machine learning and clustering algorithms.
- Fig. 7a, 7c, 7e, depicts the happiness score based on the predicted variables(Freedom, Health, Social) upon regression analysis. Fig. 7b, 7d, 7f, picturizes the visualization of model equation of predicted variables(Freedom, Health, Social).
- Table 8 shows LR got highest accuracy and Fscore without clustering while RF+Clustering proved best for our model with highest accuracy(0.89) and Fscore(0.98). Fig 8 depicts the visualization of Clustering algorithms on dataset.
- Model scored are LR=71.56%; SVM=71.41%; Clustering=71.46%; KNN=72.31%; RF=74.79%. The predicted happiness score is depicted from Fig 8b and turns out to be 5.2. RF+Clustering model got the least error=17.91%(8 features). % of variance explained=80.28. Mean Square Residual=0.24.

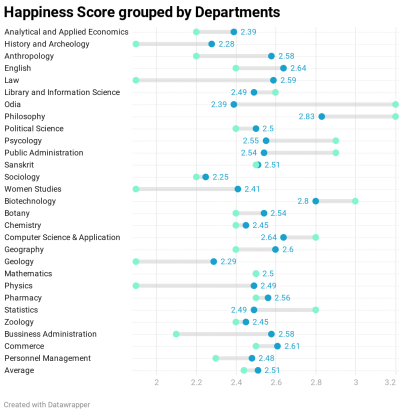
5 Conclusions and Future Scope

The ensuing model from this study establishes a new state-of-the-art to prioritize only a cluster of university students, accumulate their opinions, introspect the happiness rank department-wise using machine learning and clustering paradigms. The answer to the question whether a Utkal university student was happy or not with his/her university life, is assertive in 73% of the cases. This indicates, 69% students were satisfied with their university life. The research study of Esa[19] had dis-aggregated TM into 3 categories(meeting deadlines, WB and recreational time) and also satisfaction of school work into 5(happy with marks, enjoying studies, interesting work, coping up, resources and environment). Considering a 10% CI, age differences and gender differences had undersized impact. RF+Clustering got the highest accuracy(89%) and Fscore(0.98) and the least error(17.91%), hence turned out to be best for our study.

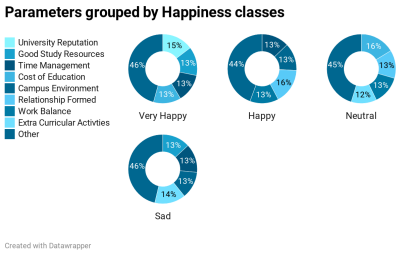
The future scope of our study can be : subsequently, we can proliferate the dataset by supplementing our online survey in other locally situated universities. Other resource crunched university students could be taken into account. A 7-scaled Likert scale could be used to make precise calculations. Experiments can be validated statistically using ANOVA and MANOVA test.

