Machine Learning Models Deployment with IBM Cloud Watson Studio

Problem Statement:

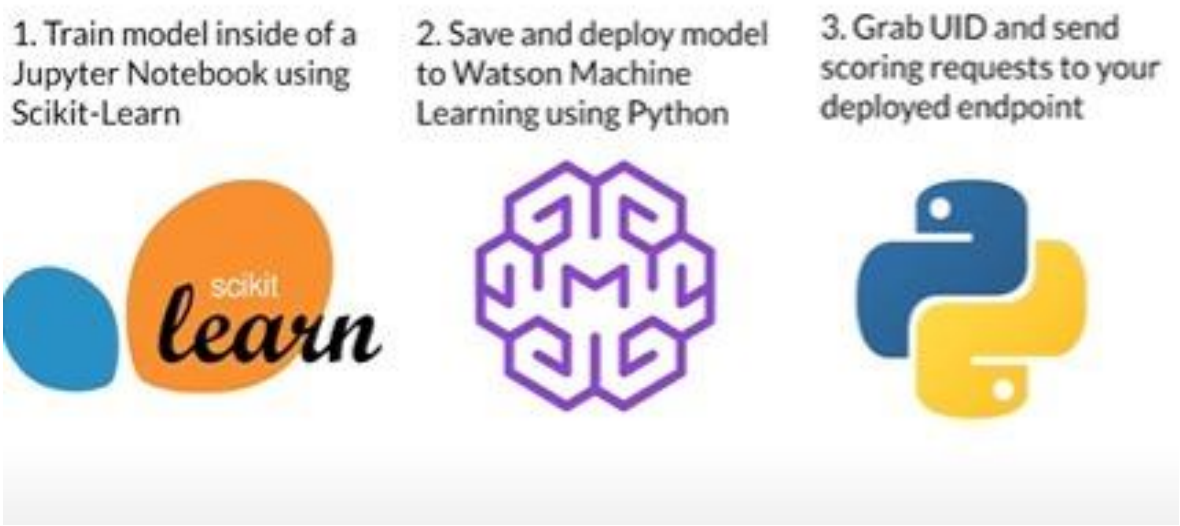
Become a wizard of predictive analytics with IBM Cloud Watson Studio. Train machine learning models to predict outcomes in real time. Deploy the models as web services and integrate them into your applications. Unlock the magic of data-driven insights and make informed decisions like never before!

Project Title: Machine Learning Model Deployment for Student Dropout Analysis with IBM Cloud Watson Studio

Problem Definition:

Student dropout rates in educational institutions have become a significant concern. Institutions invest substantial resources in nurturing students' potential, and when students drop out, it not only affects their futures but also leads to the underutilization of educational resources. The problem we aim to address is:

Approach Used:



Step 1 - Setting Up IBM Cloud and IBM Watson Machine Learning:

- Authenticate and set up an IBM Cloud account.
- Navigate to the IBM Watson Machine Learning catalog.
- Select a third-party service or app to function with Watson Machine Learning.
- Use an existing warehouse and the IBM Watson Machine Learning catalog interface.
- Suggest selecting a free plan and setting up Watson Studio.
- Set up a new deployment space, associate it with the machine learning service, and save and deploy the model.

Step 2 - Deploying with Python:

- Set up an API key and credentials for different regions
- Create a Python machine learning client using the IBM Watson Machine Learning API key and the api_client class.
- List deployment spaces and set a default deployment space.
- Explain how to save and deploy the model, specifying the model name, deployment name, and best model.

Step 3 - Model Deployment with Metadata:

- Output the model and save it to a machine learning environment.
- Specify model metadata using a dictionary and pass it through deployment.
- Save the model in the Watson Machine Learning repository.
- Pass keyword parameters, including model name, type, software specification UID, training data, and training target.
- Record the model ID for automation purposes.

Step 4 - Deploying a Model:

- Obtain the model's UID saved in a deployment space.
- Specify deployment underprop and use the command wml_client.deployment.create to deploy the saved model.
- Set the artifact_uid keyword parameter to the model's ID.
- Pass deployment properties in quantum format.

Step 5 - Scoring the Deployed Model:

- Explain how to score the deployed model.
- Grab the deployment UID and send data from a pandas data frame to it.
- Specify wml_client.deployments.get_uid and payload to make the API request.
- The payload contains input data in the form of an array representing the columns and values of the data frame.

Project Progression:

PHASE 3 - PROJECT DEVELOPMENT PART-1

- Uploading the Dataset
- Dataset cleansing
- Exploratory data analysis

PROJECT DEVELOPMENT PART-2

- Model selection

FINAL PHASE OF THE PROJECT

- Deployment of the model in IBM Cloud Watson Studio

LINK TO THE DATASET

https://docs.google.com/spreadsheets/d/1yg9IjaccCUlX347lQQGEYqVblfF-Mq8ITA4U_HbfGv4/edit?usp=sharing

About the Dataset:

This dataset offers a comprehensive perspective on students enrolled in diverse undergraduate programs provided by a higher education institution. It encompasses a wide array of information, including demographic details, socioeconomic factors, and academic performance metrics. The dataset is instrumental in conducting an analysis to identify potential predictors of student attrition and academic achievement.

The dataset is divided into multiple separate databases, each containing pertinent information recorded at the time of enrollment. This information includes details such as the application mode, marital status, chosen courses, and more. Furthermore, the dataset enables the assessment of students' overall performance at the end of each semester, considering factors like the curricular units they completed, enrolled in, evaluated, or had approved, along with the corresponding grades.

In addition to academic data, the dataset also provides economic indicators, such as the unemployment rate, inflation rate, and regional GDP. These economic factors contribute to a holistic understanding of how they might influence student attrition rates and academic outcomes.

This dataset is a valuable analytical tool, offering insights into the factors that motivate students to persist in their studies or discontinue their educational pursuits across a diverse range of academic disciplines, including agronomy, design, education, nursing, journalism, management, social services, and technology.

Column:

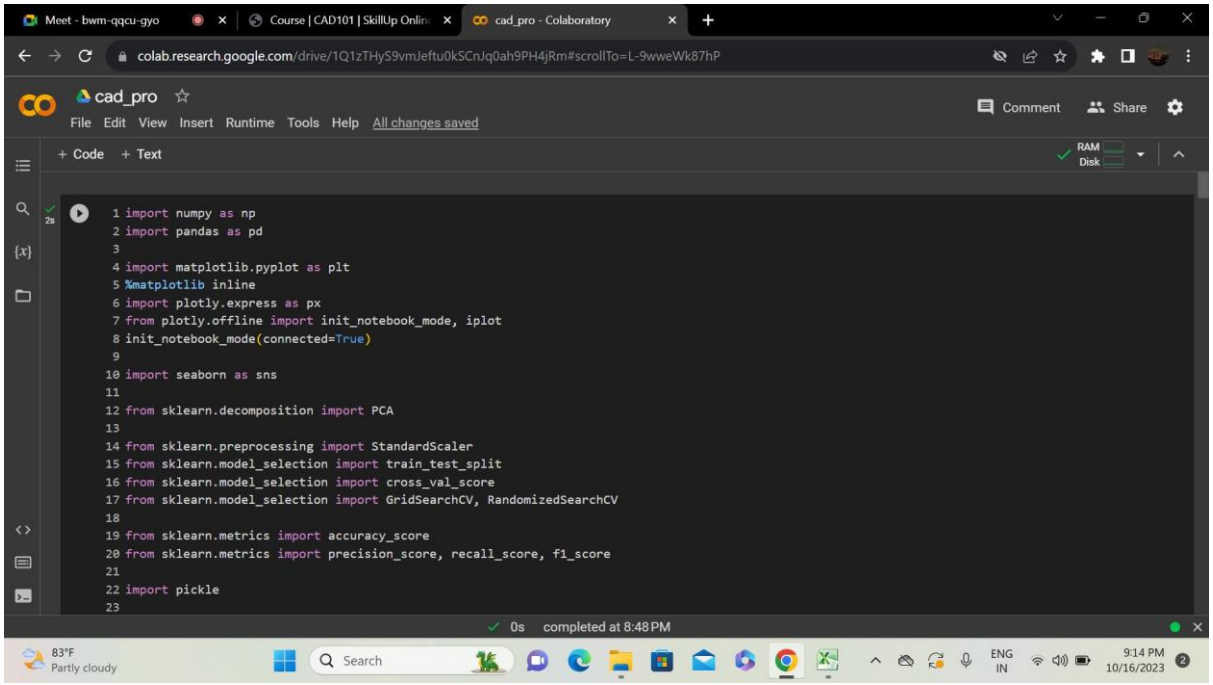
- Marital status: The marital status of the student. (Categorical)
- Application mode: The method of application used by the student. (Categorical)
- Application order: The order in which the student applied. (Numerical)
- Course: The course taken by the student. (Categorical)
- Daytime/evening attendance: Whether the student attends classes during the day or in the evening. (Categorical)

- Previous qualification: The qualification obtained by the student before enrolling in higher education. (Categorical)
- Nationality: The nationality of the student. (Categorical)
- Mother's qualification: The qualification of the student's mother. (Categorical)
- Father's qualification: The qualification of the student's father. (Categorical)
- Mother's occupation: The occupation of the student's mother. (Categorical)
- Father's occupation: The occupation of the student's father. (Categorical)
- Displaced: Whether the student is a displaced person. (Categorical)
- Educational special needs: Whether the student has any special educational needs. (Categorical)
- Debtor: Whether the student is a debtor. (Categorical)
- Tuition fees up to date: Whether the student's tuition fees are up to date. (Categorical)
- Gender: The gender of the student. (Categorical)
- Scholarship holder: Whether the student is a scholarship holder. (Categorical)
- Age at enrollment: The age of the student at the time of enrollment. (Numerical)
- International: Whether the student is an international student. (Categorical)
- Curricular units 1st sem (credited): The number of curricular units credited by the student in the first semester. (Numerical)
- Curricular units 1st sem (enrolled): The number of curricular units enrolled by the student in the first semester. (Numerical)

