Sri Lanka Institute of Information Technology

Assignment I

Data Warehouse & Business Intelligence

2022

Submitted by:

Mayurrsh.T

Y3S1.04 (DS)

Contents

1.Data set selection

Provided by: kaggle.com
Data Set Name: Melbourne Housing Snapshot
Data Set:
https://www.kaggle.com/datasets/dansbecker/melbourne-housing-snapshot

About Dataset

The chosen data source is a Kaggle based collection of transactional data. Which represents Melbourne's house-sale information. It is made up of a single CSV file containing enough data in 21 columns. The original huge CSV file has been divided into smaller sub-CSV files. like Seller Details and Property, New Identifiers are contained in the sub-CSV files. In addition, I manually changed some data records to meet the requirements.

2.PREPARATION OF DATA SOURCE

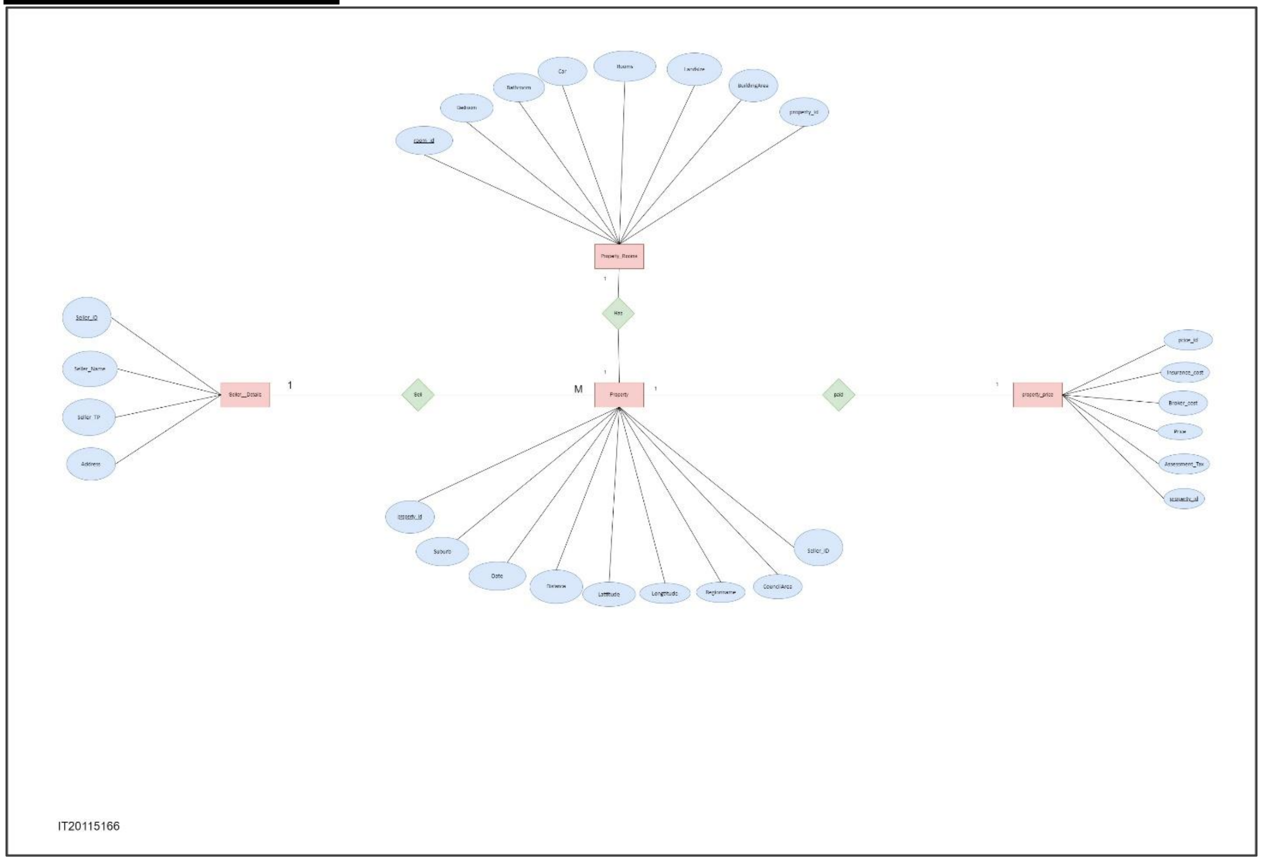
All the data sources are provided in csv format by the web site. In preparation of data sources, some changes have done for the source format. Some of the given files were converted into text files and Property Details csv files into a source database, while others were removed and added to another file.

Final State of Preparation of the source data formats before Transforming data:

- Property_Details.csv
- Property_Price.csv
- Property_Rooms.csv
- Seller.txt

This data collection contains information about house sales in Melbourne. It comprises information on the residence as well as information about the sellers who sold such properties. One csv file was split into two halves, as shown below.

ER-Diagram



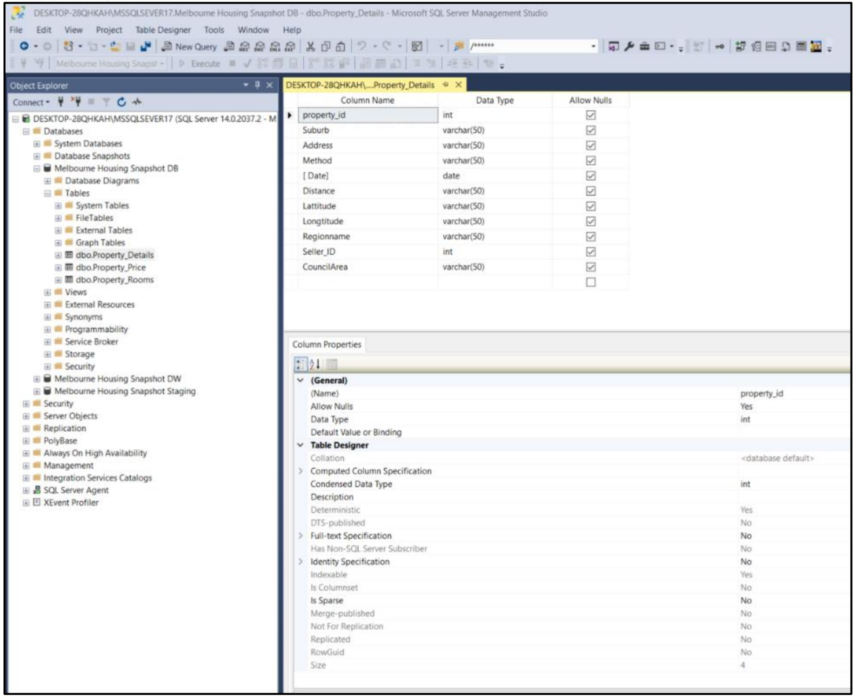
- ➢ The above diagram shows the connection between the entities in the data set
- ➢ Assumptions:
    - One Seller have many Properties.
    - There can be many campaign data sets in a single summary report
    - Many client transactions are summarized in a single summary report.

Description of the Data Set

<table>
<tr><th>Source Type</th><th>Table Name</th><th>Include</th></tr>
<tr><td>Seller.txt</td><td>Seller</td><td>
<table>
<tr><th>Column</th><th>Data type</th><th>Description</th></tr>
<tr><td>Seller_ID</td><td>nvarchar(255)</td><td>Unique id of Seller (PK)</td></tr>
<tr><td>Seller_Name</td><td>nvarchar(255)</td><td>Name of Seller</td></tr>
<tr><td>Seller_TP</td><td>nvarchar(255)</td><td>Phone number of Seller</td></tr>
<tr><td>Address</td><td>nvarchar(255)</td><td>Address of Seller</td></tr>
</table>
</td></tr>
<tr><td>Melbourne_Housing_SnapshotDB</td><td>Property_Details</td><td>
<table>
<tr><th>Column</th><th>Data type</th><th>Description</th></tr>
<tr><td>property_id</td><td>int</td><td>Unique id of Property (PK)</td></tr>
<tr><td>Suburb</td><td>nvarchar(255)</td><td>Name of Residential name</td></tr>
<tr><td>Address</td><td>nvarchar(255)</td><td>Adress of Property</td></tr>
<tr><td>Method</td><td>nvarchar(255)</td><td>Sold Method</td></tr>
<tr><td>Distance</td><td>float</td><td>Distance From Capital</td></tr>
<tr><td>Seller_ID</td><td>nvarchar(255)</td><td>Unique id of Seller (FK)</td></tr>
<tr><td>Lattitude</td><td>float</td><td>Lattitude</td></tr>
<tr><td>Longtitude</td><td>float</td><td>Longitude</td></tr>
<tr><td>Regionname</td><td>nvarchar(255</td><td>Name of Regional</td></tr>
<tr><td>CouncilArea</td><td>nvarchar(255</td><td>Governing Council for the Area</td></tr>
</table>
</td></tr>
<tr><td></td><td>Property_Price</td><td>
<table>
<tr><th>Column</th><th>Data type</th><th>Description</th></tr>
<tr><td>price_id</td><td>int</td><td>Unique id of Price (PK)</td></tr>
<tr><td>Assessment_Tax</td><td>float</td><td>Price of Aessment Tax</td></tr>
<tr><td>Broker_cost</td><td>float</td><td>Price of Broker Cost</td></tr>
<tr><td>Insurance_cost</td><td>float</td><td>Price of Insurance Cost</td></tr>
<tr><td>Price</td><td>float</td><td>Price of that Property</td></tr>
<tr><td>property_id</td><td>int</td><td>Unique id of Property (FK)</td></tr>
</table>
</td></tr>
<tr><td></td><td>Room_Count</td><td>
<table>
<tr><th>Column</th><th>Data type</th><th>Description</th></tr>
<tr><td>room_id</td><td>int</td><td>Unique id of Room (PK)</td></tr>
<tr><td>Bedroom</td><td>int</td><td>Number of bedrooms in this Property</td></tr>
<tr><td>Bathroom</td><td>int</td><td>Number of bathrooms in this Property</td></tr>
<tr><td>Car</td><td>int</td><td>Number of Cars spots in this Property</td></tr>
<tr><td>Rooms</td><td>int</td><td>Number of rooms in this Property</td></tr>
<tr><td>property_id</td><td>int</td><td>Unique id of Property (FK)</td></tr>
<tr><td>Landsize</td><td>int</td><td>Size of the Land</td></tr>
<tr><td>BuildingArea</td><td>int</td><td>Area of the Building</td></tr>
<tr><td></td><td></td><td></td></tr>
</table>
</td></tr>
</table>