Table 3: Analyzing the Research Questions

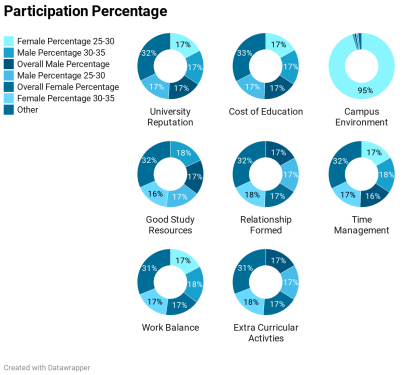
Serial	Research Question	Figure	Answer
RQ1	What is the distribution of happiness score made department-wise?	2a	The happiness range plot showed an eminent distinction in Law, Odia and Physics department. On the contrary, Zoology and Chemistry had negligible scores. Average score turned out to be 2.51.
RQ2	How is the distribution of happiness factors made based on the 4 happiness classes?	2b	The plot illustrates students were pleased with CE(averagely 25%) in all 4 classes; WB and COE showed a remarkable variation in 4 classes(11%).
RQ3	What is the overall participation percentage of students based on happiness factors?	2c	The pie chart outlines the exceptional happiness score for CE based on the overall participation of females(age 25-30) and negligible male participation. Overall percentage resulting to 17% for RF and ECA.
RQ4	What is the percentage distribution of happiness classes based on happiness factors?	2d	The pie chart portrays irrespective of the department, considering all the happiness factors, averagely 46% of students were Very Happy and Sad while Happy and Neutral resulting in 44% and 45%. Least percent(13%) for Happy and Sad classes.
RQ5	What is the departmental statistics based on all happiness factors?	3a	The plot depicts that average happiness score of 3.3 was seen in Chemistry department for TM, GSR for English, COE for Zoology and WB for Personal Management.
RQ6	What is the distribution of happiness factors based on their SD department-wise?	3b	The plot highlights the maximum standard deviation of Happiness scores(3 and 2.6) in the Botany and Philosophy department in ECA. WB in Biotechnology and Philosophy department is also catching up(2.7 and 2.3).
RQ7	How are the 'Very Happy' students grouped together based on Happiness parameters department-wise?	4a	The stacked column chart clearly depicts Commerce students were Very Happy with TM; Averagely students of Women Studies, History and Archaeology, Chemistry, Sociology and Public administration were Very Happy with UR, COE, WB and ECA respectively.
RQ8	How are the 'Happy' students grouped together based on Happiness parameters department-wise?	4b	The stacked column chart clearly depicts English and Commerce students were Happy with WB and COE. Averagely, students from Computer Sc. and Applications, Political Sc. and Geography, Sociology were Happy with WB, CE, GSR respectively.
RQ9	How are the 'Neutral' students grouped together based on Happiness parameters department-wise?	4c	The stacked column chart clearly depicts Pharmacy and Philosophy students were Neutral with RF and COE.
RQ10	How are the 'Sad' students grouped together based on Happiness parameters department-wise?	4d	The stacked column chart clearly depicts Library and Information Science students were Sad with RF. Averagely, students from Chemistry and Zoology were Sad with TM and COE respectively.
RQ11	What is the male participation statistics based on university students survey?	5a	The plot illustrates average happiness score(2.82) for CE and SD(1.41) for age 30-35. Average happiness score(2.64) for RF and SD(2.12) for age 25-30.
RQ12	What is the female participation statistics based on university students survey?	5b	The plot illustrates average happiness score(2.82) for CE and SD(1.41) for age 30-35. Average happiness score(2.64) for RF and SD(1.41) for age 25-30.



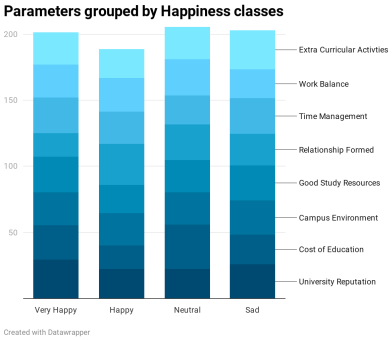
(a)



(b)



(c)

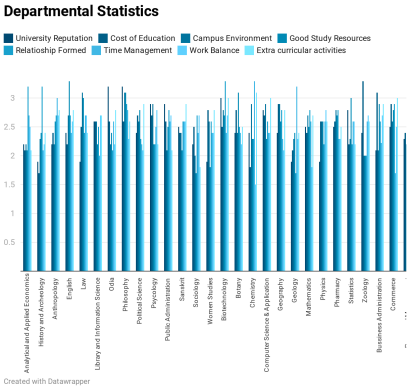


(d)

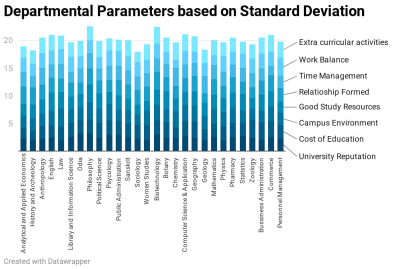
Fig. 2: Visualization of Happiness Distribution Plots

Table 4: Evaluating Regression Analysis Parameters

Dep Variable:	score	R-Squared(uncentered):	0.986
Model:	OLS	Adjusted R-Square(uncentered):	0.986
Method:	Least Squares	F-Statistic:	362.2
Date:	Fri, 20 Nov 2020	Prob (F-statistic):	7.57e-281
Time:	20:03:40	Log-Likelihood:	-307.69
No. of Observations:	312	AIC:	627.4
Df Residual:	306	BIC:	649.8
Df Model:	6	Kurtosis:	3.612
Covariance Type:	nonrobust	Skew:	0.212

