



Team Name

0109DVA - 16

Team Lead	Project Manager	Project Scribe	Project Lead
M.Nouman Arshad	Farhan Hussain	Shivansh Rajput	Farhan / Nouman

NAME	EMAIL	SPONSOR COMPANY	CLIENT
Farhan Hussain	farhanhussain2037@gmail.com	Saint Louis university	Excelerate
M.Nouman Arshad	nomi123nomiest@gmail.com	Saint Louis university	Excelerate
Shivansh Rajput	shivanshrajputop@gmail.com	Saint Louis university	Excelerate
Dia Agarwal	diaag2604@gmail.com	Saint Louis university	Excelerate
Yuva Mirsha	yuvamirshai02@gmail.com	Saint Louis university	Excelerate
Ajay Kumar Kasaudhan	a.k.kasaudhan7600@gmail.com	Saint Louis university	Excelerate

Naheed Akhtar	naheedsamoor123@gmail.com	Saint Louis university	Excelerate
Saurav Kumar	davidsauravyadav@gmail.com	Saint Louis university	Excelerate

Data Validation & Insight Summary Report

Validation Script and Explanation

Adding KPIs (Performance Metrics)

- **Connected_Calls**
Records a 1 if the outcome is a *meaningful conversation* (e.g., Interested, Pending, Visa related). Otherwise 0.
- **Disconnected_Calls**
Records a 1 if the call failed (e.g., No Response, Invalid Number). Otherwise 0.
- **Connectivity_Rate** Formula:
Connectivity Rate=Connected Calls/Connected Calls + Disconnected Calls Rounded to 2 decimals. This shows how successful outreach was in connecting with applicants.
- **Agent_Performance**
 - Good
 - Average

- Poor

```
-- New Metrics
CASE
    WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                          'Visa Related','Application In Progress',
                          'Pending Documents','Scholarship Related',
                          'Loan/Bank Document Query','I-20 Related Issue')
        THEN 1 ELSE 0 END AS "Connected_Calls",

CASE
    WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
        THEN 1 ELSE 0 END AS "Disconnected_Calls",

CASE
    WHEN (CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                      'Visa Related','Application In Progress',
                                      'Pending Documents','Scholarship Related',
                                      'Loan/Bank Document Query','I-20 Related Issue')
              THEN 1 ELSE 0 END
        +
        CASE WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
              THEN 1 ELSE 0 END) > 0
        THEN ROUND(
            (CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                      'Visa Related','Application In Progress',
                                      'Pending Documents','Scholarship Related',
                                      'Loan/Bank Document Query','I-20 Related Issue')
                  THEN 1 ELSE 0 END)::numeric
        /
        ((CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                      'Visa Related','Application In Progress',
                                      'Pending Documents','Scholarship Related',
                                      'Loan/Bank Document Query','I-20 Related Issue')
              THEN 1 ELSE 0 END)
        +
        (CASE WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
              THEN 1 ELSE 0 END)),2)
        ELSE NULL END AS "Connectivity_Rate",

CASE
    WHEN (CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                      'Visa Related','Application In Progress',
                                      'Pending Documents','Scholarship Related',
                                      'Loan/Bank Document Query','I-20 Related Issue')
              THEN 1 ELSE 0 END)
```

```

        'Pending Documents','Scholarship Related',
        'Loan/Bank Document Query','I-20 Related Issue')
    THEN 1 ELSE 0 END
+
CASE WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
    THEN 1 ELSE 0 END) > 0
THEN CASE
    WHEN ROUND((CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                            'Visa Related','Application In Progress',
                                            'Pending Documents','Scholarship Related',
                                            'Loan/Bank Document Query','I-20 Related Issue')
                    THEN 1 ELSE 0 END)::numeric
/
((CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                            'Visa Related','Application In Progress',
                            'Pending Documents','Scholarship Related',
                            'Loan/Bank Document Query','I-20 Related Issue')
                    THEN 1 ELSE 0 END)
+
(CASE WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
    THEN 1 ELSE 0 END)),2) >= 0.7
THEN 'Good'
WHEN ROUND((CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                                            'Visa Related','Application In Progress',
                                            'Pending Documents','Scholarship Related',
                                            'Loan/Bank Document Query','I-20 Related Issue')
                    THEN 1 ELSE 0 END)::numeric
/
((CASE WHEN o."Outcome" IN ('Interested','Not Interested','Pending Decision',
                            'Visa Related','Application In Progress',
                            'Pending Documents','Scholarship Related',
                            'Loan/Bank Document Query','I-20 Related Issue')
                    THEN 1 ELSE 0 END)
+
(CASE WHEN o."Outcome" IN ('No Response','Disconnected','Invalid Number')
    THEN 1 ELSE 0 END)),2) >= 0.4
THEN 'Average'
ELSE 'Poor'
END
ELSE NULL END AS "Agent_Performance"

FROM "applicant_data" a
LEFT JOIN "outreach_data" o

```

```

        THEN 1 ELSE 0 END)),2) >= 0.4
THEN 'Average'
ELSE 'Poor'
END
ELSE NULL END AS "Agent_Performance"

FROM "applicant_data" a
LEFT JOIN "outreach_data" o
ON a."App_ID" = o."Reference_ID":BIGINT
LEFT JOIN "campaign_data" c
ON o."University" = c."University";

```

Before generating insights, the dataset underwent a rigorous validation and cleaning process to ensure accuracy, consistency, and completeness:

- **Duplicate Removal:** Used App_ID as the unique identifier; retained only the latest record per applicant.
- **Invalid Records Elimination:** Removed rows with invalid IDs, missing timestamps, or critical null values.
- **Column Pruning:** Dropped high-null and redundant fields (e.g., Phone Number, Reference ID, Campaign Status).
- **Business Rule Checks:** Standardized formats for App_IDs, campaign categories, and agent performance labels.

Data Cleaning

- **DELETE WHERE LENGTH(App_ID) <> 6;**
Removes applicants with invalid IDs (not 6 digits). Ensures uniform identifiers.
- **ALTER TABLE DROP COLUMN**
Drops unnecessary columns like Phone_Number, Batch_Code, etc., which are not useful for analysis. Keeps the dataset lean.
- **UPDATE master_table SET Agent_Performance** Recalculates performance with a fallback:
 - If no calls exist, mark as "**Not Evaluated**" instead of leaving blank.
- **DELETE WHERE Received_At IS NULL OR Time_Stamp IS NULL;**
Removes incomplete or faulty records that would cause issues in analysis.

These steps ensure **clean, valid, and consistent data**.

----- Deleting those fields where the app id is more than 6 in length as it is inappropriate-----

```
DELETE FROM "master_table"  
WHERE LENGTH("App_ID)::text) <> 6;
```

----- Dropped some inappropriate column which is not in use for any analysis case-----

```
ALTER TABLE "master_table"
DROP COLUMN "Phone_Number",
DROP COLUMN "Reference_ID",
DROP COLUMN "Batch_Code",
DROP COLUMN "Followup",
DROP COLUMN "Outreach_University",
DROP COLUMN "Campaign_ID",
DROP COLUMN "Campaign_Status",
DROP COLUMN "Campaign_University";
```

----- creating some new metrics like connected, disconnected calls, and agent performance as average poor, etc.-----

```
UPDATE "master_table"
SET "Agent_Performance" = CASE
    WHEN ("Connected_Calls" + "Disconnected_Calls") > 0
    THEN CASE
        WHEN ROUND("Connected_Calls"::numeric / NULLIF(("Connected_Calls" + "Disconnected_Calls"),0), 2) >= 0.7
            THEN 'Good'
        WHEN ROUND("Connected_Calls"::numeric / NULLIF(("Connected_Calls" + "Disconnected_Calls"),0), 2) >= 0.4
            THEN 'Average'
        ELSE 'Poor'
    END
    ELSE 'Not Evaluated' -- fallback when no calls
END;
```

----- deleteing field where values are null. this can cause problems in analysis -----

```
DELETE FROM "master_table"
WHERE "Received_At" IS NULL
OR "Time_Stamp" IS NULL;
```

