

A Project Report on

UNIVERSITY ANALYTICS SYSTEM



PREPARED BY:

ASSOCIATE ID	ASSOCIATE NAME
2239859	Maaz Patel
2239836	Shantanu Subramanian
2240035	Subhan Pinjari
2239975	Tirtharaj Mitra
2240015	Bhadra J Prakash
2239959	Yanamala Shalini
2239944	Mukku Sharath Chand

GUIDED BY:

Mentor

Kavin Kumar G

Associate

AIA- Cloud Data Integration

Coach

Grace Esther S

GenC Certified Coach, Data

Human Resource – GenC

CONTENTS

• Introduction	6
• Purpose of the project	6
• About the Software System	6
• Scope of the System	6
• Functional Requirement 1: Student Dimension	9 - 11
• Functional Requirement 2: College Dimension	12 - 14
• Functional Requirement 3: Faculty Dimension	15 - 17
• Functional Requirement 4: Department Dimension	18 - 20
• Functional Requirement 5: Course Dimension	21 - 23
• Functional Requirement 6: Time Dimension	24 - 26
• Functional Requirement 7: Result Staging	27 - 30
• Functional Requirement 8: Result Fact	31 - 33
• Functional Requirement 9: Result Aggregate	34 - 36
• Integrated Workflow	37
• Integrated Task Details	37
• Design Constraints	38
• Conclusion	38

LIST OF ABBREVIATIONS

TERM	DESCRIPTION
ETL	Extraction, Transformation and Loading / Extract Transform and Load
SRC	Source
TRG	Target
SQ	Source Qualifier
WF	Workflow
JNR	Joiner
EXP	Expression
AGG	Aggregator
RTR	Router
FIL	Filter
SEQ	Sequence Generator
DIM	Dimension
F	Fact
UPD	Update Strategy
UAS	University Analytics System

LIST OF TABLES

- Table 1: Student Dimension Source Description
- Table 2: Student Target Description
- Table 3: Student Dimensions Transformations
- Table 4: College Dimension Source Description
- Table 5: College Target Description
- Table 6: College Dimension Transformations
- Table 7: Faculty Dimension Source Description
- Table 8: Faculty Target Description
- Table 9: Department Dimension Source Description
- Table 10: Department Target Description
- Table 11: Department Dimension Transformations
- Table 12: Course Dimension Source Description
- Table 13: Course Target Description
- Table 14: Course Dimension Transformations
- Table 15: Time Dimension Source Description
- Table 16: Time Target Description
- Table 17: Time Dimension Transformations
- Table 18: Staging Source Description
- Table 19: Staging Target Description
- Table 20: Result Staging Transformations
- Table 21: Result Fact Source Description
- Table 22: Result Fact Target Description
- Table 23: Result Fact Transformations
- Table 24: Result Aggregate Source Description
- Table 25: Result Aggregate Target Description
- Table 26: Result Aggregate Transformations
- Table 27: Design Constraints

LIST OF FIGURES

Fig.1 Architecture Diagram

Fig 2: Flow Chart

Fig 3: Student Target Definition

Fig. 4: Student Mapping

Fig. 5: Student Dimension Workflow

Fig. 6 Student Dimension Target Table

Fig 7: College Target Definition

Fig. 8: College Mapping

Fig. 9: College Dimension Workflow

Fig. 10: College Dimension Target Table

Fig 11: Faculty Target Definition

Fig. 12: Faculty Mapping

Fig. 13: Faculty Dimension Workflow

Fig. 14: Faculty Dimension Target Table

Fig 15: Department Target Definition

Fig. 16: Department Mapping

Fig. 17: Department Dimension Workflow

Fig. 18: Department Dimension Target Table

Fig 19: Course Target Definition

Fig. 20: Course Mapping

Fig. 21: Course Dimension Workflow

Fig. 22: Course Dimension Target Table

Fig 23: Time Target Definition

Fig. 24: Time Mapping

Fig. 25: Time Dimension Workflow

Fig. 26: Time Dimension Target Table

Fig 27: Staging Target Definition

Fig. 28: Staging Mapping

Fig. 29: Result Staging Workflow
Fig. 30: Staging Result Target Table
Fig. 31: Staging Error Target Table
Fig 32: Result Fact Target Definition
Fig. 33: Result Fact Mapping
Fig. 34: Result Fact Workflow
Fig. 35: Result Fact Target Table
Fig 36: Result Aggregate Target Definition
Fig. 37: Result Aggregate Mapping
Fig. 38: Result Aggregate Workflow
Fig. 39: Result Aggregate Target Table
Fig. 40: Integrated Workflow
Fig. 41: Task Details

Introduction

Informatica PowerCenter was used to complete the project "University Analytics System" in accordance with business requirements (ETL Tool). Informatica PowerCenter is an ETL-based data integration tool. It is popular tool for extraction, transformation, and loading. The Informatica PowerCenter components help to extract data from its source, convert it in accordance with business needs, and load it into a destination data warehouse. The " University Analytics System " system was created to organise the data of various students enrolled in the University and to calculate and aggregate the results.

Purpose of the Project

The main aim of this project is to create a University Analytics System that aids in the analysis of various possibilities for the provided data in accordance with business requirements. This report lists the functional requirements for this system along with the steps followed to create the system.

About the Software System

A University's sample data warehouse is loaded using the software system. Data from the source system is supplied as a flat file that is to be imported into a relational database in the data warehouse. The data that has been entered into the data warehouse will be used to make decisions. The following section will cover aspects related to University Analytics System (UAS) application. UAS is a system used for the analysis of various department/student/faculty performances.

The following are the modules in this proposed system

- a) Student Dimension
- b) College Dimension
- c) Department Dimension
- d) Course Dimension
- e) Faculty Dimension
- f) Time Dimension
- g) Results Staging
- h) Results Fact
- i) Results Aggregate

Scope of the System

The scope of the system is to create a University Analytics System Data Warehouse and populate Data Warehouse tables.

ARCHITECTURE DIAGRAM

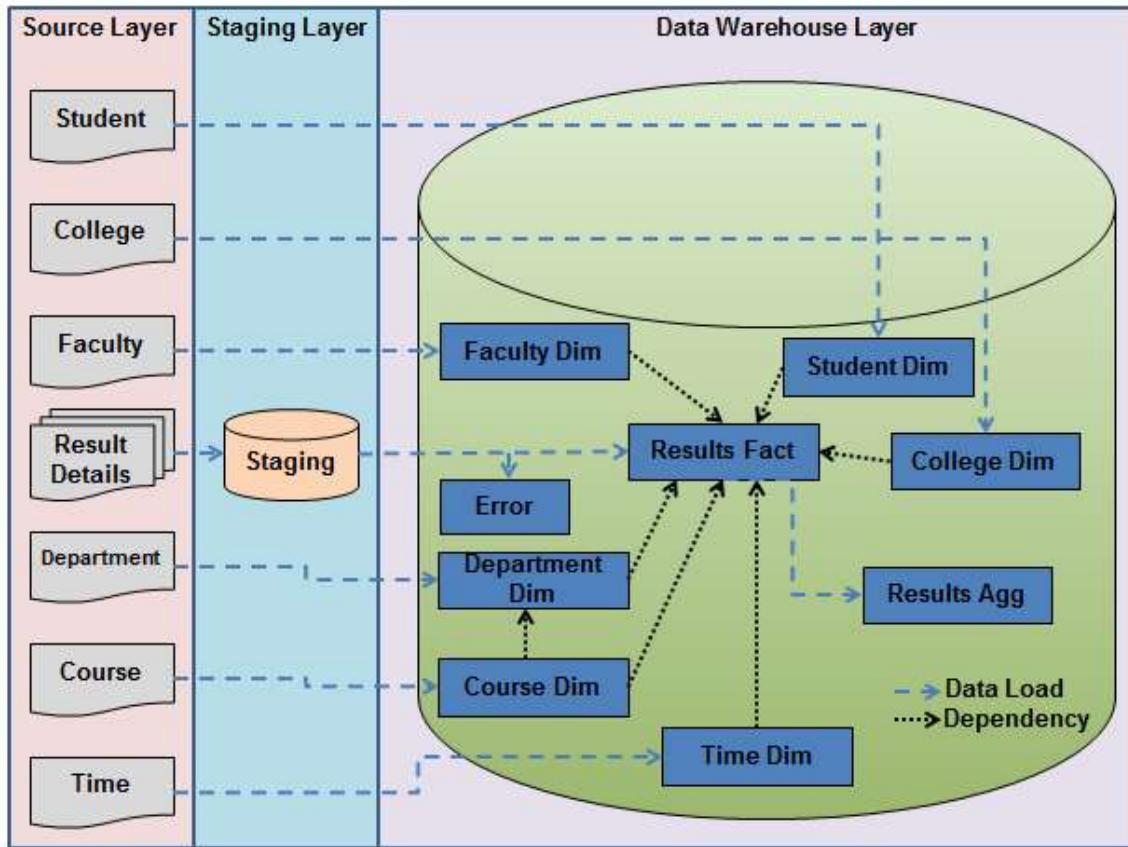


Fig.1: Architecture Diagram

FLOW CHART

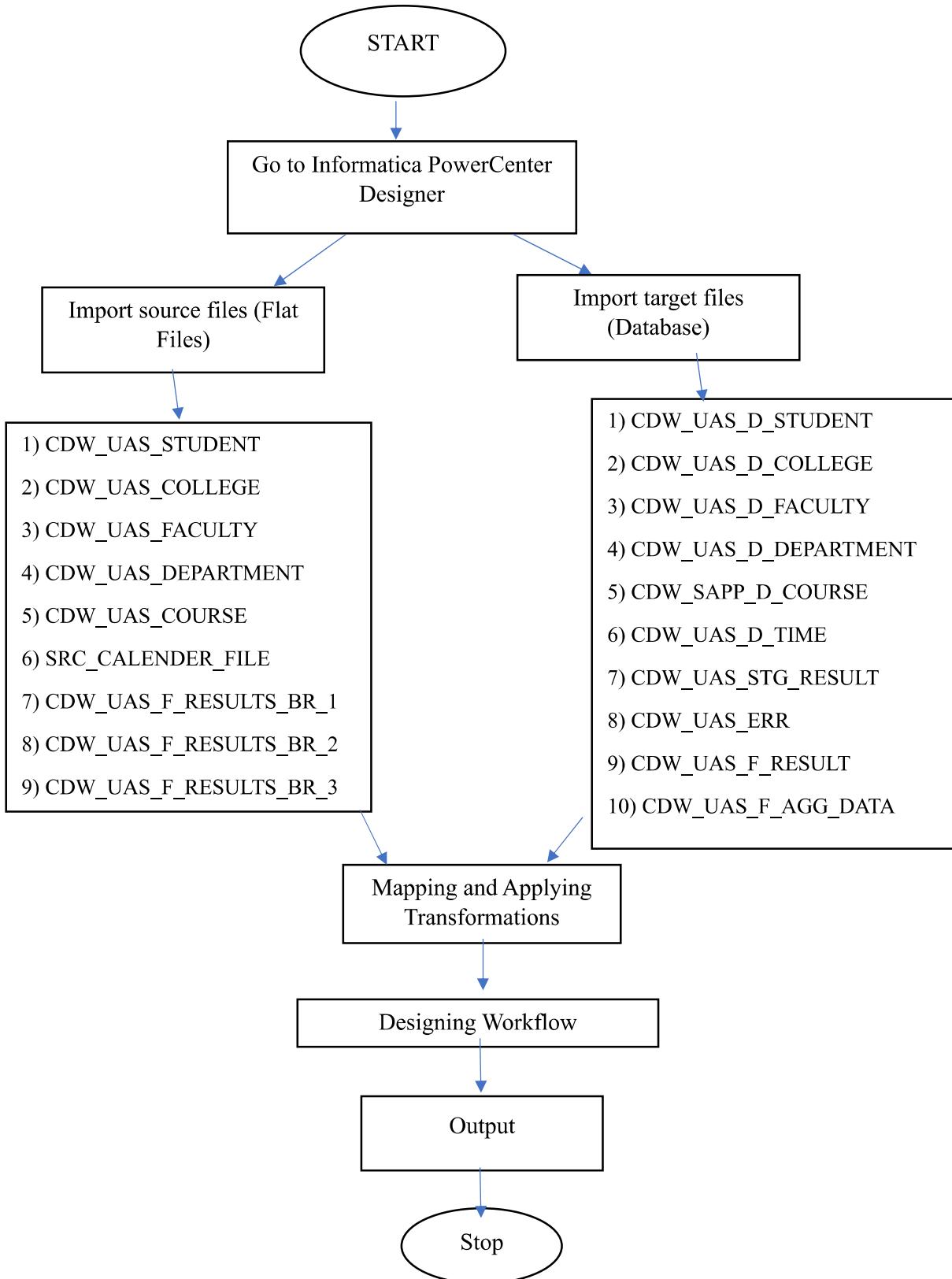


Fig 2: Flow Chart

Functional Requirement 1

STUDENT DIMENSION

SOURCE

The first functional requirement was to load the data into the Student Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_Student.txt	This is a Comma delimited file with 11 Input Fields	 CDW_UAS_Student.txt

Table 1: Student Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_D_STUDENT	It is a SCD type 2 table maintaining the created date and end date.	Dimension Table	 1_Student_Dimension.txt

