Project Report Template

INTRODUCTION

1.1 Overview

A brief description about your project:

Campus recruitment is a strategy for sourcing, engaging and hiring young talent for internship and entry-level positions. College recruiting is typically a tactic for medium- to large-sized companies with high-volume recruiting needs, but can range from small efforts (like working with university career centers to source potential candidates) to large-scale operations (like visiting a wide array of colleges and attending recruiting events throughout the spring and fall semester).Campus recruitment often involves working with university career services centers and attending career fairs to meet in-person with college students and recent graduates. Our solution revolves around the placement season of a Business School in India. Where it has various factors on candidates getting hired such as work experience, exam percentage etc., Finally it contains the status of recruitment and remuneration details.

We will be using algorithms such as KNN, SVM and ANN. We will train and testthe data with these algorithms. From this the best model is selected and saved in .pkl format. We will be doing flask integration and IBM deployment.

1.2 Purpose

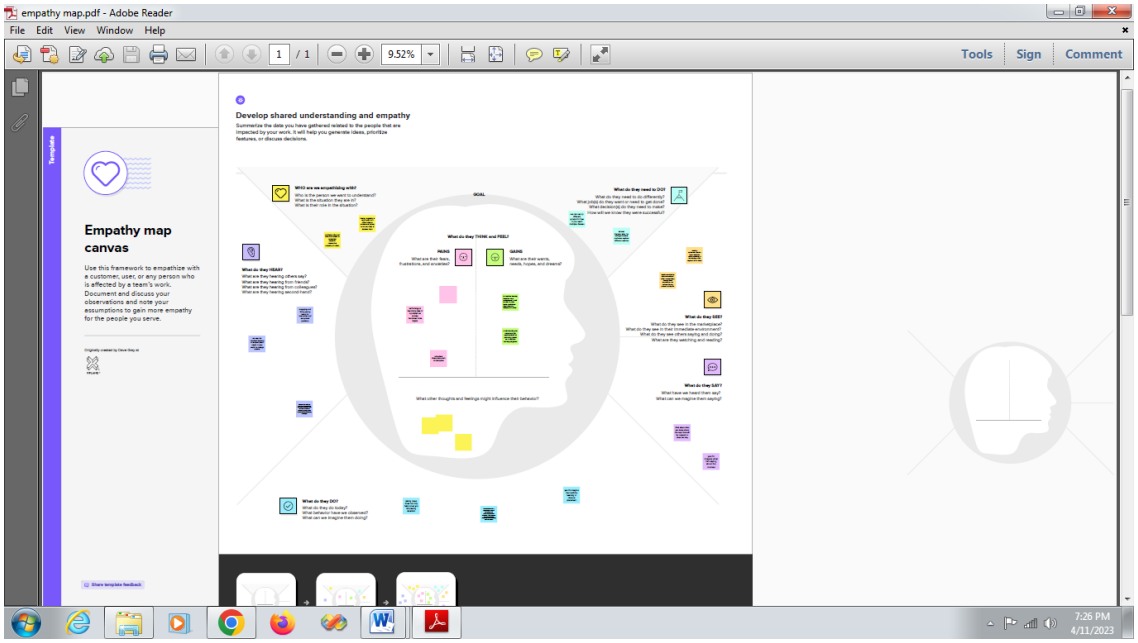
The use of this project. What can be achieved using this:

Abstract Every educational institution relies on campus placement to assist students in achieving their objectives. Machine learning classification can be used to retrieve associated data from huge student datasets.

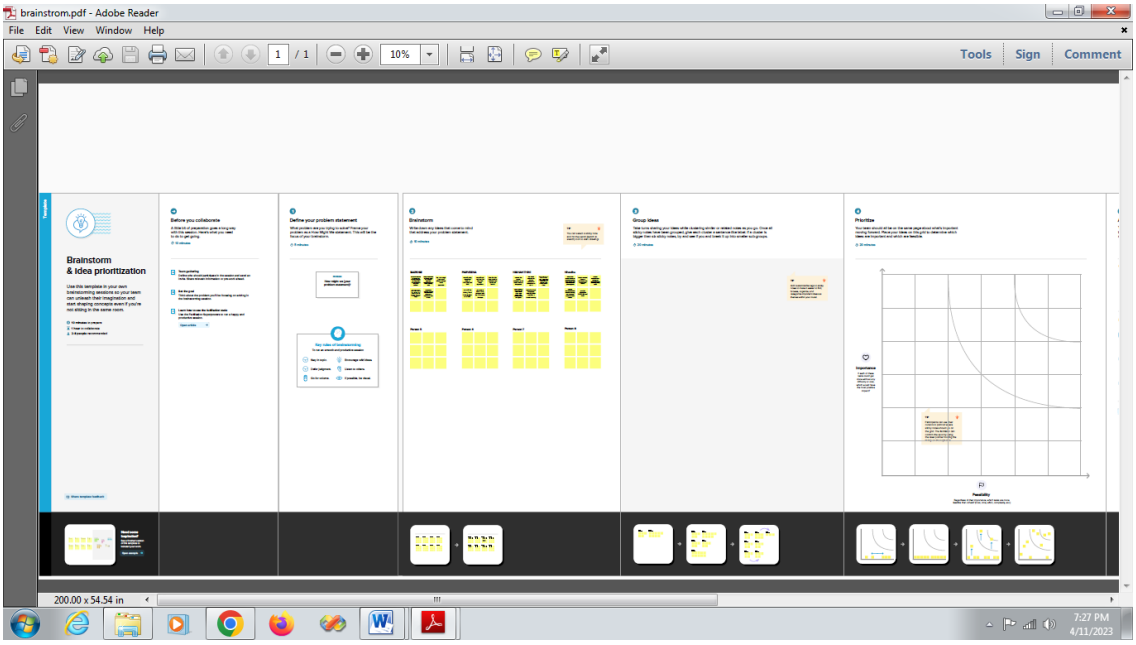
Campus placement or campus recruiting is a program conducted within universities or other educational institutions **to provide jobs to students nearing completion of their studies**. In this type of program, the educational institutions partner with corporations who wish to recruit from the student population.

Problem Definition & Design Thinking:

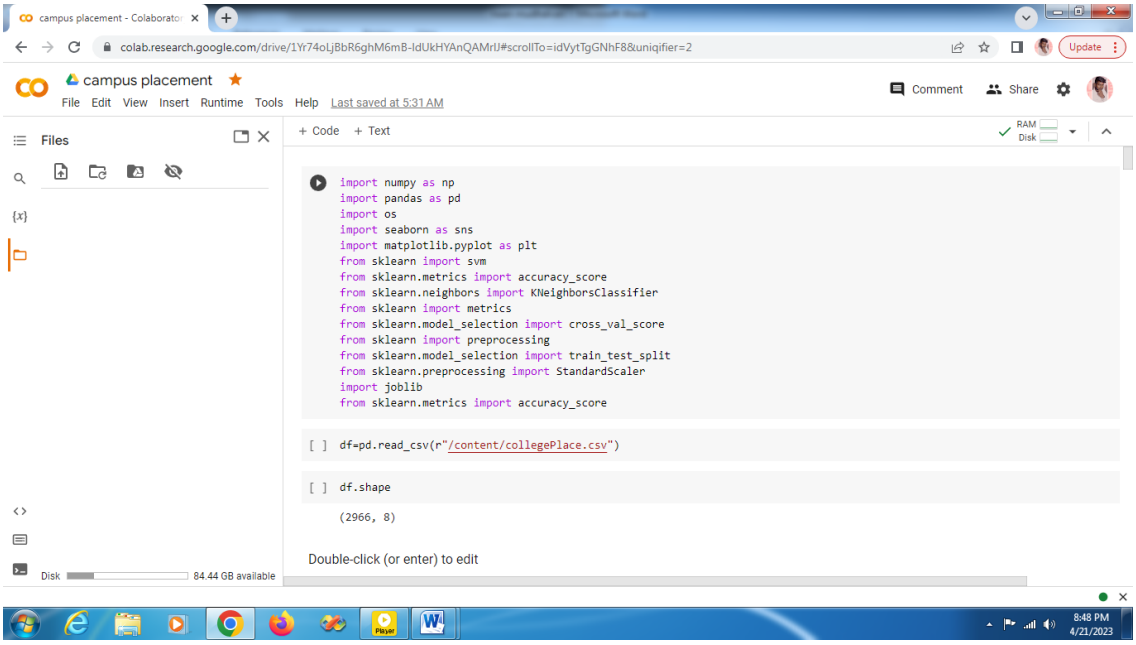
2.1 Empathy Map

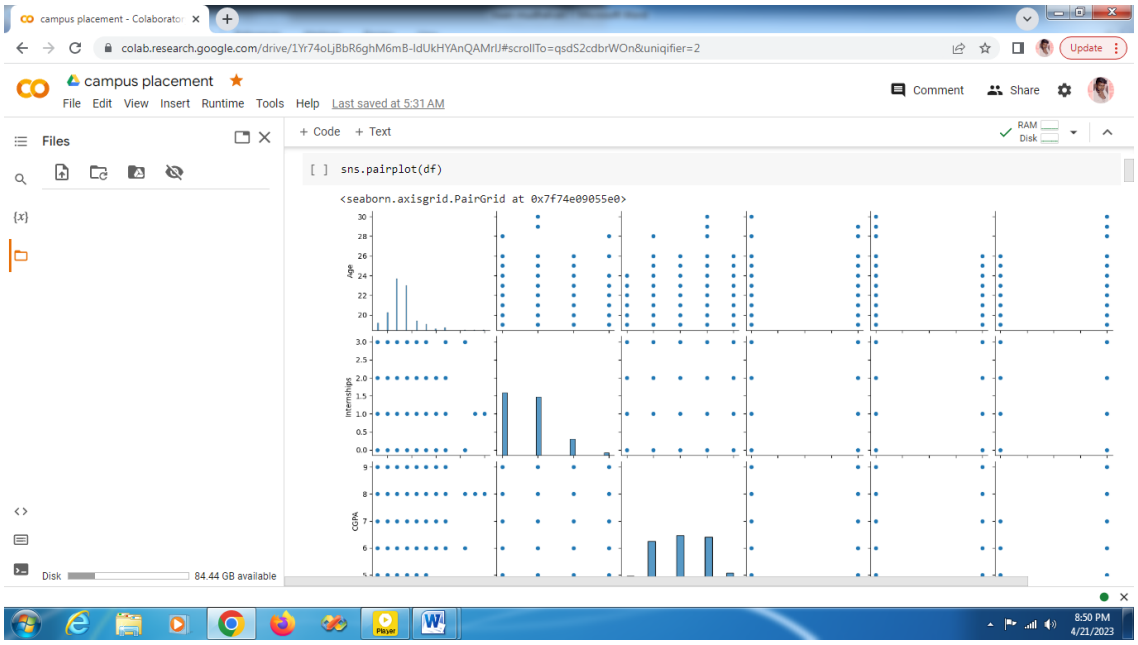


2.2 Ideation & Brainstorming Map

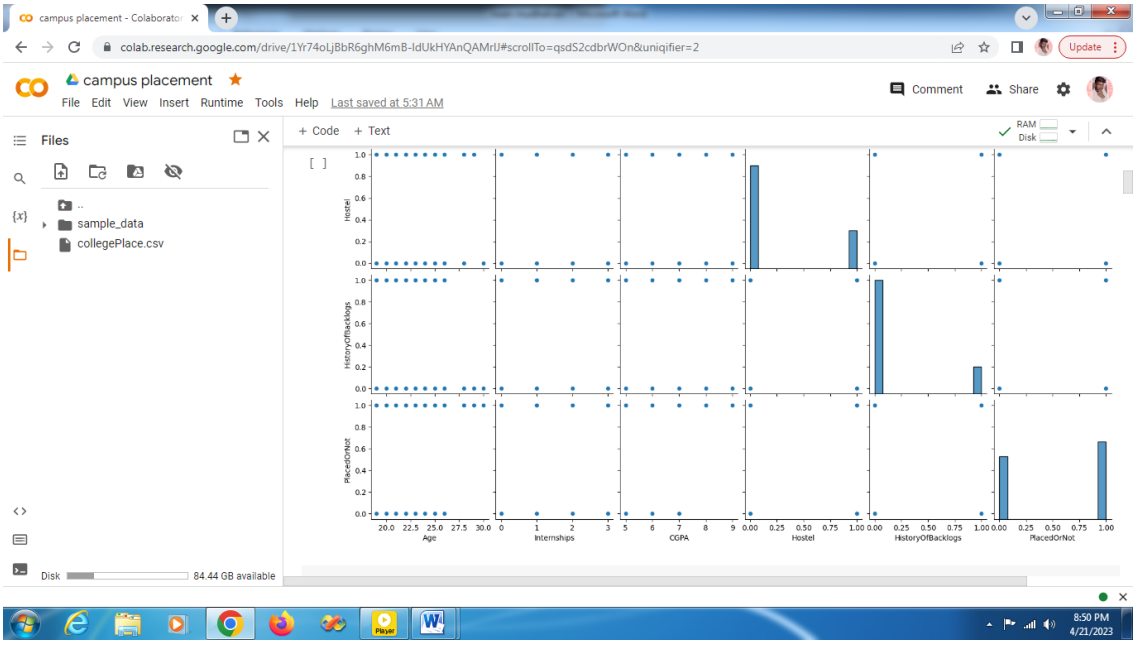


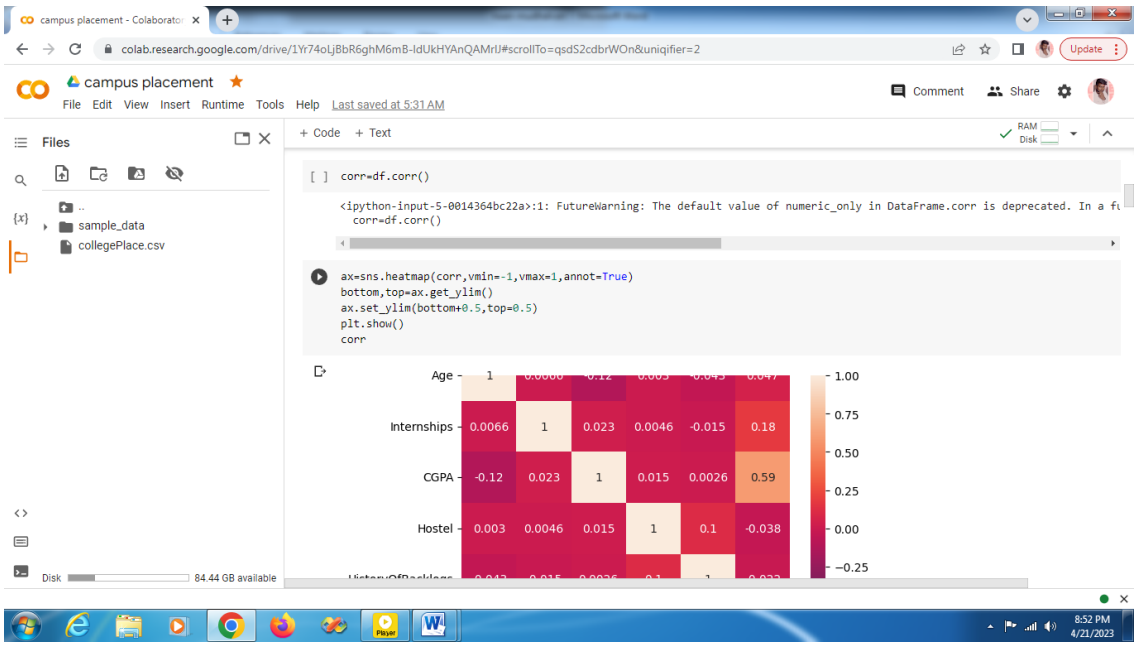
3.RESULT:


campus placement - Colaborato
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=idVytTgGNhF8&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
import numpy as np
import pandas as pd
import os
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn import svm
from sklearn.metrics import accuracy_score
from sklearn.neighbors import KNeighborsClassifier
from sklearn import metrics
from sklearn.model_selection import cross_val_score
from sklearn import preprocessing
from sklearn.model_selection import train_test_split
from sklearn.preprocessing import StandardScaler
import joblib
from sklearn.metrics import accuracy_score
[ ] df=pd.read_csv(r"/content/collegePlace.csv")
[ ] df.shape
(2966, 8)
Double-click (or enter) to edit
Disk
84.44 GB available
8:48 PM
4/21/2023

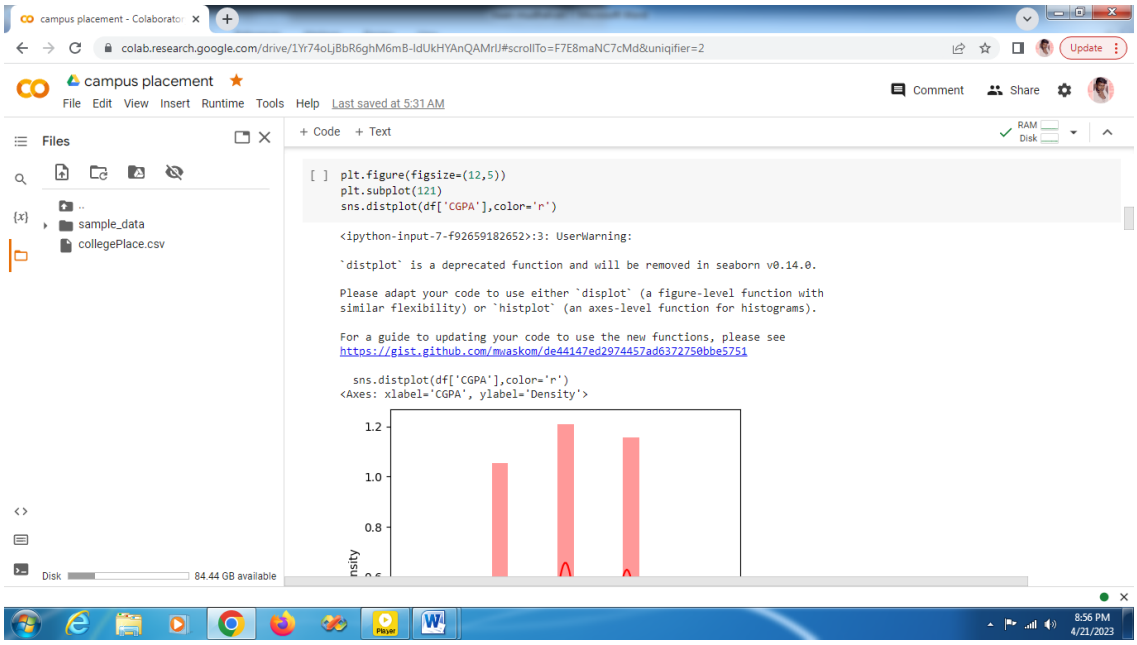

campus placement - Colaborato
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=qsdS2cdbrWOn&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
[ ] sns.pairplot(df)
<seaborn.axisgrid.PairGrid at 0x7f74e09055e0>
Age
Internships
CGPA
Disk
84.44 GB available
8:50 PM
4/21/2023