Design of Data_Source

- **Property_Details**

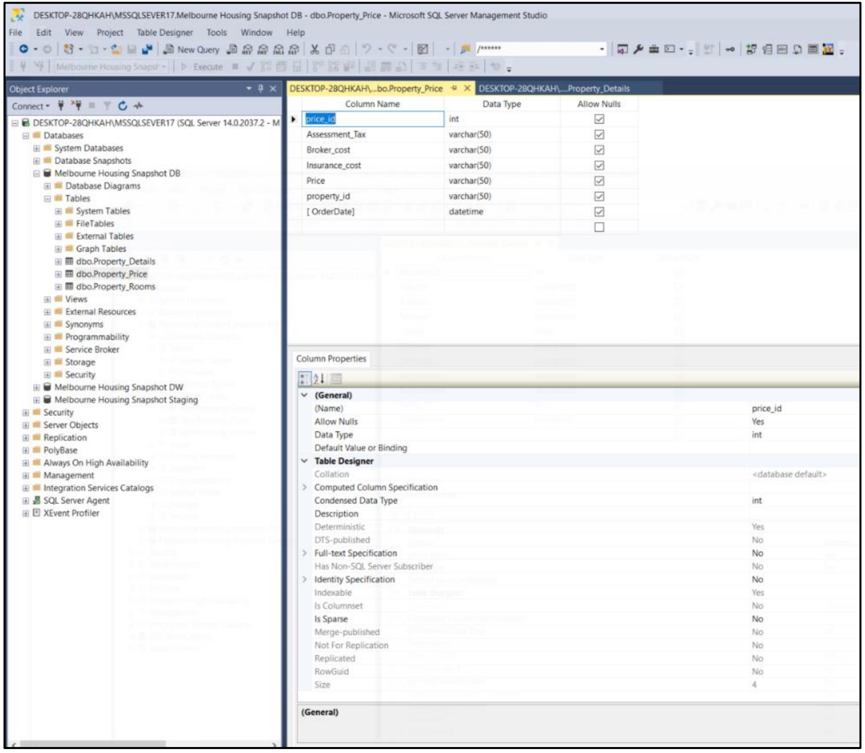

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB - dbo.Property_Details - Microsoft SQL Server Management Studio
File Edit View Project Table Designer Tools Window Help
New Query
Execute
Object Explorer
Connect
DESKTOP-28QHKAH\MSSQLSEVER17 (SQL Server 14.0.2037.2 - M
Databases
System Databases
Database Snapshots
Melbourne Housing Snapshot DB
Database Diagrams
Tables
System Tables
FileTables
External Tables
Graph Tables
dbo.Property_Details
dbo.Property_Price
dbo.Property_Rooms
Views
External Resources
Synonyms
Programmability
Service Broker
Storage
Security
Melbourne Housing Snapshot DW
Melbourne Housing Snapshot Staging
Security
Server Objects
Replication
PolyBase
Always On High Availability
Management
Integration Services Catalogs
SQL Server Agent
XEvent Profiler
DESKTOP-28QHKAH\....Property_Details

Column Name	Data Type	Allow Nulls
property_id	int	☑
Suburb	varchar(50)	☑
Address	varchar(50)	☑
Method	varchar(50)	☑
[Date]	date	☑
Distance	varchar(50)	☑
Lattitude	varchar(50)	☑
Longtitude	varchar(50)	☑
Regionname	varchar(50)	☑
Seller_ID	int	☑
CouncilArea	varchar(50)	☑
		□

Column Properties

(General)	
(Name)	property_id
Allow Nulls	Yes
Data Type	int
Default Value or Binding	
Table Designer	
Collation	<database default>
Computed Column Specification	
Condensed Data Type	int
Description	
Deterministic	Yes
DTS-published	No
Full-text Specification	No
Has Non-SQL Server Subscriber	No
Identity Specification	No
Indexable	Yes
Is Columnset	No
Is Sparse	No
Merge-published	No
Not For Replication	No
Replicated	No
RowGuid	No
Size	4

- **Property_Price**


DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB - dbo.Property_Price - Microsoft SQL Server Management Studio
File Edit View Project Table Designer Tools Window Help
New Query
Execute
Object Explorer
Connect
DESKTOP-28QHKAH\MSSQLSEVER17 (SQL Server 14.0.2037.2 - M
Databases
System Databases
Database Snapshots
Melbourne Housing Snapshot DB
Database Diagrams
Tables
System Tables
FileTables
External Tables
Graph Tables
dbo.Property_Details
dbo.Property_Price
dbo.Property_Rooms
Views
External Resources
Synonyms
Programmability
Service Broker
Storage
Security
Melbourne Housing Snapshot DW
Melbourne Housing Snapshot Staging
Security
Server Objects
Replication
PolyBase
Always On High Availability
Management
Integration Services Catalogs
SQL Server Agent
XEvent Profiler
DESKTOP-28QHKAH\...bo.Property_Price
DESKTOP-28QHKAH\....Property_Details

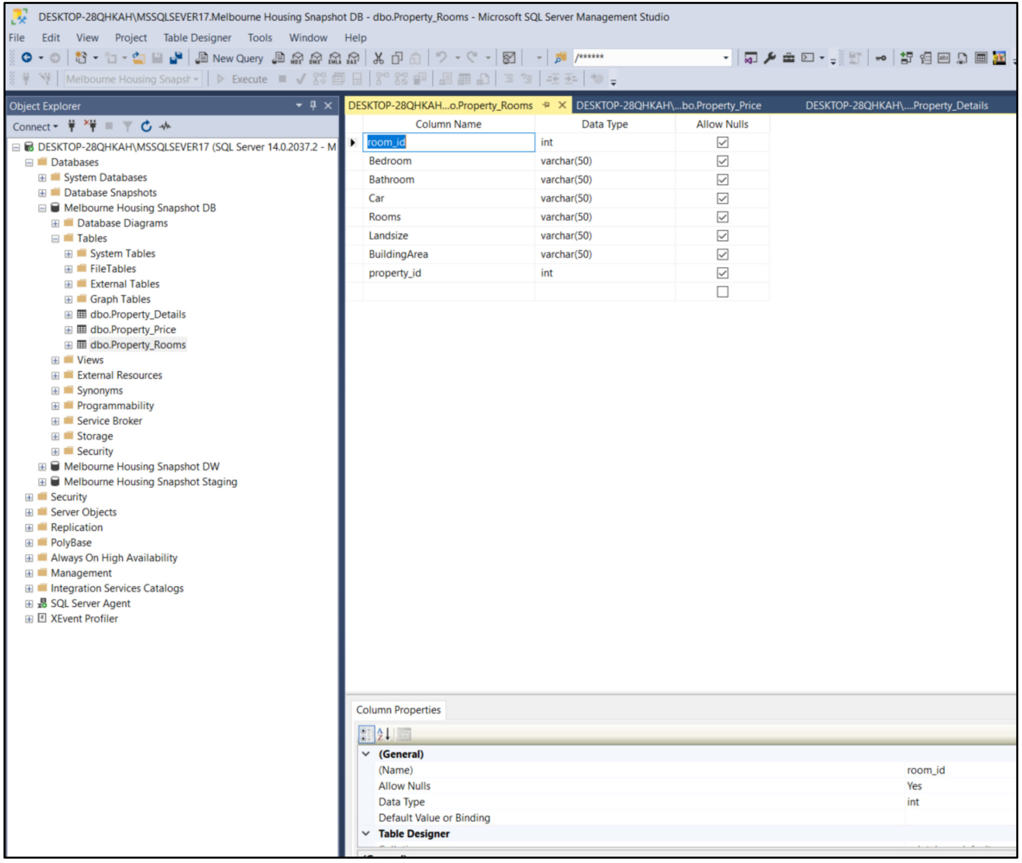
Column Name	Data Type	Allow Nulls
price_id	int	☑
Assessment_Tax	varchar(50)	☑
Broker_cost	varchar(50)	☑
Insurance_cost	varchar(50)	☑
Price	varchar(50)	☑
property_id	varchar(50)	☑
[OrderDate]	datetime	☑
		□

Column Properties

(General)	
(Name)	price_id
Allow Nulls	Yes
Data Type	int
Default Value or Binding	
Table Designer	
Collation	<database default>
Computed Column Specification	
Condensed Data Type	int
Description	
Deterministic	Yes
DTS-published	No
Full-text Specification	No
Has Non-SQL Server Subscriber	No
Identity Specification	No
Indexable	Yes
Is Columnset	No
Is Sparse	No
Merge-published	No
Not For Replication	No
Replicated	No
RowGuid	No
Size	4

(General)

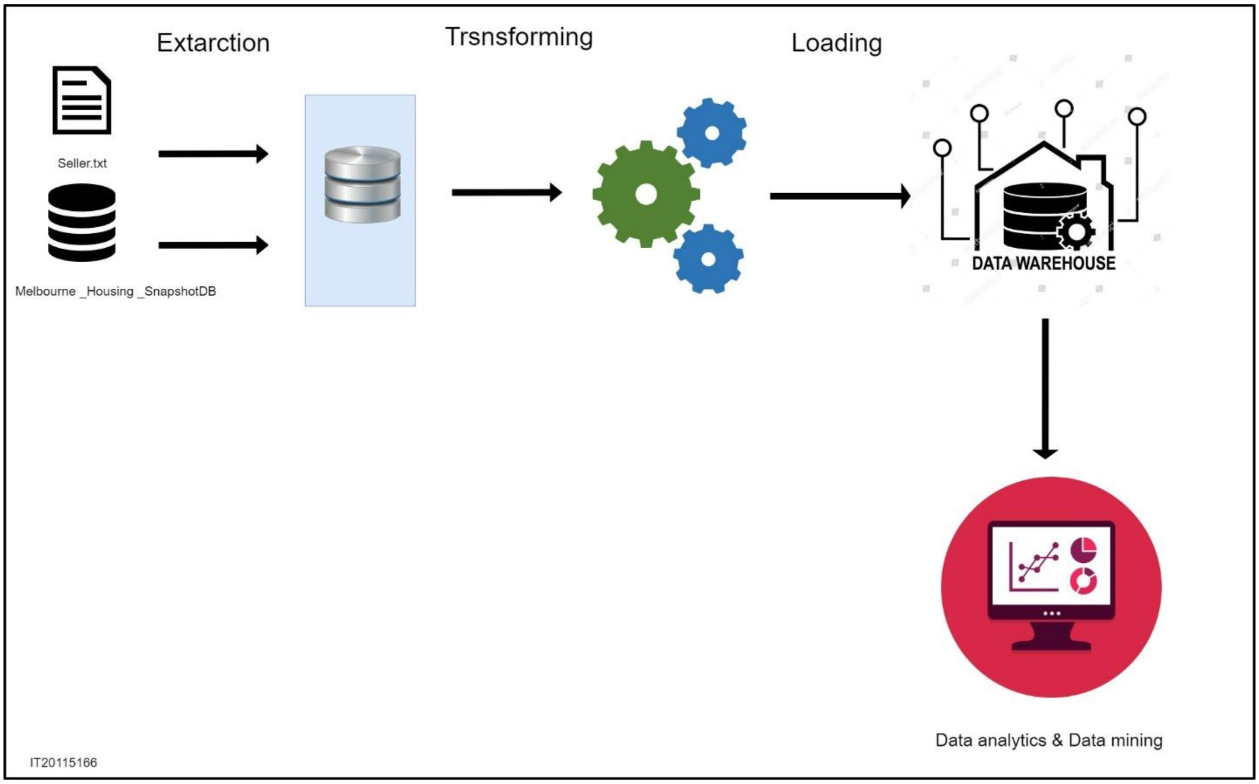
- **Property_Rooms**



Column Name	Data Type	Allow Nulls
room_id	int	☑
Bedroom	varchar(50)	☑
Bathroom	varchar(50)	☑
Car	varchar(50)	☑
Rooms	varchar(50)	☑
Landsize	varchar(50)	☑
BuildingArea	varchar(50)	☑
property_id	int	☑
		☐

