

```
In [ ]: from google.colab import drive  
drive.mount('/content/drive')
```

Mounted at /content/drive

```
In [ ]: import pandas as pd  
import numpy as np
```

```
In [ ]: # Importing the Data  
file_path = ('/content/drive/MyDrive/Data/Copy of Copy of  
DePaul_Data.xlsx')  
df = pd.read_excel(file_path)  
  
print(df.shape)  
df.head()
```

(7543, 80)

	Reference_ID	Given_Name	Last_Name	College	Major	Degree
0	45405320	Deeksha Reddy	Bhumireddy	INDIA - Osmania University - Bachelor's Degree	NaN	NaN
1	858032003	Pearl Ashok Kumar	Patel	INDIA - Manipal University Jaipur - Bachelor's...	NaN	NaN
2	902518555	Hamza	Javed	PAKISTAN - Shaheed Zulfikar Ali Bhutto Institu...	NaN	NaN
3	Hamza	NaN	Javed	PAKISTAN - Shaheed Zulfikar Ali Bhutto Institu...	NaN	NaN
4	Ronil Dhavalbhai	NaN	Thakkar	INDIA - Swarnim Institute of Technology - Swa...	NaN	NaN

5 rows × 80 columns

```
In [ ]: num_rows = df.shape[0]  
print(num_rows)
```

7543

```
In [ ]: missing_summary = pd.DataFrame({
    "Missing Count": df.isnull().sum(),
    "Missing %": (df.isnull().sum() / len(df)) * 100
})

missing_summary.sort_values("Missing %", ascending=False)
```

	Missing Count	Missing %
Degree_Type	7543	100.0
Major	7543	100.0
Citizenship	7543	100.0
Counsler	7543	100.0
Major_1st_Choice	7543	100.0
...
Recieved_At	0	0.0
University	0	0.0
Modified_At	0	0.0
Application_Source	0	0.0
Created_At	0	0.0

80 rows × 2 columns

```
In [ ]: duplicate_rows = df.duplicated().sum()
duplicate_rows
```

np.int64(0)

```
In [24]: numeric_cols = df.select_dtypes(include=["int64","float64"])

outlier_summary = {}

for col in numeric_cols.columns:
    Q1 = numeric_cols[col].quantile(0.25)
    Q3 = numeric_cols[col].quantile(0.75)
    IQR = Q3 - Q1

    lower = Q1 - 1.5 * IQR
    upper = Q3 + 1.5 * IQR

    outliers = df[(df[col] < lower) | (df[col] > upper)]
    outlier_summary[col] = len(outliers)

pd.DataFrame.from_dict(outlier_summary, orient="index", columns=["Outlier Count"])
```

Out[24]:

	Outlier Count
Major	0
Degree_Type	0
Recieved_At	1015
Counsler	0
Citizenship	0
Status	0
Major_1st_Choice	0
Street_3	0
Date_of_Birth	0
Phone_Number_2	0
Most_Recent_Released_Decision	0
RIT_Email_Created	0
Housing_Contract	0
Is_Global_Grad	0
Is_Admitted	0
SEVIS_ID	0
Official_University_email_address	0
Application_Agency_Code	0
Created_At	1015
Modified_At	1015
Reference_ID-2	0
Outstanding_Checklist_Items	0
Prior_I-20_Outreach_Detail	0
Prior_Non_I-20_Outreach_Detail	0
Most_Recent_Contact	0
Most_Recent_User_and_Date	0
Recent_GT_Form_Initiative	0

	Outlier Count
Caller_Name_2	0
Date_of_Contact_2	0
Outcome_2	0
Final_Result	0
Campaign_Id	122
Escalation_Required	0
SLU_Start_Comment	0
SLU_Start_City	0
I_901_Status	0
Created_At-2	0
Modified_At-2	0
ID	122
Category	0
Attempts	0
Intake-2	0
Status-2	0
City_and_Branch	0
Header	0
Template	0
Region-2	0
Created_At-3	0
Modified_At-3	0

```
In [ ]: # Identify and display value counts for text columns
text_cols = df.select_dtypes(include="object")

for col in text_cols.columns:
    print("\nColumn: ", col)
    print(df[col].value_counts(dropna=False).head(10))
```

Column: Reference_ID

Reference_ID	count
924093519	48
48921321	33
20367881	31
553724526	30
516679919	30
11979156	27
838525490	26
769886402	25
467992191	23
265477508	23

