

**damilare m. jinadu**

**P2733425**



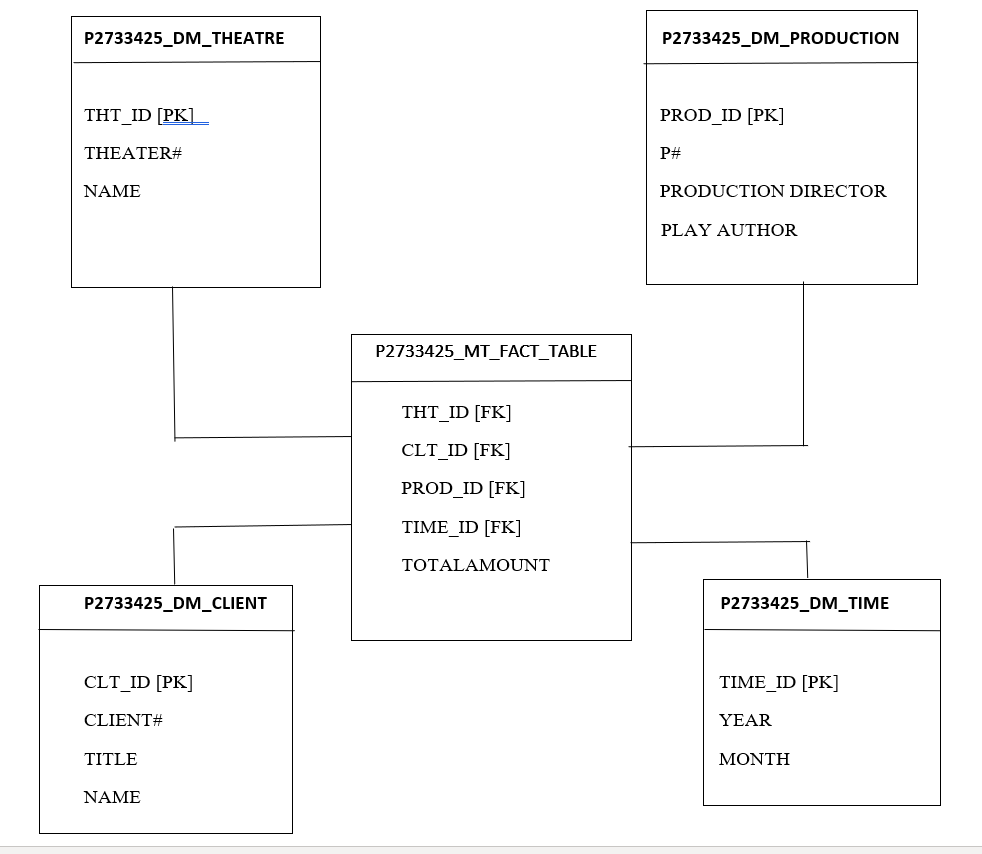
December 9, 2022

**DATA WAREHOUSE DESIGN AND OLAP**

**IMAT 5167**

**ASSESSMENT TITLE: DATA MART DESIGN**

**QUESTION 1**: **STAR SCHEMA DESIGN**

**1.1**

***Fig 1:* *Star Schema Design of MT Company Data Mart.***

**1.2**

1. The fact and dimensions and why they are chosen

The minimum dimension(s) and attribute(s) for MT Company data mart design after careful analysis of the given database design and data mart requirements are as stated below;

Dimensions:

1. P2733425\_DM\_THEATRE for Theatre.
2. P2733425\_DM\_PRODUCTION for Production.
3. P2733425\_DM\_CLIENT for Client.
4. P2733425\_DM\_TIME for Time.
5. P2733425\_MT\_FACT\_TABLE