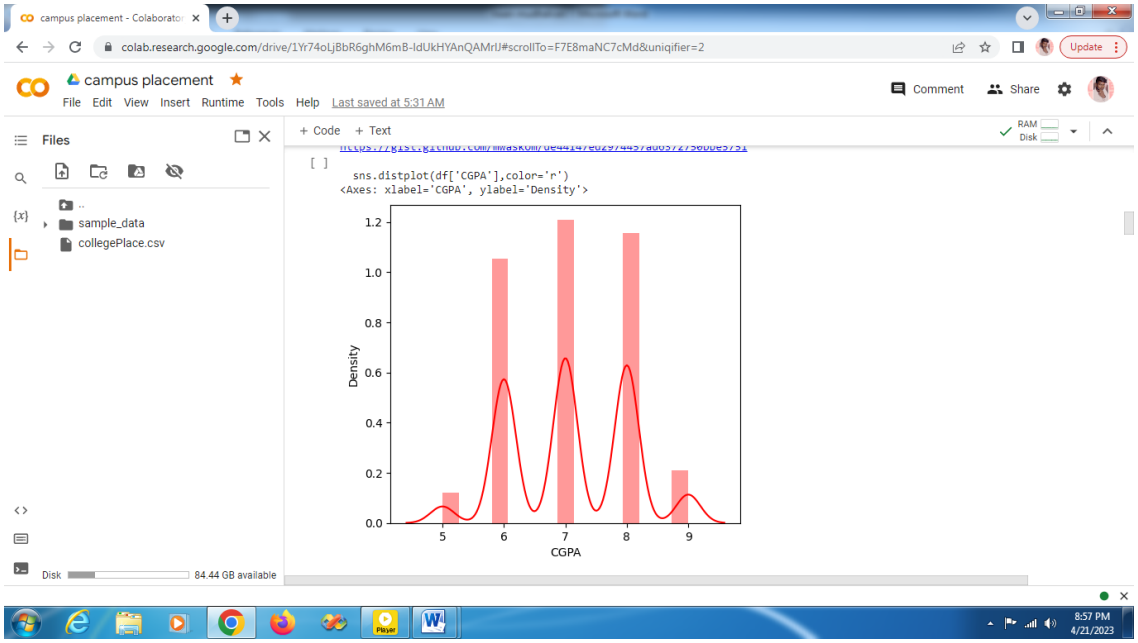
Project Report Template

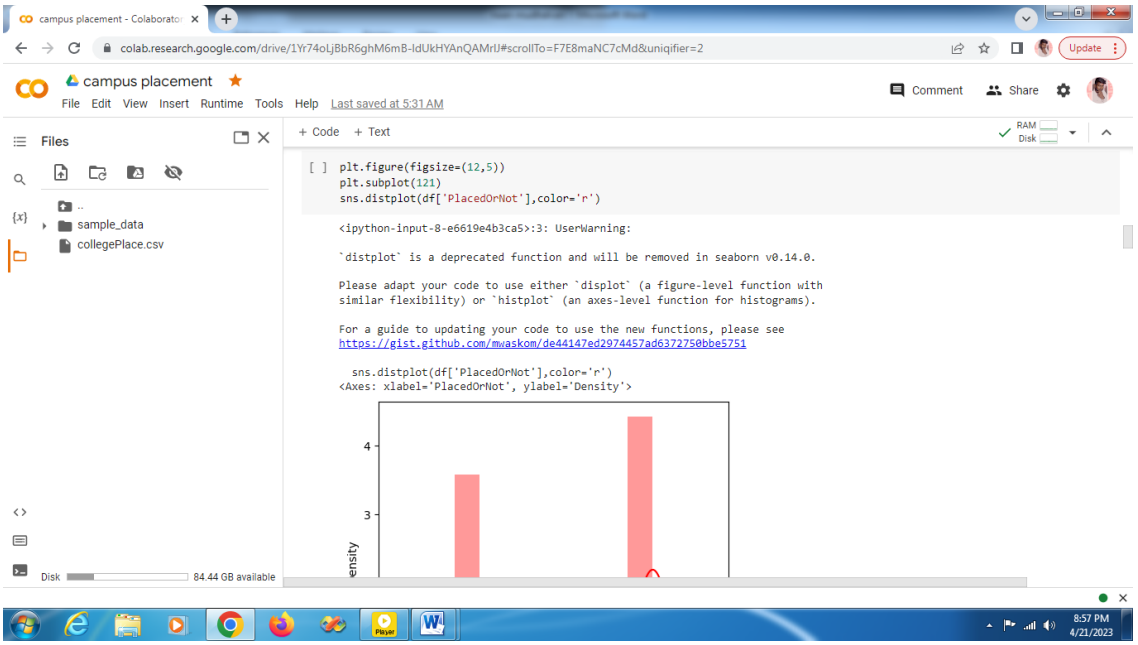


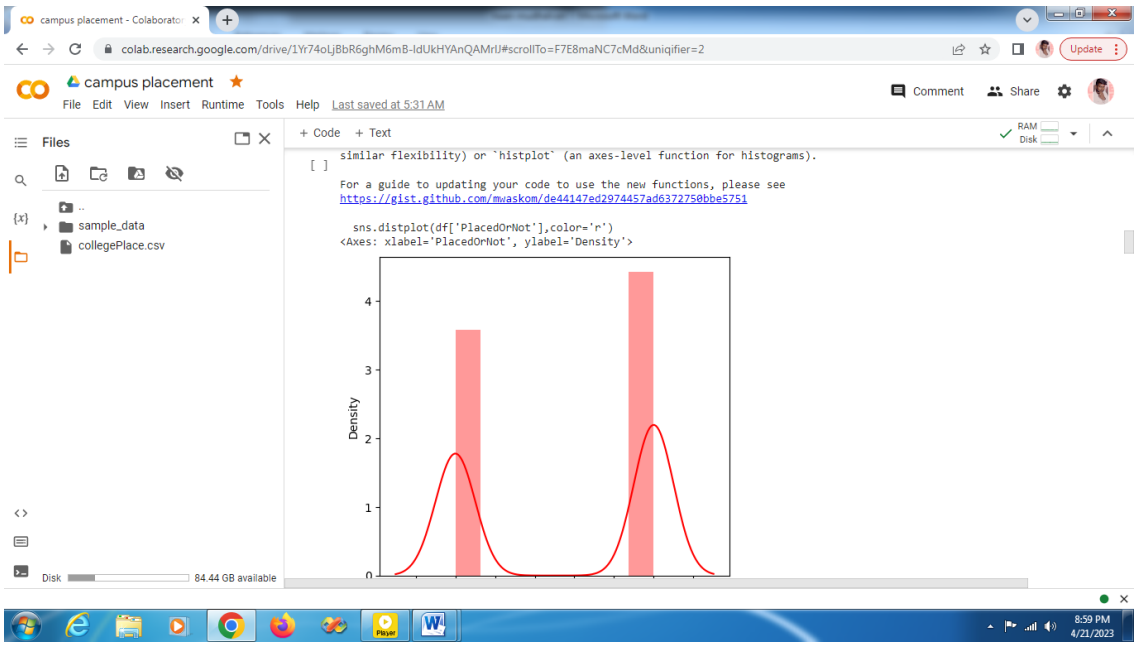


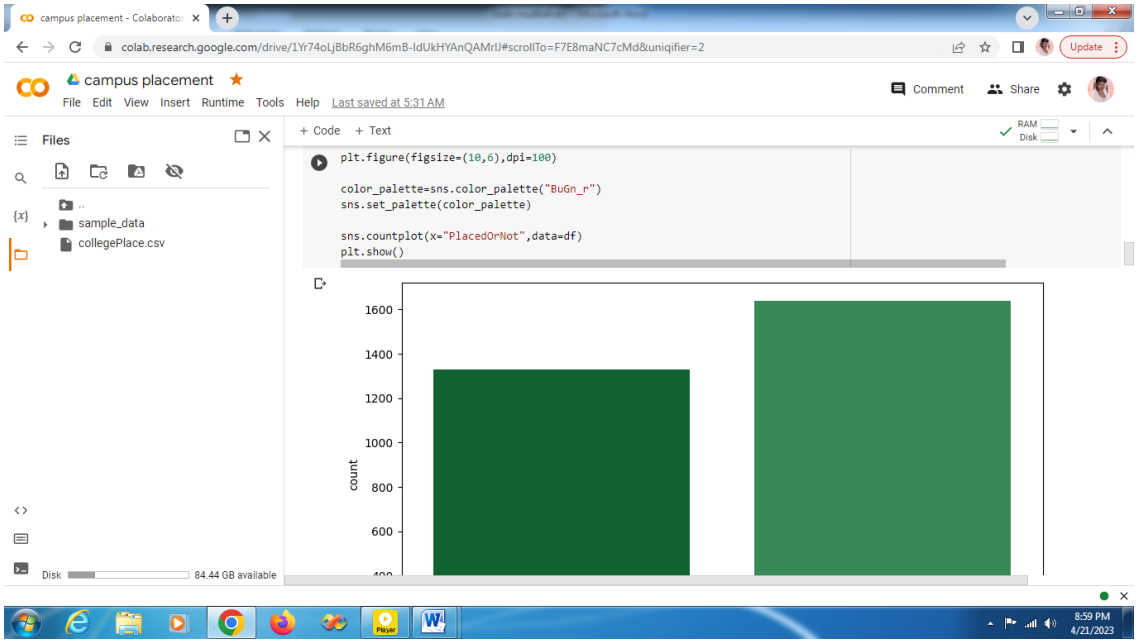
	Age	Internships	CGPA	Hostel	HistoryOfBacklogs	PlacedOrNot
Age	1.000000	0.006552	-0.119787	0.003042	-0.042586	0.046943
Internships	0.006552	1.000000	0.023496	0.004617	-0.015118	0.179334
CGPA	-0.119787	0.023496	1.000000	0.014991	0.002576	0.588648
Hostel	0.003042	0.004617	0.014991	1.000000	0.103506	-0.038182
HistoryOfBacklogs	-0.042586	-0.015118	0.002576	0.103506	1.000000	-0.022337
PlacedOrNot	0.046943	0.179334	0.588648	-0.038182	-0.022337	1.000000

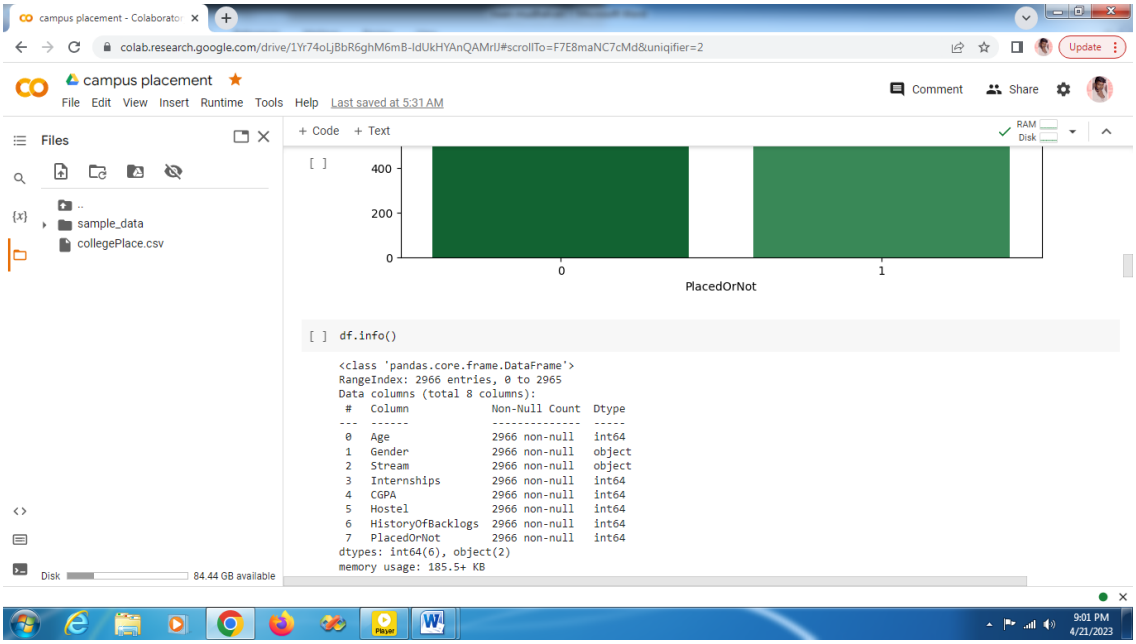


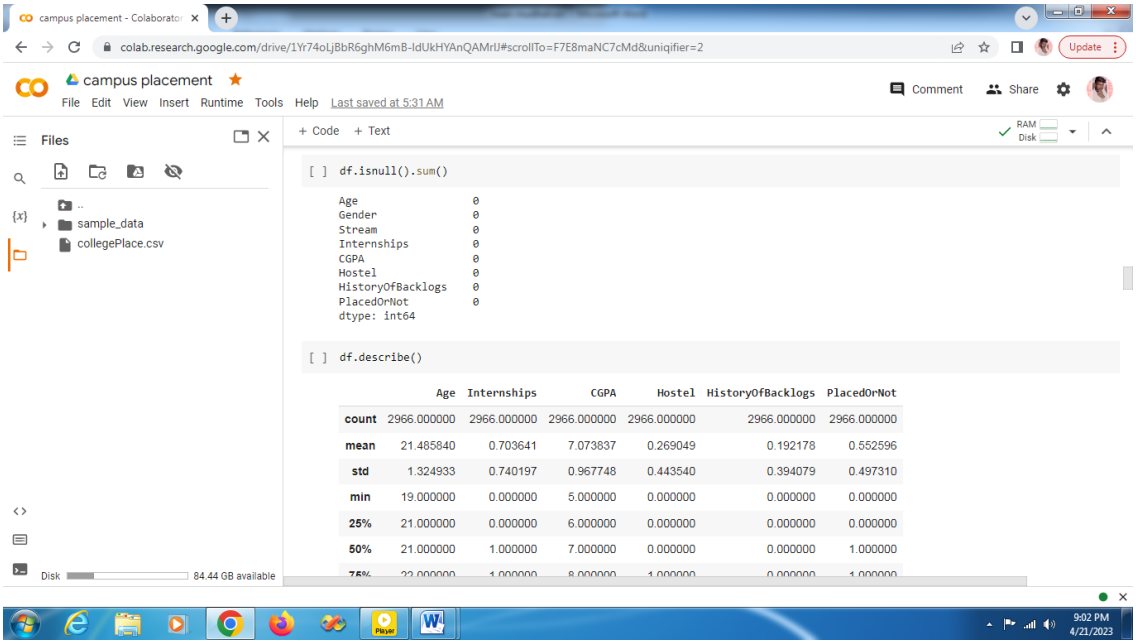

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
[ ]
sns.distplot(df['CGPA'],color='r')
<Axes: xlabel='CGPA', ylabel='Density'>
Density
CGPA
Disk 84.44 GB available
8:57 PM
4/21/2023

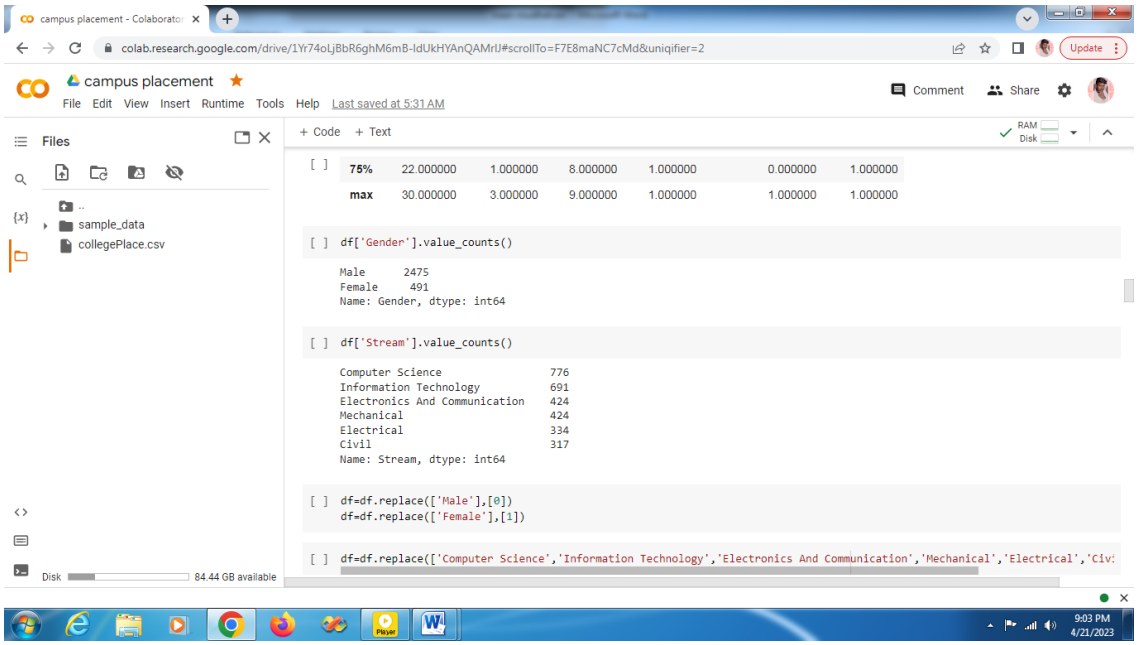

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
[ ]
plt.figure(figsize=(12,5))
plt.subplot(121)
sns.distplot(df['PlacedOrNot'],color='r')
<ipython-input-8-e6619e4b3ca5>:3: UserWarning:
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).
For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
sns.distplot(df['PlacedOrNot'],color='r')
<Axes: xlabel='PlacedOrNot', ylabel='Density'>
Disk 84.44 GB available
8:57 PM
4/21/2023

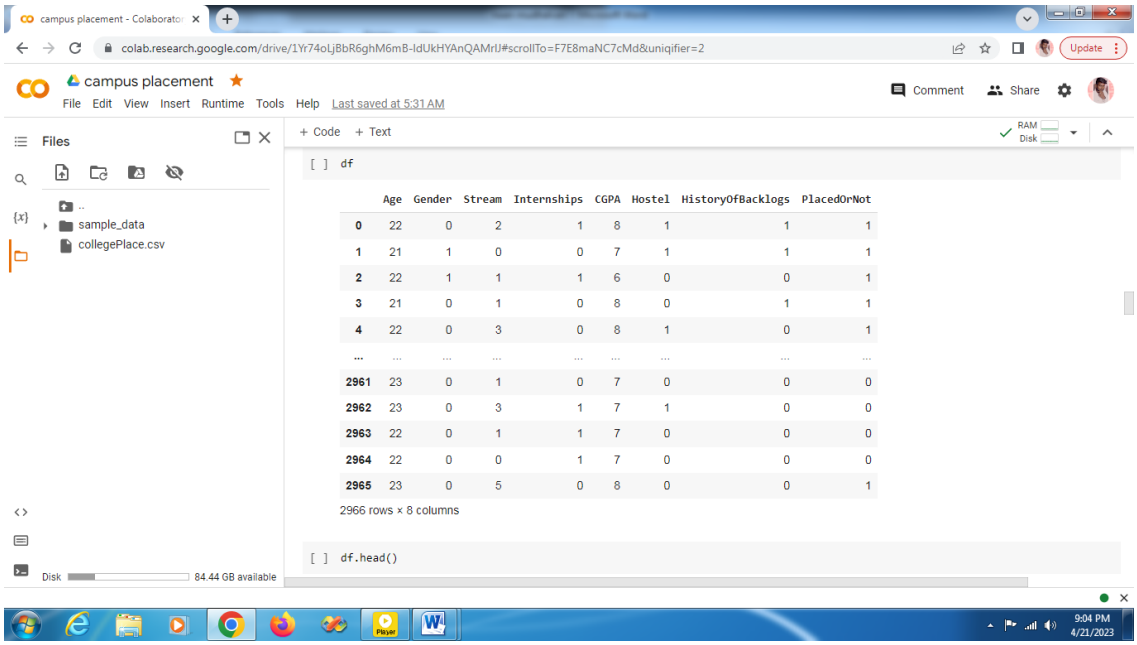

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
RAM
Disk
Files
+ Code + Text
similar flexibility) or `histplot` (an axes-level function for histograms).
For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
sns.distplot(df['PlacedOrNot'],color='r')
<Axes: xlabel='PlacedOrNot', ylabel='Density'>
sample_data
collegePlace.csv
Density
Disk
84.44 GB available
8:59 PM
4/21/2023

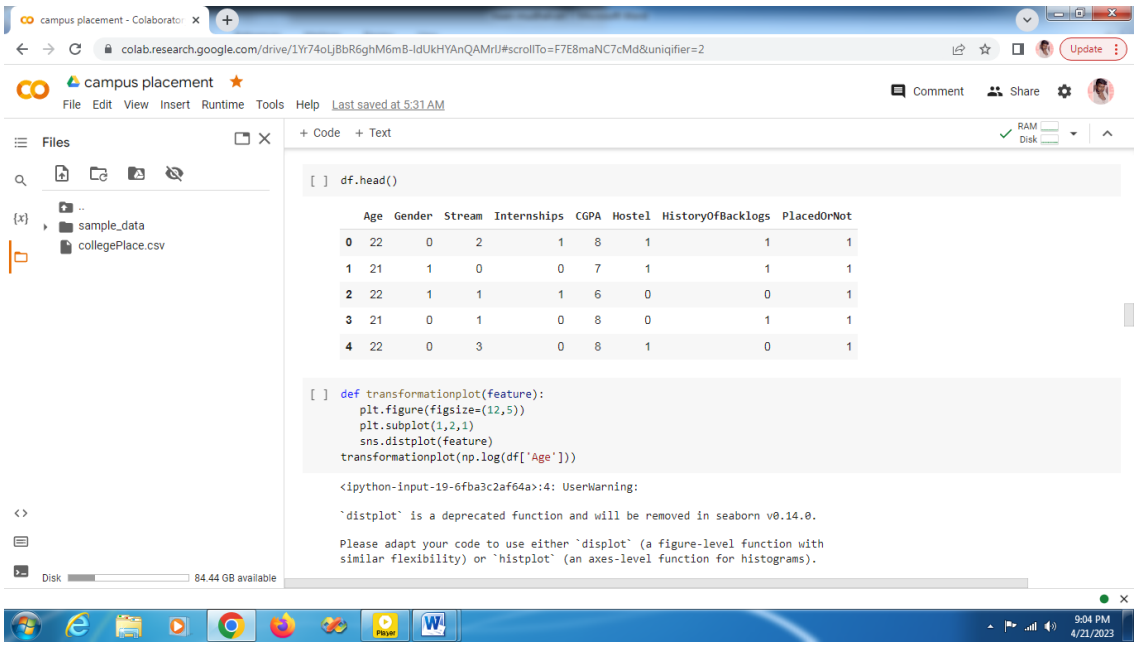

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
RAM
Disk
Files
+ Code + Text
plt.figure(figsize=(10,6),dpi=100)
color_palette=sns.color_palette("BuGn_r")
sns.set_palette(color_palette)
sns.countplot(x="PlacedOrNot",data=df)
plt.show()
sample_data
collegePlace.csv
count
Disk
84.44 GB available
8:59 PM
4/21/2023

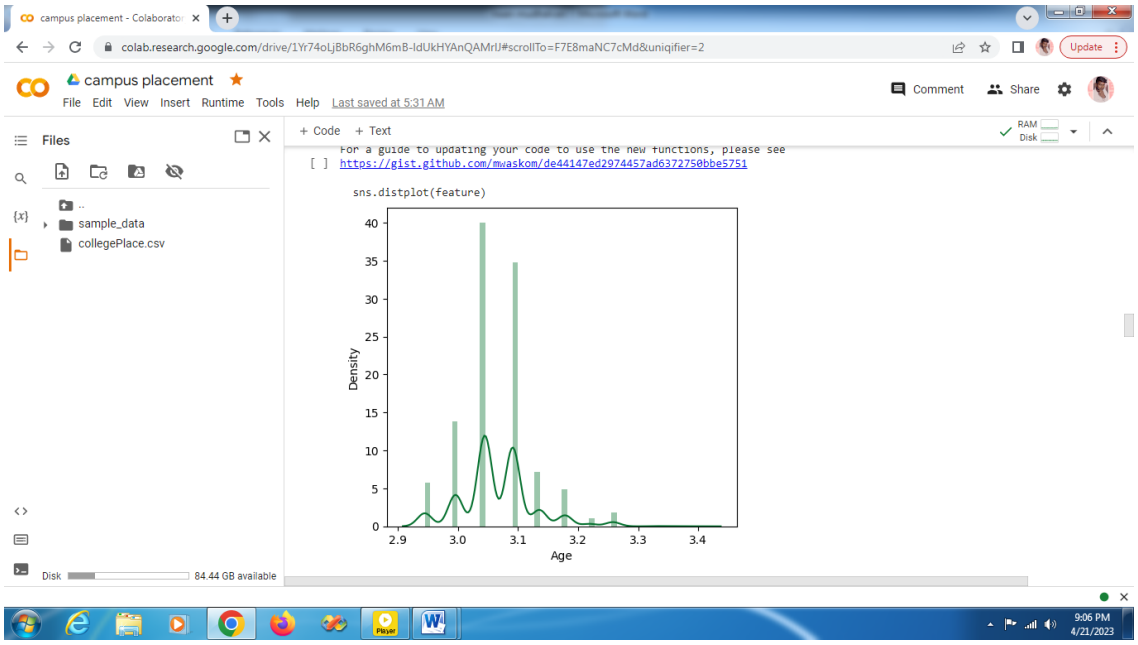

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
400
200
0
0
1
PlacedOrNot
[ ] df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2966 entries, 0 to 2965
Data columns (total 8 columns):
Column Non-Null Count Dtype
0 Age 2966 non-null int64
1 Gender 2966 non-null object
2 Stream 2966 non-null object
3 Internships 2966 non-null int64
4 CGPA 2966 non-null int64
5 Hostel 2966 non-null int64
6 HistoryOfBacklogs 2966 non-null int64
7 PlacedOrNot 2966 non-null int64
dtypes: int64(6), object(2)
memory usage: 185.5+ KB
Disk 84.44 GB available
9:01 PM
4/21/2023