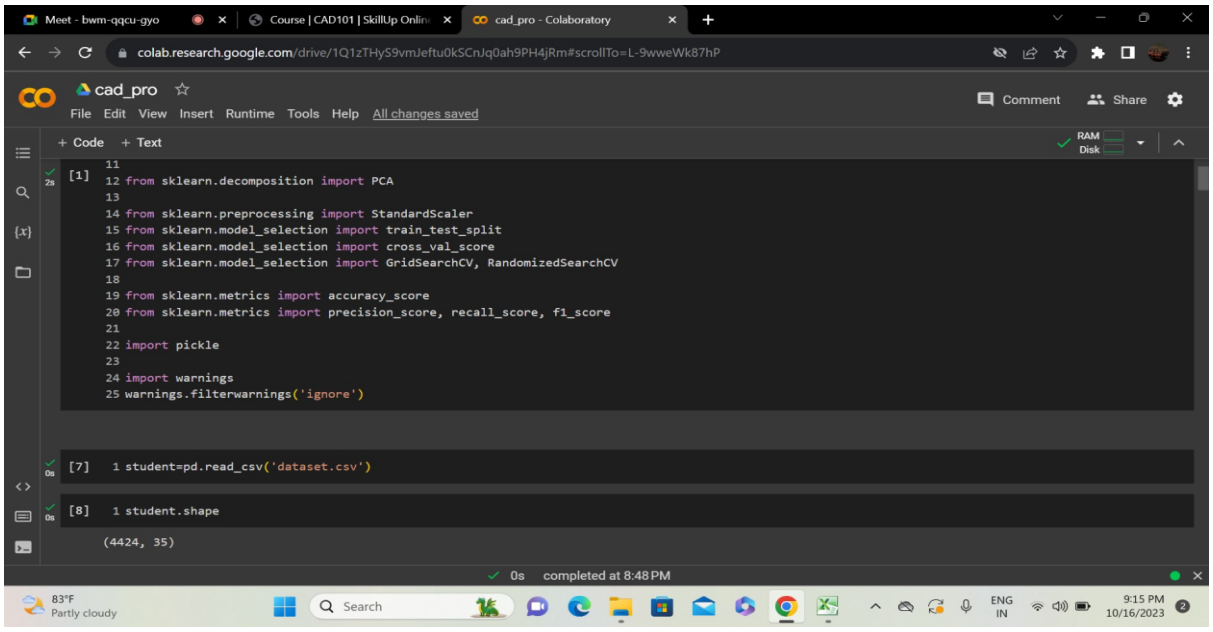
- Curricular units 1st sem (evaluations): The number of curricular units evaluated by the student in the first semester. (Numerical)
- Curricular units 1st sem (approved): The number of curricular units approved by the student in the first semester. (Numerical)

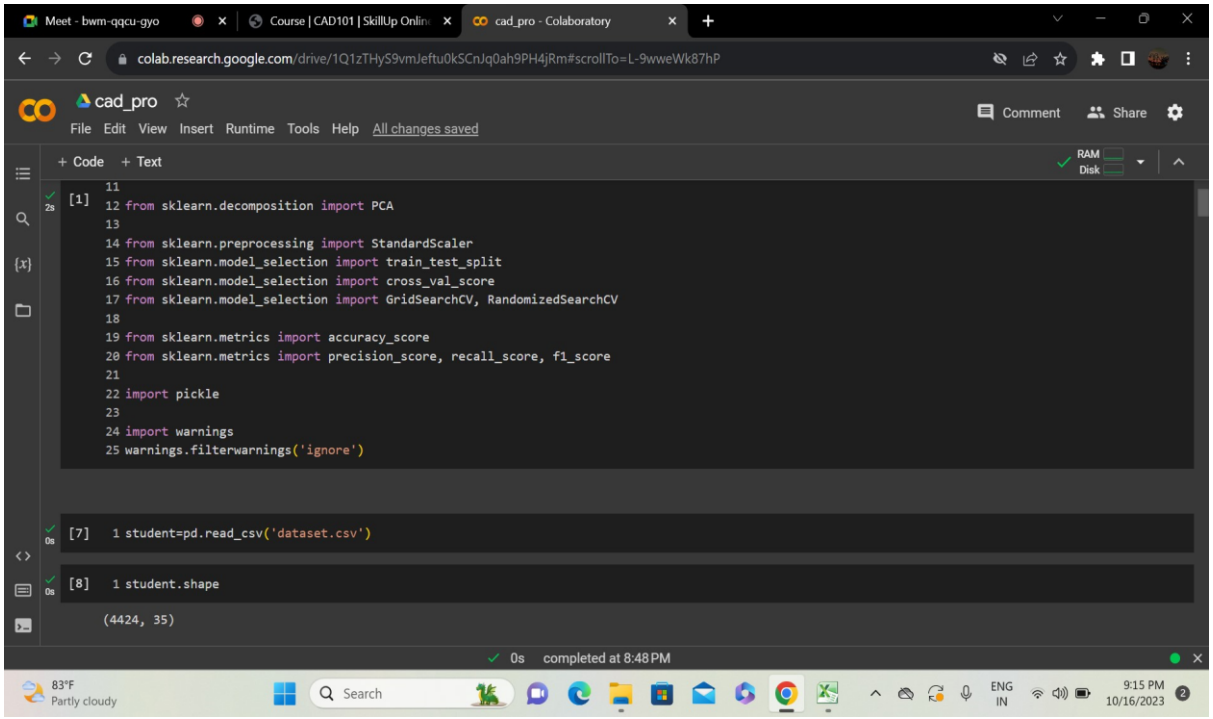
Code:

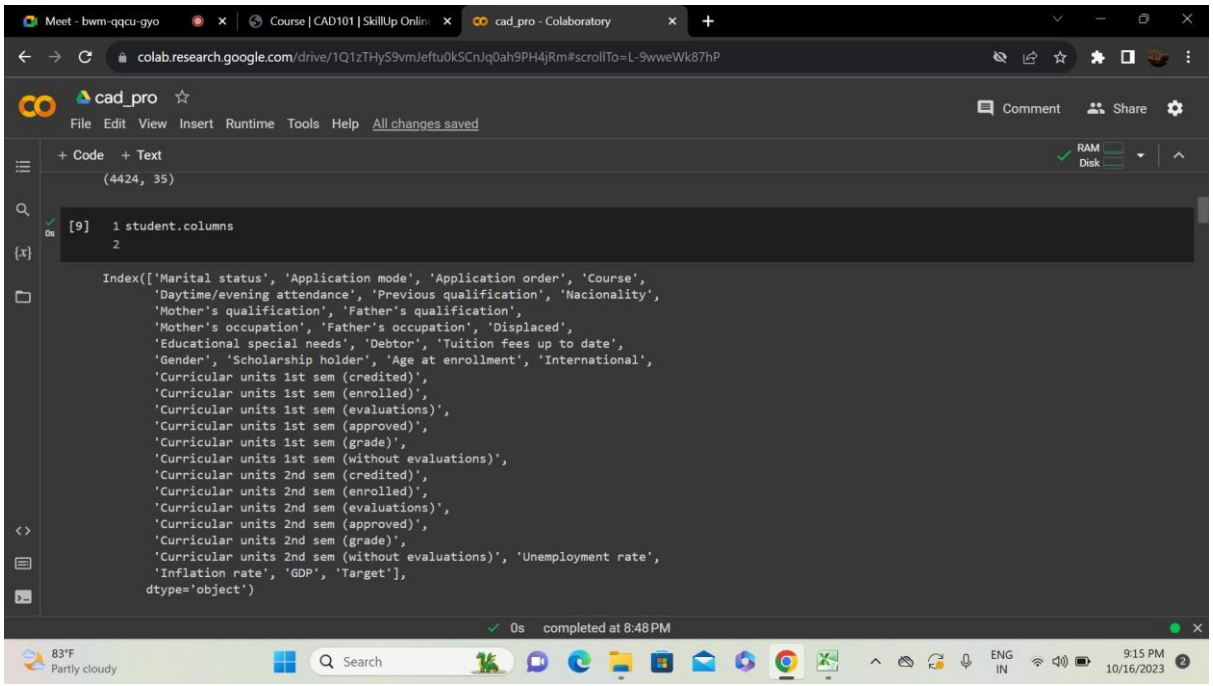
Importing Libraries:

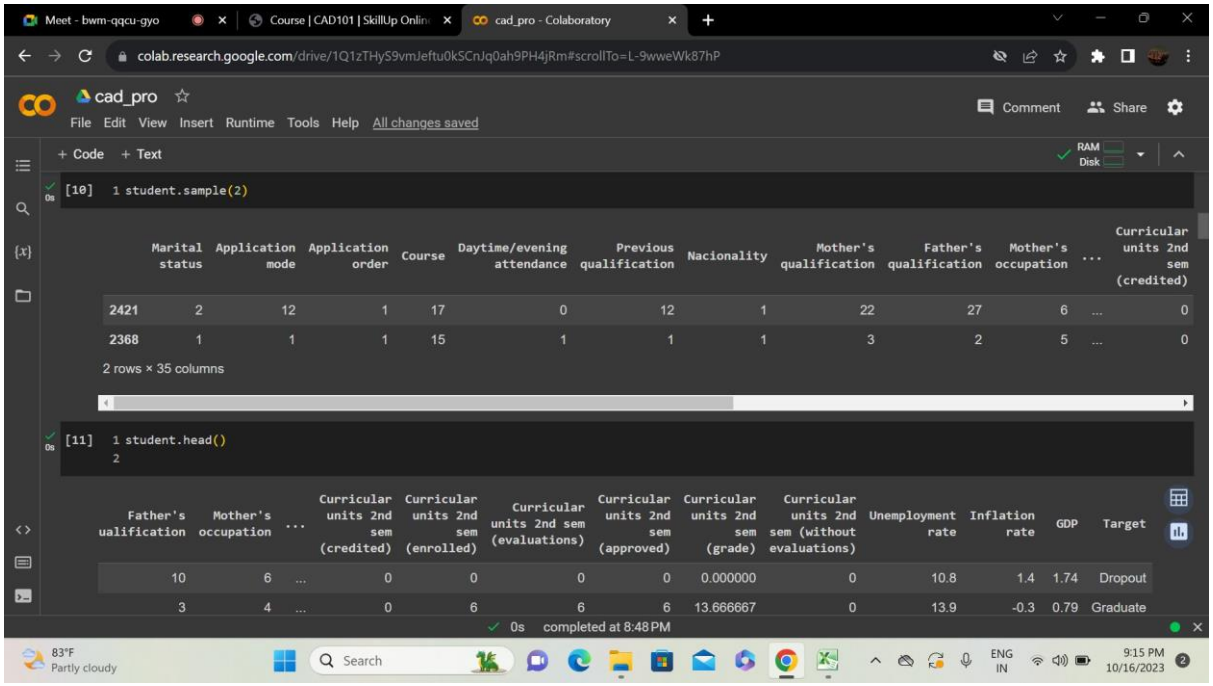


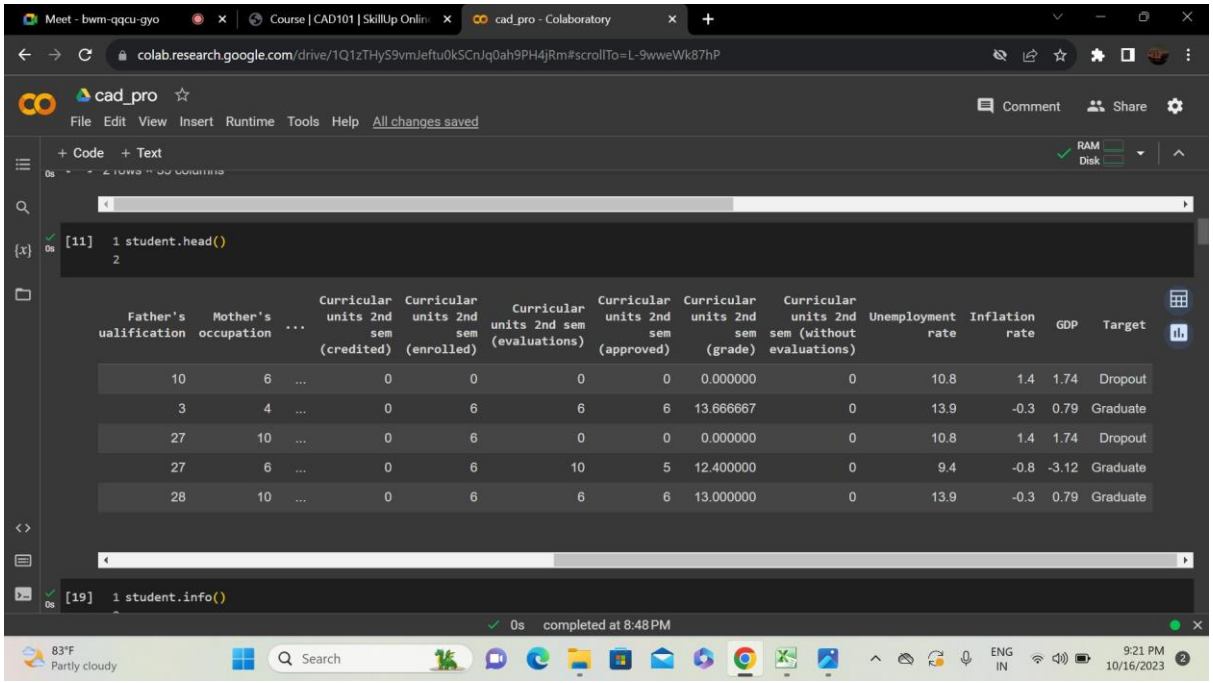
Understanding the datasets:

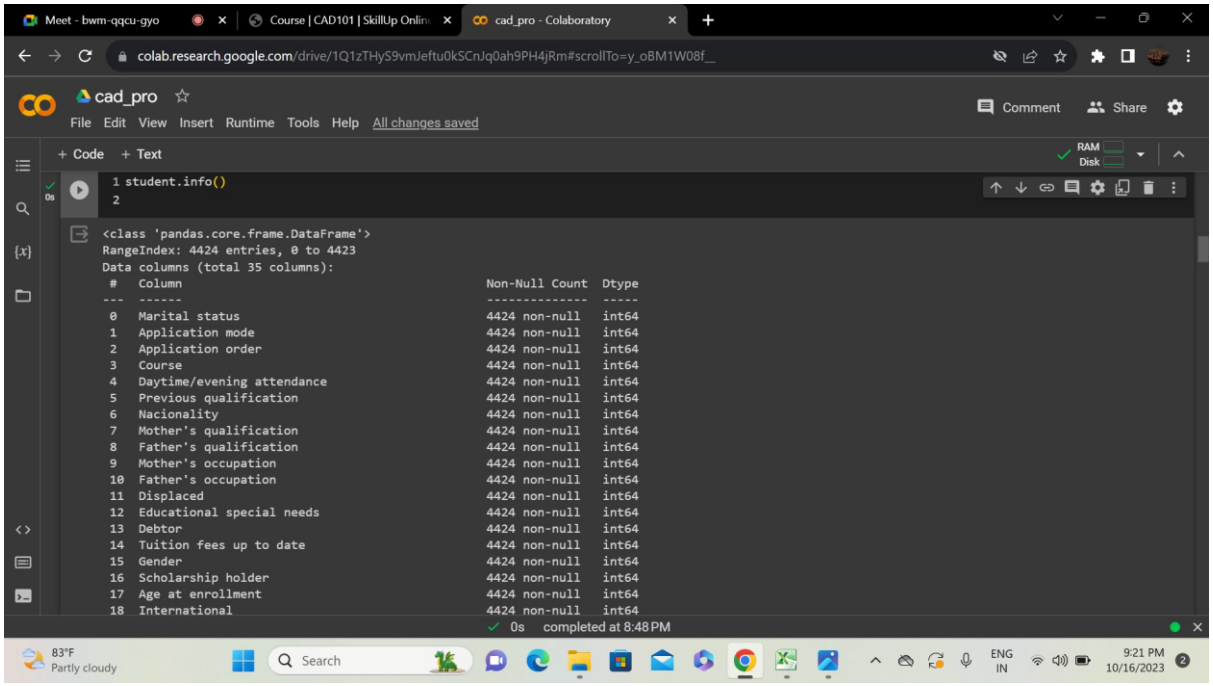


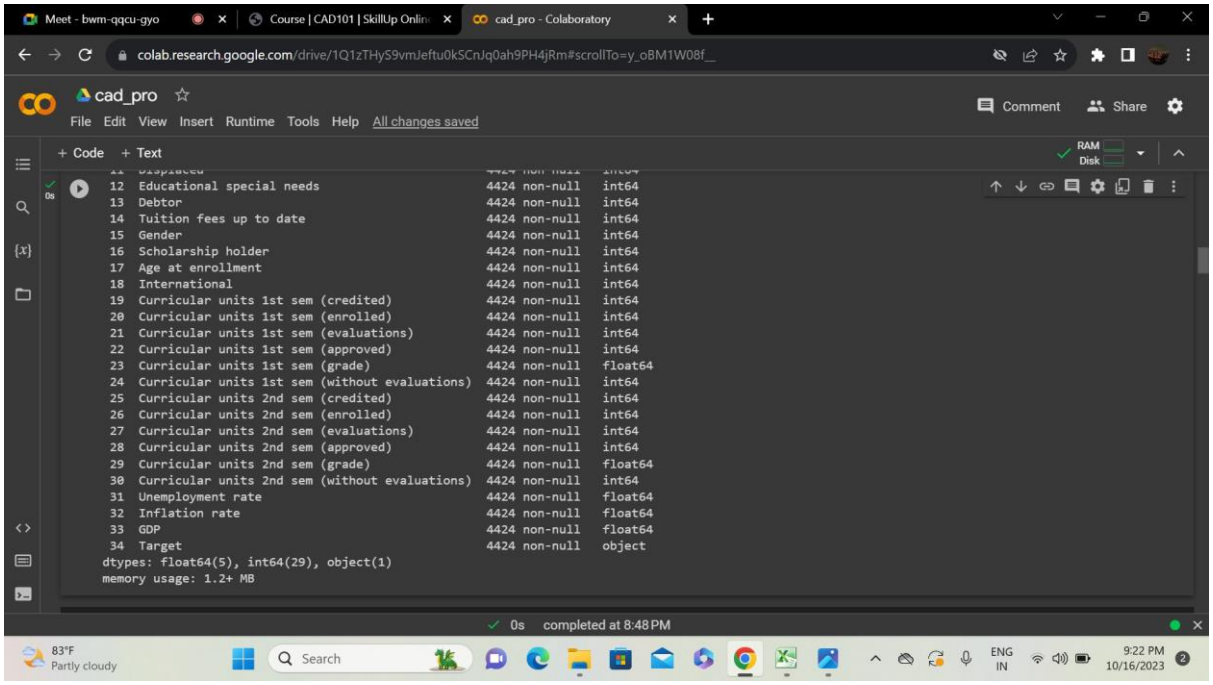

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[1] 2s
11
12 from sklearn.decomposition import PCA
13
14 from sklearn.preprocessing import StandardScaler
15 from sklearn.model_selection import train_test_split
16 from sklearn.model_selection import cross_val_score
17 from sklearn.model_selection import GridSearchCV, RandomizedSearchCV
18
19 from sklearn.metrics import accuracy_score
20 from sklearn.metrics import precision_score, recall_score, f1_score
21
22 import pickle
23
24 import warnings
25 warnings.filterwarnings('ignore')
[7] 0s
1 student=pd.read_csv('dataset.csv')
[8] 0s
1 student.shape
(4424, 35)
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:15 PM
10/16/2023

