University Campaign Marketing Insights and Visualization

Mahir Faisal Chowdhury

Presentation Structure

- Introduction
- Validation
- Findings
  - Overview & Key Metrics
  - Campaign Types
  - Calls Over Time
  - Agent Performance
  - Region Distribution
- Summary and Recommendations

Introduction

“The purpose of this presentation is to identify key trends in our outreach campaigns and to find ways to improve campaign success. In simple terms, we want to understand what makes a campaign successful, and how we can ensure more applicants get responses to their inquiries.” This presentation also contains the validation of the dashboard.”

Data Validation

Data Validation

- Data Consistency Check
  - Data types & formats
  - Correlation between key identifiers
- Completeness Verification
  - Finding missing values
- Duplicate Detection
  - Finding duplicate entries
- Business Rule Validation
  - Connected & disconnected calls matchs total calls
  - Campaign dates with valid timeframes

Data Consistency Check

Datatype of App_ID:

SELECT pg_typeof("App_ID") AS app_id_type

FROM "CleanApplicantData"

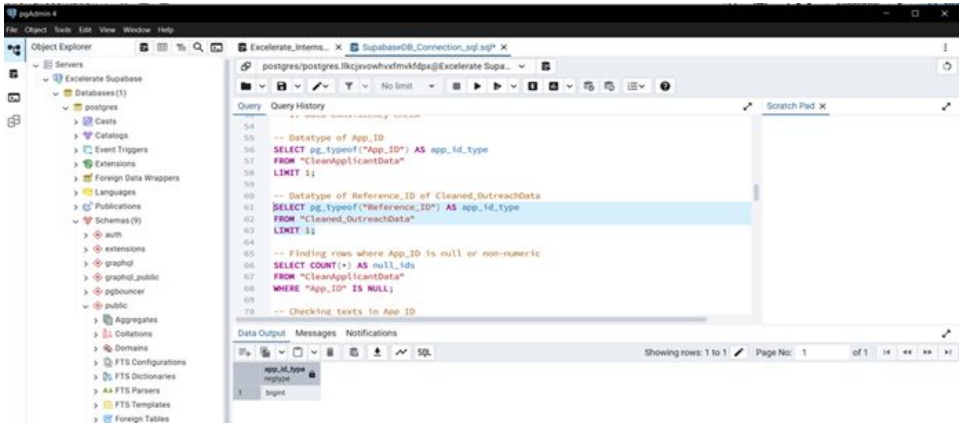
LIMIT 1;

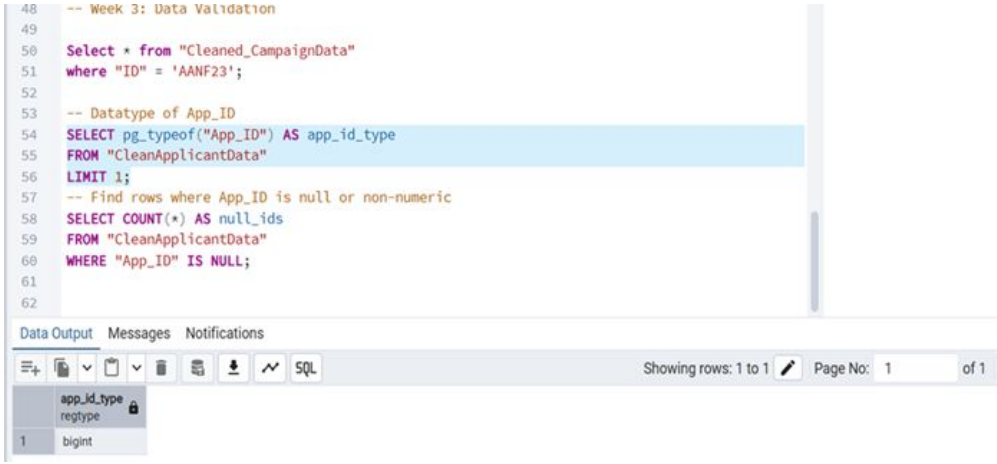
Datatype of Reference_ID of Cleaned_OutreachData:

SELECT pg_typeof("Reference_ID") AS app_id_type

FROM "Cleaned_OutreachData"

LIMIT 1;





Data Consistency Check

Finding rows where App_ID is non-numeric:

SELECT COUNT(*) AS null_ids

FROM "CleanApplicantData"

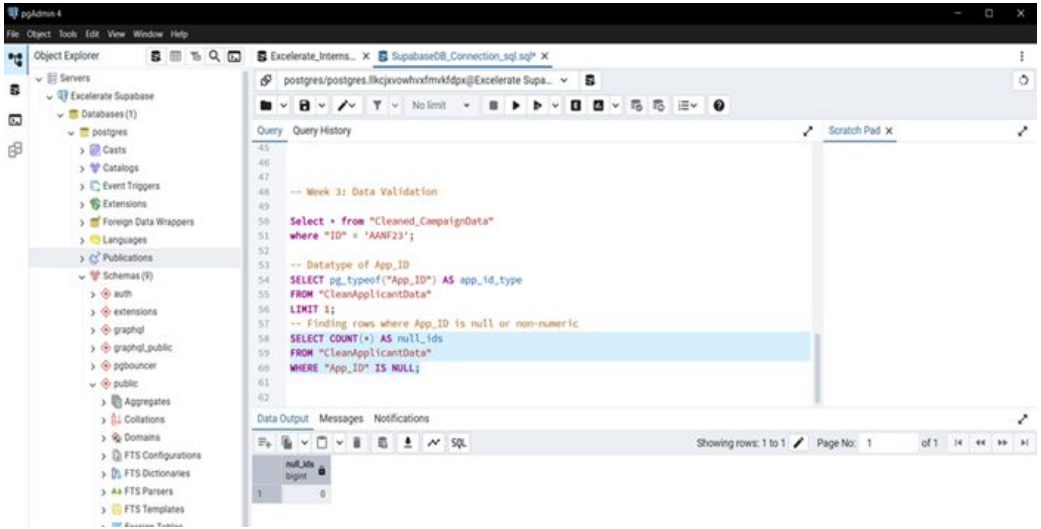
WHERE "App_ID" IS NULL;

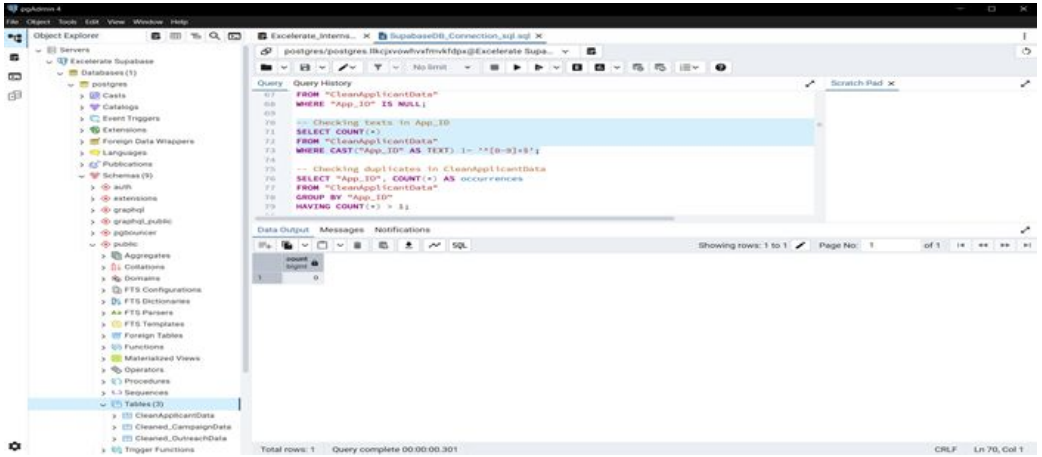
Checking texts in App_ID:

SELECT COUNT(*)

FROM "CleanApplicantData"

WHERE CAST("App_ID" AS TEXT) !~ '^[0-9]+$';





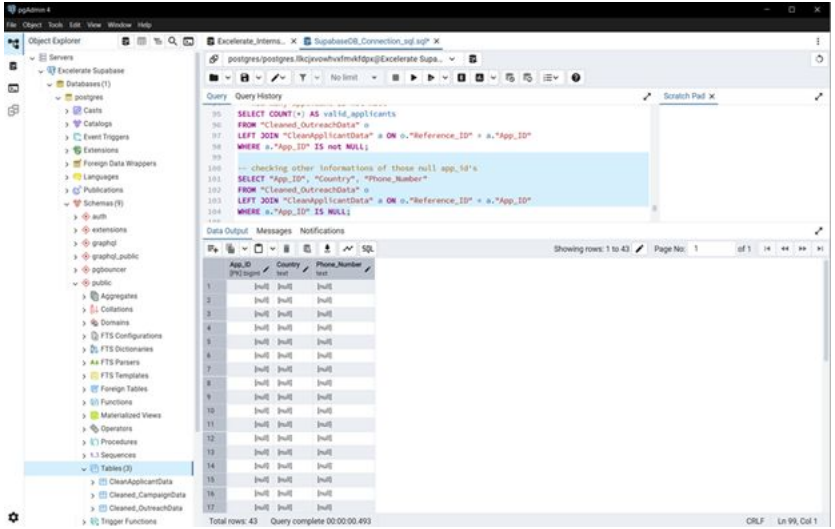
Data Consistency Check

How many applicant ID not null after joining:

```
SELECT COUNT(*) AS valid_applicants

FROM "Cleaned_OutreachData" o

LEFT JOIN "CleanApplicantData" a ON o."Reference_ID" = a."App_ID"

WHERE a."App_ID" IS not NULL;
```