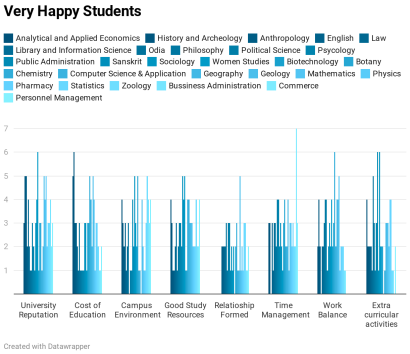


(a)

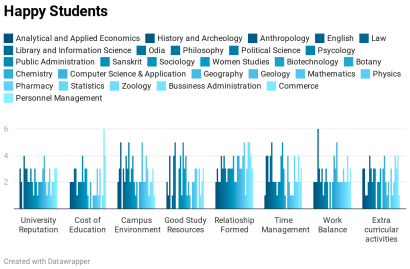


(b)

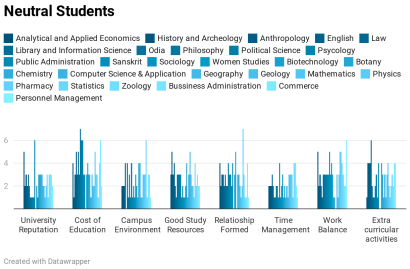
Fig. 3: Visualization of Departmental Statistics



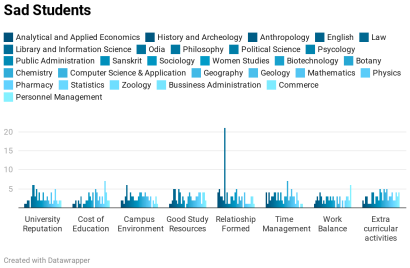
(a)



(b)

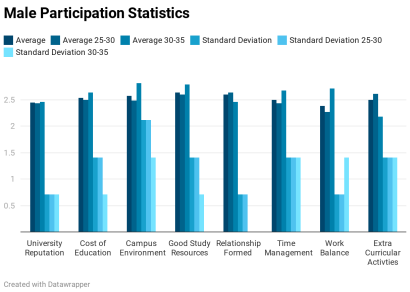


(c)

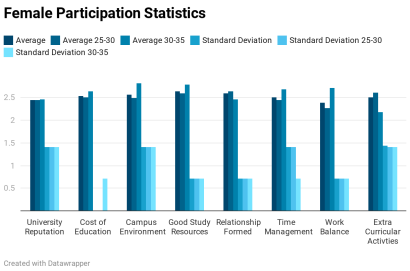


(d)

Fig. 4: Visualization of Student's Happiness Classes

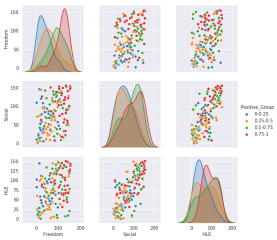


(a)

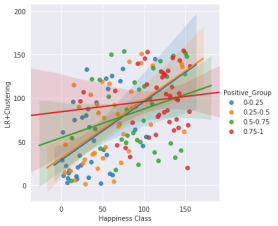


(b)

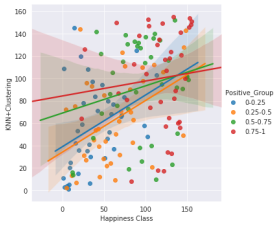
Fig. 5: Visualization of Participation Statistics



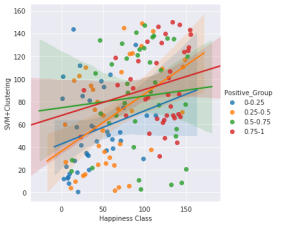
(a) Pairwise Modelling plot for predicted variables



