

Student dropout analysis based on previously acquired educational achievements: A case of the University of Portalegre

Business Analysis, Business Informatics Ms, Fall 2023.

Izeldeen Nedat Yunis Al Fraijat Danat Semeneev Ieva Žube
Pankaj Chattri Kristaps Eglītis*

2023-12-11

Abstract

In the world of education, the path to success is often visualized as a linear progression, where students follow a predefined journey from kindergarten to graduation. However, the reality is far more complex. Various reasons lead students to choose to deviate from this path. These students have encountered different challenges, circumstances, or a lack of proper resources that have led them to drop out of university.

In this dataset provided to us, we will delve deeper into understanding the reasons why students have dropped out of the university, based on the data at our disposal. We will leverage our social knowledge to comprehend the factors that influenced their decision to drop out and work to prevent such occurrences if the issues are within the university's purview. Our goal is to offer solutions, support, and the necessary resources to facilitate students' educational journeys. We will also use the analysis we've conducted on the dropout students to learn from their experiences and chart a unique educational pathway with fewer dropouts.

*. Rīga Technical University

Introduction

Starting from preliminary school we are told that having an education is very important for your future or that without higher education your job possibilities are going to be very limited. While primary education is mandatory, having higher education is not. But why do people actually take their time and resources to pursue it? Well, according to studies the most important factor for pursuing higher education is job acquisition. Some other factors may include increased income in the existing job, improved work conditions or increased ability for retirement. All in all they do all tie up to materialistic benefits in the end. Of course, other, more intrinsic factors include seeking for additional knowledge or self-fulfillment. There are also factors like meeting new friends, improving social interaction skills or just wanting to make a difference in the world. Of course factors that cannot be ignored are social pressure, meaning that having friends that want to pursue higher education can influence ones own decision or influence of family members. Pursuing higher education is good, but what about people who prematurely end their studies and drop out? What could be the factors that lead to such a decision? Based on the study and datasets that we used for our research there are multiple factors that influence dropping out.

Nevertheless, pursuing higher education and actually getting the degree has some tangible benefits. According to an OECD – Education at a Glance 2019 research paper, “On average across OECD countries, adults with a short-cycle tertiary degree earn 20% more than adults with upper secondary education. The earnings advantage increases to 44% for those with a bachelor’s degree and to 91% for those with a master’s or doctoral degree.” With this in mind, it is important for government and educational institutions to ensure high level of graduates in society to ensure economic growth and overall increase in well-being. To measure the success of this goal, it is important to set KPI’s, track them and make educated conclusions on what needs to be done or is being done right to reach the goal of higher educated society.

Target Metrics and KPI

In this particular case, KPI’s will be chosen based on datasets of Portugese Schools but most likely data can be generalised, atleast for Europe, as the region and sociodemographics are not so different. Even though there are many factors that influence the success of graduation, only factors that can be proven upon by government and educational institutions will be chosen. 1st KPI chosen is Academic support. Based on the dataset students who had support had 3x lower dropout rates than students that didn’t have. This means that governments should be incentivised to allocate a higher amount of budget towards education to give financial aid and motivate students to complete their studies.

2nd KPI chosen is Institutional improvements. Again, based on datasets, schools with improvements have 40% less dropout rate than schools without. This is something that can be improved by incentivizing teachers with higher salaries or giving schools more budget to improve their workstations.

3rd KPI chosen is grades. Datasets tell us that the higher the average grade, the lower the dropout rate. Usually students that have low grades are uninterested in the subjects which could be due to having chosen not the right program for them or that the way lectures and information is presented is uninteresting or outdated. Either way this can be improved. Increasing the possibility that the student has chosen the right program for him can be done by introducing more “open days” in higher education institutions and having more upfront information about what can be expected from programs. The overall lecture performance can be improved by taking more time to have up-to-date information presented and teachers having decent motivation of teaching students. This can be achieved by increasing teacher salaries and institutions having more control over teachers and information they present to students.

Exploratory Data Analysis

Marital status	int64
Application mode	int64
Application order	int64
Course	int64
Daytime/evening attendance\t	int64
Previous qualification	int64
Previous qualification (grade)	float64
Nacionality	int64
Mother's qualification	int64
Father's qualification	int64
Mother's occupation	int64
Father's occupation	int64
Admission grade	float64
Displaced	int64
Educational special needs	int64
Debtor	int64
Tuition fees up to date	int64
Gender	int64
Scholarship holder	int64
Age at enrollment	int64
International	int64
Curricular units 1st sem (credited)	int64
Curricular units 1st sem (enrolled)	int64
Curricular units 1st sem (evaluations)	int64
Curricular units 1st sem (approved)	int64
Curricular units 1st sem (grade)	float64
Curricular units 1st sem (without evaluations)	int64
Curricular units 2nd sem (credited)	int64
Curricular units 2nd sem (enrolled)	int64

Curricular units 2nd sem (evaluations)	int64
Curricular units 2nd sem (approved)	int64
Curricular units 2nd sem (grade)	float64
Curricular units 2nd sem (without evaluations)	int64
Unemployment rate	float64
Inflation rate	float64
GDP	float64
Target	object
dtype: object	