Outreach Reference_ID should match ApplicantData.App_ID:

```
SELECT COUNT(*) AS missing_applicants

FROM "Cleaned_OutreachData" o

LEFT JOIN "CleanApplicantData" a ON o."Reference_ID" = a."App_ID"

WHERE a."App_ID" IS NULL;
```

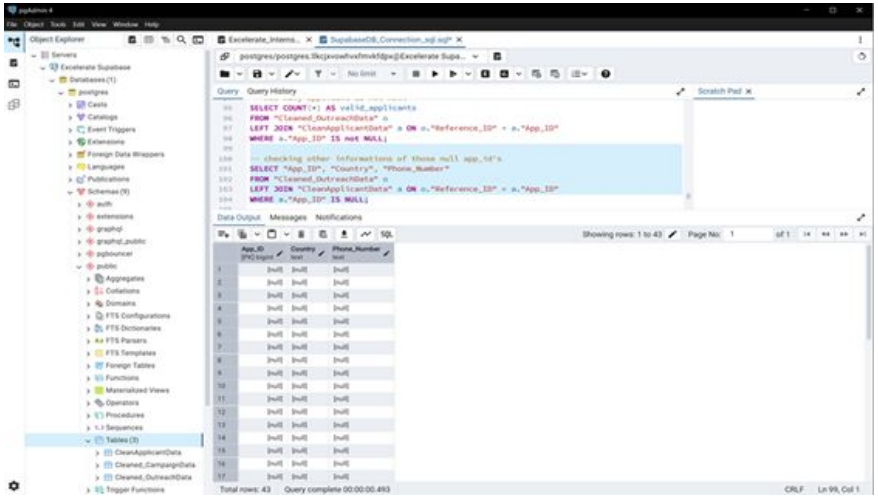


So, those null App_ID's also has null on country and other columns which shows it will not create any issue with other columns.

Data Consistency Check

checking other informations of those null app_id's:

```
SELECT "App_ID", "Country", "Phone_Number"

FROM "Cleaned_OutreachData" o

LEFT JOIN "CleanApplicantData" a ON o."Reference_ID" = a."App_ID"

WHERE a."App_ID" IS NULL;
```



1. Out of all outreach records, **42 records referenced invalid applicant IDs** that did not exist in the Applicant dataset.

2. These represent **~0.1%** of the total (42 / (33076+42)), which is statistically insignificant for overall trends.

3. For the current analysis and visualizations, these records were **excluded** to maintain data consistency.

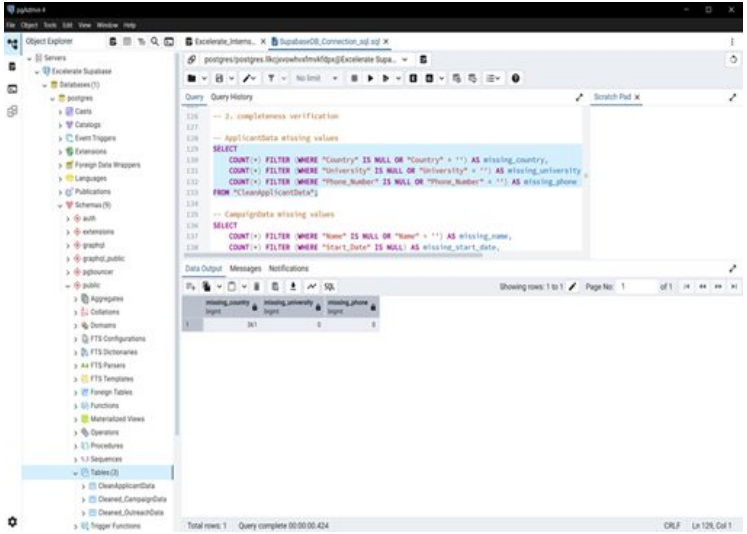
4. Future work may involve:

1.Investigating why these IDs are missing (data entry error, system issue, external applicants).
2.Either **removing them permanently** or **fixing them** if correct references can be found.

Completeness Verification

ApplicantData missing values:

```
SELECT

    COUNT(*) FILTER (WHERE "Country" IS NULL OR "Country" = '') AS missing_country,

    COUNT(*) FILTER (WHERE "University" IS NULL OR "University" = '') AS missing_university,

    COUNT(*) FILTER (WHERE "Phone_Number" IS NULL OR "Phone_Number" = '') AS missing_phone

FROM "CleanApplicantData";
```

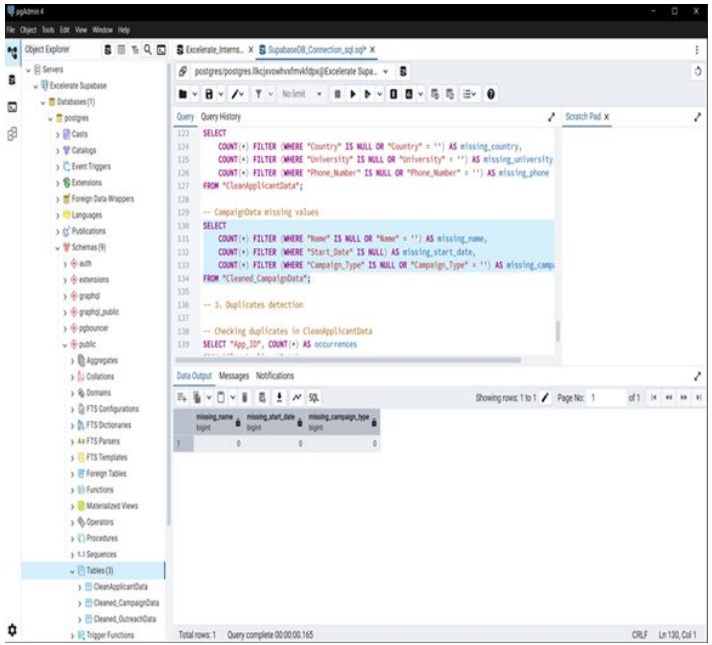


The Country column has 361 missing values, representing only 0.2% of the dataset. To account for this, a country filter was in the dashboard to ensure accurate analysis. added

Completeness Verification

CampaignData missing values:

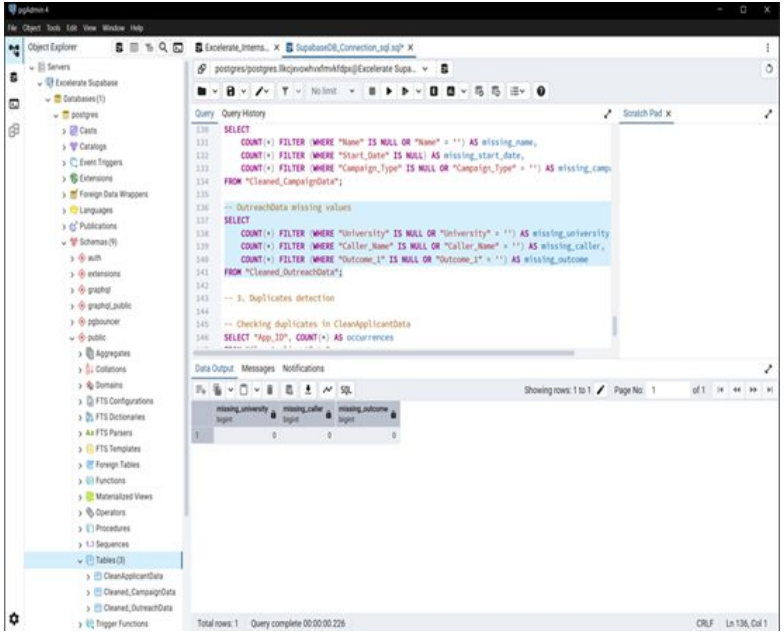
```
SELECT

    COUNT(*) FILTER (WHERE "Name" IS NULL OR "Name" = '') AS missing_name,

    COUNT(*) FILTER (WHERE "Start_Date" IS NULL) AS missing_start_date,

    COUNT(*) FILTER (WHERE "Campaign_Type" IS NULL OR "Campaign_Type" = '') AS missing_campaign_type

FROM "Cleaned_CampaignData";
```



Completeness Verification

OutreachData missing values

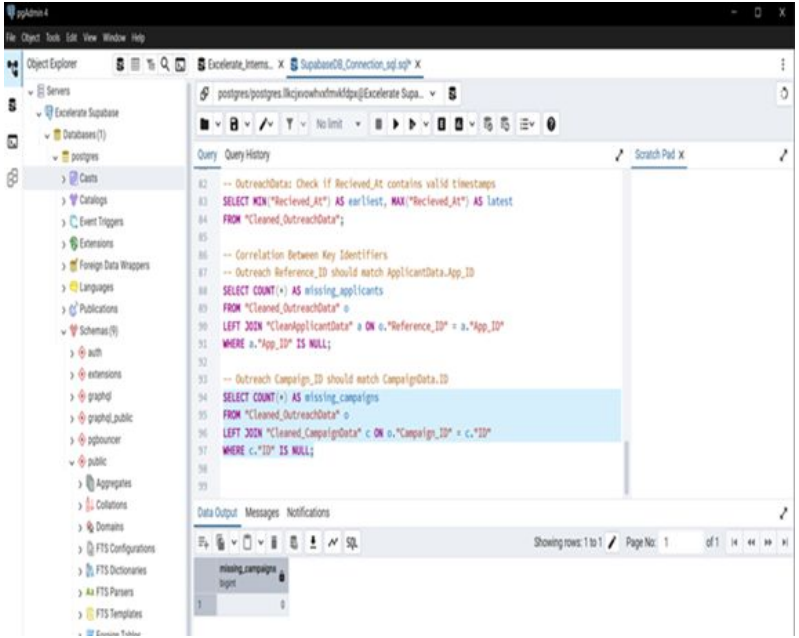
```
SELECT

        COUNT(*) FILTER (WHERE "University" IS NULL OR "University"
= '') AS missing_university,

        COUNT(*) FILTER (WHERE "Caller_Name" IS NULL OR
"Caller_Name" = '') AS missing_caller,

        COUNT(*) FILTER (WHERE "Outcome_1" IS NULL OR
"Outcome_1" = '') AS missing_outcome

FROM "Cleaned_OutreachData";
```



