



Predicting Students' Performance

Aziza Gulyamova

Business Problem

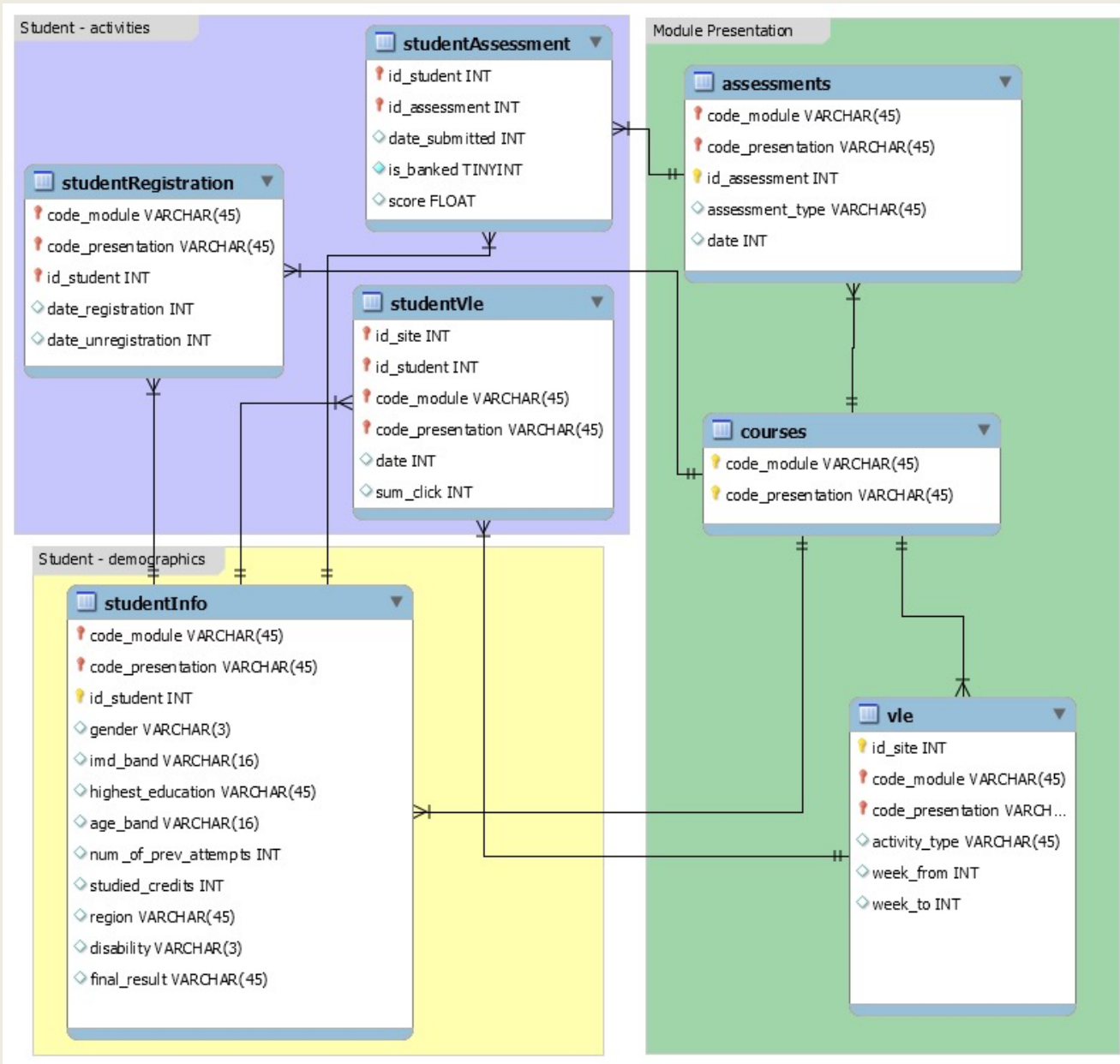
With COVID-19 accelerating online learning method, Educational Institutions are interested in increasing the success rate of the students.