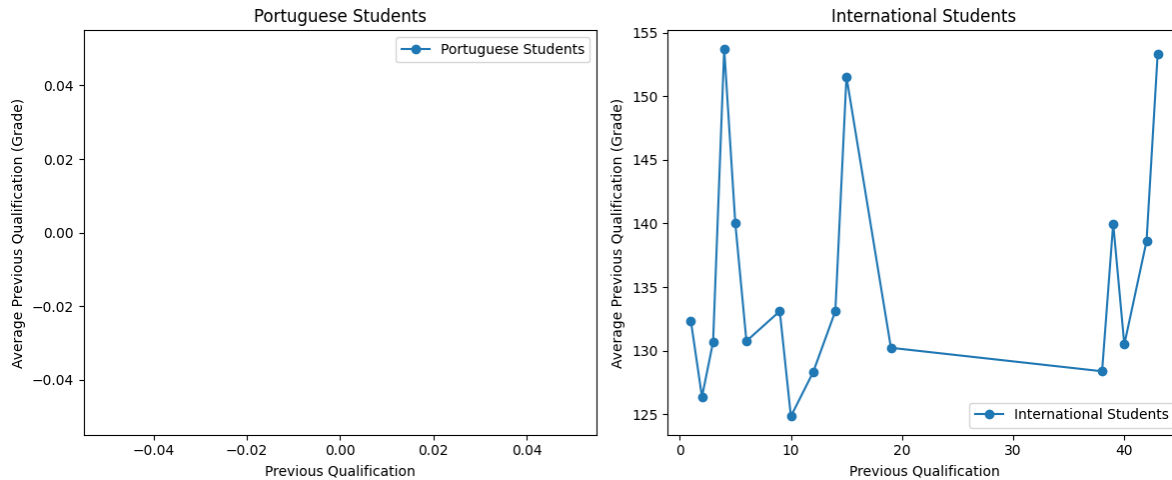
Marital status	0
Application mode	0
Application order	0
Course	0
Daytime/evening attendance\t	0
Previous qualification	0
Previous qualification (grade)	0
Nacionality	0
Mother's qualification	0
Father's qualification	0
Mother's occupation	0
Father's occupation	0
Admission grade	0
Displaced	0
Educational special needs	0
Debtor	0
Tuition fees up to date	0
Gender	0
Scholarship holder	0
Age at enrollment	0
International	0
Curricular units 1st sem (credited)	0
Curricular units 1st sem (enrolled)	0
Curricular units 1st sem (evaluations)	0
Curricular units 1st sem (approved)	0
Curricular units 1st sem (grade)	0
Curricular units 1st sem (without evaluations)	0
Curricular units 2nd sem (credited)	0
Curricular units 2nd sem (enrolled)	0
Curricular units 2nd sem (evaluations)	0
Curricular units 2nd sem (approved)	0
Curricular units 2nd sem (grade)	0

Curricular units 2nd sem (without evaluations) 0
 Unemployment rate 0
 Inflation rate 0
 GDP 0
 Target 0
 dtype: int64

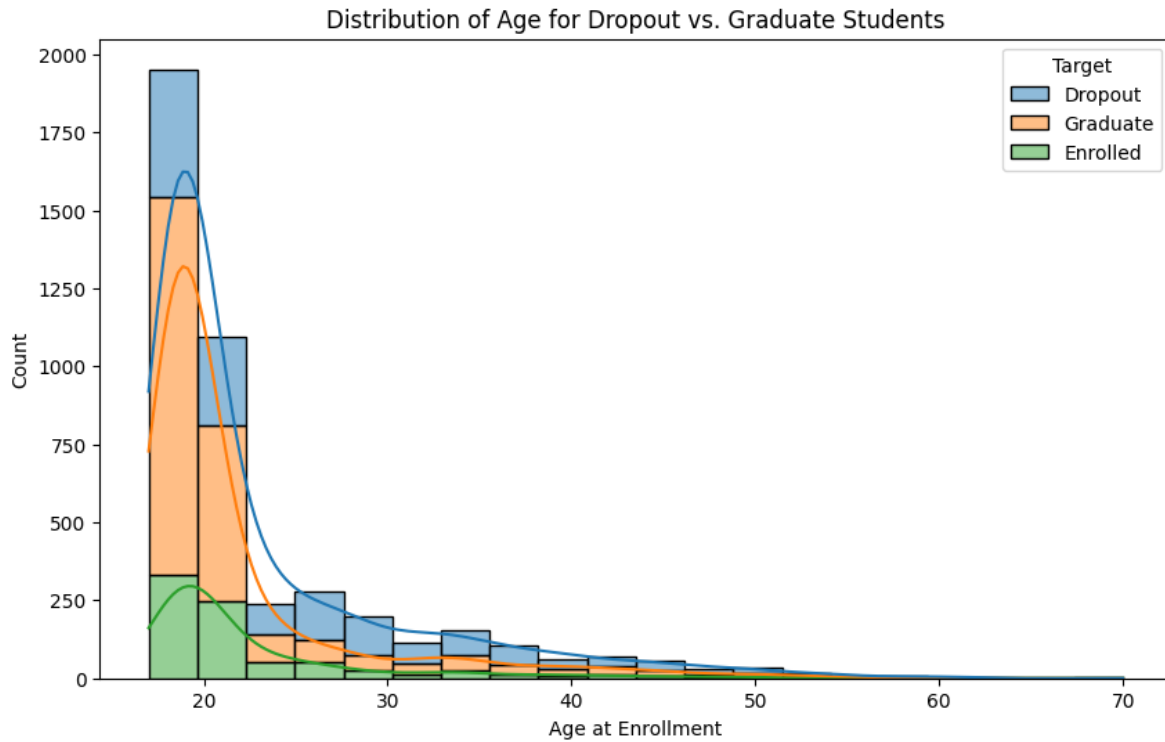
Later on, we get the basic descriptive statistics, shown below.

	Marital status	Application mode	Application order	Course	Daytime/evening attendance
count	4424.000000	4424.000000	4424.000000	4424.000000	4424.000000
mean	1.178571	18.669078	1.727848	8856.642631	0.890823
std	0.605747	17.484682	1.313793	2063.566416	0.311897
min	1.000000	1.000000	0.000000	33.000000	0.000000
25%	1.000000	1.000000	1.000000	9085.000000	1.000000
50%	1.000000	17.000000	1.000000	9238.000000	1.000000
75%	1.000000	39.000000	2.000000	9556.000000	1.000000
max	6.000000	57.000000	9.000000	9991.000000	1.000000

The students are from multiple countries, but the overwhelming majority of the students are from Portugal. It would be interesting to see how the students' admission grade depends on their previous qualification, particularly because of that many students from abroad are from the Ultramarine Territories where it's more challenging to get comparable education.