Completeness Verification

Outreach Campaign_ID should match CampaignData.ID:

```
SELECT COUNT(*) AS missing_campaigns

FROM "Cleaned_OutreachData" o

LEFT JOIN "Cleaned_CampaignData" c ON o."Campaign_ID" = c."ID"

WHERE c."ID" IS NULL;
```



Duplicate Detection

Checking duplicates in CleanApplicantData:

SELECT "App_ID", COUNT(*) AS occurrences

FROM "CleanApplicantData"

GROUP BY "App_ID"

HAVING COUNT(*) > 1;

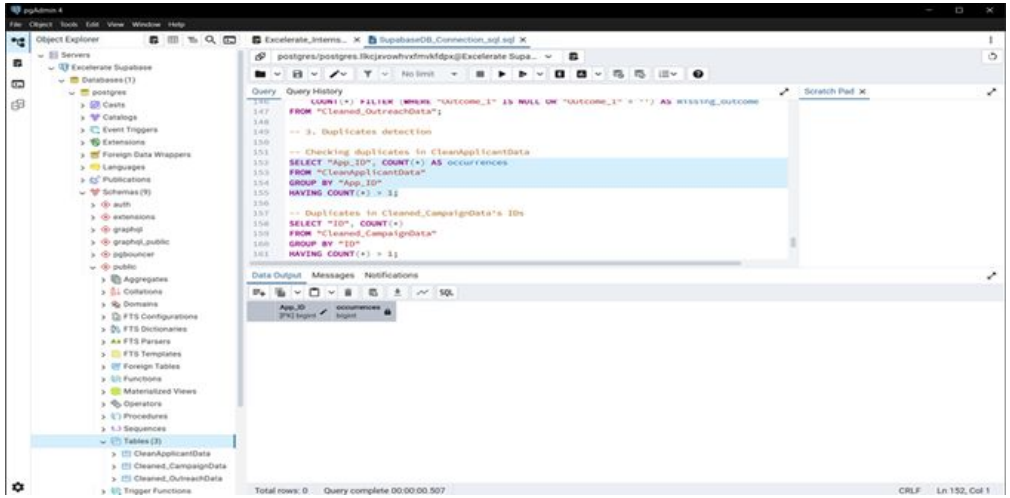
Duplicates in Cleaned_CampaignData's IDs:

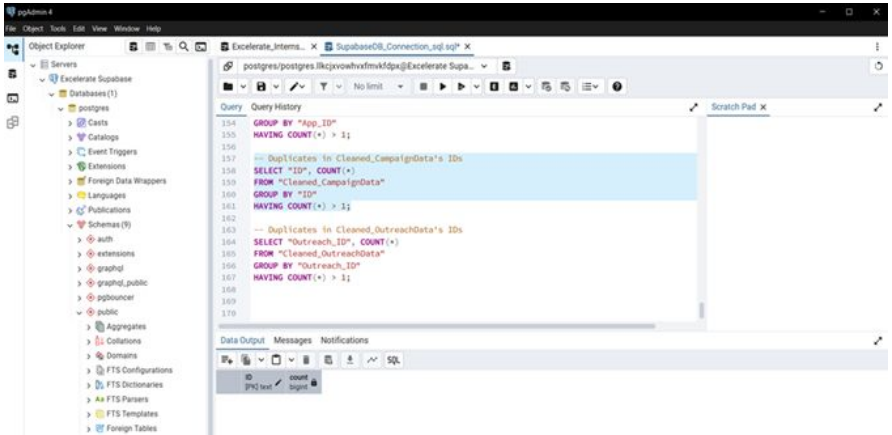
SELECT "ID", COUNT(*)

FROM "Cleaned_CampaignData"

GROUP BY "ID"

HAVING COUNT(*) > 1;





Duplicate Detection

Duplicates in Cleaned_OutreachData's IDs:

SELECT "Reference_ID", COUNT(*) as total_Duplicates

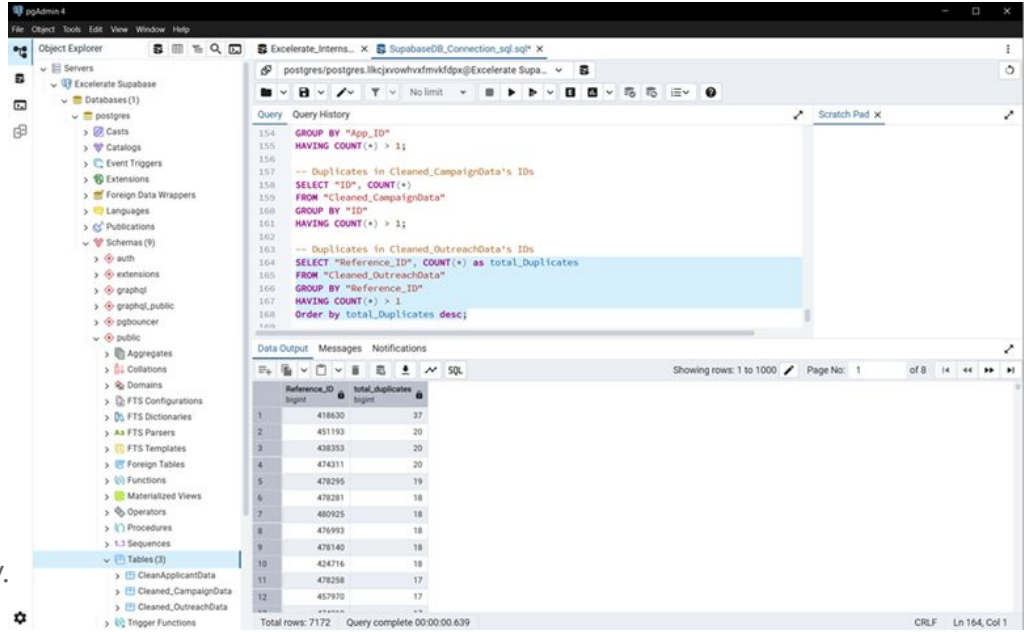
FROM "Cleaned_OutreachData"

GROUP BY "Reference_ID"

HAVING COUNT(*) > 1

Order by total_Duplicates desc;

Here, duplicates is must as “reference_Id” is foreign key.



Duplicate Detection

Total calls including duplicates

SELECT COUNT(*) AS total_calls FROM "Cleaned_OutreachData";

Unique calls (ignoring duplicates based on key fields)

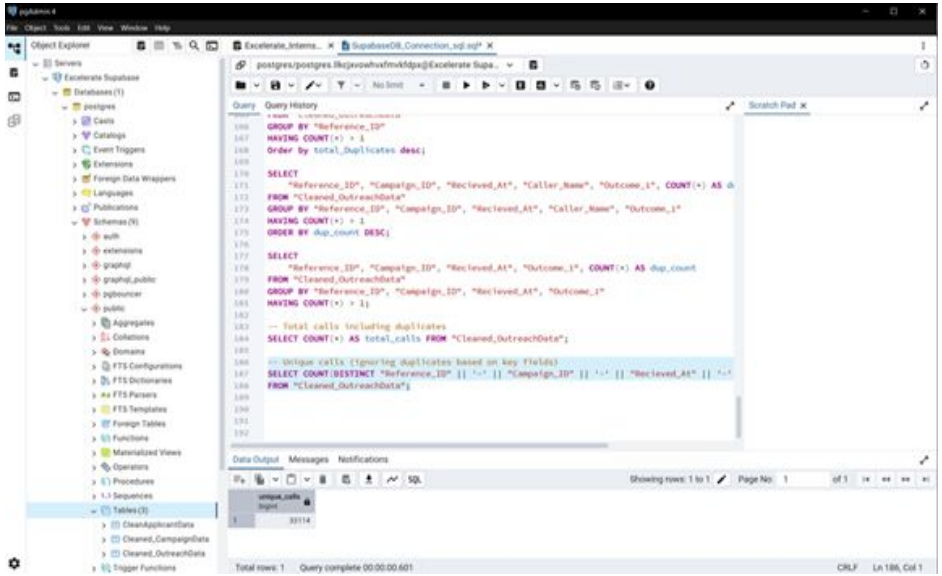
SELECT COUNT(DISTINCT "Reference_ID" || '-' || "Campaign_ID" || '-' ||

"Recieved_At" || '-' || "Outcome_1") AS unique_calls

FROM "Cleaned_OutreachData";

Total calls = 33119

Unique call = 33114



During data validation, we found 5 duplicate call records out of 33,119 total calls, representing approximately 0.015% of the dataset.

This minimal duplication has an insignificant impact on the overall analysis and visualizations. Therefore, these duplicates were not removed for the current analysis, but they can be cleaned in future iterations for complete data integrity.

