



DATA VISUALIZATION AND ASSOCIATE

Team 6 | 0704 SLU
DVA

Our Team Members



Ushna Adnan



Salman Rahobar



Mary Olurinola



Ritik Mathpal



Divya P



Muna Saha



Shoaib Salman



Syed Rayyan

Project Insights

- **Key Processes:**
 - Data cleaning & standardization
 - Exploratory Data Analysis (EDA)
 - Data validation & integration using PostgreSQL
- **Outcomes:**
 - Identification of trends and data quality issues
 - Development of an interactive dashboard showcasing
 - Enrollment and participation trends
 - Advertising campaign performance
 - Geographic distribution insights
- **Purpose:**
 - Enhance data-driven decision-making
 - Provide a clear visual narrative of our analytical journey



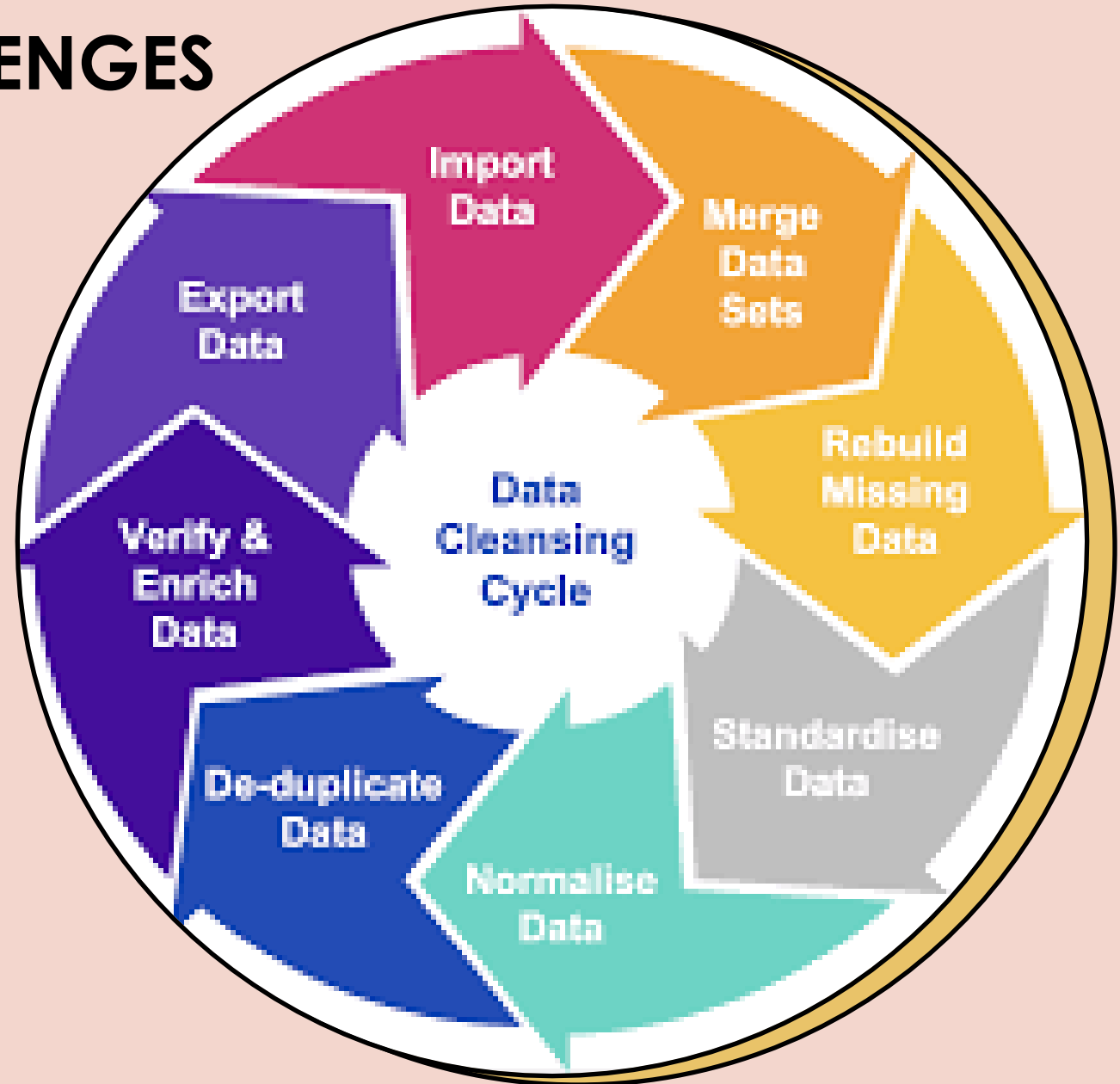
DATA CLEANING CHALLENGES

1. Missing Values

- **Cohort Size:** Missing entries imputed using median.
- **Gender:** Blank or undefined values removed.
- **Delivery Status (Marketing Data):** Contained blank entries.
- **Tracking Questions (Opportunities):** Incomplete or missing fields.

2. Inconsistent Entries

- **Status Columns** (Learner, Learner Opportunity): Included invalid or negative values.
- **Opportunity Code:** Contained duplicate or invalid codes.
- **Learner ID & Enrollment ID:** Had mismatches or redundant prefixes affecting joins.



3. Formatting Issues

- **Country, City, State:** Required standard casing and removal of extra info.
- **Opportunity & Campaign Names:** Extra spaces and inconsistent casing.
- **Email:** Needed trimming and conversion to lowercase.
- **Date Columns:** Non-standard formats across datasets (Apply Date, Birthdate, Start Dates).

4. Unreliable Data

- **Duplicate IDs:** Enrollment ID and Opportunity Code needed uniqueness validation.
- **Non-numeric Results (Marketing):** Required type conversion.
- **Currency Format (Amount Spent):** Inconsistently formatted across records.



