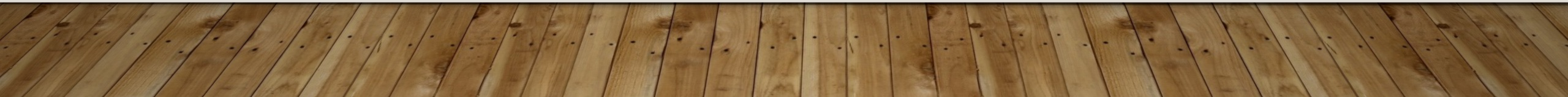# ACADEMIC DROPOUT ANALYSIS

JANUARY 2024

OBJECTIVES

A comprehensive dataset containing students and their demographics, application details, prior academic qualifications, socioeconomic background, admission scores, and academic performance will be analyzed to **determine important factors influencing students' academic success or dropout**

To achieve this:

1. Comprehensive Exploratory Analysis

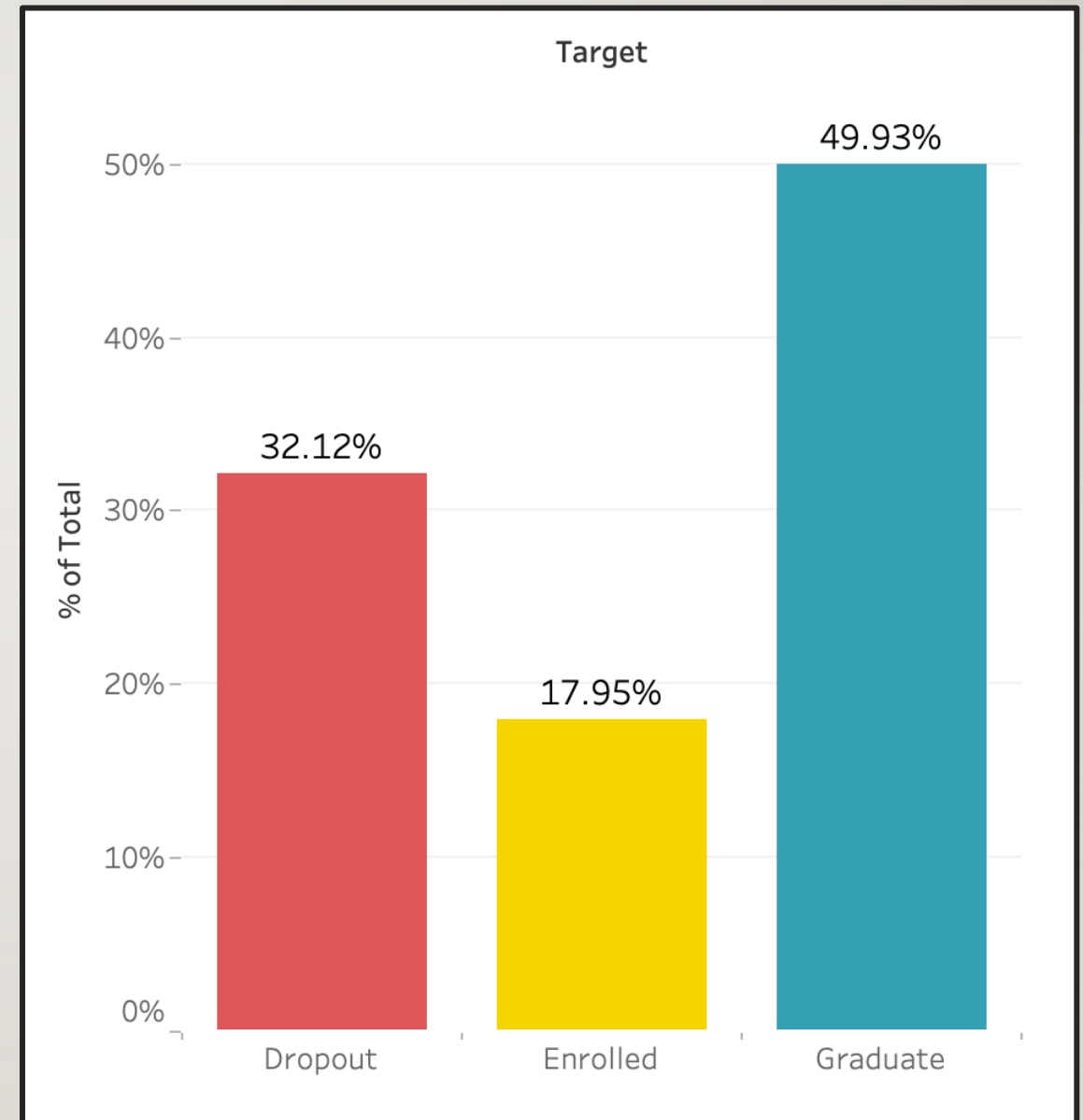2. Predictive Model determining dropout

EXPLORATORY ANALYSIS

Of the whole dataset
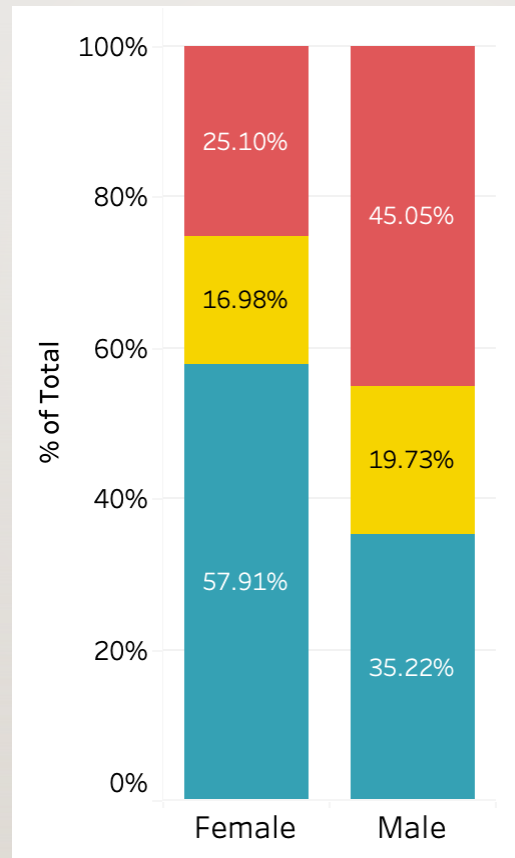
32.12 % Dropout

17.95 % Enrolled

49.93 % Graduate

DEMOGRAPHICS

Enrollment Age

Age at enrollment	Dropout	Enrolled	Graduate
16		40.00%	60.00%
18	21.01%	16.90%	62.10%
20	24.54%	22.80%	52.66%
22	35.11%	21.63%	43.26%
24	45.98%	22.77%	31.25%
26	57.84%	16.76%	25.41%
28	61.74%	13.42%	24.83%
30	63.46%	12.50%	24.04%
32	50.00%	12.26%	37.74%
34	53.21%	15.60%	31.19%
36	57.14%	10.39%	32.47%
38	55.22%	13.43%	31.34%
40	44.44%	16.67%	38.89%
42	47.37%	13.16%	39.47%
44	53.49%	13.95%	32.56%
46	56.67%	6.67%	36.67%
48	58.33%	16.67%	25.00%
50	60.87%	8.70%	30.43%
52	27.27%	9.09%	63.64%
54	58.33%	8.33%	33.33%
56	50.00%		50.00%
58	66.67%		33.33%
60	33.33%		66.67%
62			100.00%
70	100.00%		

Gender

International

International	Dropout	Target Enrolled	Graduate
No	32.20%	17.85%	49.95%
Yes	29.09%	21.82%	49.09%