Name: count, dtype: int64

Column: Given_Name

Given_Name	count
Fnu	101
Raj	53
Rahul	45
Syed Ali Mujtaba	33
Salman	31
Sejal	30
Preksha Gowda	30
Umar Javeed	27
Ishanshi	26
Dhruva Chaitanya	25

Name: count, dtype: int64

Column: Last_Name

Last_Name	count
Patel	296
Mohammed	145
Shah	67
Gupta	55
Khan	48
Kundur	48
Lnu	47
Parmar	44
Syed	41
Singh	37

Name: count, dtype: int64

Column: College

College	count
INDIA - Osmania University - Bachelor's Degree	269
INDIA - Gujarat Technological University - Bachelor's Degree	74
INDIA - University of Mumbai - Bachelor's Degree	71
NAN	71
INDIA - Visvesvaraya Technological University - Bachelor's Degree	62
INDIA - Anna University - Bachelor's Degree	60
INDIA - Jawaharlal Nehru Technological University Hyderabad - Bachelor's Degree	56
INDIA - Maharashtra State Board of Technical Education - Other\nINDIA - K J College Of Engineering Management and Research - Bachelor's	

```
Degree      48
INDIA - Savitribai Phule Pune University - Bachelor's Degree
47
INDIA - Jawaharlal Nehru Technological University, Hyderabad -
Bachelor's Degree
40
Name: count, dtype: int64
```

```
Column: Country
Country
India      4617
United states    764
Ghana      440
Nigeria     430
Pakistan     273
China       125
Nepal        81
Iran         75
Canada        69
Kenya        62
Name: count, dtype: int64
```

```
Column: University
University
DePaul University    7543
Name: count, dtype: int64
```

```
Column: Intake
Intake
Computer Science - MS: 2024 Fall (September '24)      504
Business Analytics - MS: 2024 Fall (September '24)    419
Data Science - MS: 2024 Fall (September '24)          254
Cybersecurity - MS: 2024 Fall (September '24)         181
Business Analytics - MS: 2025 Fall (September '25)    160
Artificial Intelligence - MS: 2024 Fall (September '24) 160
Health Informatics - MS: 2024 Fall (September '24)    155
Business Analytics - MS: 2025 Winter (January '25)     144
Business Analytics - MS: 2024 Winter (January '24)     143
Human-Computer Interaction - MS: 2024 Fall (September '24) 140
Name: count, dtype: int64
```

```
Column: College_1st_Choice
College_1st_Choice
Jarvis College of Computing and Digital Media      3831
Kellstadt Graduate School Of Business            2523
College of Science and Health                  269
College of Education                          208
College of Liberal Arts and Social Sciences    207
College of Communication                      87
School of Psychology                         79
School of Public Service                     77
Jarvis College of Computing and Digital Media (Fully Online) 72
College of Education (Fully Online)           35
Name: count, dtype: int64
```

```
Column: Phone_Number
Phone_Number
NaN          587
920000000000  43
919000000000  39
+91 84290 90019  33
+91 96764 94150  31
+91 80088 47896  30
+91 97899 88124  30
+91 91217 94780  27
+1 872-338-6722  26
+91 7019941170  25
Name: count, dtype: int64
```

Column: Street_1

Street_1

Omkarpark Phase-2,Rajmudra Society, Dhankawadi	48
45F Mohalla Kundigarrh Post Sadar	33
H.No:-9-4-136/146, Jamalikunta, Tolichowki	31
Annapurneshwari Krupa, Near Royal Apollo	30
6-1-306/B56 Old Cib Quarters Khairatabad	30
10-5-4/A Masab Tank	27
13 Shreenath Soc, Usmanpura Ashram Road, Ahmedabad	26
No 57,14th Cross, Mts Layout, Kengeri	25
74 Prakruti Banglows Sterling	23
60 Sainik Vihar, Pitampura	23

Name: count, dtype: int64

Column: Street_2

Street_2

NaN	4639
School Hemavathi Nagara	30
Secunderabad	30
Musheerabad	26
Satellite Town	25
City Sector F Bhopal Ahemdabad	23
Silent Zone, Opp. Airport Road, Dumas	22
Sri Ram Park Colony, Road No.8,	21
Flat 102, Mumtaz College Road,	20
Lbs Marg	20

Name: count, dtype: int64

Column: City

City

Hyderabad	975
Chicago	268
Ahmedabad	223
Chennai	206
Mumbai	204
Pune	164
Accra	158
Bengaluru	134
Karachi	108
Surat	96

Name: count, dtype: int64

Column: Region

Region

Telangana	1239
NaN	914
Maharashtra	645
IL	524
Gujarat	486
Andhra Pradesh	406
Karnataka	338
Tamil Nadu	288
Greater Accra	208
Lagos	149