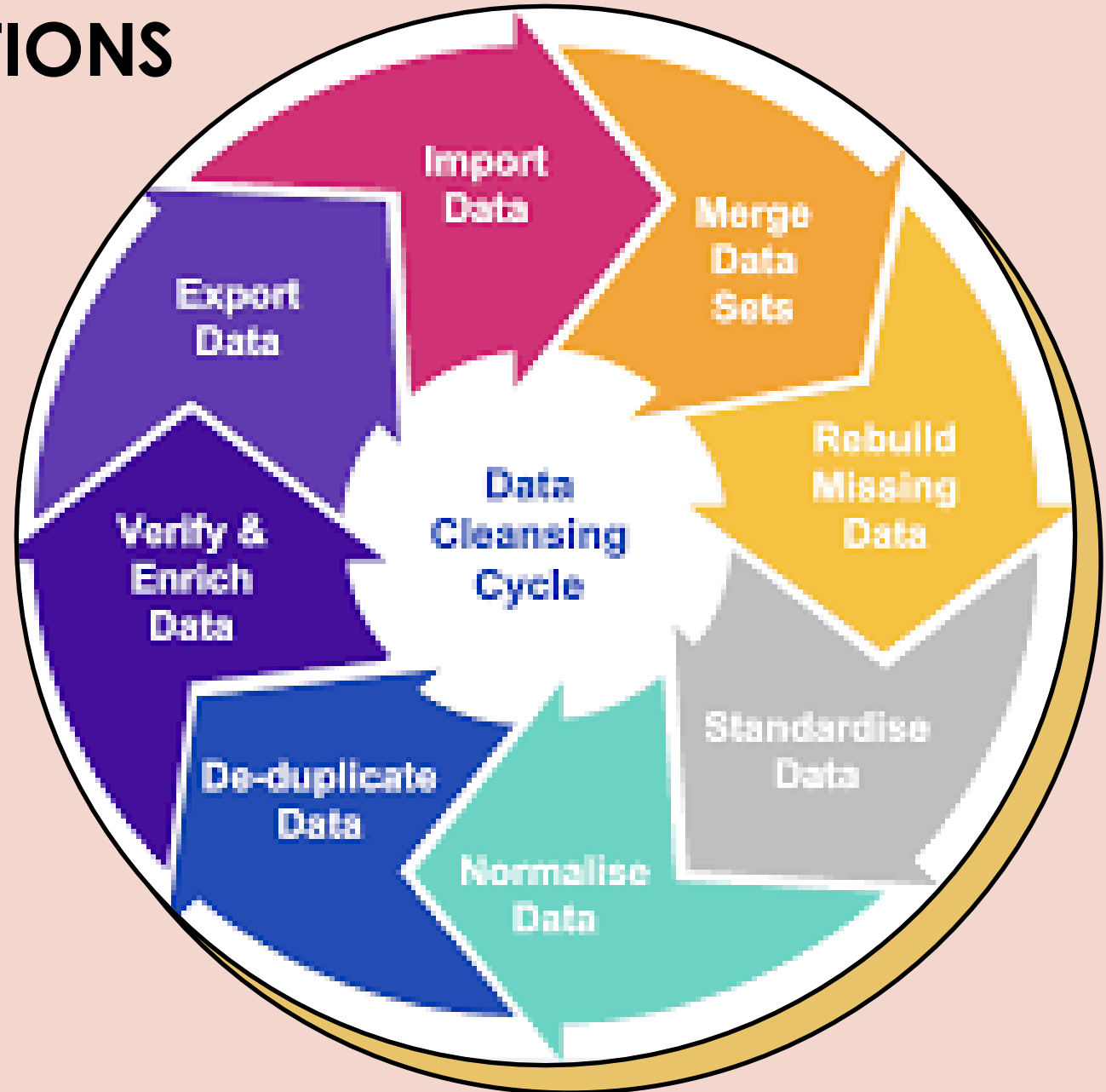
DATA CLEANING SOLUTIONS

1. Addressing Missing Values

- ✓ Imputed missing Cohort Size using median values.
- ✓ Removed blanks in Gender and Delivery Status.
- ✓ Handled nulls in tracking-related fields appropriately.

2. Resolving Inconsistent Entries

- ✓ Standardized all Status fields to remove invalid or negative values.
- ✓ Validated and de-duplicated IDs (Enrollment ID, Opportunity Code).