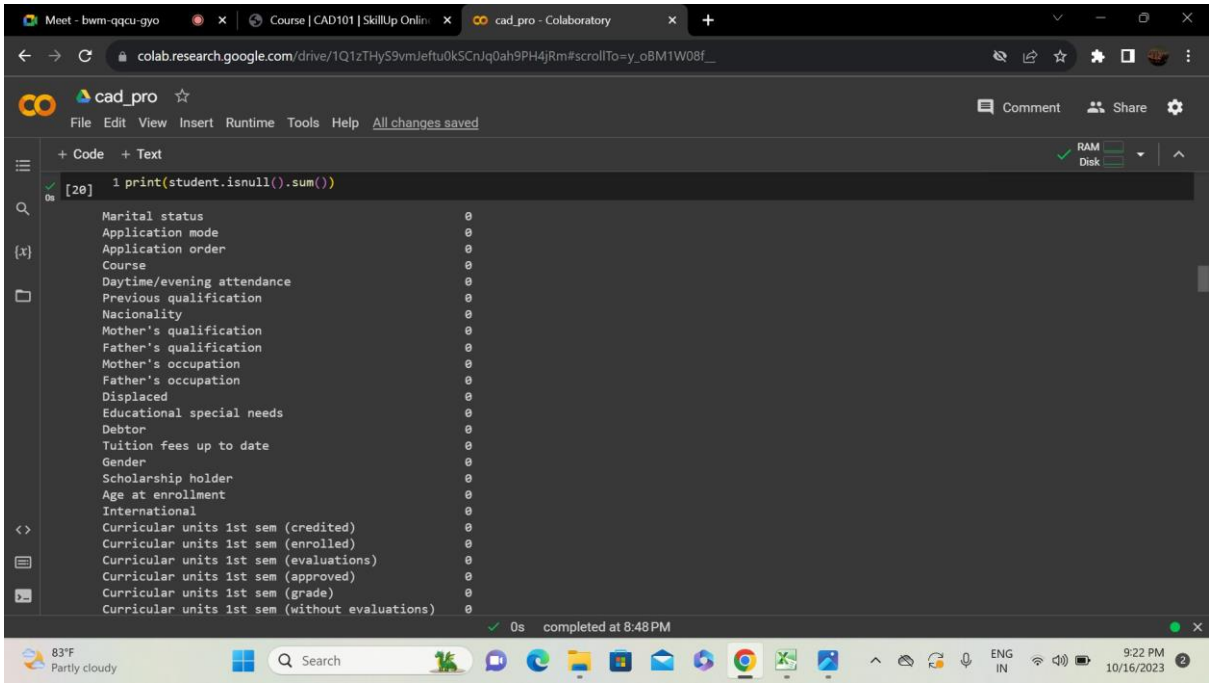

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
(4424, 35)
[9] 0s
1 student.columns
2
Index(['Marital status', 'Application mode', 'Application order', 'Course',
'Daytime/evening attendance', 'Previous qualification', 'Nacionality',
'Mother's qualification', 'Father's qualification',
'Mother's occupation', 'Father's occupation', 'Displaced',
'Educational special needs', 'Debtor', 'Tuition fees up to date',
'Gender', 'Scholarship holder', 'Age at enrollment', 'International',
'Curricular units 1st sem (credited)',
'Curricular units 1st sem (enrolled)',
'Curricular units 1st sem (evaluations)',
'Curricular units 1st sem (approved)',
'Curricular units 1st sem (grade)',
'Curricular units 1st sem (without evaluations)',
'Curricular units 2nd sem (credited)',
'Curricular units 2nd sem (enrolled)',
'Curricular units 2nd sem (evaluations)',
'Curricular units 2nd sem (approved)',
'Curricular units 2nd sem (grade)',
'Curricular units 2nd sem (without evaluations)', 'Unemployment rate',
'Inflation rate', 'GDP', 'Target'],
dtype='object')
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:15 PM
10/16/2023

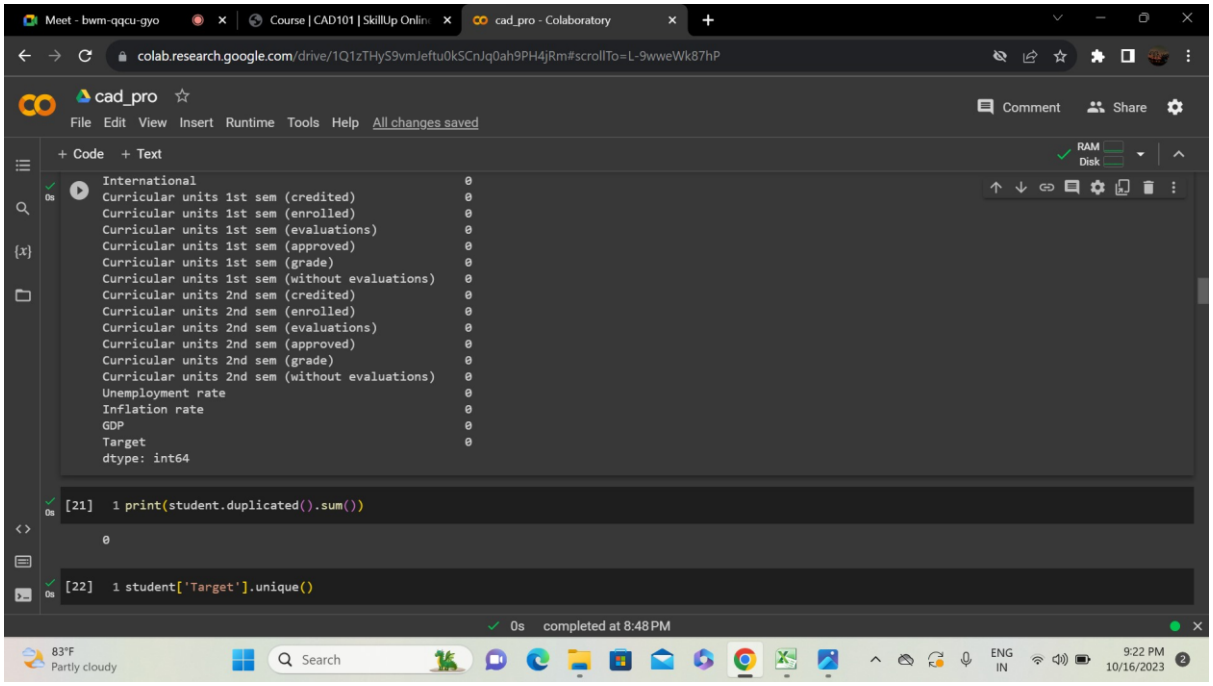

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[10] 1 student.sample(2)
Marital status | Application mode | Application order | Course | Daytime/evening attendance | Previous qualification | Nacionality | Mother's qualification | Father's qualification | Mother's occupation | ... | Curricular units 2nd sem (credited)
2421 | 2 | 12 | 1 | 17 | 0 | 12 | 1 | 22 | 27 | 6 | ... | 0
2368 | 1 | 1 | 1 | 15 | 1 | 1 | 1 | 3 | 2 | 5 | ... | 0
2 rows × 35 columns
[11] 1 student.head()
2
Father's ualification | Mother's occupation | ... | Curricular units 2nd sem (credited) | Curricular units 2nd sem (enrolled) | Curricular units 2nd sem (evaluations) | Curricular units 2nd sem (approved) | Curricular units 2nd sem (grade) | Curricular units 2nd sem (without evaluations) | Unemployment rate | Inflation rate | GDP | Target
10 | 6 | ... | 0 | 0 | 0 | 0 | 0.000000 | 0 | 10.8 | 1.4 | 1.74 | Dropout
3 | 4 | ... | 0 | 6 | 6 | 6 | 13.666667 | 0 | 13.9 | -0.3 | 0.79 | Graduate
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:15 PM
10/16/2023

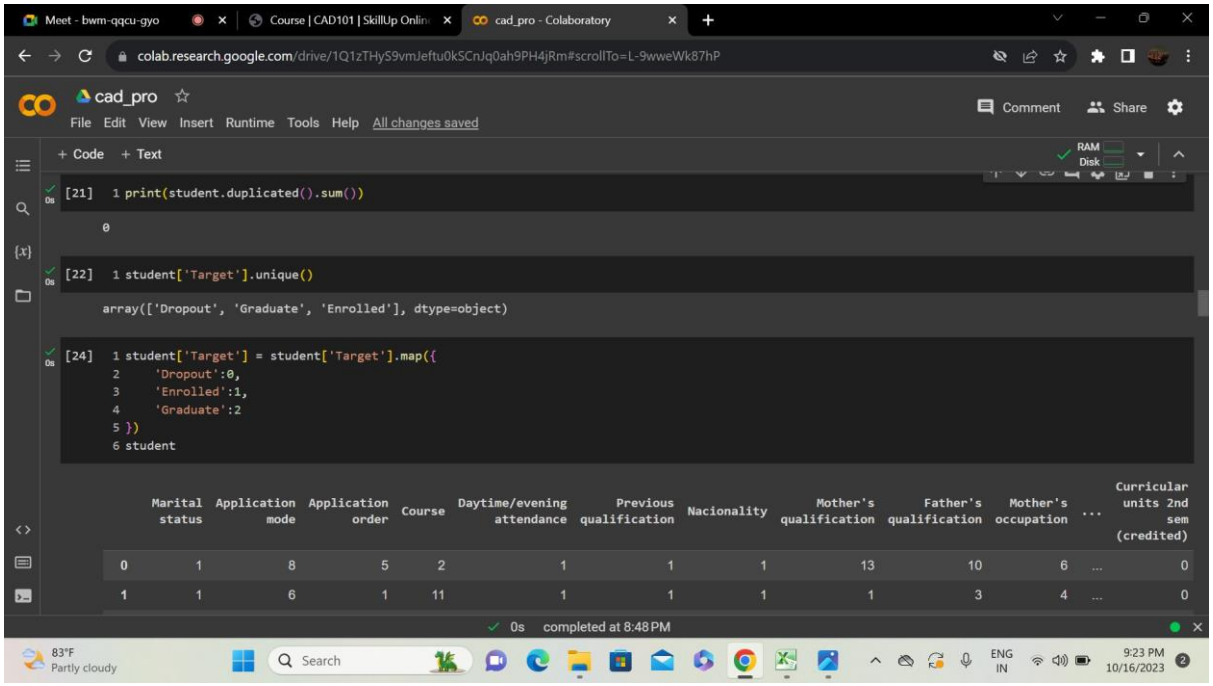

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[11] 1 student.head()
2
Father's ualification | Mother's occupation | ... | Curricular units 2nd sem (credited) | Curricular units 2nd sem (enrolled) | Curricular units 2nd sem (evaluations) | Curricular units 2nd sem (approved) | Curricular units 2nd sem (grade) | Curricular units 2nd sem (without evaluations) | Unemployment rate | Inflation rate | GDP | Target
10 | 6 | ... | 0 | 0 | 0 | 0 | 0.000000 | 0 | 10.8 | 1.4 | 1.74 | Dropout
3 | 4 | ... | 0 | 6 | 6 | 6 | 13.666667 | 0 | 13.9 | -0.3 | 0.79 | Graduate
27 | 10 | ... | 0 | 6 | 0 | 0 | 0.000000 | 0 | 10.8 | 1.4 | 1.74 | Dropout
27 | 6 | ... | 0 | 6 | 10 | 5 | 12.400000 | 0 | 9.4 | -0.8 | -3.12 | Graduate
28 | 10 | ... | 0 | 6 | 6 | 6 | 13.000000 | 0 | 13.9 | -0.3 | 0.79 | Graduate
[19] 1 student.info()
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:21 PM
10/16/2023