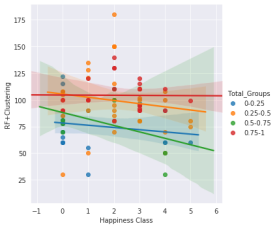
(b) LR+Clustering Modelling plot



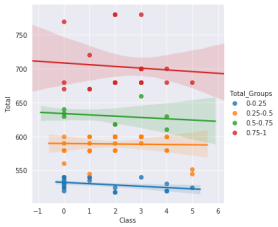
(c) KNN+Clustering Modelling plot



(d) SVM+Clustering Modelling plot



(e) RF+Clustering Modelling plot



(f) Overall Clustering plot

Fig. 6: Visualization of Modelling based on Happiness Classes

Table 5: Descriptive Statistics of Explanatory Variables

Variables	Values
Age: Student's age in years.	Mean=20(25-30years); Mean=22.97(30-35years); SD=2.82
Gender: Contrary variable, Male or Female	Males=37%; females=63%
University Reputation(UR): estimation of student's satisfaction with their university in which they are studying based on its reputation and overall rankings	Very Happy=29.64%, Happy=22.50%, Neutral=22.14%, Sad=25.71%
Cost of Education(COE): cost incurred by the students for getting higher education from educational institutions.	Very Happy=26.07%, Happy=17.86%, Neutral=33.57%, Sad=22.14%
Campus Environment(CE): environment that provides structured and regular learning opportunities, good hostel; sports facilities, conditions of security and safety for the students encircling their comfort at college.	Very Happy=25%, Happy=25.29, Neutral=24.64, Sad=25.71
Good Study Resources(GSR): resources like lab equipment, library books which helps student to achieve their academic and personal goals.	Very Happy=27.14%, Happy=21.43%, Neutral=24.64%, Sad=26.43%
Relationship Formed(RF): estimates of the bond formed between students and their faculty.	Very Happy=17.86%, Happy=31.07%, Neutral=27.14%, Sad=23.93%
Time Management(TM): estimation of time concerning work balance and university chores, meeting academic deadlines and goals, plentiful recreational time along academics.	Very Happy=26.79%, Happy=24.29%, Neutral=21.79%, Sad=27.14%
Work Balance(WB): estimation of how well a student manages his semester works regularly and remains consistent throughout.	Very Happy=25%, Happy=25.36%, Neutral=27.50%, Sad=21.79%
Extra-curricular activities(ECA): estimation of activities performed by students which falls outside their educational course curriculum.	Very Happy=24.29%, Happy=21.79%, Neutral=24.64%, Sad=29.29%
Freedom: predicted variable; estimates the privilege of students to openly think, act and interact with seniors, juniors and peers.	Very Happy=1.97%, Happy=6.47%, Neutral=4.95%, Sad=4.23%
Health: predicted variable; estimates the medical facility given to students based on health conditions.	Very Happy=8.71%, Happy=9.1%, Neutral=9.09%, Sad=9.09%
Social: predicted variable; estimates the social activities like tree plantation, promoting underprivileged students through campaigns and initiatives conducted in university.	Very Happy=13.6%, Happy=15.73%, Neutral=12.11%, Sad=11.91%

Table 6: Descriptive Statistics of Predicted variables

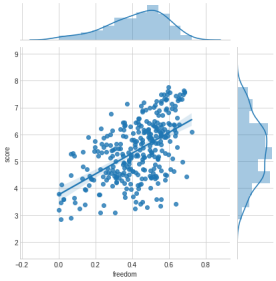
Parameters	coef	std err	t	P>—t—	0.025	0.975	count	mean	SD	min	25%	50%	75%	max	Pearson
Social	2.26	0.15	14.78	0	1.96	2.56	312	1.21	0.30	0.0	1.05	1.26	1.45	1.64	0.65
Health	1.25	0.26	4.80	0	0.73	1.76	312	0.66	0.25	0.0	0.48	0.699	0.85	1.41	0.74
Freedom	1.86	0.28	6.44	0	1.29	2.43	312	0.42	0.15	0.0	0.32	0.44	0.54	0.72	0.55
Score	-	-	-	-	-	-	312	5.39	1.11	2.85	4.51	5.37	6.17	7.76	-