Table 2: Student Target Description

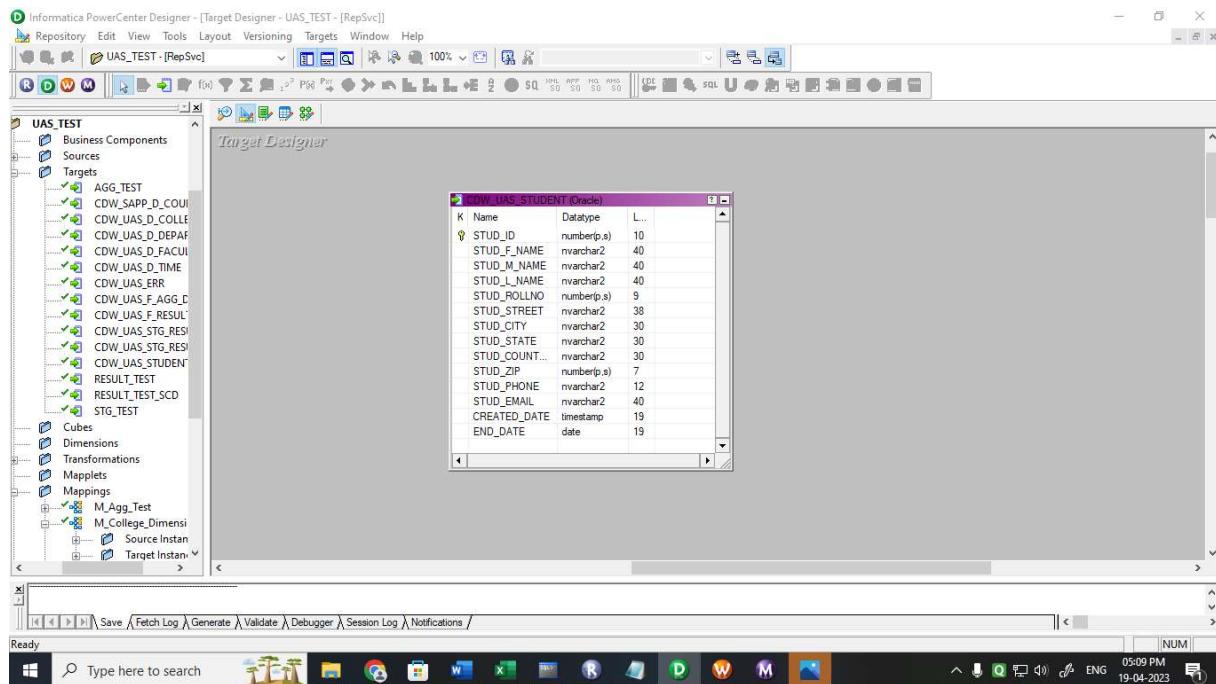


Fig 3: Student Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Used
STUD_ID	System generated based on Student SSN. Insert for new Student/ expire the old record and insert the updated record for existing Students	Sequence Generator	
STUD_F_NAME	Convert the Name to Title Case	Expression	InitCap()
STUD_M_NAME	Convert the middle name in lower case	Expression	Lower()
STUD_L_NAME	Convert the Last Name in Title Case	Expression	InitCap()
STUD_ROLLNO	Direct move (Abort the session if Roll Number is invalid. Convert to number)	Expression	IIF
STUD_STREET	Concatenate Apartment no and Street name of Student's Residence with comma as a separator	Expression	Concat
STUD_ZIP	Direct move (Convert to number)	Expression	ToInteger
STUD_PHONE	Change the format to xxx-xxx-xxxx	Expression	Concat,Substring
CREATED_DATE	SYSDATE	Expression	SYSDATE
END_DATE	Null on insert / SYSDATE - (1Sec) on update.	Update Strategy	

Table 3: Student Dimensions Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

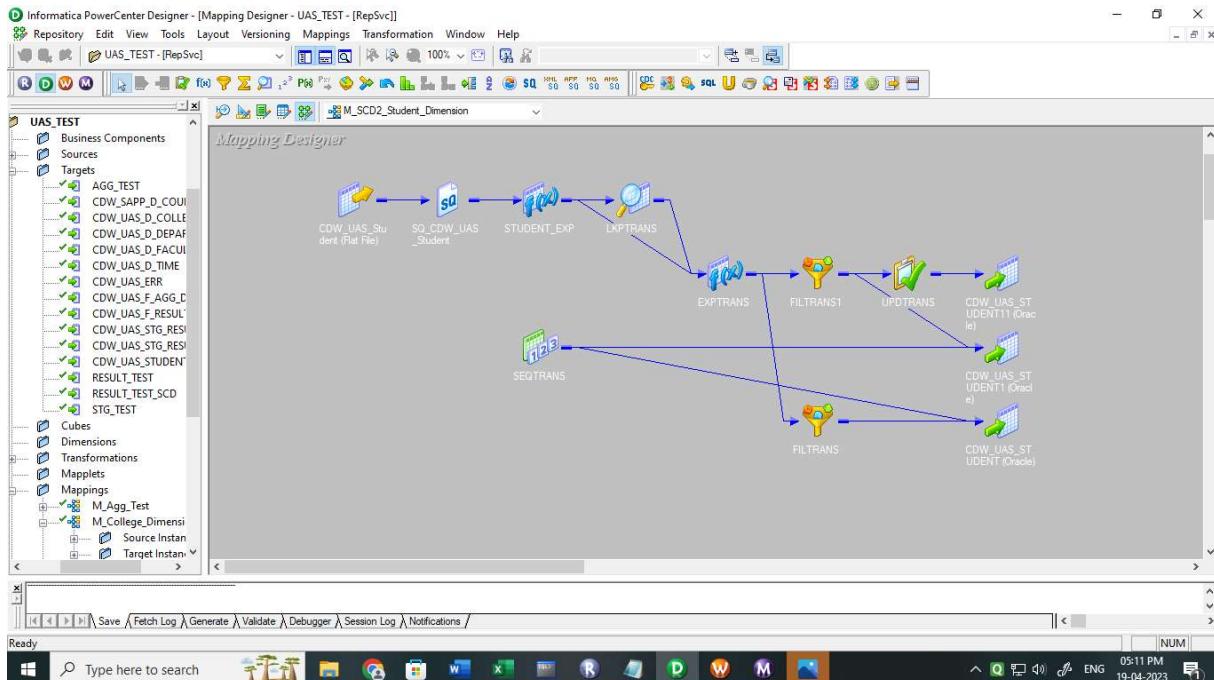


Fig. 4: Student Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

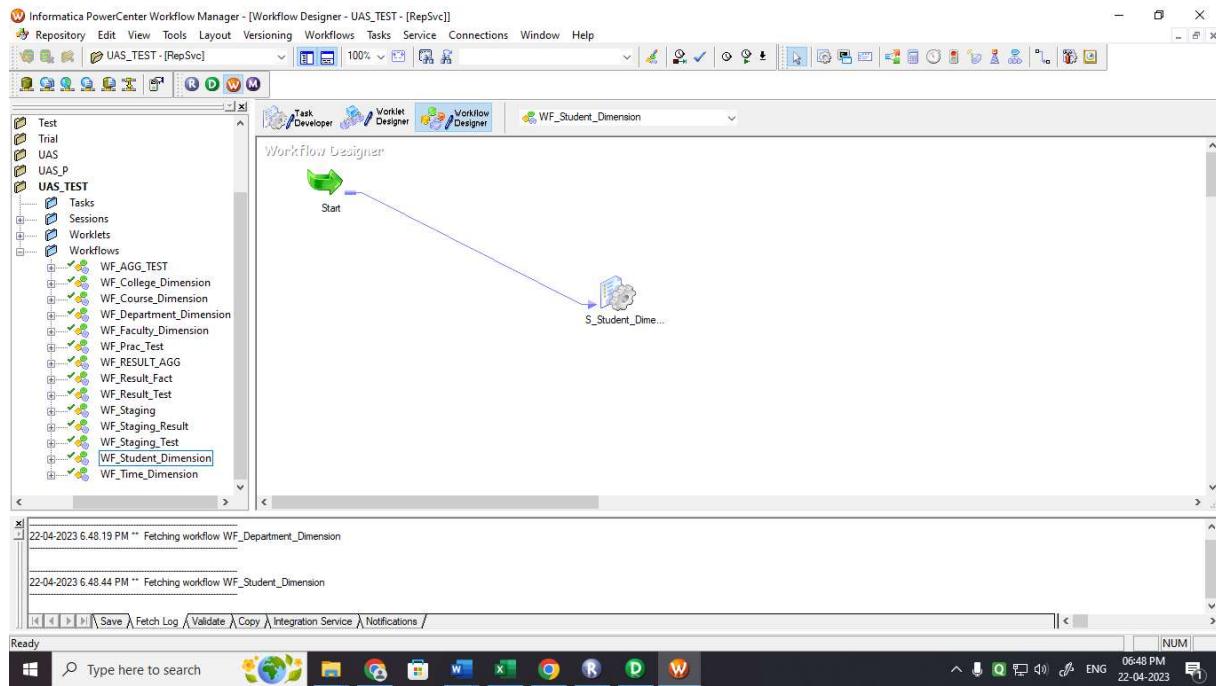


Fig. 5: Student Dimension Workflow

Target Post Execution:

The screenshot shows a "Run SQL Command Line" window. The SQL command executed is "SQL> SELECT * FROM CDW_UAS_STUDENT;". The output displays 8 rows of data from the CDW_UAS_STUDENT table, including columns such as STUD_ID, STUD_F_NAME, STUD_CITY, STUD_COUNTRY, and STUD_EMAIL. The data includes records for students Maaz, Shantanu, Subhan, Shaheen, Bhadra, Tirtharaj, Shalini, and Sharath, along with their respective details like city, state, and contact information. The bottom status bar shows "Type here to search", system icons, and "05:12 PM 19-04-2023".

STUD_ID	STUD_F_NAME	STUD_CITY	STUD_COUNTRY	STUD_EMAIL
1	Maaz	Pune	India	859 201 NIBM Road maaz.patel@cognizant.com
2	Shantanu	Bangalore	India	836 148 Karnataka Road shantanu.subramanian@cognizant.com
3	Subhan	Hyderabad	India	878 405 Vr Colony subhan.pinjari2@cognizant.com
4	Shaheen	Pallikad	India	963 12 Thotakara shaheen.muhammed@cognizant.com
5	Bhadra	Kolla	India	1012 33 Airport Road bhadra.prakash1@cognizant.com
6	Tirtharaj	Mumbai	India	945 901 Station Road tirtharaj.mitra2@cognizant.com
7	Shalini	Chennai	India	986 412 St. Joseph Road yanamala.shalini@cognizant.com
8	Sharath	Hyderabad	India	216 754 Bodhan sharath.chand@cognizant.com

Fig. 6 Student Dimension Target Table

Functional Requirement 2

COLLEGE DIMENSION

SOURCE

The functional requirement was to load the data into the College Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_College.txt	This is a Pipe delimited file with 7 Input Fields	 CDW_UAS_COLLEG E.txt

Table 4: College Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_D_College	It is a SCD type 1 table getting records from college source file.	Dimension Table	 2_College_Dimensi on.txt

Table 5: College Target Description

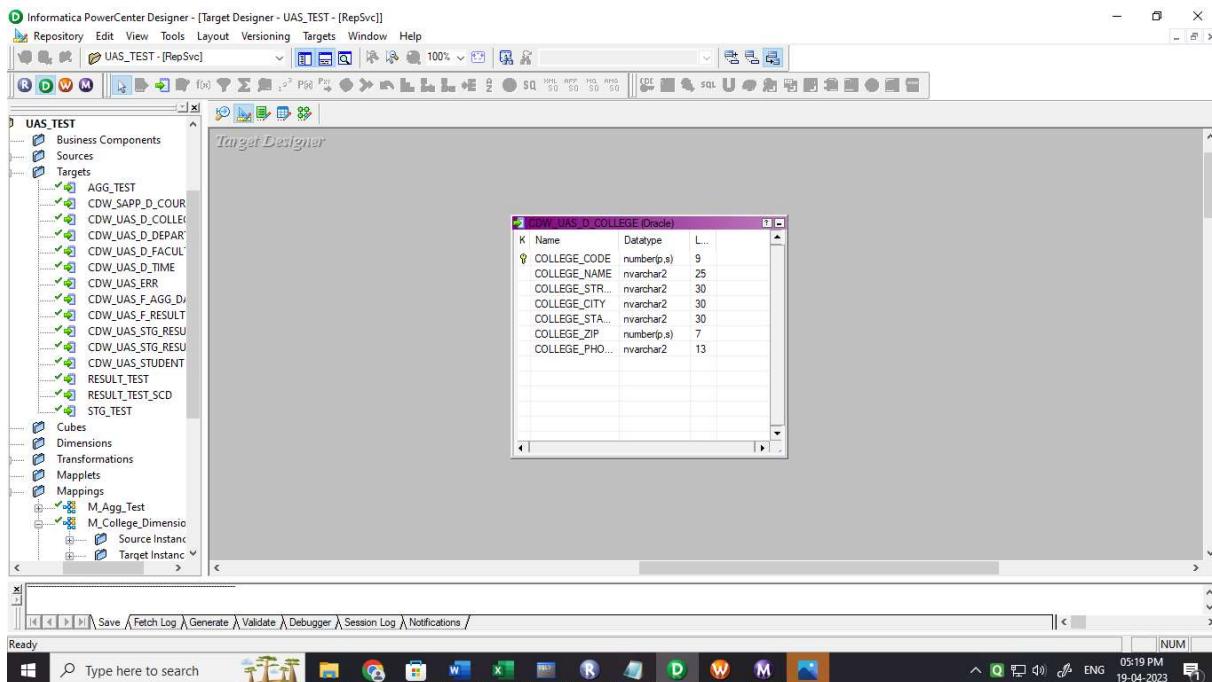


Fig 7: College Target Definition

TRANSFORMATIONS:

COLUMN NAME	MAPPING LOGIC	TRANSFORMATION NAME	FUNCTION USED
COLLEGE_CODE	Insert if new College/update the entire record for existing College	Update Strategy	DD_UPDATE or 1
COLLEGE_ZIP	If the source value is null load default value else Direct move	Expression	IIF
COLLEGE_PHONE	Change the format of phone number to (XXX)XXX-XXXX	Expression	Concat, Substring