- Academies are interested in gaining knowledge of students' outcomes ahead.
-

Business Understanding

Advantages of predicting student's performance outcome:

- Learn student's current educational status
- Navigate student in proper direction
- Improve institution's success rate



Data

Open University Learning Analytics Data:

- 7 tabular files connected via specific keys
- 3 categories of data: Demographic, Students Activities, Module Information
- 32,593 student observations, 7 modules

Data Cleaning

Observations from beginning of the course to half term – 135 days

Drop unnecessary features

Bin target variable into two categories: Pass and Fail

Merge required features into one table

Features and Target

- **Features:**

code_module

age_band

date_registration

attempts

sum_click

studied_credits

gender

disability

region

mean_score

highest_education

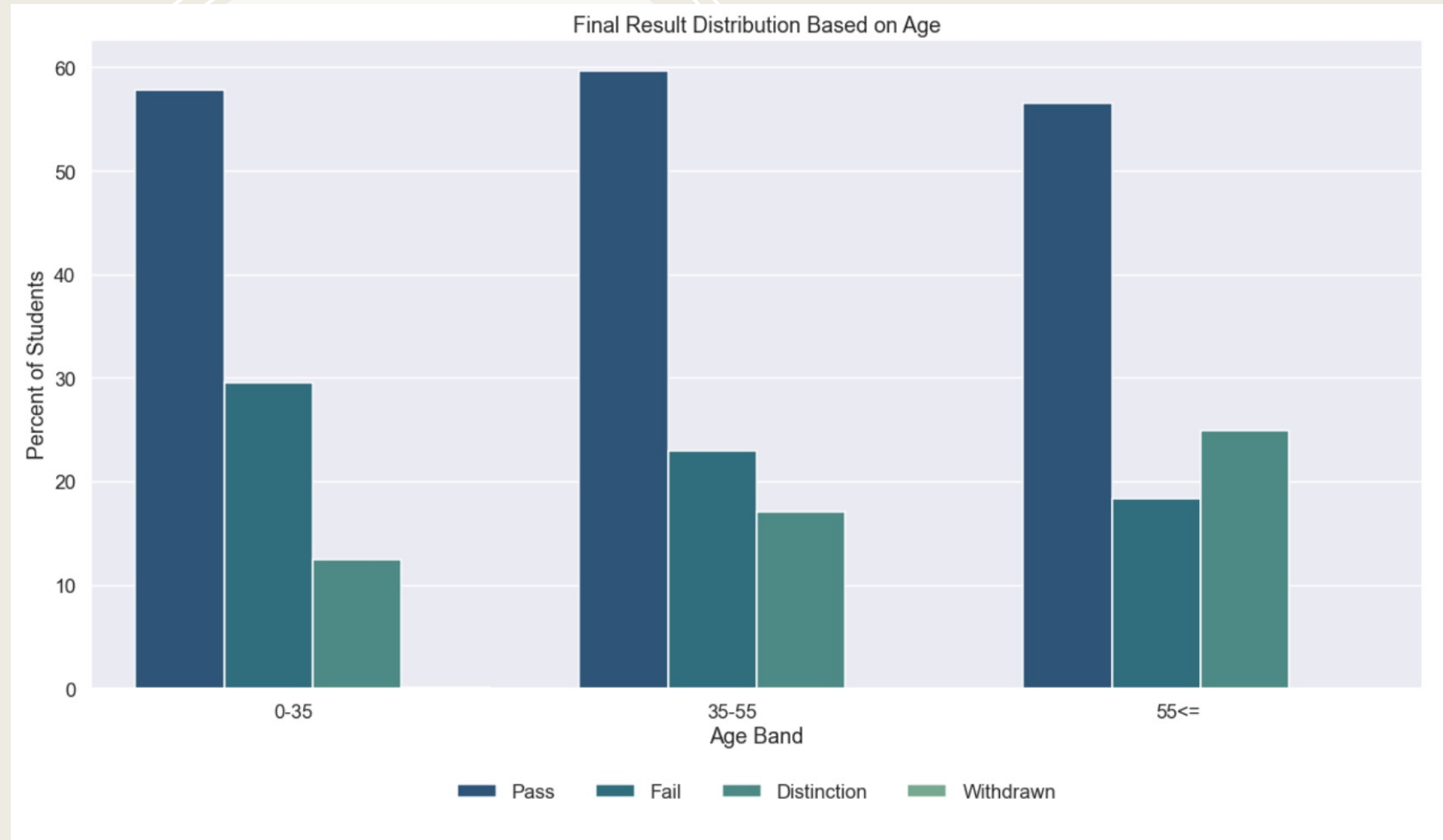
course_length

imd_band

- **Target:** final_result

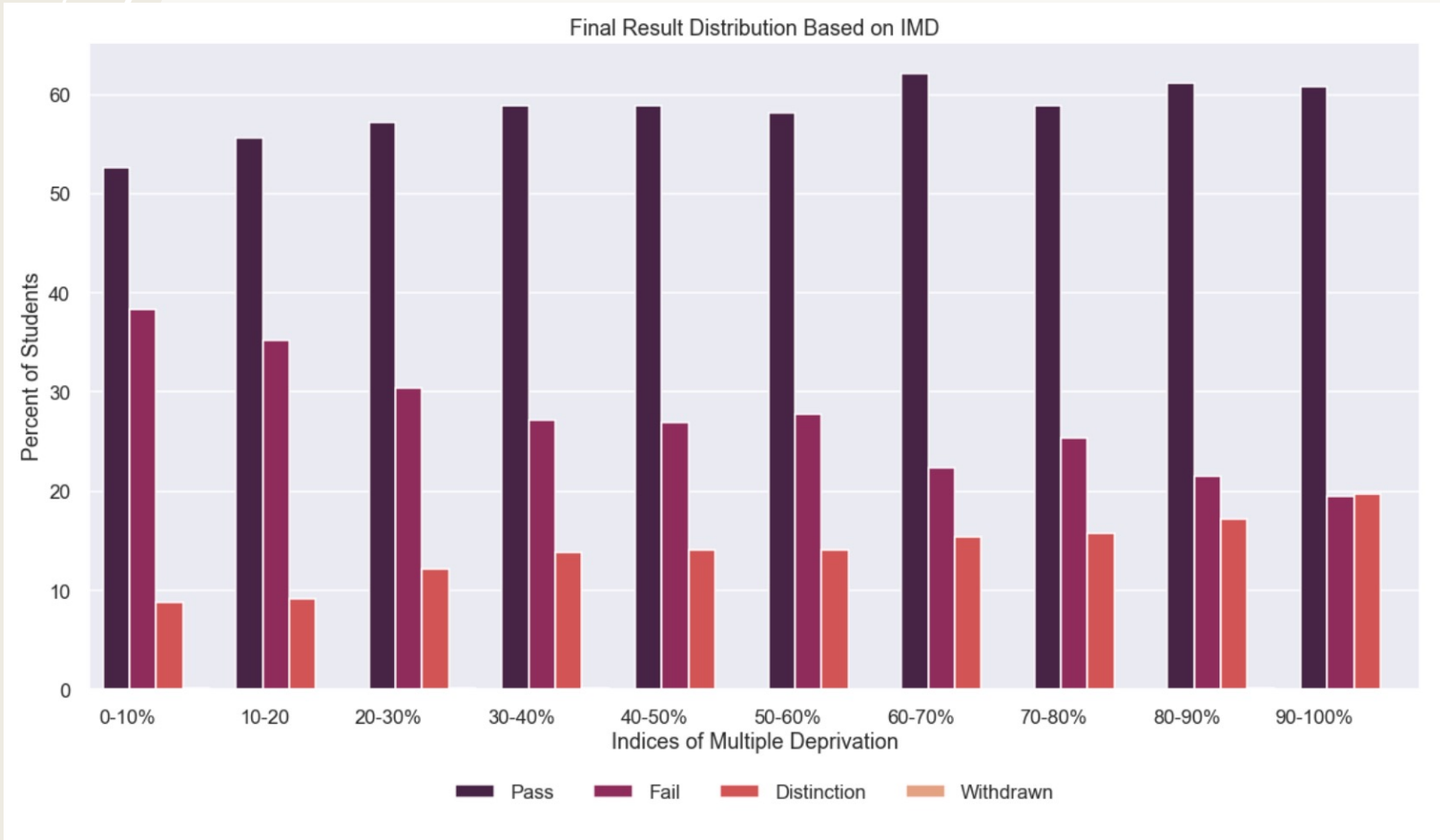
Explorations

- With age getting higher, the rate of passing gets high also



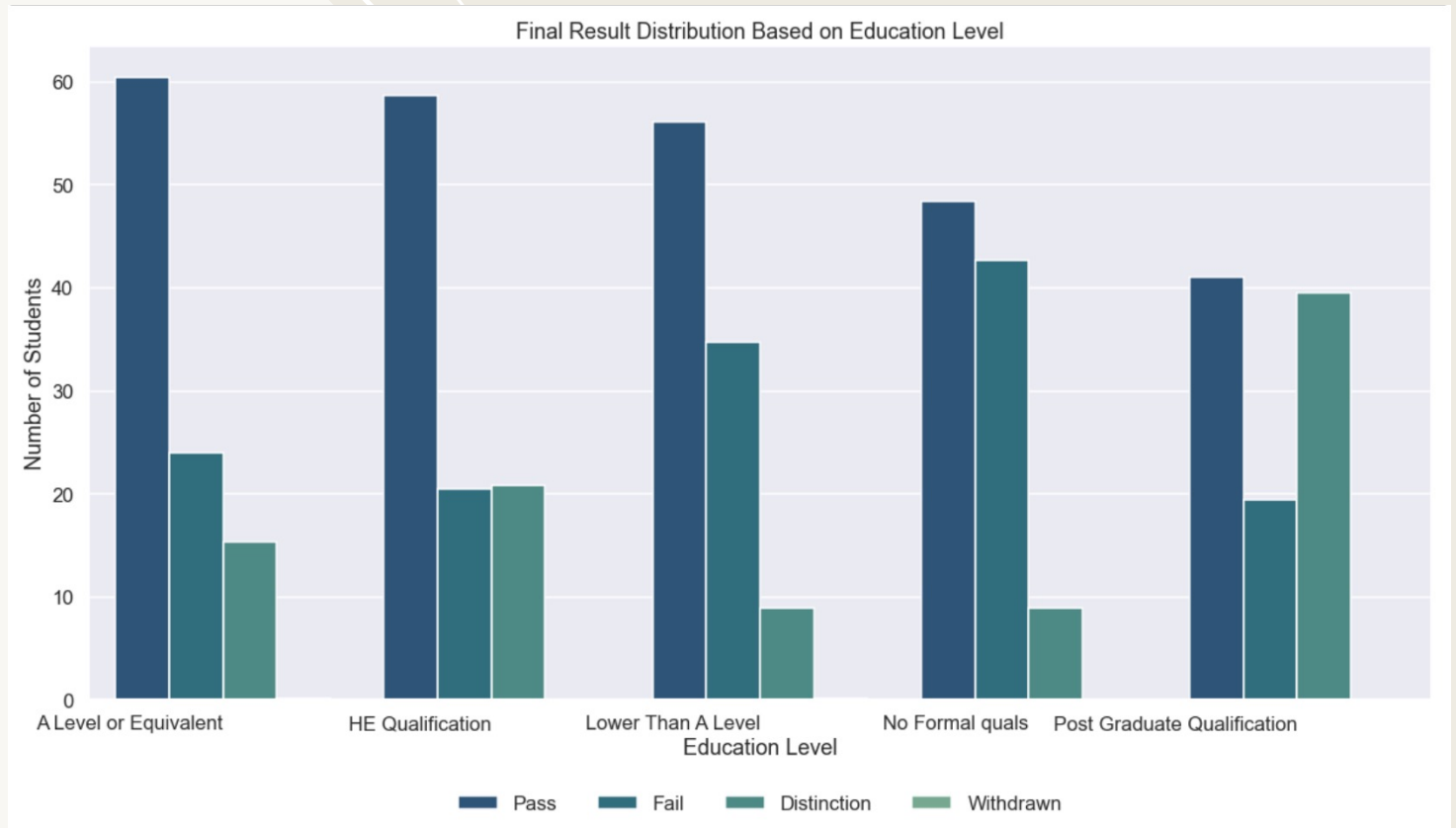
Exploration

- In all groups, the passing percentage has highest value.
- The lower the level of IMD, the higher is the fail rate.