Table 7: Estimates of Model of satisfaction with university life

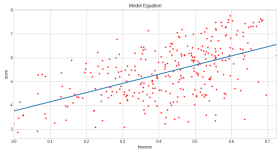
Variable	Coefficient (std err.)	y_predict	y_test
Age	0.929(0.551)	5.459	5.124
Gender	1.065(3.842)	4.592	4.297
University Reputation	0.881(0.569)	6.214	6.455
Cost of Education	0.971(0.663)	6.257	6.786
Campus Environment	0.520(1.369)	6.163	6.298
Good Study Resources	0.943(0.978)	4.852	3.819
Relationship Formed	1.02(1.222)	4.391	4.633
Time Management	0.936(0.475)	3.942	4.971
Work Balance	1.404(0.878)	5.907	5.754
Extra-curricular activities	1.447(0.955)	5.310	4.857
Freedom	0.99(0.525)	Psuedo R-Square	0.767
Health	0.99(0.696)	Model-fitting	51.022
Social	1.00(0.425)	MSE	0.26
		RMSE	0.51
		Intercept	3.36
		Coef.	2.25
		Constant	1.85
		rscore	0.63
		Train Error	0.528
		Test Error	0.470

Table 8: Empirical Results obtained from the Machine Learning Algorithms and Clustering

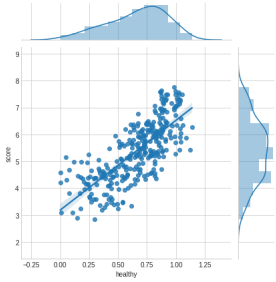
	Without Clustering					Clustering				
	F1-Score	Accuracy	AUC	Train score	Test Score	F1-score	Accuracy	AUC	Train Score	Test Score
LR	0.88	0.875	0.96	0.84	0.875	0.90	0.87	0.90	0.84	0.87
SVM	0.90	0.84	0.94	0.83	0.77	0.85	0.82	0.93	0.82	0.70
KNN	0.86	0.81	0.92	0.81	0.84	0.84	0.78	0.91	0.84	0.71
RF	0.84	0.87	0.96	0.99	0.875	0.98	0.89	0.91	0.81	0.82



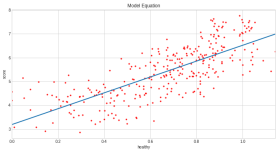
(a) Scatter-plot for Happiness Score and Freedom



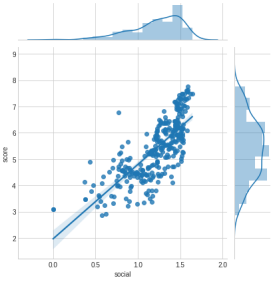
(b) Modeling for Happiness score vs Freedom



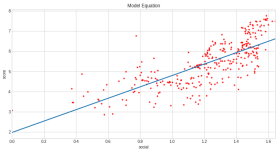
(c) Scatter-plot for Happiness Score and Health



(d) Modeling for Happiness score vs Health



(e) Scatter-plot for Happiness Score and Social



(f) Modeling for Happiness score vs Social

Fig. 7: Visualization of Predicted variables using Regression Analysis for Student's Happiness