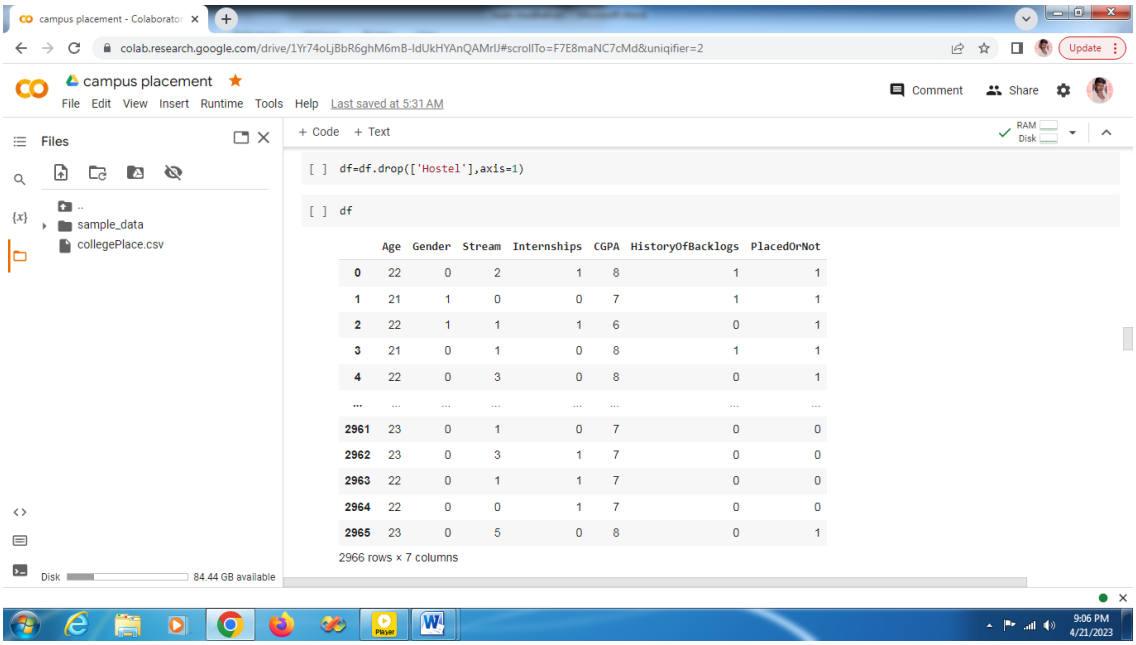

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ] df.isnull().sum()
Age 0
Gender 0
Stream 0
Internships 0
CGPA 0
Hostel 0
HistoryOfBacklogs 0
PlacedOrNot 0
dtype: int64
[ ] df.describe()
Age Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
count 2966.000000 2966.000000 2966.000000 2966.000000 2966.000000 2966.000000
mean 21.485840 0.703641 7.073837 0.269049 0.192178 0.552596
std 1.324933 0.740197 0.967748 0.443540 0.394079 0.497310
min 19.000000 0.000000 5.000000 0.000000 0.000000 0.000000
25% 21.000000 0.000000 6.000000 0.000000 0.000000 0.000000
50% 21.000000 1.000000 7.000000 0.000000 0.000000 1.000000
Disk 84.44 GB available
9:02 PM
4/21/2023

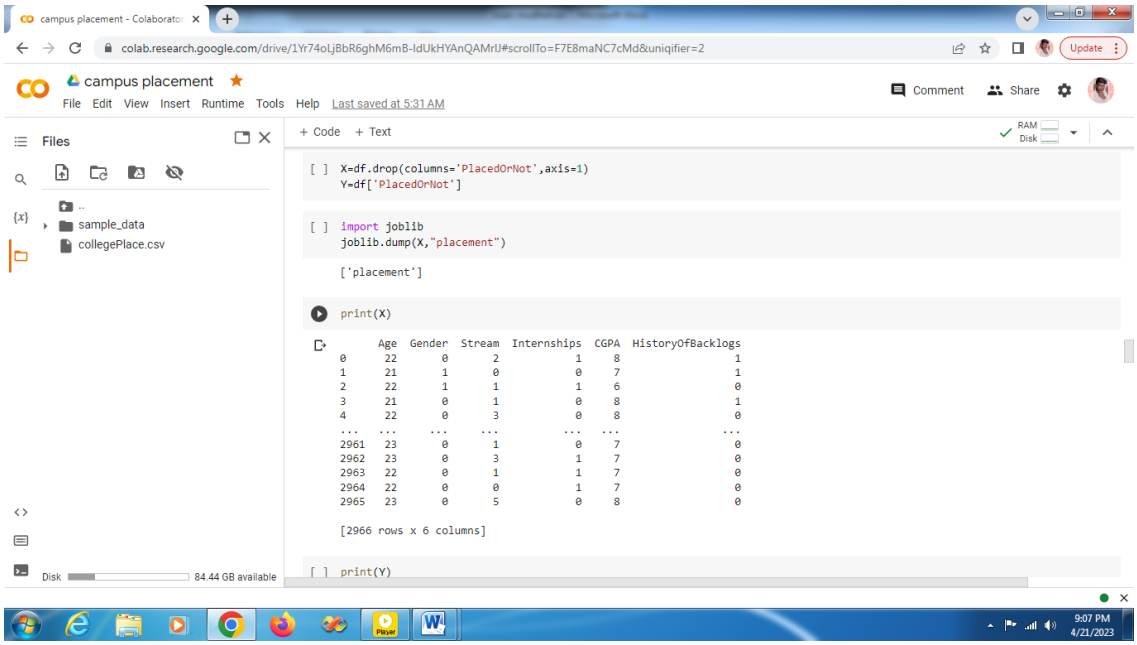

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
75% 22.000000 1.000000 8.000000 1.000000 0.000000 1.000000
max 30.000000 3.000000 9.000000 1.000000 1.000000 1.000000
df['Gender'].value_counts()
Male 2475
Female 491
Name: Gender, dtype: int64
df['Stream'].value_counts()
Computer Science 776
Information Technology 691
Electronics And Communication 424
Mechanical 424
Electrical 334
Civil 317
Name: Stream, dtype: int64
df=df.replace(['Male'],[0])
df=df.replace(['Female'],[1])
df=df.replace(['Computer Science','Information Technology','Electronics And Communication','Mechanical','Electrical','Civ
Disk
84.44 GB available
9:03 PM
4/21/2023

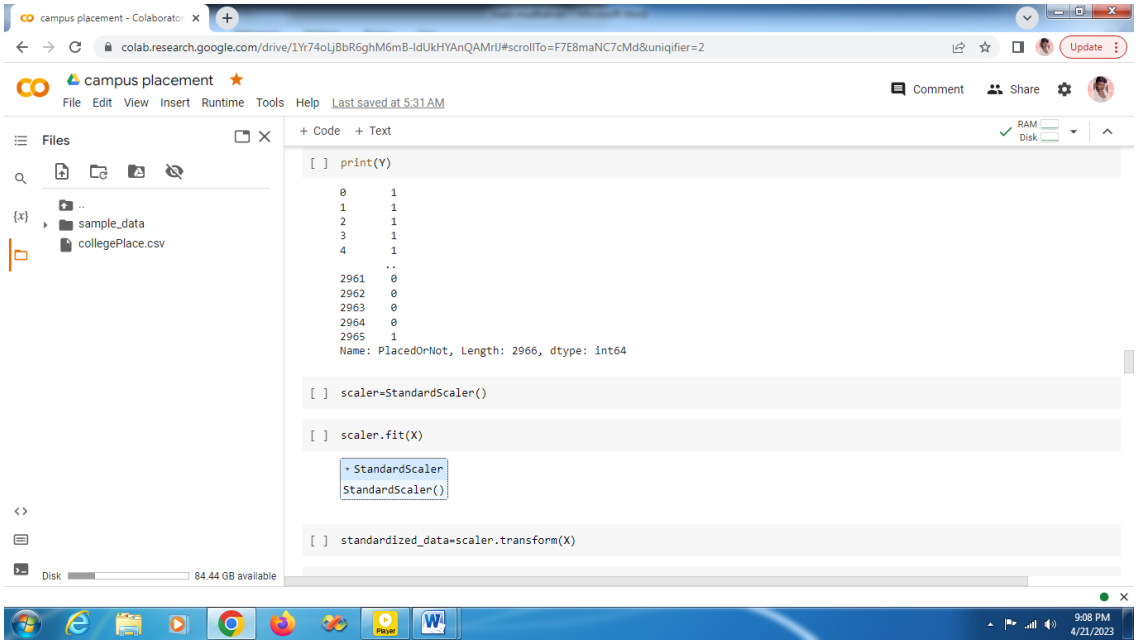

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
df
Age Gender Stream Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
0 22 0 2 1 8 1 1 1
1 21 1 0 0 7 1 1 1
2 22 1 1 1 6 0 0 1
3 21 0 1 0 8 0 1 1
4 22 0 3 0 8 1 0 1
2961 23 0 1 0 7 0 0 0
2962 23 0 3 1 7 1 0 0
2963 22 0 1 1 7 0 0 0
2964 22 0 0 1 7 0 0 0
2965 23 0 5 0 8 0 0 1
2966 rows × 8 columns
df.head()
Disk
84.44 GB available
9:04 PM
4/21/2023

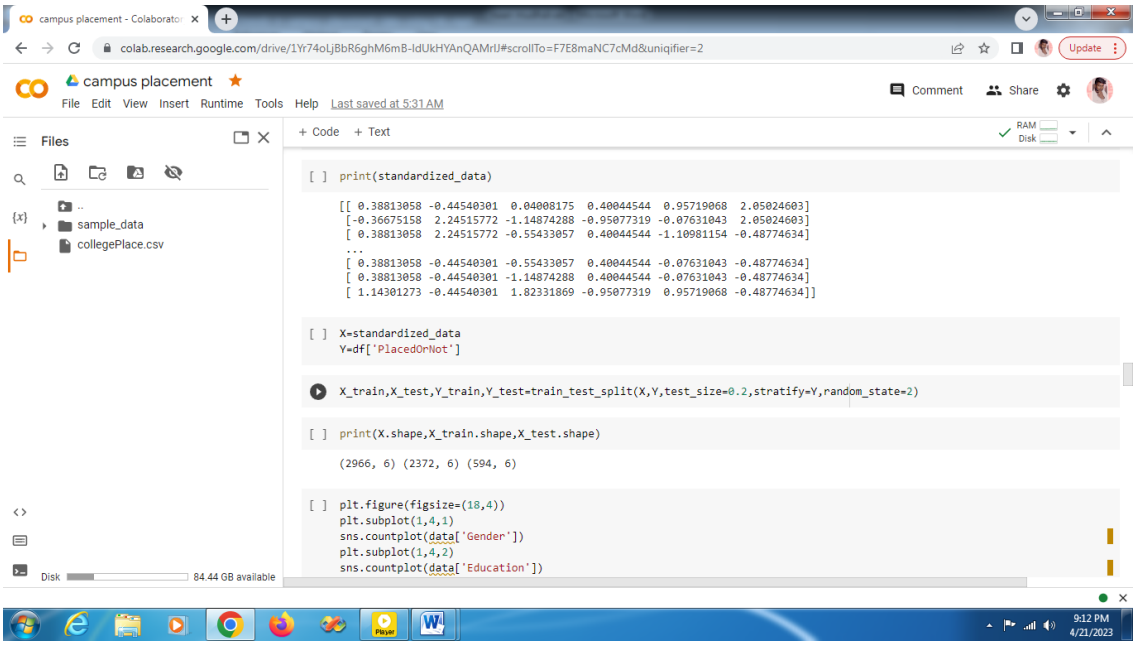

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ] df.head()
Age Gender Stream Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
0 22 0 2 1 8 1 1 1
1 21 1 0 0 7 1 1 1
2 22 1 1 1 6 0 0 1
3 21 0 1 0 8 0 1 1
4 22 0 3 0 8 1 0 1
[ ] def transformationplot(feature):
plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
sns.distplot(feature)
transformationplot(np.log(df['Age']))
<ipython-input-19-6fba3c2af64a>:4: UserWarning:
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).
Disk
84.44 GB available
9:04 PM
4/21/2023

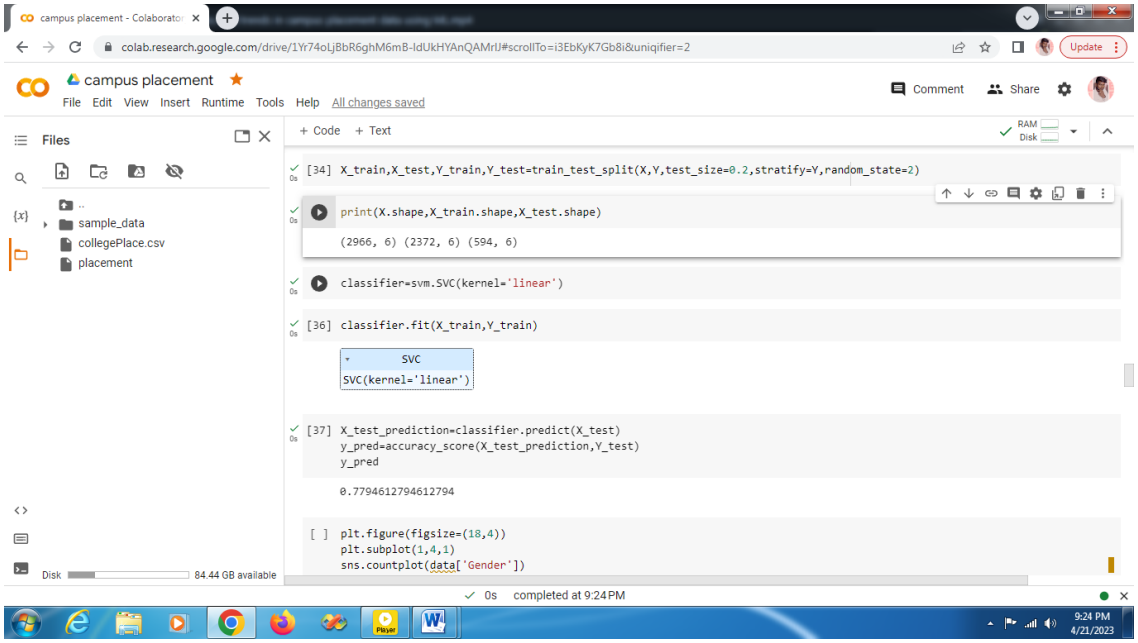

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
+ Code + Text
RAM
Disk
For a guide to updating your code to use the new functions, please see
[ ] https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
sns.distplot(feature)
sample_data
collegePlace.csv
Density
Age
0
5
10
15
20
25
30
35
40
2.9
3.0
3.1
3.2
3.3
3.4
Disk
84.44 GB available
9:06 PM
4/21/2023

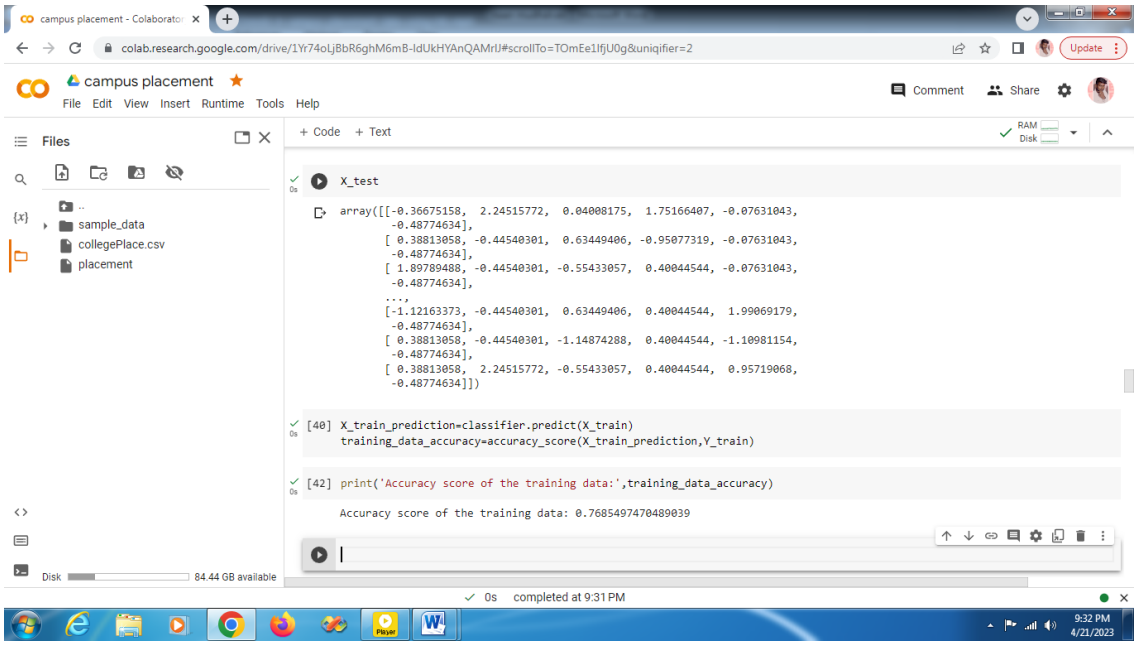

campus placement - Colaborato
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ] df=df.drop(['Hostel'],axis=1)
[ ] df
Age Gender Stream Internships CGPA HistoryOfBacklogs PlacedOrNot
0 22 0 2 1 8 1 1
1 21 1 0 0 7 1 1
2 22 1 1 1 6 0 1
3 21 0 1 0 8 1 1
4 22 0 3 0 8 0 1
... ...
2961 23 0 1 0 7 0 0
2962 23 0 3 1 7 0 0
2963 22 0 1 1 7 0 0
2964 22 0 0 1 7 0 0
2965 23 0 5 0 8 0 1
2966 rows × 7 columns
Disk 84.44 GB available
9:06 PM
4/21/2023

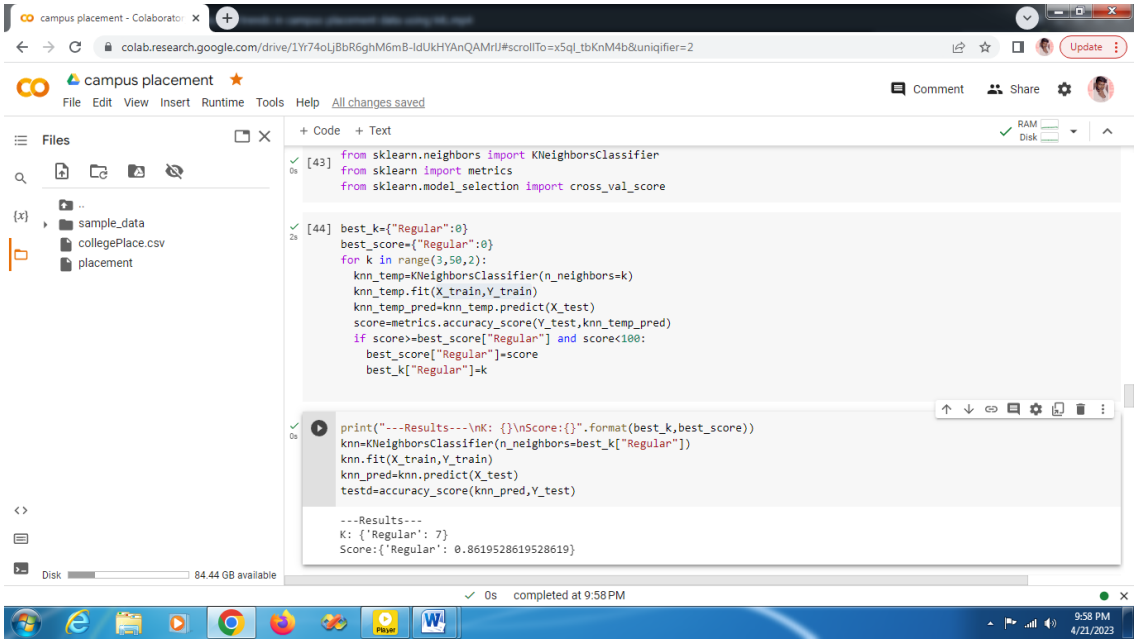

campus placement - Colaborato
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ] X=df.drop(columns='PlacedOrNot',axis=1)
Y=df['PlacedOrNot']
[ ] import joblib
joblib.dump(X,"placement")
['placement']
print(X)
Age Gender Stream Internships CGPA HistoryOfBacklogs
0 22 0 2 1 8 1
1 21 1 0 0 7 1
2 22 1 1 1 6 0
3 21 0 1 0 8 1
4 22 0 3 0 8 0
... ...
2961 23 0 1 0 7 0
2962 23 0 3 1 7 0
2963 22 0 1 1 7 0
2964 22 0 0 1 7 0
2965 23 0 5 0 8 0
[2966 rows x 6 columns]
[ ] print(Y)
Disk 84.44 GB available
9:07 PM
4/21/2023

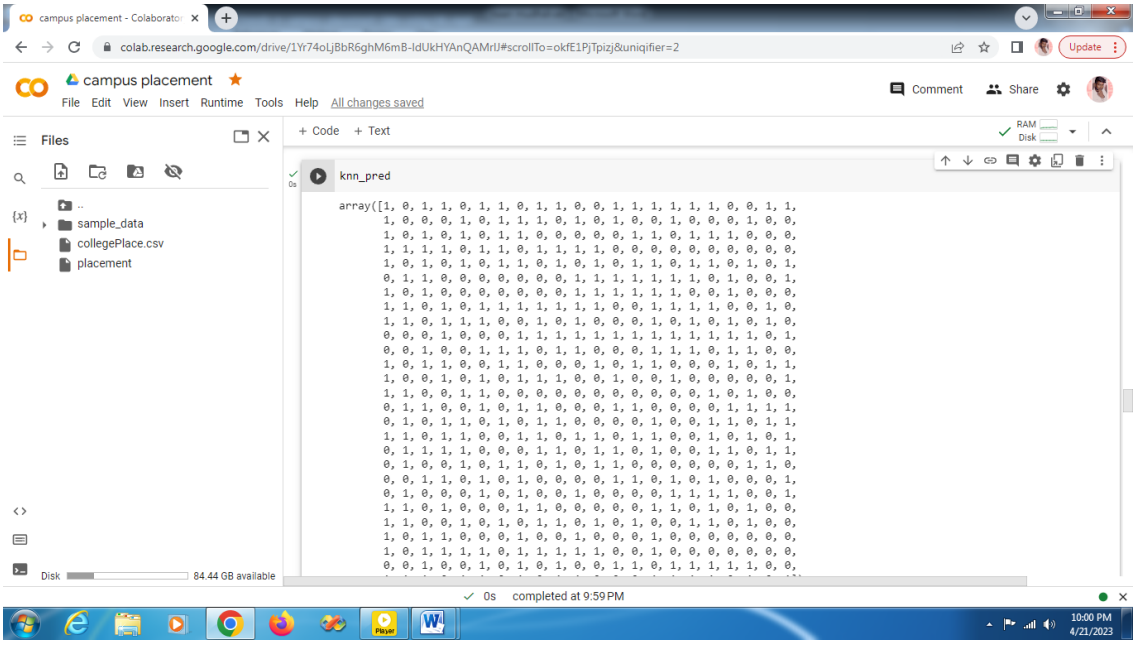

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
[ ] print(Y)
0 1
1 1
2 1
3 1
4 1
..
2961 0
2962 0
2963 0
2964 0
2965 1
Name: PlacedOrNot, Length: 2966, dtype: int64
[ ] scaler=StandardScaler()
[ ] scaler.fit(X)
StandardScaler
StandardScaler()
[ ] standardized_data=scaler.transform(X)
Disk 84.44 GB available
9:08 PM
4/21/2023