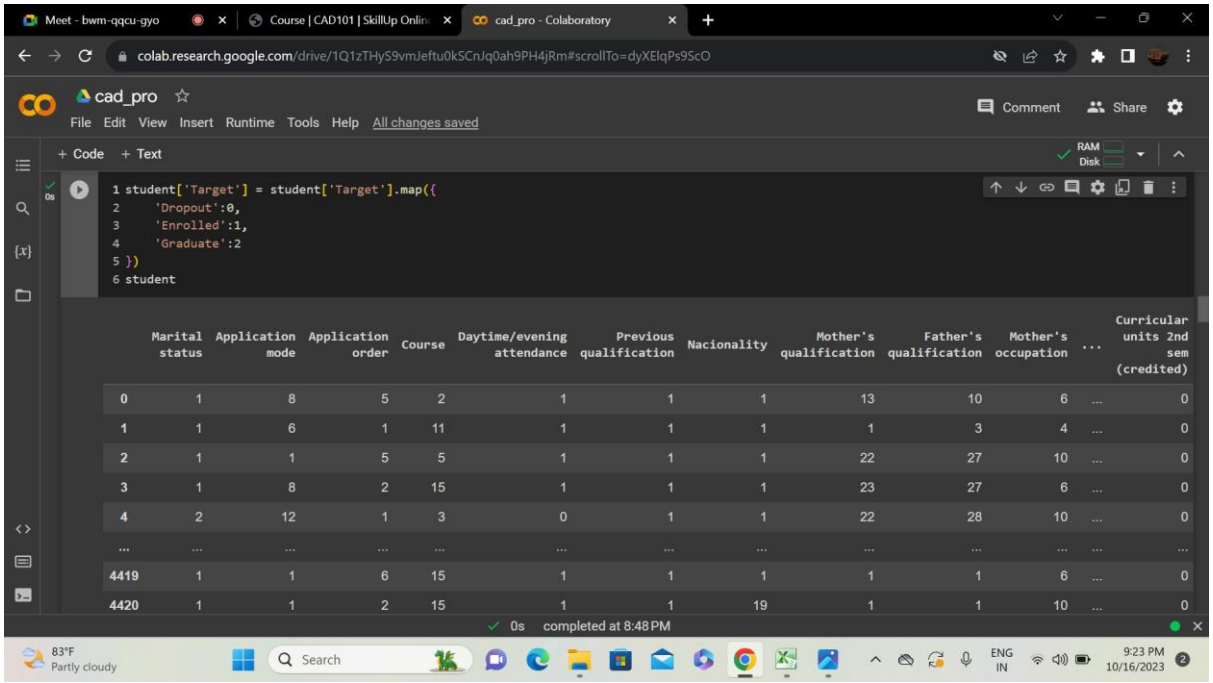

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=y_oBM1W08f_
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
1 student.info()
2
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 4424 entries, 0 to 4423
Data columns (total 35 columns):
Column Non-Null Count Dtype
--- ------ -------------- -----
0 Marital status 4424 non-null int64
1 Application mode 4424 non-null int64
2 Application order 4424 non-null int64
3 Course 4424 non-null int64
4 Daytime/evening attendance 4424 non-null int64
5 Previous qualification 4424 non-null int64
6 Nacionality 4424 non-null int64
7 Mother's qualification 4424 non-null int64
8 Father's qualification 4424 non-null int64
9 Mother's occupation 4424 non-null int64
10 Father's occupation 4424 non-null int64
11 Displaced 4424 non-null int64
12 Educational special needs 4424 non-null int64
13 Debtor 4424 non-null int64
14 Tuition fees up to date 4424 non-null int64
15 Gender 4424 non-null int64
16 Scholarship holder 4424 non-null int64
17 Age at enrollment 4424 non-null int64
18 International 4424 non-null int64
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:21 PM
10/16/2023

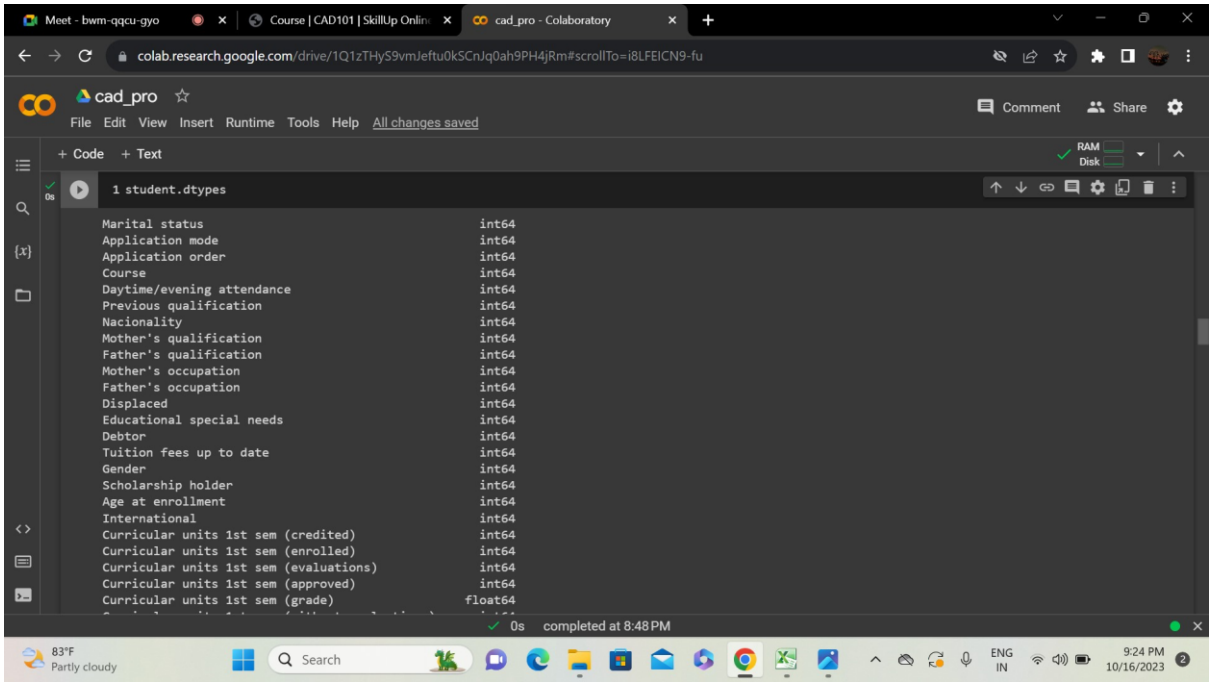

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=y_oBM1W08f_
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
12 Educational special needs 4424 non-null int64
13 Debtor 4424 non-null int64
14 Tuition fees up to date 4424 non-null int64
15 Gender 4424 non-null int64
16 Scholarship holder 4424 non-null int64
17 Age at enrollment 4424 non-null int64
18 International 4424 non-null int64
19 Curricular units 1st sem (credited) 4424 non-null int64
20 Curricular units 1st sem (enrolled) 4424 non-null int64
21 Curricular units 1st sem (evaluations) 4424 non-null int64
22 Curricular units 1st sem (approved) 4424 non-null int64
23 Curricular units 1st sem (grade) 4424 non-null float64
24 Curricular units 1st sem (without evaluations) 4424 non-null int64
25 Curricular units 2nd sem (credited) 4424 non-null int64
26 Curricular units 2nd sem (enrolled) 4424 non-null int64
27 Curricular units 2nd sem (evaluations) 4424 non-null int64
28 Curricular units 2nd sem (approved) 4424 non-null int64
29 Curricular units 2nd sem (grade) 4424 non-null float64
30 Curricular units 2nd sem (without evaluations) 4424 non-null int64
31 Unemployment rate 4424 non-null float64
32 Inflation rate 4424 non-null float64
33 GDP 4424 non-null float64
34 Target 4424 non-null object
dtypes: float64(5), int64(29), object(1)
memory usage: 1.2+ MB
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:22 PM
10/16/2023

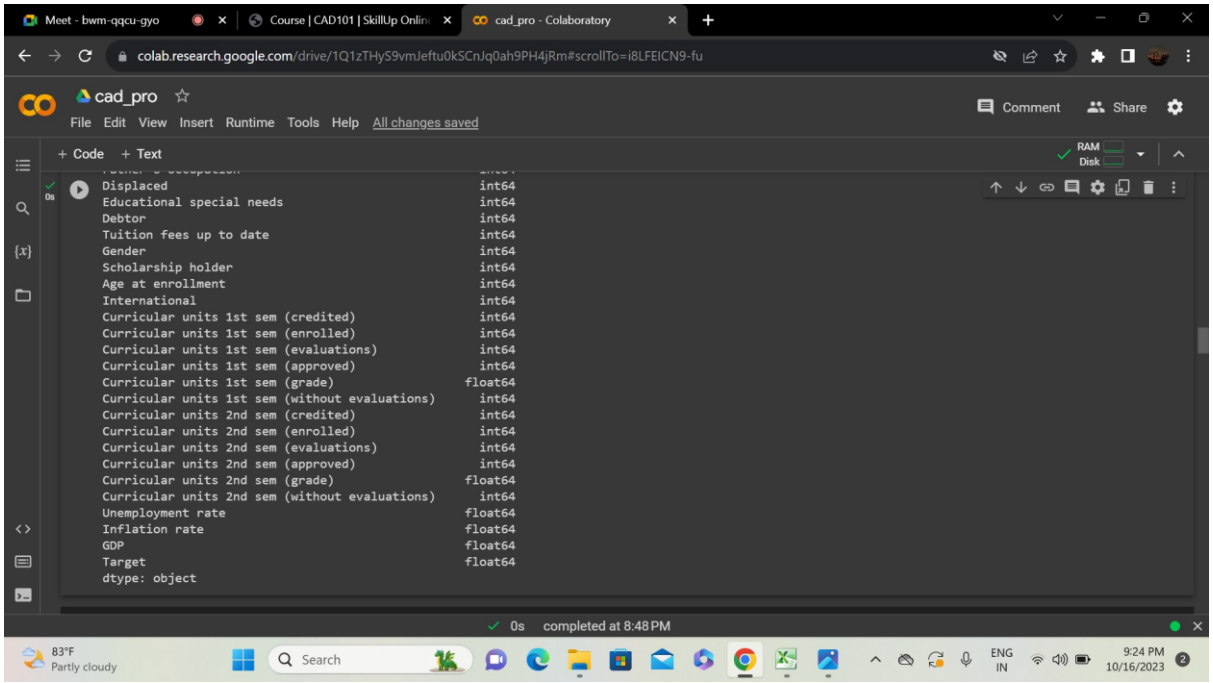

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=y_oBM1W08f_
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[20] 1 print(student.isnull().sum())
Marital status 0
Application mode 0
Application order 0
Course 0
Daytime/evening attendance 0
Previous qualification 0
Nacionality 0
Mother's qualification 0
Father's qualification 0
Mother's occupation 0
Father's occupation 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
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:22 PM
10/16/2023