Daytime/Evening Attendance

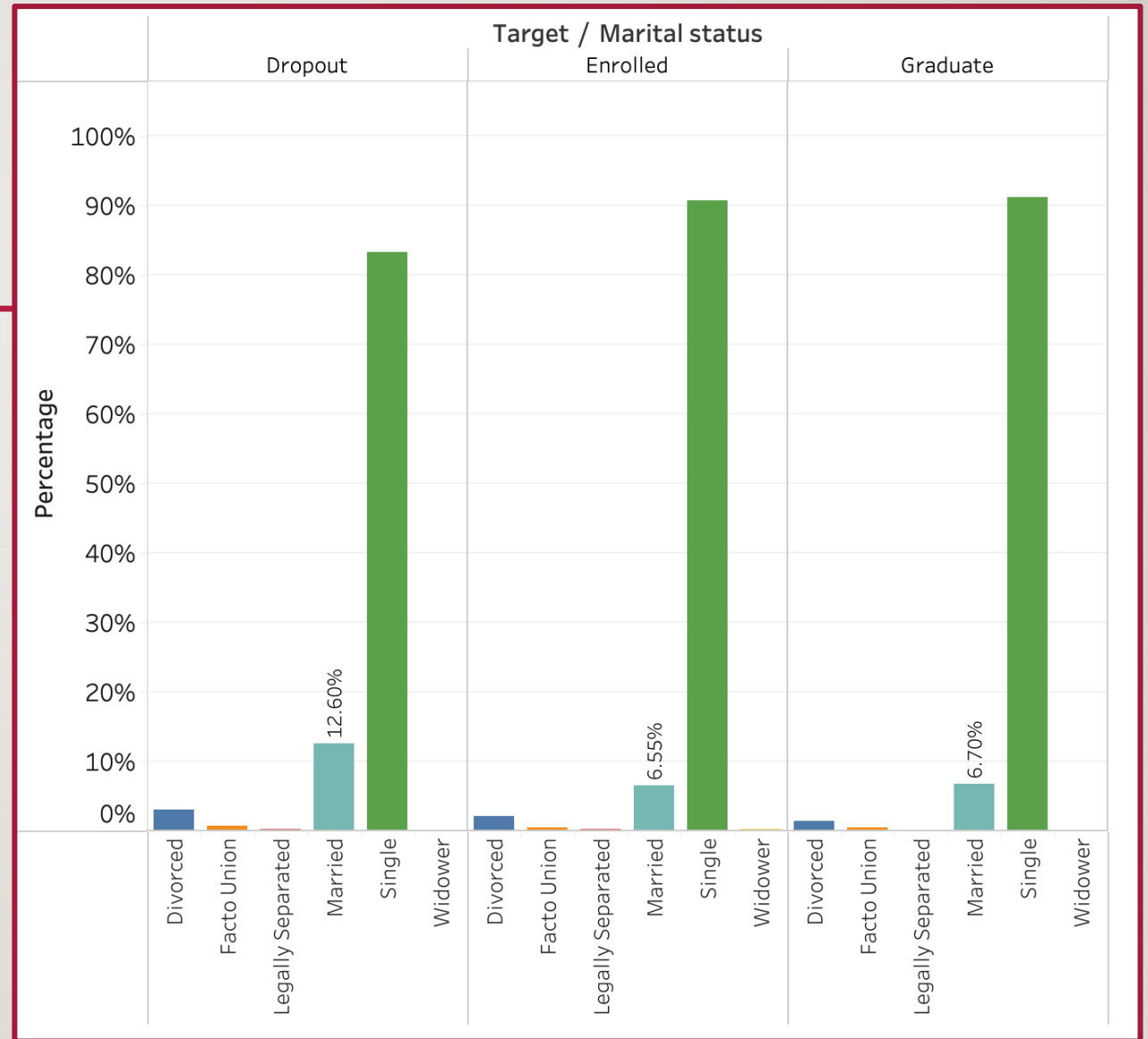| D/E Attendance	Dropout	Target Enrolled	Graduate
Daytime	30.80%	18.24%	50.95%
Evening	42.86%	15.53%	41.61%

DEMOGRAPHICS

Marital Status

Insights

Approximately double **married students** who are dropouts versus married students who are enrolled or graduates

Higher dropout rates for **older** students, **males** students, and **evening attendance** students