Business Rule Validation

Statistical Record of Total Calls:

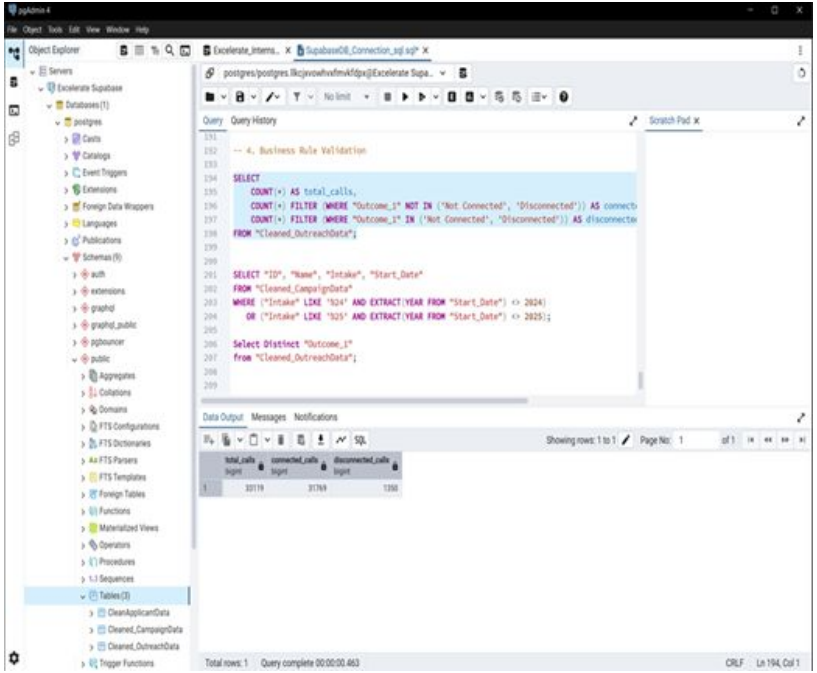
```
SELECT

        COUNT(*) AS total_calls,

        COUNT(*) FILTER (WHERE "Outcome_1" NOT IN ('Not Connected', 'Disconnected')) AS
connected_calls,

        COUNT(*) FILTER (WHERE "Outcome_1" IN ('Not Connected', 'Disconnected')) AS
disconnected_calls

FROM "Cleaned_OutreachData";
```

Here, Connected calls = 31769, Disconnected calls = 1350, and Total calls = 33119

For business rule validation, we classified all call outcomes as either Connected or Disconnected.

‘Not Connected’ and ‘Disconnected’ outcomes were grouped as Disconnected, while all other outcomes were considered Connected.

This allowed us to validate total calls against connected and disconnected counts for consistency



Business Rule Validation

Statistical Record of Total Calls:

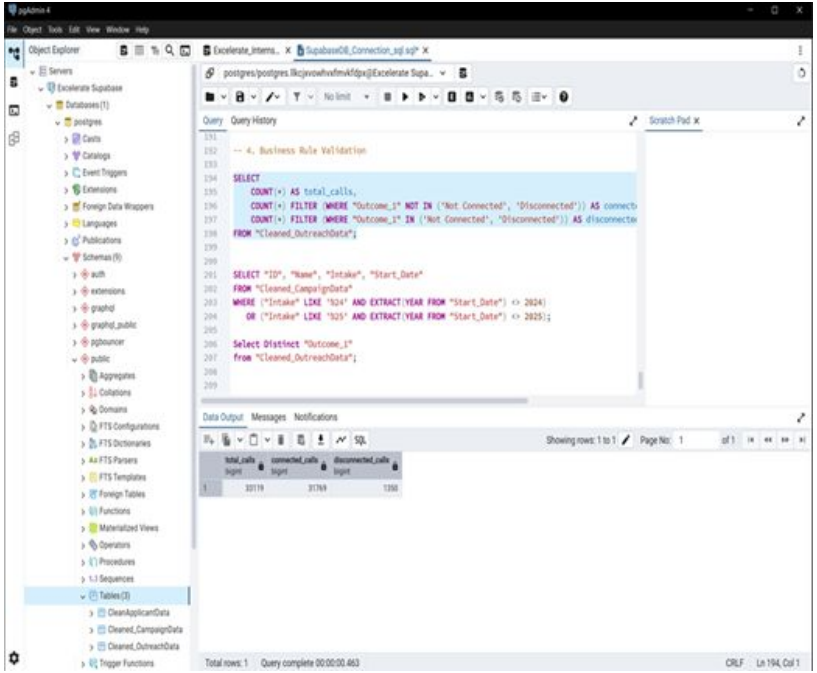
```
SELECT

        COUNT(*) AS total_calls,

        COUNT(*) FILTER (WHERE "Outcome_1" NOT IN ('Not Connected', 'Disconnected')) AS
connected_calls,

        COUNT(*) FILTER (WHERE "Outcome_1" IN ('Not Connected', 'Disconnected')) AS
disconnected_calls

FROM "Cleaned_OutreachData";
```

Here, Connected calls = 31769, Disconnected calls = 1350, and Total calls = 33119

For business rule validation, we classified all call outcomes as either Connected or Disconnected.

‘Not Connected’ and ‘Disconnected’ outcomes were grouped as Disconnected, while all other outcomes were considered Connected.

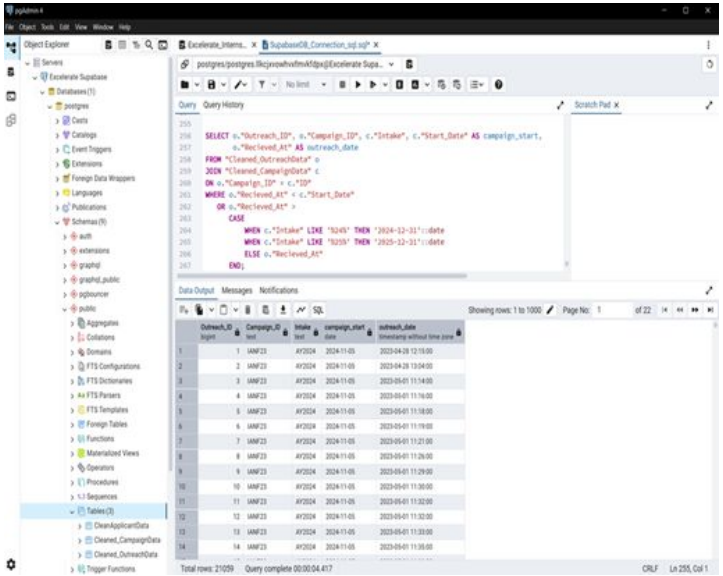
This allowed us to validate total calls against connected and disconnected counts for consistency



Business Rule Validation

Validation of Campaign and Outreach Timeframes:

```
SELECT o."Outreach_ID", o."Campaign_ID", c."Intake", c."Start_Date" AS campaign_start,
        o."Recieved_At" AS outreach_date
FROM "Cleaned_OutreachData" o
JOIN "Cleaned_CampaignData" c
ON o."Campaign_ID" = c."ID"
WHERE o."Recieved_At" < c."Start_Date"
  OR o."Recieved_At" >
        CASE
        WHEN c."Intake" LIKE '%24%' THEN '2024-12-31'::date
        WHEN c."Intake" LIKE '%25%' THEN '2025-12-31'::date
        ELSE o."Recieved_At"
        END;
```



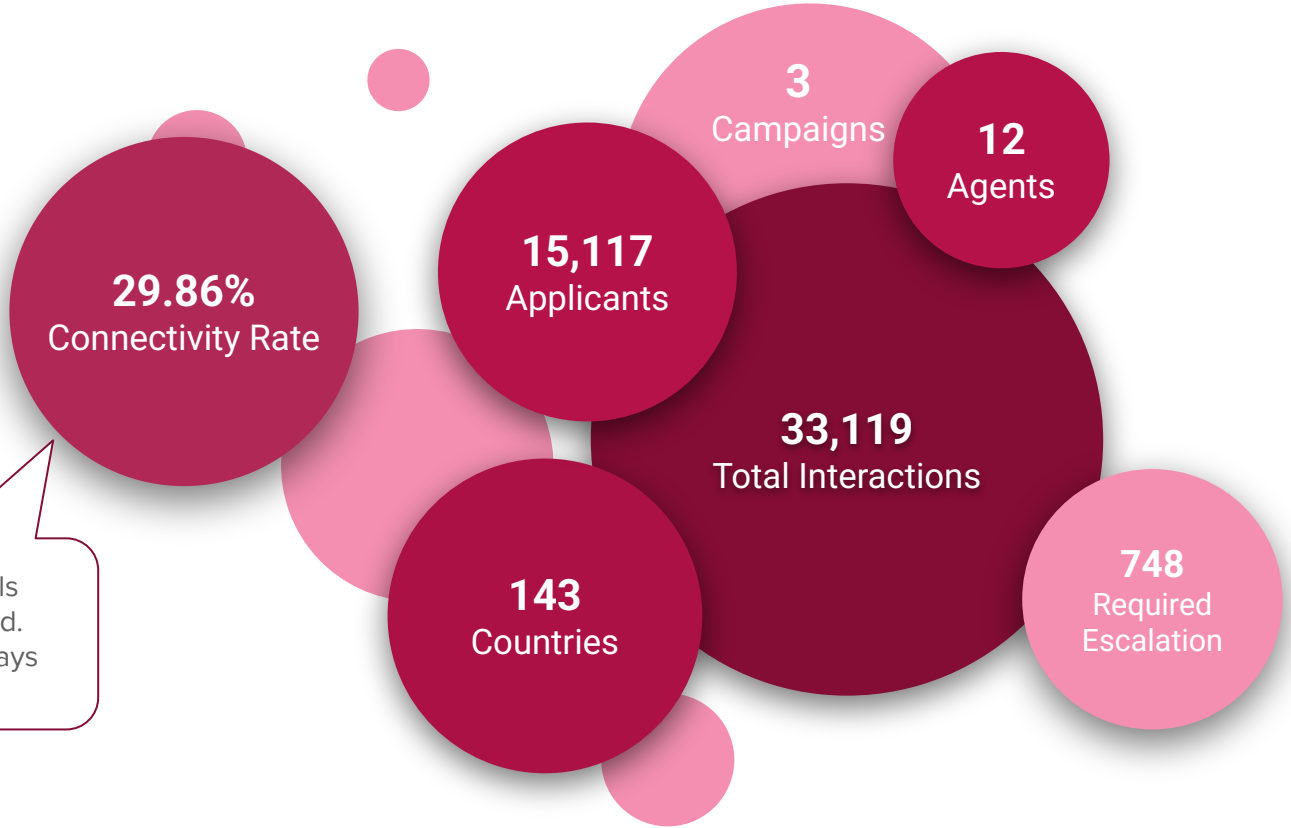
Campaigns for intakes FA24 and SP25 were planned for 2024 and 2025, respectively. Outreach and marketing activities began earlier, with some Start_Dates in 2023, which is consistent with standard marketing practice and does not indicate a data error.
Business rule validation confirms that no campaign Start_Date exceeds its intake year range.