Validation & Quality Checks

- **SELECT COUNT(*)**
Counts how many valid records remain after cleaning.
- **Duplicate Check (App_ID)**
Identifies applicants who appear more than once.
- **Duplicate Removal**
Keeps only the **latest entry per applicant (based on Time_Stamp)** and deletes older ones.
- **Missing Value Audit**
Counts missing values column-wise (App_ID, Country, University, Outcome, Campaign_Name). Helps measure data completeness.
- **Connectivity Validation**
Ensures all rates are between 0 and 1 (no invalid calculations).
- **Performance Labels Check**
Lists distinct Agent_Performance values (should be Good, Average, Poor, Not Evaluated).
- **Stage / Outcome / Campaign Intake Checks**
Lists all unique values in these columns to detect unexpected or invalid entries.

```
----- Performing Validation-----
-- Count total rows in cleaned table
```

```
SELECT COUNT(*) AS total_records
FROM "master_table";
```

```
-- Check duplicate rows based on unique identifier
```

```
SELECT "App_ID", COUNT(*)
FROM "master_table"
GROUP BY "App_ID"
HAVING COUNT(*) > 1;
```

```
-- Remove duplicate rows from master_table_cleaned
```

```
DELETE FROM "master_table"
WHERE ctid NOT IN (
    SELECT ctid
    FROM (
        SELECT ctid,
               ROW_NUMBER() OVER (
                   PARTITION BY "App_ID"
                   ORDER BY "Time_Stamp" DESC      -- keep latest row
               ) AS rn
        FROM "master_table"
    ) t
    WHERE rn = 1
);
```

```
-- Count missing values column-wise
SELECT
    SUM(CASE WHEN "App_ID" IS NULL THEN 1 ELSE 0 END) AS missing_app_id,
    SUM(CASE WHEN "Applicant_Country" IS NULL THEN 1 ELSE 0 END) AS missing_country,
    SUM(CASE WHEN "Applicant_University" IS NULL THEN 1 ELSE 0 END) AS missing_university,
    SUM(CASE WHEN "Outcome" IS NULL THEN 1 ELSE 0 END) AS missing_outcome,
    SUM(CASE WHEN "Campaign_Name" IS NULL THEN 1 ELSE 0 END) AS missing_campaign
FROM "master_table";
```

```
----Checking Connectivity Rate is between 0 and 1
SELECT COUNT(*) AS invalid_connectivity
FROM "master_table"
WHERE "Connectivity_Rate" < 0 OR "Connectivity_Rate" > 1;
```

---- Checking Agen performance label

```
SELECT DISTINCT "Agent_Performance"
FROM "master_table";
```

-- Checking for unexpected Stage values

```
SELECT DISTINCT "Stage"
FROM "master_table";
```

-- Check distinct campaign statuses

```
SELECT DISTINCT "Campaign_Intake"
FROM "master_table";
```

-- List unique call outcomes

```
SELECT DISTINCT "Outcome"
FROM "master_table";
```

Impact of Validation

- Significant reduction in inflated counts:
 - *Application in Progress* dropped from **11,178 - 209**.
 - *Pending Document* dropped from **7,475 - 192**.
- For example, Applicant ID **432571** reduced from **69 records - 1 clean record(removed duplicacy)**, eliminating duplicates that distorted connectivity analysis.
- **No call data was lost:** Disconnected Calls remained at 0 after validation.

New Metrics Introduced

- **Connected Calls & Disconnected Calls** → enabled accurate connectivity rate calculation (92%).
- **Agent Performance Categories** → standardized into Good, Average, Poor, or Not Evaluated.

Assurance of Reliability

- Final dataset is **clean, consistent, and validated** for analysis.
- All insights in this report are derived from a **single-source, trustworthy master table**, ensuring reliability for decision-making.

Major Findings

- High Connectivity Rate: Out of **29,210 calls**, **26,749** were connected while only **2,461** were disconnected (**~92% connectivity rate**).
- Campaign Performance: GR GS Call Campaign generated the highest volume of calls and outcomes, followed by SP25 and FA24 campaigns. Some outcomes were 'Not Provided', indicating gaps in data logging.
- Application Stage Distribution: Completed applications dominate, but significant portions remain Incomplete (**16.7%**), New Inquiry (**8.3%**), In Progress (**8.3%**), and Not Enrolled (**8.3%**). 'No Information' cases are high (**50%**).
- Geographic Distribution: Largest applicant base from India, followed by Ghana, Bangladesh, Nigeria, Pakistan, and Nepal. Emerging interest from South Africa, Iran, UK, and Egypt.

- Agent Performance: Certain agents like Rudra appear frequently across multiple stages, suggesting high activity but uneven documentation.

Trends

- Growing Call Volumes: Peaks in **July and December 2024** indicate seasonal recruitment drives.
- Regional Concentration: Heavy dependence on South Asian countries (**India, Bangladesh, Pakistan, Nepal**).
- Data Gaps: High levels of '**No Information**' outcomes suggest inconsistencies in applicant tracking.

Actionable Recommendations

- Improve Data Tracking: Reduce '**No Information**' cases with stricter protocols and CRM integration.
- Optimize Campaigns: Invest more in high-performing campaigns; analyze weaker ones.
- Agent Training & Accountability: Train agents for accurate data entry; implement KPIs based on outcomes.
- Expand Geographic Outreach: Diversify efforts into **Africa** and the **Middle East**.
- Stage Conversion Focus: Prioritize 'In Progress' and 'Incomplete' applicants with follow-ups to boost conversion

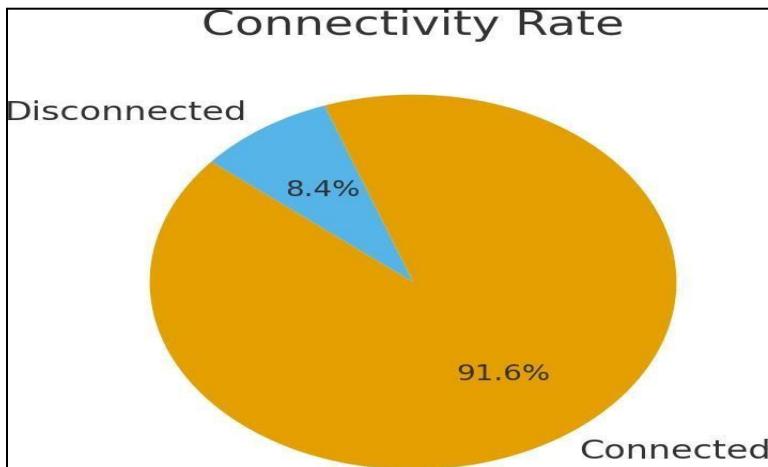
Summary

The master table brings together applicant, outreach, and campaign datasets into a **single, comprehensive view**, enabling a complete understanding of interactions and performance. By calculating key metrics such as connected calls, disconnected calls, connectivity rate, and agent performance, the table provides:

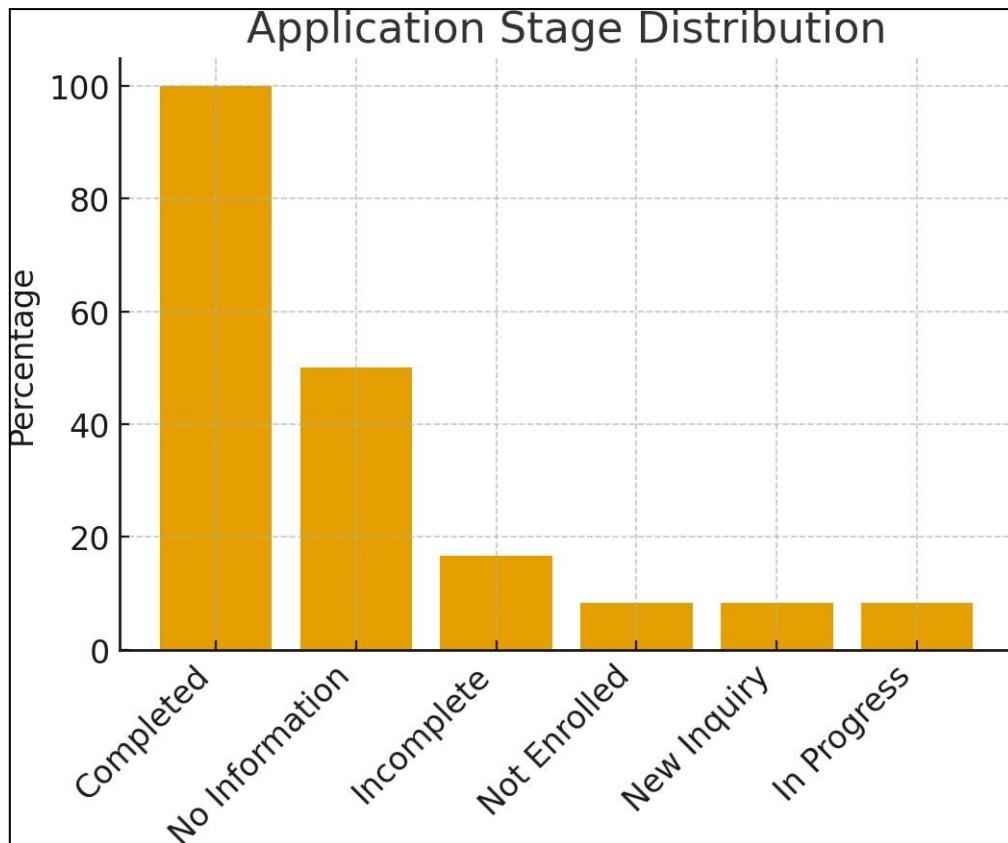
- Insightful measures of outreach effectiveness.
- Thorough data cleaning and validation including removal of duplicates, invalid IDs, null timestamps, and unnecessary columns
- Ensuring the dataset is accurate, consistent, and reliable.
- Overall, this master table serves as a solid foundation for analyzing agent efficiency
- Optimizing campaigns, and supporting informed decision-making across the outreach process

