



FINAL MASTER TABLE

Team-1



Team Member Details:

NAME	EMAIL ID
TARINEE	tarinee@s.amity.edu
PRASHANT KASHYAP	prashantkashyap5579@gmail.com
VARUN JOSE MADANU	varunjosemadanu@gmail.com
MUSTAFA OUN	wnvyvcgc@gmail.com
ANKIT RAJ	226301029@gkv.ac.in
ANCHAL KEDIA	anchalkedia45@gmail.com
USHWA SHARIQ	uswashariq@gmail.com
VAISHNAVI GATTAGONI	gattagonivaishnavi@gmail.com

Table Overview:

- **Name of the table:** Master_Table
- **Purpose:** The master table was created to integrate data from five interconnected sources—user data, cohort data, Cognito data, learner opportunity data, and opportunity data—into one cohesive structure. This unified dataset ensures efficient querying, maintains data consistency, and supports in-depth analysis. An additional dataset, the marketing dataset, was excluded from the master table as it has no direct or indirect relationship with the others; it will be analysed and visualized independently.
- **Data Source Summary:** User data, cognito data, cohort data, learner opportunity and opportunity data.

Scheme Design:

For each column in the master table, document:

SR. NO	COLUMN NO.	DATA TYPE	SOURCE FIELD
1.	cd_user_id	Text	Cognito Data
2.	email	Text	Cognito Data
3.	gender	Text	Cognito Data
4.	UserCreateDate	Date	Cognito Data
5.	UserLastModifiedDate	Date	Cognito Data
6.	city	Text	Cognito Data
7.	state	Text	Cognito Data
8.	ud_user_id	Text	User Data
9.	ud_learner_id	Text	User Data
10.	country	Text	User Data
11.	degree	Text	User Data
12.	institution	Text	User Data
13.	major	Text	User Data
14.	enrollment_id	Text	Learner Opportunity Data
15.	learner_id	Text	Learner Opportunity Data
16.	assigned_cohort	Text	Learner Opportunity Data
17.	apply_date	Date	Learner Opportunity Data
18.	status	Text	Learner Opportunity Data
19.	opportunity_id	Text	Opportunity Data
20.	Opportunity_name	Text	Opportunity Data
21.	category	Text	Opportunity Data
22.	opportunity_code	Text	Opportunity Data
23.	cohort_code	Text	Cohort Data
24.	start_date	Date	Cohort Data
25.	end_date	Date	Cohort Data
26.	size	Numeric	Cohort Data
27.	generation_name	Text	Master_Table
28.	birthdate_clean	Date	Cognito Data
29.	User_Activity_Portal	Numeric	Master_Table
30.	User_engage_opp	Numeric	Master_Table
31.	User_comp_gap	Numeric	Master_Table

Source Table Mapping:

SR. NO	COLUMN NAME 1	COULMN NAME 2	SOURCE 1	SOURCE 2
1.	user_id	user_id	Cognito Data	User Data
2.	learner_id	enrollment_id	User Data	Learner Opportunity Data
3.	Learner_id	Opportunity_id	Learner Opportunity Data	Opportunity Data
4.	Assigned_cohort	Cohort_code	Learner Opportunity Data	Cohort Data

Primary Key:

user_id (cognito data) , learner_id (user data)

Foreign Key:

learner_id (user data), enrolment_id (learner opportunity data), opportunity_id (opportunity data), assigned_cohort (cohort data).

Join Table Conditions or Logic:

CREATE TABLE MasterData AS

SELECT

cd.user_id AS cd_user_id,
cd.email,
cd.gender,
cd.UserCreateDate,
cd.UserLastModifiedDate,
cd.birthdate,
cd.city,
cd.zip,
cd.state,

ud.user_id AS ud_user_id,
ud.learner_id AS ud_learner_id,
ud.country,
ud.degree,
ud.institution,
ud.major,

lo.enrollment_id,
lo.learner_id,
lo.assigned_cohort,
lo.apply_date,
lo.status,

od.opportunity_id,
od.opportunity_name,
od.category,
od.opportunity_code,

od.tracking_questions,

coh.cohort_id,
coh.cohort_code,
coh.start_date,
coh.end_date,
coh.size

FROM "CognitoData" cd
JOIN "UserData" ud
 ON cd.user_id = ud.user_id
JOIN "LearnerOpportunity" lo
 ON lo.enrollment_id = ud.learner_id
JOIN "OpportunityData" od
 ON lo.learner_id = od.opportunity_id
JOIN "CohortData" coh
 ON coh.cohort_code = lo.assigned_cohort;

END