Table 6: College Dimension Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

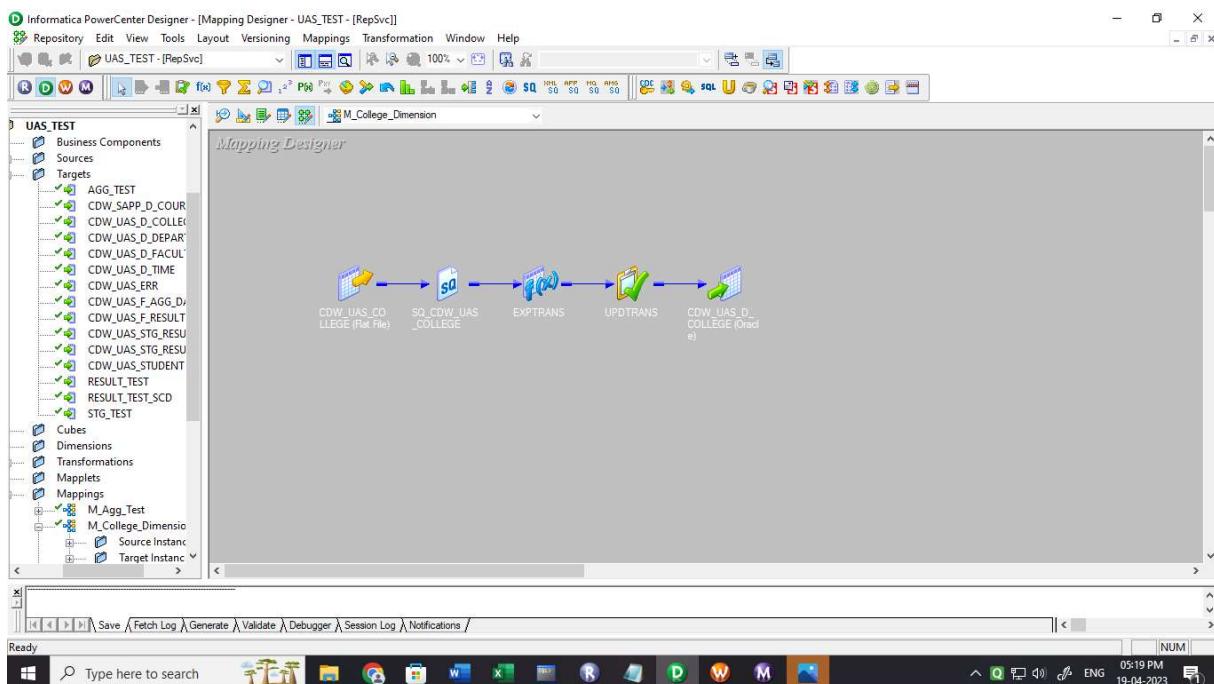


Fig. 8: College Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

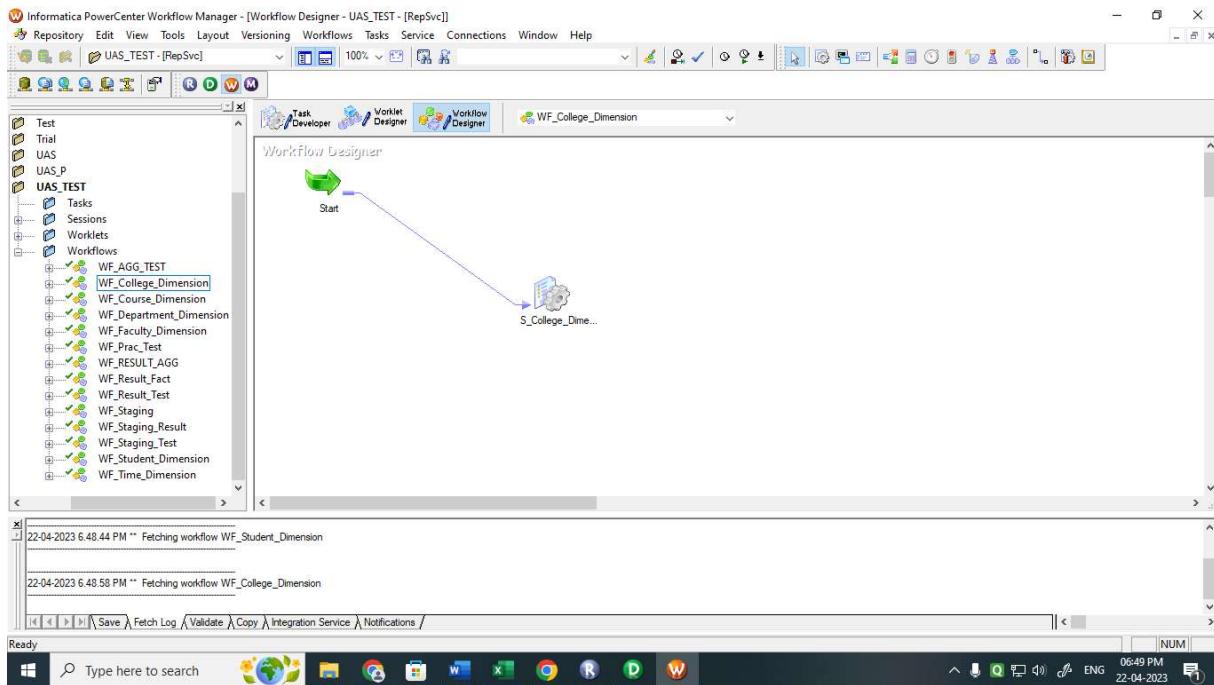


Fig. 9: College Dimension Workflow

Target Post Execution:

The screenshot shows a SQL command line interface. The command "SELECT * FROM CDW_UAS_D_COLLEGE;" was run, resulting in 9 rows selected. The output is a table with columns: COLLEGE_CODE, COLLEGE_NAME, COLLEGE_STREET, COLLEGE_CITY, COLLEGE_STATE, COLLEGE_ZIP, and COLLEGE_PHONE. The data includes various college details such as SAE Kondhwa, People Education Society, etc., across different states like Maharashtra, Karnataka, Kerala, etc. The Windows taskbar at the bottom shows the date and time as "19-04-2023 05:22 PM".

COLLEGE_CODE	COLLEGE_NAME	COLLEGE_STREET	COLLEGE_CITY	COLLEGE_STATE	COLLEGE_ZIP	COLLEGE_PHONE
223	SAE Kondhwa	Katraj Road	Pune	Maharashtra	411048	(997)025-4625
77	People Education Society	2nd Cross	Bangalore	Karnataka	560050	(882)672-7364
33	SREENIDHI	GHATKESAR	HYDERABAD	TELANGANA	501301	(276)403-9424
25	TKM CE	KILIKOLLUR	KOLLAM	KERALA	691005	(041)224-5953
38	Sir Agnel	Andheri East Road	Mumbai	Maharashtra	400081	(022)275-0394
52	SRM IST	KTR	Chennai	TamilNadu	603203	(044)274-5510
42	IIM Bangalore	Airport road	Bangalore	Karnataka	560022	(896)336-9645
41	IIM Ahmedabad	Station road	Ahmedabad	Gujarat	670022	(986)302-3654
46	IIM Indore	IIM Road	Indore City	Madhya Pradesh	670022	(630)236-5498

Fig. 10: College Dimension Target Table

Functional Requirement 3

FACULTY DIMENSION

SOURCE

The functional requirement was to load the data into the Faculty Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_Faculty.txt	This is a Pipe delimited file with 4 Input Fields	 CDW_UAS_FACULTY.txt

Table 7: Faculty Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_D_Faculty	It is a SCD type 1 table getting records from faculty source file.	Dimension Table	 3_Faculty_Dimension.txt

Table 8: Faculty Target Description

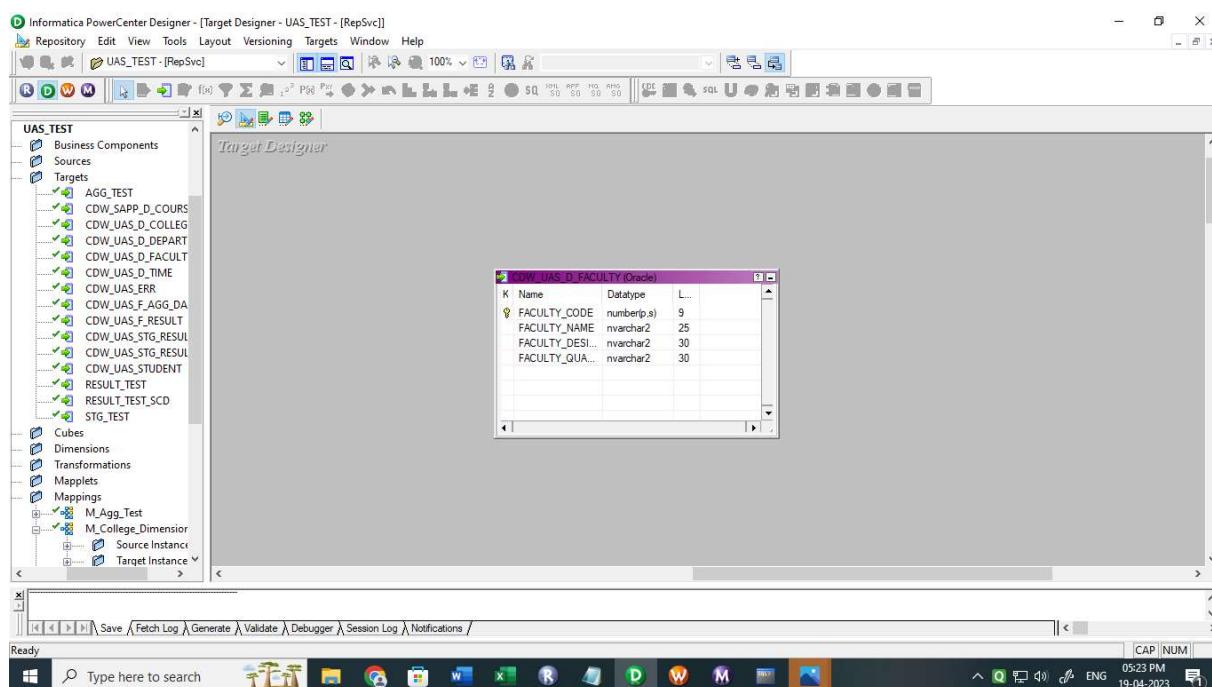


Fig 11: Faculty Target Definition

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

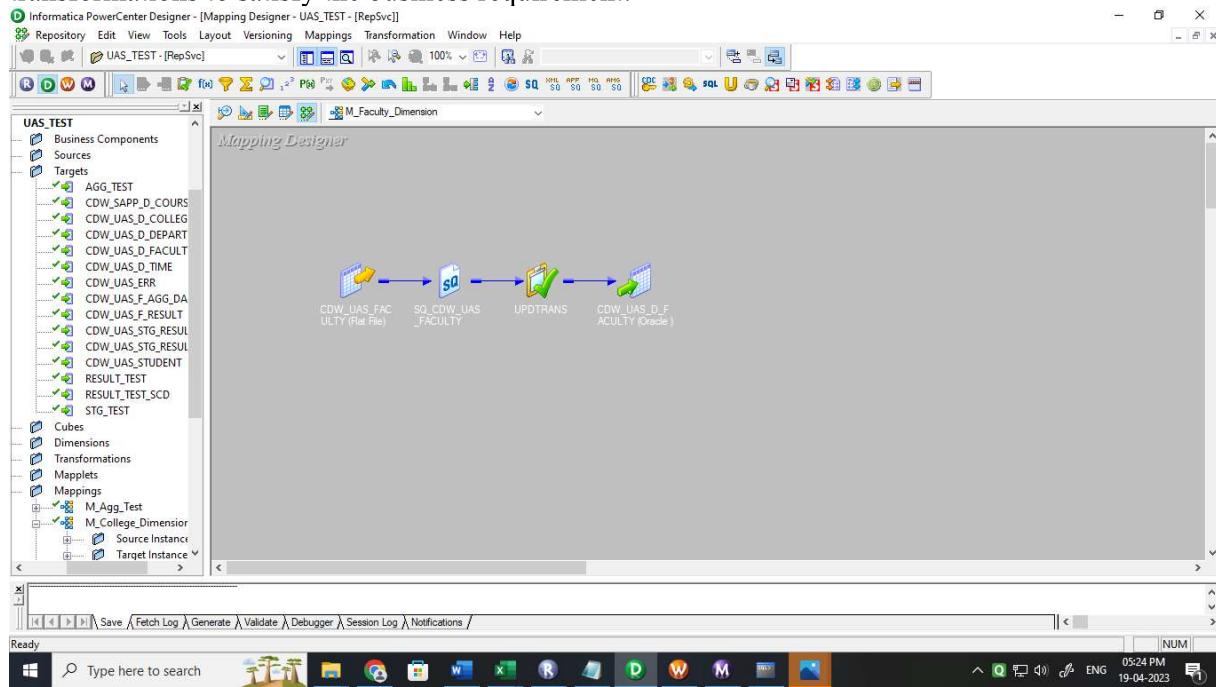


Fig. 12: Faculty Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

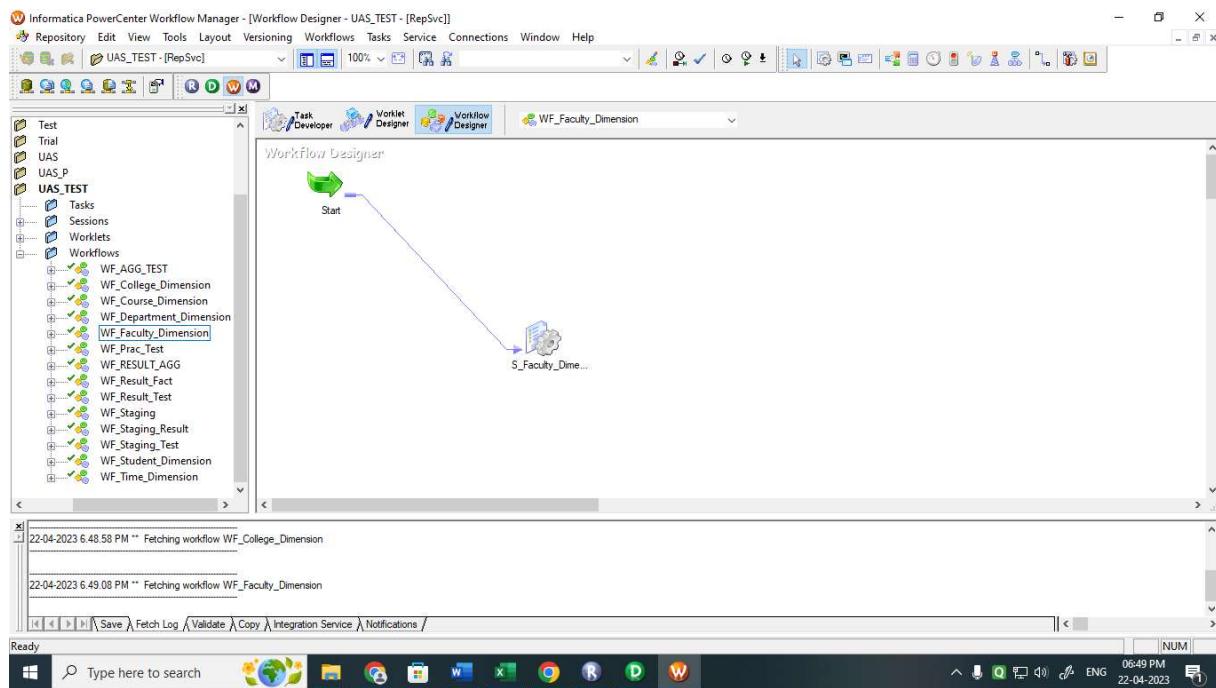


Fig. 13: Faculty Dimension Workflow

Target Post Execution:

```
Run SQL Command Line
```

```
SQL> SELECT * FROM CDW_UAS_D_faculty;
```