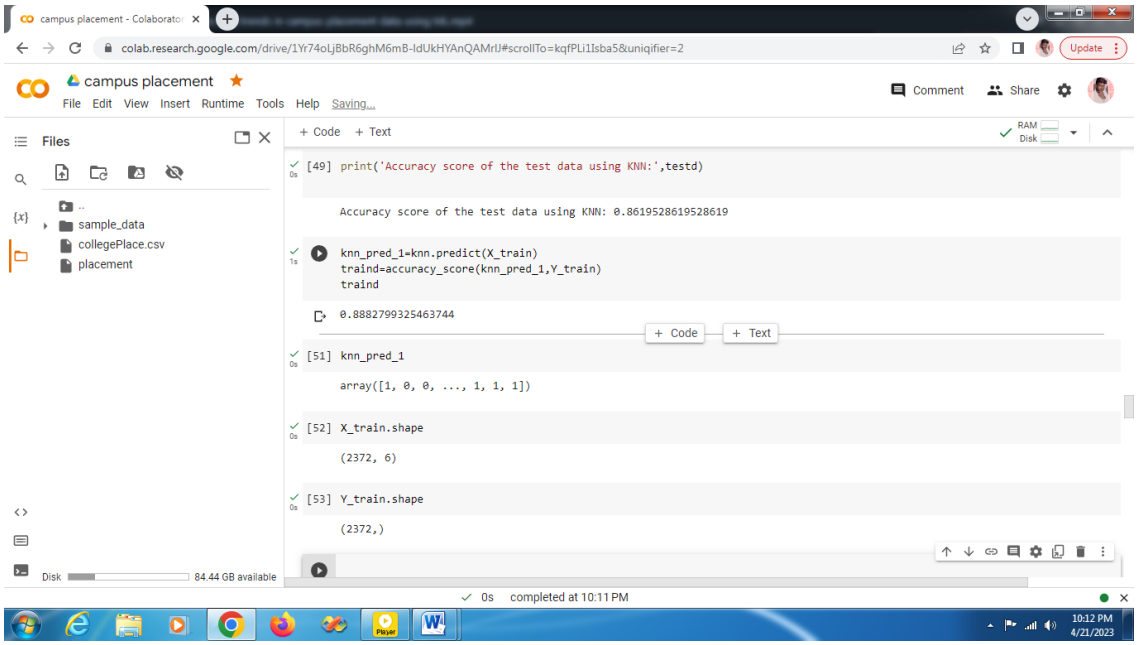

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
[ ] print(standardized_data)
[[0.38813058 -0.44540301 0.04008175 0.40044544 0.95719068 2.05024603]
[-0.36675158 2.24515772 -1.14874288 -0.95077319 -0.07631043 2.05024603]
[0.38813058 2.24515772 -0.55433057 0.40044544 -1.10981154 -0.48774634]
...
[0.38813058 -0.44540301 -0.55433057 0.40044544 -0.07631043 -0.48774634]
[0.38813058 -0.44540301 -1.14874288 0.40044544 -0.07631043 -0.48774634]
[1.14301273 -0.44540301 1.82331869 -0.95077319 0.95719068 -0.48774634]]
[ ] X=standardized_data
Y=df['PlacedOrNot']
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
[ ] print(X.shape,X_train.shape,X_test.shape)
(2966, 6) (2372, 6) (594, 6)
[ ] plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
plt.subplot(1,4,2)
sns.countplot(data['Education'])
Disk 84.44 GB available
9:12 PM
4/21/2023

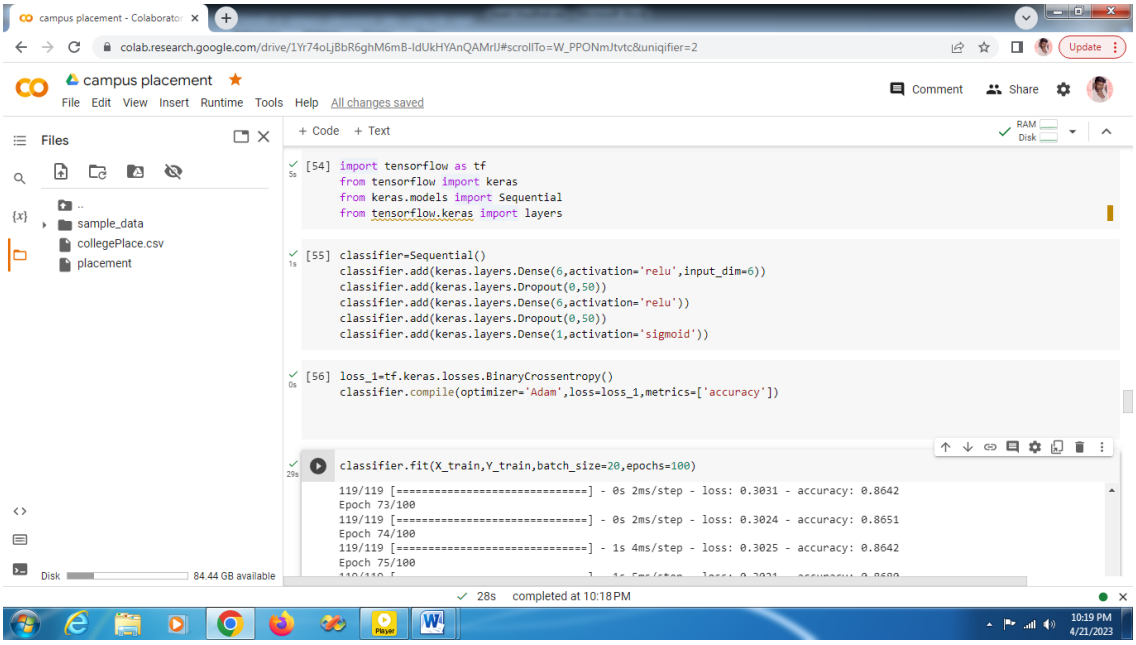

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=i3EbKyK7Gb8i&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Update
Files
sample_data
collegePlace.csv
placement
+ Code + Text
[34] X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
print(X.shape,X_train.shape,X_test.shape)
(2966, 6) (2372, 6) (594, 6)
classifier=svm.SVC(kernel='linear')
[36] classifier.fit(X_train,Y_train)
SVC
SVC(kernel='linear')
[37] X_test_prediction=classifier.predict(X_test)
y_pred=accuracy_score(X_test_prediction,Y_test)
y_pred
0.7794612794612794
[ ] plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
Disk 84.44 GB available
0s completed at 9:24 PM
9:24 PM
4/21/2023

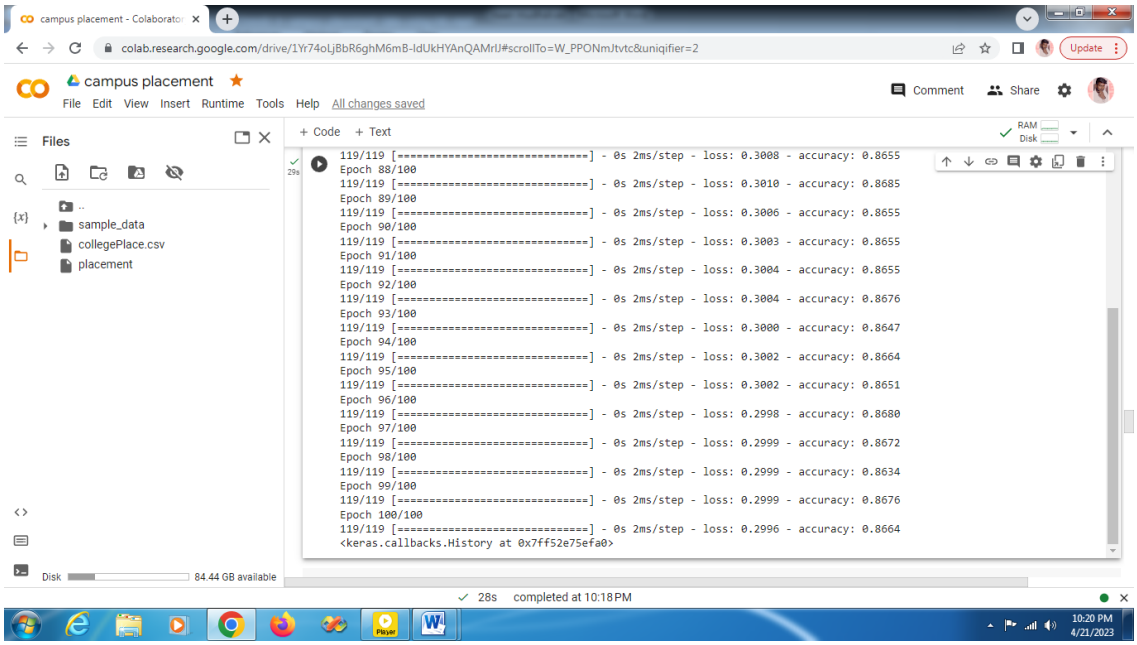

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=TOmEe1IfjU0g&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help
Comment
Share
Update
Files
sample_data
collegePlace.csv
placement
+ Code + Text
X_test
array([[-0.36675158, 2.24515772, 0.04008175, 1.75166407, -0.07631043,
-0.48774634],
[0.38813058, -0.44540301, 0.63449406, -0.95077319, -0.07631043,
-0.48774634],
[1.89789488, -0.44540301, -0.55433057, 0.40044544, -0.07631043,
-0.48774634],
...,
[-1.12163373, -0.44540301, 0.63449406, 0.40044544, 1.99069179,
-0.48774634],
[0.38813058, -0.44540301, -1.14874288, 0.40044544, -1.10981154,
-0.48774634],
[0.38813058, 2.24515772, -0.55433057, 0.40044544, 0.95719068,
-0.48774634]])
[40] X_train_prediction=classifier.predict(X_train)
training_data_accuracy=accuracy_score(X_train_prediction,Y_train)
[42] print('Accuracy score of the training data:',training_data_accuracy)
Accuracy score of the training data: 0.7685497470489039
Disk 84.44 GB available
0s completed at 9:31 PM
9:32 PM
4/21/2023

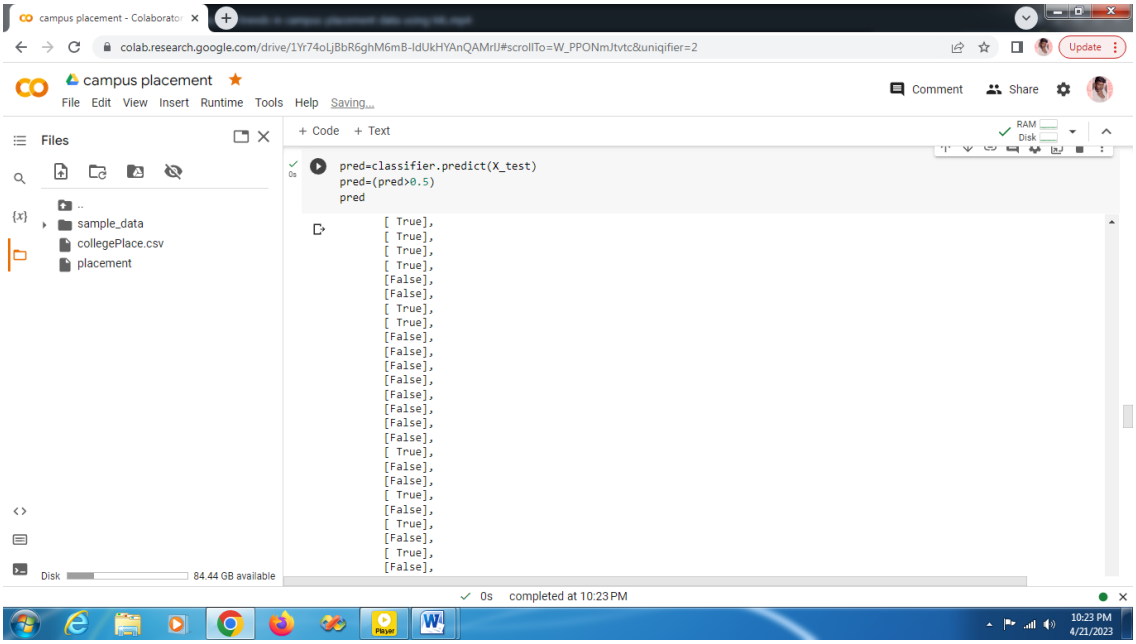

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=x5ql_tbKnM4b&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Update
Files
sample_data
collegePlace.csv
placement
+ Code + Text
RAM
Disk
[43] from sklearn.neighbors import KNeighborsClassifier
from sklearn import metrics
from sklearn.model_selection import cross_val_score
[44] best_k={"Regular":0}
best_score={"Regular":0}
for k in range(3,50,2):
knn_temp=KNeighborsClassifier(n_neighbors=k)
knn_temp.fit(X_train,Y_train)
knn_temp_pred=knn_temp.predict(X_test)
score=metrics.accuracy_score(Y_test,knn_temp_pred)
if score>=best_score["Regular"] and score<100:
best_score["Regular"]=score
best_k["Regular"]=k
print("---Results---\nK: {}\nScore:{}".format(best_k,best_score))
knn=KNeighborsClassifier(n_neighbors=best_k["Regular"])
knn.fit(X_train,Y_train)
knn_pred=knn.predict(X_test)
testd=accuracy_score(knn_pred,Y_test)
---Results---
K: {'Regular': 7}
Score:{'Regular': 0.8619528619528619}
Disk
84.44 GB available
0s completed at 9:58 PM
9:58 PM
4/21/2023

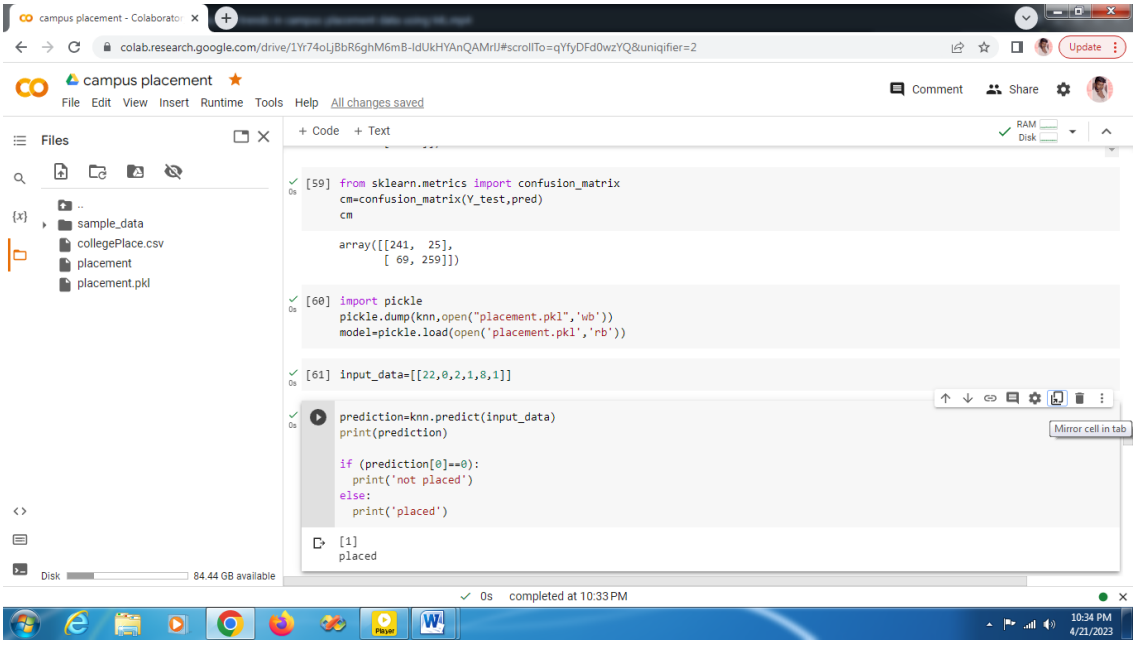

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=okfE1PjTpizj&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Update
Files
sample_data
collegePlace.csv
placement
+ Code + Text
RAM
Disk
knn_pred
array([1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1,
1, 0, 0, 0, 1, 0, 1, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0,
1, 0, 1, 0, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0, 1, 1, 1, 0, 0, 0,
1, 1, 1, 1, 0, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0,
1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 0, 1,
0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 1,
1, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0,
1, 1, 0, 1, 0, 1, 1, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 1, 1, 0, 0, 1, 0,
1, 1, 0, 1, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1, 0,
0, 0, 0, 1, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 0, 1,
0, 0, 1, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0, 0, 1, 1, 1, 0, 1, 1, 1, 0, 0,
1, 0, 1, 1, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1,
1, 0, 0, 1, 0, 1, 0, 1, 1, 1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 0, 1,
1, 1, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0,
0, 1, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 1, 1,
0, 1, 0, 1, 1, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 1, 0, 1, 1, 1,
1, 1, 0, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 1, 0, 1,
0, 1, 1, 1, 1, 0, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 0, 1, 1,
0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 1, 0, 1, 1, 0, 0, 0, 0, 0, 1, 1, 0,
0, 0, 1, 0, 1, 0, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 0, 1,
0, 1, 0, 0, 0, 1, 0, 1, 1, 0, 0, 1, 0, 0, 0, 0, 1, 1, 1, 1, 0, 0, 1,
1, 1, 0, 1, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0,
1, 1, 0, 0, 1, 0, 1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 1, 0, 0,
1, 0, 1, 1, 0, 0, 0, 1, 0, 0, 1, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0,
1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0,
0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 1, 0, 0, 1, 1, 1, 1, 1, 0, 0,
Disk
84.44 GB available
0s completed at 9:59 PM
10:00 PM
4/21/2023

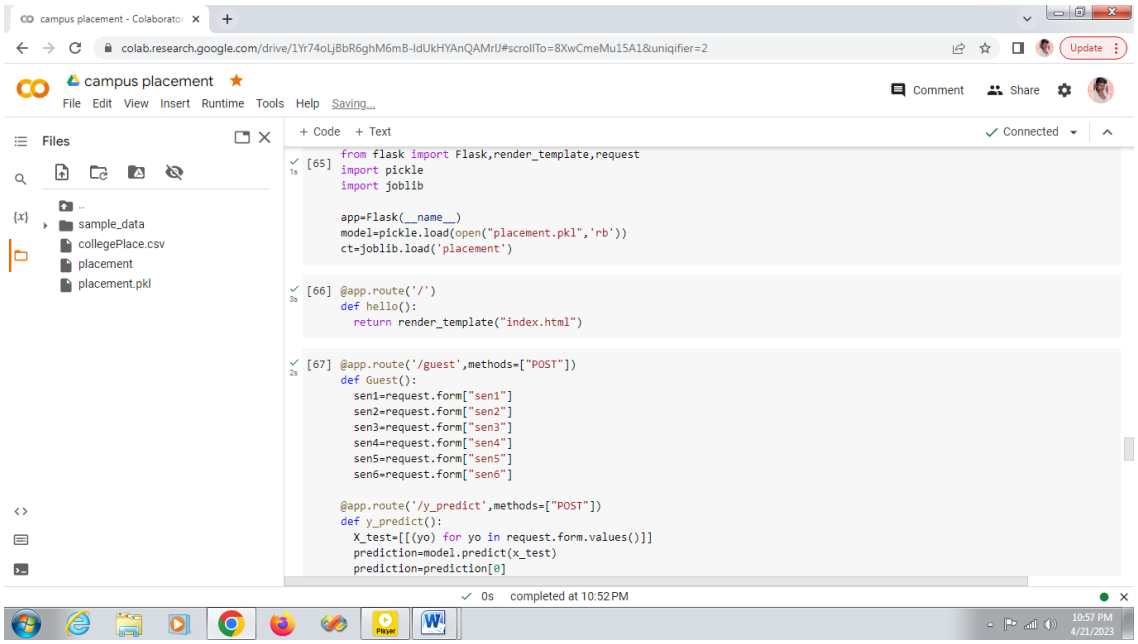

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=kqfPLi1Isba5&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
sample_data
collegePlace.csv
placement
[49] print('Accuracy score of the test data using KNN:',testd)
Accuracy score of the test data using KNN: 0.8619528619528619
knn_pred_1=knn.predict(X_train)
traind=accuracy_score(knn_pred_1,Y_train)
traind
0.8882799325463744
[51] knn_pred_1
array([1, 0, 0, ..., 1, 1, 1])
[52] X_train.shape
(2372, 6)
[53] Y_train.shape
(2372,)
Disk 84.44 GB available
0s completed at 10:11 PM
10:12 PM
4/21/2023