What we are trying to find?

- Connectivity rate and general overlook of data
- Distribution of call outcomes
- Analyse agent
- Region distribution
- Recommendations to boost performance

Overview & Key Metrics

Summary of Data - KPIs

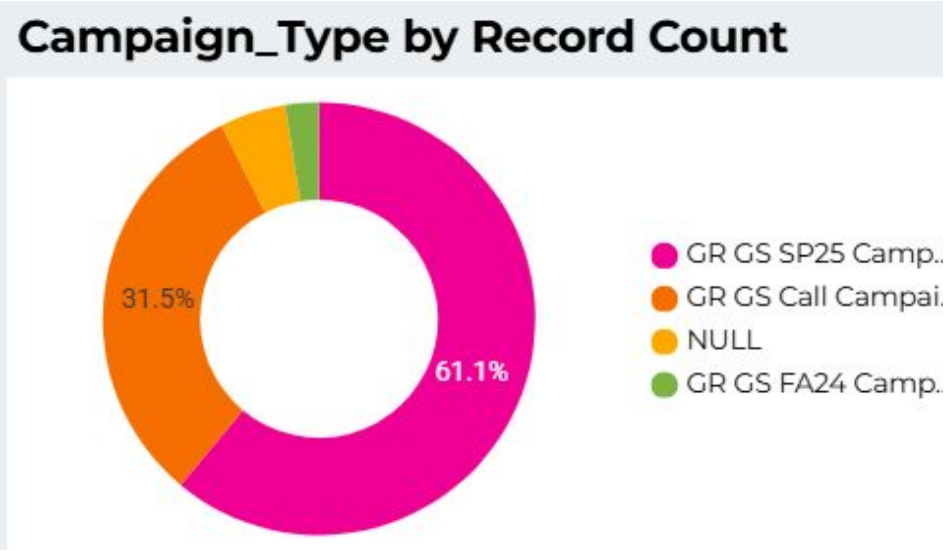


Majority of the calls were disconnected. We will explore ways of reducing this.

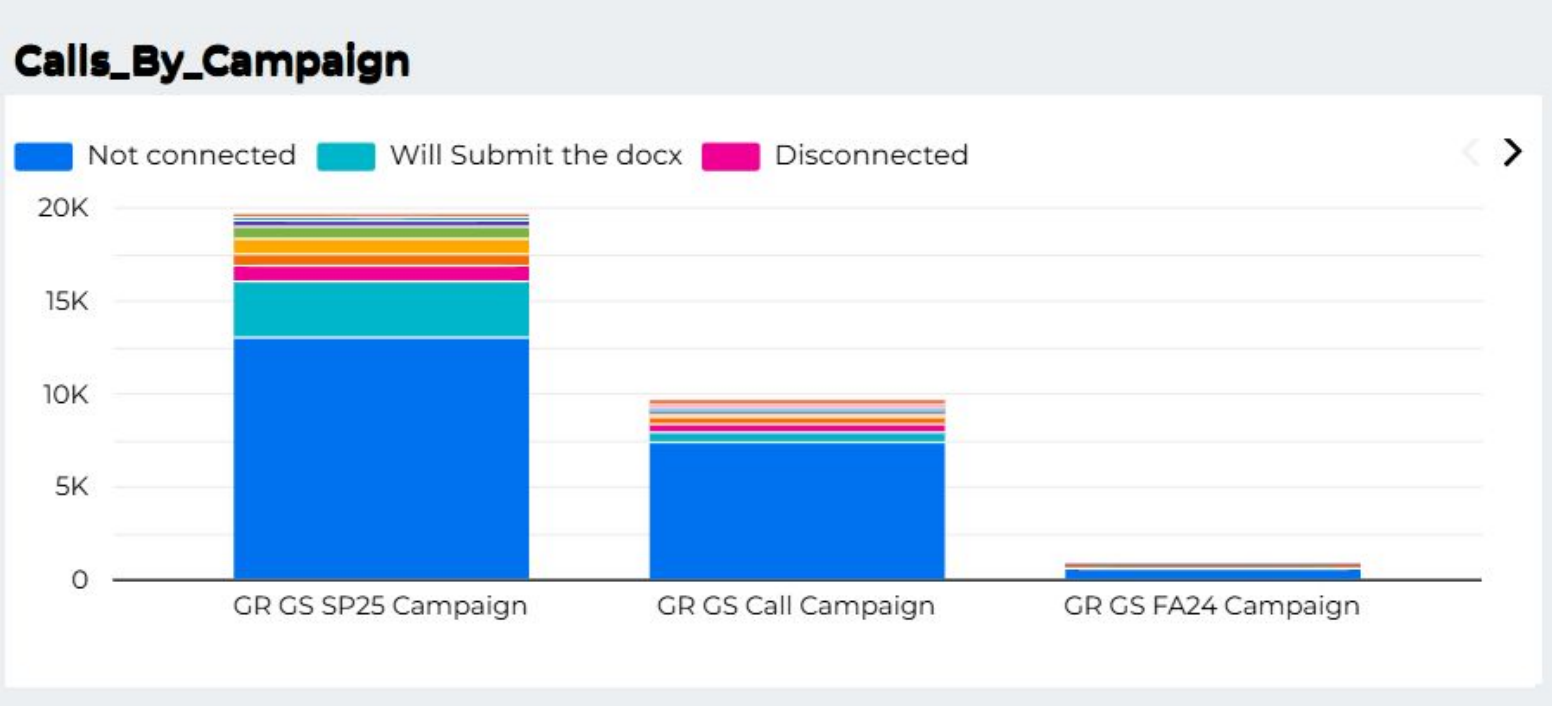
Campaign Types

Overview of Campaign Volume

“The GR GS SP25 campaign accounted for most of the calls and was the most successful in attracting applicants. However, some calls were not attributed to any known campaign, which shows a gap in campaign tracking.”



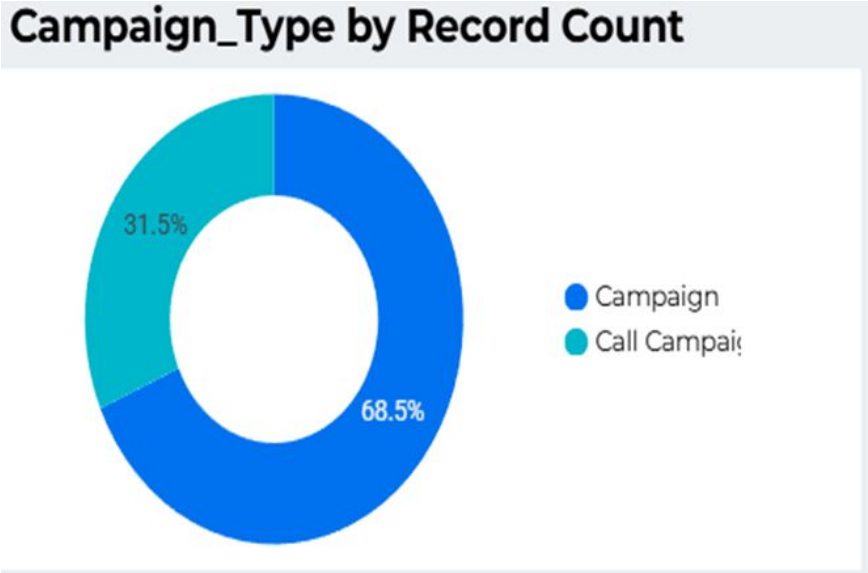
Call by Campaign



“Looking at call outcomes, we found that across all campaigns, the majority of calls were not connected. This highlights a major challenge in outreach effectiveness.”

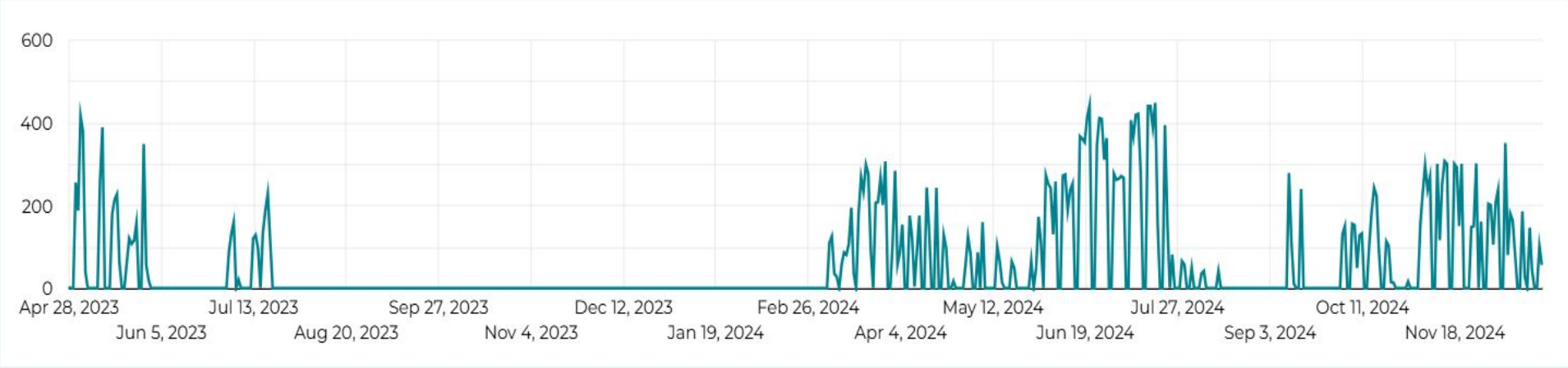
Campaign Type by Record Count

“When comparing campaign types, we see that Call Campaigns had about 31% of applicants, while Normal Campaigns had 69%. This indicates that applicants prefer engaging more during Normal Campaigns.”