FAMILY BACKGROUND

Students who have a parent who is not working or is deceased are much more likely to dropout
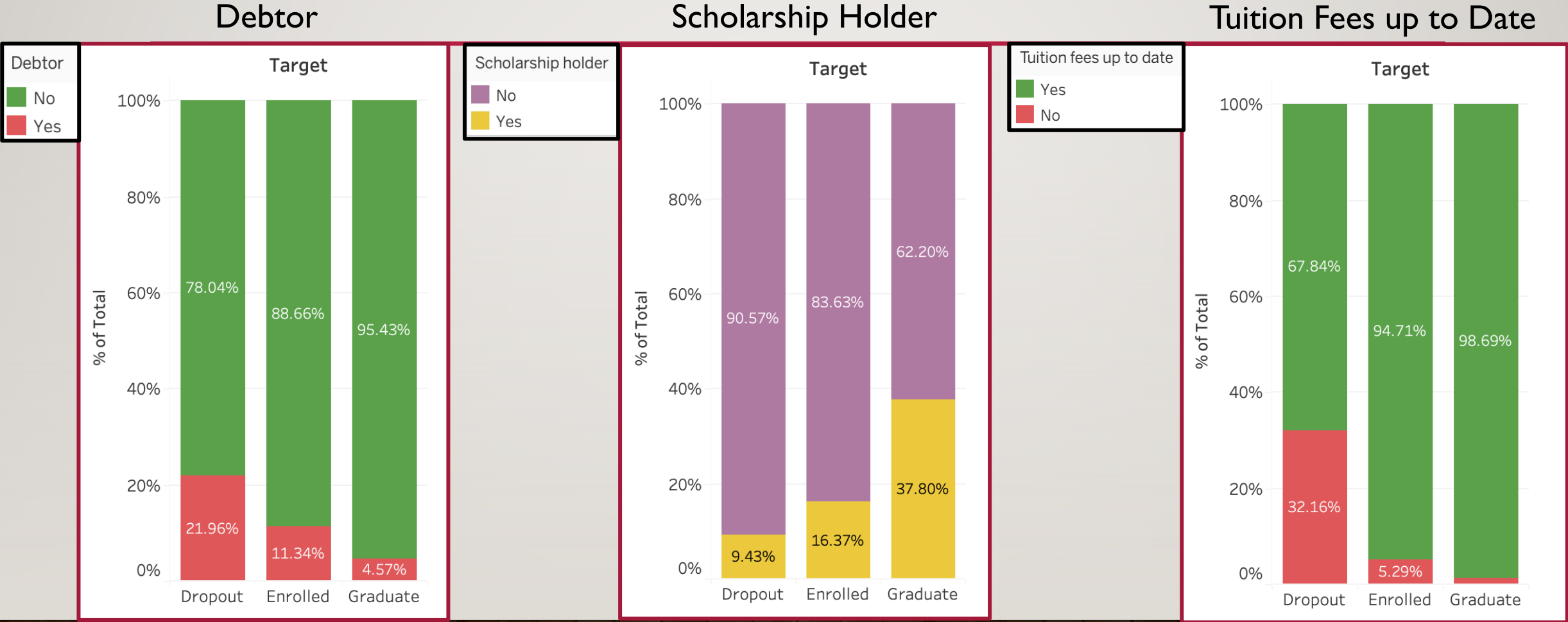
Mother's Occupation	Target		
	Dropout	Enrolled	Graduate
Null	68.75%	0.69%	30.56%
Administrative staff	29.43%	17.74%	52.83%
Farmers and Skilled Workers in Agriculture, Fisheries, & Forestry	29.41%	17.65%	52.94%
Installation and Machine Operators and Assembly Workers	31.07%	16.74%	52.19%
Intermediate Level Technicians and Professions	30.35%	17.99%	51.65%
Personal Services, Security and Safety Workers, and Sellers	28.57%	15.38%	56.04%
Representatives of the Legislative Power and Executive Bodies	32.08%	24.53%	43.40%
Skilled Workers in Industry, Construction, and Craftsmen	41.67%	19.44%	38.89%
Specialists in Intellectual and Scientific Activities	27.07%	22.51%	50.43%
Street vendors (except food) and street service providers	38.46%	24.73%	36.81%
Student	38.24%	14.71%	47.06%
Unskilled Workers	25.00%	50.00%	25.00%