The fact table garner data from the dimension tables

1. **Why the attributes are chosen**

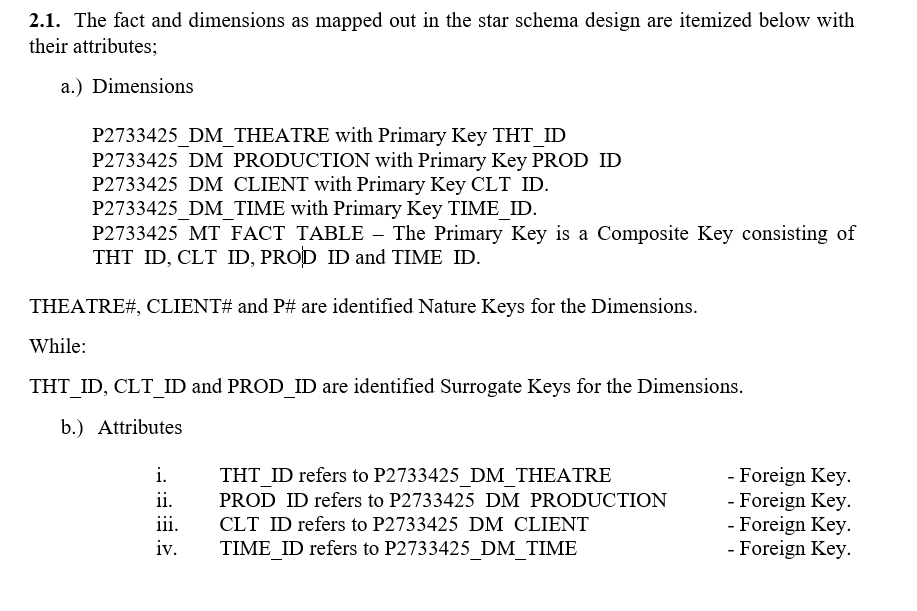
The attributes are chosen in order to ensure the Data Mart’s easy handling of prospective changes in OLTP due to changes in business requirements.

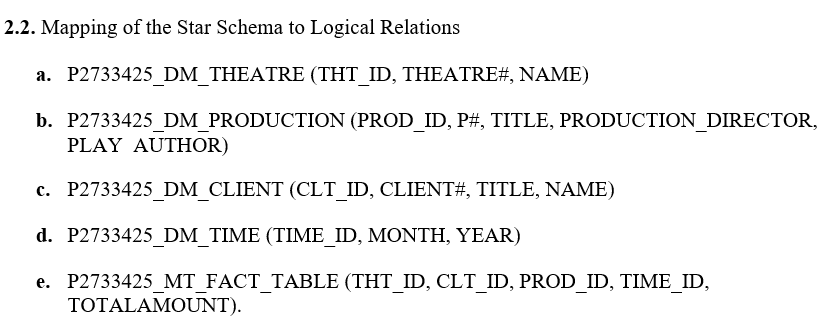
The primary keys from the Online Transaction Processing (OLTP) Database are represented on the Dimension tables as Nature Keys while the foreign keys represented on the Fact table are the primary keys from the Dimension tables.

1. **Granularity**

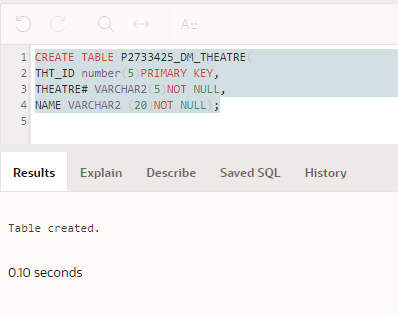
The granularity for each row of the Fact Table (i.e., P2733425\_MT\_FACT\_TABLE) is “The total amount of sales generated for each year, for each production, by one Theatre for one Client in a month”.

**QUESTION 2: DERIVATION OF LOGICAL RELATIONS**

****

The FACT TABLE contains only the fact (i.e., the Total amount of Sales and the associated foreign keys which represents the Primary Keys in the associated Dimension Tables which contains only required attributes from the given Database.

**QUESTION** **3: CREATION OF TABLES WITH INTEGRITY RULES**

**3.1 SQL code for each table creation:**

**a.** CREATE TABLE P2733425\_DM\_THEATRE (

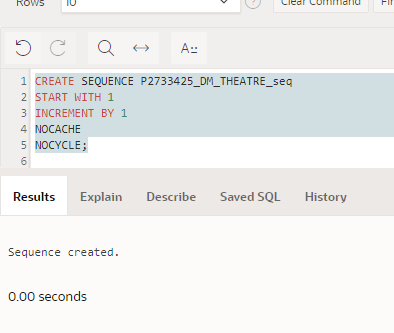
THT\_ID number (5) PRIMARY KEY,

THEATRE# VARCHAR2(5) NOT NULL,

NAME VARCHAR2 (20) NOT NULL);

***Fig 2: P2733425\_DM\_THEATRE TABLE***

b. CREATE SEQUENCE P2733425\_DM\_THEATRE\_seq

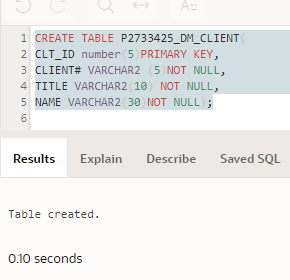
START WITH 1

INCREMENT BY 1

NOCACHE

NOCYCLE;