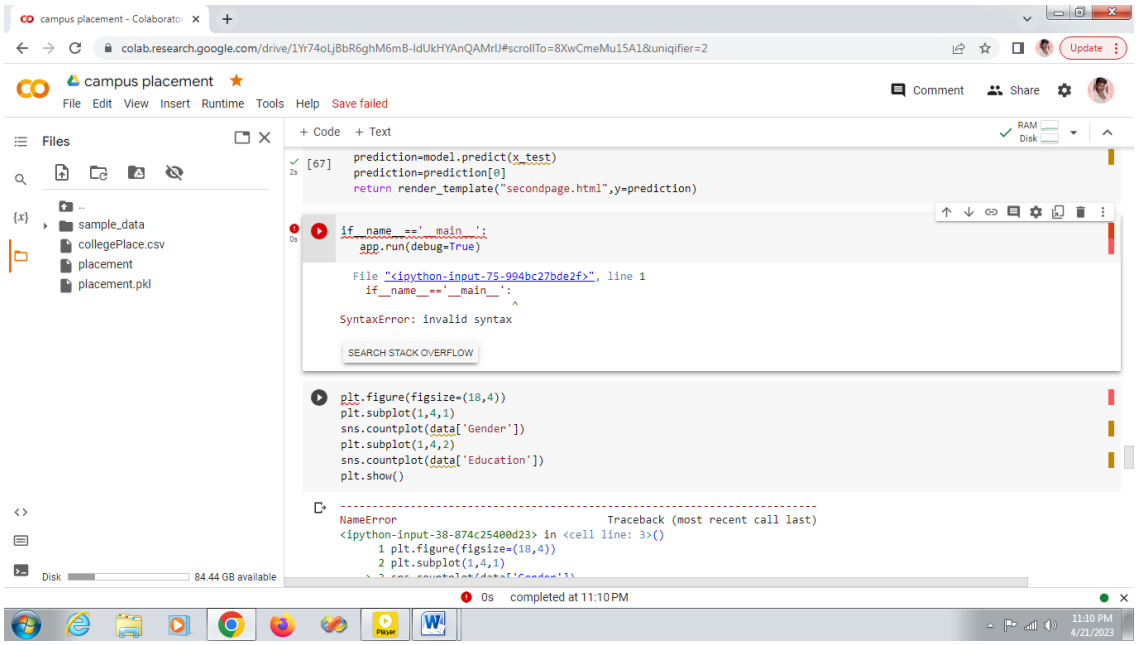

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=W_PPONmJtvtc&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
sample_data
collegePlace.csv
placement
[54] import tensorflow as tf
from tensorflow import keras
from keras.models import Sequential
from tensorflow.keras import layers
[55] classifier=Sequential()
classifier.add(keras.layers.Dense(6,activation='relu',input_dim=6))
classifier.add(keras.layers.Dropout(0,50))
classifier.add(keras.layers.Dense(6,activation='relu'))
classifier.add(keras.layers.Dropout(0,50))
classifier.add(keras.layers.Dense(1,activation='sigmoid'))
[56] loss_1=tf.keras.losses.BinaryCrossentropy()
classifier.compile(optimizer='Adam',loss=loss_1,metrics=['accuracy'])
classifier.fit(X_train,Y_train,batch_size=20,epochs=100)
119/119 [==============================] - 0s 2ms/step - loss: 0.3031 - accuracy: 0.8642
Epoch 73/100
119/119 [==============================] - 0s 2ms/step - loss: 0.3024 - accuracy: 0.8651
Epoch 74/100
119/119 [==============================] - 1s 4ms/step - loss: 0.3025 - accuracy: 0.8642
Epoch 75/100
Disk 84.44 GB available
28s completed at 10:18 PM
10:19 PM
4/21/2023


campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrIJ#scrollTo=W_PPONmJtvtc&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
placement
119/119 [==========] - 0s 2ms/step - loss: 0.3008 - accuracy: 0.8655
Epoch 88/100
119/119 [==========] - 0s 2ms/step - loss: 0.3010 - accuracy: 0.8685
Epoch 89/100
119/119 [==========] - 0s 2ms/step - loss: 0.3006 - accuracy: 0.8655
Epoch 90/100
119/119 [==========] - 0s 2ms/step - loss: 0.3003 - accuracy: 0.8655
Epoch 91/100
119/119 [==========] - 0s 2ms/step - loss: 0.3004 - accuracy: 0.8655
Epoch 92/100
119/119 [==========] - 0s 2ms/step - loss: 0.3004 - accuracy: 0.8676
Epoch 93/100
119/119 [==========] - 0s 2ms/step - loss: 0.3000 - accuracy: 0.8647
Epoch 94/100
119/119 [==========] - 0s 2ms/step - loss: 0.3002 - accuracy: 0.8664
Epoch 95/100
119/119 [==========] - 0s 2ms/step - loss: 0.3002 - accuracy: 0.8651
Epoch 96/100
119/119 [==========] - 0s 2ms/step - loss: 0.2998 - accuracy: 0.8680
Epoch 97/100
119/119 [==========] - 0s 2ms/step - loss: 0.2999 - accuracy: 0.8672
Epoch 98/100
119/119 [==========] - 0s 2ms/step - loss: 0.2999 - accuracy: 0.8634
Epoch 99/100
119/119 [==========] - 0s 2ms/step - loss: 0.2999 - accuracy: 0.8676
Epoch 100/100
119/119 [==========] - 0s 2ms/step - loss: 0.2996 - accuracy: 0.8664
<keras.callbacks.History at 0x7ff52e75efa0>
Disk
84.44 GB available
28s completed at 10:18 PM
10:20 PM
4/21/2023


campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrIJ#scrollTo=W_PPONmJtvtc&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
placement
pred=classifier.predict(X_test)
pred=(pred>0.5)
pred
[True],
[True],
[True],
[True],
[False],
[False],
[True],
[True],
[False],
[False],
[False],
[False],
[False],
[False],
[False],
[False],
[False],
[True],
[False],
[False],
[True],
[False],
[True],
[False],
[True],
[False],
Disk
84.44 GB available
0s completed at 10:23 PM
10:23 PM
4/21/2023


campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=qYfyDFd0wzYQ&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
placement
placement.pkl
[59] from sklearn.metrics import confusion_matrix
cm=confusion_matrix(Y_test,pred)
cm
array([[241, 25],
[69, 259]])
[60] import pickle
pickle.dump(knn,open("placement.pkl",'wb'))
model=pickle.load(open('placement.pkl','rb'))
[61] input_data=[[22,0,2,1,8,1]]
prediction=knn.predict(input_data)
print(prediction)
if (prediction[0]==0):
print('not placed')
else:
print('placed')
Mirror cell in tab
[1]
placed
Disk
84.44 GB available
0s completed at 10:33 PM
10:34 PM
4/21/2023


campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=8XwCmeMu15A1&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Saving...
Comment
Share
Files
+ Code + Text
Connected
sample_data
collegePlace.csv
placement
placement.pkl
[65] from flask import Flask,render_template,request
import pickle
import joblib
app=Flask(__name__)
model=pickle.load(open("placement.pkl",'rb'))
ct=joblib.load('placement')
[66] @app.route('/')
def hello():
return render_template("index.html")
[67] @app.route('/guest',methods=["POST"])
def Guest():
sen1=request.form["sen1"]
sen2=request.form["sen2"]
sen3=request.form["sen3"]
sen4=request.form["sen4"]
sen5=request.form["sen5"]
sen6=request.form["sen6"]
@app.route('/y_predict',methods=["POST"])
def y_predict():
X_test=[[(yo) for yo in request.form.values()]]
prediction=model.predict(x_test)
prediction=prediction[0]
0s completed at 10:52 PM
10:57 PM
4/21/2023



4. ADVANTAGES & DISADVANTAGES:

DISADVANTAGES:

- They used normal conditional statements of getting information
- They don't have proper algorithms for analyzing data for new placements

ADVANTAGES :

Predict the results and we then compare the efficiency of the algorithms, which is based on the dataset. This model helps the position cell at intervals a corporation to spot the potential students and concentrate to and improve their technical and social skills.

5.APPLICATIONS:

To identify the talented and qualified students in the college. To create promising career opportunities for students in reputed corporate companies. To select candidates who are suitable for the current job roles without any biased behavior.

To provide recruitment to students. To provide exposure to BE pursuing students. To have good relations with the recruiters. Managing Recruiters correspondence and feedback's.

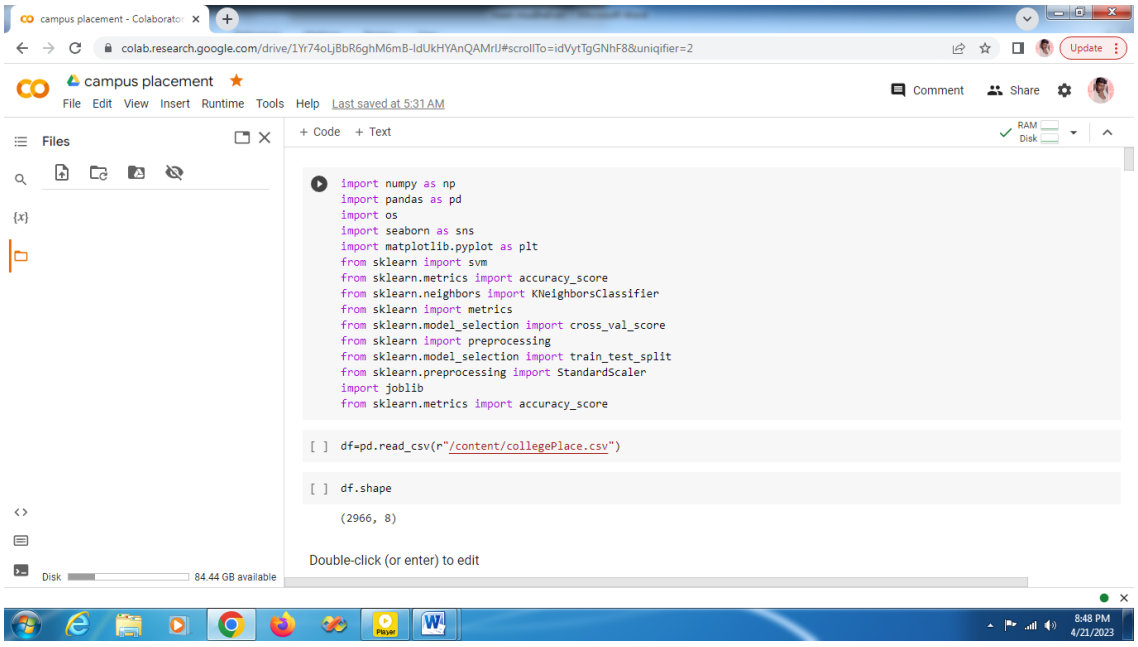
6.CONCLUSION:

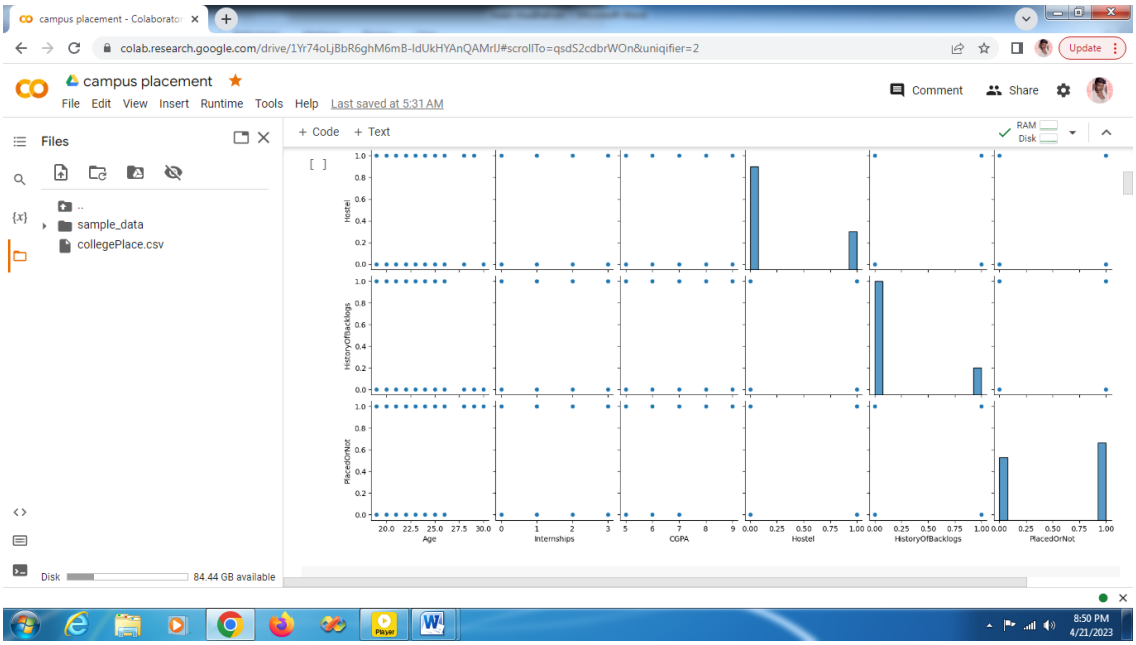
Predicting the placement of a student gives an idea to the Placement Office as well as the student on where they stand. Not all companies look for similar talents. If the strengths and weaknesses of the students are identified it would benefit the student in getting placed. The placement Office can work on identifying the weaknesses of the students and take measures of improvement so that the students can overcome the weakness and perform to the best of their abilities. Thus the key lies in assessing the capabilities of the student in the right areas and subjecting them to the right training

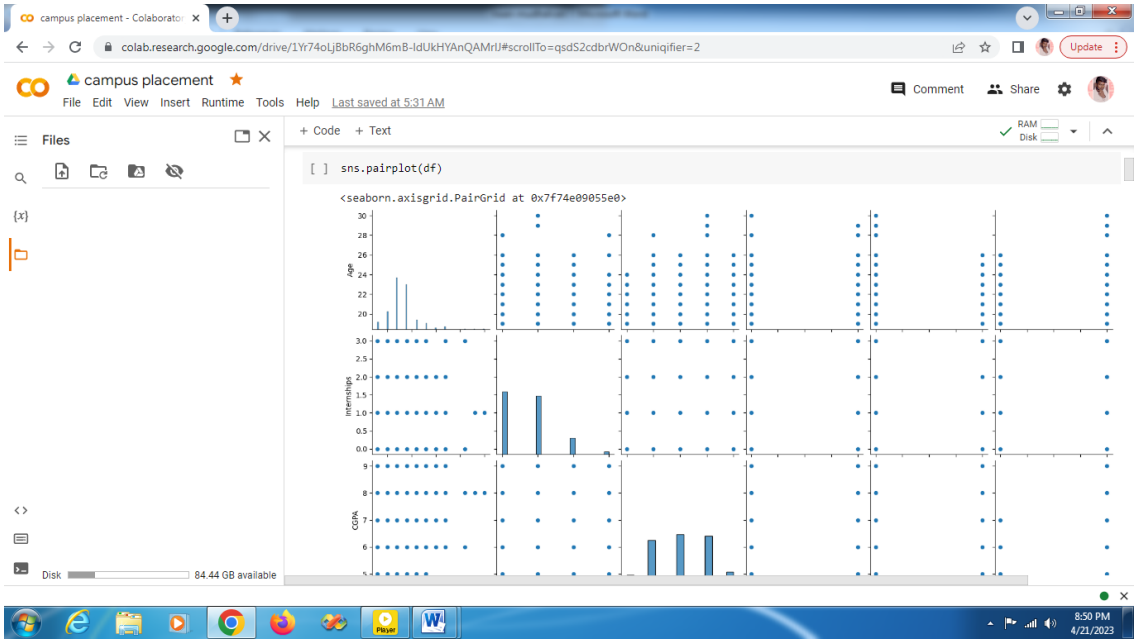
7. FUTURE SCOPE

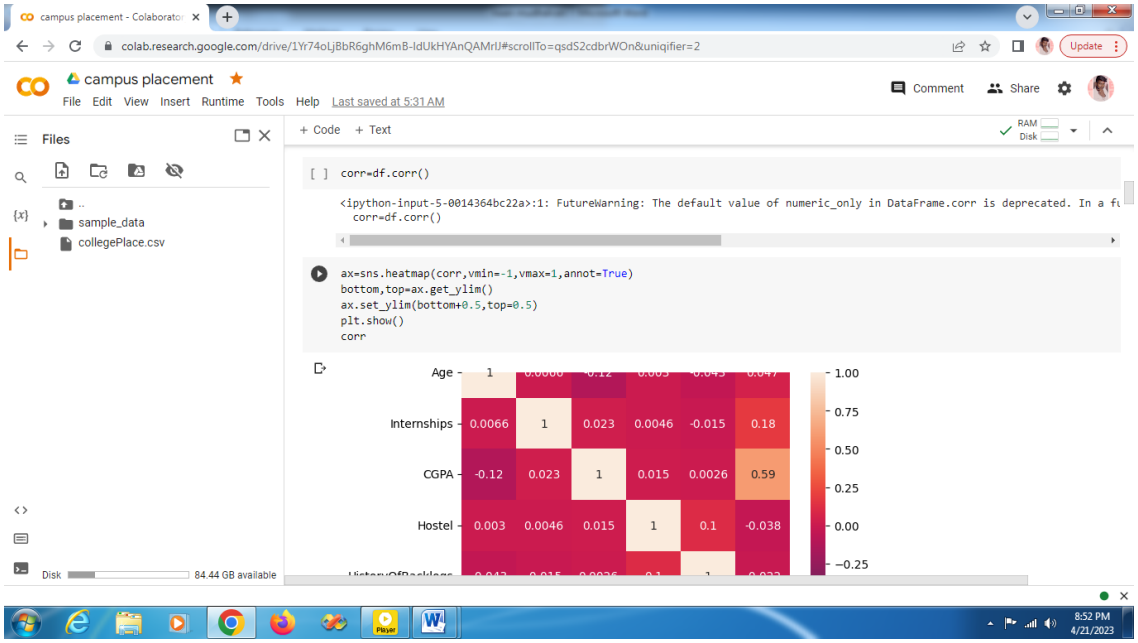
The above work was carried out with respect to the Information Science Engineering and Computer Science Engineering branch of our college. Further work can be carried out by applying other algorithms that could lead to improvement in results, also different Skill sets for above mentioned branches and on data of different streams of engineering. The key to this would be to identify the aspects of these engineering branches that would test the student's knowledge in getting placed in a core company for their respective branches.

8.APPENDIX

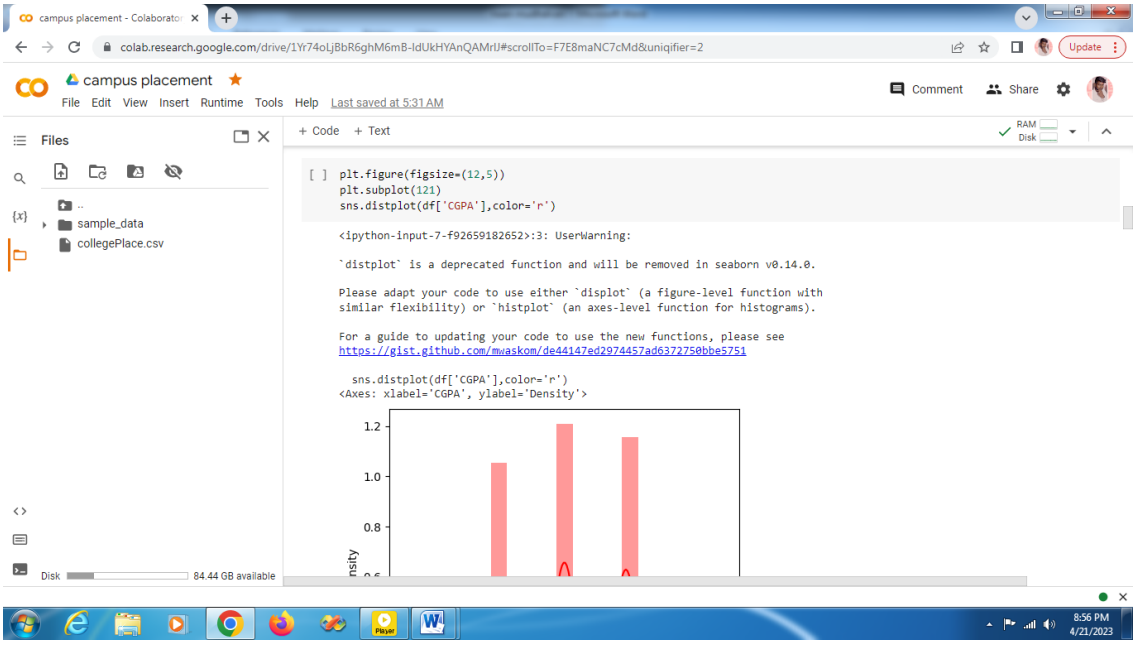


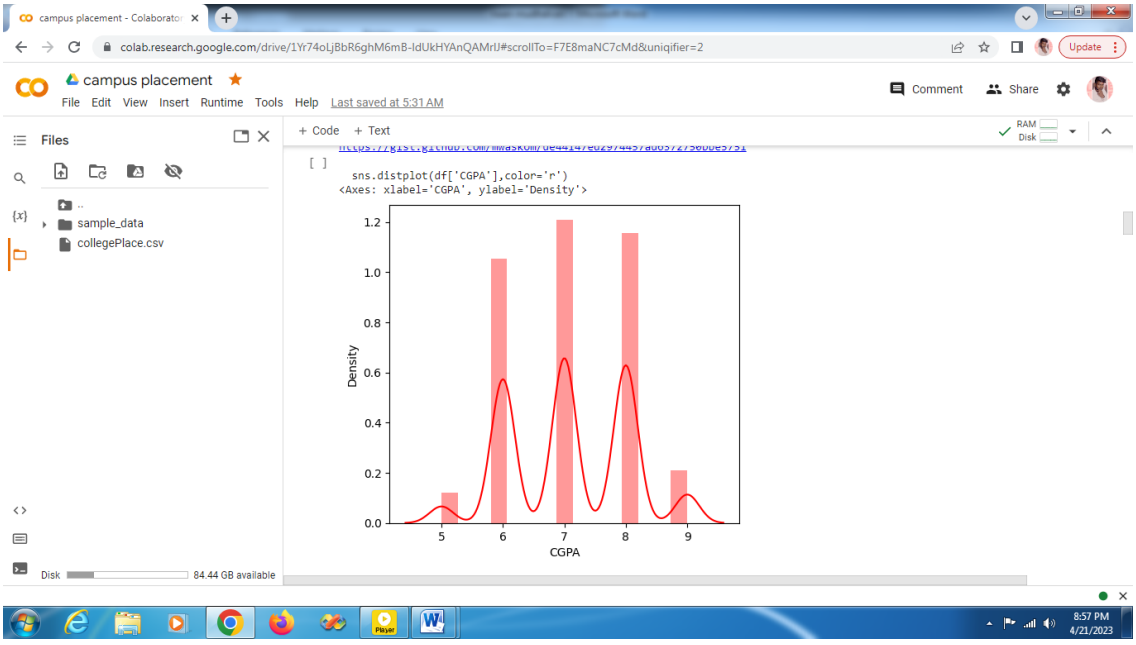

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=qsdS2cdbrWOn&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
Hostel
HistoryOfBacklogs
PlacedOrNot
Age
Internships
CGPA
Hostel
HistoryOfBacklogs
PlacedOrNot
Disk 84.44 GB available
8:50 PM
4/21/2023

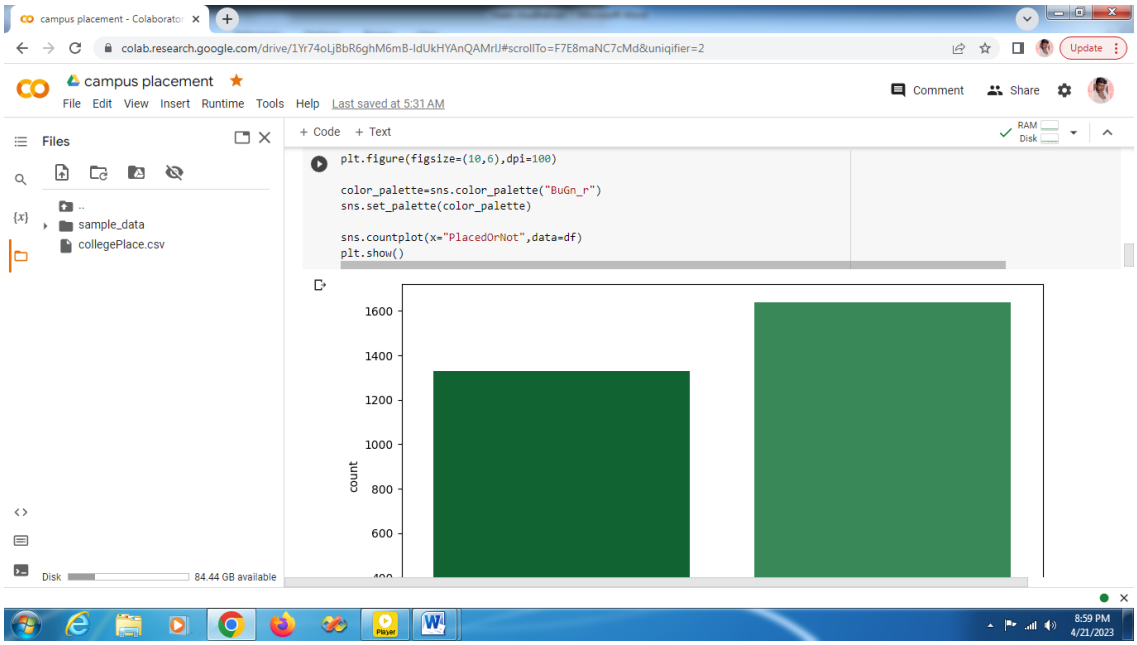

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=qsdS2cdbrWOn&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
[ ] sns.pairplot(df)
<seaborn.axisgrid.PairGrid at 0x7f74e09055e0>
Age
Internships
CGPA
Disk 84.44 GB available
8:50 PM
4/21/2023

