Designed by: Robin B.

STUDENT PERFORMANCE - Overview

Number of Students	Average Age	Average Number of Enrolled Curricular Units (1st Semester)	Average Numberof Approved Curricular Units (1st Semester)
4,424	23.27	6.27	4.71

Student Status: All

Course: All

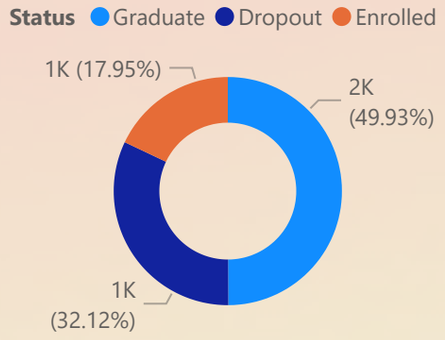
Gender: All

International: All

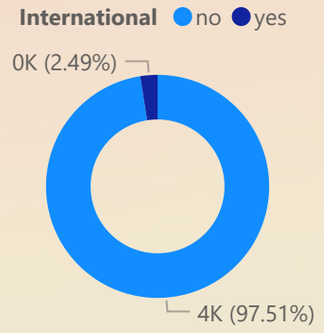
Marital status: All

Scholarship hol...: All

Proportion of Students by Status



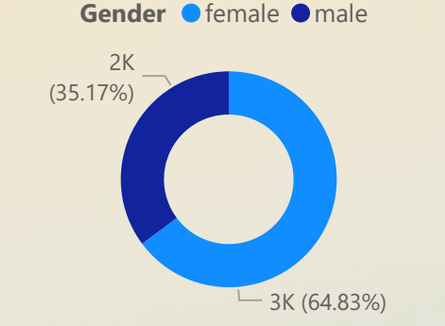
Proportion of International Students



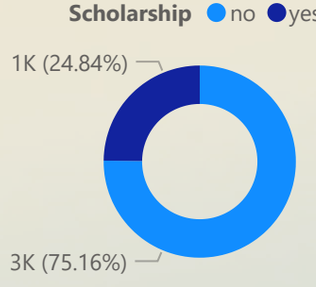
Number of Records by Marital Status

Marital status	Number of Students	% Grand Total
Single	3,919	88.58%
Married	379	8.57%
Divorced	91	2.06%
Facto union	25	0.57%
Legally separated	6	0.14%
Widower	4	0.09%

Proportion of Students by Gender



Proportion of Students with Scholarships



Number of Records by Age Group

Age Group	Number of Students	% Grand Total
17-25	3,379	76.38%
26-35	653	14.76%
36-45	279	6.31%
46-55	100	2.26%
56-65	12	0.27%
Over 65	1	0.02%

Designed by: Robin B.

STUDENT PERFORMANCE - Dropout Evaluation

Student Status
All

Course
All

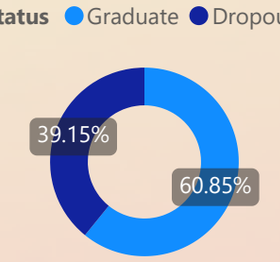
Gender
All

International
All

Marital status
All

Scholarship hol...
All

Proportion of Dropouts/Graduates



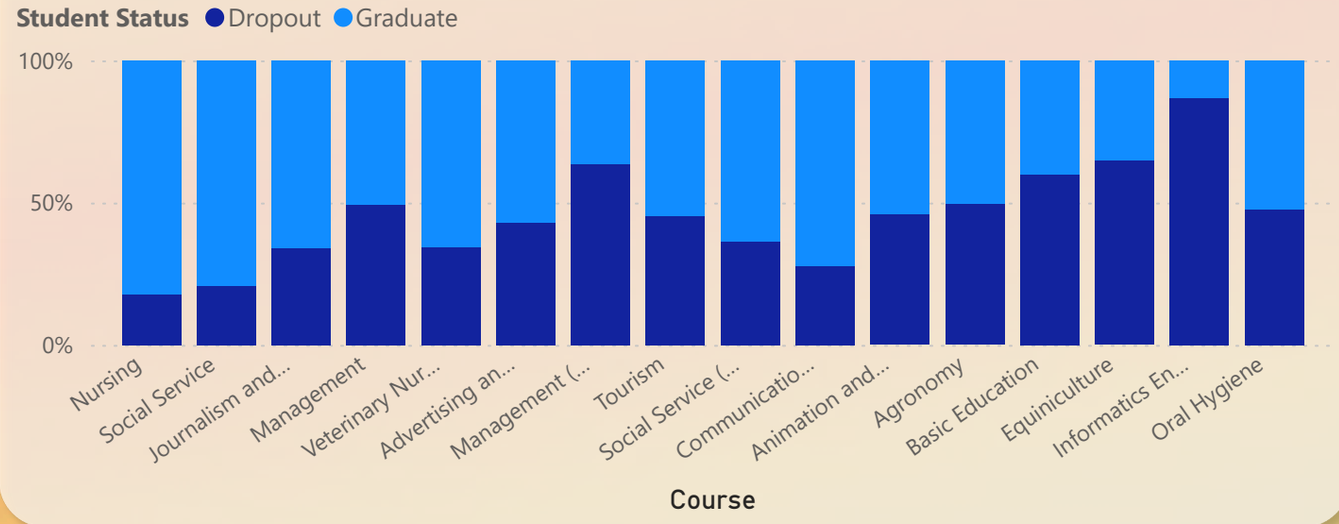
Dropout Rate by Age

Age Group	Number of Students	Dropout Rate
17-25	2,726	30.70%
26-35	559	67.26%
36-45	242	60.33%
46-55	90	61.11%
56-65	12	50.00%
Over 65	1	100.00%