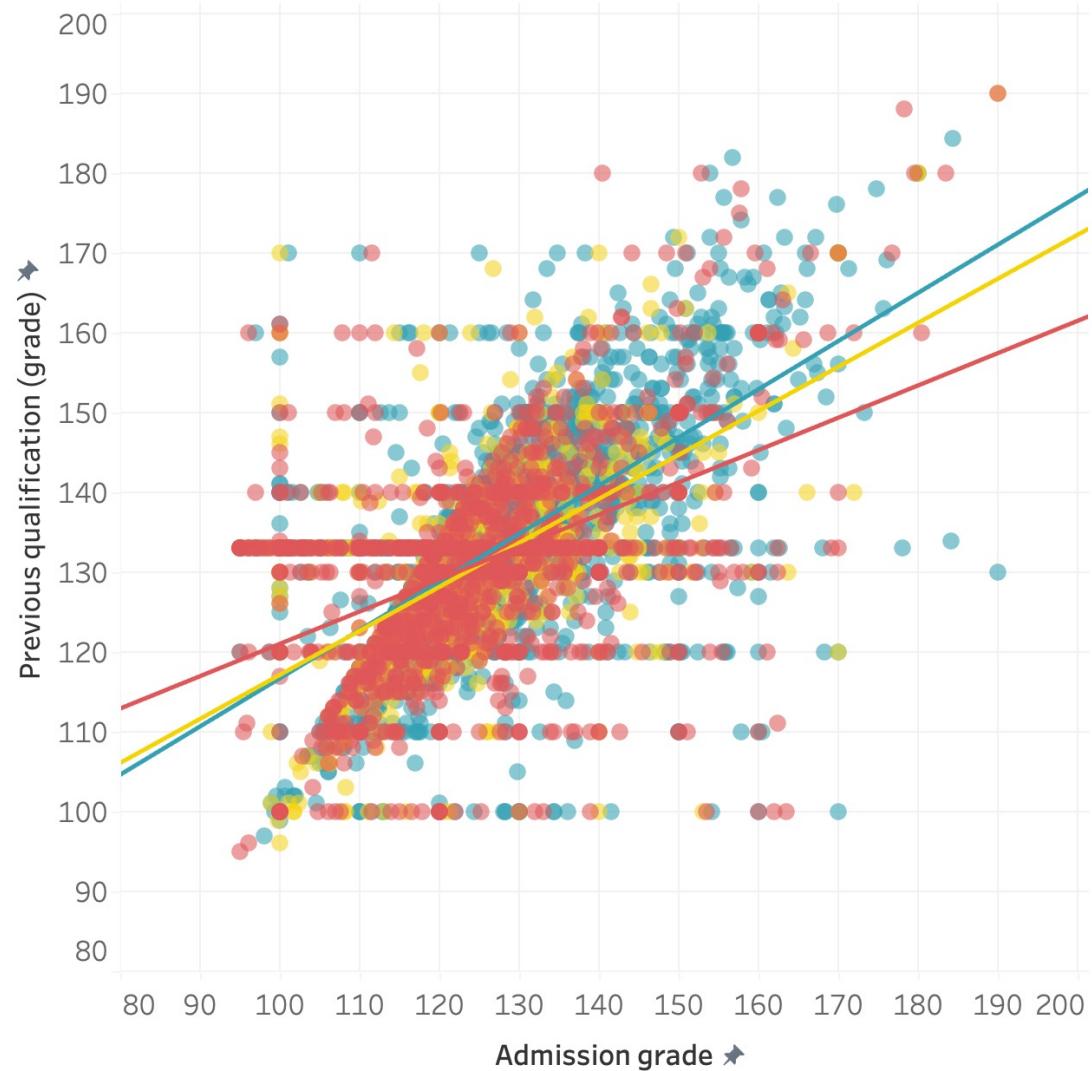| Father's Occupation	Target		
	Dropout	Enrolled	Graduate
Null	64.06%	2.34%	33.59%
Administrative staff	28.68%	20.74%	50.58%
Farmers and Skilled Work..	27.63%	16.67%	55.71%
Installation and Machine ..	31.98%	16.14%	51.88%
Intermediate Level Techni..	36.01%	19.17%	44.82%
Personal Services, Securit..	28.51%	19.83%	51.65%
Representatives of the Le..	35.53%	25.38%	39.09%
Skilled Workers in Industr..	29.56%	17.30%	53.14%
Specialists in Intellectual ..	29.69%	16.41%	53.91%
Street vendors (except fo..	36.72%	24.29%	38.98%
Student	35.82%	17.16%	47.01%
Unskilled Workers	31.95%	20.30%	47.74%