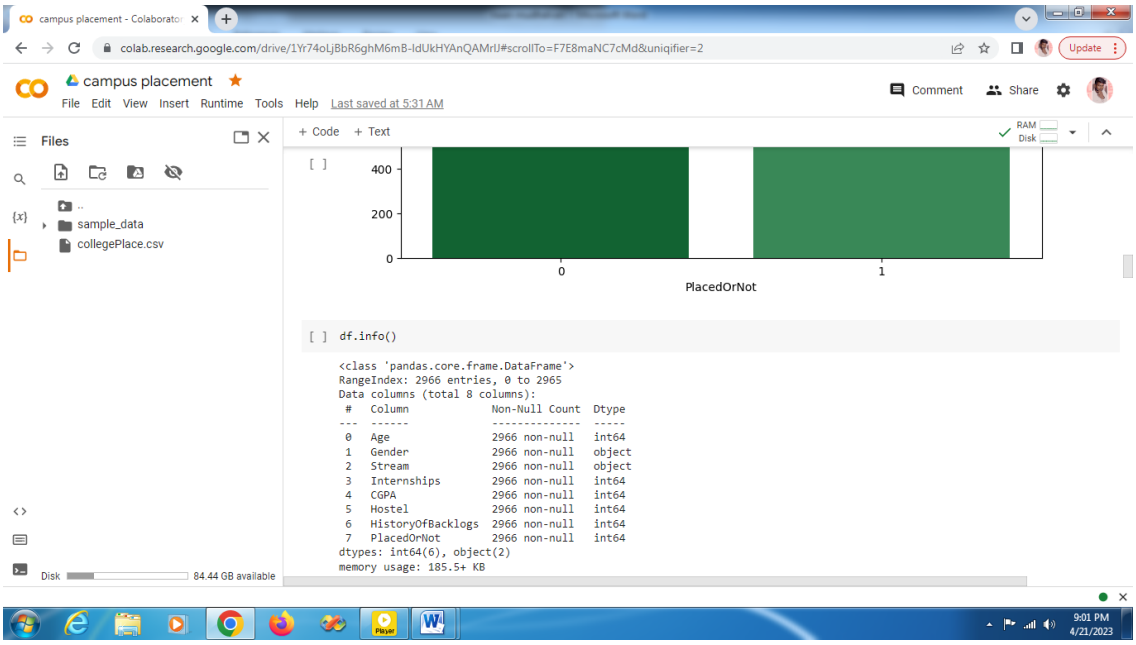

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=qsdS2cdbrWOn&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
[ ] corr=df.corr()
<ipython-input-5-0014364bc22a>:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a fu
corr=df.corr()
ax=sns.heatmap(corr,vmin=-1,vmax=1,annot=True)
bottom,top=ax.get_ylim()
ax.set_ylim(bottom+0.5,top=0.5)
plt.show()
corr
Age
Internships
CGPA
Hostel
0.0066 1 0.023 0.0046 -0.015 0.18
-0.12 0.023 1 0.015 0.0026 0.59
0.003 0.0046 0.015 1 0.1 -0.038
1.00
0.75
0.50
0.25
0.00
−0.25
Disk
84.44 GB available
8:52 PM
4/21/2023

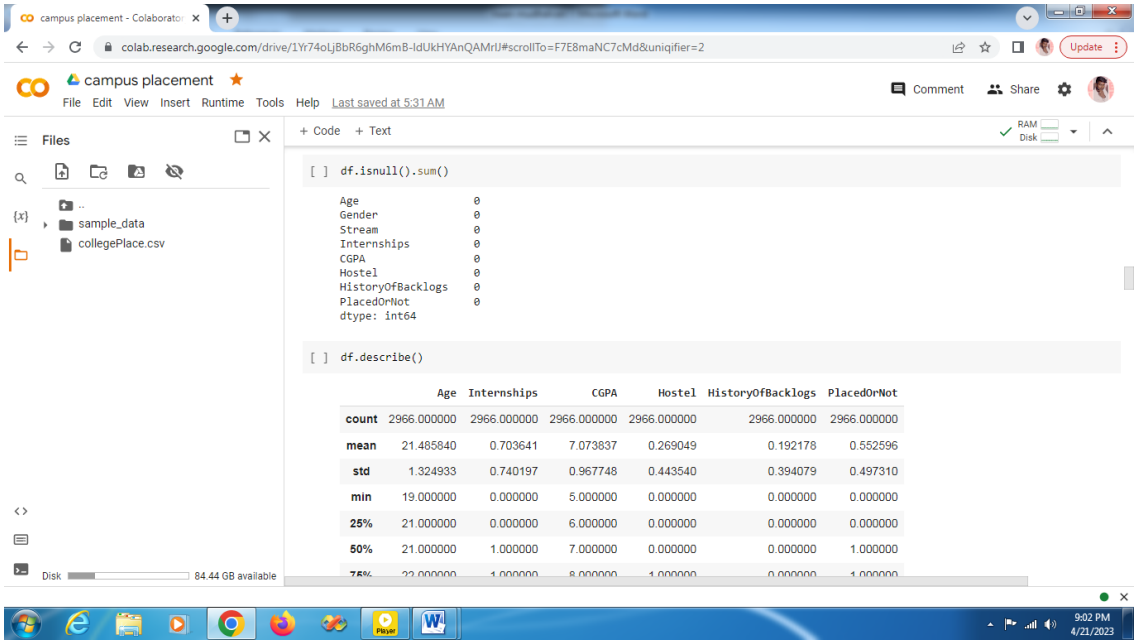

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=qsdS2cdbrWOn&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
HistoryOfBacklogs -0.043 -0.015 0.0026 0.1 1 -0.022
PlacedOrNot 0.047 0.18 0.59 -0.038 -0.022 1
−0.25
−0.50
−0.75
−1.00
Age
Internships
CGPA
Hostel
HistoryOfBacklogs
PlacedOrNot
Age Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
Age 1.000000 0.006552 -0.119787 0.003042 -0.042586 0.046943
Internships 0.006552 1.000000 0.023496 0.004617 -0.015118 0.179334
CGPA -0.119787 0.023496 1.000000 0.014991 0.002576 0.588648
Hostel 0.003042 0.004617 0.014991 1.000000 0.103506 -0.038182
HistoryOfBacklogs -0.042586 -0.015118 0.002576 0.103506 1.000000 -0.022337
PlacedOrNot 0.046943 0.179334 0.588648 -0.038182 -0.022337 1.000000
Disk
84.44 GB available
8:53 PM
4/21/2023

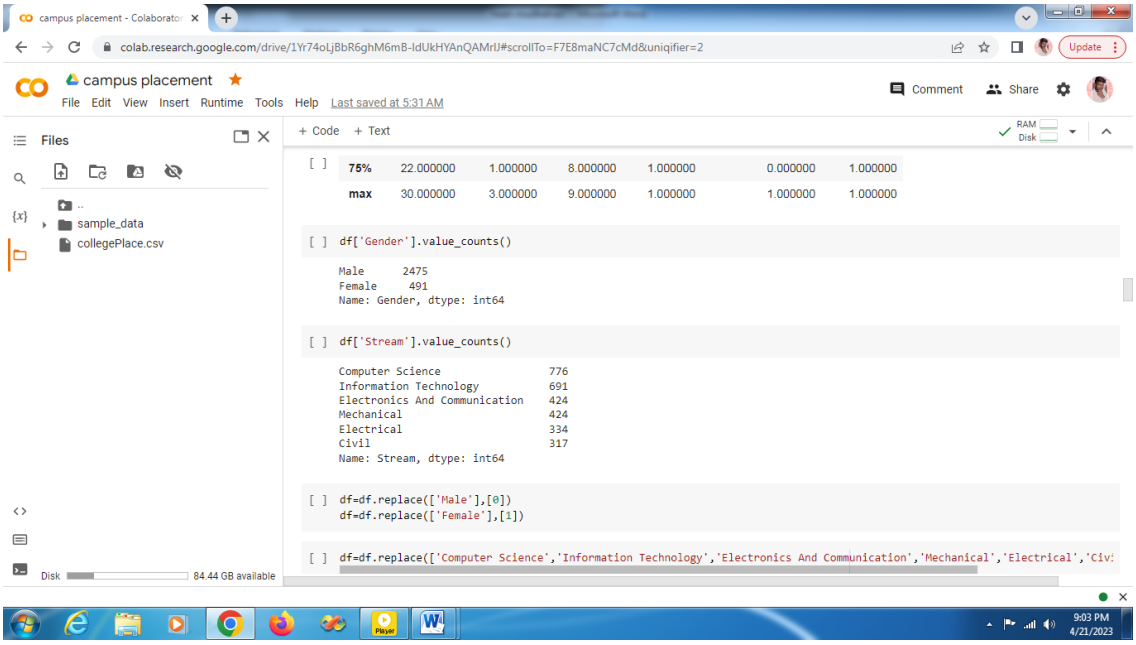

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ] plt.figure(figsize=(12,5))
plt.subplot(121)
sns.distplot(df['CGPA'],color='r')
<ipython-input-7-f92659182652>:3: UserWarning:
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).
For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
sns.distplot(df['CGPA'],color='r')
<Axes: xlabel='CGPA', ylabel='Density'>
1.2
1.0
0.8
Disk
84.44 GB available
8:56 PM
4/21/2023

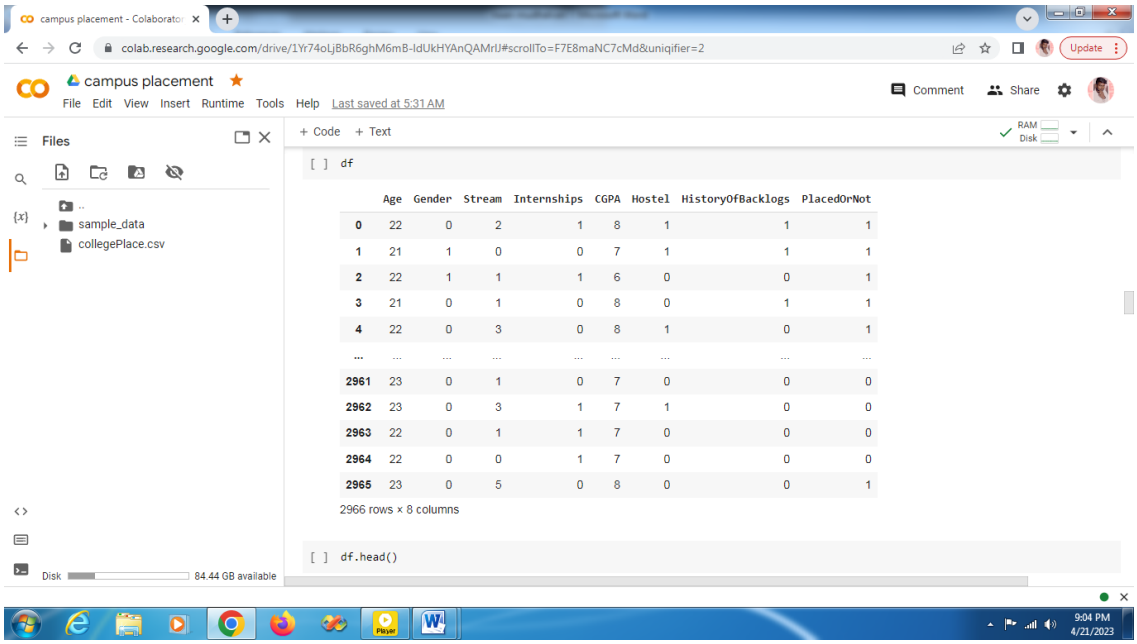

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
[ ]
sns.distplot(df['CGPA'],color='r')
<Axes: xlabel='CGPA', ylabel='Density'>
Density
1.2
1.0
0.8
0.6
0.4
0.2
0.0
5
6
7
8
9
CGPA
Disk
84.44 GB available
8:57 PM
4/21/2023

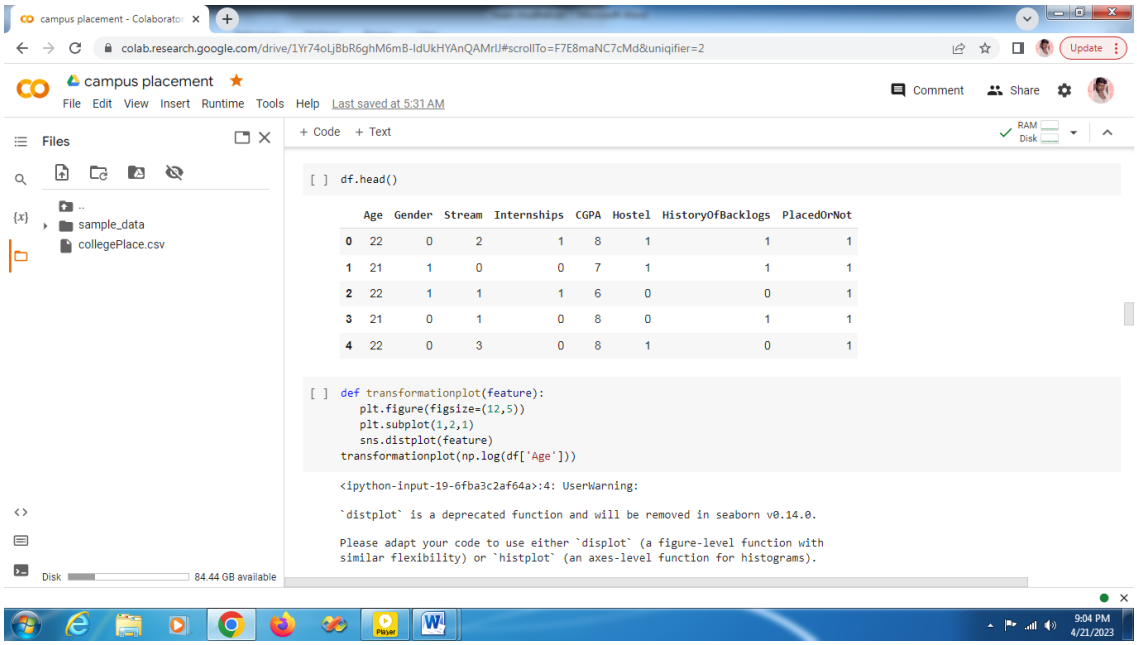

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
RAM
Disk
Files
+ Code + Text
sample_data
collegePlace.csv
plt.figure(figsize=(10,6),dpi=100)
color_palette=sns.color_palette("BuGn_r")
sns.set_palette(color_palette)
sns.countplot(x="PlacedOrNot",data=df)
plt.show()
1600
1400
1200
1000
800
600
count
Disk
84.44 GB available
8:59 PM
4/21/2023

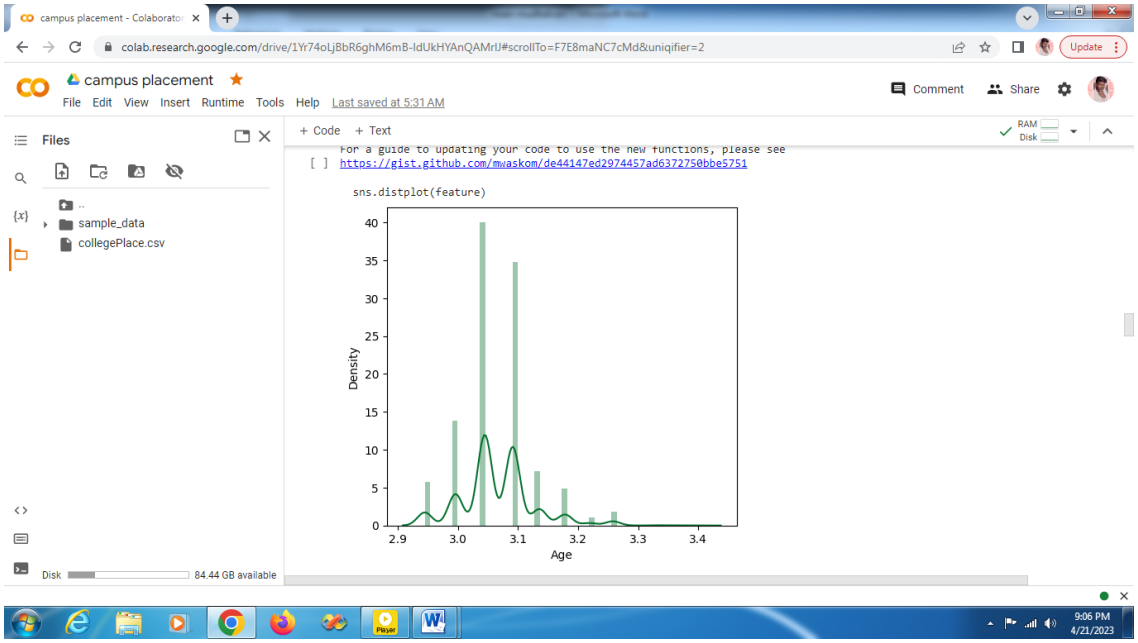

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
RAM
Disk
Files
+ Code + Text
sample_data
collegePlace.csv
400
200
0
0
1
PlacedOrNot
df.info()
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2966 entries, 0 to 2965
Data columns (total 8 columns):
Column Non-Null Count Dtype
0 Age 2966 non-null int64
1 Gender 2966 non-null object
2 Stream 2966 non-null object
3 Internships 2966 non-null int64
4 CGPA 2966 non-null int64
5 Hostel 2966 non-null int64
6 HistoryOfBacklogs 2966 non-null int64
7 PlacedOrNot 2966 non-null int64
dtypes: int64(6), object(2)
memory usage: 185.5+ KB
Disk
84.44 GB available
9:01 PM
4/21/2023

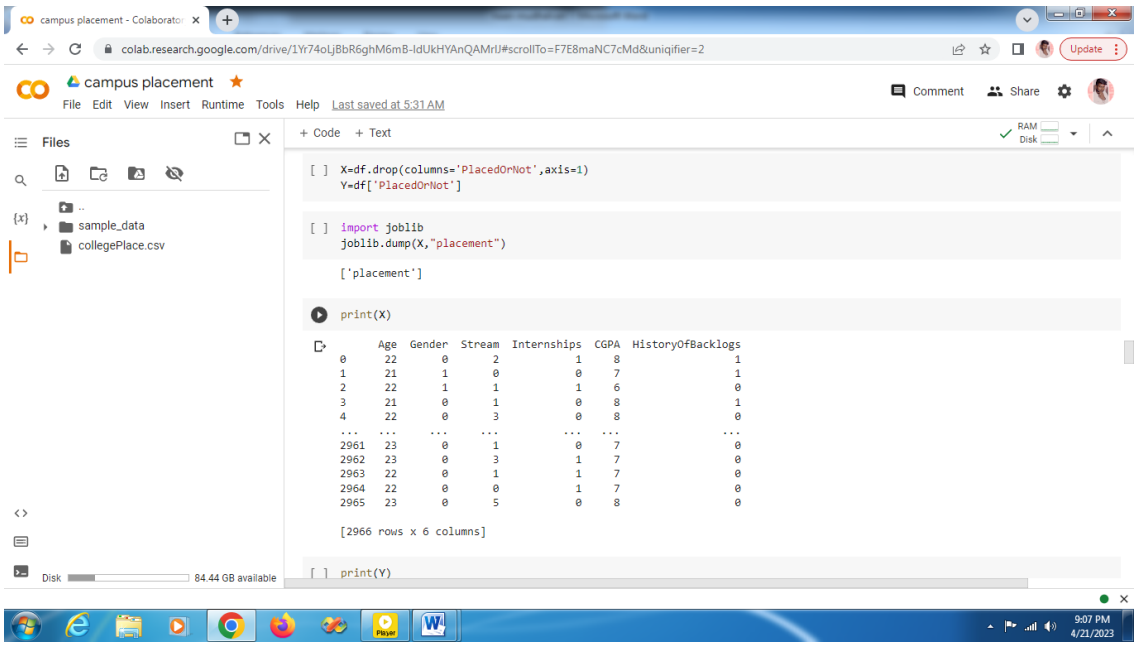

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
RAM
Disk
[ ] df.isnull().sum()
Age 0
Gender 0
Stream 0
Internships 0
CGPA 0
Hostel 0
HistoryOfBacklogs 0
PlacedOrNot 0
dtype: int64
[ ] df.describe()
Age Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
count 2966.000000 2966.000000 2966.000000 2966.000000 2966.000000 2966.000000
mean 21.485840 0.703641 7.073837 0.269049 0.192178 0.552596
std 1.324933 0.740197 0.967748 0.443540 0.394079 0.497310
min 19.000000 0.000000 5.000000 0.000000 0.000000 0.000000
25% 21.000000 0.000000 6.000000 0.000000 0.000000 0.000000
50% 21.000000 1.000000 7.000000 0.000000 0.000000 1.000000
Disk
84.44 GB available
9:02 PM
4/21/2023