(General)	
(Name)	room_id
Allow Nulls	Yes
Data Type	int
Default Value or Binding	
Table Designer	

3.SOLUTION ARCHITECTURE



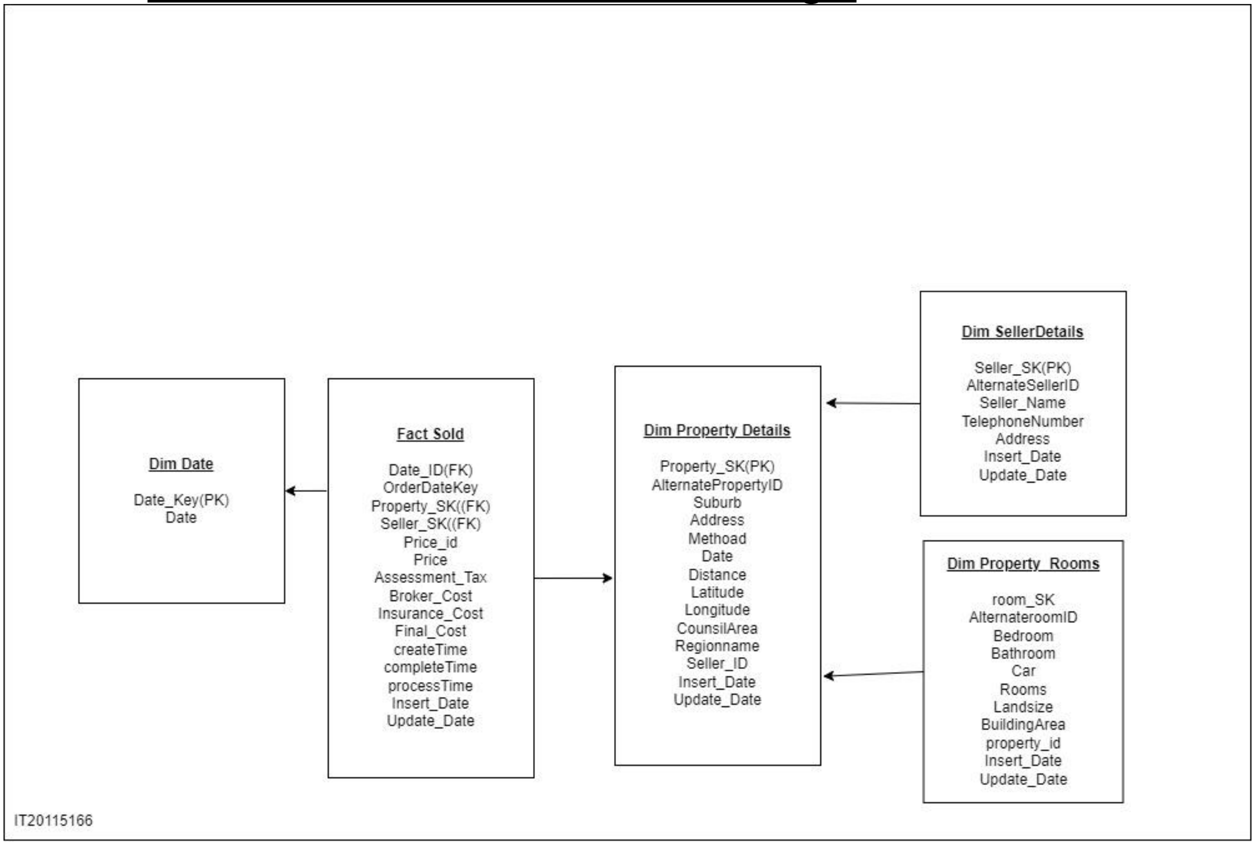
(As the figure 2 shows for the ETL processing, initially)

- **Seller**: Text File
- **Melbourne _Housing _SnapshotDB**: Source Database,

We can handle data from various sources and transform it to business insights to make decisions, analyze data, and produce reports using diverse procedures, structures, and technologies. This will also give the data a new dimension

4.Data warehouse Design & Development

I. Data warehouse Table design



Dimention Name	**Dimention Attributes**	**Data Type**	**Key column**	**Derived Logic**
Dim_Property_ Details	Property_SK	int	Primary key	Auto increment
	AlternatePropertyID	int		
	Suburb	varchar (50)		
	Address	varchar (50)		
	Method	varchar (50)		
	Date	varchar (50)		
	Diatance	varchar (50)		

	Lattitude	varchar(50)		
	Longitude	varchar(50)		
	Seller_ID	int		
	Regionname	varchar(50)		
	Insert_Date	datetime		System Datetime
	ModifiedDatedate	datetime		System Datetime
Dim_SellerDetails	Seller_SK	int	Primary key	
	AlternateSellerID	int		
	Seller_Name	nvarchar(50)		
	Telephone_Number	nvarchar(50)		
	Address	nvarchar(50)		
	Insert_Date	datetime		System Datetime
	Update_Date	datetime		System Datetime
DimDate	DateKey	int	Primary key	
	Date	datetime		
	FullDateUK	char(10)		
	FullDateUSA	char(10)		
	DayOfMonth	varchar(4)		
	DaySuffix	varchar(9)		
	DayName	varchar(9)		
	More….			
Fact_Sold	Seller_SK	int	foreign key	
	Property_SK	int	foreign	

			key	
	Date_ID	int	foreign key	
	Price	float		
	price_id	int		
	Assessment_Tax	varchar(50)		
	Broker_Cost	varchar(50)		
	createTime	datetime		
	completeTime	datetime		
	processTime	datetime		
	Insurance_Cost	float		
	Final_Cost	float		Price+Assessment_Tax+Broker_Cost+Insuarence_Cost
	Insert_Date	datetime		System Datetime
	Update_Date	datetime		System Datetime
DimProperty_Rooms	room_SK	int		
	AlternateroomID	int		
	Bedroom	varchar(50)		
	Bathroom	varchar(50		
	Car	varchar(50		
	Rooms	varchar(50		
	landsize	varchar(50		
	BuildingArea	varchar(50		
	property_id	int		
	Insert_Date	datetime		System Datetime
	Insert_Date	datetime		System Datetime

Calculation

: Final Cost =(Price+ Assessment_Tax+ Broker_Cost+ Insurance_Cost)

I. Assumptions

- dbo.DimDate is added to the Data Warehouse for better performance.
- dbo. Property_Price is used in creating the fact table

II. Slowly changing dimensions

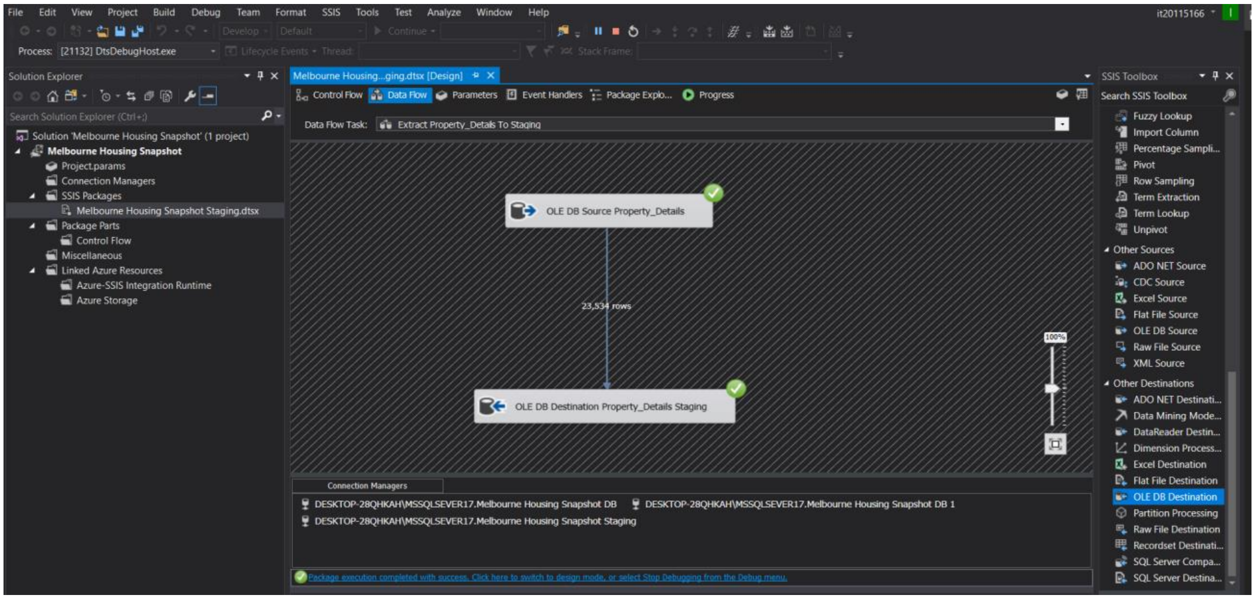
- Customer Details were considered as a slowly changing dimension

• **Dimension table**	**Attributes**
Dim_SellerDetails	Telephone_Number (changing attribute) Address (Hostorical)