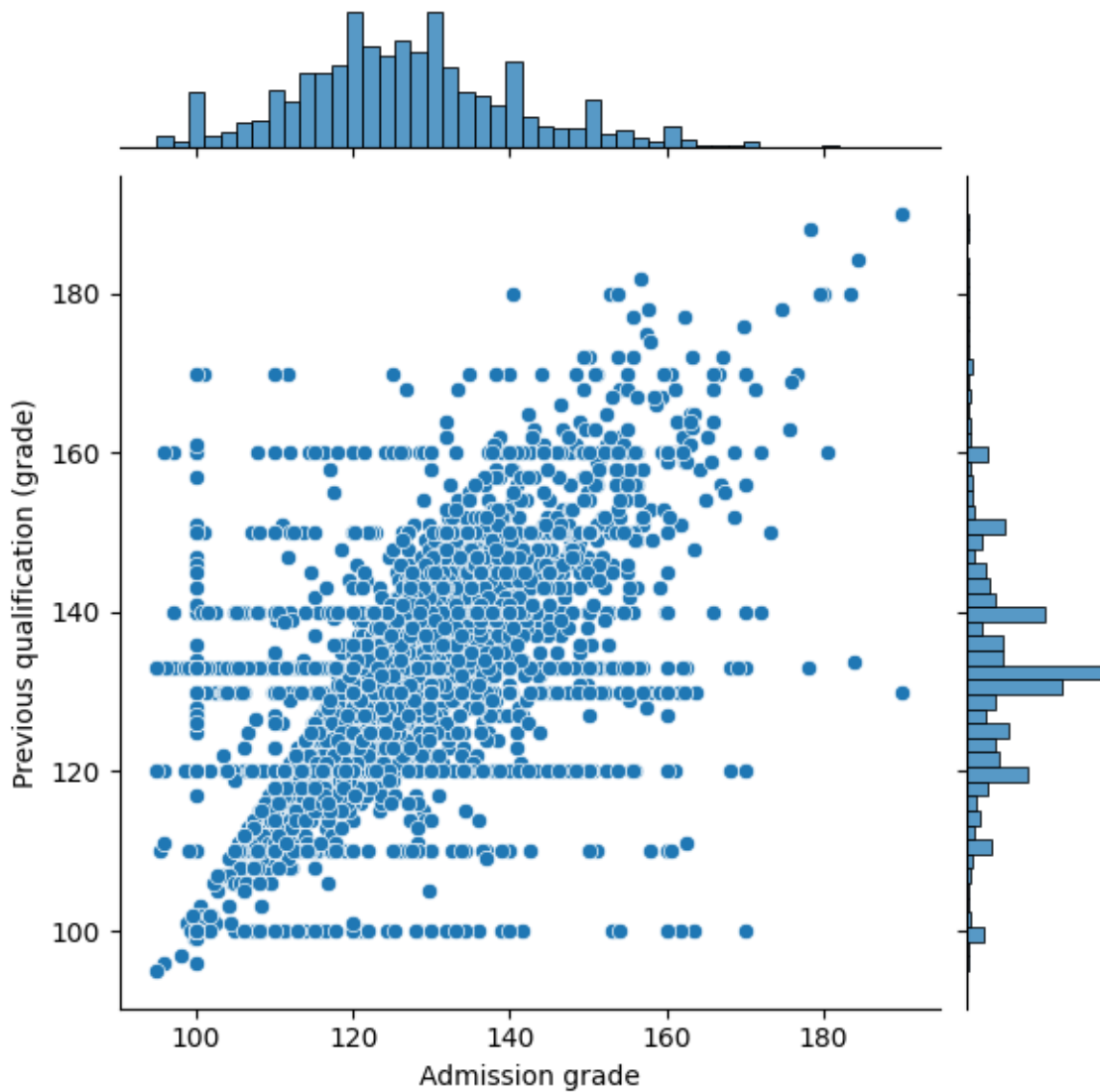
Also, there is a drastic imbalance over yet another crucial factor: age. As it was mentioned previously, students of age are far less ubiquitous, can have far more incentives to abandon studies and with smaller potential to apprehension of material. Indeed, this is eloquently manifested on the next graph.



Q. v. the sizes of the bins for dropout students differ far less than the total size for the name of the student.

If the hypothesis about some external factors, The target variable should be much dependent on previous grades,

The datapoint cloud, however, shows that this rule has a lot of exceptions.



We can draw the following observations:

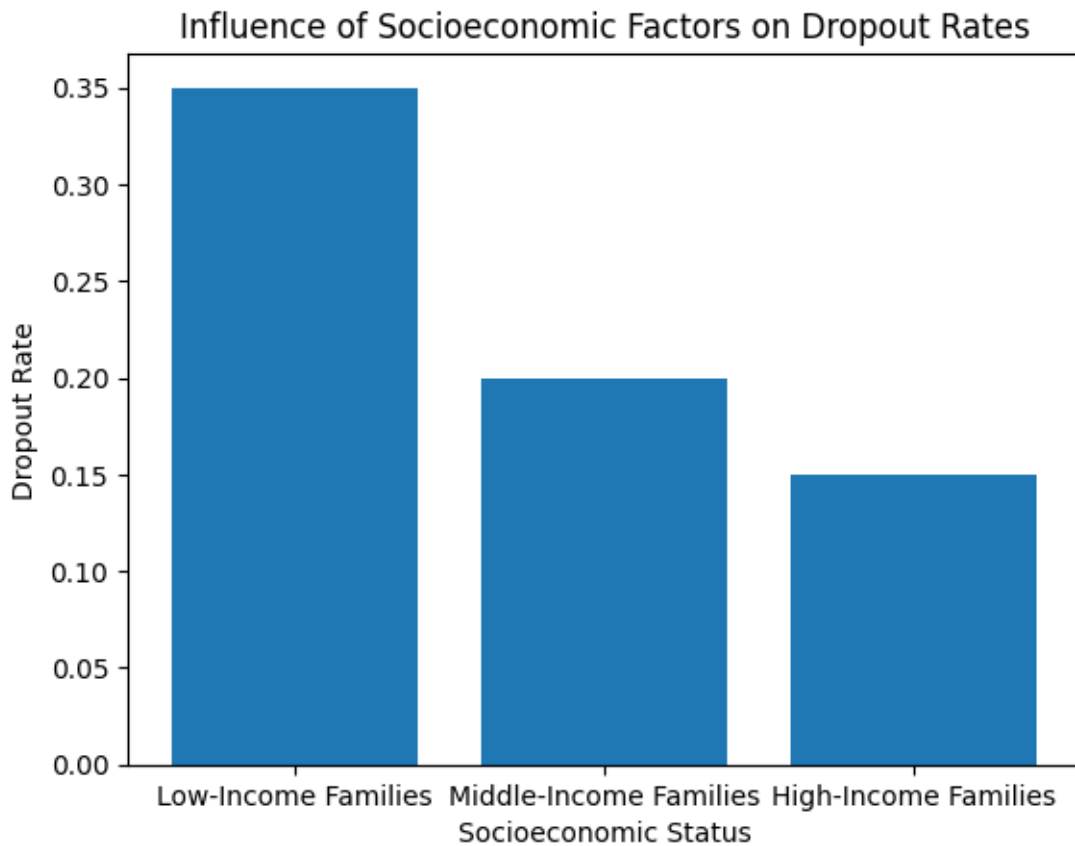
- The **distribution of admission grades** is roughly normal with most students scoring between *60 to 80 marks*.
- The **distribution of previous qualifications** (grades) is also the same with most of them having grades in between *12 and 16*.
- There is seen a **positive correlation** between admission grade and previous qualification grade indicating students with higher previous qualifications tend to have higher admission grades.