Visual Insights

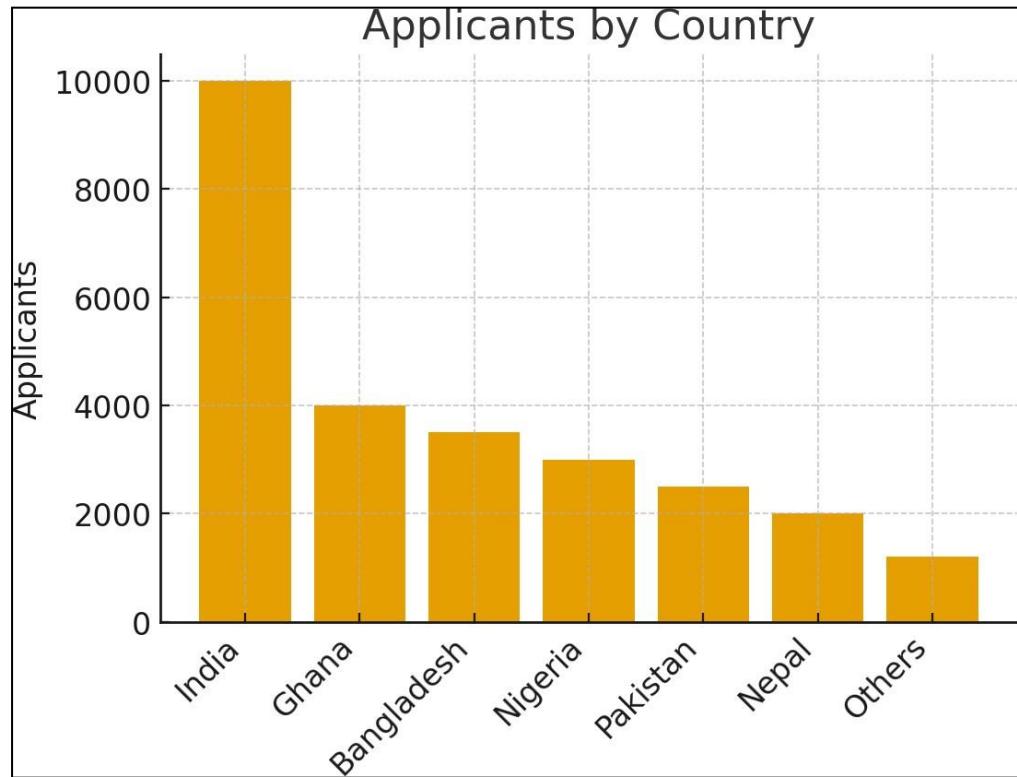
Connectivity Rate



Application Stage Distribution

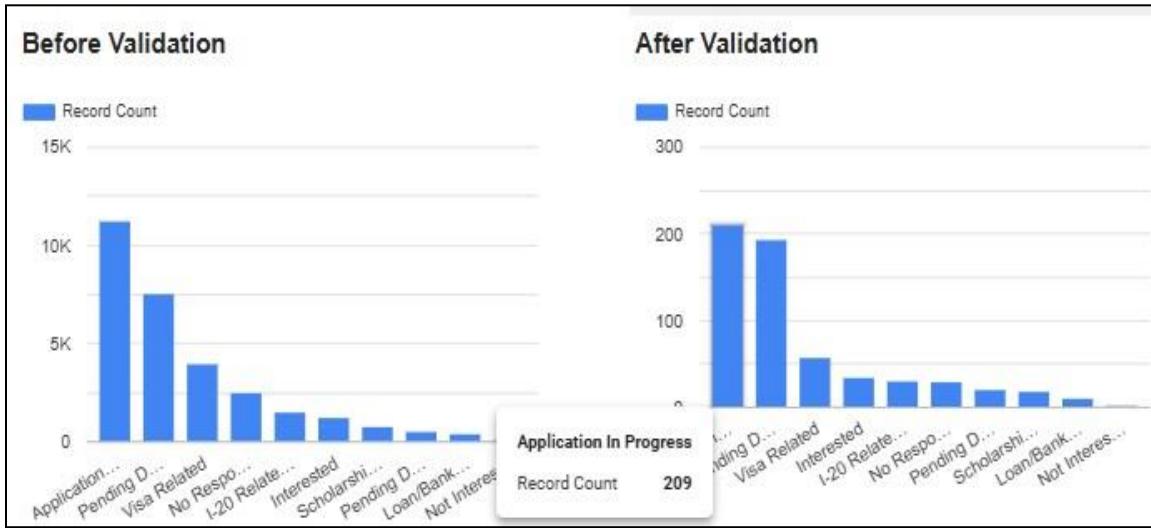


Applicants by Country



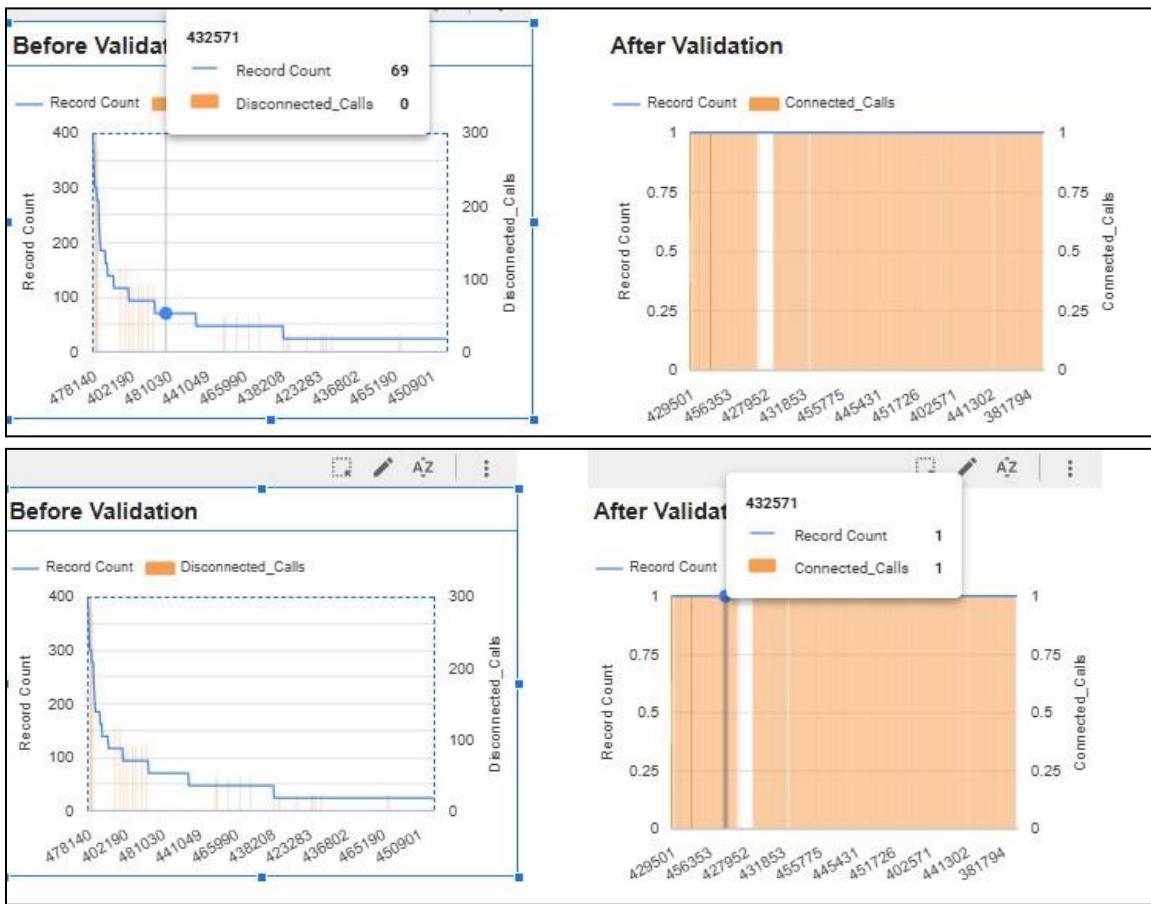
Applicant Progress (Before vs After Validation)





Here we can see, In the **"Application in Progress"** outcome, the record count dropped from **11,178** to **209** after validation. It shows that the duplicacy has been removed.