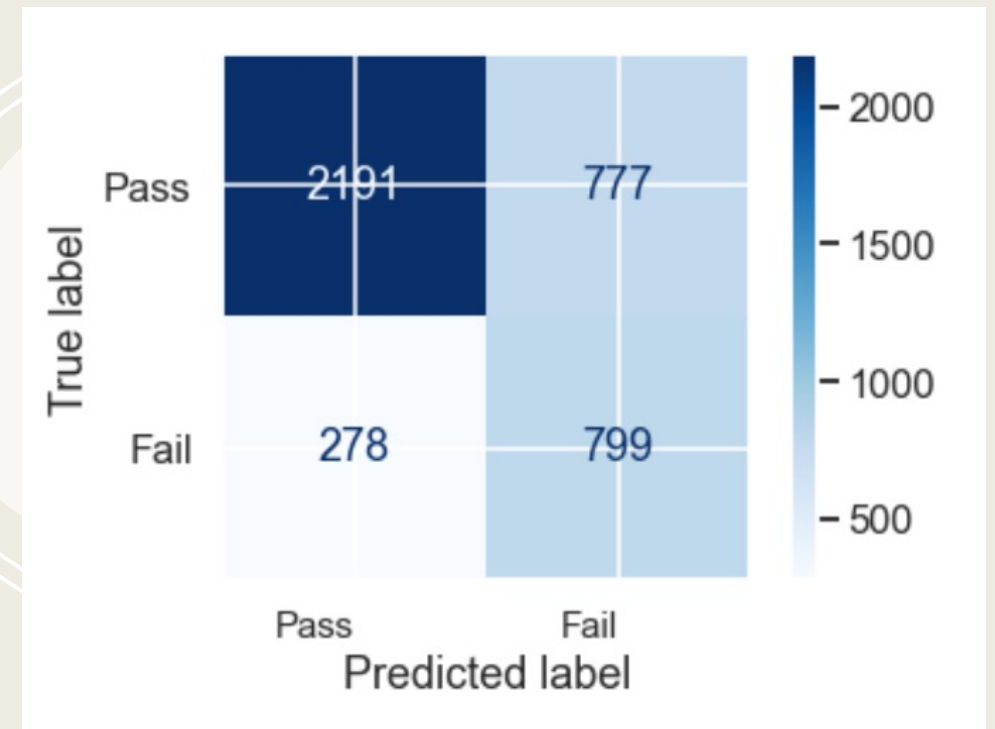
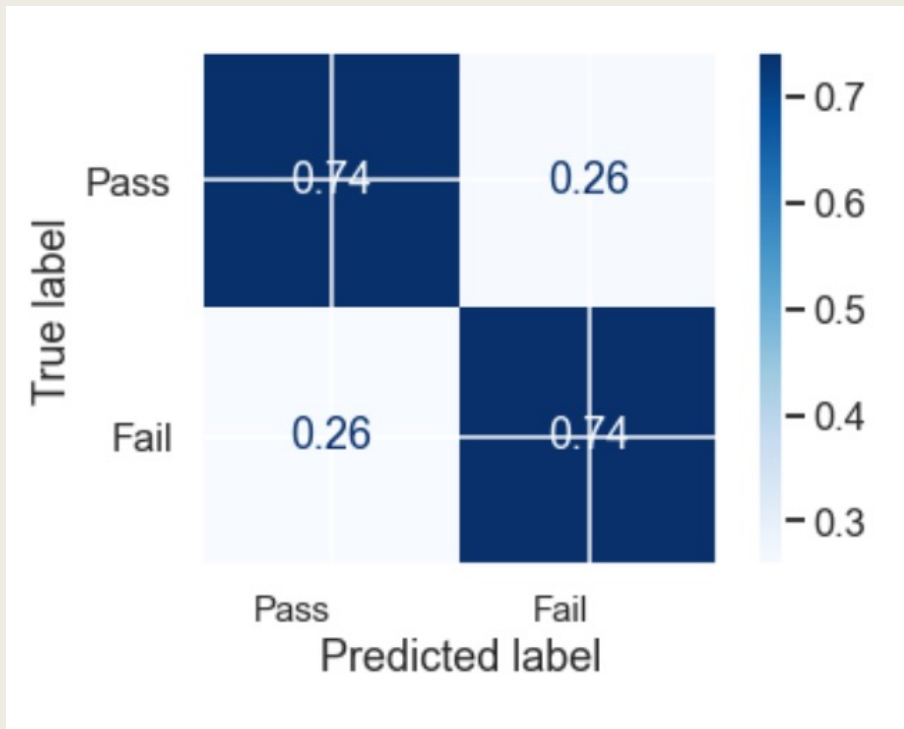


Explorations

- Students with **no formal qualifications** have lowest rate of passing.
- Students with **A level of education or lower** tend to withdraw.



	Pass	Fail	Withdraw	Distinct	Success	Failure
A Level or Equivalent	60.49	24.02	0.11	15.39	75.88	24.13
HE Qualification	58.67	20.47	0.06	20.79	79.46	20.53
Lower Than A Level	56.16	34.79	0.11	8.94	65.1	34.9
No Formal quals	48.41	42.68	0.0	8.92	57.33	42.68
Post Graduate Qualification	41.01	19.42	0.0	39.57	80.58	19.42



Modeling

- Best performed model – **Logistic Regression** with **Recall Score of 0.73** and **Accuracy Score of 0.73**.
- The model predicted **74% of failing students as failing**, and only **26%** were **falsely predicted as passing**.



Provide **additional study sources and tools**.



Advise students to **closely interact with online study materials**



Determine optimal course length.



Provide excessive support for **students** who are in regions **with low IMD**.



Advise students in lower age range to **closely communicate with students that are older**.

Recommendations

Thank You!

GitHub: https://github.com/zyazyaG/Students_Performance

Email: agulyamova14@gmail.com