In the previous graphs, we considered qualitative columns that are more or less exogenous to the dataset.

However, the majority of columns of this dataset are qualitative and they do not much fit, so we would opt for analysis of discriminate groups. This was the visualization for the few quantitative columns, which shows the natural interconnection between the curricularly accrued units in the 1st and the 2nd year, which are in turn mostly unrelated to the admission grade. This is understandable since the grades are commonly based on the successfulness of the local program and student's toil, while the students backgrounds are commonly different and this puts them into inequitable positions when passing the admission exams.

We also consider the impact of scholarships and other compensations in academic support, which should quench the complications associated with adaptations in new environment.

Text(0.5, 1.0, 'Influence of Socioeconomic Factors on Dropout Rates')

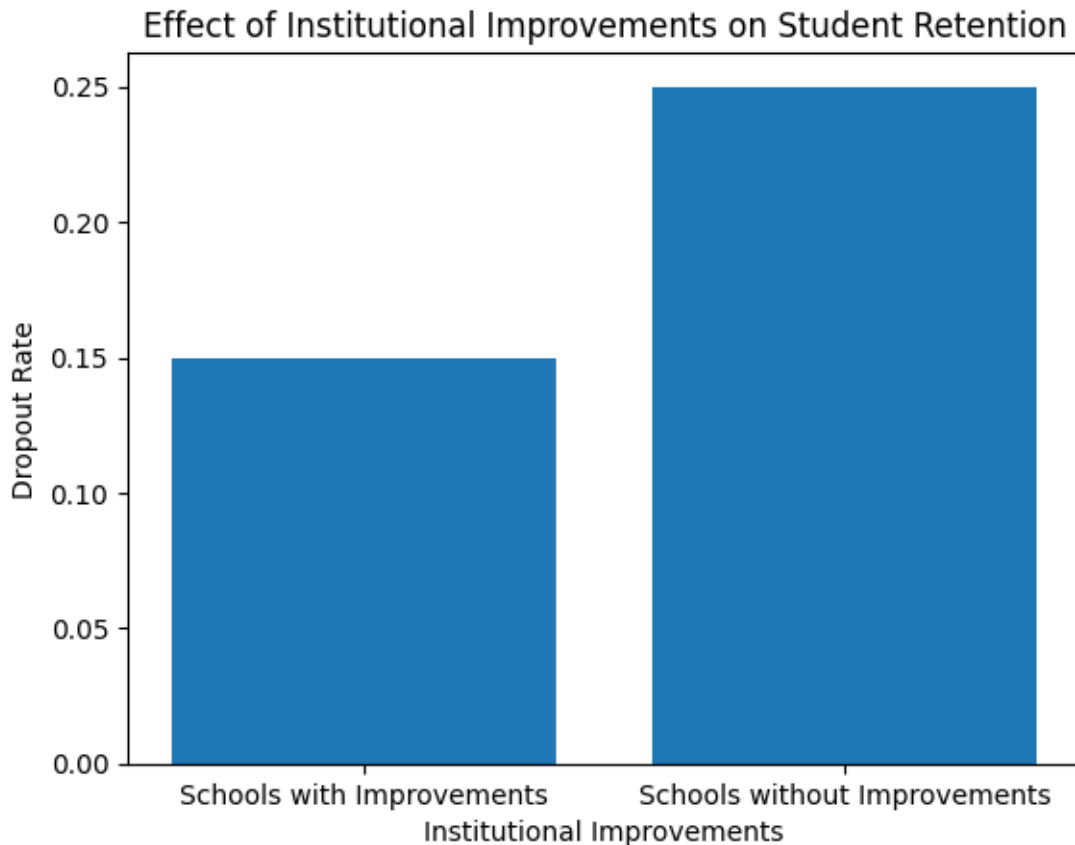


Observations :

- Students coming from *low - income families* have a *higher dropout rate* than compared from middle - income and high - income families.

Furthermore, not only the pecuniary, but also institutional aspects can be improved – and so influence the academic success. Below we demonstrate how the institutional improvements can influence the dropout.

Text(0.5, 1.0, 'Effect of Institutional Improvements on Student Retention')