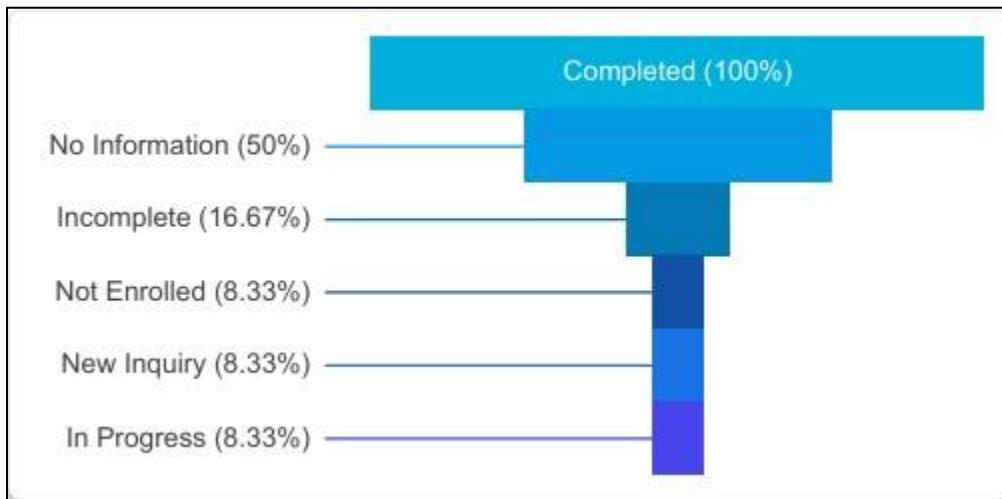
Applicant with Connected & Disconnected calls



Observations:

- For Applicant ID 432571, record count was reduced from 69 to 1 after validation.
- Disconnected Calls remained 0, confirming no actual call data was lost.
- The reduction shows that 68 duplicate records were removed and only unique applicants remain.

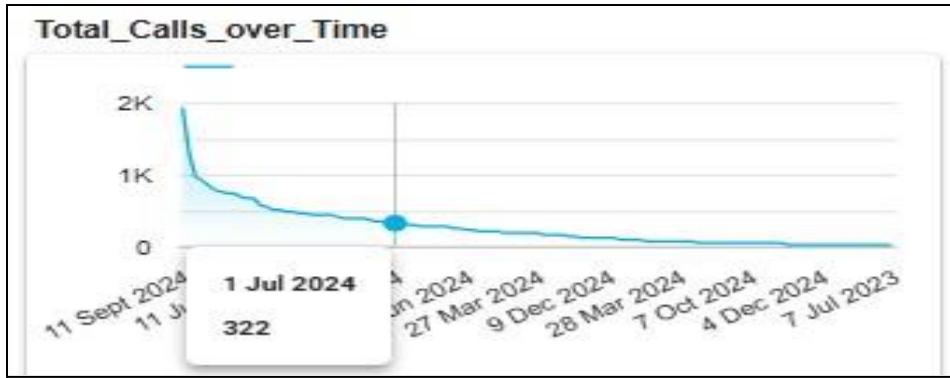
Stage-Status by Total Calls



Observations:

- **Highest Drop-off :** 50% records fall under “**No Information**” - indicates major data capture gap.
- **Validation Impact :** Duplicates & invalids removed - huge reduction in inflated counts.
- **Balanced Lower Categories :** “**Not Enrolled**,” “**New Inquiry**,” and “**In Progress**” each ~8.33%.
- **True Representation :** Clean dataset ensures funnel now reflects actual applicant journey.
- **Next Step Focus :** Prioritize reducing “**No Information**” & “**Incomplete**” stages to improve enrollment.

Total Calls Over Time



Observations:

- **High Starting Point:**

At the beginning (around Sept 2023), the total calls were very high (above 2,000).

- **Sharp Decline:**

Shortly after, there's a steep downward trend, meaning the number of calls reduced significantly.

- **Gradual Decrease:**

From early 2024 onwards, the line smooths out, indicating calls were steadily decreasing but at a slower rate.

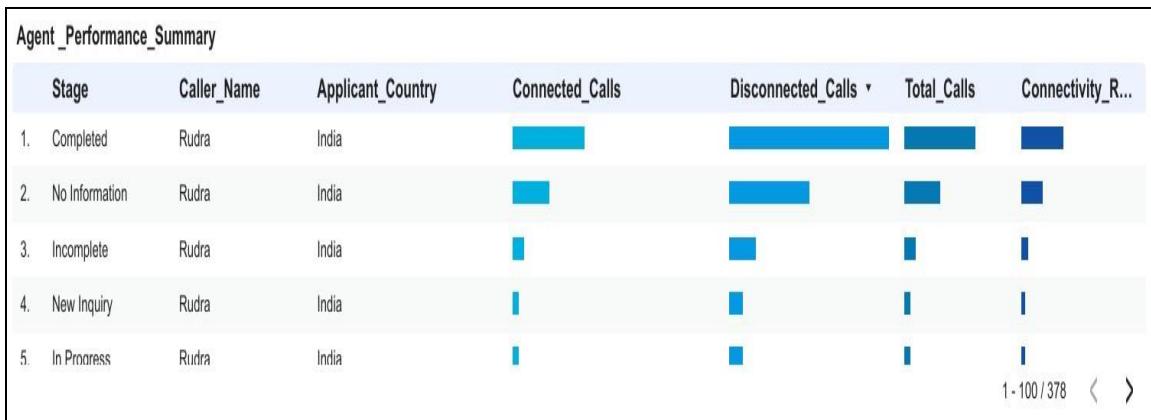
- Data Point Highlighted (1 July 2024):

- On 1 July 2024, the total calls recorded were 322.

- This suggests a huge drop compared to the initial values.

Agent Performance Summary

This is a performance summary dashboard for an agent, displaying how calls were distributed across different stages. The table has both numbers (hidden behind bars) and horizontal bar visualizations to quickly compare call volumes.



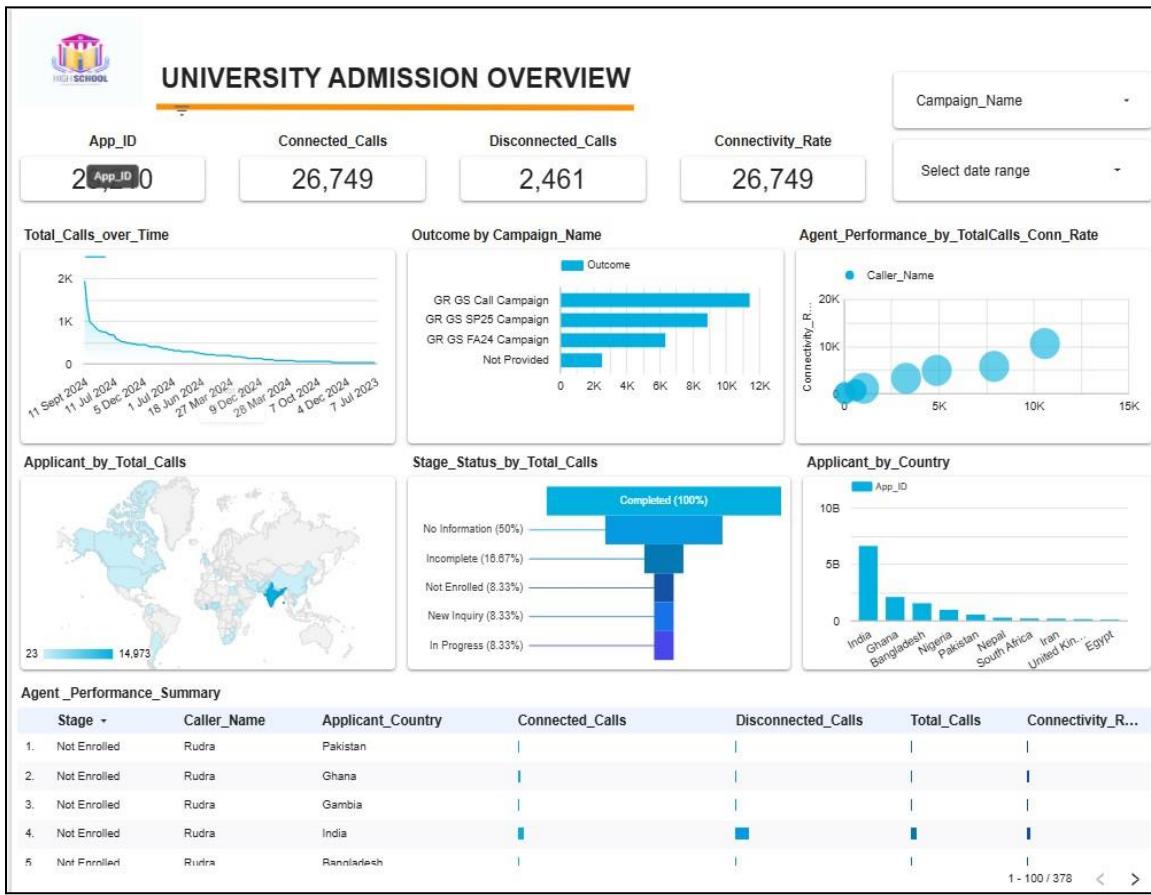
Observations:

Stage – The current status of calls:

1. Completed
2. No Information
3. Incomplete
4. New Inquiry
5. In Progress
6. **Caller_Name** - The agent handling calls
7. **Applicant_Country** - The applicant's location (*India* in all cases here).
8. **Connected_Calls** - Calls where the applicant actually picked up and engaged. The length of the cyan bar = number of connected calls.
9. **Disconnected_Calls** - Calls attempted but not connected (e.g., missed, declined). The blue bar represents disconnected calls.
10. **Total_Calls** - Sum of connected + disconnected calls.
The bar here combines the contributions from both categories.
11. **Connectivity_Rate** - (likely Connectivity_Rate) – Ratio of connected calls to total calls.
Darker blue bar shows higher connectivity rate.

Final Dashboard Observation after validation

This dashboard gives a full admissions performance view: it tracks call activity, agent efficiency, applicant geography, campaign effectiveness, and admission funnel progress. It highlights strong outreach but also gaps in applicant follow-through (many not enrolled or incomplete).



- **KPIs:** 2,490 applications, ~26.7K connected calls, 2.4K disconnected calls.
- **Call Trends:** High outreach at the start (Sept 2023), declining over time.
- **Campaigns:** Most calls came from the *GR GS Call Campaign*, followed by SP25 and FA24.
- **Agents:** Some agents handled >10K calls with strong connectivity (bubble chart).
- **Applicants:** Majority from India, with notable numbers from Ghana, Bangladesh, Nigeria, and Pakistan.
- **Stages:** Half of applicants had “No Information”; ~17% incomplete; only a small portion moved to “Enrolled.”
- **Agent Summary:** Rudra managed applicants across multiple countries, but many ended in “Not Enrolled.”

Overall: Strong call activity, but low conversion to enrollment, with India as the top applicant source.

Data Validation Report

Presenter

Farhan Hussain - Shivansh Rajput
M.Nouman Arshad - Yuva Mirsha
Dia Agarwal - Naheed Akhtar
Saurav Kumar



DATA VALIDATION REPORT

PREPARED FOR DATA QUALITY AND INSIGHTS

- Presentation Objective: Showcase the process of validating, cleaning, and preparing the master dataset for reliable analysis and reporting.
- Purpose of Validation : The purpose of validation is to ensure data accuracy, consistency, and completeness within the Master Table, enabling reliable analysis and informed decision-making throughout the project lifecycle.
- Data Sources Overview : The master table consolidates data from multiple sources including internal databases, third-party providers, and real-time streams, ensuring comprehensive coverage and consistency for validation processes.

The screenshot shows a database interface with three tables displayed sequentially:

- Application Table:** Columns: App_ID (bigint), Applicant_Country (text), Applicant_University (text). Data: Rows 1-6 show repeated entries for App_ID 12345, Applicant_Country India, and Applicant_University Illinois Institute of Technolo... . Total rows: 2467563. Query complete 00:00:19.346.
- Campaign Table:** Columns: ID (text), Name (text), Category (text), Intake (text). Data: Rows 1-6 show entries for various campaigns: AANF23, AND23, BPNAN..., BPNND..., DANE24, and CTKANF... . All categories are Post Admission and intake is AY2024.
- Call_Records Table:** Columns: Reference_ID (double precision), Received_At (date), University (text), Caller_Name (text). Data: Rows 1-6 show calls received at various dates (e.g., 2024-03-20, 2024-11-05) from different callers (Rudra, Shailja, Isha) to the Illinois Institute of Technology.

DATA VALIDATION REPORT – MASTER TABLE

INTRODUCTION



Purpose

Ensure accuracy, consistency, and completeness of the master dataset, establishing a trustworthy foundation for further analysis and reporting.



Original Size

The dataset initially contained approximately 24 lakh rows, representing a vast pool of applicant and campaign information.



Data Sources

Data compiled from Applicant Data, Outreach Data, and Campaign Data to build a comprehensive master_table.



Current Focus

Validation and cleaning of the master_table to prepare it for downstream insights and decision-making.

RECORD COUNT

BEFORE AND AFTER CLEANING



Before Cleaning

The master dataset contained nearly 24 lakh rows, encompassing applicant, outreach, and campaign information in its raw form

Total rows: 2467563

Query complete 00:00:25.569



Impact

The resulting dataset is more manageable and computationally efficient while preserving accuracy and integrity.



After Cleaning

The dataset size reduced significantly after duplicate removal, invalid ID elimination, and deletion of records with excessive null values

Total rows: 594

Query complete 00:00:02.865



Validation

This record count comparison highlights the importance of rigorous validation in maintaining dataset usability for analysis.

DUPLICATE HANDLING

ENSURING UNIQUE APPLICANT RECORDS



Duplicate Check

Duplicates were identified using App_ID as the unique identifier across the dataset.



Multiple Entries

Several applicants had multiple rows recorded, often due to repeated outreach or updates in application details.



Retention Rule

For each applicant, only the latest record was preserved based on the Time_Stamp field, ensuring recency of information.



Outcome

All duplicate records were successfully removed, leaving a single, validated entry per applicant.

MISSING VALUE TREATMENT

ADDRESSING NULL AND INCOMPLETE DATA

- Dropped High-Null Columns: Columns such as Phone_Number, Reference_ID, Batch_Code, Followup, Outreach_University, Campaign_ID, Campaign_Status, and Campaign_University were eliminated due to excessive null values.
- Removed Null Timestamps: Rows with missing values in critical timestamp fields (Received_At / Time_Stamp) were deleted to maintain temporal consistency.
- Key Columns Validation: Ensured completeness in essential attributes including App_ID, Applicant_Country, Applicant_University, Outcome, and Campaign_Name.
- Improved Data Quality: The dataset now emphasizes completeness, eliminating ambiguity from partial or missing records.



VALIDATIONS PROOF

VISUALS AND QUERIES SCRIPT

Data Output Messages Notifications							
Showing rows: 1 to 1000 / 2468 / Page No: 1 of 2468							
App_ID	Applicant_Country	Applicant_University	Phone_Number	Reference_ID	Received_At	Time_Stamp	Caller_Name
1	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
2	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
3	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
4	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
5	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
6	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
7	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
8	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
9	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
10	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
11	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
12	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00
13	12345	India	Illinois Institute of Technolo...	8019011222	12345	2023-04-28	13:04:00