Name: count, dtype: int64

Column: Postal

Postal

NaN	226
233	198
500008	112
0	81
500028	65
500004	58
411043	48
500036	45
500034	38
110034	34

Name: count, dtype: int64

Column: Email_ID

Email_ID	
rajkundur.rk@gmail.com	48
syedalig1020@gmail.com	33
salmanawaise707@gmail.com	31
sejalmudiraj12@gmail.com	30
sppreksha.india@gmail.com	30
umar.amf555@gmail.com	27
ishanshipatel12@gmail.com	26
dhruvachaitanyaagopal25082002@gmail.com	25
ayush593sach@gmail.com	23
krishvekariya116@gmail.com	23

Name: count, dtype: int64

Column: Comments

Comments	
Evaluation Completed	
1642	
NaN	
1625	
No Record	
159	
Email sent to study group team	
145	
Transcript not on slate	
139	
Student will upload official semester wise marksheets by today via mail.\n February 27, 2024Call not answered\n February 28, 2024I have applied for the official transcript and will receive it soon. And I will upload it immediately.\n February 27, 2024Dear Raj,\n\nWe hope this email finds you well.\n\nYour application to DePaul University has been received. Before we can start reviewing your transcripts you will need to provide your complete semesterwise / yearly undergraduate transcripts.\n\nWe kindly request that you send us your complete semesterwise / yearly undergraduate transcripts at your earliest convenience. After they are received, our team of professionals will begin the transcript evaluation process. You can either reply to this email with the transcripts attached or upload them on your Grad Gateway portal by clicking the belowmentioned link.\n\nhttps://grad.depaul.edu/portal/status \n\nWe appreciate your cooperation and look forward to receiving your transcripts soon.\n\nBest Regards, \nJagdeep Singh\nDePaul Transcript Evaluation Team\n February 27, 2024The attached file is my unofficial transcript including all semester wise details year wise.I have already applied for official transcript as soon as i receive my official transcript, I will upload that immediately.	
48	
Email sent to study group team.	
45	
Evaluation completed	
38	
Student will submit the Prefinal semester marksheets only on May month\n April 29, 2024Dear Syed,\n\nWe hope this email finds you well.\n\nYour application to DePaul University has been received. Before we can start reviewing your transcripts you will need to provide your Prefinals university authenticated undergraduate transcripts.\n\nWe kindly request that you send us your prefinals university authenticated undergraduate transcripts at your earliest convenience. After they are received, our team of professionals will begin the transcript evaluation process. You can either reply to this email with the transcripts attached or upload them on your Grad Gateway portal by clicking the belowmentioned link.\n\nhttps://grad.depaul.edu/portal/status \n\nWe appreciate your cooperation and look forward to receiving your transcripts soon.\n\nBest Regards, \nJagdeep Singh\nDePaul Transcript Evaluation Team\n March 17, 2024Need Prefinal semester marksheets	
33	
student's 7th semester exam will held in March' 24\n February 26,	

2024Dear Jagdeep,\nGreetings for the day!\n\nThank you for replying to me about the transcript evaluation, Presently I have marksheets of my Bachelor's degree till Semester 6th & for the 7th Semester mark sheet it will take some time till April month.\nAs soon as I receive my 7thsemester mark sheet, I will send the documents for the evaluation process.\n\nHope you will understand as I am still pursuing my Final Year Bachelor.\nThanks & Regards\nSalman Awaise.\nFebruary 20, 2024Dear Salman,\n\nWe hope this email finds you well.\n\nYour application to DePaul University has been received. Before we can start reviewing your transcripts you will need to provide your official 6th & 7th semester undergraduate transcripts.\n\nWe kindly request that you send us your official 6th & 7th semester undergraduate transcripts at your earliest convenience. After they are received, our team of professionals will begin the transcript evaluation process. You can either reply to this email with the transcripts attached or upload them on your Grad Gateway portal by clicking the belowmentioned link.\n\nhttps://grad.depaul.edu/portal/status\n\nWe appreciate your cooperation and look forward to receiving your transcripts soon.\n\nBest Regards, \nJagdeep Singh\nDePaul Transcript Evaluation Team\nJanuary 18, 2024Course inprogress. Need final year marksheets\nAugust 12, 2024Please find the attached copy for my 7th sem memo.\nAugust 21, 2024Evaluation Completed 31