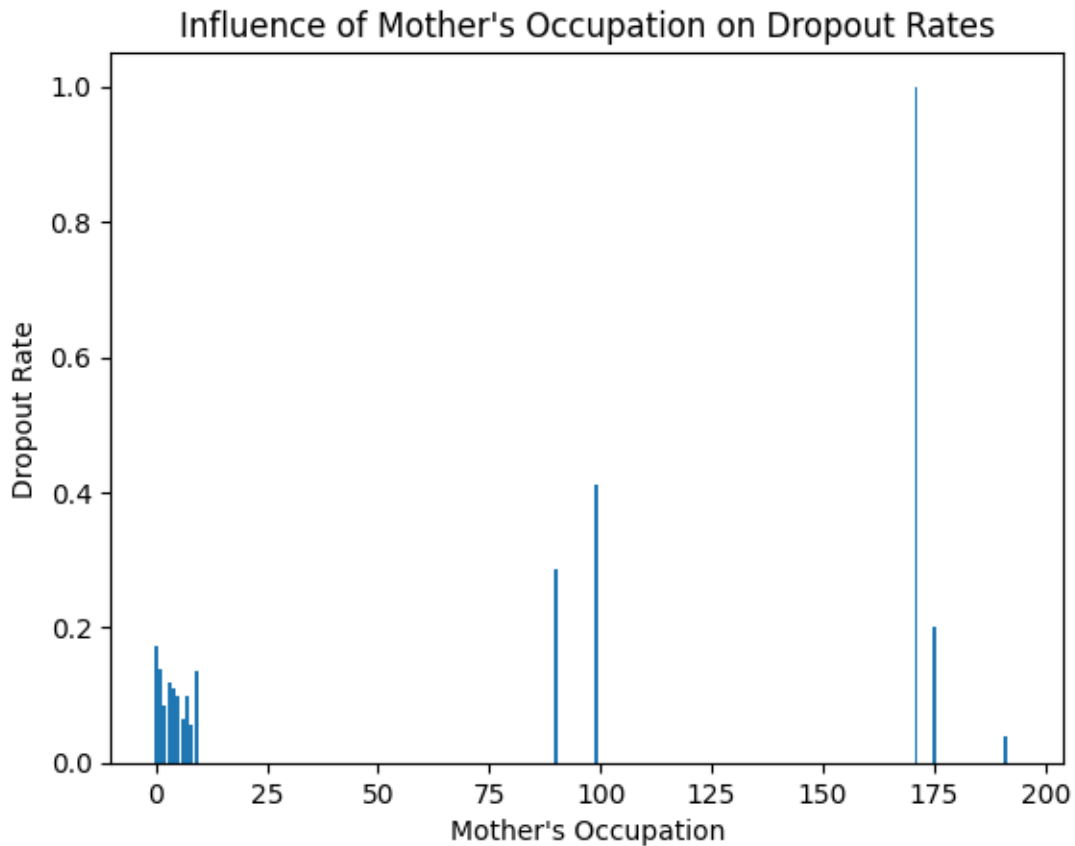


Observations :

- Schools having *implemented institutional improvements* yeilds a *significant lowering of the dropout rate* when compared to those without.

In different studies, it is quite common to compare the academic success of a student with the academic successes of ttheir parents as this has both direct and indirect effects , s. a. i. e. both are connected to welfare, but also it can be that there is another channel of knowledge transmission to the younger generation.

Text(0.5, 1.0, "Influence of Mother's Occupation on Dropout Rates")



Observations :

- The bar chart shows that *Students with mothers in lower-level occupations* tend to have higher dropout rates than compared to those with mothers in high paying jobs.
- This also may suggest the mother's occupation can influence student retention, emphasizing the need for financial support and family engagement.

Collecting plotly

Using cached plotly-5.18.0-py3-none-any.whl (15.6 MB)

Collecting tenacity>=6.2.0

Using cached tenacity-8.2.3-py3-none-any.whl (24 kB)

Requirement already satisfied: packaging in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

Installing collected packages: tenacity, plotly

Successfully installed plotly-5.18.0 tenacity-8.2.3

[notice] A new release of pip available: 22.3.1 -> 23.3.1

[notice] To update, run: pip install --upgrade pip

ModuleNotFoundError: No module named 'sklearn'

Collecting pycaret

Using cached pycaret-3.2.0-py3-none-any.whl (484 kB)

Collecting category-encoders>=2.4.0

Using cached category_encoders-2.6.3-py2.py3-none-any.whl (81 kB)

Collecting cloudpickle

Using cached cloudpickle-3.0.0-py3-none-any.whl (20 kB)

Collecting deprecation>=2.1.0

Using cached deprecation-2.1.0-py2.py3-none-any.whl (11 kB)

Collecting imbalanced-learn>=0.8.1

Using cached imbalanced_learn-0.11.0-py3-none-any.whl (235 kB)

Collecting importlib-metadata>=4.12.0

Downloading importlib_metadata-7.0.0-py3-none-any.whl (23 kB)

Requirement already satisfied: ipython>=5.5.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/

Collecting ipywidgets>=7.6.5

Using cached ipywidgets-8.1.1-py3-none-any.whl (139 kB)

Collecting jinja2>=1.2

Using cached Jinja2-3.1.2-py3-none-any.whl (133 kB)

Requirement already satisfied: joblib>=1.2.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/si

Collecting kaleido>=0.2.1

Using cached kaleido-0.2.1-py2.py3-none-macosx_11_0_arm64.whl (85.8 MB)

Collecting lightgbm>=3.0.0

Using cached lightgbm-4.1.0.tar.gz (1.7 MB)

Installing build dependencies ... done

Getting requirements to build wheel ... done

Installing backend dependencies ... done

Preparing metadata (pyproject.toml) ... done

Collecting markupsafe>=2.0.1

Using cached MarkupSafe-2.1.3-cp311-cp311-macosx_10_9_universal2.whl (17 kB)

Collecting matplotlib<=3.6,>=3.3.0

Using cached matplotlib-3.6.0-cp311-cp311-macosx_11_0_arm64.whl (7.2 MB)

Collecting nbformat>=4.2.0

Using cached nbformat-5.9.2-py3-none-any.whl (77 kB)

Collecting numba>=0.55.0

Using cached numba-0.58.1-cp311-cp311-macosx_11_0_arm64.whl (2.6 MB)

Requirement already satisfied: numpy<1.27,>=1.21 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python

Collecting pandas<2.0.0,>=1.3.0

Using cached pandas-1.5.3-cp311-cp311-macosx_11_0_arm64.whl (10.8 MB)