Total rows: 2467563 Query complete 00:00:16.433 CRLF Ln 1, Col 1

Data Output Messages Notifications							
Showing rows: 1 to 594 / 594 / Page No: 1 of 1							
App_ID	Applicant_Country	Applicant_University	Received_At	Time_Stamp	Caller_Name	Outcome	Remark
1	421192	India	Illinois Institute of Technology	2024-04-20 18:22:00	Rudra	No Response	not connected
2	471044	Ghana	Illinois Institute of Technology	2024-04-05 18:15:00	Rudra	No Response	student will join sp25 sessions
3	425901	India	Illinois Institute of Technology	2024-04-28 18:54:00	Rudra	Pending Documents	will submit the docx
4	459595	India	Illinois Institute of Technology	2024-04-05 12:47:00	Rudra	Interested	not interested to it
5	565548	India	Illinois Institute of Technology	2023-09-12 12:47:00	Poppy	Application in Progress	not interested
6	367708	India	Illinois Institute of Technology	2023-09-03 13:51:00	Namrata	I-20 Related Issue	completed application
7	373465	India	Illinois Institute of Technology	2023-09-11 16:16:00	Shalja	No Response	reschedule
8	244742	India	Illinois Institute of Technology	2023-09-12 14:20:00	Poppy	Pending Documents	completed application
9	366971	India	Illinois Institute of Technology	2023-09-12 13:05:00	Shalja	I-20 Related Issue	completed application
10	248519	India	Illinois Institute of Technology	2023-09-12 17:08:00	Poppy	Application in Progress	not interested
11	366637	India	Illinois Institute of Technology	2023-09-12 17:33:00	Poppy	Pending Documents	will submit the docx
12	356664	South Africa	Illinois Institute of Technology	2023-09-25 18:08:00	Poppy	Pending Documents	will submit the docx
13	364554	India	Illinois Institute of Technology	2023-09-29 14:58:00	Shalja	I-20 Related Issue	completed application
14	355548	India	Illinois Institute of Technology	2023-07-03 16:54:00	Shalja	Visa Related	ready to pay the deposit
15	351960	Pakistan	Illinois Institute of Technology	2023-07-05 17:49:00	Isha	Application in Progress	ready to pay the deposit
16	350602	Pakistan	Illinois Institute of Technology	2023-07-07 18:20:00	Poppy	Visa Related	Not Provided
17	450155	India	Illinois Institute of Technology	2023-09-18 15:50:00	Prayval	Pending Documents	want to defer
18	351188	India	Illinois Institute of Technology	2023-09-01 13:56:00	Shalja	Interested	not interested
19	424987	India	Illinois Institute of Technology	2024-04-21 14:13:00	Rudra	Pending Decision	Not Provided
20	410407	India	Illinois Institute of Technology	2024-04-21 14:15:00	Rudra	Pending Decision	will confirm later
21	427742	India	Illinois Institute of Technology	2024-04-21 14:16:00	Rudra	Pending Decision	will confirm later
22	423223	India	Illinois Institute of Technology	2024-04-21 14:28:00	Rudra	Pending Decision	will confirm later
23	447745	India	Illinois Institute of Technology	2024-04-28 13:19:00	Prayval	Pending Documents	completed application
24	450228	India	Illinois Institute of Technology	2024-04-29 18:08:00	Prayval	Pending Documents	completed application
25	450204	India	Illinois Institute of Technology	2024-04-29 18:28:00	Rudra	Pending Documents	will submit the docx
26	400191	India	Illinois Institute of Technology	2024-04-29 19:01:00	Shalja	Pending Documents	Not Provided

Total rows: 594 Query complete 00:00:02.865 CRLF Ln 1, Col 1

Before Validation

App_ID	Record Count
1.	474311
2.	438353
3.	478295
4.	476993
5.	480925
6.	478281
7.	480914

1 - 100 / 12210 < >

After_Validation

App_ID	Record Count
1.	429501
2.	365548
3.	367708
4.	375465
5.	366971
6.	346519
7.	366637

1 - 100 / 594 < >

Dataset Validation: Record Count Before vs After



```
DELETE FROM "master_table"
WHERE LENGTH("App_ID")::text) <> 6;
```

Data Output Messages Notifications

UPDATE 594

Query returned successfully in 2 secs 738 msec.

```
ALTER TABLE "master_table"
DROP COLUMN "Phone_Number",
DROP COLUMN "Reference_ID",
DROP COLUMN "Batch_Code",
DROP COLUMN "Followup",
DROP COLUMN "Outreach_University",
DROP COLUMN "Campaign_ID",
DROP COLUMN "Campaign_Status",
DROP COLUMN "Campaign_University";
```

#	App_ID	Applicant	Applicant_University	Received_At	Time_Stamp	Caller_Name	Outcome	Remark	Campaign_Name	Campaign_Category	Campaign_JobId	Start_Date	End_Date
1	427192	India	Illinois Institute of Technolo...	2024-05-20	10:22:00	Rudra	No Response	not connected	GR GS SP25 Campa...	Pre Admission	A10024	2024-10-18	No
2	471044	Ghana	Illinois Institute of Technolo...	2024-05-15	18:51:00	Rudra	No Response	student will join sp25 sess...	GR GS FA24 Campa...	Post Admission	A10024	2024-09-11	On
3	426501	India	Illinois Institute of Technolo...	2024-05-20	18:34:00	Rudra	Pending Documents	will submit the docx	GR GS FA24 Campa...	Pre Admission	A10024	2024-09-06	On
4	450350	India	Illinois Institute of Technolo...	2024-05-15	12:47:00	Rudra	Interested	not interested to it	GR GS Call Campaign	Post Admission	A10024	2024-07-05	On
5	365548	India	Illinois Institute of Technolo...	2023-05-12	12:47:00	Poppy	Application In Progress	not interested	GR GS Call Campaign	Pre Admission	A10024	2023-04-28	In
6	357708	India	Illinois Institute of Technolo...	2023-05-20	13:51:00	Namrata	100 Related Issue	completed application	Not Provided	Post Admission	A10024	2024-07-17	No
7	375465	India	Illinois Institute of Technolo...	2023-05-11	16:16:00	Shalja	No Response	reschedule	GR GS FA24 Campa...	Post Admission	A10024	2024-09-20	On
8	349472	India	Illinois Institute of Technolo...	2023-05-12	12:40:00	Poppy	Pending Documents	completed application	GR GS Call Campaign	Pre Admission	A10024	2023-05-16	No
9	366871	India	Illinois Institute of Technolo...	2023-05-12	13:05:00	Shalja	100 Related Issue	completed application	GR GS FA24 Campa...	Post Admission	A10024	2024-09-20	On
10	344619	India	Illinois Institute of Technolo...	2023-05-12	17:00:00	Poppy	Application In Progress	not interested	GR GS SP25 Campa...	Pre Admission	A10024	2024-10-09	In
11	366857	India	Illinois Institute of Technolo...	2023-05-12	17:33:00	Poppy	Pending Documents	will submit the docx	GR GS SP25 Campa...	Pre Admission	A10024	2024-10-07	On
12	355608	South...	Illinois Institute of Technolo...	2023-05-20	10:08:00	Poppy	Pending Documents	will submit the docx	GR GS FA24 Campa...	Post Admission	A10024	2024-09-11	On
13	365524	India	Illinois Institute of Technolo...	2023-05-20	14:56:00	Shalja	100 Related Issue	completed application	GR GS SP25 Campa...	Pre Admission	A10024	2024-10-09	In
14	355648	India	Illinois Institute of Technolo...	2023-05-03	16:54:00	Shalja	Via Related	ready to pay the deposit	GR GS Call Campaign	Post Admission	A10024	2023-07-05	On
15	351961	Pakistan	Illinois Institute of Technolo...	2023-05-05	17:49:00	Isha	Application In Progress	ready to pay the deposit	Not Provided	Post Admission	A10024	2024-07-17	On

Total rows 594 - Query complete 00:00:50.838 CRLF - Ln 3 Col 1

```
DELETE FROM "master_table"
WHERE "Received_At" IS NULL
OR "Time_Stamp" IS NULL;
```