Calls Over Time

Calls Over Time



The spring season is the most season active.

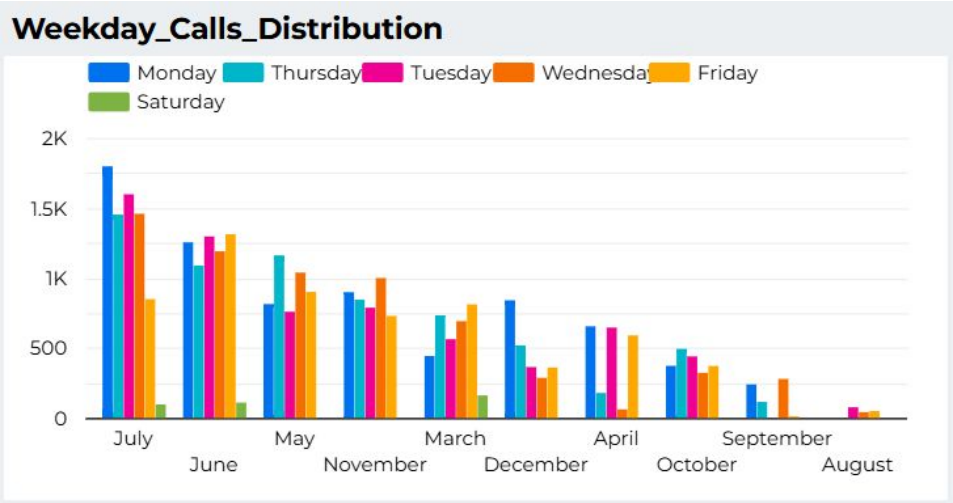
Overall there is significant call activity in the months of February to July and from September to November.

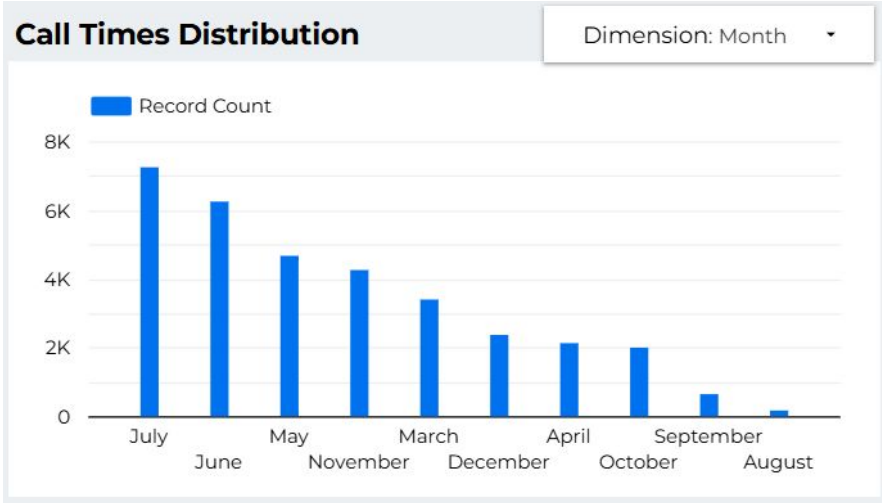
The call activity correlates with the campaign dates.

Weekday Calls Distribution

Call Times Distribution

Similarly this chart shows that June and July are the most active months in terms of calls. August has the least interactions.



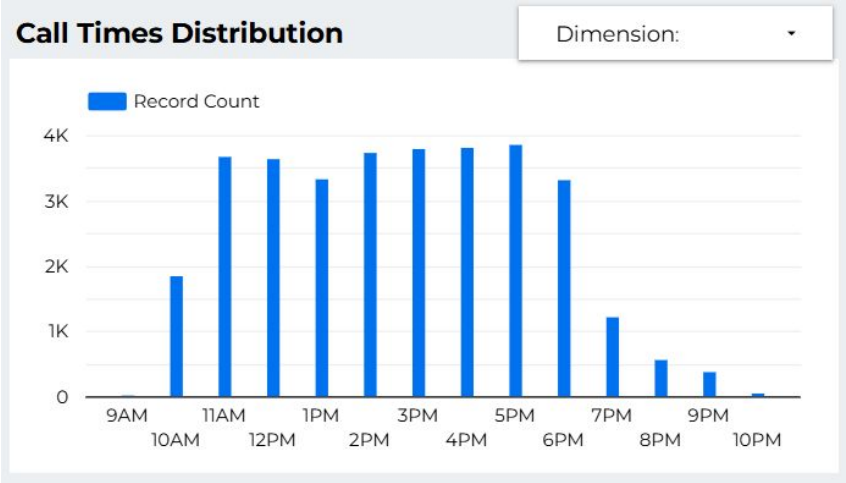


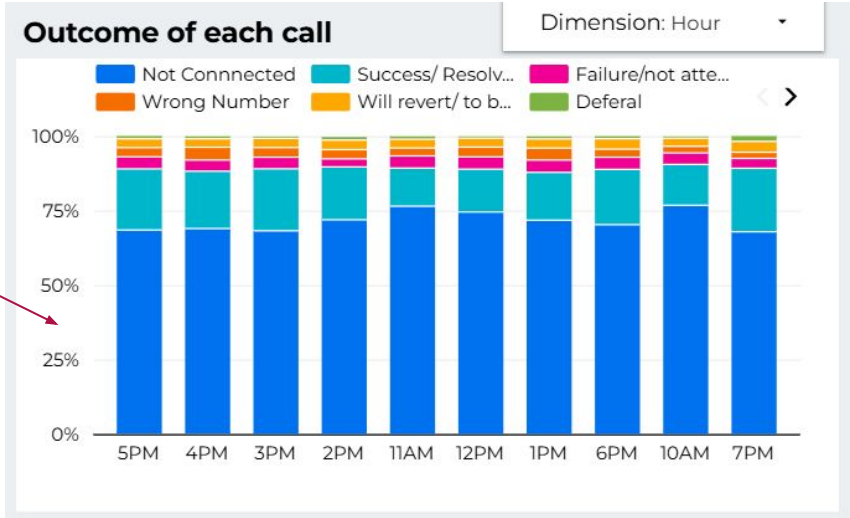
Call Times Distribution

Calls come in from 10 AM to 10PM.

Highest Call activity is between 11AM and 6PM.

The outcome distribution varies slightly over call hours. However more calls are unconnected between 10AM and 12 PM.

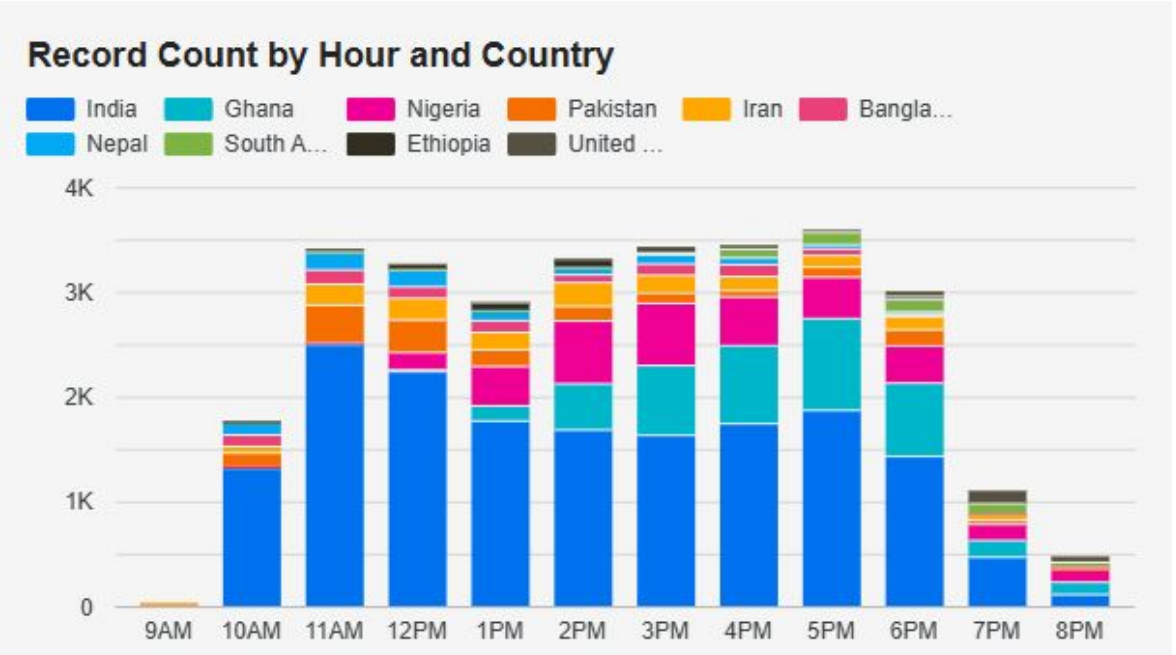




Call times by Country

This shows the variation in call activity per country for each hour.