FINANCIAL BACKGROUND

Higher % of debtors, unpaid tuition fees, and non-scholarship holders in dropout students
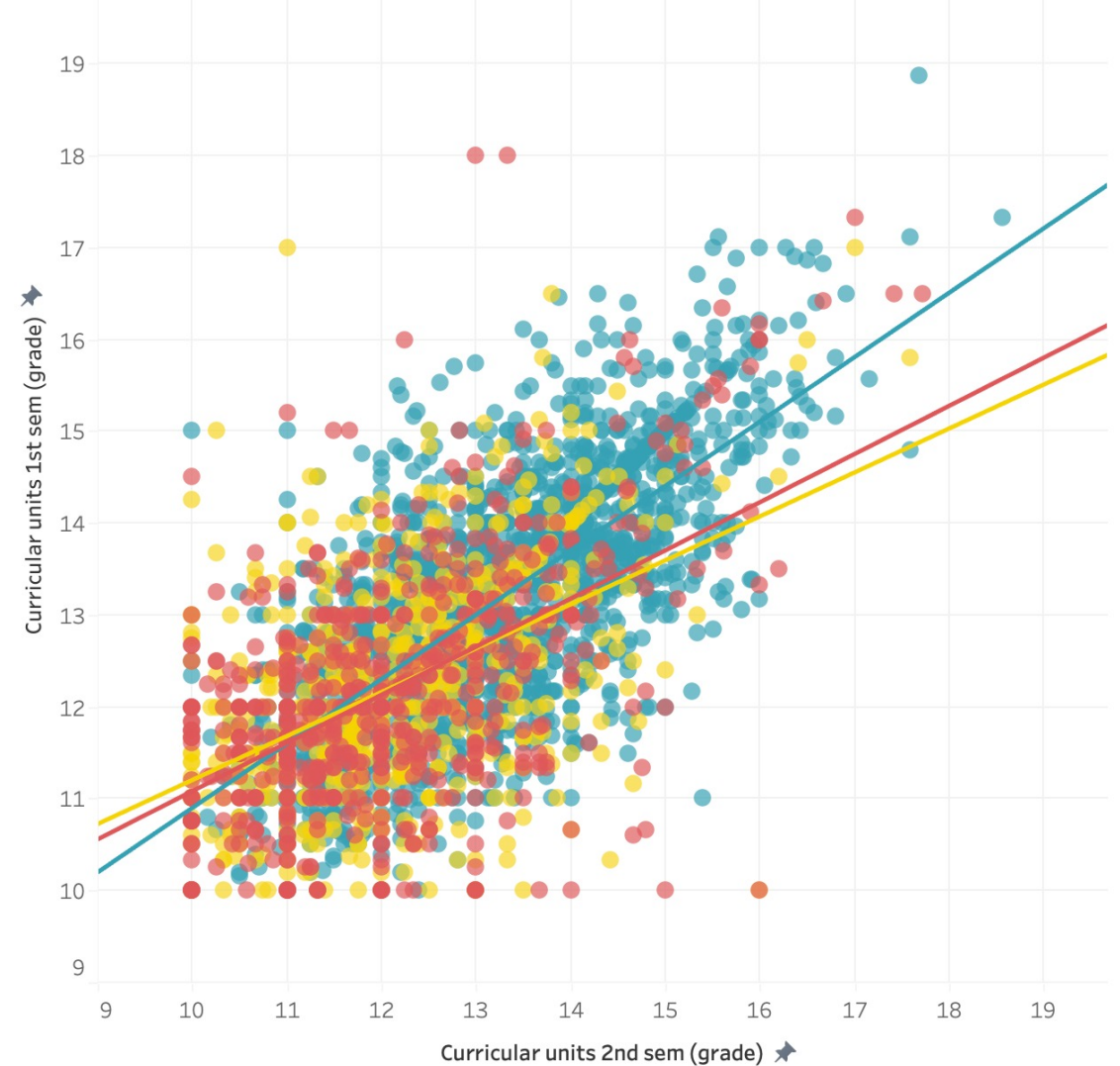
ACADEMIC