***Fig 3: P2733425\_DM\_THEATRE TABLE SEQUENCE***

1. CREATE TABLE P2733425\_DM\_CLIENT(

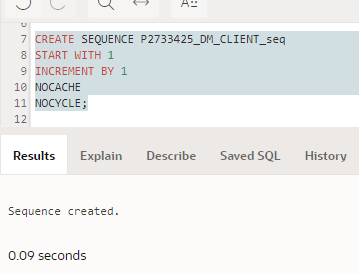
CLT\_ID number(5)PRIMARY KEY,

CLIENT# VARCHAR2 (5)NOT NULL,

TITLE VARCHAR2(10) NOT NULL,

NAME VARCHAR2(30)NOT NULL);

***Fig 4: P2733425\_DM\_CLIENT TABLE***

1. CREATE SEQUENCE P2733425\_DM\_CLIENT\_seq

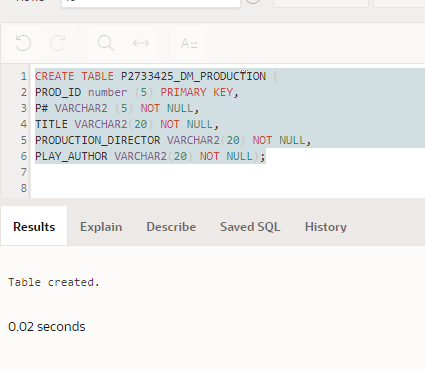
START WITH 1

INCREMENT BY 1

NOCACHE

NOCYCLE;

***Fig 5: P2733425\_DM\_CLIENT TABLE SEQUENCE***

1. CREATE TABLE P2733425\_DM\_PRODUCTION (

PROD\_ID number (5) PRIMARY KEY,

P# VARCHAR2 (5) NOT NULL,

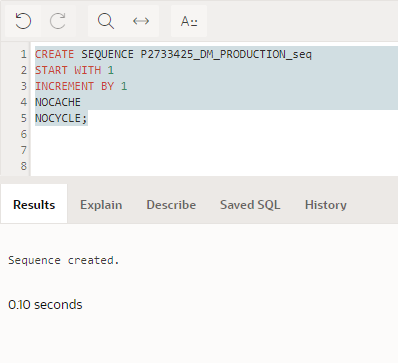
TITLE VARCHAR2(20) NOT NULL,

PRODUCTION\_DIRECTOR VARCHAR2(20) NOT NULL,

PLAY\_AUTHOR VARCHAR2 (20) NOT NULL);

***Fig 6: P2733425\_DM\_PRODUCTION TABLE***

1. CREATE SEQUENCE P2733425\_DM\_PRODUCTION\_seq

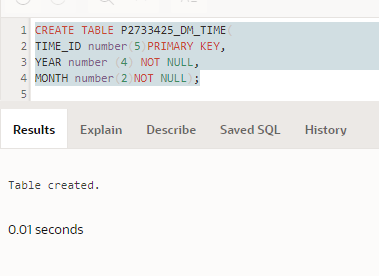
START WITH 1

INCREMENT BY 1

NOCACHE

NOCYCLE;

***Fig 7: P2733425\_DM\_PRODUCTION TABLE SEQUENCE***



1. CREATE TABLE P2733425\_DM\_TIME(

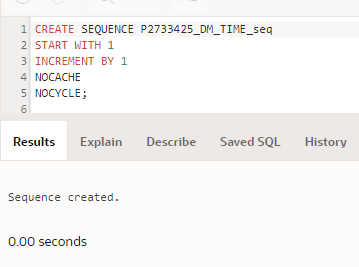
TIME\_ID number(5)PRIMARY KEY,

YEAR number (4) NOT NULL,

MONTH number(2)NOT NULL);

***Fig 8: P2733425\_DM\_TIME TABLE***

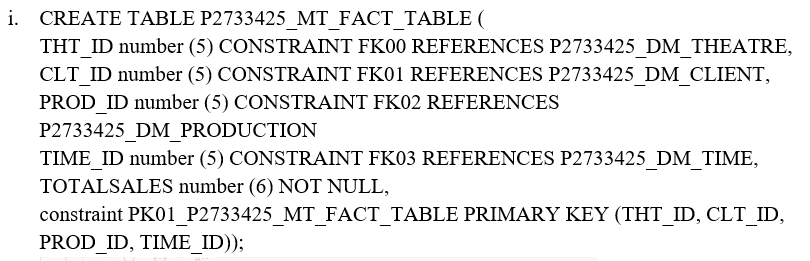
1. CREATE SEQUENCE P2733425\_DM\_TIME\_seq