5.ETL Development

I. Data Extraction & Load into Staging Tables

Property_Details



Data Flow Task: Extract Property_Details To Staging

OLE DB Source Property_Details

23,534 rows

OLE DB Destination Property_Details Staging

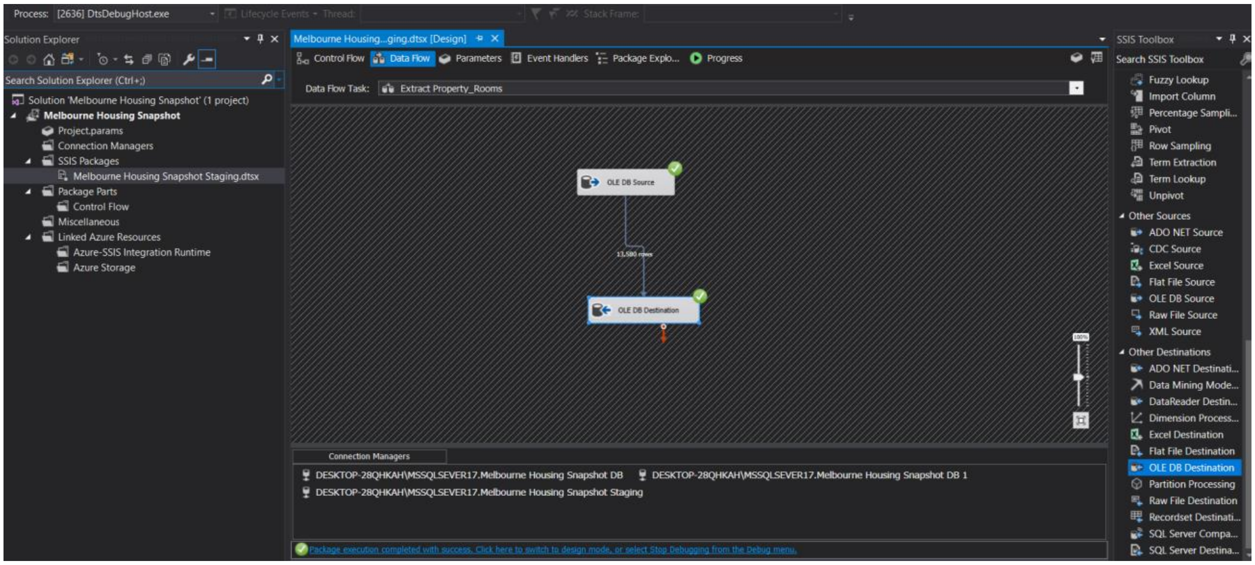
Connection Managers

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot Staging

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

(Property Details is extracted from Property Details the table in the source database and inserted to the Property Details Staging table)

Property_Rooms



Data Flow Task: Extract Property_Rooms

OLE DB Source

13,580 rows

OLE DB Destination

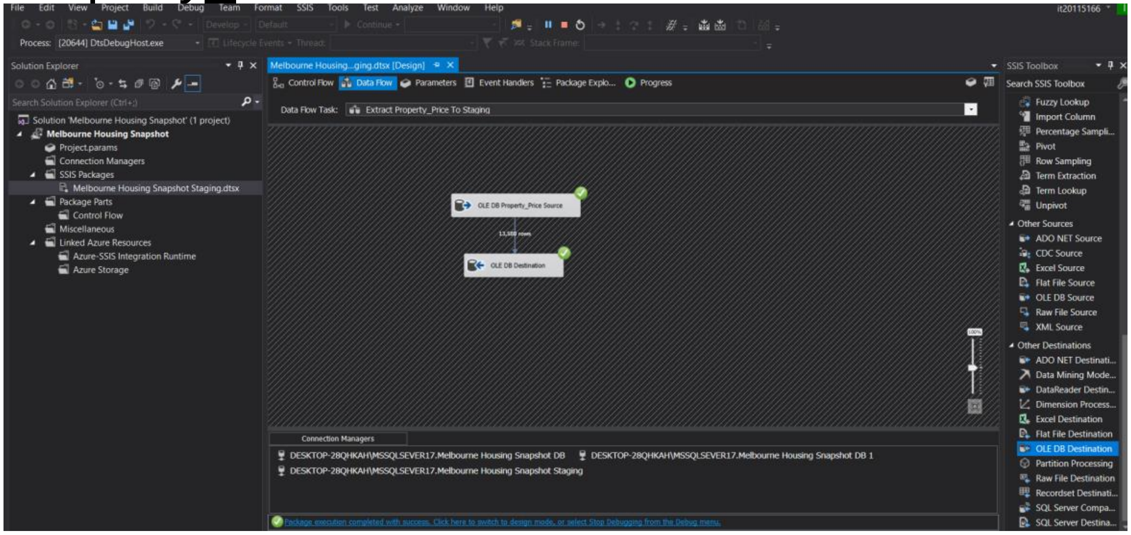
Connection Managers

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot Staging

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

(Place Details is extracted from Property Rooms the table in the source database and inserted to the Property Place Staging table)

Property_Price


File Edit View Project Build Debug Team Format SSIS Tools Test Analyze Window Help
Process: [20644] DtsDebugHost.exe
Solution Explorer
Melbourne Housing...ging.dtsx [Design]
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Progress
SSIS Toolbox
Search SSIS Toolbox
OLE DB Property_Price Source
13,580 rows
OLE DB Destination
Connection Managers

Solution 'Melbourne Housing Snapshot' (1 project)
Melbourne Housing Snapshot
Project.params
Connection Managers
SSIS Packages
Melbourne Housing Snapshot Staging.dtsx
Package Parts
Control Flow
Miscellaneous
Linked Azure Resources
Azure-SSIS Integration Runtime
Azure Storage

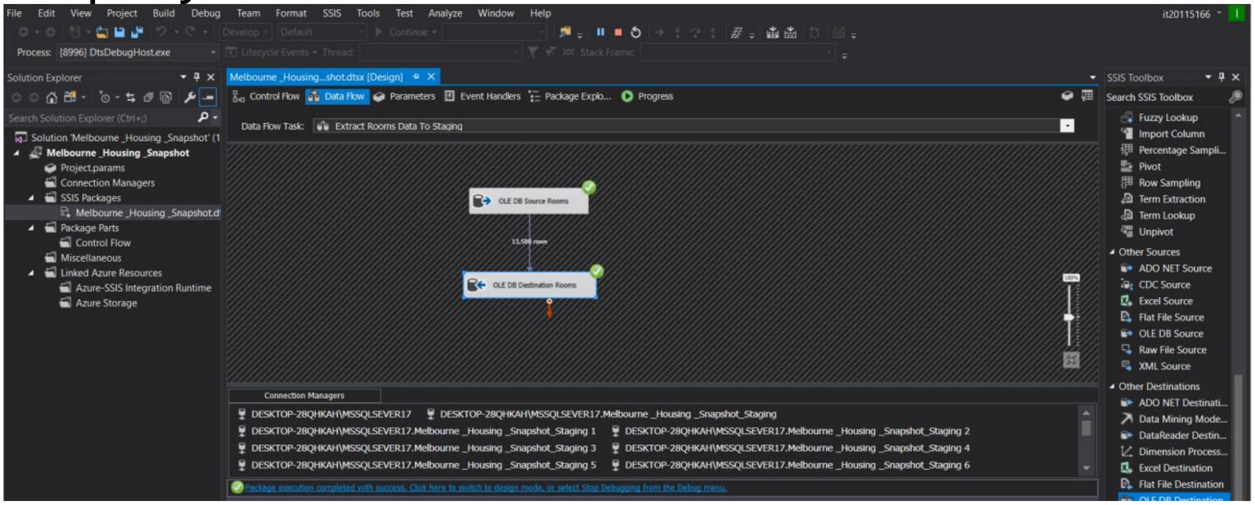
Data Flow Task: Extract Property_Price To Staging

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot Staging

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

(Price Details is extracted from Property_Price the table in the source database and inserted to the Property Price Staging table)

Property Rooms



Solution 'Melbourne _Housing _Snapshot' (1
Melbourne _Housing _Snapshot
Project.params
Connection Managers
SSIS Packages
Melbourne _Housing _Snapshot.d
Package Parts
Control Flow
Miscellaneous
Linked Azure Resources
Azure-SSIS Integration Runtime
Azure Storage

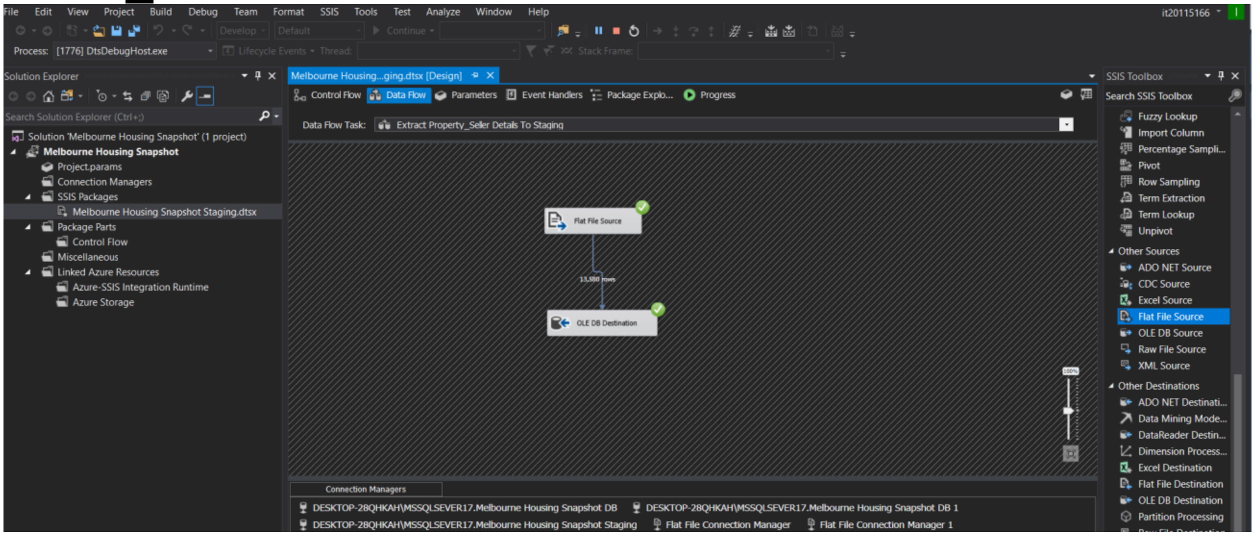
Data Flow Task: Extract Rooms Data To Staging

DESKTOP-28QHKAH\MSSQLSEVER17
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 2
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 3
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 4
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 5
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging 6

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

(Room Details is extracted from room_count the table in the source database and inserted to the Room Count Staging table)

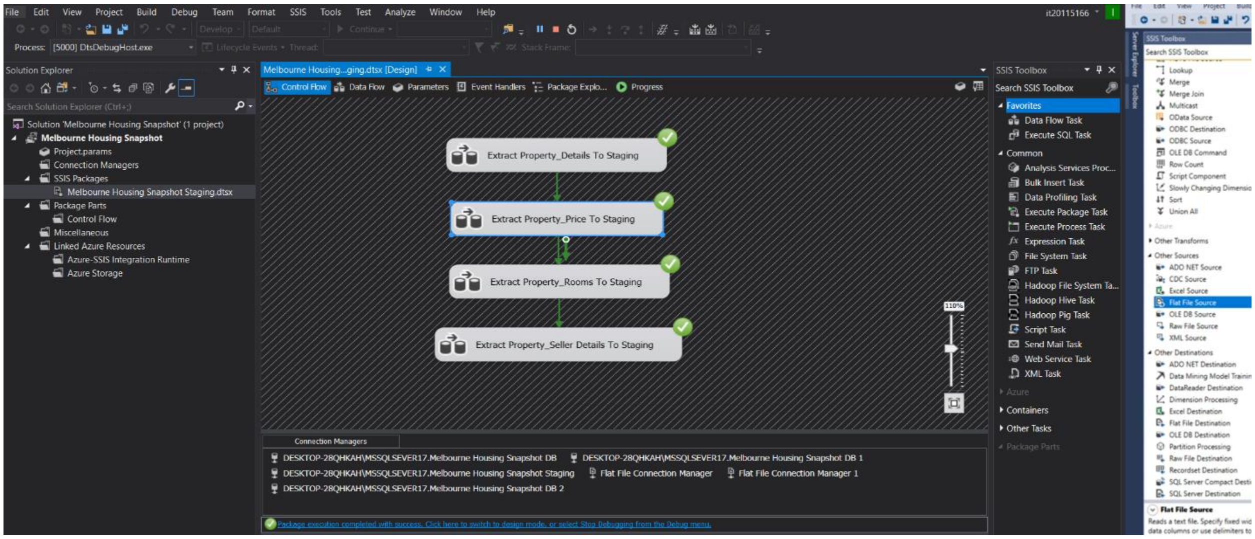
Seller_Details



Data Flow Task: Extract Property_Seller Details To Staging

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot Staging
Flat File Connection Manager
Flat File Connection Manager 1