Data Output Messages Notifications

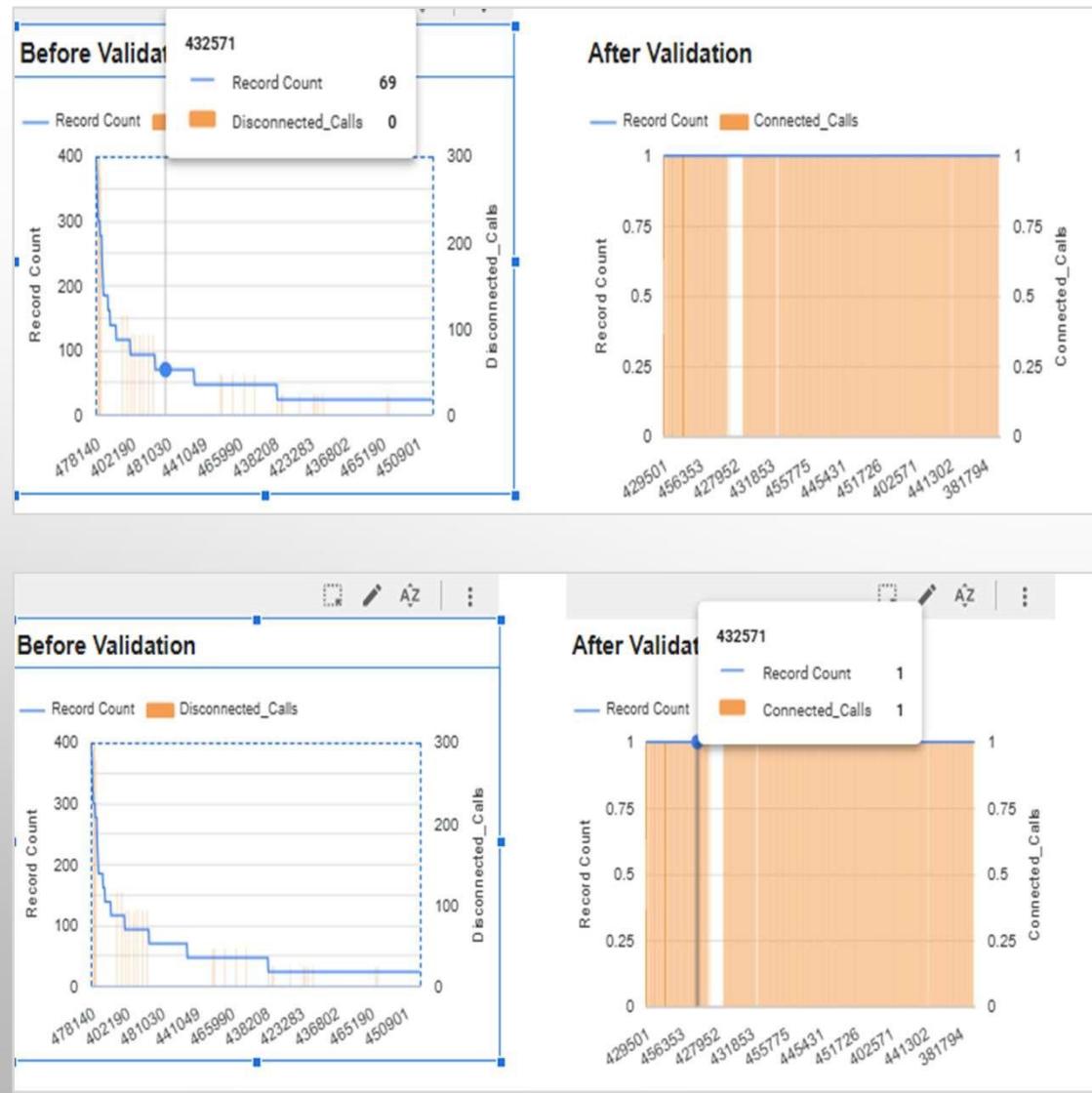
DELETE 192

Query returned successfully in 2 secs 518 msec.



Difference Between Before & After Validation

- In the “Application in Progress” outcome, record count dropped from 11,178 to 209 after validation.
- In the “Pending Document” outcome, record count reduced from 7,475 to 192.
- This significant reduction occurred due to removal of duplicate records and invalid entries, which were previously inflating counts and hampering accurate analysis.



Difference Between Before & After Validation

- For Applicant ID 432571, record count reduced from 69 to 1 after validation.
- Disconnected Calls remained 0, confirming no actual call data was lost.
- The reduction shows that 69 duplicate records were removed.
- These duplicates could have distorted connectivity analysis and inflated applicant activity.
- Post-validation, each applicant ID now has a single, accurate record, ensuring reliability.

BUSINESS RULE CHECKS

VALIDATING DATA AGAINST DEFINED STANDARDS

- App_ID Format: Verified that each App_ID adheres to the required 6-digit structure, ensuring uniqueness and standardization.
 - Connectivity Rate: Checked that all connectivity rates fall strictly between 0 and 1, preventing invalid metrics.
 - Agent Performance: Validated that performance categories are limited to Good, Average, Poor, or Not Evaluated, ensuring consistency.
 - Campaign Validation: Cross-checked Campaign Intake, Stage, and Outcomes against predefined categories for accuracy and compliance.

ACTIONS TAKEN

STEPS IMPLEMENTED FOR DATA VALIDATION

- Column Pruning: Dropped inappropriate and high-null columns to reduce redundancy and noise.
- Invalid ID Removal: Filtered out records with invalid App_IDs that did not meet required standards.
- Duplicate Elimination: Ensured only the latest applicant records were retained by removing duplicates.
- Timestamp Validation: Removed rows with missing or invalid timestamps, maintaining chronological integrity.
- New Metrics Created: Introduced Connected Calls, Disconnected Calls, Connectivity Rate, and Agent Performance measures for richer insights.

CONCLUSION

VALIDATED DATASET READY FOR INSIGHTS

- Clean and Consistent: Final dataset free of duplicates, invalid IDs, and high-null columns, ensuring reliability.
 - Validated for Analysis: Key fields validated and business rules applied to guarantee usability for reporting.
 - Foundation for Insights: Dataset now supports meaningful visualizations, accurate metrics, and actionable insights.
 - Decision-Ready: Ensures strong basis for informed decisions, strategy development, and operational improvements.