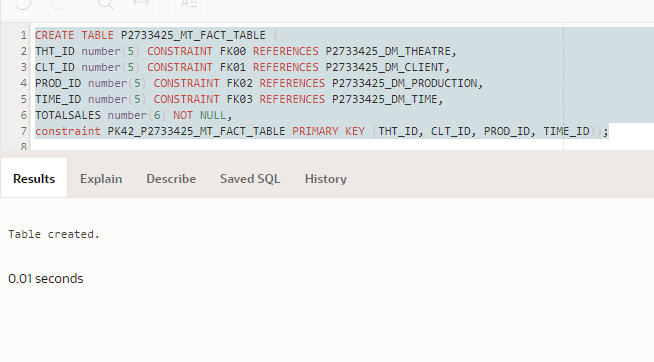
START WITH 1

INCREMENT BY 1

NOCACHE

NOCYCLE;

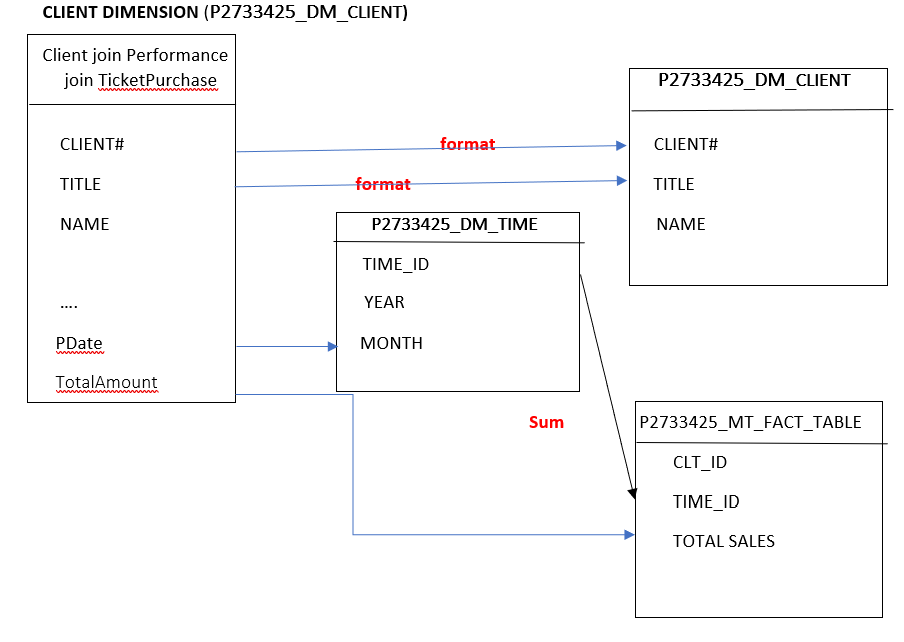
 ***Fig 9: P2733425\_DM\_TIME TABLE SEQUENCE***



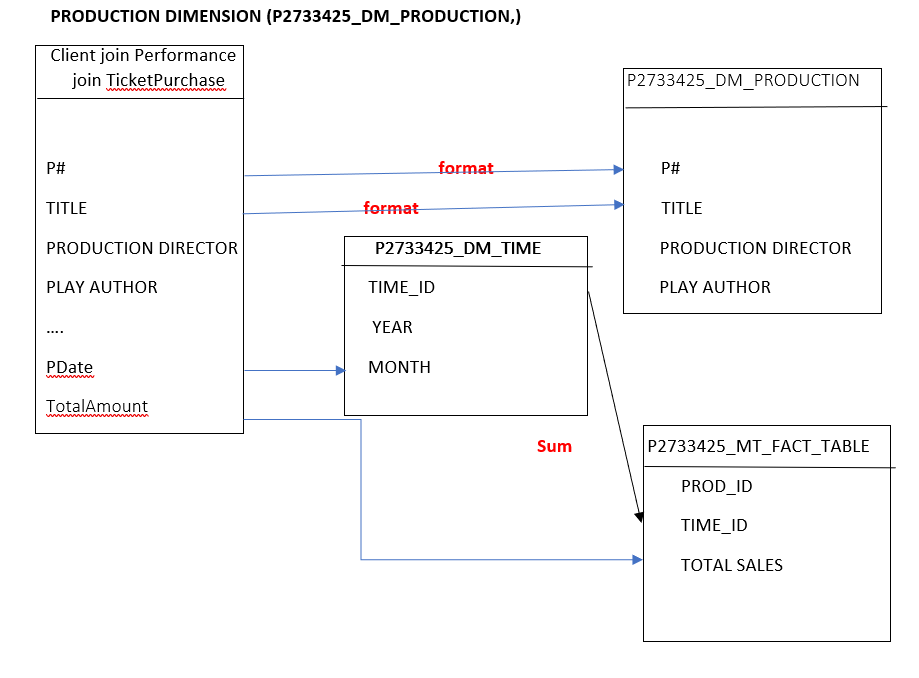
***Fig 10: P2733425\_MT\_FACT\_ TABLE***