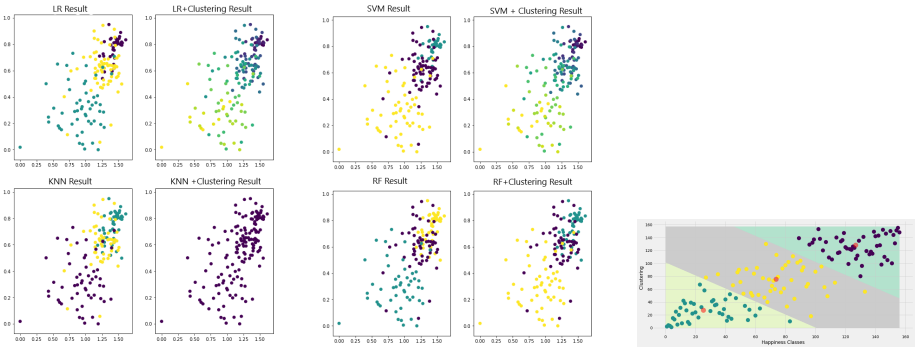


Fig. 8: Visualization of Clustering Results based happiness scores

References

1. Seligman, Martin EP. "Positive psychology, positive prevention, and positive therapy." Handbook of positive psychology 2, no. 2002 (2002): 3-12.
2. Diener, Ed. "Subjective well-being: The science of happiness and a proposal for a national index." American psychologist 55, no. 1 (2000): 34.
3. Fujita, Frank, and Ed Diener. "Life satisfaction set point: stability and change." Journal of personality and social psychology 88, no. 1 (2005): 158.
4. Kahneman, Daniel, Edward Diener, and Norbert Schwarz, eds. Well-being: Foundations of hedonic psychology. Russell Sage Foundation, 1999.
5. Forgeard, Marie JC, Eranda Jayawickreme, Margaret L. Kern, and Martin EP Seligman. "Doing the right thing: Measuring wellbeing for public policy." International journal of wellbeing 1, no. 1 (2011).
6. Haybron, Daniel M. "Mood propensity as a constituent of happiness: A rejoinder to Hill." Journal of Happiness Studies 11, no. 1 (2010): 19-31.
7. Kimhy, David, Philippe Delespaul, Hongshik Ahn, Shengnan Cai, Marina Shikhman, Jeffrey A. Lieberman, Dolores Malaspina, and Richard P. Sloan. "Concurrent measurement of "real-world" stress and arousal in individuals with psychosis: assessing the feasibility and validity of a novel methodology." Schizophrenia bulletin 36, no. 6 (2010): 1131-1139.
8. Davis, James H., F. David Schoorman, and Lex Donaldson. "Davis, Schoorman, and Donaldson reply: The distinctiveness of agency theory and stewardship theory." (1997): 611-613.
9. Deci, Edward L., and Richard M. Ryan. "Intrinsic motivation." The corsini encyclopedia of psychology (2010): 1-2.
10. Frey, Bruno S., and Alois Stutzer. "What can economists learn from happiness research?." Journal of Economic literature 40, no. 2 (2002): 402-435.
11. Lyard, Florent, Fabien Lefevre, Thierry Letellier, and Olivier Francis. "Modelling the global ocean tides: modern insights from FES2004." Ocean dynamics 56, no. 5-6 (2006): 394-415.
12. Glenn, Norval. "Is the apparent U-shape of well-being over the life course a result of inappropriate use of control variables? A commentary on Blanchflower and Oswald (66: 8, 2008, 1733–1749)." Social science medicine 69, no. 4 (2009): 481-485.
13. Lyubomirsky, Sonja, and Heidi S. Lepper. "A measure of subjective happiness: Preliminary reliability and construct validation." Social indicators research 46, no. 2 (1999): 137-155.
14. Veenhoven, Ruut. "Happiness." The psychologist (2003).
15. Kahneman, Daniel, and Jason Riis. "Living, and thinking about it: Two perspectives on life." The science of well-being 1 (2005): 285-304.
16. Daniel, Kahneman, and Riis Jason. "Living, and thinking about it: Two perspectives on life." (2005): 285-304.
17. Arrindell, Willem A., José Heesink, and Jan A. Feij. "The satisfaction with life scale (SWLS): Appraisal with 1700 healthy young adults in The Netherlands." Personality and individual differences 26, no. 5 (1999): 815-826.
18. Shimai, Satoshi, Keiko Otake, Narisuke Utsuki, Akira Ikemi, and Sonja Lyubomirsky. "Development of a Japanese version of the Subjective Happiness Scale (SHS), and examination of its validity and reliability." [Nihon koshu eisei zasshi] Japanese journal of public health 51, no. 10 (2004): 845.
19. Mangeloja, Esa, and Tatu Hirvonen. "What makes university students happy." International Review of Economics Education 6, no. 2 (2007): 27-41.

20. Chan, Grace, Paul W. Miller, and MoonJoong Tcha. "Happiness in university education." International Review of Economics Education 4, no. 1 (2005): 20-45.
21. Clark, Andrew E., and Andrew J. Oswald. "Satisfaction and comparison income." Journal of public economics 61, no. 3 (1996): 359-381.
22. Hartog, Joop, and Hessel Oosterbeek. "Health, wealth and happiness: why pursue a higher education?." Economics of Education Review 17, no. 3 (1998): 245-256.
23. Yong, Fung L., and John D. McIntyre. "Comparison of self-concepts of students identified as gifted and regular students." Perceptual and motor skills 73, no. 2 (1991): 443-446.
24. Terry, Tara, and E. Scott Huebner. "The relationship between self-concept and life satisfaction in children." Social Indicators Research 35, no. 1 (1995): 39-52.
25. Zeidner, Moshe. "“Don't worry—be happy”: The sad state of happiness research in gifted students." High Ability Studies (2020): 1-18.