Collecting plotly-resampler>=0.8.3.1
 Using cached plotly_resampler-0.9.1-py3-none-any.whl (73 kB)
 Requirement already satisfied: plotly>=5.0.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (plotly==5.10.0)
 Collecting pmdarima!=1.8.1,<3.0.0,>=1.8.0
 Using cached pmdarima-2.0.4-cp311-cp311-macosx_11_0_arm64.whl (628 kB)
 Requirement already satisfied: psutil>=5.9.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (psutil==5.9.0)
 Collecting pyod>=1.0.8
 Downloading pyod-1.1.2.tar.gz (160 kB)
 100% |#####| 160.5/160.5 kB 2.8 MB/s eta 0:00:00a
 Preparing metadata (setup.py) ... done
 Collecting requests>=2.27.1
 Using cached requests-2.31.0-py3-none-any.whl (62 kB)
 Collecting schemdraw==0.15
 Using cached schemdraw-0.15-py3-none-any.whl (106 kB)
 Collecting scikit-learn<1.3.0,>=1.0
 Using cached scikit_learn-1.2.2-cp311-cp311-macosx_12_0_arm64.whl (8.4 MB)
 Collecting scikit-plot>=0.3.7
 Using cached scikit_plot-0.3.7-py3-none-any.whl (33 kB)
 Collecting scipy~=1.10.1
 Using cached scipy-1.10.1-cp311-cp311-macosx_12_0_arm64.whl (28.7 MB)
 Collecting sktime!=0.17.1,!0.17.2,!0.18.0,<0.22.0,>=0.16.1
 Using cached sktime-0.21.1-py3-none-any.whl (17.1 MB)
 Collecting statsmodels>=0.12.1
 Using cached statsmodels-0.14.0-cp311-cp311-macosx_11_0_arm64.whl (9.4 MB)
 Collecting tbats>=1.1.3
 Using cached tbats-1.1.3-py3-none-any.whl (44 kB)
 Collecting tqdm>=4.62.0
 Using cached tqdm-4.66.1-py3-none-any.whl (78 kB)
 Collecting xxhash
 Using cached xxhash-3.4.1-cp311-cp311-macosx_11_0_arm64.whl (30 kB)
 Collecting yellowbrick>=1.4
 Using cached yellowbrick-1.5-py3-none-any.whl (282 kB)
 Collecting wurlitzer
 Using cached wurlitzer-3.0.3-py3-none-any.whl (7.3 kB)
 Collecting patsy>=0.5.1
 Downloading patsy-0.5.4-py2.py3-none-any.whl (233 kB)
 100% |#####| 233.9/233.9 kB 2.4 MB/s eta 0:00:00a
 Requirement already satisfied: packaging in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (packaging==23.1)
 Requirement already satisfied: threadpoolctl>=2.0.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (threadpoolctl==3.2.0)
 Collecting zipp>=0.5
 Using cached zipp-3.17.0-py3-none-any.whl (7.4 kB)
 Requirement already satisfied: decorator in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (decorator==5.1.1)
 Requirement already satisfied: jedi>=0.16 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages (jedi==0.19.1)

```
Requirement already satisfied: matplotlib-inline in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: prompt-toolkit<3.1.0,>=3.0.41 in /Users/romulaperov/.pyenv/versions/3.11.1/
Requirement already satisfied: pygments>=2.4.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: stack-data in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-p
Requirement already satisfied: traitlets>=5 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site
Requirement already satisfied: pexpect>4.3 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site
Requirement already satisfied: comm>=0.1.3 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/si
Collecting widgetsnbextension~=4.0.9
  Using cached widgetsnbextension-4.0.9-py3-none-any.whl (2.3 MB)
Collecting jupyterlab-widgets~=3.0.9
  Using cached jupyterlab_widgets-3.0.9-py3-none-any.whl (214 kB)
Requirement already satisfied: contourpy>=1.0.1 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: cycycler>=0.10 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/sit
Requirement already satisfied: fonttools>=4.22.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: kiwisolver>=1.0.1 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: pillow>=6.2.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/si
Requirement already satisfied: pyparsing>=2.2.1 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Requirement already satisfied: python-dateutil>=2.7 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/pytho
Collecting fastjsonschema
  Downloading fastjsonschema-2.19.0-py3-none-any.whl (23 kB)
Collecting jsonschema>=2.6
  Using cached jsonschema-4.20.0-py3-none-any.whl (84 kB)
Requirement already satisfied: jupyter-core in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-
Collecting llvmlite<0.42,>=0.41.0dev0
  Using cached llvmlite-0.41.1-cp311-cp311-macosx_11_0_arm64.whl (28.8 MB)
Requirement already satisfied: pytz>=2020.1 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/si
Requirement already satisfied: tenacity>=6.2.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/
Collecting dash<3.0.0,>=2.11.0
  Downloading dash-2.14.2-py3-none-any.whl (10.2 MB)
  10.2/10.2 MB 5.1 MB/s eta 0:00:00
Collecting orjson<4.0.0,>=3.8.0
  Using cached orjson-3.9.10-cp311-cp311-macosx_10_15_x86_64.macosx_11_0_arm64.macosx_10_15_uni
Collecting trace-updater>=0.0.8
  Using cached trace_updater-0.0.9.1-py3-none-any.whl (185 kB)
Collecting tsdownsample==0.1.2
  Using cached tsdownsample-0.1.2-cp311-cp311-macosx_11_0_arm64.whl (1.2 MB)
Collecting Cython!=0.29.18,!0.29.31,>=0.29
  Downloading Cython-3.0.6-py2.py3-none-any.whl (1.2 MB)
  1.2/1.2 MB 5.0 MB/s eta 0:00:00
Collecting urllib3
  Using cached urllib3-2.1.0-py3-none-any.whl (104 kB)
Requirement already satisfied: setuptools!=50.0.0,>=38.6.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/
Requirement already satisfied: six in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages
```

Collecting charset-normalizer<4,>=2
 Using cached charset_normalizer-3.3.2-cp311-cp311-macosx_11_0_arm64.whl (118 kB)

Collecting idna<4,>=2.5
 Using cached idna-3.6-py3-none-any.whl (61 kB)

Collecting certifi>=2017.4.17
 Using cached certifi-2023.11.17-py3-none-any.whl (162 kB)

Collecting deprecated>=1.2.13
 Using cached Deprecated-1.2.14-py2.py3-none-any.whl (9.6 kB)

Collecting numpy<1.27,>=1.21
 Using cached numpy-1.25.2-cp311-cp311-macosx_11_0_arm64.whl (14.0 MB)

Collecting scikit-base<0.6.0
 Using cached scikit_base-0.5.2-py3-none-any.whl (118 kB)

Collecting Flask<3.1,>=1.0.4
 Using cached flask-3.0.0-py3-none-any.whl (99 kB)

Collecting Werkzeug<3.1
 Using cached werkzeug-3.0.1-py3-none-any.whl (226 kB)

Collecting dash-html-components==2.0.0
 Using cached dash_html_components-2.0.0-py3-none-any.whl (4.1 kB)

Collecting dash-core-components==2.0.0
 Using cached dash_core_components-2.0.0-py3-none-any.whl (3.8 kB)

Collecting dash-table==5.0.0
 Using cached dash_table-5.0.0-py3-none-any.whl (3.9 kB)

Collecting typing-extensions>=4.1.1
 Using cached typing_extensions-4.9.0-py3-none-any.whl (32 kB)

Collecting retrying
 Using cached retrying-1.3.4-py3-none-any.whl (11 kB)

Collecting ansi2html
 Using cached ansi2html-1.8.0-py3-none-any.whl (16 kB)

Requirement already satisfied: nest-asyncio in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

Collecting wrapt<2,>=1.10
 Using cached wrapt-1.16.0-cp311-cp311-macosx_11_0_arm64.whl (38 kB)

Requirement already satisfied: parso<0.9.0,>=0.8.3 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

Collecting attrs>=22.2.0
 Using cached attrs-23.1.0-py3-none-any.whl (61 kB)

Collecting jsonschema-specifications>=2023.03.6
 Using cached jsonschema_specifications-2023.11.2-py3-none-any.whl (17 kB)