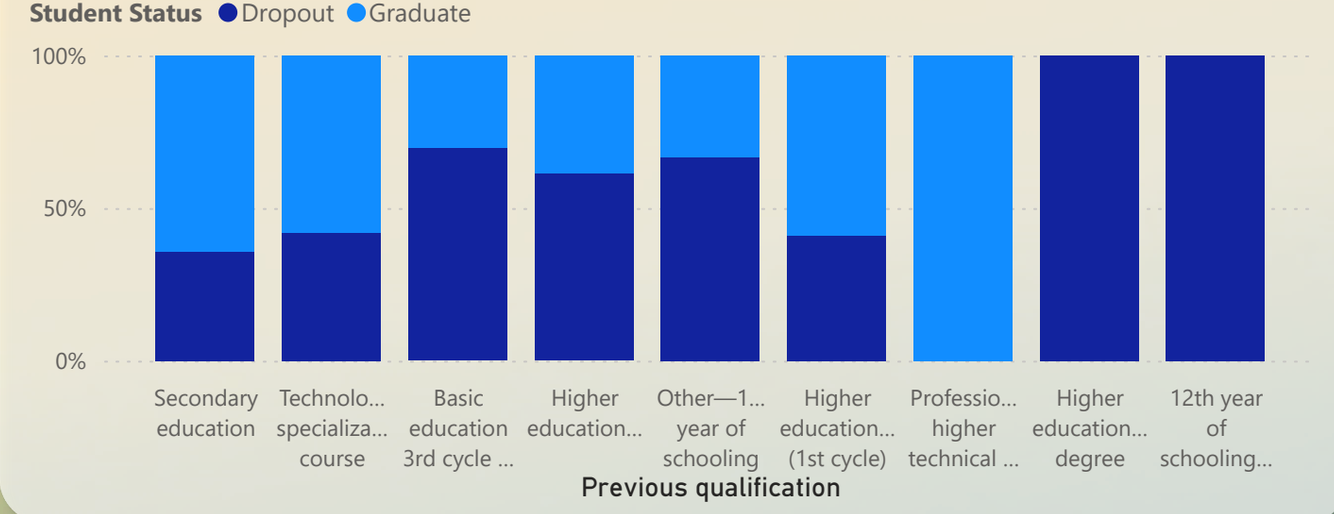
Dropout Rate by Number of Approved Units

Number of approved curricular Units	Number of Students	Dropout Rate
0	647	88.10%
1-5	1,206	53.65%
6-10	1,612	10.73%
11-15	125	23.20%
16-20	35	2.86%
Over 20	5	20.00%

Proportion of Dropouts/Graduates by Course



Proportion of Dropouts/Graduates by Previous Qualification



Designed by: Robin B.

STUDENT PERFORMANCE - Predictive Model

Objective: Utilize machine learning to predict whether a student will drop out or graduate while omitting second semester fields.

Student Status: All

Course: All

Gender: All

International: All

Marital status: All

Scholarship hol...: no

Data Overview

# of fields in entire Dataset:	35
# of fields used for the model:	15

Fields Used for Predicting Student Status

1. Marital status
2. Course
3. Daytime/evening attendance"
4. Previous qualification
5. Nationality
6. Displaced
7. Debtor
8. Tuition fees up to date
9. Gender
10. Scholarship holder
11. Age at enrollment
12. International
13. Curricular units 1st sem (approved)
14. Curricular units 1st sem (enrolled)
15. Target (Student Status)

Field with the Strongest Correlation to Student Status

Curricular units 1st sem (approved):
- correlation of **0.529** with student status.

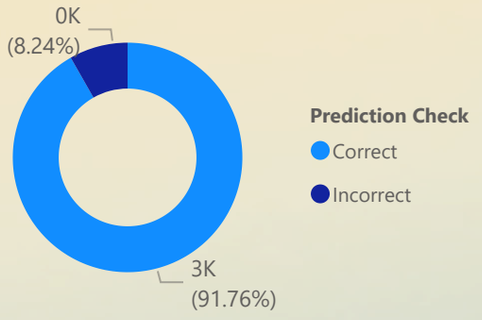
ML Model Results

Using the 15 selected fields, we compared the Support Vector Classifier, Gradient Boosting Classifier, and Random Forest Classifier to predict the actual student status within the dataset. Below, you'll find the accuracy scores for each of these three ML models.

- Support Vector Classifier: **71% Accuracy Score**
- Gradient Boosting Classifier: **76% Accuracy Score**
- Random Forest Classifier: **92% Accuracy Score**

The Random Forest Classifier model was used to make final predictions. These predictions were joined with our data to enable us to compare the predicted values with the actual values. (See below):

Proportion of Correct/Incorrect Predictions for Student Status



Number of Correct/Incorrect Predictions by Number of Approved Units

Curricular Units Range (approved)	Correct	Incorrect
0	592	52
1-5	1,275	120
6-10	1,032	102
11-15	120	
16-20	27	
Over 20	5	