Name: count, dtype: int64

Column: Admit_Date

Admit_Date

NaN	1754
2024-12-03 00:00:00	80
2024-12-06 00:00:00	80
03/18/2024	79
2024-05-03 00:00:00	67
2024-03-04 00:00:00	66
02/27/2024	63
2024-06-05 00:00:00	61
2024-12-01 00:00:00	59
2024-10-04 00:00:00	59

Name: count, dtype: int64

Column: Application_Source

Application_Source

Non Study Group	6528
-----------------	------

STUDY GROUP	1015
-------------	------

Name: count, dtype: int64

Column: Recieved_At-2

Recieved_At-2

NaN	3194
2024-01-04 00:00:00	47
2024-03-05 00:00:00	45
2024-04-01 00:00:00	44
2024-02-07 00:00:00	41
2024-06-03 00:00:00	41
2024-03-06 00:00:00	40
2024-06-05 00:00:00	40
2024-05-03 00:00:00	40
03/18/2024 00:00:00	39

Name: count, dtype: int64

Column: University-2

University-2

DePaul University	4349
-------------------	------

NaN	3194
-----	------

Name: count, dtype: int64

Column: Date_of_Contact

Date_of_Contact

NaN	3194
-----	------

2024-01-04 00:00:00	47
---------------------	----

```
2024-03-05 00:00:00      45
2024-04-01 00:00:00      44
2024-02-07 00:00:00      41
2024-06-03 00:00:00      41
2024-03-06 00:00:00      40
2024-06-05 00:00:00      40
2024-05-03 00:00:00      40
03/18/2024 00:00:00      39
Name: count, dtype: int64

Column: Caller_Name
Caller_Name
NaN      3194
Jagdeep   2085
Swarna    1214
Ayesha    855
Lavnaya   195
Name: count, dtype: int64

Column: Outcome_1
Outcome_1
NaN          3194
Follow up     1375
Missing 0 Transcript (NOT Study Group) 1177
Issue ID's     450
Study Group ID 397
Transcript received on mail 318
NOT Study Group 277
Transcript received 79
Missing >1 Transcript (NOT Study Group) 72
Replied to student 53
Name: count, dtype: int64

Column: Remark
Remark
NaN
6869
Call not answered
544
Email sent
103
call not answered
9
Will submit the transcripts soon
3
study group
2
student not able to login in Depaul application portal and will share
marksheets on mail      2
will share marksheets on mail
2
Student is from US and have US transcripts
2
Transcript received on mail and uploaded
1
Name: count, dtype: int64

Column: Remark_2
Remark_2
NaN
7459
Discussed about the transcript
6
Discussed about the transcripts
5
Discussed about the transcripts.
3
Student just finished his final sem exams, he is still not sure about
the results.           2
```

```
Student will submit Provisional or graduation certificate today via  
mail. 2  
Student will submit the semester 8 marksheets & graduation certificate  
in August month 2  
Student will submit the final semester marksheets in august.  
2  
Student will submit Complete Bachelor marksheets in a month  
2  
Student will submit 7 sem marksheets today and 8 sem marksheets after 2  
months. 2  
Name: count, dtype: int64
```

```
Column: Slate_Form_Filled  
Slate_Form_Filled  
Yes 4336  
NaN 3194  
No 13  
Name: count, dtype: int64
```

```
Column: Name  
Name  
NaN 3316  
Email To Student 1007  
Call To Student 960  
Sent For Evaluation 960  
Complete Case 549  
Email From Student 413  
Email Sent To Study Group Team 211  
Call From Student 83  
Whatsapp From Student 24  
Whatsapp To Student 20  
Name: count, dtype: int64
```

```
Column: University-3  
University-3  
DePaul University 4349  
NaN 3194  
Name: count, dtype: int64
```

```
Column: Degree_Type-2  
Degree_Type-2  
grad 4349  
NaN 3194  
Name: count, dtype: int64
```

```
Column: Type  
Type  
Calling 4349  
NaN 3194  
Name: count, dtype: int64
```

```
In [ ]: # Quick Summary
quality_report = {
    "Rows": df.shape[0],
    "Columns": df.shape[1],
    "Duplicate Rows": df.duplicated().sum(),
    "Total Missing Values": df.isnull().sum().sum(),
    "Total Numeric Outliers": sum(outlier_summary.values())
}

quality_report
```

```
{'Rows': 7543,
 'Columns': 80,
 'Duplicate Rows': np.int64(0),
 'Total Missing Values': np.int64(371664),
 'Total Numeric Outliers': 3289}
```

Exported with [runcell](#) — convert notebooks to HTML or PDF anytime at [runcell.dev](#).