(Seller Details is extracted from Seller.txt the table in the source database and inserted to the Seller Staging table)



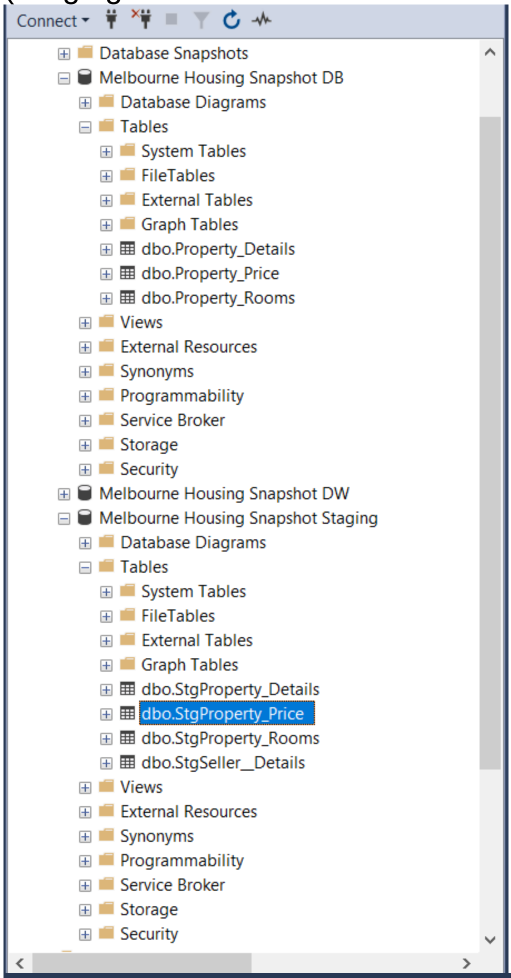
Extract Property_Details To Staging
Extract Property_Price To Staging
Extract Property_Rooms To Staging
Extract Property_Seller Details To Staging
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Progress
SSIS Toolbox
Connection Managers

DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot Staging
Flat File Connection Manager
Flat File Connection Manager 1
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne Housing Snapshot DB 2

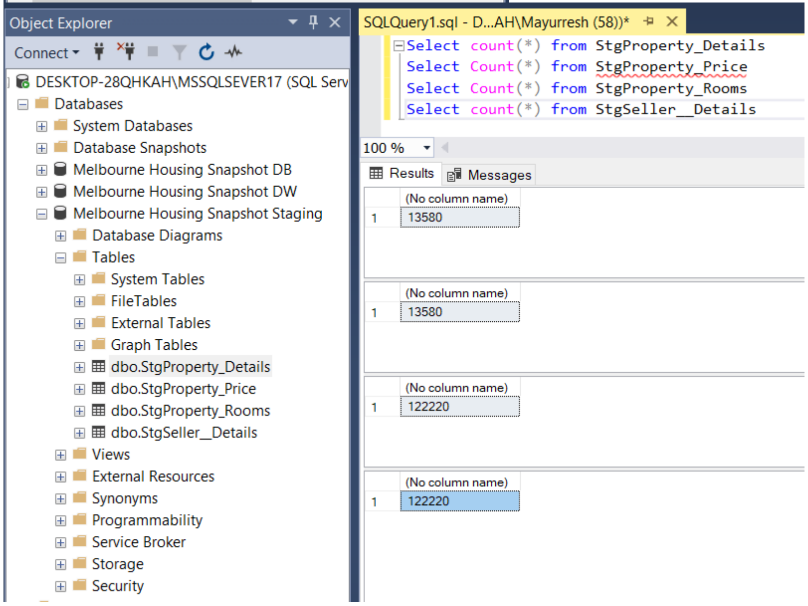
Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

(The Control Flow of ‘Extract Data and Load into Staging’ Step can illustrate as the give figure)

(Staging Tables created and values inserted)



```
⊞ Database Snapshots
⊟ Melbourne Housing Snapshot DB
    ⊞ Database Diagrams
    ⊟ Tables
        ⊞ System Tables
        ⊞ FileTables
        ⊞ External Tables
        ⊞ Graph Tables
        ⊞ dbo.Property_Details
        ⊞ dbo.Property_Price
        ⊞ dbo.Property_Rooms
    ⊞ Views
    ⊞ External Resources
    ⊞ Synonyms
    ⊞ Programmability
    ⊞ Service Broker
    ⊞ Storage
    ⊞ Security
⊞ Melbourne Housing Snapshot DW
⊟ Melbourne Housing Snapshot Staging
    ⊞ Database Diagrams
    ⊟ Tables
        ⊞ System Tables
        ⊞ FileTables
        ⊞ External Tables
        ⊞ Graph Tables
        ⊞ dbo.StgProperty_Details
        ⊞ dbo.StgProperty_Price
        ⊞ dbo.StgProperty_Rooms
        ⊞ dbo.StgSeller__Details
    ⊞ Views
    ⊞ External Resources
    ⊞ Synonyms
    ⊞ Programmability
    ⊞ Service Broker
    ⊞ Storage
    ⊞ Security
```



```
DESKTOP-28QHKAH\MSSQLSEVER17 (SQL Serv
⊟ Databases
    ⊞ System Databases
    ⊞ Database Snapshots
    ⊞ Melbourne Housing Snapshot DB
    ⊞ Melbourne Housing Snapshot DW
    ⊟ Melbourne Housing Snapshot Staging
        ⊞ Database Diagrams
        ⊟ Tables
            ⊞ System Tables
            ⊞ FileTables
            ⊞ External Tables
            ⊞ Graph Tables
            ⊞ dbo.StgProperty_Details
            ⊞ dbo.StgProperty_Price
            ⊞ dbo.StgProperty_Rooms
            ⊞ dbo.StgSeller__Details
        ⊞ Views
        ⊞ External Resources
        ⊞ Synonyms
        ⊞ Programmability
        ⊞ Service Broker
        ⊞ Storage
        ⊞ Security
```

```sql
Select count(*) from StgProperty_Details
Select Count(*) from StgProperty_Price
Select Count(*) from StgProperty_Rooms
Select count(*) from StgSeller__Details
```

	(No column name)
1	13580

	(No column name)
1	13580

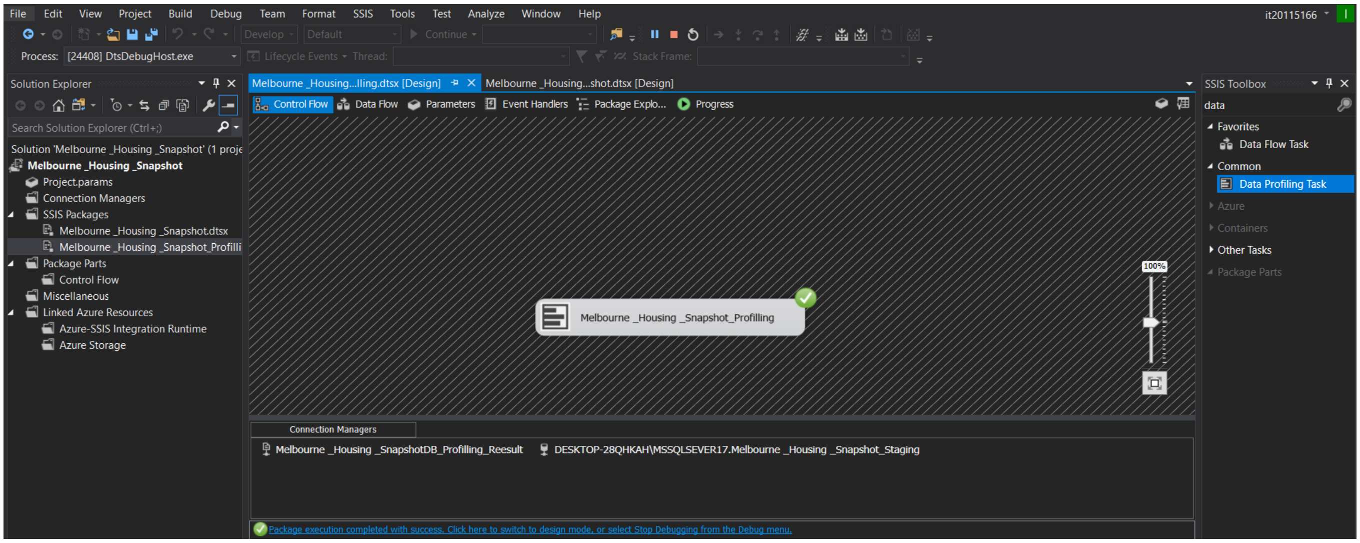
	(No column name)
1	122220

	(No column name)
1	122220

(Available are In Staging

II. Data Profiling

Data Profiling provides the means of analyzing large amount of data using different kind of processes. In this step, null values, repeated values and quality of the data is checked.