**QUESTION 4: DATA SOURCE IDENTIFICATION, DATA EXTRACTION, TRANSFORMATION AND LOADING**

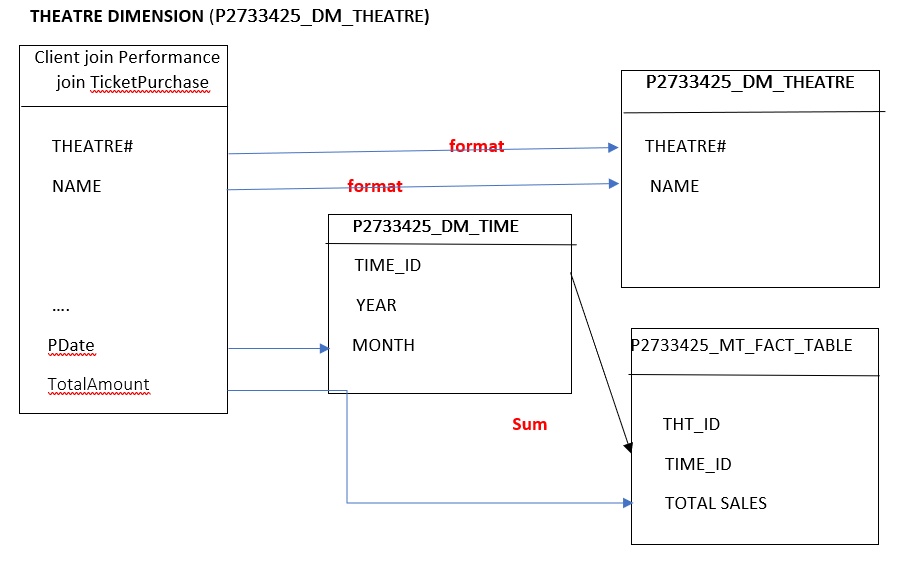
**4.1 DATA SOURCE MAPPING**

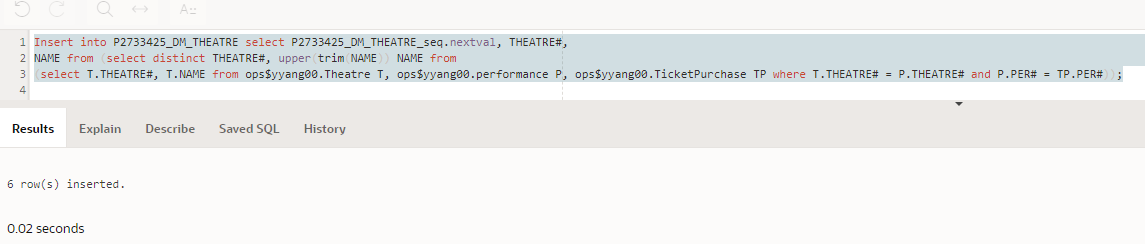
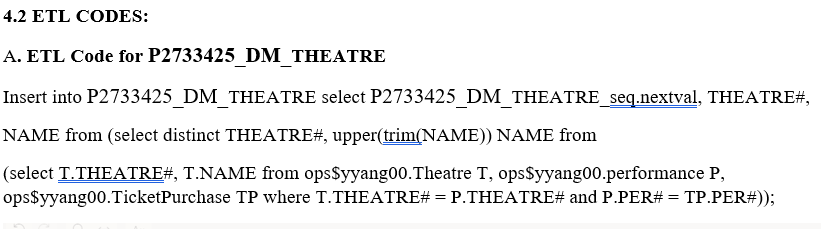