FACULTY_CODE	FACULTY_NAME	FACULTY_DESIGNATION	FACULTY_QUALIFICATION
183112001	Santosh S Kumar	Assistant Professor	Master of Science
183122002	Rajesh Kanakala	Sports Teacher	Master of Science
183132002	Jai Rajasekhar	Assistant Director	Master of Business
183142003	Ravi Varma	Senior Professor	Post Graduation
183152006	Vikram Sara Bhai	Physical education Teacher	Bachelor of Science
183162005	Suresh Raina	Vice President	Master of Science
183172001	Vinod Pusala	Assistant Professor	Post Graduation
183182004	Harini Yadhav	Senior Lecture	Master of Science
183192006	Priya H Bhavani	Teacher	Bachelor of Arts

```
9 rows selected.
```

```
SQL>
```



Fig. 14: Faculty Dimension Target Table

Functional Requirement 4

DEPARTMENT DIMENSION

SOURCE

The functional requirement was to load the data into the Department Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_Department.txt	This is a Pipe delimited file with 4 Input Fields	 CDW_UAS_DEPARTMENT.txt

Table 9: Department Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_D_Department	It is a SCD type 1 table getting records from department source file and generates new field department_id.	Dimension Table	 4_Department_Dimension.txt

Table 10: Department Target Description

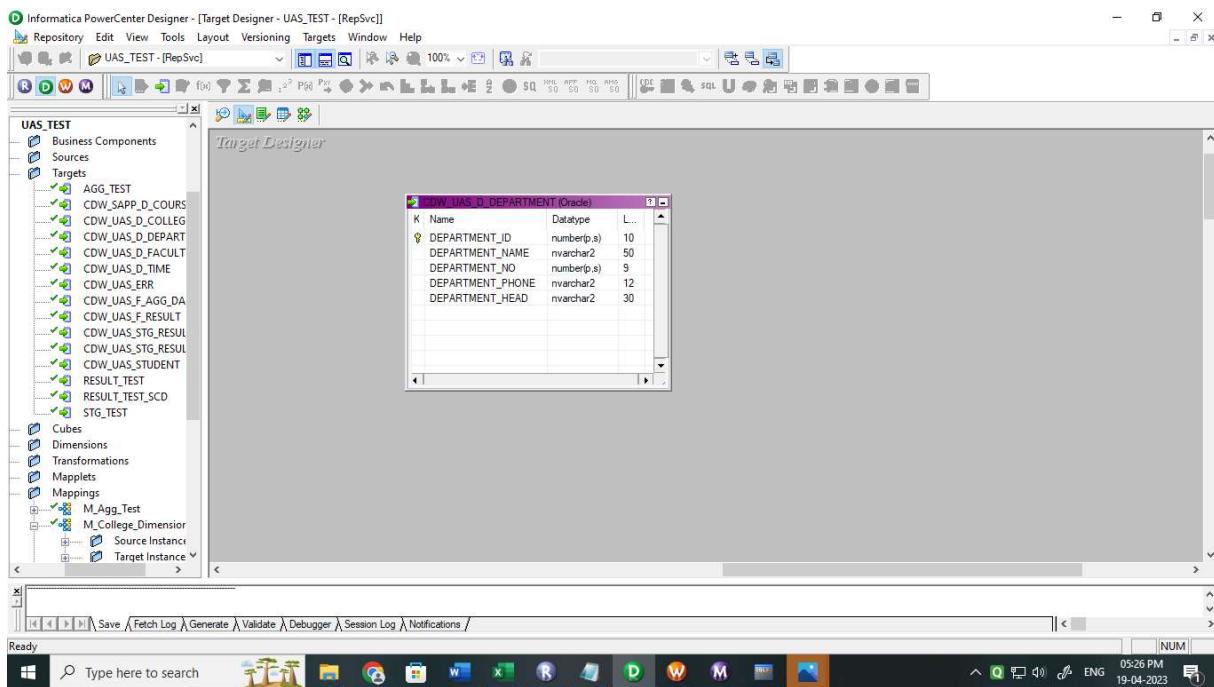


Fig 15: Department Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Used
DEPARTMENT_ID	System generated based on Department No.Insert for new Department/update the entire record for existing Departments	Sequence Generator	
DEPARTMENT_NAME	If a percentage symbol is present in the name of Department remove and load else direct move	Expression	IIF, InStr
DEPARTMENT_NO	Direct move(Abort the session if Number is invalid)	Expression	IIF
DEPARTMENT_PHONE	Standardize the phone number to XXX-XXX-XXXX	Expression	Concat, Substring

Table 11: Department Dimension Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

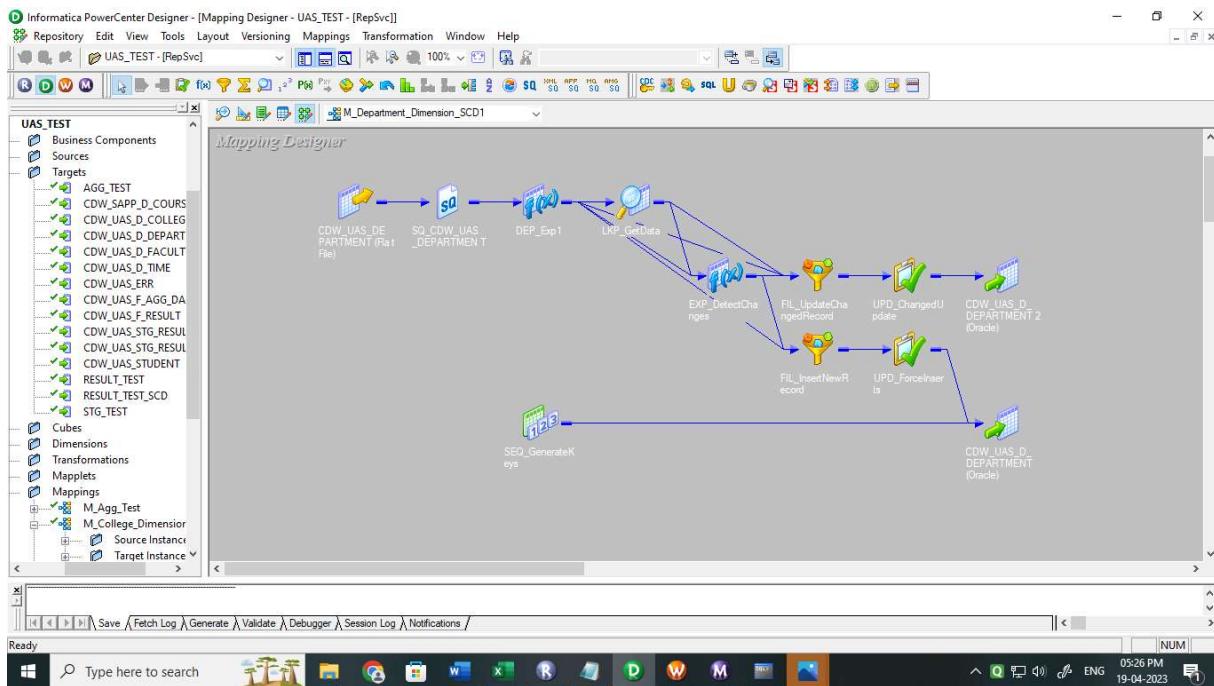


Fig. 16: Department Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

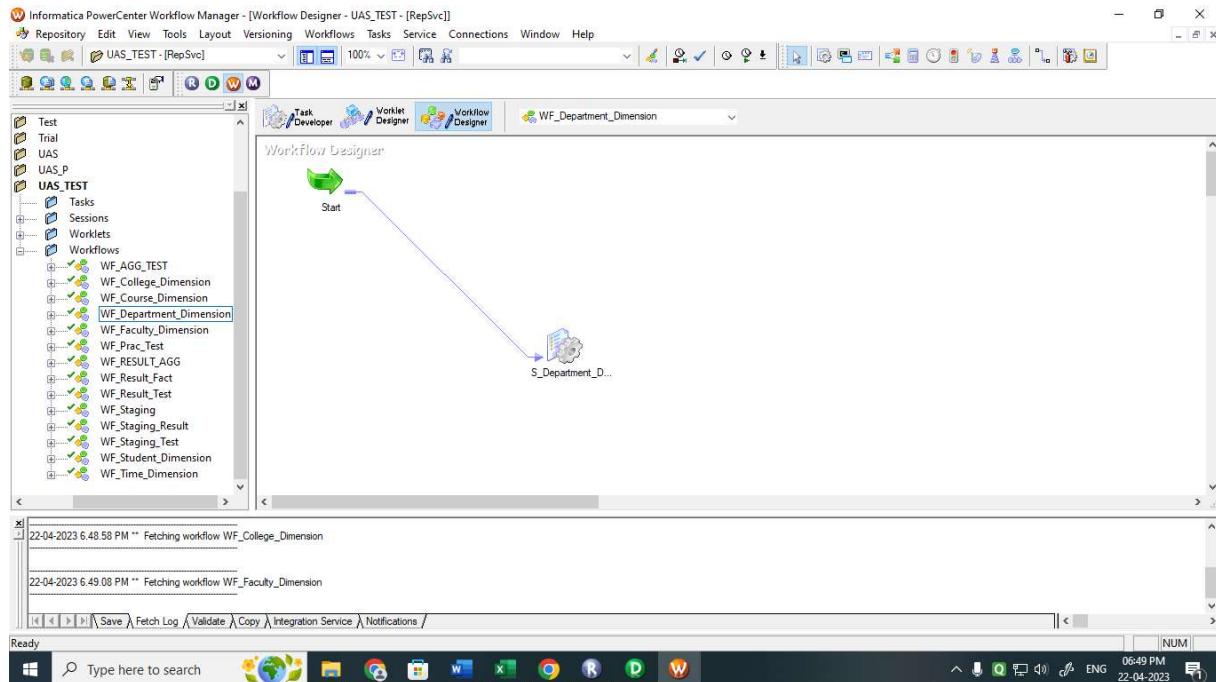


Fig. 17: Department Dimension Workflow

Target Post Execution:

The screenshot shows a SQL Command Line window. The command entered is "SELECT * FROM CDW_UAS_D_department;". The output displays 8 rows of data from the target table:

DEPARTMENT_ID	DEPARTMENT_NAME	DEPARTMENT_NO	DEPARTMENT_P	DEPARTMENT_HEAD
1	DEPARTMENT OF CIVIL	2022	228-963-2	JOHN CREED
2	DEPARTMENT OF CHEMICAL	2023	224-987-2	SMITH HOLLAND
3	DEPARTMENT OF COMPUTER SCIENCE	2024	225-536-1	CLERK HARLOW
4	DEPARTMENT OF ELECTRONICS	2025	365-478-9	JOSEPH BEYONCE
5	DEPARTMENT OF ELECTRICAL	2026	896-542-8	WILL TATE
6	DEPARTMENT OF MECHANICAL	2027	225-647-2	SKY ANDREW
7	DEPARTMENT OF PRODUCTION	2028	225-478-9	TOM TRISTAN
8	DEPARTMENT OF ARCHITECTURE	2029	224-987-8	CRUISE J FERGUNSON

8 rows selected.

Fig. 18: Department Dimension Target Table

Functional Requirement 5

COURSE DIMENSION

SOURCE

The functional requirement was to load the data into the Course Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_Course.txt	This is a Comma delimited file with 4 Input Fields	 CDW_UAS_COURSE.txt

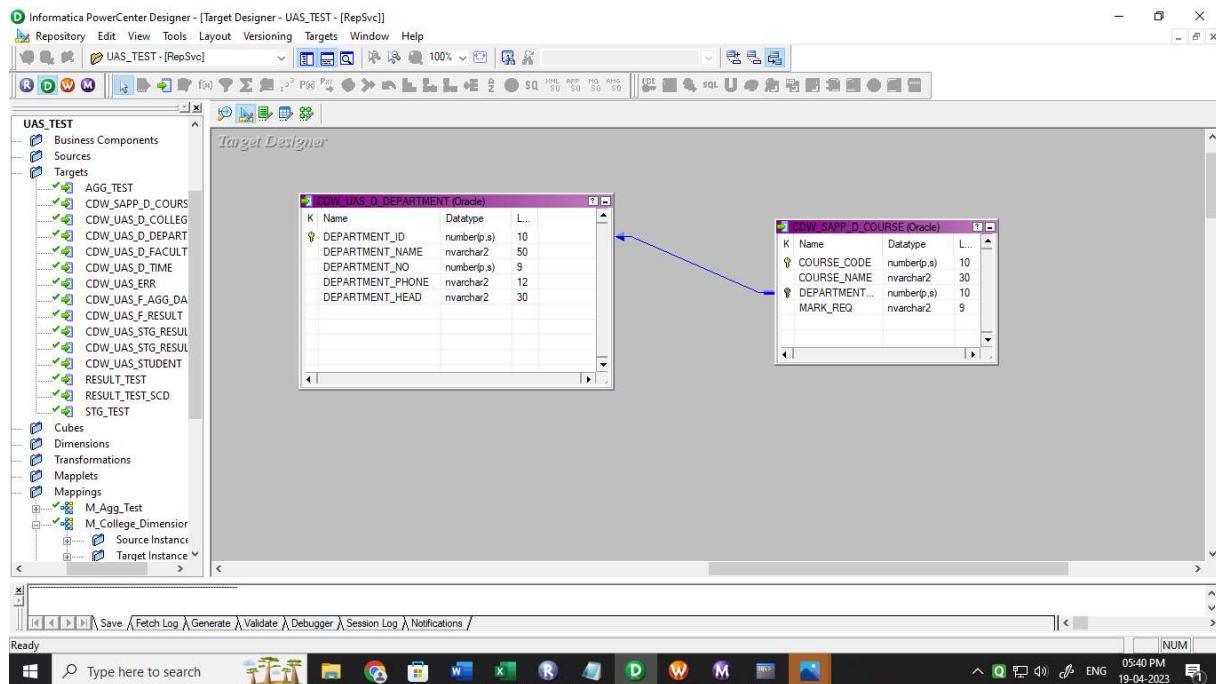
Table 12: Course Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_SAPP_Course	It is a SCD type 1 table getting records from course source file and field department_id from department table	Dimension Table	 5_Course_Dimension.txt