Melbourne _Housing _SnapshotDB_Profilling_Reesult

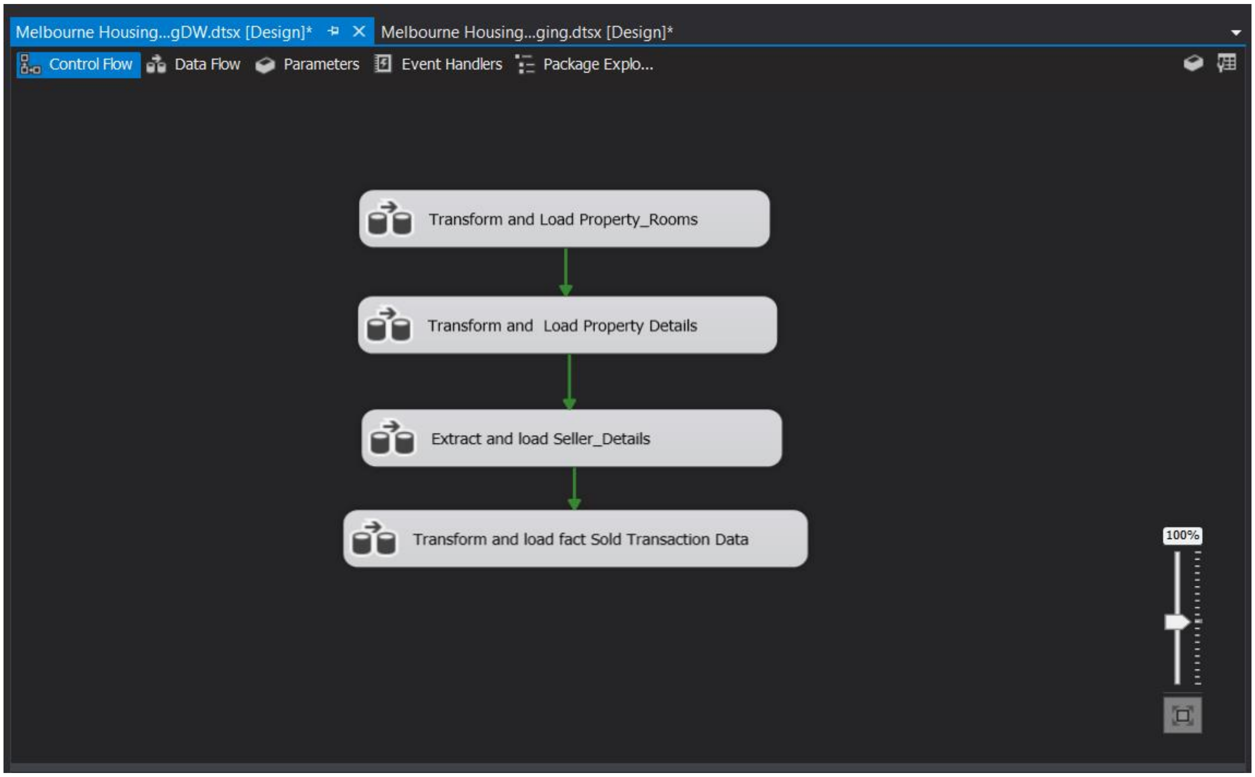
DESKTOP-28QHKAH\MSSQLSEVER17.Melbourne _Housing _Snapshot_Staging

Package execution completed with success. Click here to switch to design mode, or select Stop Debugging from the Debug menu.

- ❖ Each staging table is profiled and saved in a specific folder.
- ❖ As the figure shows, after the Staging step doing this task shows the things what the developer has to consider about the data which are stored in staging table and the developer is able to identify the issues with staging data by data profiling (such as null values).
- ❖ The diagram depicted the entirety of Data Profiling as it relates to Staging.

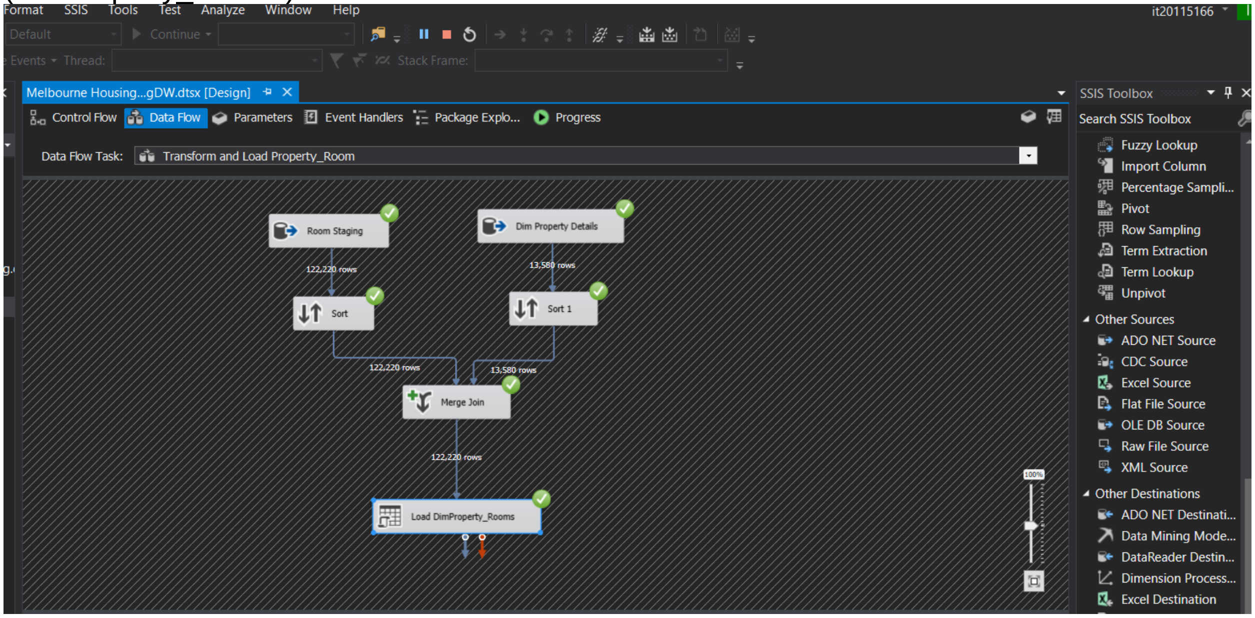
III. Data Transformation and Loading

- Data Transformation is developed according to the dimensional modeling designed above.

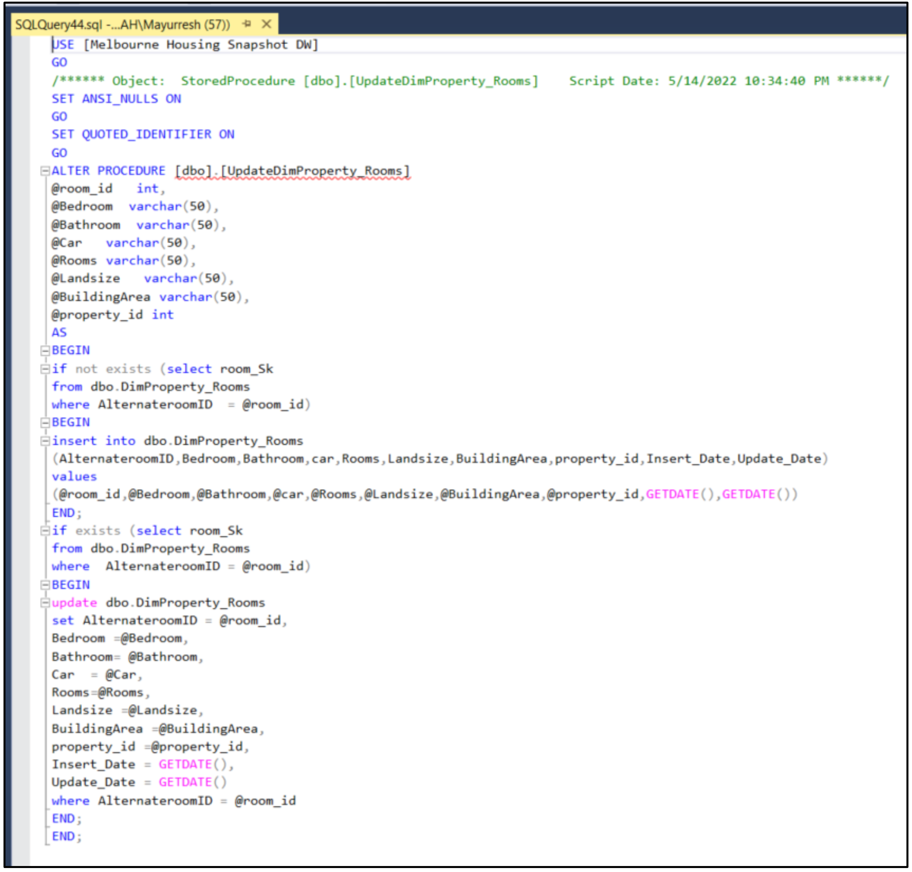

Melbourne Housing...gDW.dtsx [Design]*
Melbourne Housing...ging.dtsx [Design]*
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Transform and Load Property_Rooms
Transform and Load Property Details
Extract and load Seller_Details
Transform and load fact Sold Transaction Data
100%

- In this step, the Dimension Tables created in Melbourne Housing Snapshot_DW are loaded with the data of relevant staging tables

(DimProperty_Rooms)



- ❖ Property Room data is loaded to the DimProperty_Room.
- ❖ Sort and merge transformation tasks are used.

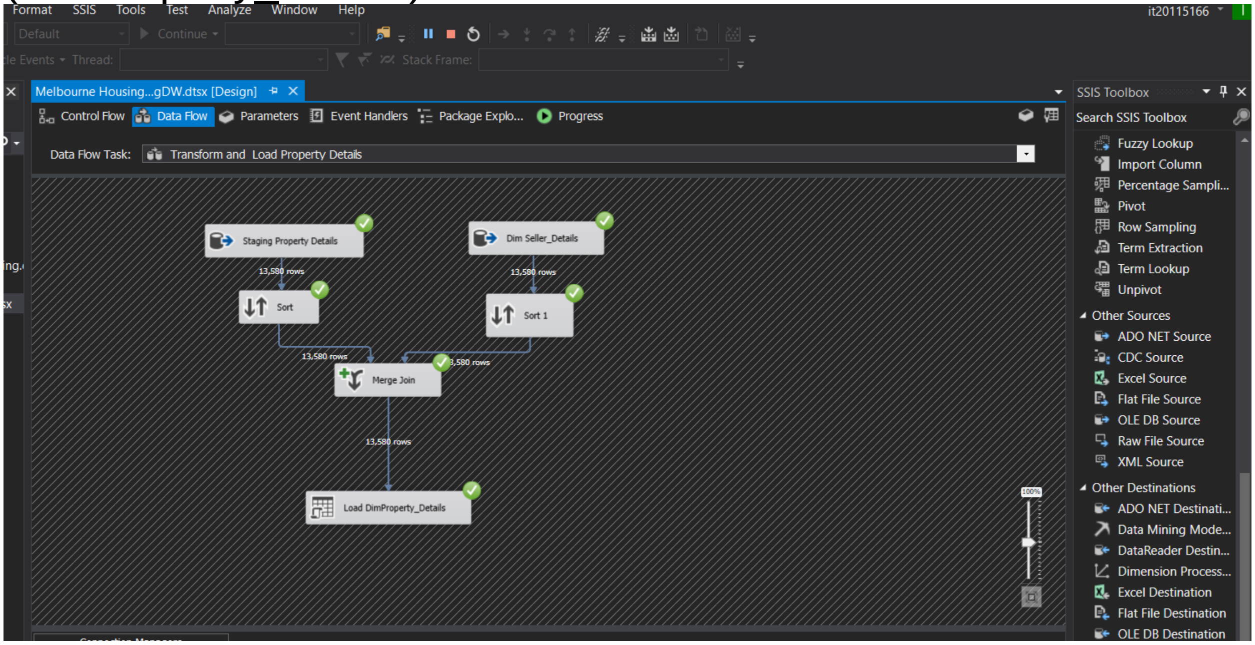


SQLQuery44.sql -...AH\Mayurresh (57))