***Fig 11: P2733425\_DM\_CLIENT DATA SOURCE MAPPING***

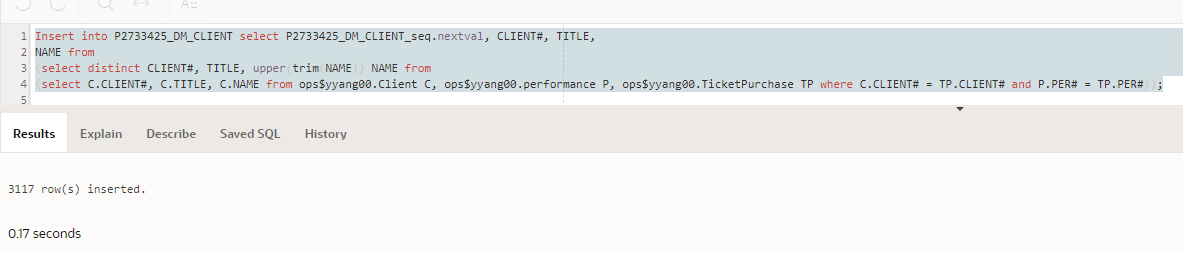
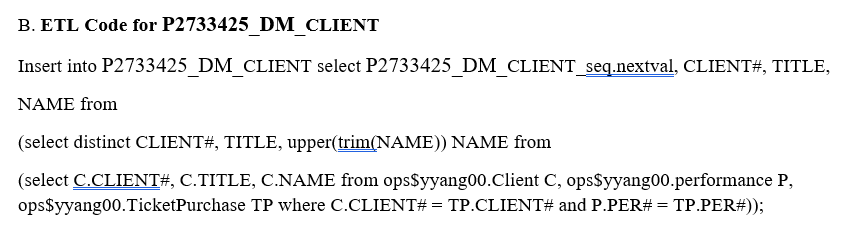
****

***Fig 12: P2733425\_DM\_PRODUCTION DATA SOURCE MAPPING***

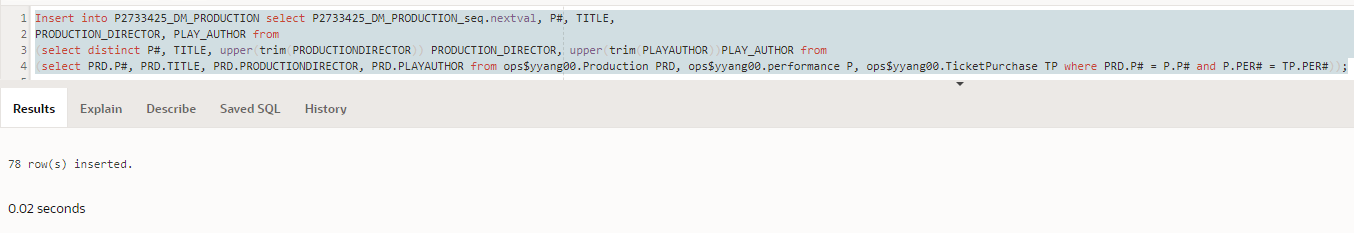
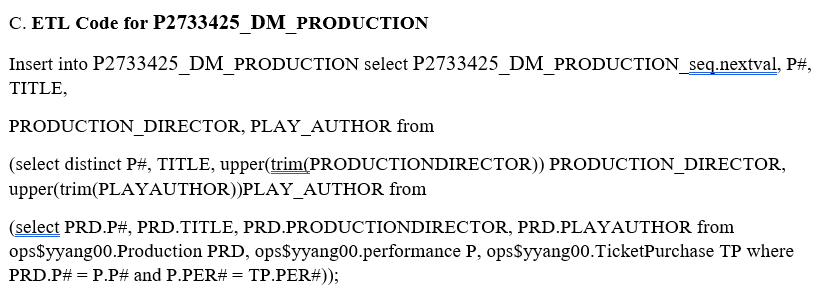
 ***Fig 13: P2733425\_DM\_THEATRE DATA SOURCE MAPPING***