- ✓ Aligned Learner ID formats across datasets for consistent joins.



3. Fixing Formatting Issues

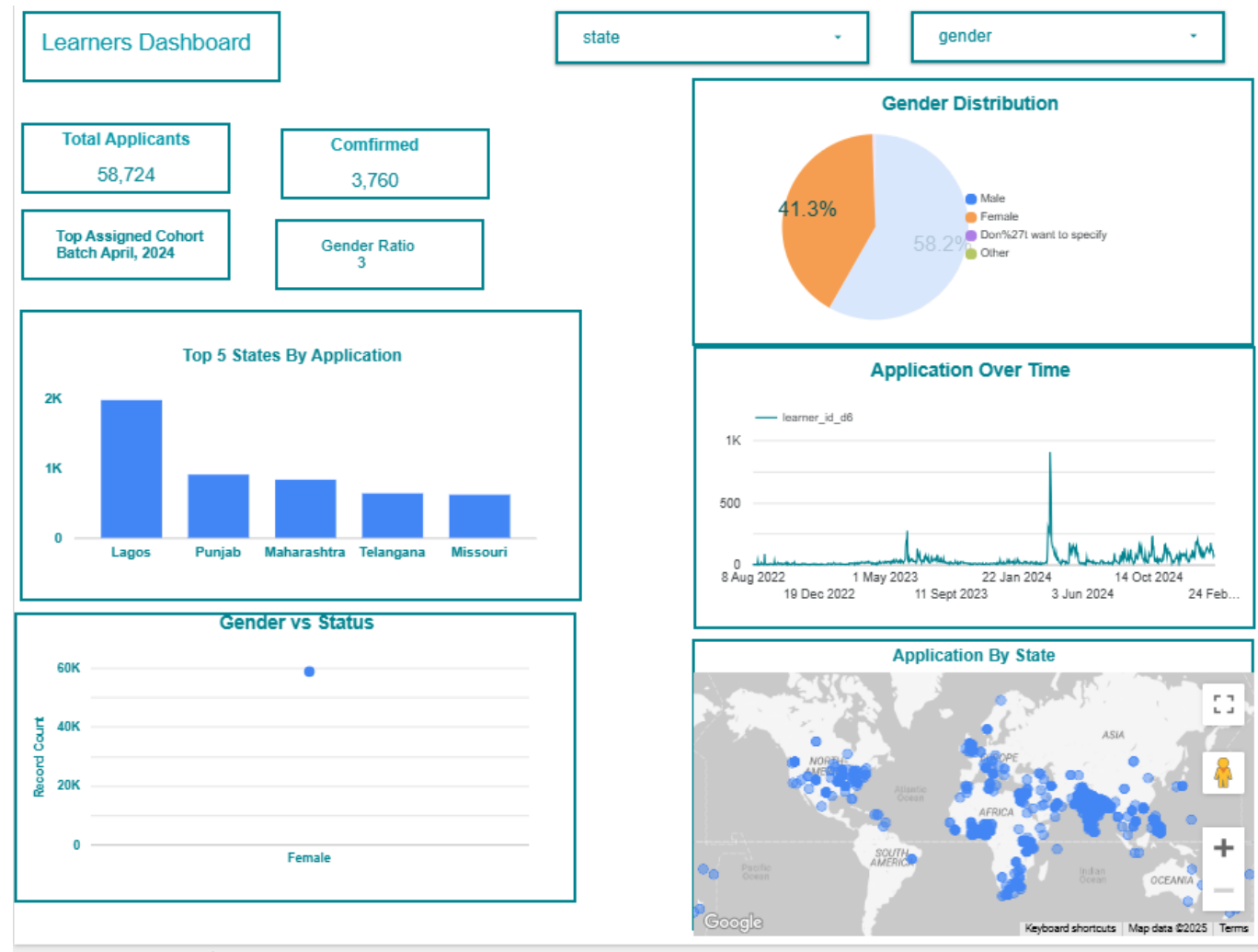
- ✓ **Standardized text** (Country, City, State, Names) with proper casing and trimmed spaces.
- ✓ **Cleaned IDs** (e.g., removed prefixes like "Learner#", "Opportunity#").
- ✓ **Converted all date fields** to a uniform datetime format.
- ✓ **Formatted emails** to lowercase and trimmed spaces.
- ✓ **Ensured numeric conversion** for Results and Amount Spent.