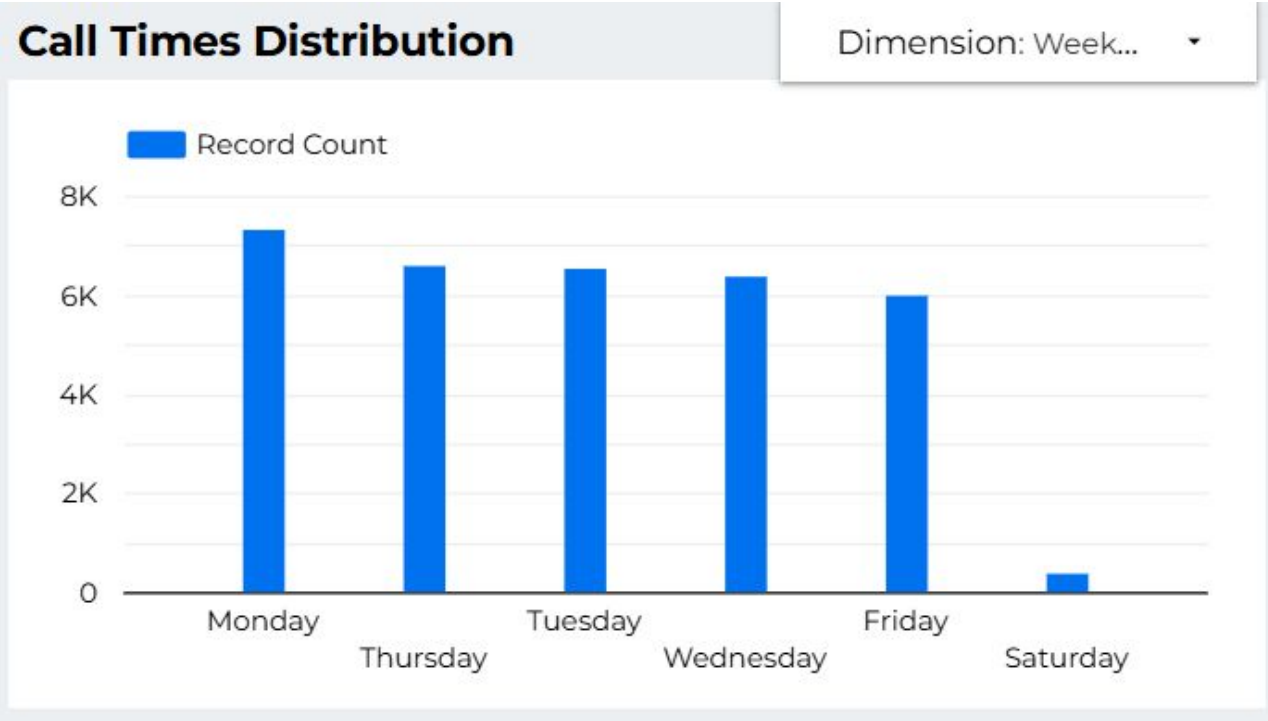
Leaning into this trend can help agents optimize call hours for the various time zones to maximize call success.



Daily call distribution

Weekdays have similar call activity, however Monday has the most.

Weekends have very little activity



Agent Performance

Analyse agent performance by Call Volume, Connectivity rate and Success rate.

Overview of Method

In order to analyse call outcomes, the results were standardized by grouping them into a smaller number of categories:

- Not Connected (Disconnected, voicemail, not connected)
- Success/resolved
- Failure/ not attending
- Wrong Number
- Will revert/to be solved
- Deferral
- Processing Visa

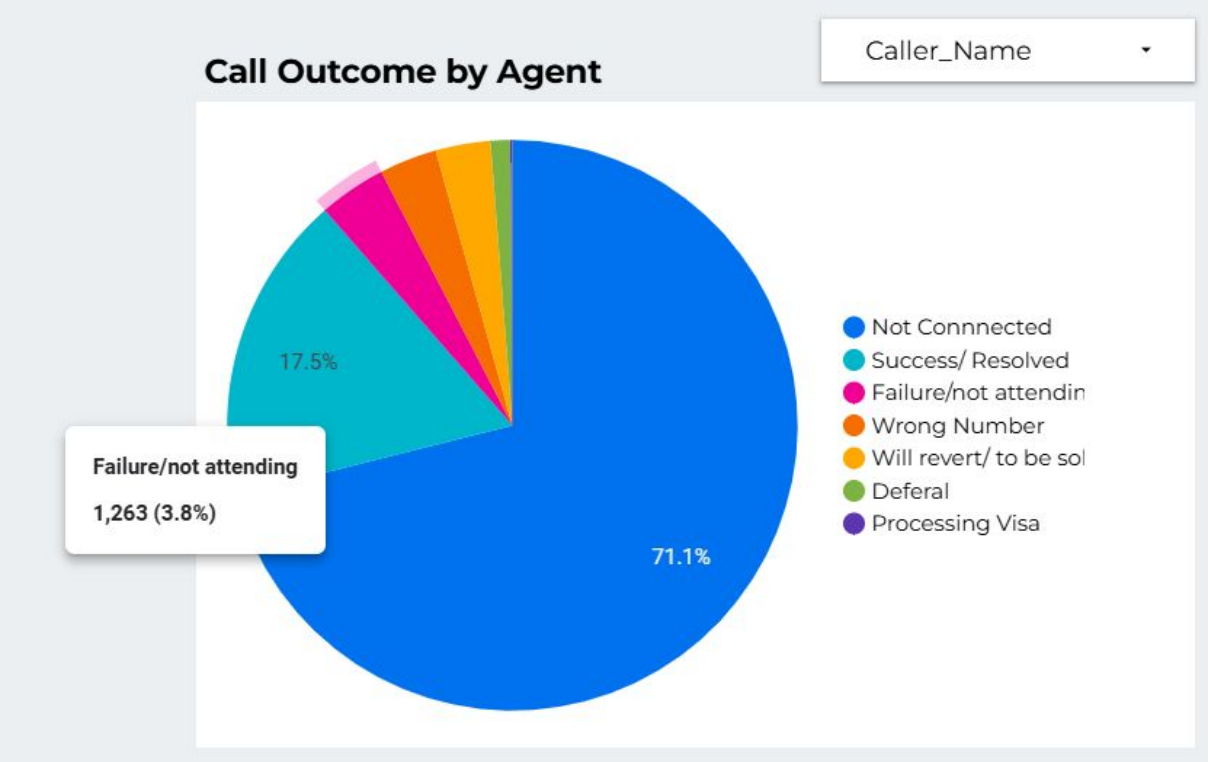
To evaluate agent performance, the tested interactive graph shows the distribution of call outcomes for each agent.

Success/resolved outcome includes an assurance of future university attendance, connectivity and resolved inquired .

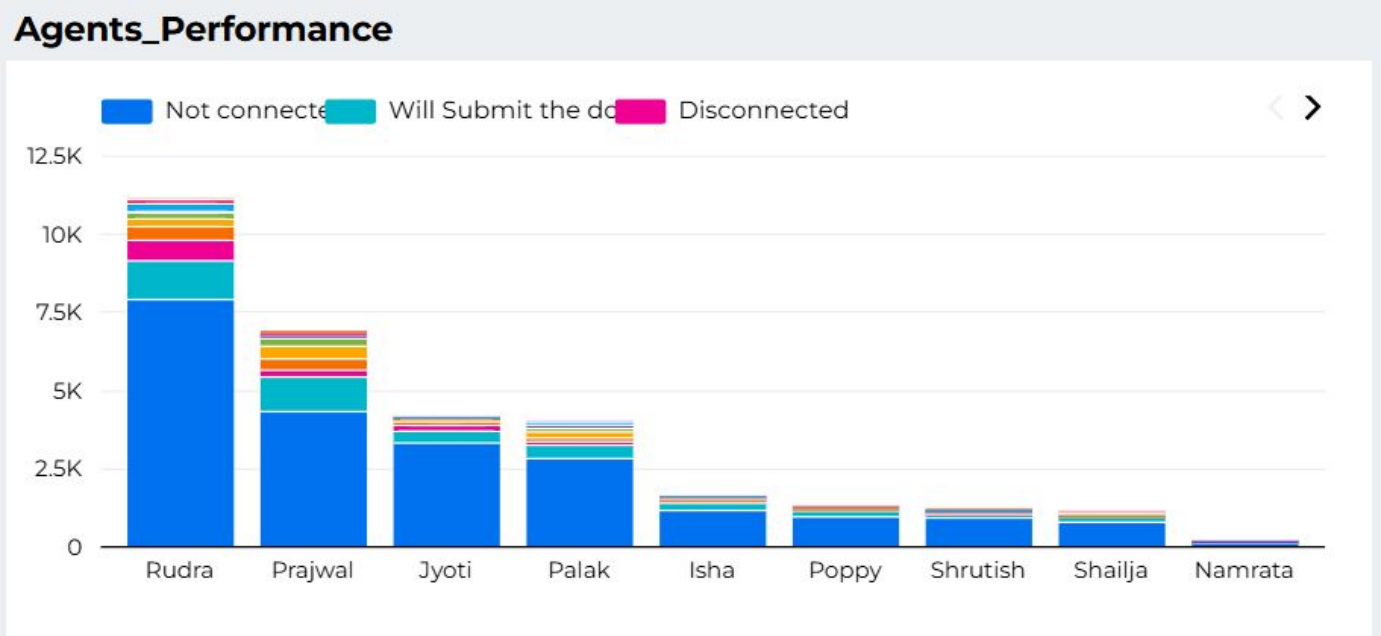
Overview of Call Outcomes

71.1% of calls were didn't connected.

3.8% of calls were about inability to attend the university or loss of interest.