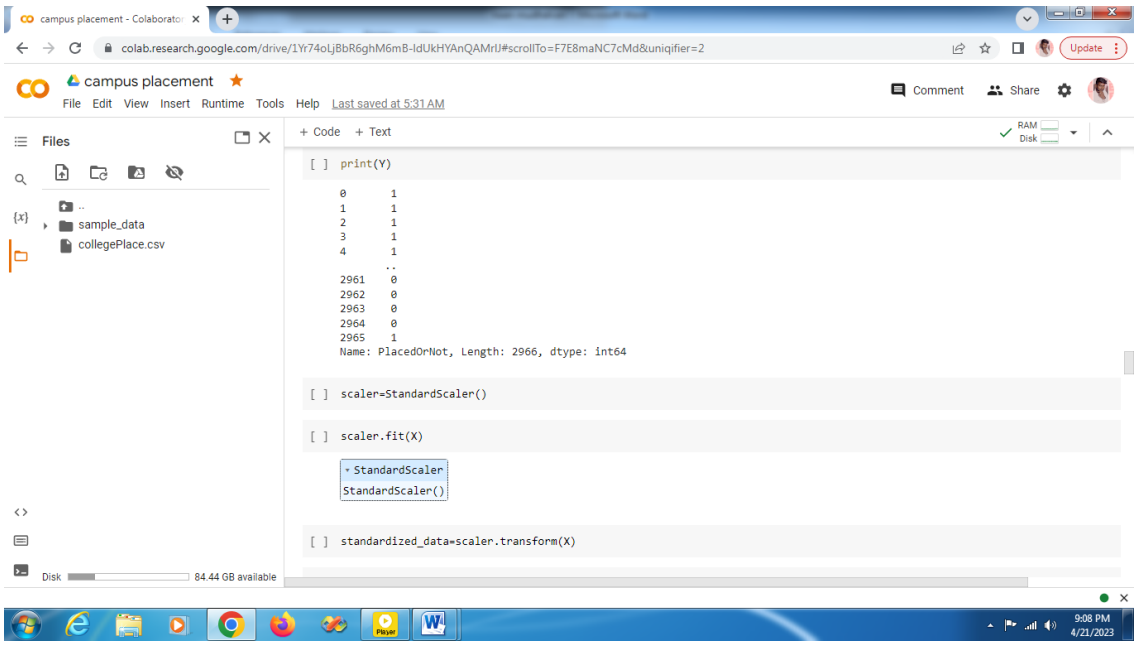

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-IdUkHYAnQAMrIJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Update
Files
sample_data
collegePlace.csv
+ Code + Text
RAM
Disk
75% 22.000000 1.000000 8.000000 1.000000 0.000000 1.000000
max 30.000000 3.000000 9.000000 1.000000 1.000000 1.000000
[ ] df['Gender'].value_counts()
Male 2475
Female 491
Name: Gender, dtype: int64
[ ] df['Stream'].value_counts()
Computer Science 776
Information Technology 691
Electronics And Communication 424
Mechanical 424
Electrical 334
Civil 317
Name: Stream, dtype: int64
[ ] df=df.replace(['Male'],[0])
df=df.replace(['Female'],[1])
[ ] df=df.replace(['Computer Science','Information Technology','Electronics And Communication','Mechanical','Electrical','Civ
Disk
84.44 GB available
9:03 PM
4/21/2023

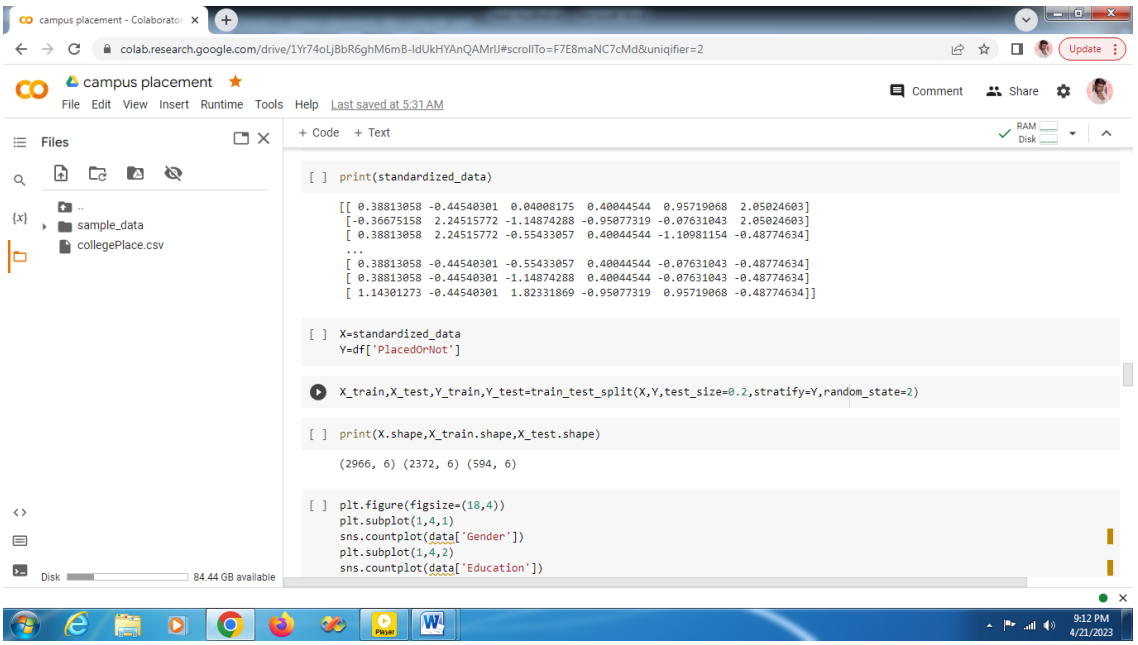

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrIJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment Share
Files
+ Code + Text
sample_data
collegePlace.csv
[ ] df
Age Gender Stream Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
0 22 0 2 1 8 1 1 1
1 21 1 0 0 7 1 1 1
2 22 1 1 1 6 0 0 1
3 21 0 1 0 8 0 1 1
4 22 0 3 0 8 1 0 1
... ...
2961 23 0 1 0 7 0 0 0
2962 23 0 3 1 7 1 0 0
2963 22 0 1 1 7 0 0 0
2964 22 0 0 1 7 0 0 0
2965 23 0 5 0 8 0 0 1
2966 rows × 8 columns
[ ] df.head()
Disk 84.44 GB available
9:04 PM
4/21/2023

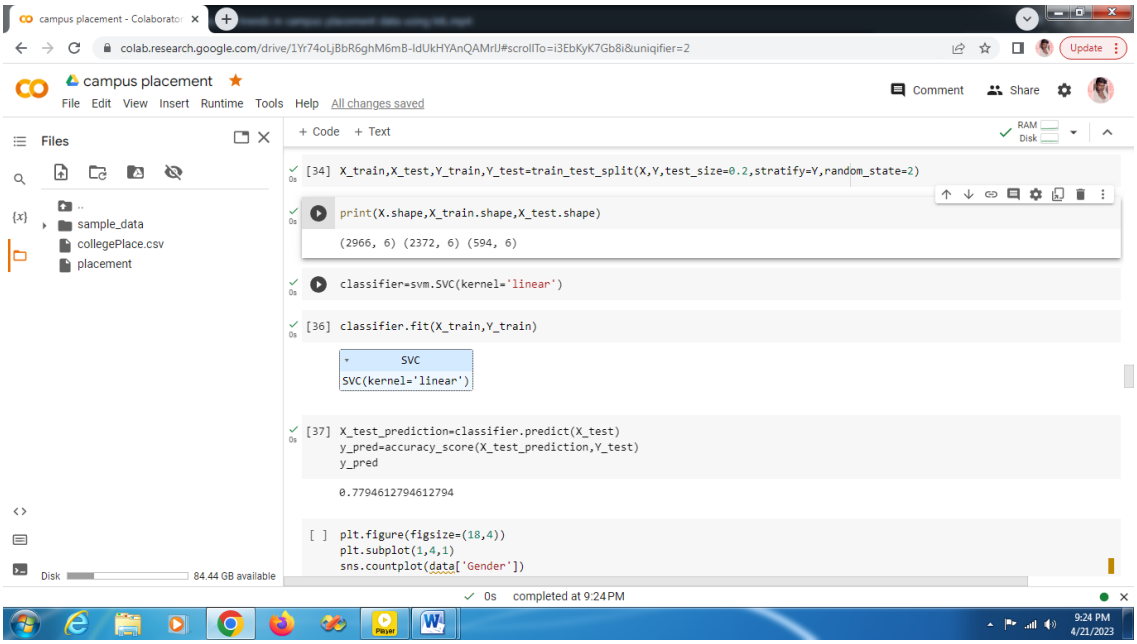

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrIJ#scrollTo=F7E8maNC7cMd&uniqifier=2
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment Share
Files
+ Code + Text
sample_data
collegePlace.csv
[ ] df.head()
Age Gender Stream Internships CGPA Hostel HistoryOfBacklogs PlacedOrNot
0 22 0 2 1 8 1 1 1
1 21 1 0 0 7 1 1 1
2 22 1 1 1 6 0 0 1
3 21 0 1 0 8 0 1 1
4 22 0 3 0 8 1 0 1
[ ] def transformationplot(feature):
plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
sns.distplot(feature)
transformationplot(np.log(df['Age']))
<ipython-input-19-6fba3c2af64a>:4: UserWarning:
`distplot` is a deprecated function and will be removed in seaborn v0.14.0.
Please adapt your code to use either `displot` (a figure-level function with
similar flexibility) or `histplot` (an axes-level function for histograms).
Disk 84.44 GB available
9:04 PM
4/21/2023

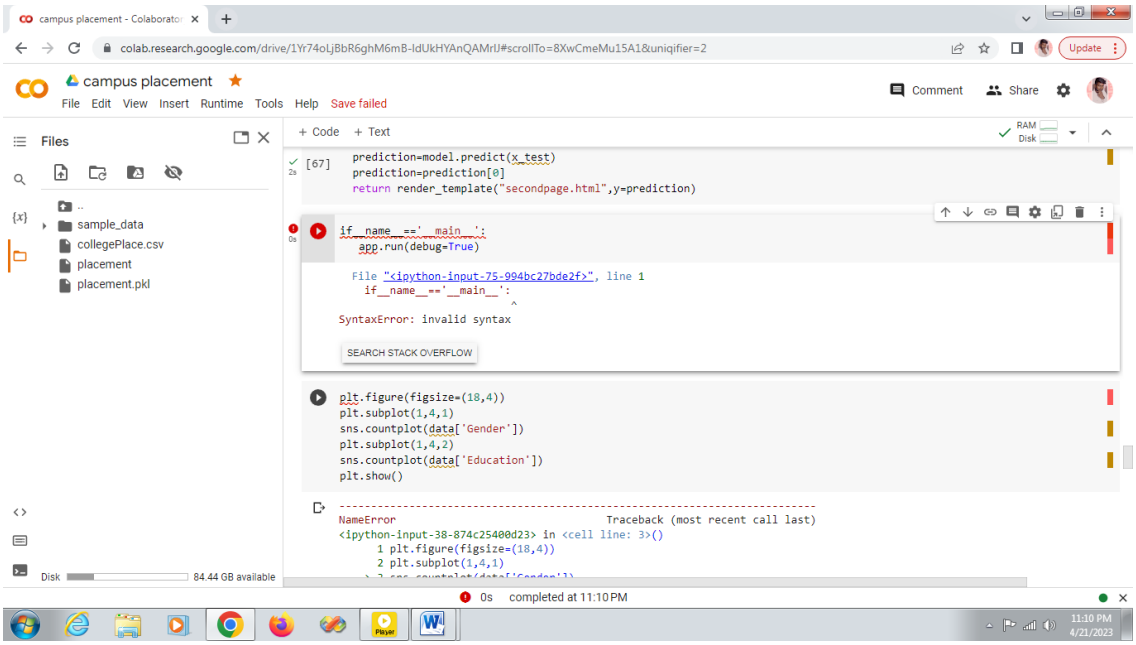

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
For a guide to updating your code to use the new functions, please see
[ ] https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751
sns.distplot(feature)
sample_data
collegePlace.csv
Density
Age
Disk 84.44 GB available
9:06 PM
4/21/2023

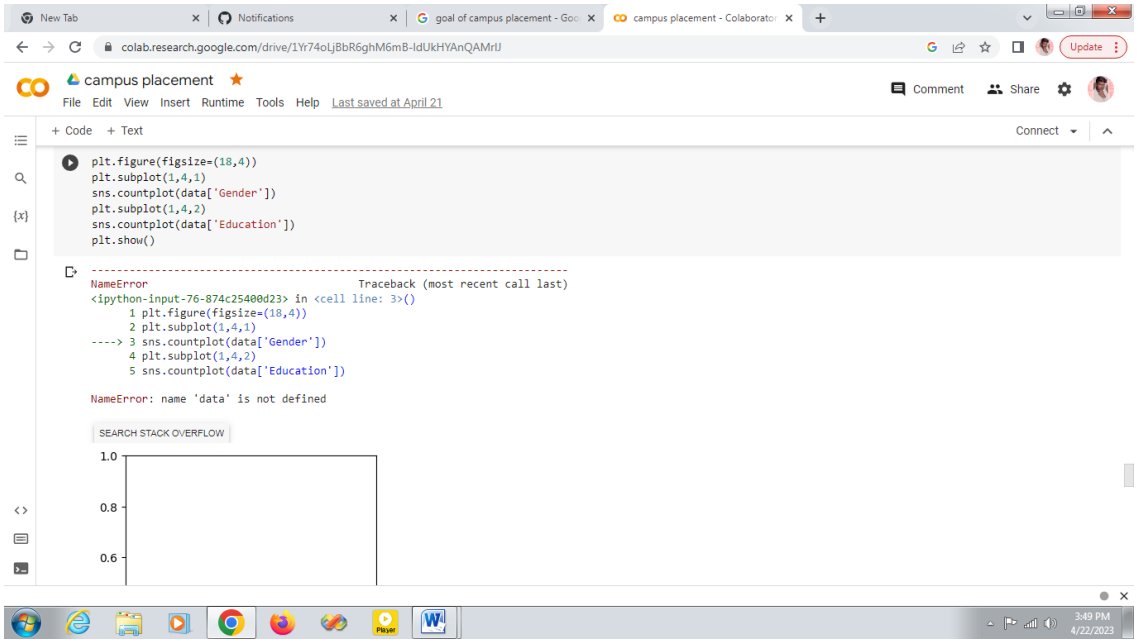

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
+ Code + Text
RAM
Disk
[ ] X=df.drop(columns='PlacedOrNot',axis=1)
Y=df['PlacedOrNot']
[ ] import joblib
joblib.dump(X,"placement")
['placement']
print(X)
Age Gender Stream Internships CGPA HistoryOfBacklogs
0 22 0 2 1 8 1
1 21 1 0 0 7 1
2 22 1 1 1 6 0
3 21 0 1 0 8 1
4 22 0 3 0 8 0
... ...
2961 23 0 1 0 7 0
2962 23 0 3 1 7 0
2963 22 0 1 1 7 0
2964 22 0 0 1 7 0
2965 23 0 5 0 8 0
[2966 rows x 6 columns]
[ ] print(Y)
sample_data
collegePlace.csv
Disk 84.44 GB available
9:07 PM
4/21/2023

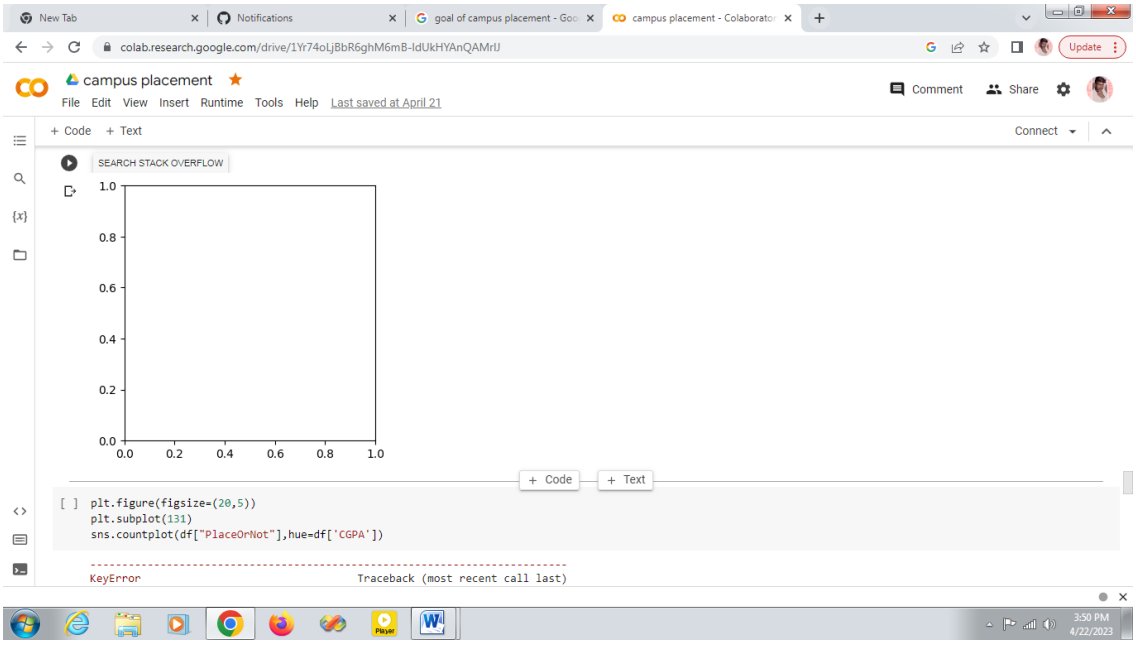

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
sample_data
collegePlace.csv
+ Code + Text
RAM
Disk
[ ] print(Y)
0 1
1 1
2 1
3 1
4 1
..
2961 0
2962 0
2963 0
2964 0
2965 1
Name: PlacedOrNot, Length: 2966, dtype: int64
[ ] scaler=StandardScaler()
[ ] scaler.fit(X)
StandardScaler
StandardScaler()
[ ] standardized_data=scaler.transform(X)
Disk 84.44 GB available
9:08 PM
4/21/2023

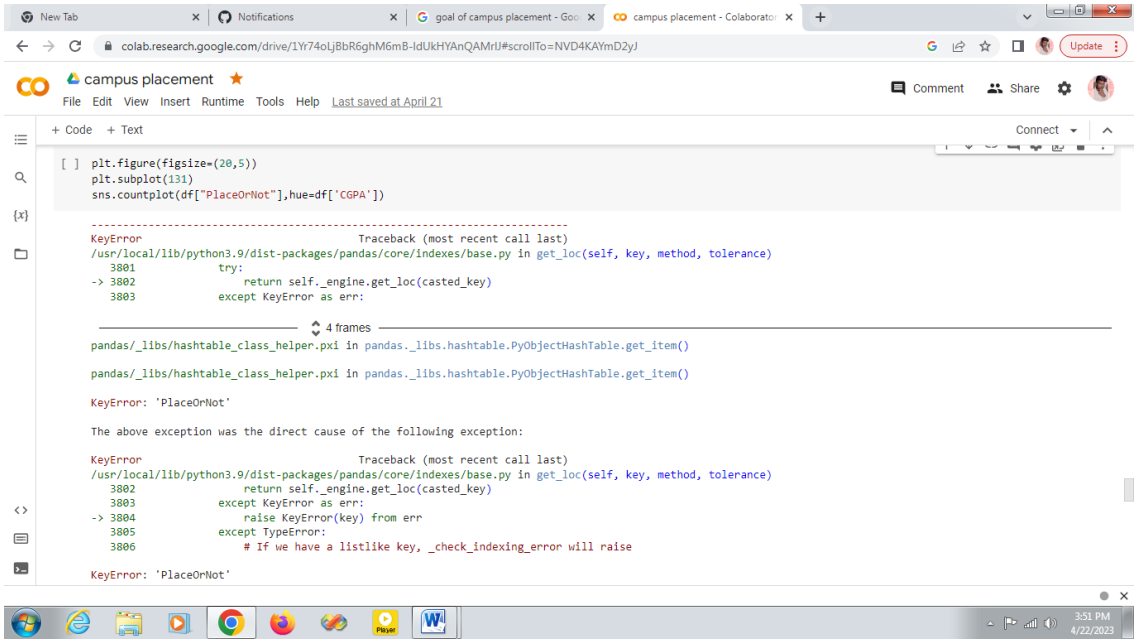

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrlJ#scrollTo=F7E8maNC7cMd&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Last saved at 5:31 AM
Comment
Share
Files
sample_data
collegePlace.csv
+ Code + Text
RAM
Disk
[ ] print(standardized_data)
[[0.38813058 -0.44540301 0.04008175 0.40044544 0.95719068 2.05024603]
[-0.36675158 2.24515772 -1.14874288 -0.95077319 -0.07631043 2.05024603]
[0.38813058 2.24515772 -0.55433057 0.40044544 -1.10981154 -0.48774634]
...
[0.38813058 -0.44540301 -0.55433057 0.40044544 -0.07631043 -0.48774634]
[0.38813058 -0.44540301 -1.14874288 0.40044544 -0.07631043 -0.48774634]
[1.14301273 -0.44540301 1.82331869 -0.95077319 0.95719068 -0.48774634]]
[ ] X=standardized_data
Y=df['PlacedOrNot']
X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
[ ] print(X.shape,X_train.shape,X_test.shape)
(2966, 6) (2372, 6) (594, 6)
[ ] plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
plt.subplot(1,4,2)
sns.countplot(data['Education'])
Disk 84.44 GB available
9:12 PM
4/21/2023

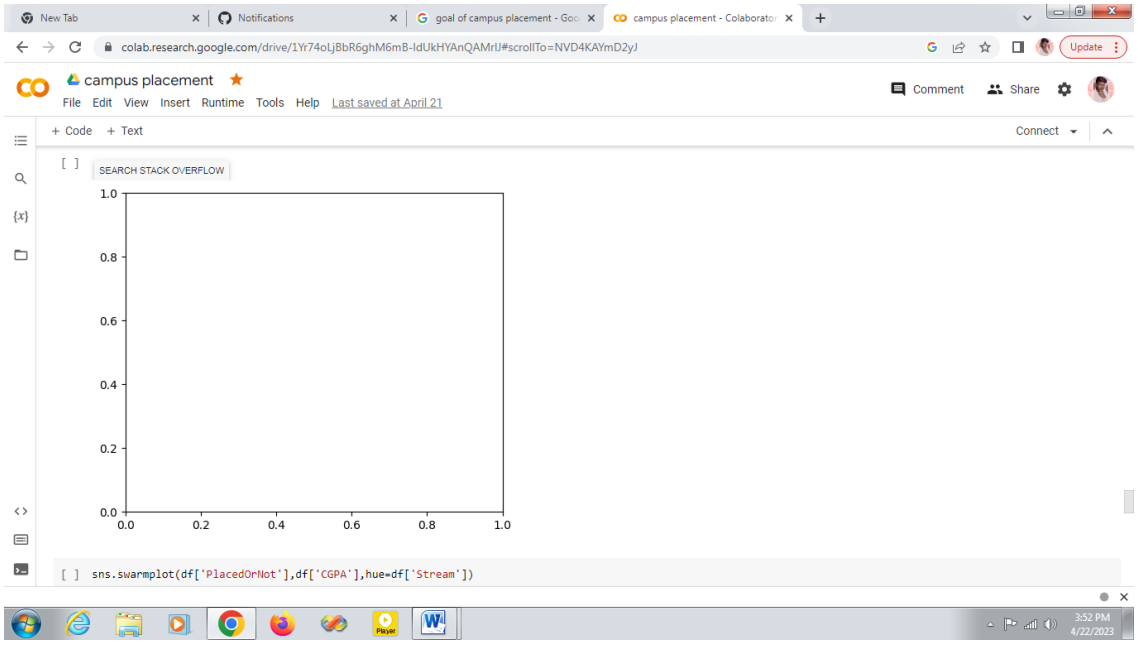

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrJ#scrollTo=i3EbKyK7Gb8i&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help All changes saved
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
placement
[34] X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
print(X.shape,X_train.shape,X_test.shape)
(2966, 6) (2372, 6) (594, 6)
classifier=svm.SVC(kernel='linear')
[36] classifier.fit(X_train,Y_train)
SVC
SVC(kernel='linear')
[37] X_test_prediction=classifier.predict(X_test)
y_pred=accuracy_score(X_test_prediction,Y_test)
y_pred
0.7794612794612794
[ ] plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
Disk
84.44 GB available
0s completed at 9:24 PM
9:24 PM
4/21/2023