4. Addressing Unreliable Data

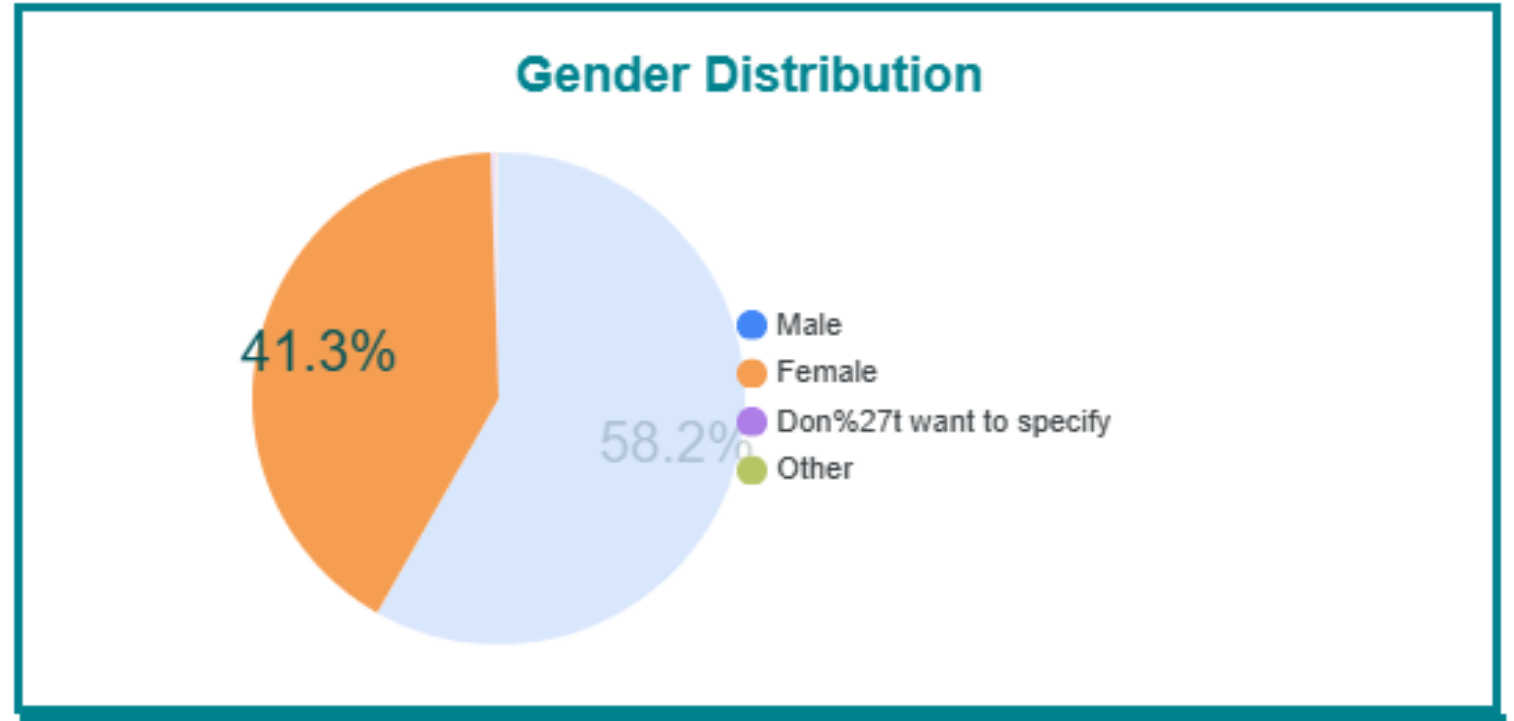
- ✓ **Ensured uniqueness** of key identifiers for reliable joins and tracking.
- ✓ **Validated numerical and currency fields** for consistency and accuracy.