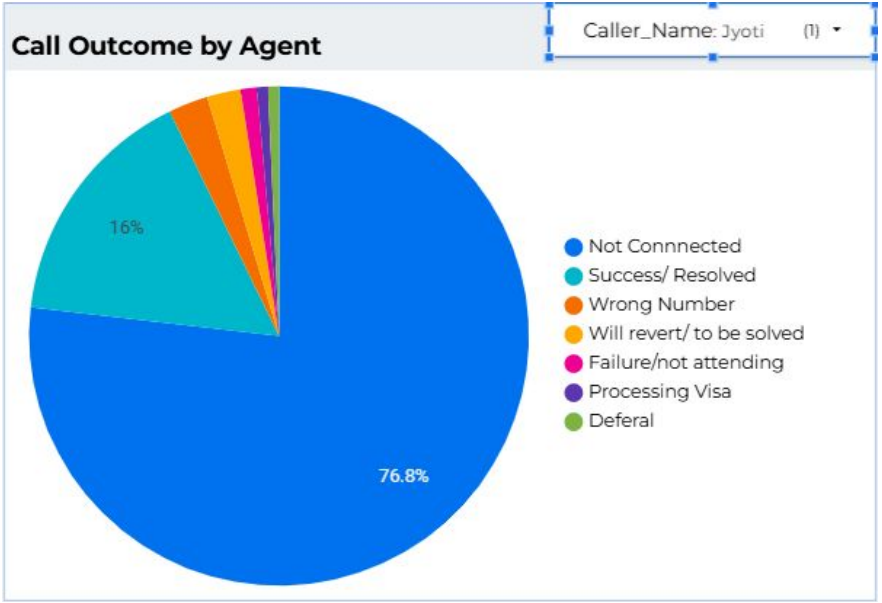
Agent Performance by volume of calls



Rudra serves the most calls while Namrata has the least.

Rudra had more positive communications with document submission

Agent performance by Proportion of Disconnected Calls



Among agents who serviced over 200 calls, Prajwal has the least percentage of unconnected calls, whiles Jyoti has the most proportion of unconnected calls.

Agent Performance by Success rate

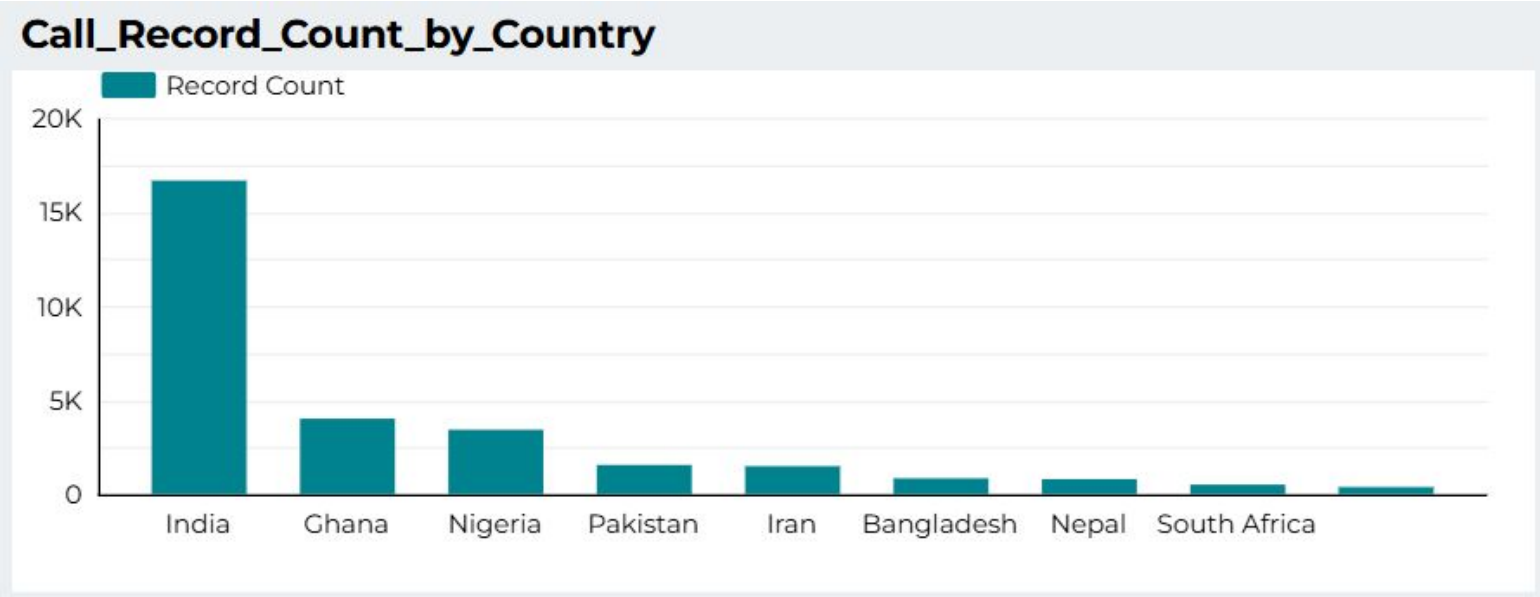
Shalija had the highest success rate for agents over 200 calls.

Also, had lowest proportion of wrong number + not connected outcomes.

Success is denoted by a assurance of future university attendance (and connectivity).

Region Distribution

Country Distribution



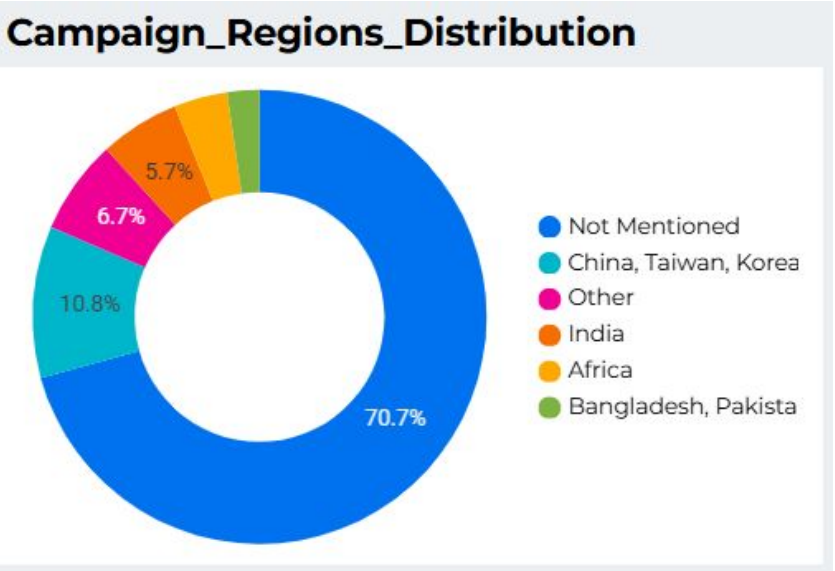
India, Ghana and Nigeria had the highest number of applicants in Illinois Institute of Technology's marketing campaigns

Agents should be well aware of the needs of students from these countries.

Campaign Regions

Normal campaign (70%) does not have region informations. And that is why more than 70% has “not mentioned” in region column.

On the other hand, 30% of the call campaign have region information which is why 11% (China, Taiwan, Korea), 7% other and 6% India are found where these campaigns were held.



Summary and Recommendations

We will look at the key takeaways, and explore recommendations to improve campaign success and data collection.

What are the key takeaways?

- Only 29.86% of the calls were connected. While 70.14% failed.
- GR GS SP25 generated the highest call volume.
- 31.5% of calls were not attributed to any campaign.
- Spring period had highest monthly call volume.
- The Fall period also has significant call activity.
- Volume drops around May.
- 11AM to 6AM has the highest call volume in a day.
- More calls are unconnected between 10AM and 12 PM.
- Rudra handled the most calls. Namrata handled fewest calls.
- Shalija had the highest success rate.
- India accounted for majority of the calls

What are the key takeaways? - Validation

Findings

- June and July are the most busy period of the year in this marketing campaign.
- 42 outreach records had no matching applicant IDs (orphan records).
- Duplicate outreach entries inflated total calls by 5 records (0.01%).
- Connected + Disconnected calls = Total calls (business rule validated).
- Campaign Start_Dates began in 2023 for FA24 and 2024 for SP25, consistent with marketing practice.
- No campaign exceeded its intake year timeframe.

Recommendations

- Improve connectivity rates
  - Adjust agent schedules & methods to account for peak hours with lower connectivity rates (10am to 12pm)
- Enhance Campaign Tracking to reduce null or uncategorized calls.
  - Include categorizes for miscellaneous calls for further insight in the future.
  - More effort should be put in to ensure that the calls are accurately categorized. Also, if needed more categorizes should be included to capture the miscellaneous calls.
- Agent performance management
  - Reward agents with high success rates to extend this outcome amongst more agents.
- Optimize call timings for different countries.
- Ensure agents are familiar with the needs of Indian, Ghanaian, Nigerian and Pakistan applicants.
- More resources can be allocated to the spring season as it seems popular with applicant. It has a high ROI.

Recommendations to handle high volume seasons

The busiest times are May, June July, and November.

Part time workers can be hired to work during periods of high international demand. To maximize returns, part time workers' ideal work days on Mondays, Tuesdays and Thursdays in June, July and November.

Specifically trained additional agents should be assigned to India and West Africa.

Also we can pivot working hours to meet varying international demand.

Recommendations to boost Performance

- Provide bonuses or other commissions to agents which will improve productivity.
- Continue early marketing since 2023 outreach supported 2024 intakes.
- Focus campaigns in high-engagement regions for better ROI.
- Engage more with Indian applicants as they are showing more interest.

Future Work

- Address missing applicant data through better collection methods.
- Remove duplicates before reporting to avoid inflated metrics.
- Improve applicant linkage for outreach to reduce orphan records.
- Removing data of those 43 Reference_ID’s which is connected with 43 null App_IDs.