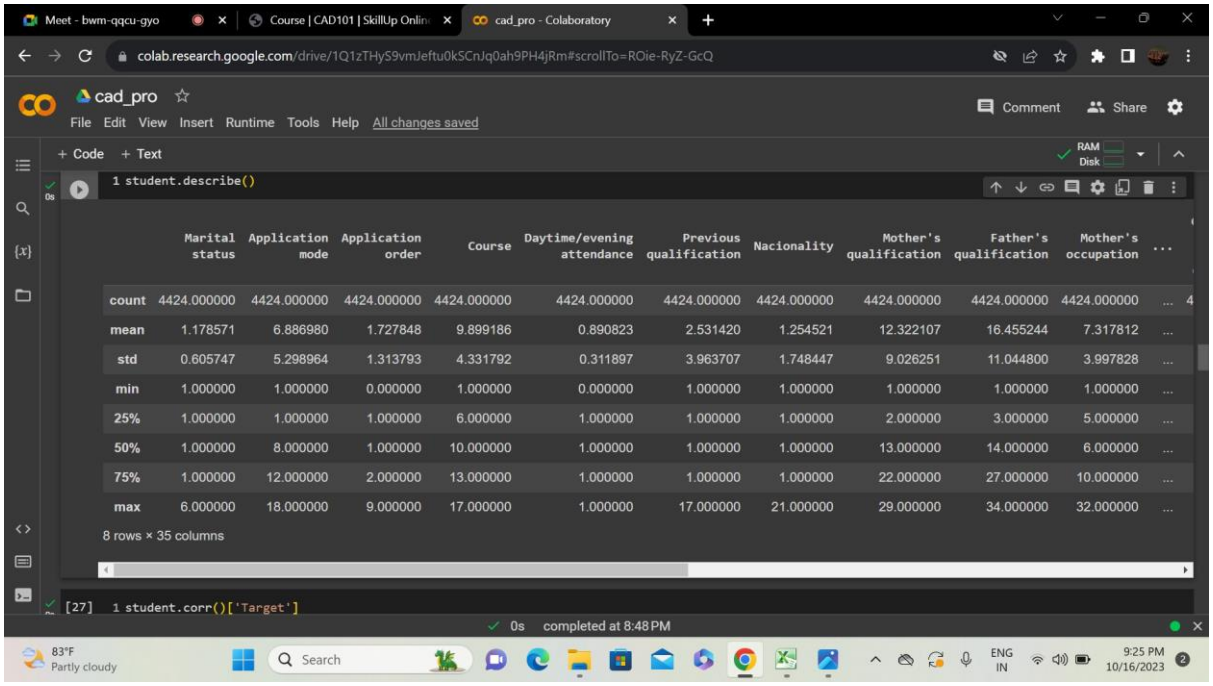

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
International 0
Curricular units 1st 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
[21] 1 print(student.duplicated().sum())
0
[22] 1 student['Target'].unique()
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:22 PM
10/16/2023

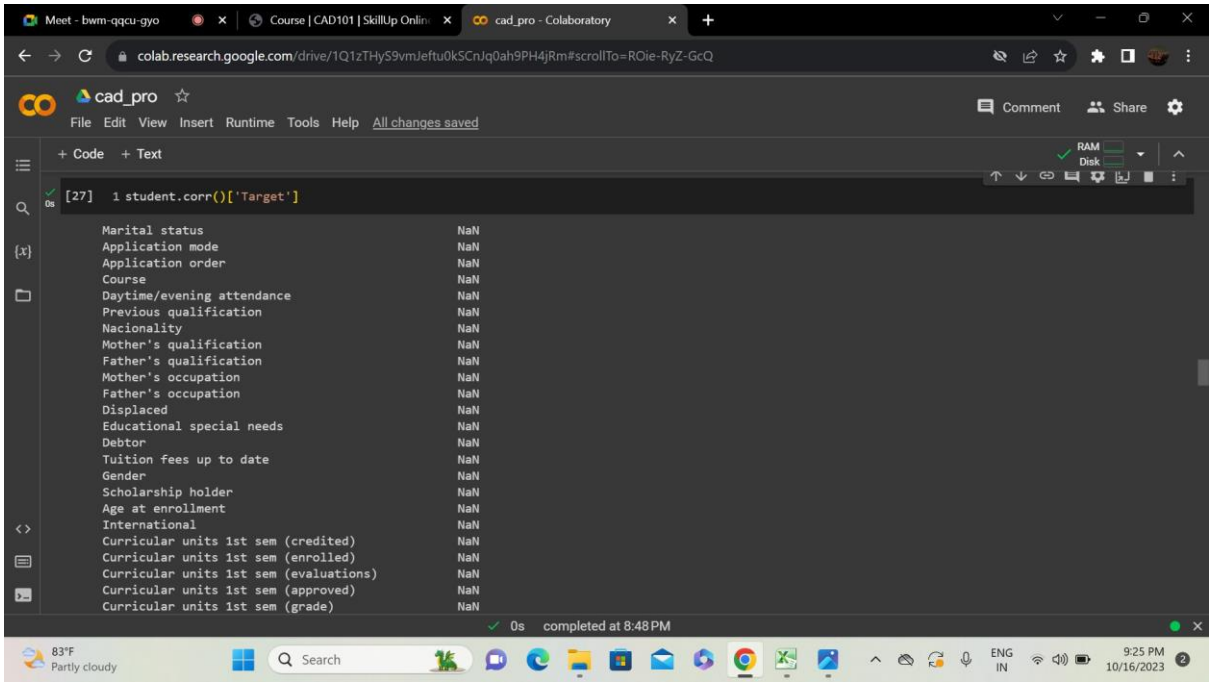

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=L-9wweWk87hP
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[21] 1 print(student.duplicated().sum())
0
[22] 1 student['Target'].unique()
array(['Dropout', 'Graduate', 'Enrolled'], dtype=object)
[24] 1 student['Target'] = student['Target'].map({
2 'Dropout':0,
3 'Enrolled':1,
4 'Graduate':2
5 })
6 student
Marital status
Application mode
Application order
Course
Daytime/evening attendance
Previous qualification
Nacionality
Mother's qualification
Father's qualification
Mother's occupation
Curricular units 2nd sem (credited)
0 1 8 5 2 1 1 1 13 10 6 ... 0
1 1 6 1 11 1 1 1 1 3 4 ... 0
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:23 PM
10/16/2023

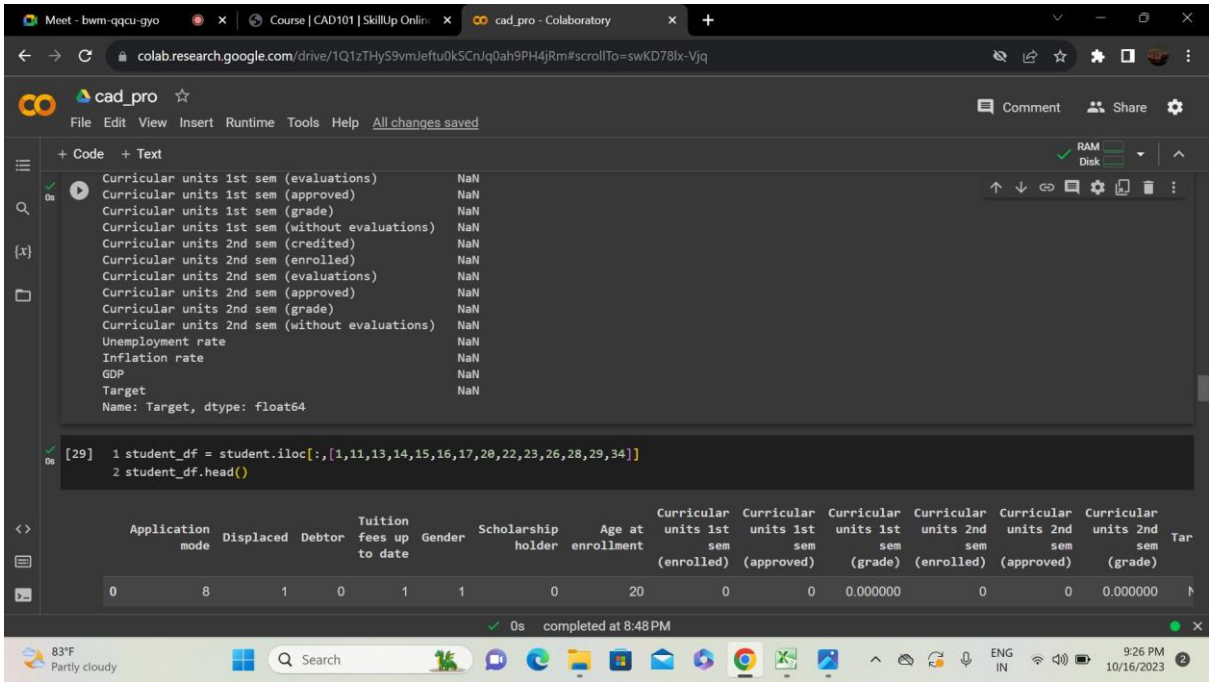

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=dyXElqPs9ScO
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
1 student['Target'] = student['Target'].map({
2 'Dropout':0,
3 'Enrolled':1,
4 'Graduate':2
5 })
6 student
Marital status
Application mode
Application order
Course
Daytime/evening attendance
Previous qualification
Nacionality
Mother's qualification
Father's qualification
Mother's occupation
Curricular units 2nd sem (credited)
0 1 8 5 2 1 1 1 13 10 6 ... 0
1 1 6 1 11 1 1 1 1 3 4 ... 0
2 1 1 5 5 1 1 1 22 27 10 ... 0
3 1 8 2 15 1 1 1 23 27 6 ... 0
4 2 12 1 3 0 1 1 22 28 10 ... 0
4419 1 1 6 15 1 1 1 1 1 6 ... 0
4420 1 1 2 15 1 1 19 1 1 10 ... 0
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:23 PM
10/16/2023