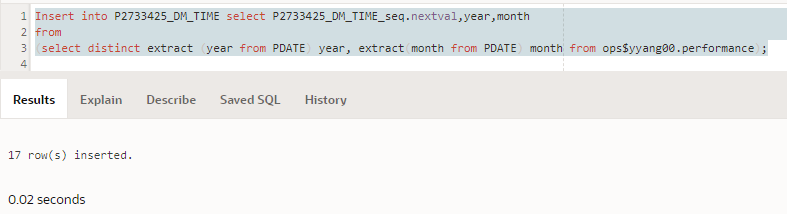
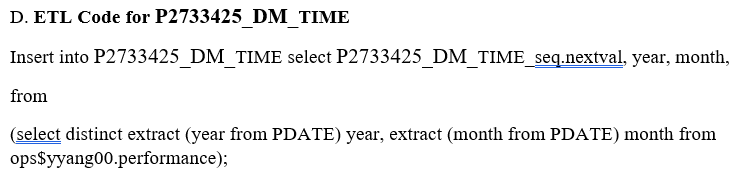
***Fig 14: EVIDENCE OF P2733425\_DM\_THEATRE ETL CODE***



***Fig 15: EVIDENCE OF P2733425\_DM\_CLIENT ETL CODE***

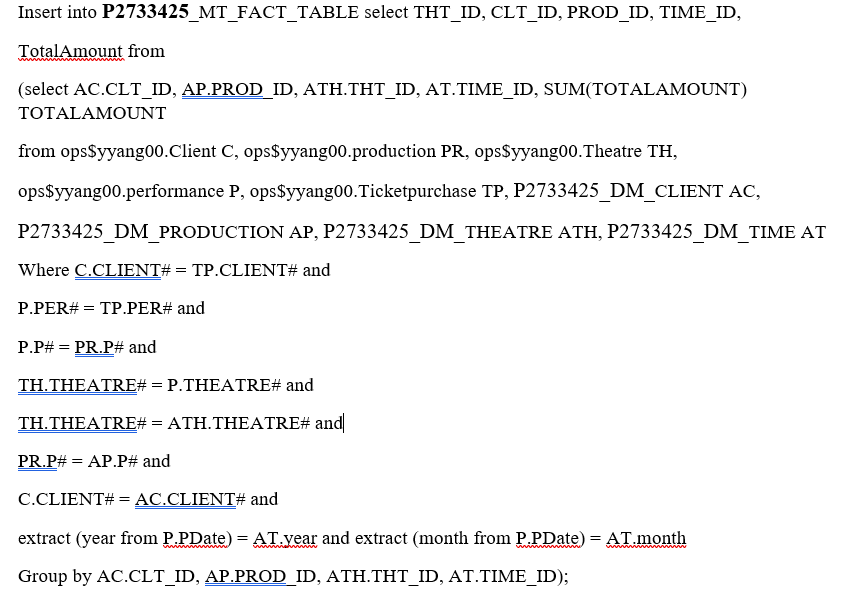


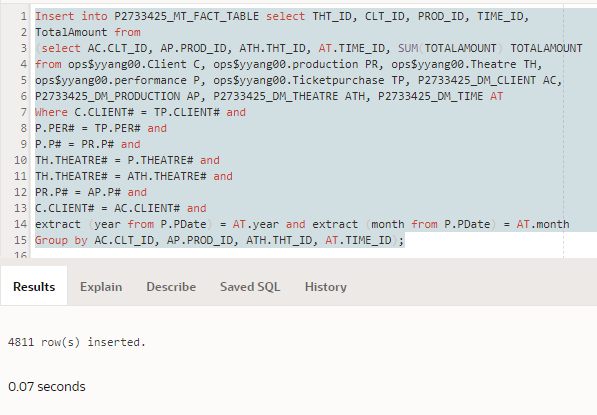
***Fig 16: EVIDENCE OF P2733425\_DM\_PRODUCTION ETL CODE***



***Fig 17: EVIDENCE OF P2733425\_DM\_TIME ETL CODE***

1. **ETL Code for P2733425\_MT\_FACT\_TABLE**





***Fig 18: EVIDENCE OF* P2733425\_MT\_FACT\_TABLE ETL CODE**

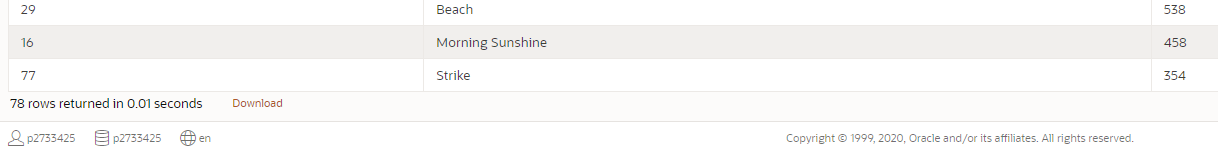
**QUESTION 5: JUSTIFICATON OF THE DATA MART DESIGN AND COMPARISON OF DATA MART AND OLTP**