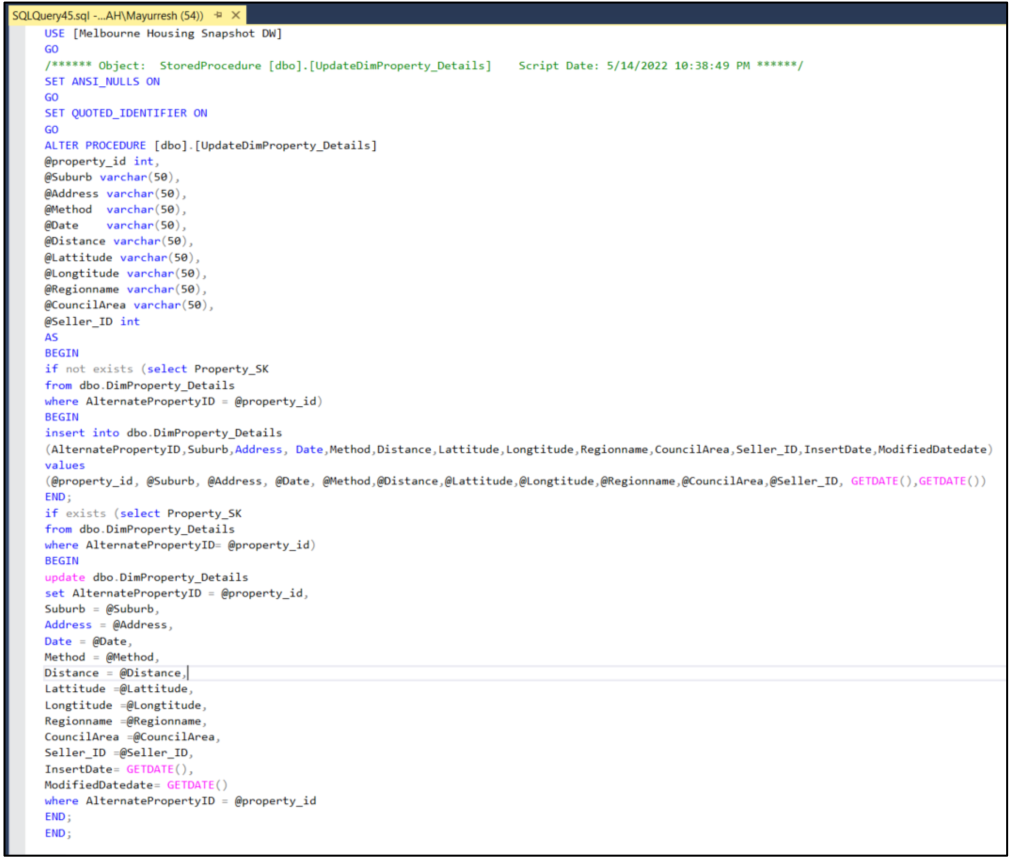
```sql
USE [Melbourne Housing Snapshot DW]
GO
/****** Object:  StoredProcedure [dbo].[UpdateDimProperty_Rooms]    Script Date: 5/14/2022 10:34:40 PM ******/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[UpdateDimProperty_Rooms]
@room_id    int,
@Bedroom  varchar(50),
@Bathroom  varchar(50),
@Car    varchar(50),
@Rooms varchar(50),
@Landsize   varchar(50),
@BuildingArea varchar(50),
@property_id int
AS
BEGIN
if not exists (select room_Sk
from dbo.DimProperty_Rooms
where AlternateroomID  = @room_id)
BEGIN
insert into dbo.DimProperty_Rooms
(AlternateroomID,Bedroom,Bathroom,car,Rooms,Landsize,BuildingArea,property_id,Insert_Date,Update_Date)
values
(@room_id,@Bedroom,@Bathroom,@car,@Rooms,@Landsize,@BuildingArea,@property_id,GETDATE(),GETDATE())
END;
if exists (select room_Sk
from dbo.DimProperty_Rooms
where  AlternateroomID = @room_id)
BEGIN
update dbo.DimProperty_Rooms
set AlternateroomID = @room_id,
Bedroom =@Bedroom,
Bathroom= @Bathroom,
Car   = @Car,
Rooms=@Rooms,
Landsize =@Landsize,
BuildingArea =@BuildingArea,
property_id =@property_id,
Insert_Date = GETDATE(),
Update_Date = GETDATE()
where AlternateroomID = @room_id
END;
END;
```

- ❖ Update DimProperty_Room procedure is used to check whether the data inserted or not.

(DimProperty_Details)



- ❖ Property Details data is loaded to the DimProperty_Details.
- ❖ Sort and merge transformation tasks are used.

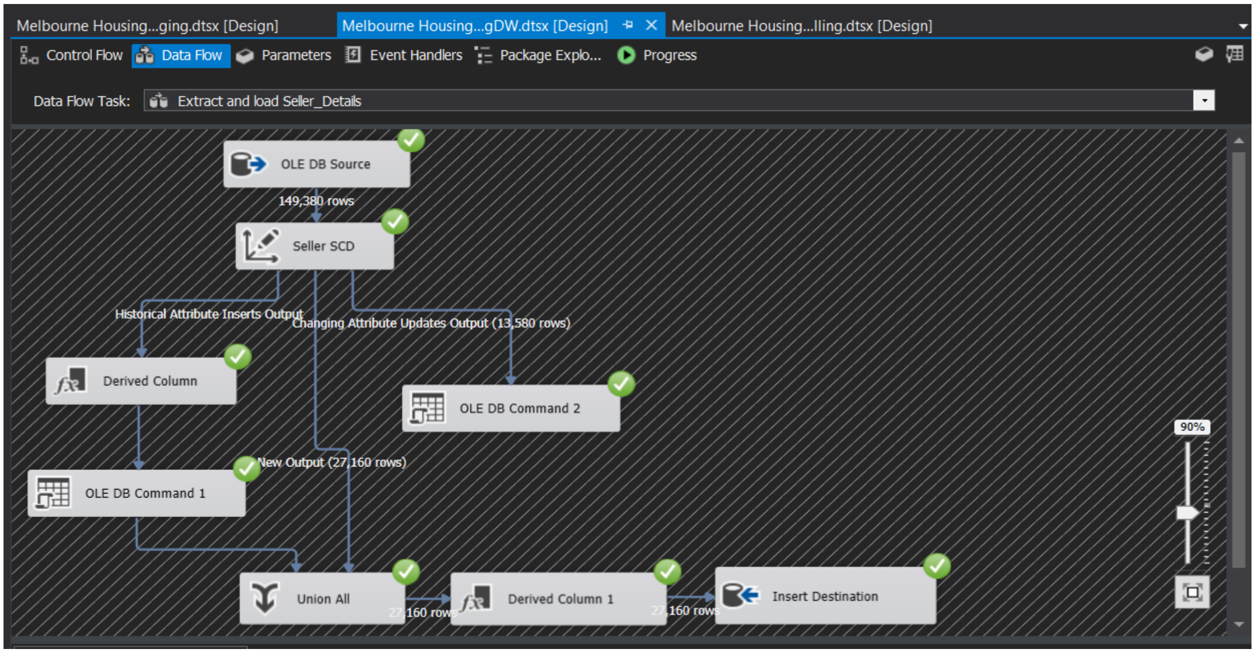


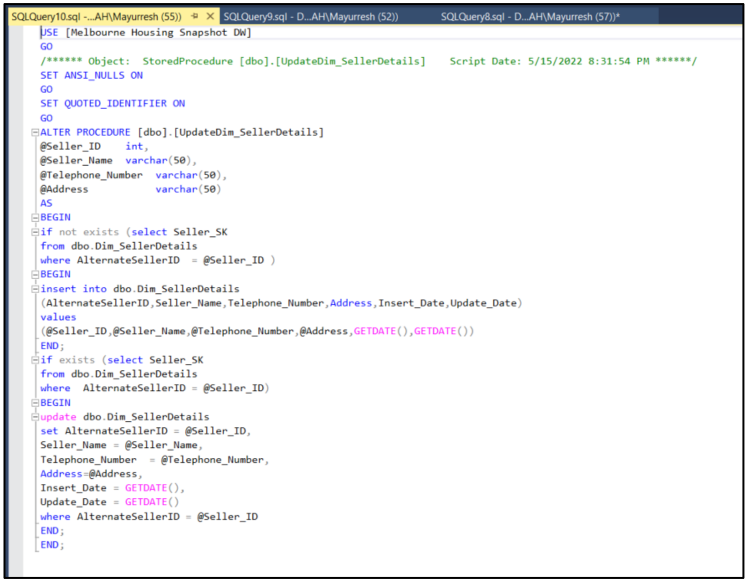
```sql
USE [Melbourne Housing Snapshot DW]
GO
/****** Object:  StoredProcedure [dbo].[UpdateDimProperty_Details]    Script Date: 5/14/2022 10:38:49 PM ******/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[UpdateDimProperty_Details]
@property_id int,
@Suburb varchar(50),
@Address varchar(50),
@Method  varchar(50),
@Date    varchar(50),
@Distance varchar(50),
@Lattitude varchar(50),
@Longtitude varchar(50),
@Regionname varchar(50),
@CouncilArea varchar(50),
@Seller_ID int
AS
BEGIN
if not exists (select Property_SK
from dbo.DimProperty_Details
where AlternatePropertyID = @property_id)
BEGIN
insert into dbo.DimProperty_Details
(AlternatePropertyID,Suburb,Address, Date,Method,Distance,Lattitude,Longtitude,Regionname,CouncilArea,Seller_ID,InsertDate,ModifiedDatedate)
values
(@property_id, @Suburb, @Address, @Date, @Method,@Distance,@Lattitude,@Longtitude,@Regionname,@CouncilArea,@Seller_ID, GETDATE(),GETDATE())
END;
if exists (select Property_SK
from dbo.DimProperty_Details
where AlternatePropertyID= @property_id)
BEGIN
update dbo.DimProperty_Details
set AlternatePropertyID = @property_id,
Suburb = @Suburb,
Address = @Address,
Date = @Date,
Method = @Method,
Distance = @Distance,
Lattitude =@Lattitude,
Longtitude =@Longtitude,
Regionname =@Regionname,
CouncilArea =@CouncilArea,
Seller_ID =@Seller_ID,
InsertDate= GETDATE(),
ModifiedDatedate= GETDATE()
where AlternatePropertyID = @property_id
END;
END;
```

- ❖ Update DimProperty_Details procedure is used to check whether the data inserted or not

Loading Slowly Changing Dimension

- DimSellerDetails is the slowly changing dimension in this dimensional modeling.
- In order to load data to Dimension table, the slowly changing dimensions (historical) have two specific columns as StartDate & End Date to ensure that the data is valid at the moment.
- slowly changing dimension wizard let the developer to select the Dimension table, Business keys of the dimension and what would be the slowly changing attributes.
- The below mentioned columns were set as changing attributes:
    - Telephone_Number : Changing Attribute
    - Adresss :Historical Attribute