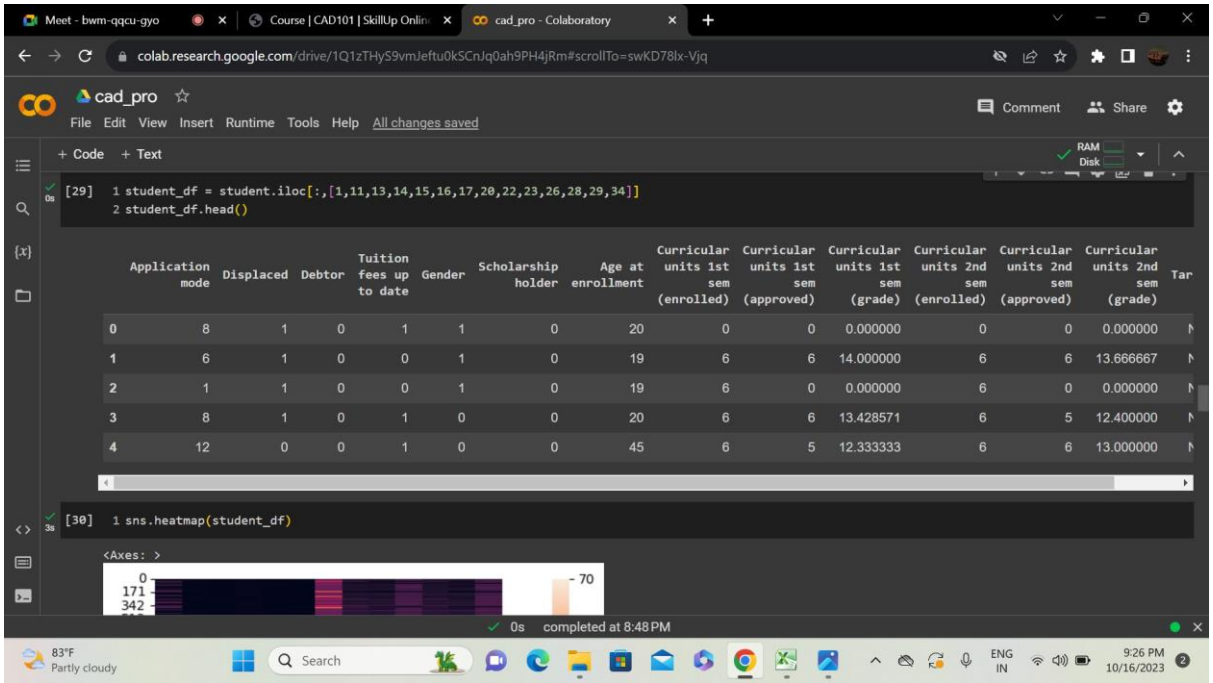

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=i8LFEICN9-fu
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
1 student.dtypes
Marital status int64
Application mode int64
Application order int64
Course int64
Daytime/evening attendance int64
Previous qualification int64
Nacionality int64
Mother's qualification int64
Father's qualification int64
Mother's occupation int64
Father's occupation int64
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
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:24 PM
10/16/2023

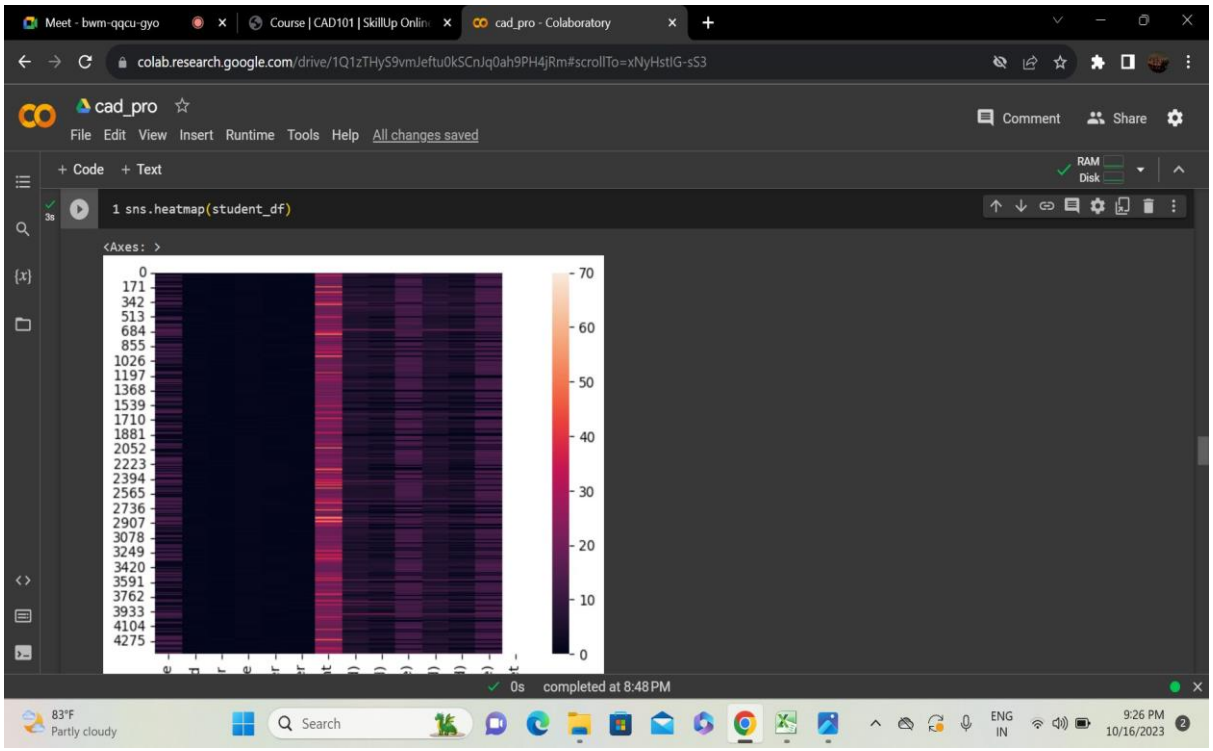

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=i8LFEICN9-fu
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Displaced int64
Educational special 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 float64
dtype: object
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:24 PM
10/16/2023


Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=ROie-RyZ-GcQ
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
1 student.describe()
Marital status
Application mode
Application order
Course
Daytime/evening attendance
Previous qualification
Nacionality
Mother's qualification
Father's qualification
Mother's occupation
count 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000 4424.000000
mean 1.178571 6.886980 1.727848 9.899186 0.890823 2.531420 1.254521 12.322107 16.455244 7.317812
std 0.605747 5.298964 1.313793 4.331792 0.311897 3.963707 1.748447 9.026251 11.044800 3.997828
min 1.000000 1.000000 0.000000 1.000000 0.000000 1.000000 1.000000 1.000000 1.000000 1.000000
25% 1.000000 1.000000 1.000000 6.000000 1.000000 1.000000 1.000000 2.000000 3.000000 5.000000
50% 1.000000 8.000000 1.000000 10.000000 1.000000 1.000000 1.000000 13.000000 14.000000 6.000000
75% 1.000000 12.000000 2.000000 13.000000 1.000000 1.000000 1.000000 22.000000 27.000000 10.000000
max 6.000000 18.000000 9.000000 17.000000 1.000000 17.000000 21.000000 29.000000 34.000000 32.000000
8 rows × 35 columns
[27] 1 student.corr()['Target']
0s completed at 8:48 PM
83°F Partly cloudy
Search
ENG IN
9:25 PM 10/16/2023


Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=ROie-RyZ-GcQ
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[27] 1 student.corr()['Target']
Marital status NaN
Application mode NaN
Application order NaN
Course NaN
Daytime/evening attendance NaN
Previous qualification NaN
Nacionality NaN
Mother's qualification NaN
Father's qualification NaN
Mother's occupation NaN
Father's occupation NaN
Displaced NaN
Educational special needs NaN
Debtor NaN
Tuition fees up to date NaN
Gender NaN
Scholarship holder NaN
Age at enrollment NaN
International NaN
Curricular units 1st sem (credited) NaN
Curricular units 1st sem (enrolled) NaN
Curricular units 1st sem (evaluations) NaN
Curricular units 1st sem (approved) NaN
Curricular units 1st sem (grade) NaN
0s completed at 8:48 PM
83°F Partly cloudy
Search
ENG IN
9:25 PM 10/16/2023