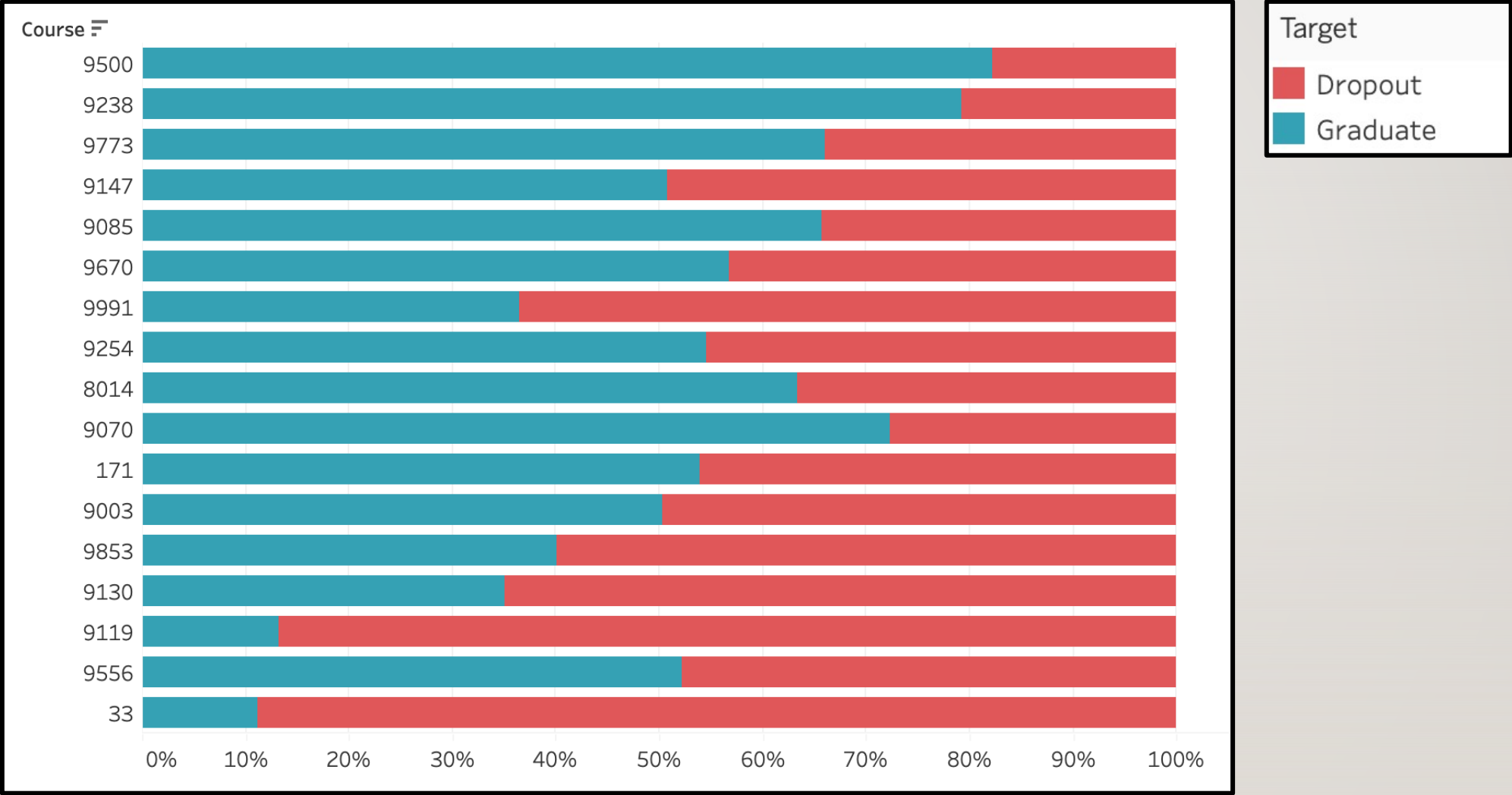## Prev. Qualification Grade vs Admissions Grade