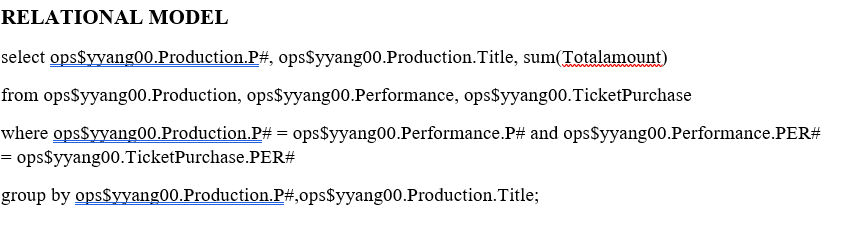
**SQL Codes for the required queries: DATA MART**

select PD.PROD\_ID, PD.Title, sum(TotalSales)

from P2733425\_DM\_PRODUCTION PD, P2733425\_MT\_FACT\_TABLE FT

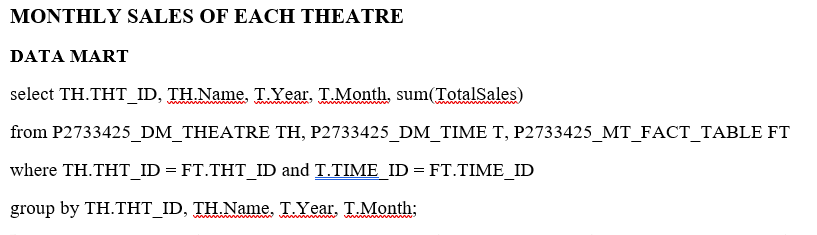
where PD.PROD\_ID = FT.PROD\_ID

group by PD.PROD\_ID, PD.Title; ****

***Fig 19: EVIDENCE OF SQL CODES FOR DATA MART***

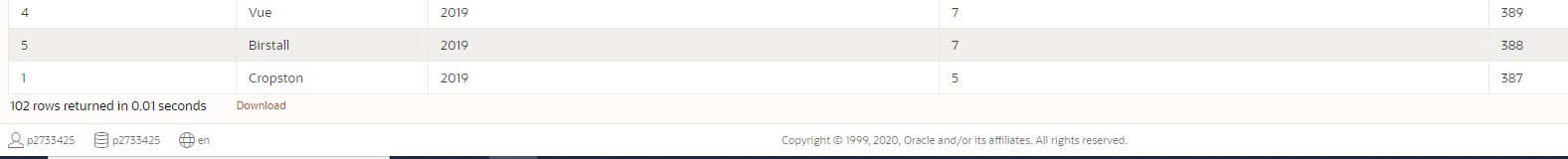
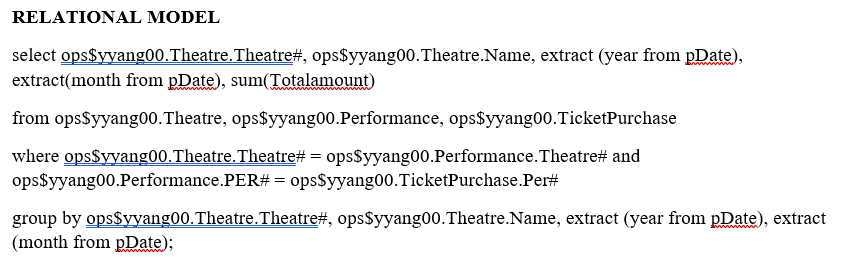


***Fig 20: EVIDENCE OF SQL CODES FOR OLTP***



***Fig 21: EVIDENCE OF SQL CODES FOR MONTHLY SALES OF EACH THEATRE FROM***

***DATA MART***



***Fig 22: EVIDENCE OF SQL CODES FOR MONTHLY SALES OF EACH THEATRE FROM***

***OLTP***

**DATAMART (FACT TABLE)**

select \* from P2733425\_MT\_FACT\_TABLE FT;

**RELATIONAL MODEL (FACT TABLE)**

Select \* from ops$yyang00.TicketPurchase;

**5.3.**

The Data Mart satisfies the requirements of MT Company. It’s used for extracting the essential dimensions and attributes from the available database (i.e. a subset of relational database).

The Data Mart is important for storing essential information in one table, i.e., the fact table, which makes querying easier.

However, the relational database requires a high-level query to join different tables, columns and rows. This may make extraction of useful data extremely difficult, however Data Mart facilitates easy organization of data for analysis, hence Data Mart analysis is most efficient for easy analysis and extraction of data for informed decision.