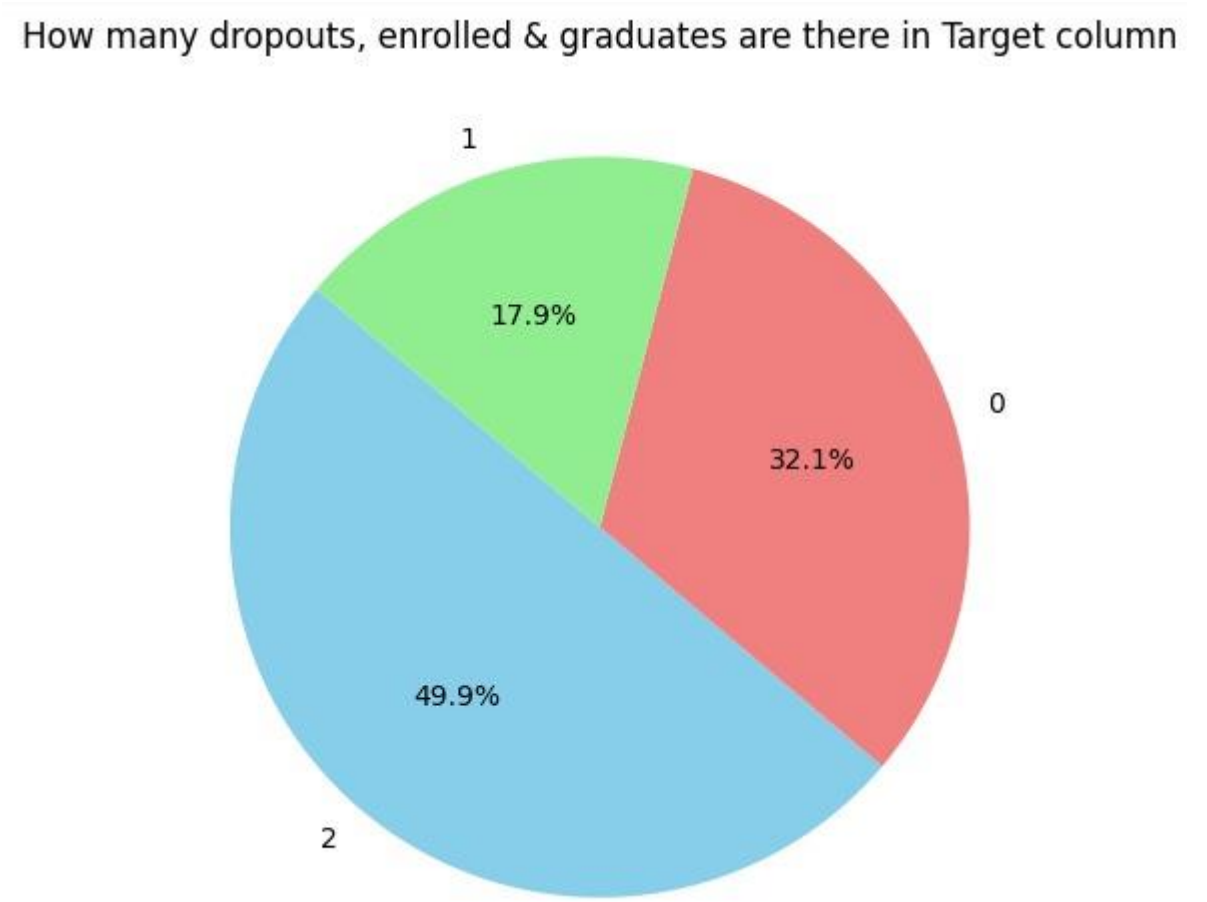

Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=swKD78lx-Vjq
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
Curricular units 1st sem (evaluations) NaN
Curricular units 1st sem (approved) NaN
Curricular units 1st sem (grade) NaN
Curricular units 1st sem (without evaluations) NaN
Curricular units 2nd sem (credited) NaN
Curricular units 2nd sem (enrolled) NaN
Curricular units 2nd sem (evaluations) NaN
Curricular units 2nd sem (approved) NaN
Curricular units 2nd sem (grade) NaN
Curricular units 2nd sem (without evaluations) NaN
Unemployment rate NaN
Inflation rate NaN
GDP NaN
Target NaN
Name: Target, dtype: float64
[29] 1 student_df = student.iloc[:,[1,11,13,14,15,16,17,20,22,23,26,28,29,34]]
2 student_df.head()
Application mode
Displaced
Debtor
Tuition fees up to date
Gender
Scholarship holder
Age at enrollment
Curricular units 1st sem (enrolled)
Curricular units 1st sem (approved)
Curricular units 1st sem (grade)
Curricular units 2nd sem (enrolled)
Curricular units 2nd sem (approved)
Curricular units 2nd sem (grade)
Tar
0 8 1 0 1 1 0 20 0 0 0.000000 0 0 0.000000
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:26 PM
10/16/2023


Meet - bwm-qqcu-gyo
Course | CAD101 | SkillUp Onlin
cad_pro - Colaboratory
colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm#scrollTo=swKD78lx-Vjq
cad_pro
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
+ Code + Text
RAM
Disk
[29] 1 student_df = student.iloc[:,[1,11,13,14,15,16,17,20,22,23,26,28,29,34]]
2 student_df.head()
Application mode
Displaced
Debtor
Tuition fees up to date
Gender
Scholarship holder
Age at enrollment
Curricular units 1st sem (enrolled)
Curricular units 1st sem (approved)
Curricular units 1st sem (grade)
Curricular units 2nd sem (enrolled)
Curricular units 2nd sem (approved)
Curricular units 2nd sem (grade)
Tar
0 8 1 0 1 1 0 20 0 0 0.000000 0 0 0.000000
1 6 1 0 0 1 0 19 6 6 14.000000 6 6 13.666667
2 1 1 0 0 1 0 19 6 0 0.000000 6 0 0.000000
3 8 1 0 1 0 0 20 6 6 13.428571 6 5 12.400000
4 12 0 0 1 0 0 45 6 5 12.333333 6 6 13.000000
[30] 1 sns.heatmap(student_df)
<Axes: >
0
171
342
70
0s completed at 8:48 PM
83°F
Partly cloudy
Search
ENG
IN
9:26 PM
10/16/2023



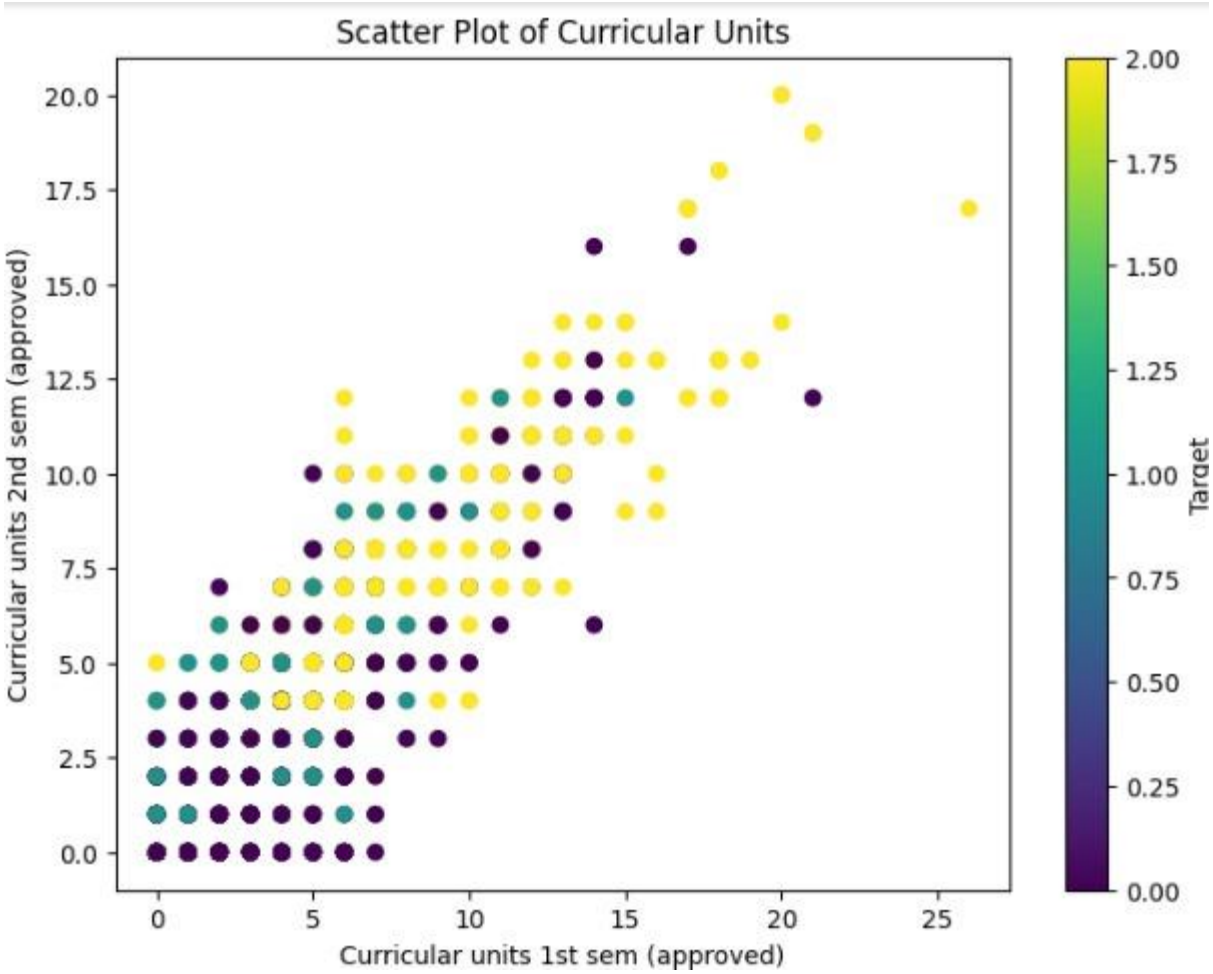
Exploratory Data Analysis:

```
[26] 1 import matplotlib.pyplot as plt
     2 import pandas as pd
     3 x = student_df['Target'].value_counts().index
     4 y = student_df['Target'].value_counts().values
     5 plt.figure(figsize=(6, 6))
     6 plt.pie(y, labels=x, autopct='%1.1f%%', startangle=140, colors=['skyblue', 'lightcoral', 'lightgreen'])
     7 plt.title('How many dropouts, enrolled & graduates are there in Target column')
     8 plt.show()
```




```python
import matplotlib.pyplot as plt
x = student_df['Curricular units 1st sem (approved)']
y = student_df['Curricular units 2nd sem (approved)']
colors = student_df['Target']
plt.figure(figsize=(8, 6))
plt.scatter(x, y, c=colors, cmap='viridis')
plt.xlabel('Curricular units 1st sem (approved)')
plt.ylabel('Curricular units 2nd sem (approved)')
plt.title('Scatter Plot of Curricular Units')
plt.colorbar(label='Target')
plt.show()

```



LINK TO THE GOOGLE COLAB NOTEBOOK:

https://colab.research.google.com/drive/1Q1zTHyS9vmJeftu0kSCnJq0ah9PH4jRm?usp=sharing

CONCLUSION:

In conclusion, this project harnesses the power of machine learning and IBM Cloud to tackle the issue of student dropout in educational institutions. By deploying predictive models, we enable these institutions to make informed decisions and take proactive steps to improve student retention rates, ultimately ensuring that students have a better chance at a successful academic journey. This work underscores the importance of data-driven insights in addressing real-world challenges in the educational sector.