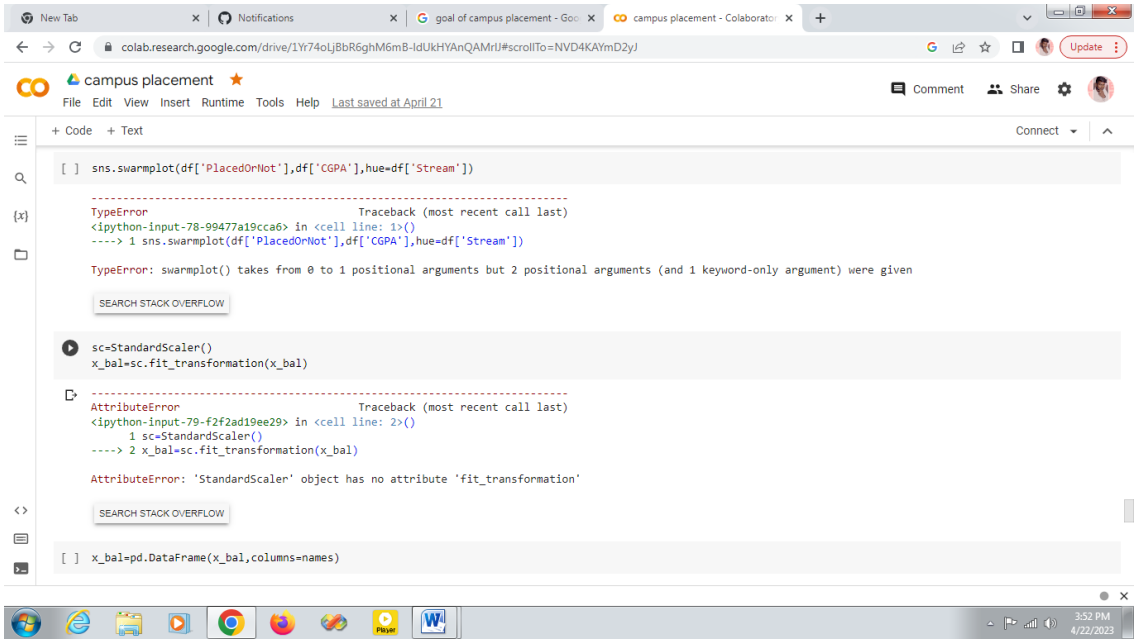

campus placement - Colaborator
colab.research.google.com/drive/1Yr74oLjBbR6ghM6mB-ldUkHYAnQAMrJ#scrollTo=8XwCmeMu15A1&uniqifier=2
Update
campus placement
File Edit View Insert Runtime Tools Help Save failed
Comment
Share
Files
+ Code + Text
RAM
Disk
sample_data
collegePlace.csv
placement
placement.pkl
[67] prediction=model.predict(x_test)
prediction=prediction[0]
return render_template("secondpage.html",y=prediction)
if __name__=='__main__':
app.run(debug=True)
File "<ipython-input-75-994bc27bde2f>", line 1
if__name__=='__main__':
^
SyntaxError: invalid syntax
SEARCH STACK OVERFLOW
plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
plt.subplot(1,4,2)
sns.countplot(data['Education'])
plt.show()
NameError Traceback (most recent call last)
<ipython-input-38-874c25400d23> in <cell line: 3>()
1 plt.figure(figsize=(18,4))
2 plt.subplot(1,4,1)
Disk
84.44 GB available
0s completed at 11:10 PM
11:10 PM
4/21/2023

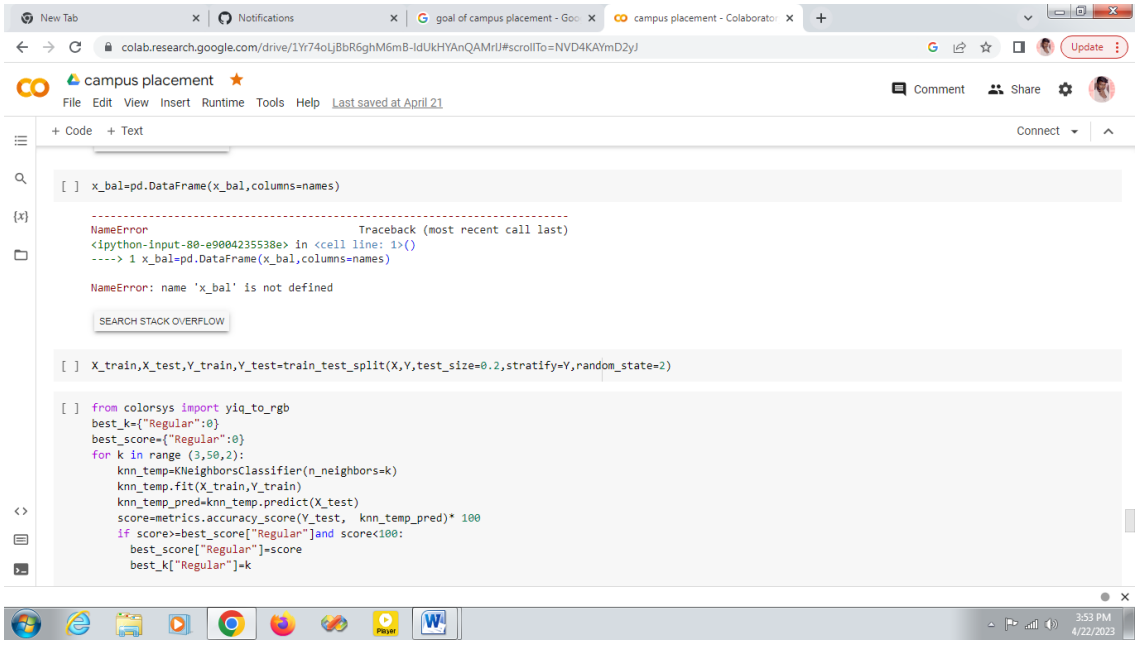

campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment
Share
Connect
+ Code
+ Text
plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(data['Gender'])
plt.subplot(1,4,2)
sns.countplot(data['Education'])
plt.show()
NameError Traceback (most recent call last)
<ipython-input-76-874c25400d23> in <cell line: 3>()
1 plt.figure(figsize=(18,4))
2 plt.subplot(1,4,1)
----> 3 sns.countplot(data['Gender'])
4 plt.subplot(1,4,2)
5 sns.countplot(data['Education'])
NameError: name 'data' is not defined
SEARCH STACK OVERFLOW
1.0
0.8
0.6

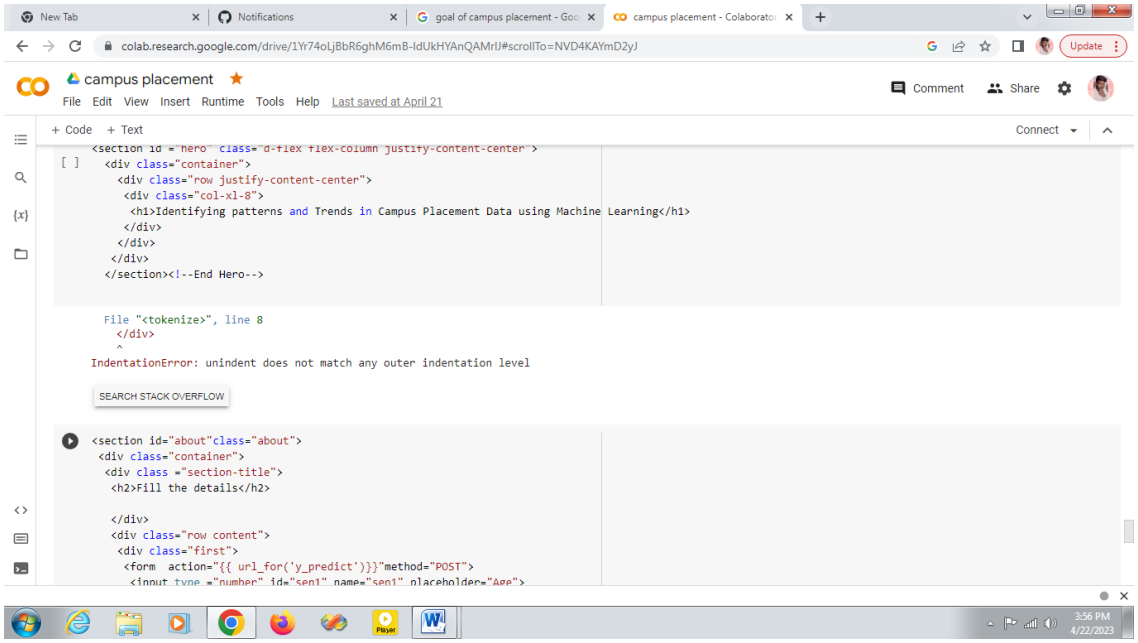

campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment
Share
Connect
+ Code
+ Text
SEARCH STACK OVERFLOW
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0
+ Code
+ Text
[ ] plt.figure(figsize=(20,5))
plt.subplot(131)
sns.countplot(df["PlaceOrNot"],hue=df['CGPA'])
KeyError Traceback (most recent call last)


campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment Share
+ Code + Text
Connect
[ ] plt.figure(figsize=(20,5))
plt.subplot(131)
sns.countplot(df["PlaceOrNot"],hue=df['CGPA'])
KeyError Traceback (most recent call last)
/usr/local/lib/python3.9/dist-packages/pandas/core/indexes/base.py in get_loc(self, key, method, tolerance)
3801 try:
-> 3802 return self._engine.get_loc(casted_key)
3803 except KeyError as err:
4 frames
pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()
pandas/_libs/hashtable_class_helper.pxi in pandas._libs.hashtable.PyObjectHashTable.get_item()
KeyError: 'PlaceOrNot'
The above exception was the direct cause of the following exception:
KeyError Traceback (most recent call last)
/usr/local/lib/python3.9/dist-packages/pandas/core/indexes/base.py in get_loc(self, key, method, tolerance)
3802 return self._engine.get_loc(casted_key)
3803 except KeyError as err:
-> 3804 raise KeyError(key) from err
3805 except TypeError:
3806 # If we have a listlike key, _check_indexing_error will raise
KeyError: 'PlaceOrNot'


campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment Share
+ Code + Text
Connect
[ ]
SEARCH STACK OVERFLOW
1.0
0.8
0.6
0.4
0.2
0.0
0.0 0.2 0.4 0.6 0.8 1.0
[ ] sns.swarmplot(df['PlacedOrNot'],df['CGPA'],hue=df['Stream'])


campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment
Share
Connect
+ Code + Text
[ ] sns.swarmplot(df['PlacedOrNot'],df['CGPA'],hue=df['Stream'])
TypeError Traceback (most recent call last)
<ipython-input-78-99477a19cca6> in <cell line: 1>()
----> 1 sns.swarmplot(df['PlacedOrNot'],df['CGPA'],hue=df['Stream'])
TypeError: swarmplot() takes from 0 to 1 positional arguments but 2 positional arguments (and 1 keyword-only argument) were given
SEARCH STACK OVERFLOW
sc=StandardScaler()
x_bal=sc.fit_transformation(x_bal)
AttributeError Traceback (most recent call last)
<ipython-input-79-f2f2ad19ee29> in <cell line: 2>()
1 sc=StandardScaler()
----> 2 x_bal=sc.fit_transformation(x_bal)
AttributeError: 'StandardScaler' object has no attribute 'fit_transformation'
SEARCH STACK OVERFLOW
[ ] x_bal=pd.DataFrame(x_bal,columns=names)


campus placement
File Edit View Insert Runtime Tools Help Last saved at April 21
Comment
Share
Connect
+ Code + Text
[ ] x_bal=pd.DataFrame(x_bal,columns=names)
NameError Traceback (most recent call last)
<ipython-input-80-e9004235538e> in <cell line: 1>()
----> 1 x_bal=pd.DataFrame(x_bal,columns=names)
NameError: name 'x_bal' is not defined
SEARCH STACK OVERFLOW
[ ] X_train,X_test,Y_train,Y_test=train_test_split(X,Y,test_size=0.2,stratify=Y,random_state=2)
[ ] from colorsys import yiq_to_rgb
best_k={"Regular":0}
best_score={"Regular":0}
for k in range (3,50,2):
knn_temp=KNeighborsClassifier(n_neighbors=k)
knn_temp.fit(X_train,Y_train)
knn_temp_pred=knn_temp.predict(X_test)
score=metrics.accuracy_score(Y_test, knn_temp_pred)* 100
if score>=best_score["Regular"]and score<100:
best_score["Regular"]=score
best_k["Regular"]=k

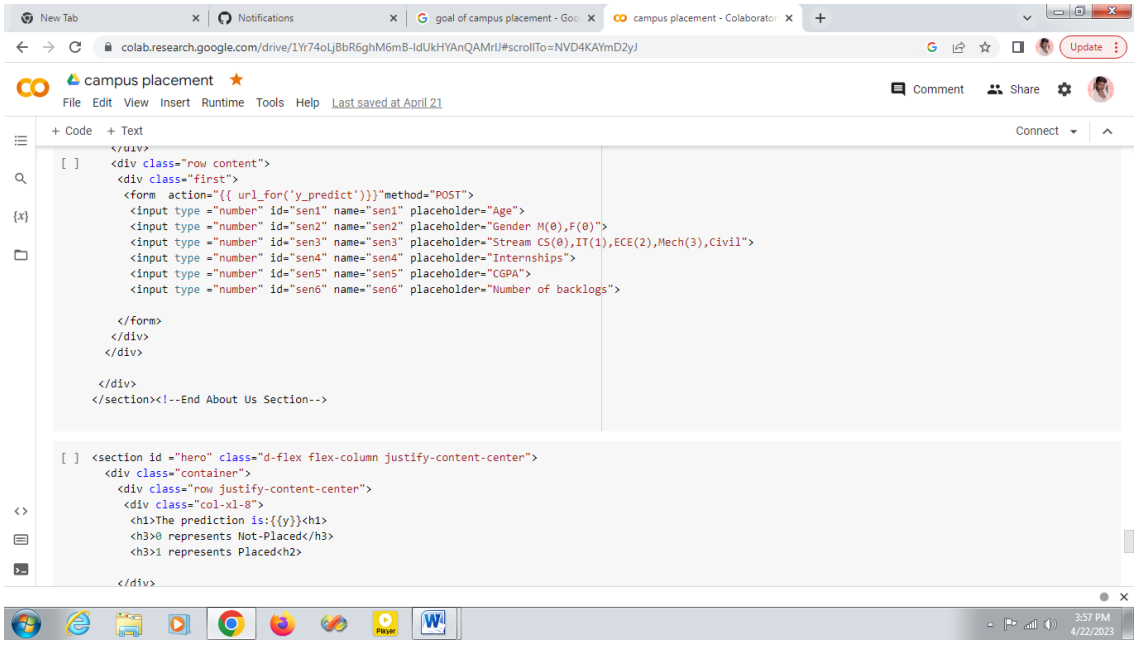


```
<section id="hero" class="d-flex flex-column justify-content-center">
  <div class="container">
    <div class="row justify-content-center">
     <div class="col-xl-8">
      <h1>Identifying patterns and Trends in Campus Placement Data using Machine Learning</h1>
     </div>
    </div>
   </div>
  </section><!--End Hero-->
```

```
  File "<tokenize>", line 8
    </div>
    ^
IndentationError: unindent does not match any outer indentation level
```

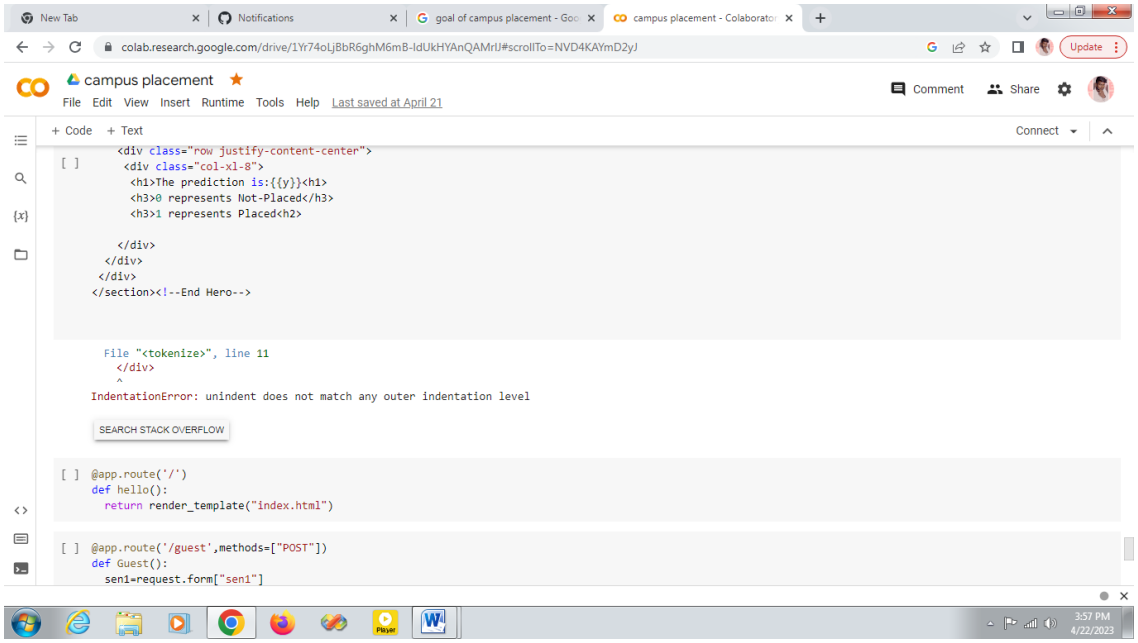
SEARCH STACK OVERFLOW

```
<section id="about"class="about">
 <div class="container">
  <div class ="section-title">
   <h2>Fill the details</h2>

  </div>
  <div class="row content">
   <div class="first">
    <form  action="{{ url_for('y_predict')}}"method="POST">
     <input type ="number" id="sen1" name="sen1" placeholder="Age">
```



```
    </div>
   <div class="row content">
    <div class="first">
     <form  action="{{ url_for('y_predict')}}"method="POST">
      <input type ="number" id="sen1" name="sen1" placeholder="Age">
      <input type ="number" id="sen2" name="sen2" placeholder="Gender M(0),F(0)">
      <input type ="number" id="sen3" name="sen3" placeholder="Stream CS(0),IT(1),ECE(2),Mech(3),Civil">
      <input type ="number" id="sen4" name="sen4" placeholder="Internships">
      <input type ="number" id="sen5" name="sen5" placeholder="CGPA">
      <input type ="number" id="sen6" name="sen6" placeholder="Number of backlogs">

     </form>
    </div>
   </div>

  </div>
 </section><!--End About Us Section-->
```

```
<section id ="hero" class="d-flex flex-column justify-content-center">
  <div class="container">
    <div class="row justify-content-center">
     <div class="col-xl-8">
      <h1>The prediction is:{{y}}<h1>
      <h3>0 represents Not-Placed</h3>
      <h3>1 represents Placed<h2>

     </div>
```



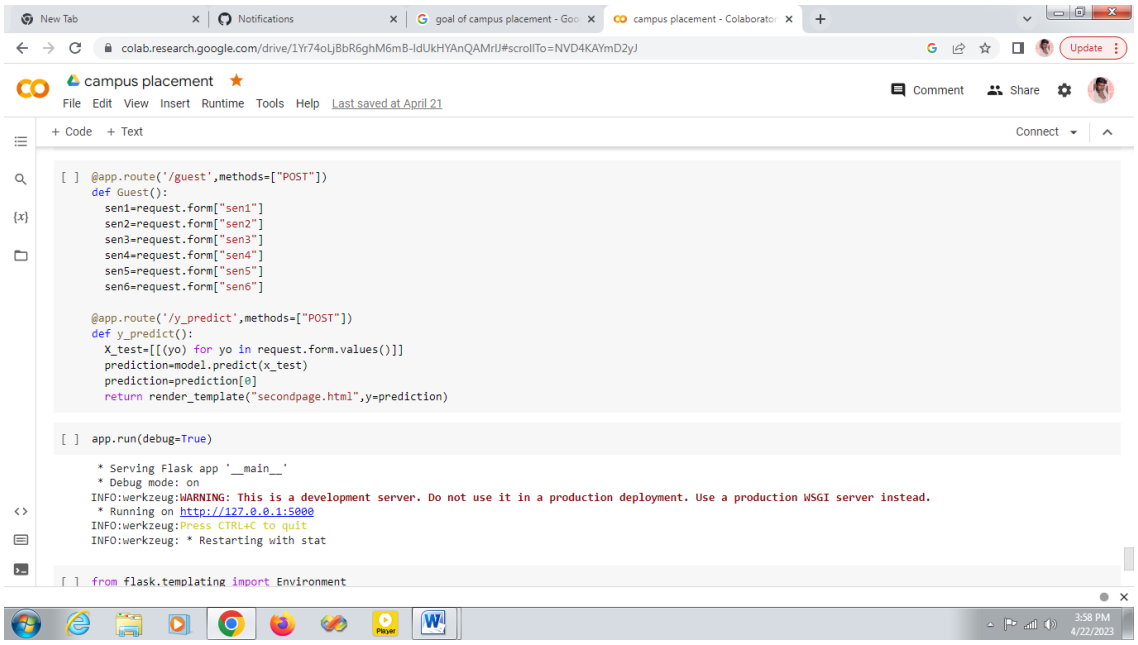
```
      <div class="row justify-content-center">
       <div class="col-xl-8">
        <h1>The prediction is:{{y}}<h1>
        <h3>0 represents Not-Placed</h3>
        <h3>1 represents Placed<h2>

       </div>
      </div>
     </div>
    </section><!--End Hero-->
```

```
  File "<tokenize>", line 11
    </div>
    ^
IndentationError: unindent does not match any outer indentation level
```

SEARCH STACK OVERFLOW

```
@app.route('/')
def hello():
  return render_template("index.html")
```

```
@app.route('/guest',methods=["POST"])
def Guest():
  sen1=request.form["sen1"]
```



```
@app.route('/guest',methods=["POST"])
def Guest():
  sen1=request.form["sen1"]
  sen2=request.form["sen2"]
  sen3=request.form["sen3"]
  sen4=request.form["sen4"]
  sen5=request.form["sen5"]
  sen6=request.form["sen6"]

@app.route('/y_predict',methods=["POST"])
def y_predict():
  X_test=[[(yo) for yo in request.form.values()]]
  prediction=model.predict(x_test)
  prediction=prediction[0]
  return render_template("secondpage.html",y=prediction)
```

```
app.run(debug=True)
```

```
 * Serving Flask app '__main__'
 * Debug mode: on
INFO:werkzeug:WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
INFO:werkzeug:Press CTRL+C to quit
INFO:werkzeug: * Restarting with stat
```

```
from flask.templating import Environment
```

Project Report Template