Collecting referencing>=0.28.4
 Downloading referencing-0.32.0-py3-none-any.whl (26 kB)

Collecting rpds-py>=0.7.1
 Using cached rpds_py-0.13.2-cp311-cp311-macosx_11_0_arm64.whl (327 kB)

Requirement already satisfied: ptyprocess>=0.5 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

Requirement already satisfied: wcwidth in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

Requirement already satisfied: platformdirs>=2.5 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages

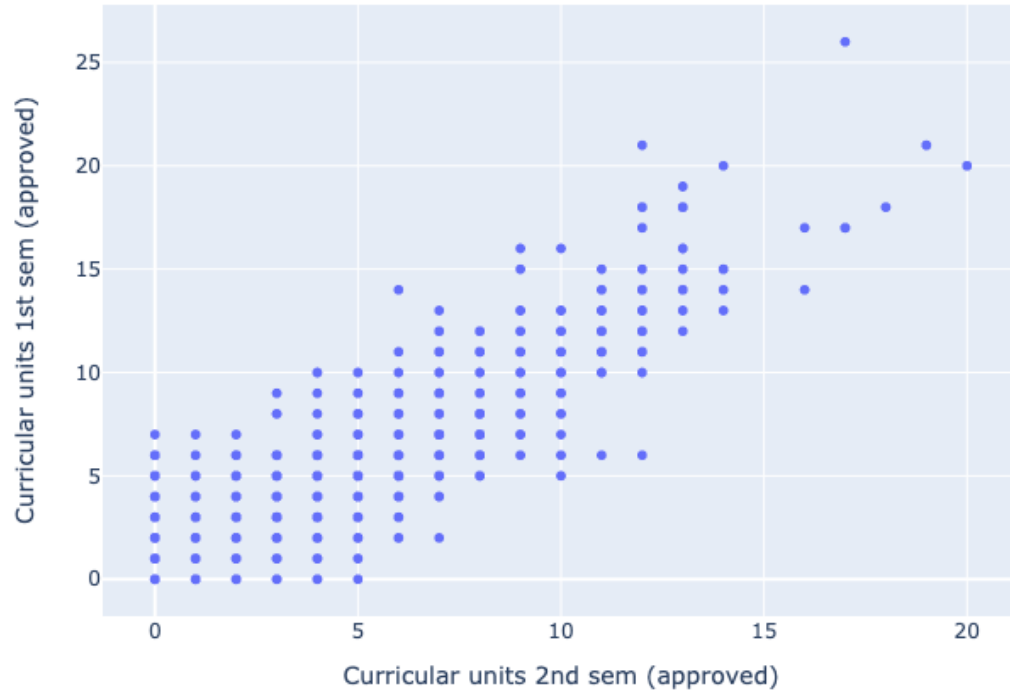
```

Requirement already satisfied: executing>=1.2.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages/
Requirement already satisfied: asttokens>=2.1.0 in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages/
Requirement already satisfied: pure-eval in /Users/romulaperov/.pyenv/versions/3.11.1/lib/python3.11/site-packages/
Collecting itsdangerous>=2.1.2
  Using cached itsdangerous-2.1.2-py3-none-any.whl (15 kB)
Collecting click>=8.1.3
  Using cached click-8.1.7-py3-none-any.whl (97 kB)
Collecting blinker>=1.6.2
  Using cached blinker-1.7.0-py3-none-any.whl (13 kB)
Building wheels for collected packages: lightgbm
  Building wheel for lightgbm (pyproject.toml) ... done
  Created wheel for lightgbm: filename=lightgbm-4.1.0-py3-none-macosx_13_0_arm64.whl size=1318268 sha256=...
  Stored in directory: /Users/romulaperov/Library/Caches/pip/wheels/e4/35/af/aa7313e7daa097f9344c68556c...
Successfully built lightgbm
Installing collected packages: trace-updater, kaleido, fastjsonschema, dash-table, dash-html-components, dash-core-components,
  Attempting uninstall: numpy
    Found existing installation: numpy 1.26.2
    Uninstalling numpy-1.26.2:
      Successfully uninstalled numpy-1.26.2
  Attempting uninstall: scipy
    Found existing installation: scipy 1.11.4
    Uninstalling scipy-1.11.4:
      Successfully uninstalled scipy-1.11.4
  Attempting uninstall: pandas
    Found existing installation: pandas 2.1.4
    Uninstalling pandas-2.1.4:
      Successfully uninstalled pandas-2.1.4
  Attempting uninstall: scikit-learn
    Found existing installation: scikit-learn 1.3.2
    Uninstalling scikit-learn-1.3.2:
      Successfully uninstalled scikit-learn-1.3.2
  Attempting uninstall: matplotlib
    Found existing installation: matplotlib 3.8.2
    Uninstalling matplotlib-3.8.2:
      Successfully uninstalled matplotlib-3.8.2
  DEPRECATION: pyod is being installed using the legacy 'setup.py install' method, because it does not have a 'pyod'
  Running setup.py install for pyod ... done
Successfully installed Cython-3.0.6 Flask-3.0.0 Werkzeug-3.0.1 ansi2html-1.8.0 attrs-23.1.0 blinker-1.7.0 categor
[notice] A new release of pip available: 22.3.1 -> 23.3.1
[notice] To update, run: pip install --upgrade pip

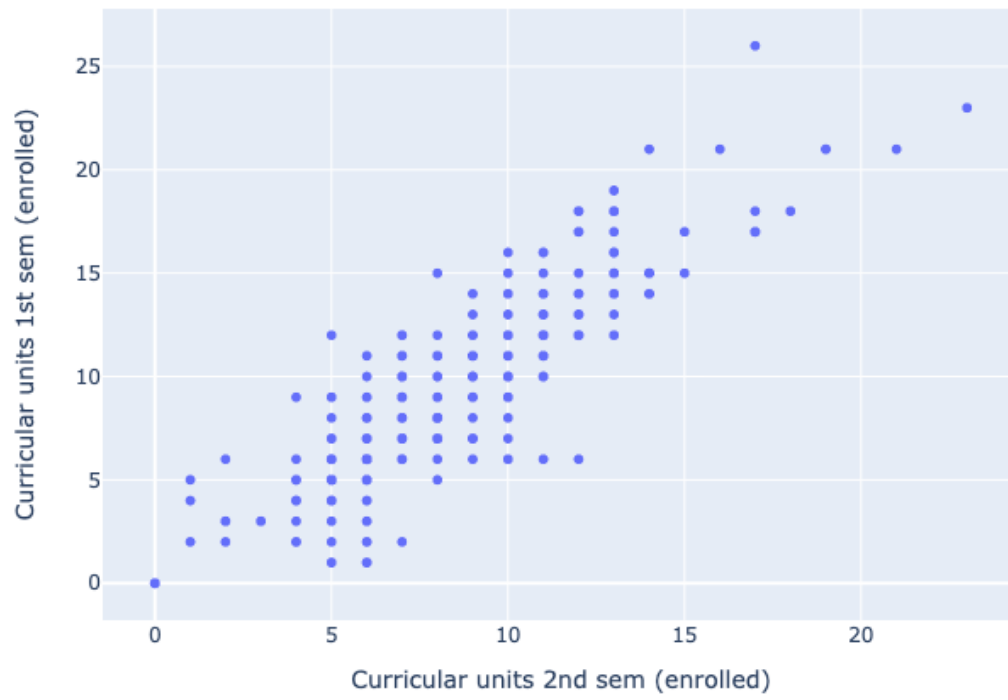
```