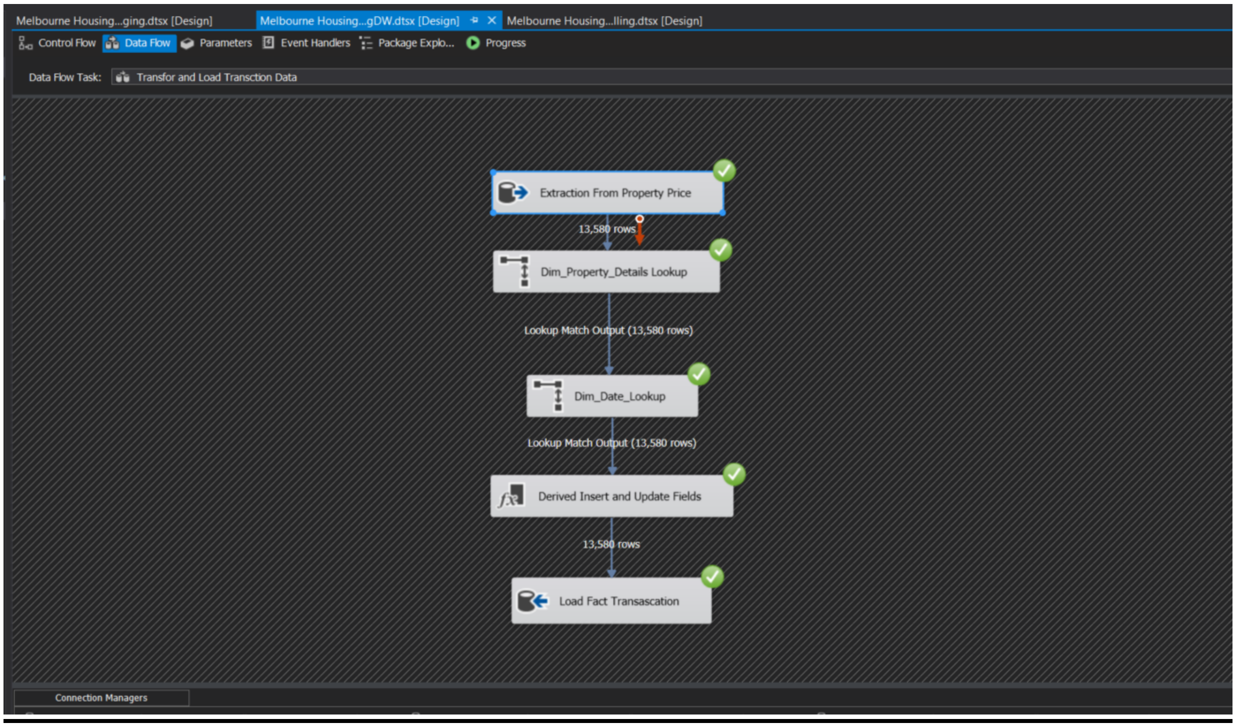


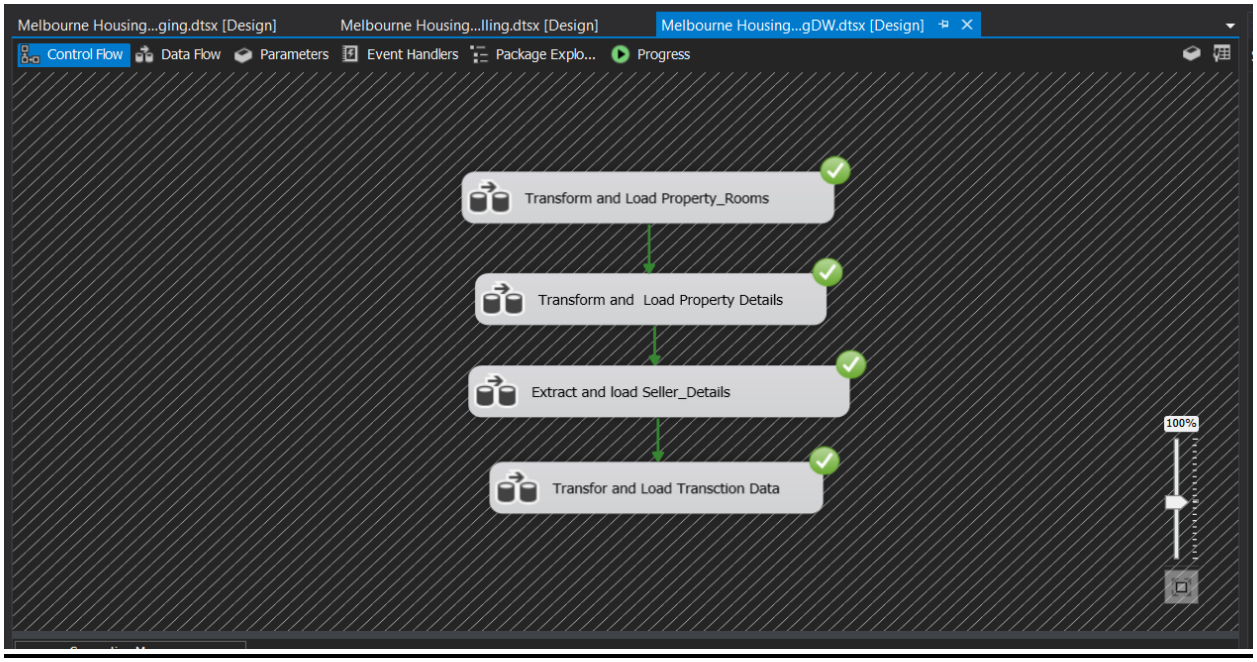

SQLQuery10.sql -...AH\Mayurresh (55))
SQLQuery9.sql - D...AH\Mayurresh (52))
SQLQuery8.sql - D...AH\Mayurresh (57))*

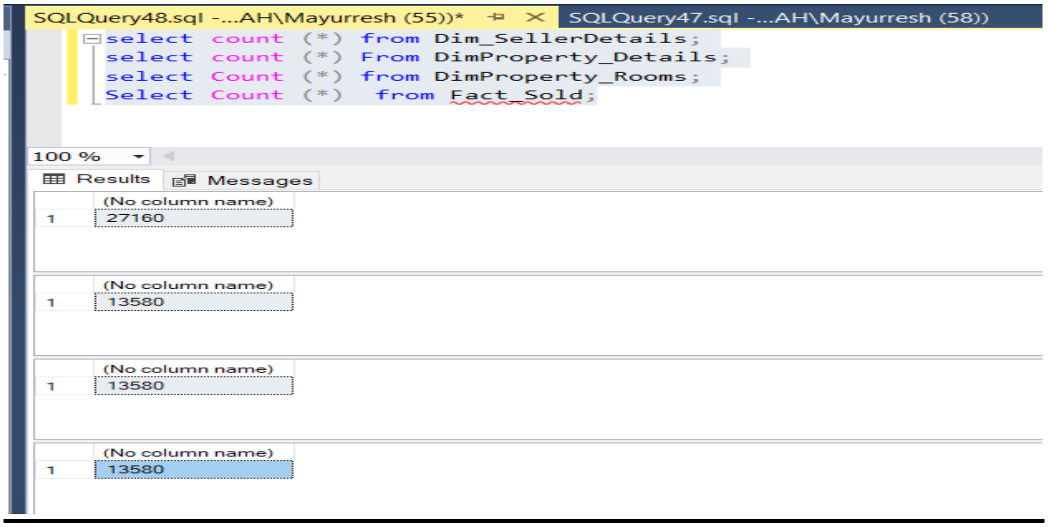
```sql
USE [Melbourne Housing Snapshot DW]
GO
/****** Object:  StoredProcedure [dbo].[UpdateDim_SellerDetails]    Script Date: 5/15/2022 8:31:54 PM ******/
SET ANSI_NULLS ON
GO
SET QUOTED_IDENTIFIER ON
GO
ALTER PROCEDURE [dbo].[UpdateDim_SellerDetails]
@Seller_ID    int,
@Seller_Name  varchar(50),
@Telephone_Number  varchar(50),
@Address           varchar(50)
AS
BEGIN
if not exists (select Seller_SK
from dbo.Dim_SellerDetails
where AlternateSellerID  = @Seller_ID )
BEGIN
insert into dbo.Dim_SellerDetails
(AlternateSellerID,Seller_Name,Telephone_Number,Address,Insert_Date,Update_Date)
values
(@Seller_ID,@Seller_Name,@Telephone_Number,@Address,GETDATE(),GETDATE())
END;
if exists (select Seller_SK
from dbo.Dim_SellerDetails
where  AlternateSellerID = @Seller_ID)
BEGIN
update dbo.Dim_SellerDetails
set AlternateSellerID = @Seller_ID,
Seller_Name = @Seller_Name,
Telephone_Number  = @Telephone_Number,
Address=@Address,
Insert_Date = GETDATE(),
Update_Date = GETDATE()
where AlternateSellerID = @Seller_ID
END;
END;
```

Load data to Fact table

- The final step of Transformation & Loading is load data to fact table. According to the dimensional model, TransactionStaging table is used to insert values into DimTransaction table.
- After loading to all the dimensions, lastly data was loaded to the fact table. The below steps were followed:
  - ❖ Data extracted from the StgProperty_Price staging.
  - ❖ Join operation is done for the Dim_Property_Details Lookup.
  - ❖ Join operation is done for the Dim_Date_Lookup.
  - ❖ insert and modified date were derived.
  - ❖ Fact details loaded to the Fact_Load table.


Melbourne Housing...ging.dtsx [Design]
Melbourne Housing...gDW.dtsx [Design]
Melbourne Housing...lling.dtsx [Design]
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Progress
Data Flow Task:
Transfor and Load Transction Data
Extraction From Property Price
13,580 rows
Dim_Property_Details Lookup
Lookup Match Output (13,580 rows)
Dim_Date_Lookup
Lookup Match Output (13,580 rows)
Derived Insert and Update Fields
13,580 rows
Load Fact Transascation
Connection Managers


Melbourne Housing...ging.dtsx [Design]
Melbourne Housing...lling.dtsx [Design]
Melbourne Housing...gDW.dtsx [Design]
Control Flow
Data Flow
Parameters
Event Handlers
Package Explo...
Progress
Transform and Load Property_Rooms
Transform and Load Property Details
Extract and load Seller_Details
Transfor and Load Transction Data
100%


SQLQuery48.sql -...AH\Mayurresh (55))*
SQLQuery47.sql -...AH\Mayurresh (58))
100 %
Results
Messages

```
select count (*) from Dim_SellerDetails;
select count (*) From DimProperty_Details;
select Count (*) from DimProperty_Rooms;
Select Count (*)  from Fact_Sold;
```

```
   (No column name)
1  27160

   (No column name)
1  13580

   (No column name)
1  13580

   (No column name)
1  13580
```

All Details are loaded