Table 13: Course Target Description



The screenshot shows the Informatica PowerCenter Designer interface. The left pane displays the project structure under 'UAS_TEST' with nodes like AGG_TEST, CDW_SAPP_D_COURS, CDW_UAS_D_COLLEG, etc. The central pane is titled 'Target Designer' and contains two tables:

- CDW_UAS_D_DEPARTMENT (Oracle)** with columns:

K	Name	Datatype	L...
Y	DEPARTMENT_ID	number(p,s)	10
	DEPARTMENT_NAME	nvarchar2	50
	DEPARTMENT_NO	number(p,s)	9
	DEPARTMENT_PHONE	nvarchar2	12
	DEPARTMENT_HEAD	nvarchar2	30
- CDW_UAS_D_COURSE (Oracle)** with columns:

K	Name	Datatype	L...
Y	COURSE_CODE	number(p,s)	10
	COURSE_NAME	nvarchar2	30
	DEPARTMENT...	number(p,s)	10
	MARK_REQ	nvarchar2	9

Fig 19: Course Target Definition

TRANSFORMATION:

Column Name	Mapping Logic	Transformation Name	Function Used
COURSE_CODE	Insert for new Course/update the entire record for existing Course(Abort the session if Code is invalid)	Update Strategy	DD_UPDATE or 1
COURSE_NAME	Trim the trailing spaces and load the data	Expression	RTrim
DEPARTMENT_ID	Join with Department table, based on No from Department Table and load the corresponding Department ID	Joiner	Join Department table based on condition: Department_No(Course)=Department_No(Department)

Table 14: Course Dimension Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

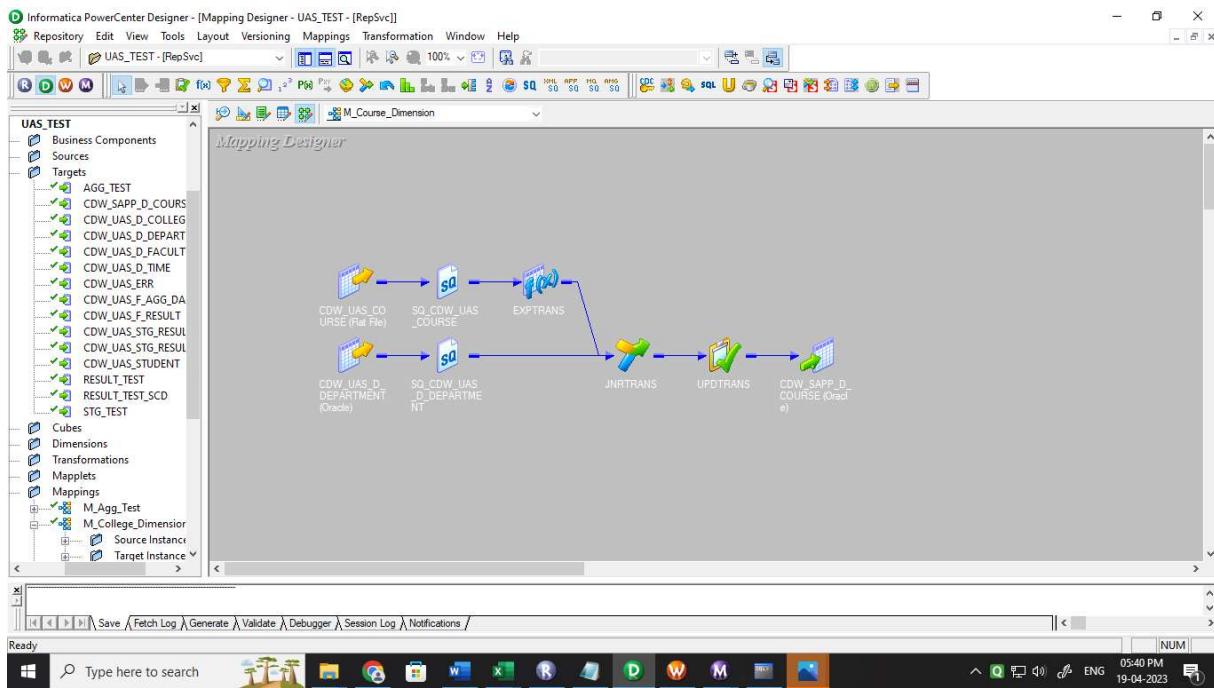


Fig. 20: Course Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

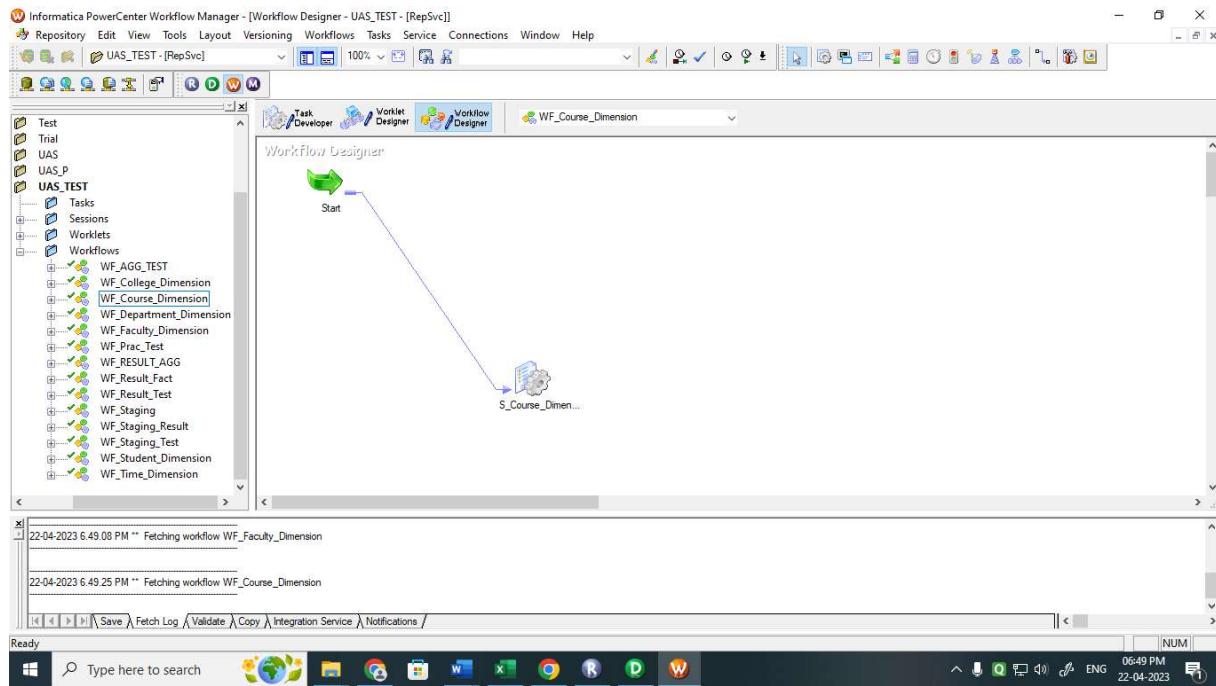


Fig. 21: Course Dimension Workflow

Target Post Execution:

```
Run SQL Command Line
SQL> SELECT * FROM CDW_sapp_D_course;
      COURSE_CODE COURSE_NAME          DEPARTMENT_ID MARK_REQ
-----  -----
2568303 ENGINEERING MECHANICS      1 40
2568404 ENGINEERING CHEMISTRY      2 35
2568305 COMPUTER GRAPHICS        3 35
2568204 ELECTRONICS & COMMUNICATION 4 40
2568401 ANALOG ELECTRONICS       5 40
2568100 FLUID MECHANICS          6 35
2568301 THEORY OF MACHINES       7 40
2568205 URBANS PLANNING           8 35
8 rows selected.
SQL>
```

Fig. 22: Course Dimension Target Table

Functional Requirement 6

TIME DIMENSION

SOURCE

The functional requirement was to load the data into the Time Table from a source flat file i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
SRC_Calender_File.txt	This is a file with 1 Input Fields.	 SRC_CALENDAR_FIL E.txt

Table 15: Time Dimension Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET TABLE
CDW_UAS_D_Time	It is a table which gets records from SRC_Calender_File and transforms the time key to different fields.	Dimension Table	 6_Time_Dimension.t xt

Table 16: Time Target Description

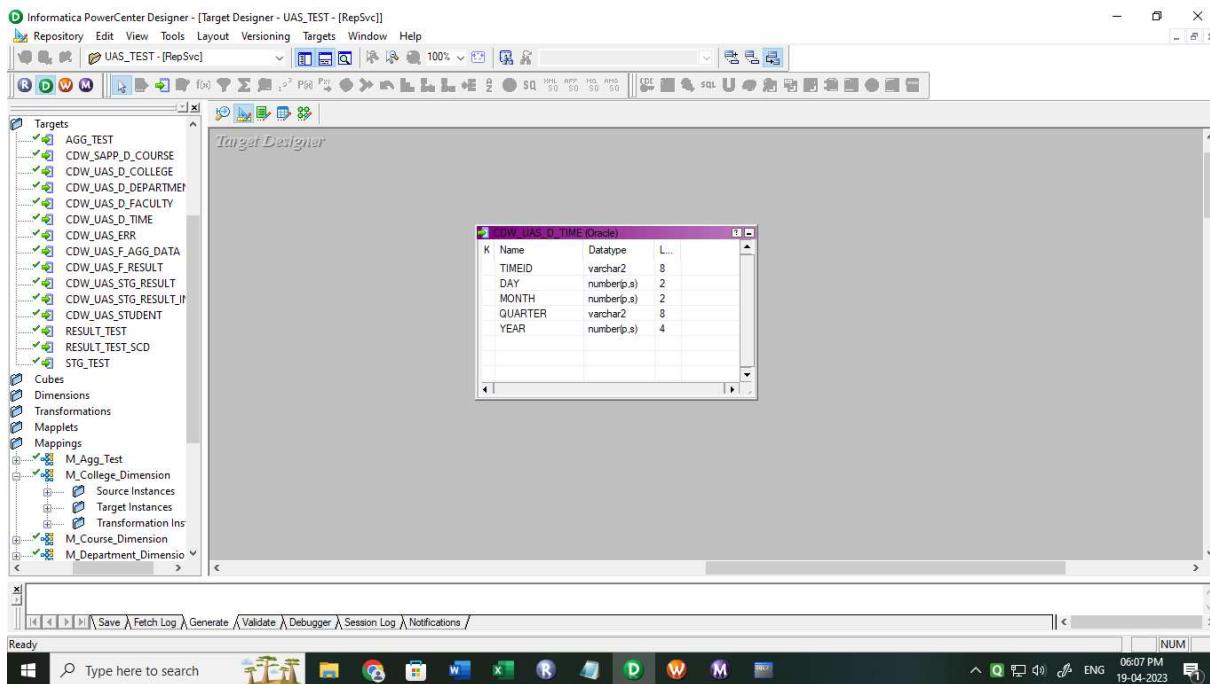


Fig 23: Time Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Used
DAY	Get the date part of TIME_ID	Expression	GET_DATE_PART, TO_DATE
MONTH	Get the Month part of TIME_ID	Expression	GET_DATE_PART, TO_DATE
QUARTER	Calculate using the MONTH column.	Expression	GET_DATE_PART, TO_DATE
YEAR	Get the year part of TIME_ID	Expression	GET_DATE_PART, TO_DATE