Cleaned & Validated Master Table

Global Project Status Report - Q3 2024										
Project ID		Project Details			Team & Resources			Performance Metrics		
Category	ID	Name	Type	Description	Manager	Lead Dev	Team Size	Budget (USD)	Progress (%)	Completion Date
Software Development	PJ-001	Project Alpha	Web Application	Developing a new web application for internal use.	J. Doe	A. Patel	15	120,000	85%	2024-09-30
Software Development	PJ-002	Project Beta	Mobile App	Developing a mobile application for Android and iOS.	J. Doe	B. Singh	10	80,000	70%	2024-11-30
Hardware Engineering	PJ-003	Project Gamma	Hardware Design	Designing a new hardware component for a specific industry.	M. Lee	C. Kumar	12	150,000	60%	2024-10-31
Hardware Engineering	PJ-004	Project Delta	Hardware Assembly	Assembling and testing hardware components.	M. Lee	D. Patel	18	90,000	45%	2024-12-31
Infrastructure	PJ-005	Project Epsilon	Cloud Migration	Migrating existing infrastructure to the cloud.	N. Patel	E. Singh	10	70,000	30%	2024-11-15
Infrastructure	PJ-006	Project Zeta	Network Upgrade	Upgrading network infrastructure across multiple locations.	N. Patel	F. Kumar	15	100,000	20%	2024-12-15
Research & Development	PJ-007	Project Eta	AI Research	Researching and developing AI algorithms.	G. Singh	H. Patel	8	60,000	15%	2025-01-31
Research & Development	PJ-008	Project Theta	Machine Learning	Developing machine learning models for various applications.	G. Singh	I. Kumar	10	80,000	10%	2025-02-28
Testing	PJ-009	Project Iota	QA Testing	Performing quality assurance testing on software and hardware.	K. Patel	J. Kumar	12	40,000	90%	2024-10-01
Testing	PJ-010	Project Kappa	UAT Testing	Performing user acceptance testing on completed projects.	K. Patel	L. Singh	10	30,000	80%	2024-11-01
Support	PJ-011	Project Lambda	Customer Support	Providing support to clients and users.	M. Singh	N. Patel	8	20,000	95%	2024-12-01
Support	PJ-012	Project Mu	Technical Support	Providing technical support to clients and users.	M. Singh	O. Kumar	10	30,000	90%	2024-12-15
Training	PJ-013	Project Nu	Employee Training	Providing training to employees on new software and tools.	P. Patel	Q. Singh	10	10,000	100%	2024-12-31
Training	PJ-014	Project Xi	Client Training	Providing training to clients on new software and tools.	P. Patel	R. Kumar	12	15,000	95%	2025-01-15
Marketing	PJ-015	Project Omicron	Marketing Campaigns	Planning and executing marketing campaigns.	S. Singh	T. Patel	10	50,000	75%	2024-10-15
Marketing	PJ-016	Project Pi	Social Media	Managing social media presence and engagement.	S. Singh	U. Kumar	8	30,000	60%	2024-11-15
Marketing	PJ-017	Project Rho	Email Marketing	Planning and executing email marketing campaigns.	V. Patel	W. Singh	10	20,000	50%	2024-12-15
Marketing	PJ-018	Project Sigma	Content Marketing	Creating and publishing content for various platforms.	V. Patel	X. Kumar	12	30,000	40%	2025-01-15
Logistics	PJ-019	Project Tau	Supply Chain	Managing supply chain and logistics operations.	Y. Patel	Z. Singh	10	40,000	80%	2024-11-01
Logistics	PJ-020	Project Upsilon	Inventory Management	Managing inventory levels and distribution.	Y. Patel	A. Kumar	12	30,000	70%	2024-12-01
Logistics	PJ-021	Project Phi	Delivery Services	Providing delivery services to clients.	B. Singh	C. Patel	10	20,000	60%	2024-12-15
Logistics	PJ-022	Project Chi	Warehouse Management	Managing warehouse operations and storage.	B. Singh	D. Kumar	12	30,000	50%	2025-01-15
Customer Support	PJ-023	Project Psi	Client Support	Providing support to clients.	E. Patel	F. Singh	10	15,000	90%	2024-10-01
Customer Support	PJ-024	Project Omega	Feedback Management	Managing client feedback and suggestions.	E. Patel	G. Kumar	12	20,000	80%	2024-11-01
Customer Support	PJ-025	Project Epsilon	Retention Programs	Developing retention programs for clients.	H. Patel	I. Singh	10	10,000	70%	2024-12-01
Customer Support	PJ-026	Project Zeta	Referral Programs	Developing referral programs for clients.	H. Patel	J. Kumar	12	15,000	60%	2024-12-15
Customer Support	PJ-027	Project Iota	Retention Analysis	Performing analysis on client retention rates.	K. Patel	L. Singh	10	10,000	50%	2025-01-15
Customer Support	PJ-028	Project Kappa	Referral Analysis	Performing analysis on referral rates.	K. Patel	M. Kumar	12	15,000	40%	2025-02-28
Product Development	PJ-029	Project Lambda	New Product Research	Researching and developing new products.	N. Patel	O. Singh	10	80,000	70%	2024-10-01
Product Development	PJ-030	Project Mu	New Product Design	Designing new products.	N. Patel	P. Kumar	12	100,000	60%	2024-11-01
Product Development	PJ-031	Project Nu	New Product Testing	Testing new products.	Q. Patel	R. Singh	10	50,000	50%	2024-12-01
Product Development	PJ-032	Project Xi	New Product Launch	Planning and launching new products.	Q. Patel	S. Kumar	12	70,000	40%	2024-12-15
Product Development	PJ-033	Project Omicron	New Product Support	Providing support for new products.	T. Patel	U. Singh	10	30,000	30%	2025-01-15
Product Development	PJ-034	Project Pi	New Product Iterations	Iterating on new products.	T. Patel	V. Kumar	12	50,000	20%	2025-02-28
Product Development	PJ-035	Project Rho	New Product Feedback	Managing feedback for new products.	W. Patel	X. Singh	10	20,000	10%	2025-03-15
Product Development	PJ-036	Project Sigma	New Product Retention	Managing retention for new products.	W. Patel	Y. Kumar	12	30,000	0%	2025-04-15
Product Development	PJ-037	Project Tau	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2025-05-15
Product Development	PJ-038	Project Upsilon	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2025-06-15
Product Development	PJ-039	Project Phi	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2025-07-15
Product Development	PJ-040	Project Chi	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2025-08-15
Product Development	PJ-041	Project Psi	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2025-09-15
Product Development	PJ-042	Project Omega	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2025-10-15
Product Development	PJ-043	Project Epsilon	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2025-11-15
Product Development	PJ-044	Project Zeta	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2025-12-15
Product Development	PJ-045	Project Iota	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2026-01-15
Product Development	PJ-046	Project Kappa	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2026-02-15
Product Development	PJ-047	Project Lambda	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2026-03-15
Product Development	PJ-048	Project Mu	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2026-04-15
Product Development	PJ-049	Project Nu	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2026-05-15
Product Development	PJ-050	Project Xi	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2026-06-15
Product Development	PJ-051	Project Omicron	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2026-07-15
Product Development	PJ-052	Project Pi	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2026-08-15
Product Development	PJ-053	Project Rho	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2026-09-15
Product Development	PJ-054	Project Sigma	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2026-10-15
Product Development	PJ-055	Project Tau	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2026-11-15
Product Development	PJ-056	Project Upsilon	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2026-12-15
Product Development	PJ-057	Project Phi	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2027-01-15
Product Development	PJ-058	Project Chi	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2027-02-15
Product Development	PJ-059	Project Psi	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2027-03-15
Product Development	PJ-060	Project Omega	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2027-04-15
Product Development	PJ-061	Project Epsilon	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2027-05-15
Product Development	PJ-062	Project Zeta	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2027-06-15
Product Development	PJ-063	Project Iota	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2027-07-15
Product Development	PJ-064	Project Kappa	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2027-08-15
Product Development	PJ-065	Project Lambda	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2027-09-15
Product Development	PJ-066	Project Mu	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2027-10-15
Product Development	PJ-067	Project Nu	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2027-11-15
Product Development	PJ-068	Project Xi	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2027-12-15
Product Development	PJ-069	Project Omicron	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2028-01-15
Product Development	PJ-070	Project Pi	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2028-02-15
Product Development	PJ-071	Project Rho	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2028-03-15
Product Development	PJ-072	Project Sigma	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2028-04-15
Product Development	PJ-073	Project Tau	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2028-05-15
Product Development	PJ-074	Project Upsilon	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2028-06-15
Product Development	PJ-075	Project Phi	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2028-07-15
Product Development	PJ-076	Project Chi	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2028-08-15
Product Development	PJ-077	Project Psi	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2028-09-15
Product Development	PJ-078	Project Omega	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2028-10-15
Product Development	PJ-079	Project Epsilon	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2028-11-15
Product Development	PJ-080	Project Zeta	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2028-12-15
Product Development	PJ-081	Project Iota	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2029-01-15
Product Development	PJ-082	Project Kappa	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2029-02-15
Product Development	PJ-083	Project Lambda	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2029-03-15
Product Development	PJ-084	Project Mu	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2029-04-15
Product Development	PJ-085	Project Nu	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2029-05-15
Product Development	PJ-086	Project Xi	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2029-06-15
Product Development	PJ-087	Project Omicron	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2029-07-15
Product Development	PJ-088	Project Pi	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2029-08-15
Product Development	PJ-089	Project Rho	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2029-09-15
Product Development	PJ-090	Project Sigma	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2029-10-15
Product Development	PJ-091	Project Tau	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2029-11-15
Product Development	PJ-092	Project Upsilon	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2029-12-15
Product Development	PJ-093	Project Phi	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2030-01-15
Product Development	PJ-094	Project Chi	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2030-02-15
Product Development	PJ-095	Project Psi	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2030-03-15
Product Development	PJ-096	Project Omega	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2030-04-15
Product Development	PJ-097	Project Epsilon	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2030-05-15
Product Development	PJ-098	Project Zeta	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2030-06-15
Product Development	PJ-099	Project Iota	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2030-07-15
Product Development	PJ-100	Project Kappa	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2030-08-15
Product Development	PJ-101	Project Lambda	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2030-09-15
Product Development	PJ-102	Project Mu	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2030-10-15
Product Development	PJ-103	Project Nu	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2030-11-15
Product Development	PJ-104	Project Xi	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2030-12-15
Product Development	PJ-105	Project Omicron	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2031-01-15
Product Development	PJ-106	Project Pi	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2031-02-15
Product Development	PJ-107	Project Rho	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2031-03-15
Product Development	PJ-108	Project Sigma	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	12	50,000	0%	2031-04-15
Product Development	PJ-109	Project Tau	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	10	30,000	0%	2031-05-15
Product Development	PJ-110	Project Upsilon	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	12	50,000	0%	2031-06-15
Product Development	PJ-111	Project Phi	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	10	20,000	0%	2031-07-15
Product Development	PJ-112	Project Chi	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	12	70,000	0%	2031-08-15
Product Development	PJ-113	Project Psi	New Product Support	Providing support for new products.	Z. Patel	Z. Singh	10	30,000	0%	2031-09-15
Product Development	PJ-114	Project Omega	New Product Iterations	Iterating on new products.	Z. Patel	Z. Singh	12	50,000	0%	2031-10-15
Product Development	PJ-115	Project Epsilon	New Product Feedback	Managing feedback for new products.	Z. Patel	Z. Singh	10	20,000	0%	2031-11-15
Product Development	PJ-116	Project Zeta	New Product Retention	Managing retention for new products.	Z. Patel	Z. Singh	12	30,000	0%	2031-12-15
Product Development	PJ-117	Project Iota	New Product Launch	Planning and launching new products.	Z. Patel	Z. Singh	10	70,000	0%	2032-01-15
Product Development	PJ-118	Project Kappa								

RECOMMENDATIONS

- AUTOMATE VALIDATION PIPELINES – SET UP RECURRING SCRIPTS TO CONTINUOUSLY DETECT AND HANDLE DUPLICATES, NULLS, AND INVALID RECORDS.
- ESTABLISH DATA QUALITY METRICS – DEFINE KPI'S SUCH AS DUPLICATE RATIO, NULL PERCENTAGE, AND ERROR FREQUENCY TO MONITOR DATASET HEALTH.
- IMPLEMENT STRONG DATA ENTRY CONTROLS – ENFORCE FIELD VALIDATIONS (E.G., APP_ID FORMAT, MANDATORY TIMESTAMP) TO REDUCE ERRORS AT THE SOURCE.
- MAINTAIN AUDIT & TRACEABILITY LOGS – DOCUMENT EVERY VALIDATION STEP (RECORDS DROPPED, RULES APPLIED) FOR ACCOUNTABILITY AND REPRODUCIBILITY.
- STANDARDIZE DATA GOVERNANCE PRACTICES – CREATE CLEAR POLICIES FOR DATA OWNERSHIP, VALIDATION RULES, AND PERIODIC QUALITY CHECKS ACROSS ALL DATASETS.



**THANK
YOU!**