

Data Dictionary - OutReach Dataset

Column Name	Description	Example Value	Data Type
Reference_ID	Unique identifier for each record or call entry	47944	Object
Received_At	Timestamp when the call or record was received	2024-12-16 16:23:47	Datetime
University	Name of the university associated with the record	Illinois Institute of Technology	Object
Caller_Name	Name of the caller or student contacted	Jyoti	Object
Outcome_1	Primary outcome of the call (e.g., Connected, Not Connected)	Not connected	Object
Remark	Remarks or additional notes recorded during the call	No Remarks	Object
Campaign_ID	Identifier for the campaign or project under which the call was made	SP25AI2S	Object
Escalation_Required	Indicates whether the call required escalation (Yes/No)	No	Object

Data Dictionary - Application Dataset

Column Name	Data Type	Description	Notes / Changes Made
App_ID	Text / String	Unique identifier for each applicant	1 missing value removed; all remaining records unique
Country	Text / String	Applicant's country of residence	No missing values
University	Text / String	Name of the university applicant applied to	No missing values
Phone_Number	Text / String	Original phone number as provided by applicant	Mixed format; cleaned in Phone_Number_Clean
Phone_Number_Clean	Text / String	Standardized 10-digit phone number (country code removed)	Non-numeric characters removed; country code stripped; last 10 digits kept

Data Dictionary - Campaign Dataset

Column Name	Data Type	Description	Notes / Changes Made
ID	Text	Unique identifier for each campaign	No issues found
Name	Text	Name of the campaign including session info	No issues found
Category	Text	Indicates the type of campaign (Pre/Post Admission)	Verified values — consistent
LIntake	Text	Academic year or intake term for the campaign	No cleaning required
University	Text	University name associated with the campaign	All entries consistent
Status	Text	Current status of the campaign	All records marked as 'Completed'
Start_Date	Date	Date when the campaign started	Standardized to YYYY-MM-DD format

Here is the link for the Final cleaned dataset

Final joined dataset

[clean_data.xlsx](#)

Outreach data

[Cleaned_Data_outreach.xlsx](#)

Application data

[Clean_application.xlsx](#)

Campaign Dataset

[Clean_campaign.csv](#)