Table 17: Time Dimension Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

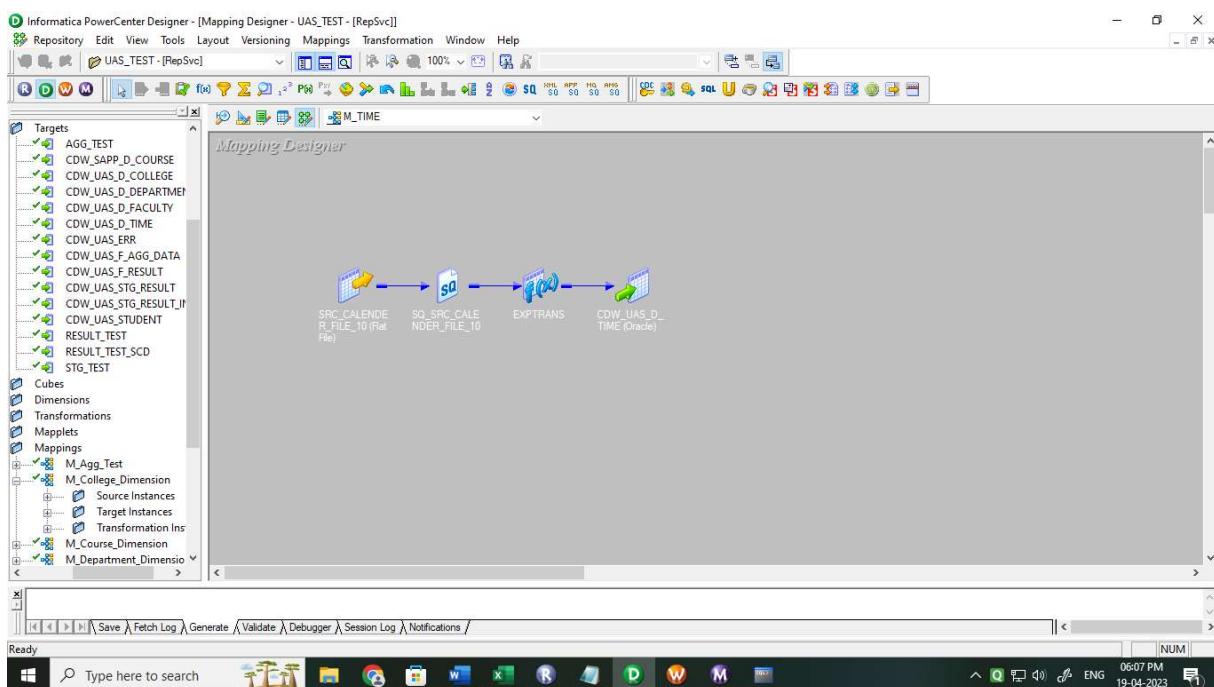


Fig. 24: Time Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

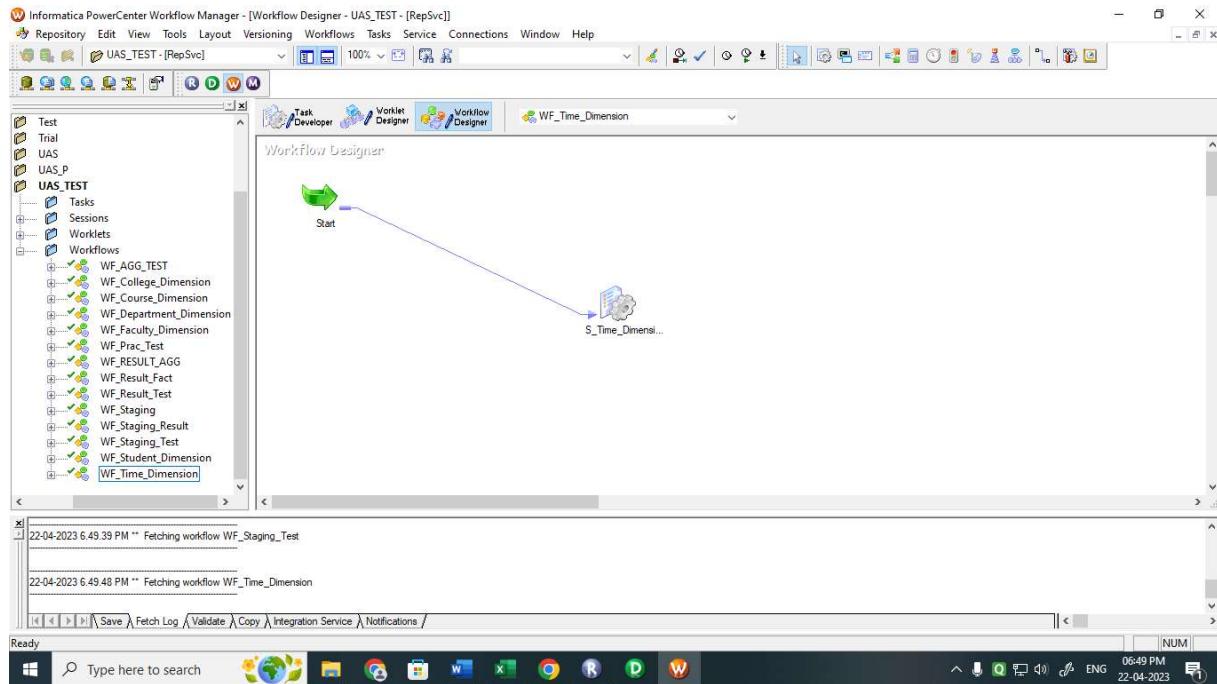


Fig. 25: Time Dimension Workflow

Target Post Execution:

```
Run SQL Command Line
SQL> SELECT * FROM CDM_uas_D_time;
TIMEID      DAY     MONTH QUARTER      YEAR
-----  -----  -----  -----  -----
12-01-23      12      1   QUART-1    2023
24-01-23      24      1   QUART-1    2023
10-02-23      10      2   QUART-1    2023
19-02-23      19      2   QUART-1    2023
28-02-23      28      2   QUART-1    2023
07-03-23       7      3   QUART-1    2023
16-03-23      16      3   QUART-1    2023
24-03-23      24      3   QUART-1    2023
12-04-23      12      4   QUART-2    2023
19-04-22      19      4   QUART-2    2022
30-04-23      30      4   QUART-2    2023
11 rows selected.
SQL>
```

Fig. 26: Time Dimension Target Table

Functional Requirement 7

RESULT STAGING

SOURCE

The functional requirement was to load the data into the Staging Table from 3 source files i.e.

SOURCE FILE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_F_RESULTS_BR_1.txt	This is a Comma delimited file with 9 Input Fields.	 CDW_UAS_F_RESULTS_BR_1.txt
CDW_UAS_F_RESULTS_BR_2.txt	This is a Comma delimited file with 9 Input Fields.	 CDW_UAS_F_RESULTS_BR_2.txt
CDW_UAS_F_RESULTS_BR_3.txt	This is a Fixed Width file with 9 Input Fields.	 CDW_UAS_F_RESULTS_BR_3.txt

Table 18: Staging Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_Stg_Result	It is a table which gets records from the three input files and lookup from different dimensions.	Staging Table	 7_1_Stg_Result.txt
CDW_UAS_ERR	It is a table which maintains the error records of the staging area.	Error Table	 7_2_Stg_Err.txt

Table 19: Staging Target Description

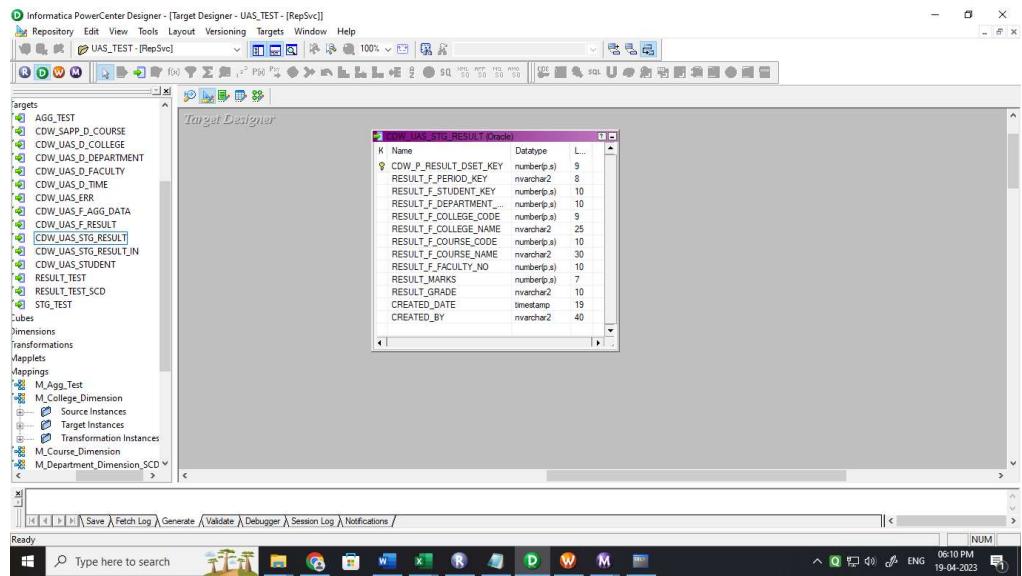


Fig 27: Staging Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Used
CDW_P_RESULT_DSET_KEY	seq NUMBER	Sequence Generator	
RESULT_F_PERIOD_KEY	Look up from Time dim table and load the TIMEID	Look Up	
RESULT_F_PERIOD_KEY	Look up from Time dim table and load the TIMEID	Look Up	
RESULT_F_PERIOD_KEY	Look up from Time dim table and load the TIMEID	Look Up	
RESULT_F_Student_KEY	look up from Student dim table.	Look Up	
RESULT_F_DEPARTMENT_KEY	look up from Department dim table	Look Up	
RESULT_F_COLLEGE_CODE	Look up College Code using College name from the College Dimension Table	Look Up	
RESULT_F_COURSE_CODE	Look up Course Code using Course Name from the Course Dimension Table	Look Up	
RESULT_F_FACULTY_NO	Look up Faculty Code using Faculty Name from the Faculty Dimension Table	Look Up	
RESULT_GRADE	Need to be calculated from the Course table data (If the Result Marks is greater than the Marks Required then PASS else FAIL).	Expression	IIF
CREATED_DATE	sysdate	Expression	SYSDATE
CREATED_BY	workflow name	Expression	\$PMWorkflowName

Table 20: Result Staging Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

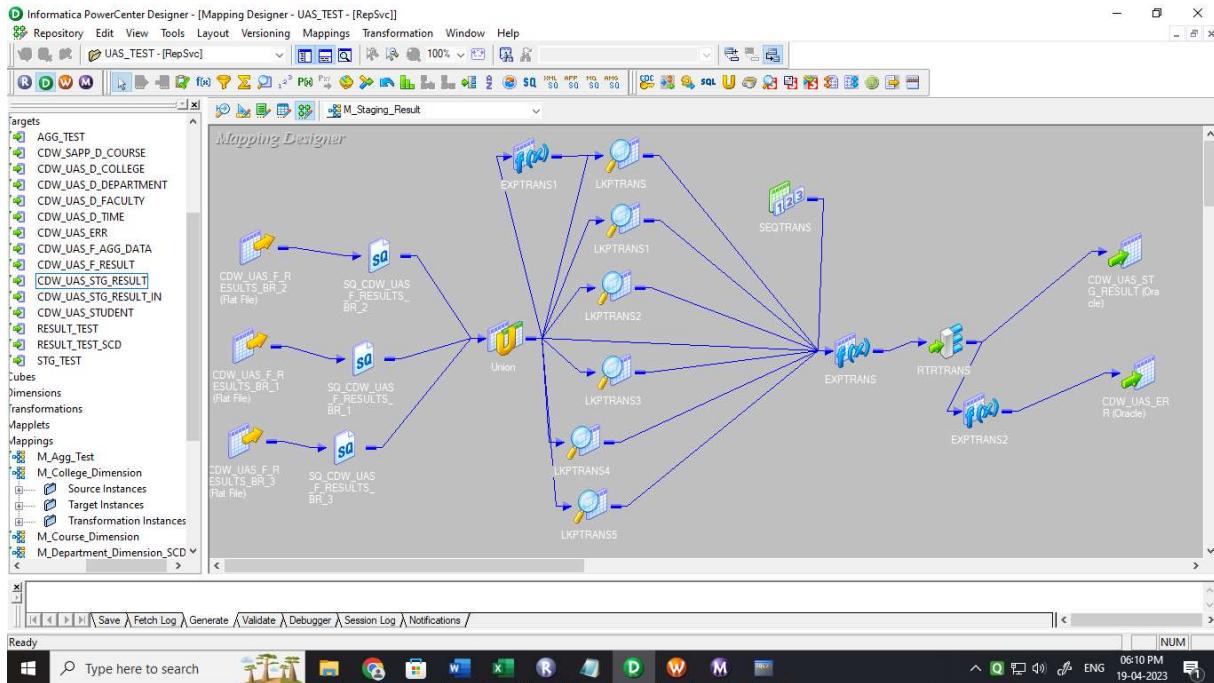


Fig. 28: Staging Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

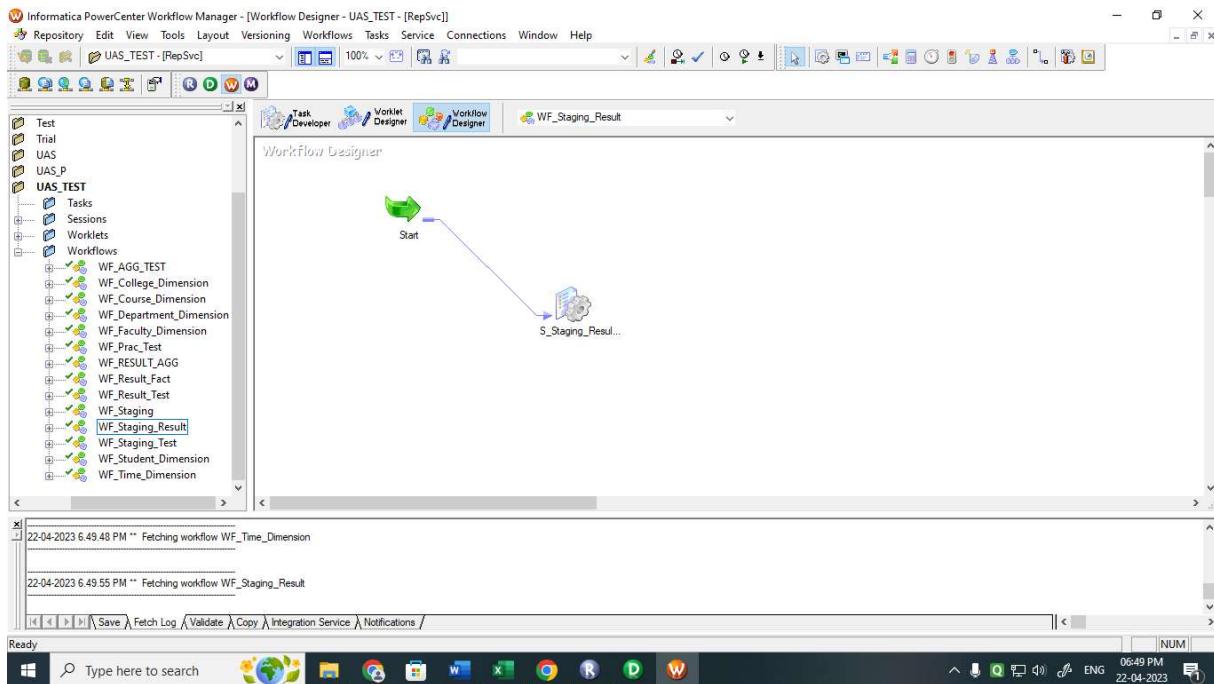


Fig. 29: Result Staging Workflow

Target Post Execution:

Run SQL Command Line

```
SQL> SELECT * FROM CDW_uas_stg_result;
```

					RESULT_F_COURSE_CODE	RESULT_F_COURSE_NAME	
							CREATED_BY
1	30-04-23	183112001	90 PASS	19-APR-23 06.11.56 PM	223 SAE Kondhwa	2568303 ENGINEERING MECHANICS	WF_Staging_Result
2	30-04-23	183122002	91 PASS	19-APR-23 06.11.56 PM	77 People Education Society	2568404 ENGINEERING CHEMISTRY	WF_Staging_Result
3	30-04-23	183112001	66 PASS	19-APR-23 06.11.56 PM	25 TKM CE	2568303 ENGINEERING MECHANICS	WF_Staging_Result
4	12-04-23	183182004	38 FAIL	19-APR-23 06.11.56 PM	52 SRM IST	2568204 ELECTRONICS & COMMUNICA	WF_Staging_Result
5	10-02-23		28 FAIL	19-APR-23 06.11.56 PM	SREENIDHI	COMPUTER GRAPHICS	WF_Staging_Result
7	30-04-23	183152006	34 FAIL	19-APR-23 06.11.56 PM	25 TKM CE	2568401 ANALOG ELECTRONICS	WF_Staging_Result
8	30-04-23	183162005	55 PASS	19-APR-23 06.11.56 PM	38 Sir Agnel	2568100 FLUID MECHANICS	WF_Staging_Result

7 rows selected.