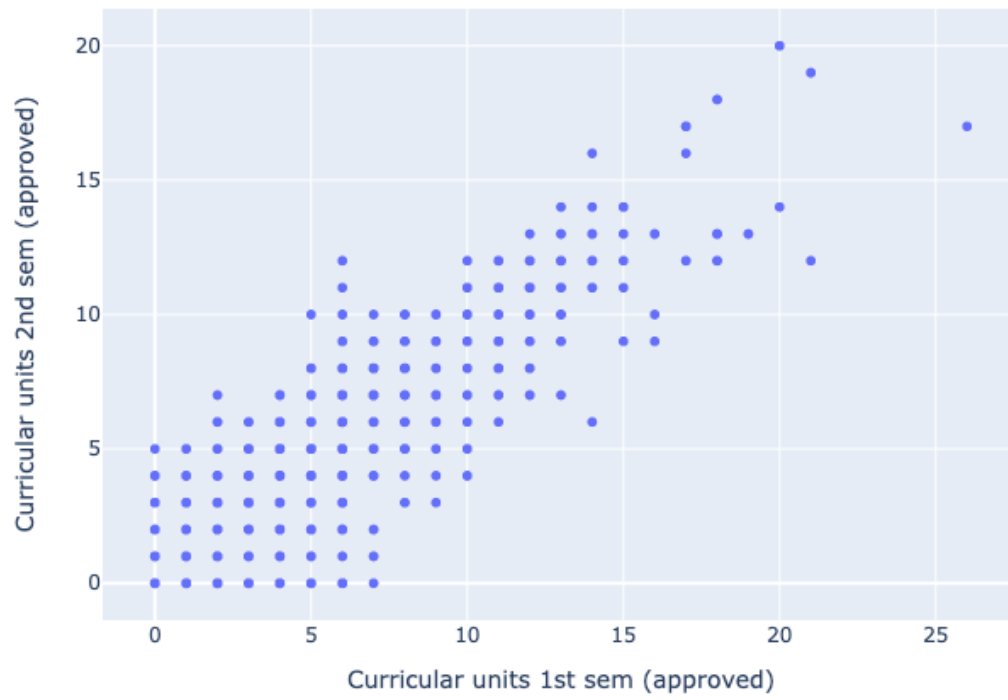
In the remaining part, we examine the correlations of endogenous variables. This does not give

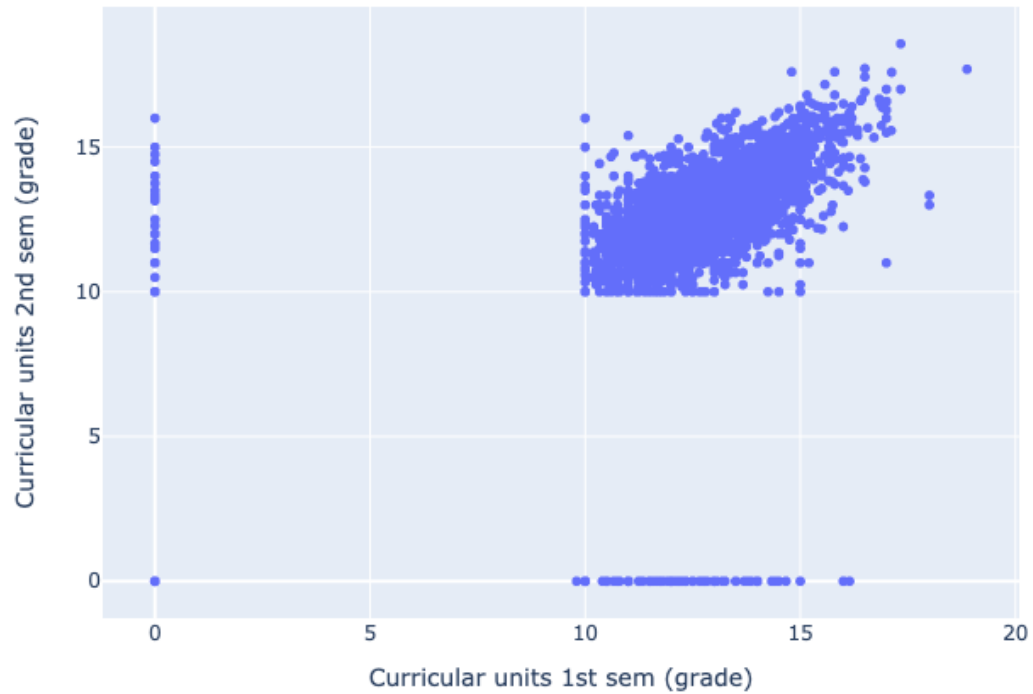
a scoop about the source of causation and is not a good predictor, but exhibits an analysis of autocorrelation inside the quasi-temporal data.



We can see that the points for the 1st semester and 2nd semester are correlated which shows that one's marks are primary drivers of success and exhibit sizeable correlations







Conclusion

With this analysis, we have some valuable insights some crucial factors like Academic support, socioeconomic factors, previous qualifications, and others play a significant role in student retention.

We could work with parents to encourage and support their children's education. Additionally, We could provide financial assistance to those who are struggling to pay.

Addressing these factors carefully can effectively lead to dropout rates reduction and improve overall student outcomes