1st Semester vs 2nd Semester Grades

ACADEMIC

ECONOMY

Lower Dropout Rates when GDP is higher

GDP	Dropout	Target Enrolled	Graduate
-5	35.01%	17.38%	47.61%
-4	32.65%	20.45%	46.90%
-2	33.65%	15.04%	51.31%
-1	38.40%	12.71%	48.90%
0	33.30%	19.04%	47.66%
1	29.79%	18.56%	51.65%
2	28.26%	16.18%	55.56%
3	27.72%	20.92%	51.36%

Lower percentage of students enrolled when unemployment rate is higher

Unemployment rate	Dropout	Target Enrolled	Graduate
7	35.38%	20.84%	43.78%
8	27.72%	20.92%	51.36%
9	32.65%	20.45%	46.90%
10	34.67%	17.14%	48.19%
11	28.26%	16.18%	55.56%
12	28.70%	17.71%	53.59%
13	30.26%	16.41%	53.33%
15	35.01%	17.38%	47.61%
16	38.40%	12.71%	48.90%

PREDICTIVE MODEL

TWO CLASSIFICATION MODELS

Model 1

- All variables in dataset, **including** 1st and 2nd semester performance

Model 2

- All variables in dataset, **excluding** 1st and 2nd semester performance
- Made for incoming students to predict dropout, in order to take preventative measures if necessary

MODEL 1

76.5% accuracy

Features	Importance
Curricular units 2nd sem approved	47.021838
Tuition fees up to date	5.381794
Curricular units 2nd sem grade	3.576575
Curricular units 1st sem evaluations	3.161188
Admission grade	2.810703
Previous qualification grade	2.766756
Curricular units 1st sem approved	2.445071
Age at enrollment	2.434991
Curricular units 2nd sem enrolled	2.350735
Curricular units 2nd sem evaluations	2.331269

MODEL 1

76.5 % accuracy

Features	Importance
Curricular units 2nd sem approved	47.021838
Tuition fees up to date	5.381794
Curricular units 2nd sem grade	3.576575
Curricular units 1st sem evaluations	3.161188
Admission grade	2.810703
Previous qualification grade	2.766756
Curricular units 1st sem approved	2.445071
Age at enrollment	2.434991
Curricular units 2nd sem enrolled	2.350735
Curricular units 2nd sem evaluations	2.331269

MODEL 2

For incoming students

64.2 % accuracy

Features	Importance
Tuition fees up to date	15.364829
Admission grade	9.007001
Age at enrollment	7.216262
Previous qualification grade	6.636007
Scholarship holder	6.560176
GDP	4.448183
Unemployment rate	3.366316
Course_9500	2.585017
Inflation rate	2.410330
Course_9119	2.042362

LIMITATIONS

LIMITATIONS

Classification model's accuracy was not very strong, however there are many outside factors that may affect a student's dropout or academic success which may have not been recorded

Data to Consider:
- Student Mental Health, Family/Community Support, School Environment, Peer Influences, Extracurricular Activities

One Hot Encoding on Categorical Variables in Predictive Models
- Likely diminished the accuracy of both models