SQL>

Fig. 30: Staging Result Target Table

Run SQL Command Line

```
SQL> SELECT * FROM CDW_uas_err;
```

					RESULT_F_COURSE_CODE	RESULT_F_COURSE_NAME	
							CREATED_BY
1	30-04-23	183112001	90 PASS	19-APR-23 06.11.56 PM	223 SAE Kondhwa	2568303 ENGINEERING MECHANICS	WF_Staging_Result
2	30-04-23	183122002	91 PASS	19-APR-23 06.11.56 PM	77 People Education Society	2568404 ENGINEERING CHEMISTRY	WF_Staging_Result
3	30-04-23	183112001	66 PASS	19-APR-23 06.11.56 PM	25 TKM CE	2568303 ENGINEERING MECHANICS	WF_Staging_Result
4	12-04-23	183182004	38 FAIL	19-APR-23 06.11.56 PM	52 SRM IST	2568204 ELECTRONICS & COMMUNICA	WF_Staging_Result
5	10-02-23		28 FAIL	19-APR-23 06.11.56 PM	SREENIDHI	COMPUTER GRAPHICS	WF_Staging_Result
7	30-04-23	183152006	34 FAIL	19-APR-23 06.11.56 PM	25 TKM CE	2568401 ANALOG ELECTRONICS	WF_Staging_Result
8	30-04-23	183162005	55 PASS	19-APR-23 06.11.56 PM	38 Sir Agnel	2568100 FLUID MECHANICS	WF_Staging_Result

7 rows selected.

```
SQL> SELECT * FROM CDW_uas_err;
```

					ERR_COLUMN_NM
					RESULT_F_PERIOD_KEY
			6	6 INVALID DATE	

SQL>

Fig. 31: Staging Error Target Table

Functional Requirement 8

RESULT FACT

SOURCE

The functional requirement was to load the data into the Result Fact Table from Result staging table i.e.

SOURCE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_Stg_Result	This table contains the records from staging area.	

Table 21: Result Fact Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_F_Result	It is a Fact table which gets all records from the Staging table and transforms the data type of Period Key to date type.	Fact Table	 8_Fact_Result.txt

Table 22: Result Fact Target Description

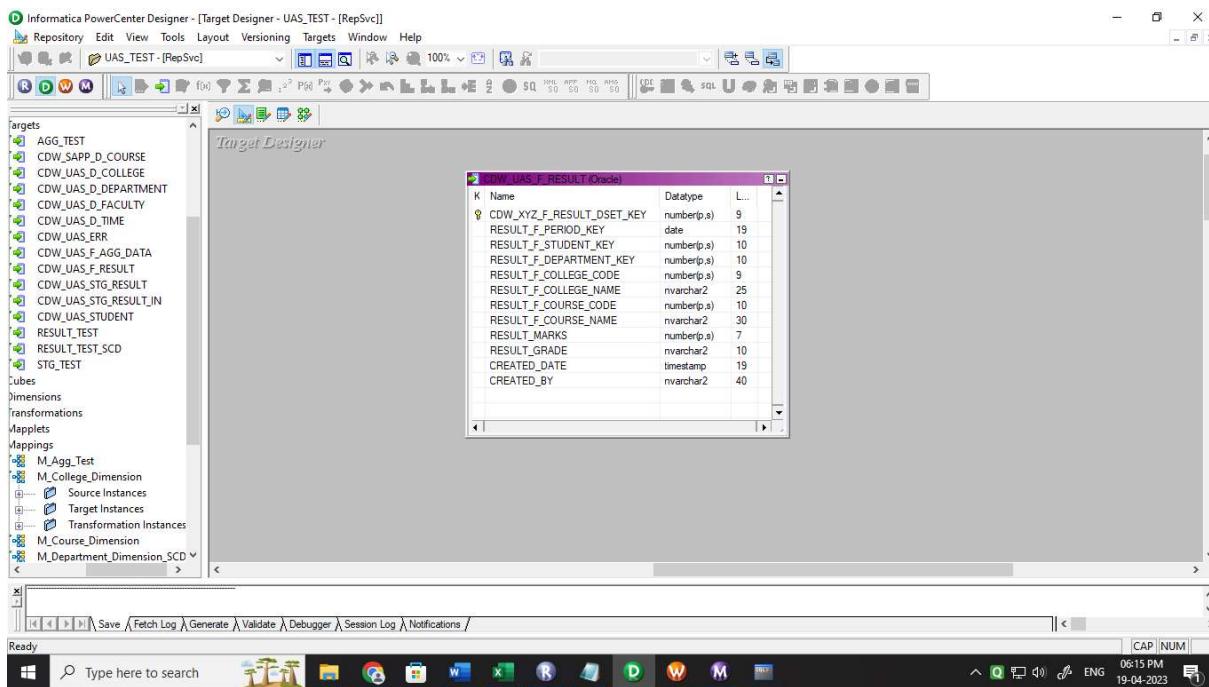


Fig 32: Result Fact Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Name
RESULT_F_PERIOD_KEY	Direct Move (Convert to date datatype)	Expression	ToDate

Table 23: Result Fact Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

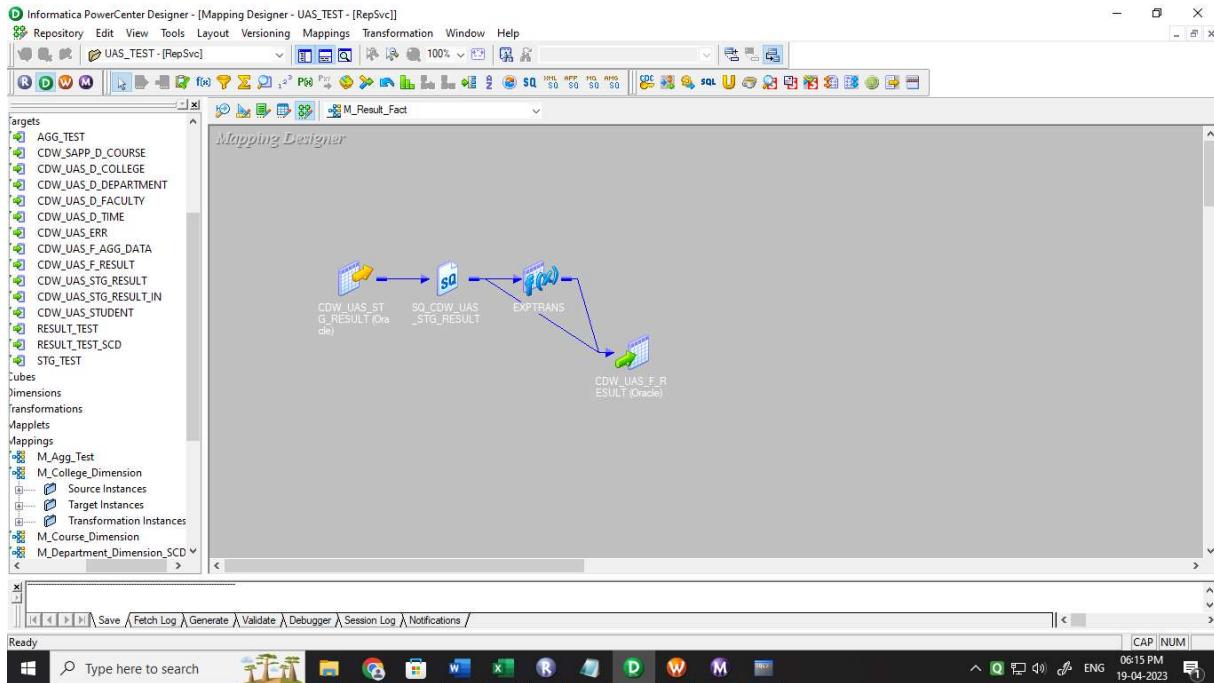


Fig. 33: Result Fact Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

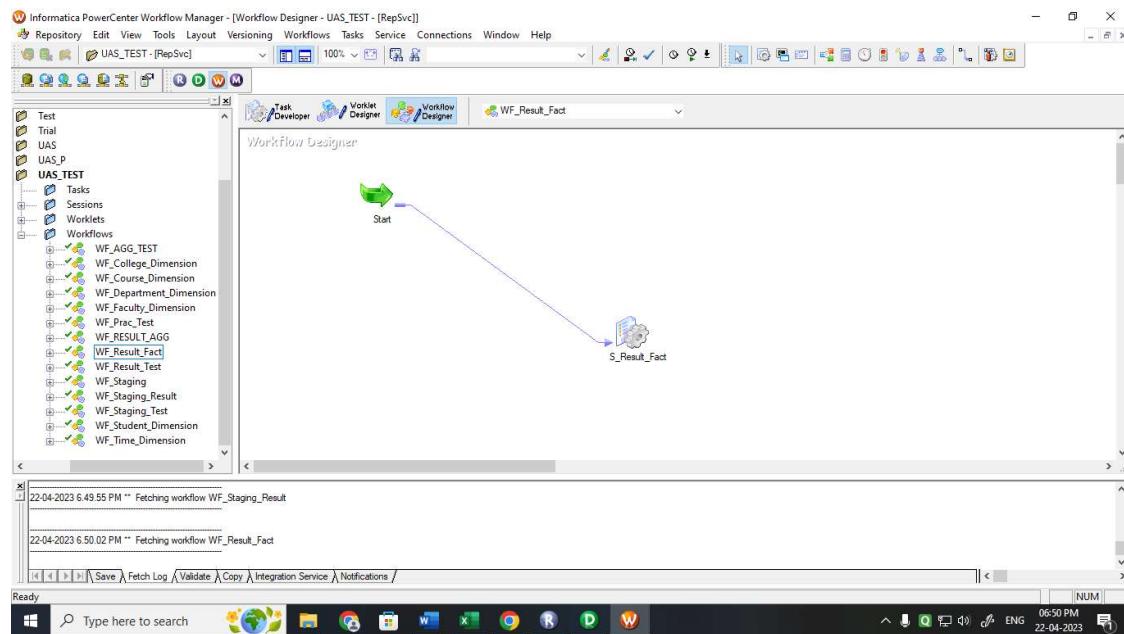


Fig. 34: Result Fact Workflow

Target Post Execution:

The screenshot shows the Informatica PowerCenter Designer interface. A "Preview Data" dialog box is open, connected to the "UAS_TEST" database via ODBC. The dialog shows data for the table "CDW_UAS_F_RESULT". The data is as follows:

CDW_XYZ...	RESULT_F...	RESULT_F...	RESULT_F...	RESULT_F...	RESULT_F...	RESULT_F...	RESULT_F...	RESULT_M...	RESULT_G...	CREATED...	CREATED...
1	4/30/2023	1	1	223	SAE Kond...	2568303	ENGINEER...	90	PASS	4/19/2023...	WF_Stagi...
2	4/30/2023	2	2	77	People Ed...	2568404	ENGINEER...	91	PASS	4/19/2023...	WF_Stagi...
3	4/30/2023	4	1	25	TMM CE	2568303	ENGINEER...	66	PASS	4/19/2023...	WF_Stagi...
4	4/12/2023	7	4	52	SRM IST	2568204	ELECTRO...	38	FAIL	4/19/2023...	WF_Stagi...
5	2/10/2023	3	3	NULL	SREENIDH...	NULL	COMPUTE...	28	FAIL	4/19/2023...	WF_Stagi...
7	4/30/2023	5	5	25	TMM CE	2568401	ANALOG E...	34	FAIL	4/19/2023...	WF_Stagi...
8	4/30/2023	6	6	38	Sir Agnel	2568100	FLUID ME...	35	PASS	4/19/2023...	WF_Stagi...

Fig. 35: Result Fact Target Table

Functional Requirement 9

RESULT AGGREGATE

SOURCE

The functional requirement was to transform and then load the data into the Result Aggregate Table from Result fact table i.e.

SOURCE NAME	DESCRIPTION	SOURCE FILE
CDW_UAS_F_Result	It is the result fact table containing the marks and result_grade.	

Table 24: Result Aggregate Source Description

TARGET

The target table has been generated by creating a schema from the data provided in the requirements document.

TARGET NAME	DESCRIPTION	TARGET TYPE	TARGET SCHEMA
CDW_UAS_F_AGG_Data	It is a Aggregate table which gets the data from the Fact table which is aggregated as per the business requirements before loading into aggregate table.	Fact Table	 9_Agg_Result.txt

Table 25: Result Aggregate Target Description

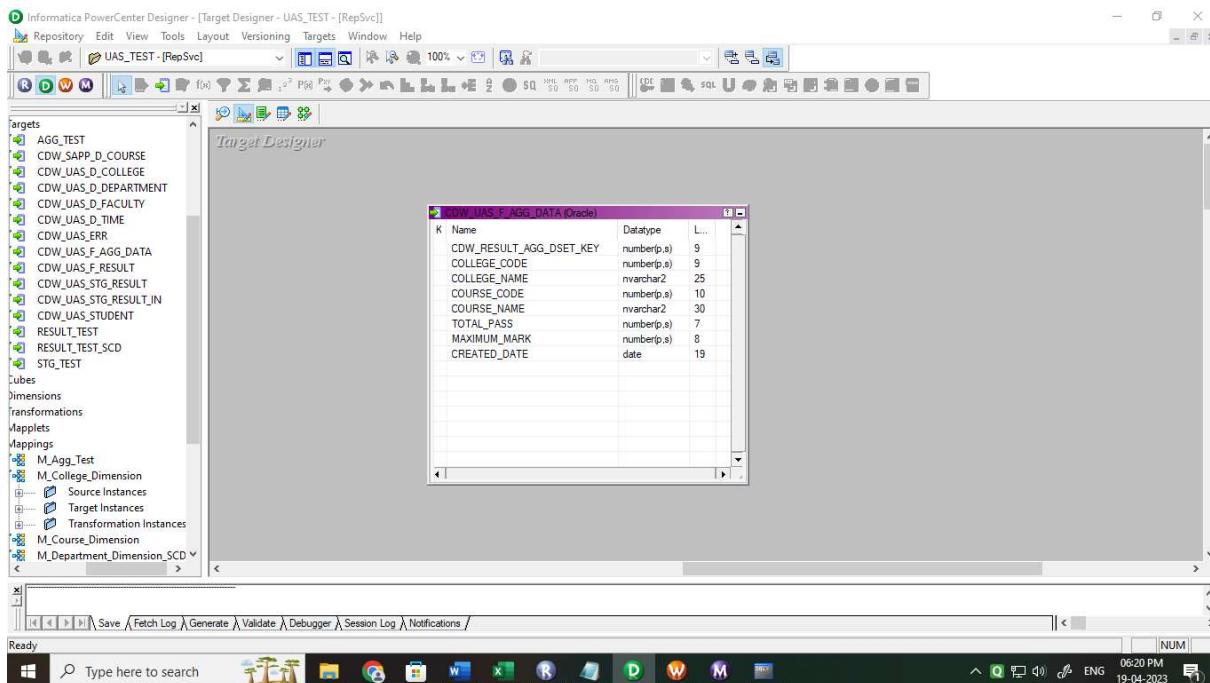


Fig 36: Result Aggregate Target Definition

TRANSFORMATIONS:

Column Name	Mapping Logic	Transformation Name	Function Used
CDW_RESULT_AGG_DSET_KEY	Sequence generated number	Sequence Generator	
TOTAL PASS	Count of the total students passes in a particular Course evaluated in each College for the current reporting period.	Aggregator	Count()
MAXIMUM_MARK	Maximum mark scored in a particular Course in each college for the current reporting period.	Aggregator	Max()
CREATED_DATE	Sysdate	Aggregator	SYSDATE

Table 26: Result Aggregate Transformations

MAPPING

The below figure shows the data flow from source to the target table by applying the necessary transformations to satisfy the business requirement.