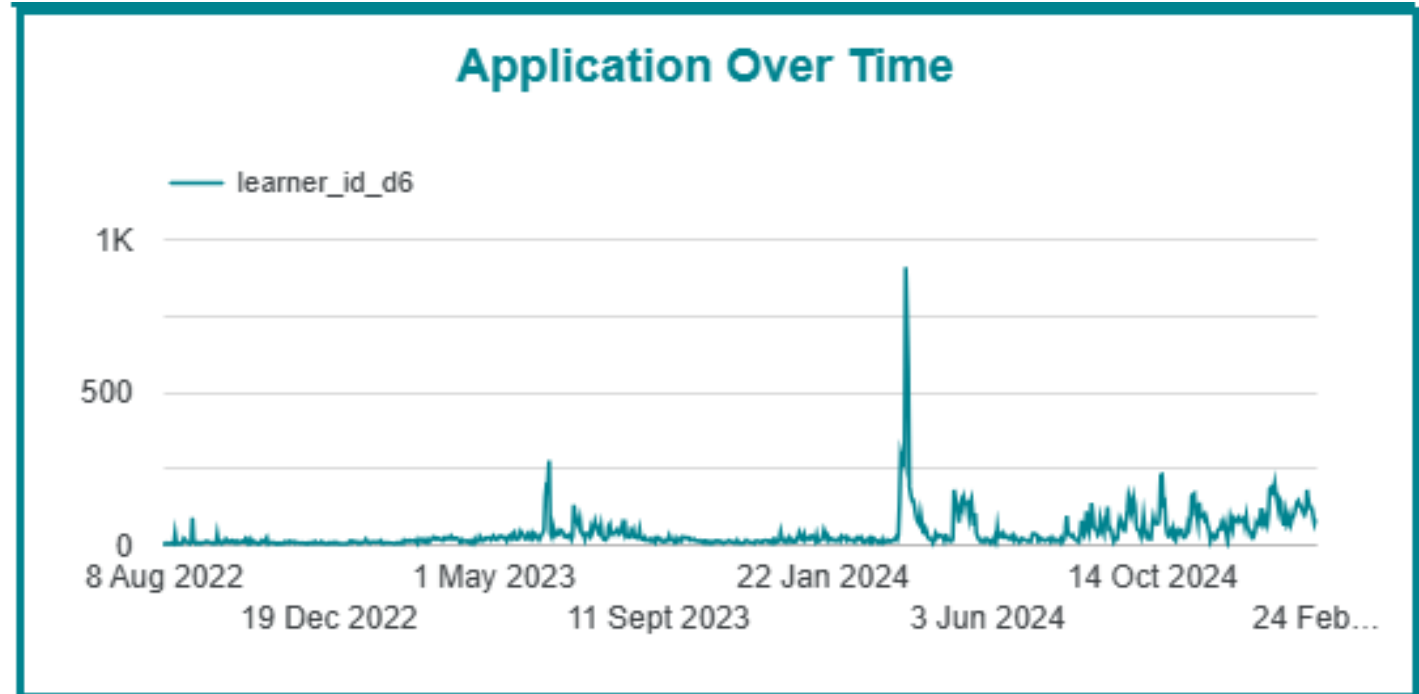
DASHBOARD:



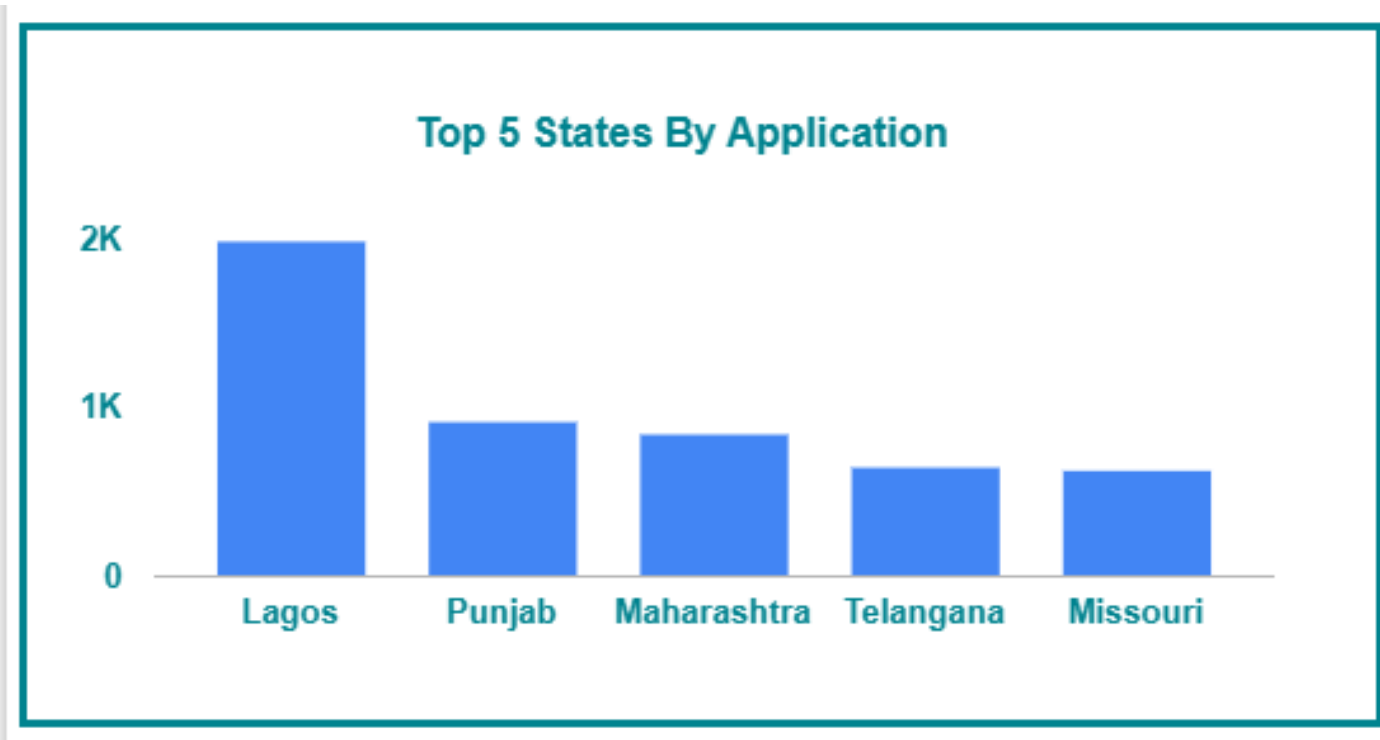
GENDER DISTRIBUTION:



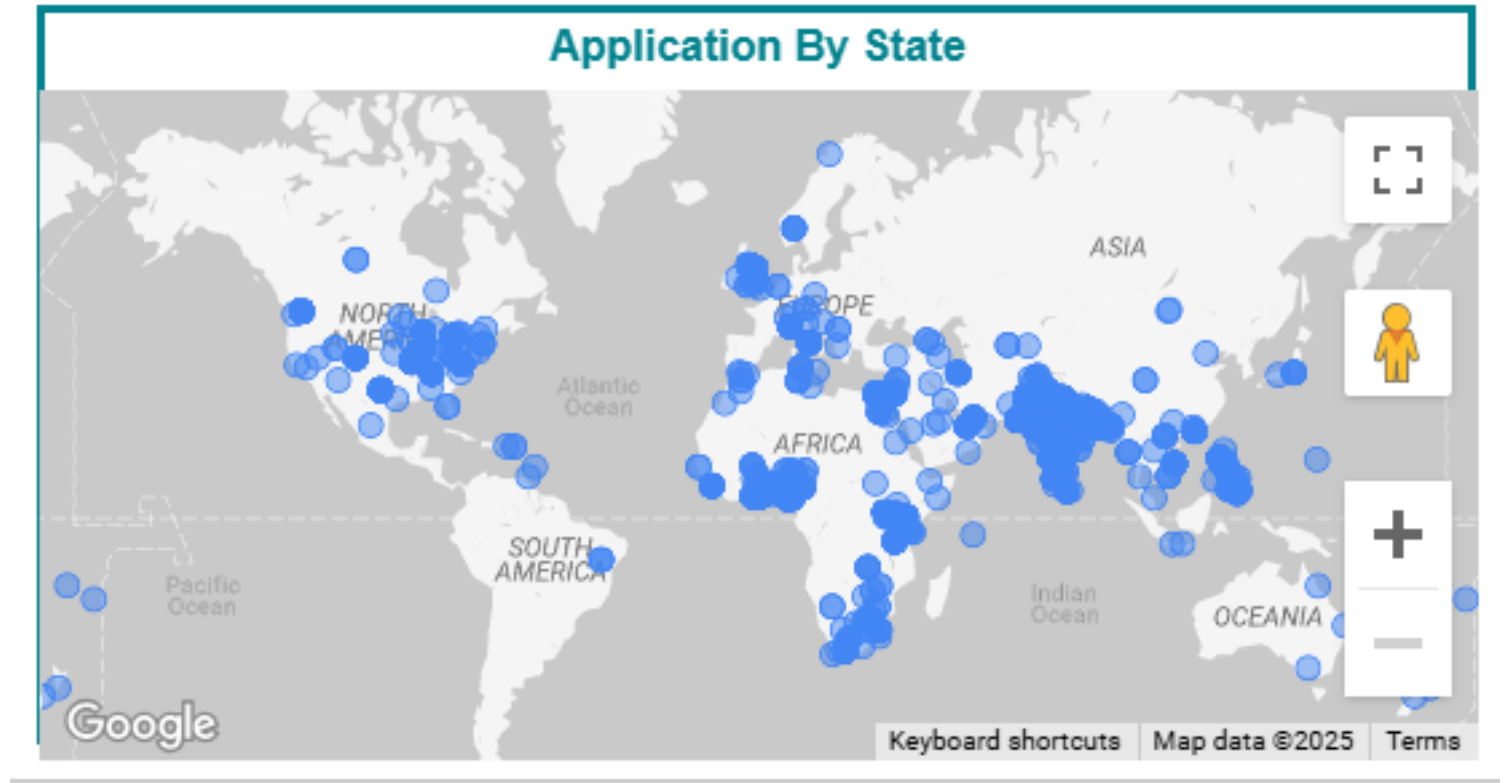
APPLOCATION OVER TIME:



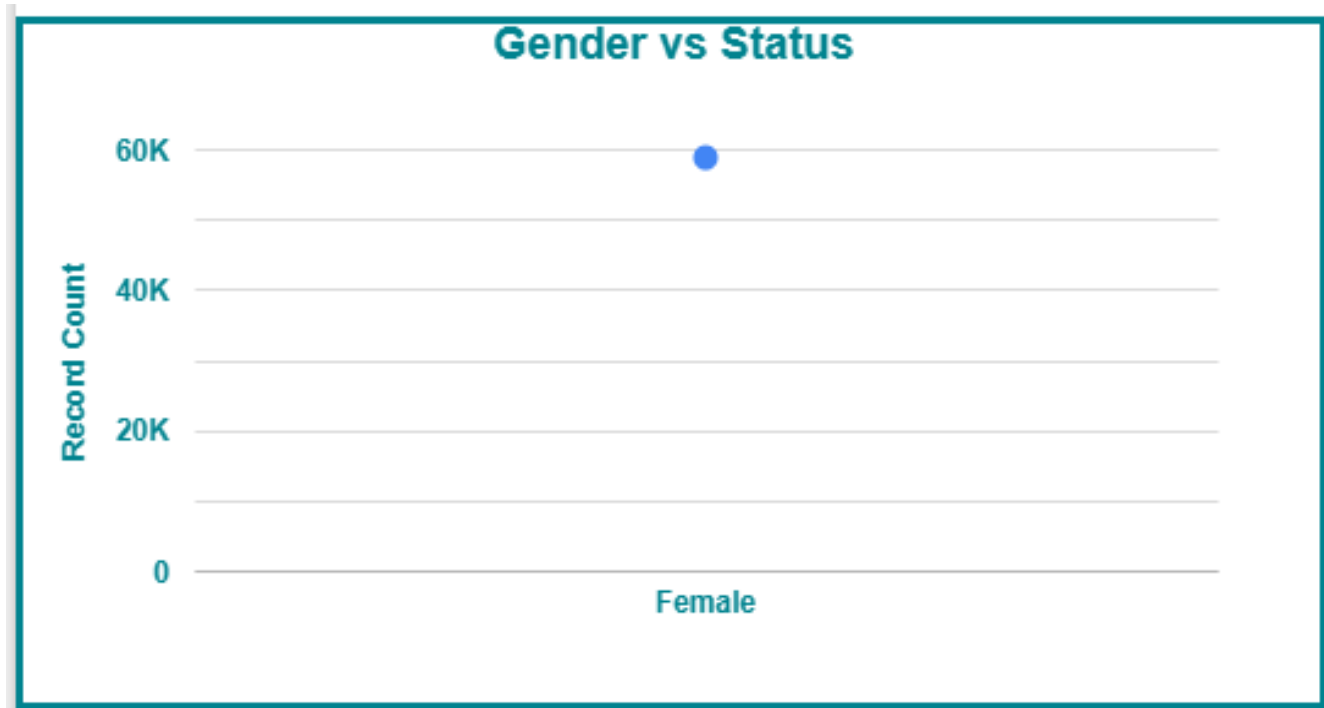
TOP 5 STATES BY APPLICATION:



APPLICATION BY STATE:



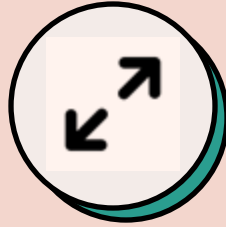
GENDER VS STATUS:





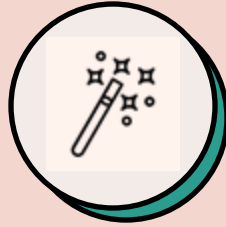
Our teamwork in data cleaning, processing, and visualization was crucial in revealing important insights into enrollment patterns, learner demographics, and platform interactions. By organizing and standardizing the data, we maintained consistency, accuracy, and reliability in our analyses. The results underscored significant learner engagement in South Asia and Africa, a substantial number of learners already holding degrees, and promising opportunities for targeted marketing campaigns and course enhancements. Through our structured approach, we streamlined data workflows, supporting better decision-making and boosting reporting efficiency.

Recommendations:



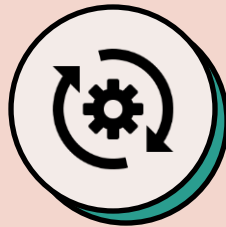
Expand Targeted Marketing

Since India, Nigeria, and Pakistan show the highest enrolments, consider region-specific marketing campaigns and partnerships with local institutions.



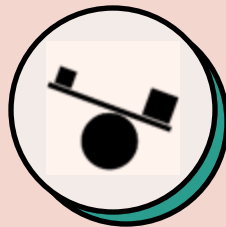
Enhance Course Offerings

With a significant number of learners holding Bachelor's and Graduate degrees, introducing advanced certifications and specialized programs can increase engagement.



Optimize Data Collection

Implement automated data cleaning procedures to maintain accuracy and consistency in future datasets.



Leverage Visual Analytics

Develop more interactive dashboards to track learner progress, cohort trends, and the effectiveness of marketing campaigns.



Improve Advertising Strategies

Analyse campaign performance data to refine targeting strategies and optimize ad spending for better conversion rates.

Thank you