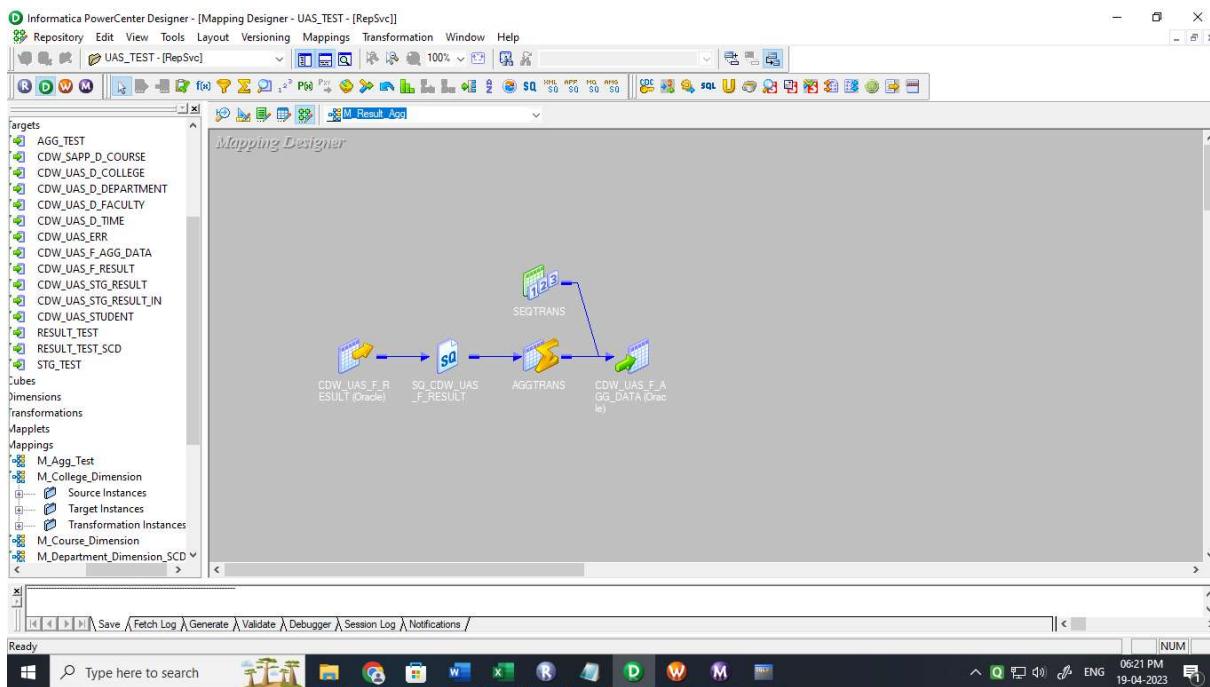


Fig. 37: Result Aggregate Mapping

Workflow

Below Figure shows the workflow which executes the mappings created in the designer.

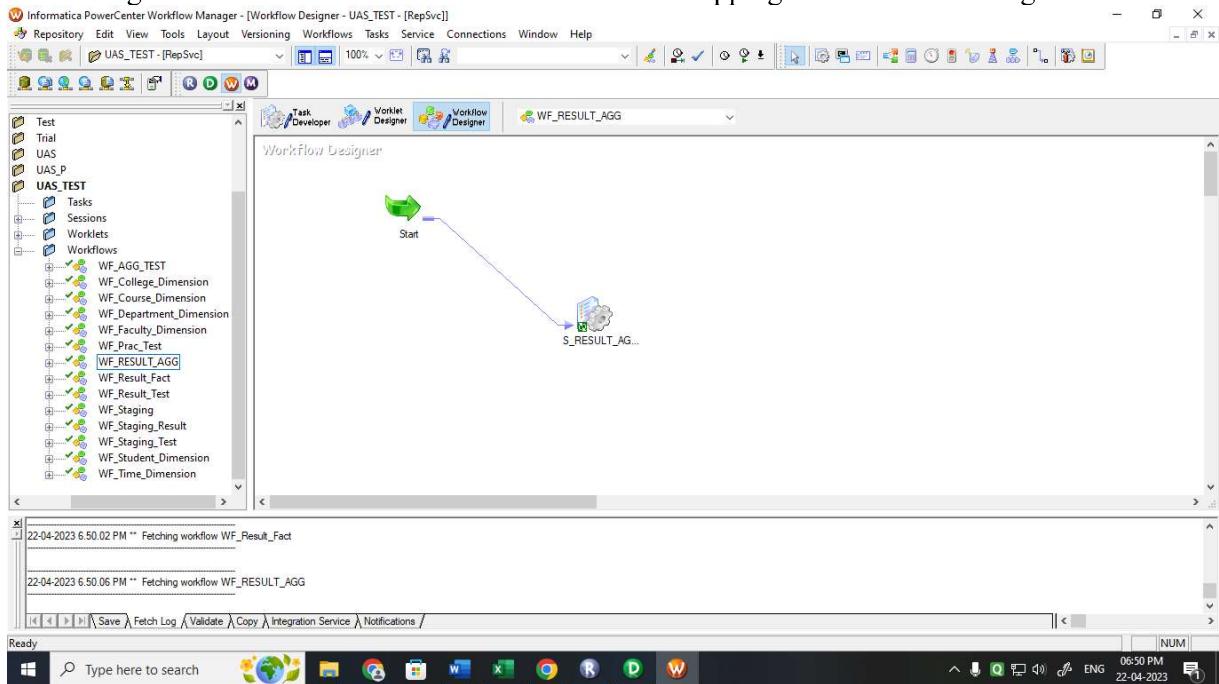


Fig. 38: Result Aggregate Workflow

Target Post Execution:

CDW_RESULT_AGG_DSET_KEY	COLLEGE_CODE	COLLEGE_NAME	COURSE_CODE	COURSE_NAME	TOTAL_PASS	MAXIMUM_MARK	CREATED_DATE
1	25	TKM CE	2568303	ENGINEERING MECHANICS	1	66	19-APR-23
2	25	TKM CE	2568401	ANALOG ELECTRONICS	0	34	19-APR-23
3	38	Sir Agnel	2568100	FLUID MECHANICS	1	55	19-APR-23
4	52	SRM IST	2568204	ELECTRONICS & COMMUNICATION	0	38	19-APR-23
5	77	People Education Society	2568404	ENGINEERING CHEMISTRY	1	91	19-APR-23
6	223	SAE Kondhwa	2568303	ENGINEERING MECHANICS	1	96	19-APR-23

Fig. 39: Result Aggregate Target Table

INTEGRATED WORKFLOW

To run the UAS the workflows have to be combined together to execute them as per the requirements. This can be achieved either by creating worklets or reusable tasks. The worklets or tasks are then combined and conditions are given for them to execute one after the other or based on success of previous tasks.

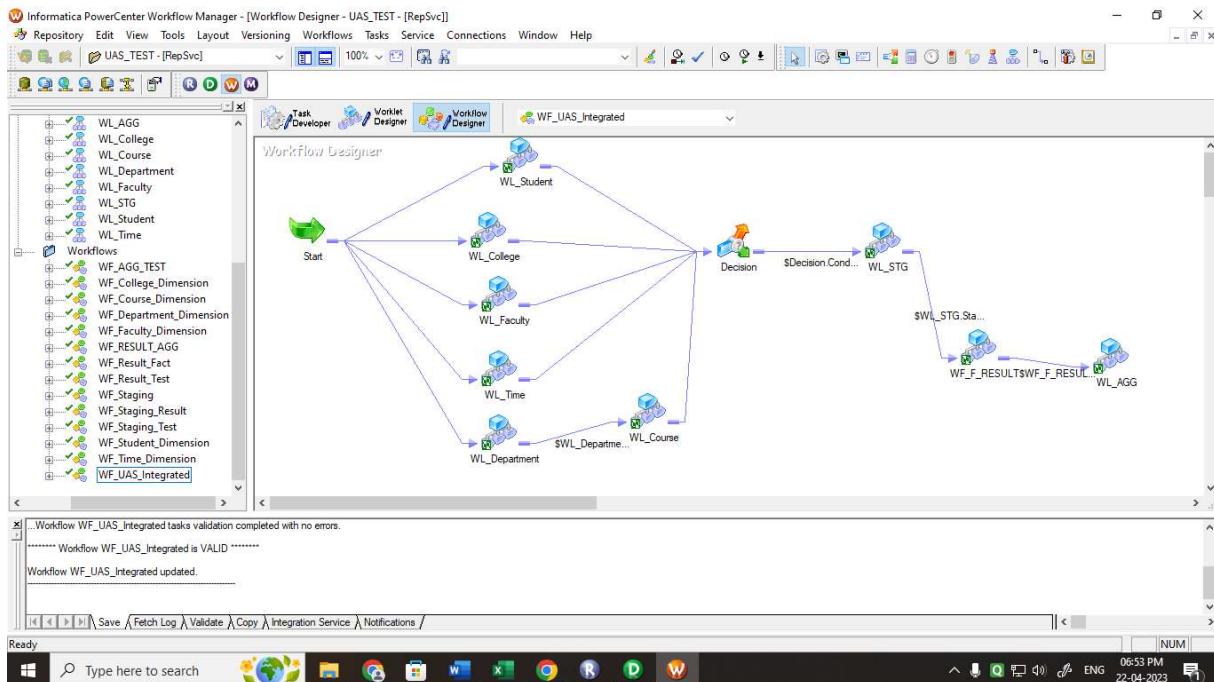


Fig. 40: Integrated Workflow

TASK VIEW AND SESSION STATISTICS

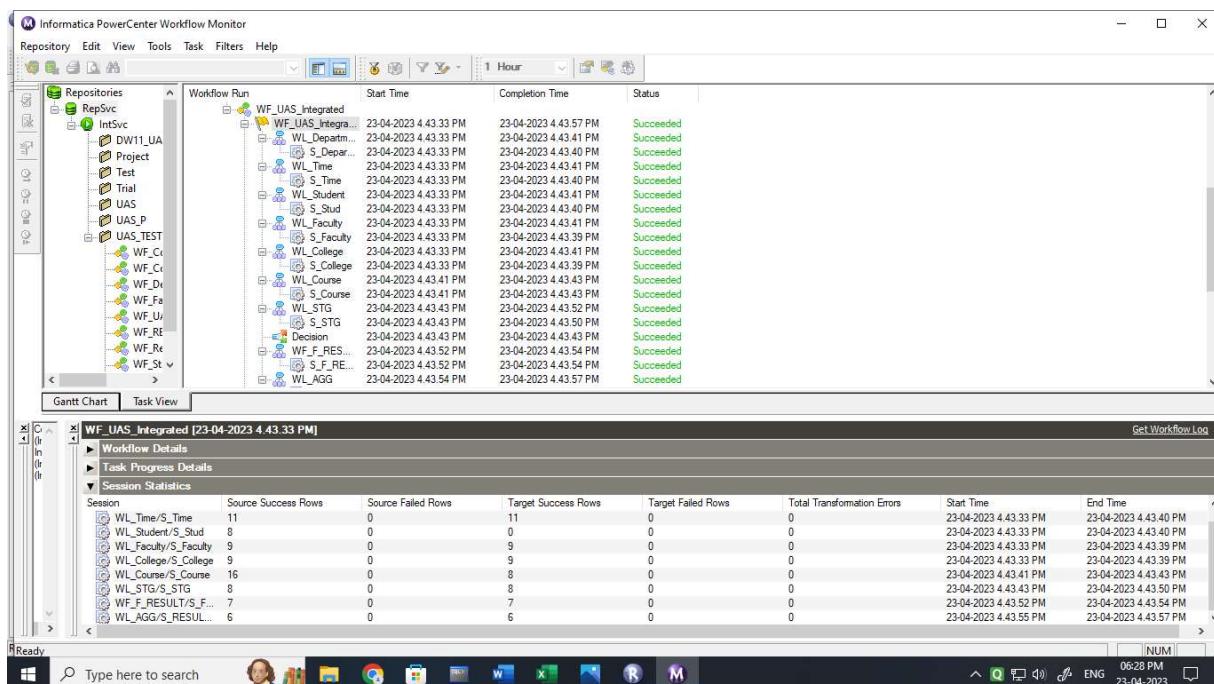


Fig. 41: Task Details

DESIGN CONSTRAINTS

CONSTRAINT	SUCCESS	IMPLEMENTATION
Course table should get loaded only when a valid department exist for the course in department dimension table.	Yes	Same column name (Department_No) having matching values in both Course and Department tables.
A student record getting loaded in to the fact table should have corresponding entry in student dimension table.	Yes	Student_Id(Surrogate key) loaded to Fact Table.
Date of result getting loaded in to the fact table should have corresponding entry in Time dimension table.	Yes	Time_Key loaded to the Fact table.
The ETL data flow schedule should accommodate the order of dimension table loading before the fact table load.	Yes	Created Integrated workflow and scheduled the workflows of dimensions in parallel and the loading the fact table only when all dimensions are succeeded.

Table 27: Design Constraints

CONCLUSION

All business requirements were taken into account while developing the "**University Analytics System**". By creating 10 target tables viz. Six dimension tables, Two Staging tables (Stg_Result and Error table) a Fact table and an Aggregate table, the data was loaded into the dimensions and passed through the staging layer and subsequently it was loaded into the Fact table. Aggregation was performed on the data from the Fact table and the it was loaded into the aggregate table. By successful loading of data and testing the given restrictions, the project's main goals have been accomplished. The system has been operated only in accordance with the functional specifications outlined in the system requirements document. The system was thus successful in implementing all the requirements and can be executed to be used as "University Analytics System".