Project Notebook

IMPLEMENTING BUSINESS INTELLIGENCE FOR GADGET-MART

Apoorva Balasubramanian

Table of Contents

1. Introduction

1.1 Overview

Gadget-Mart is a multinational retail company that mainly deals with electronic products. The company is spread across different geographies across the world and has several product lines. This company does its business through online and stores. The Sales Team of Gadget-Mart would like to have data out of all their stores (and online sales) to be converted into useful information based on which they can analyze the Sales performance which will aid them in making vital decisions for allocating budgets.

.

1.2 About the Project

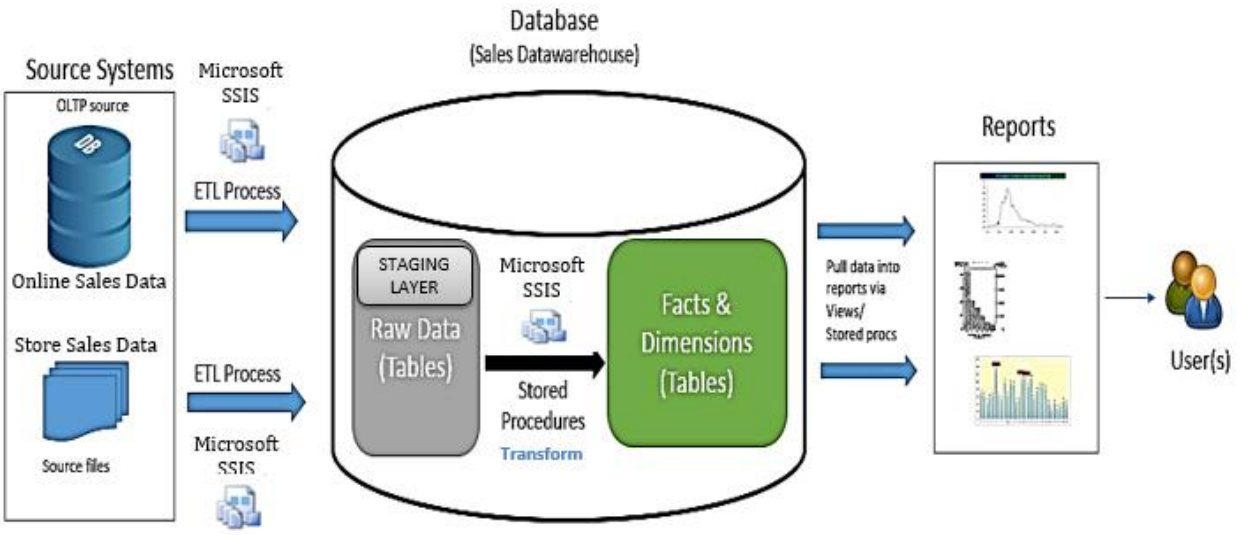
We have built a Sales Data Platform which will accumulate data across all the systems (Online Sales and Store Sales) on Daily basis and consolidate in a Data warehouse into Facts and Dimensions. We have develop reports that will project this vast data into user friendly dashboard/charts that will help stakeholders to take decisions.

.

2. Architecture and Workflow

The company has many stores across the world that provide data of their daily transactions in flat files whereas the online sale data in captured in one or more OLTP systems. Our goal is to consolidate all the heterogeneous datasets into a single data warehouse from where we can build reports/dashboards for users.

The below figure provides a high level overview of the data flow and important components in our system.



3. Components and Functionalities:

3.1 Source:

Transaction data of stores (in files) and online sales data comes from OLTP system. (This data is simulated from Microsoft AdventureWorks Database).

3.2 ETL Process (Extraction Transform Load):

Extracts the data from source systems and load it into staging tables. We have used Microsoft SQL SERVER INTEGRATION SERVICES to implement the ETL process.
All tables in the destination will be truncated and then loaded.

NAME	TYPE OF ARTIFACT	DESCRIPTION
Online_To_Staging.dtsx	SSIS Package	This package imports the data from online OLTP system to the Staging layer
Store_To_Staging.dtsx	SSIS Package	This package imports the data from the flat files (Store Sales data) to the Staging layer

3.3 Staging tables:

They hold the raw source data. Staging tables will be named under separate schema called "STAGING". E.g. "Staging. Sales". This is to logically differentiate the Staging tables.
We have implemented default values to track the 'Staging Created Date' and 'Staging Created By'

Staging Tables
Staging.CountryMaster
Staging.CustomerMaster
Staging.OnlineSalesOrderDetail
Staging.OnlineSalesOrderHeader
Staging.ProductCategory
Staging.ProductMaster
Staging.SalesTerritory
Staging.StoreSales

3.4 Dimensions and Facts:

These are our main tables where-in our master/Lookup data (in Dimensions) and Measures like Sales data (in Facts) will reside.

1) All the dimension have the Identity column which will generate the DW keys.
2) All the dimensions tables will have primary key.
3) All keys in Facts (FK) refer to the Dimensions (PK)

TABLE NAME	TYPE	DESCRIPTION
DimCountry	Dimension	This table has master data of all countries
DimCustomer	Dimension	This table has master data of all customers
DimDate	Dimension	This table has the date keys for all the calendar dates. This table will be used to map the corresponding month/quarter/year for any date
DimLocation	Dimension	This table has master data of all locations at a city level
DimProduct	Dimension	This table has master data of all the products and their respective product groups
FactOnlineSales	Fact	This table has transactional data of Online Sales
FactStoreSales	Fact	This table has transactional data of Store Sales

3.5 Stored Procedures to transform the data:

We have created stored procedures that transform the data from Raw staging tables to the respective facts and dimensions. E.g. sp_LoadDimCountry will load the master country data into DimCountry.

In each procedure, following functionalities are implemented:

1) We are using variables to store the 'created date' and 'created by' attributes.
2) We use the in-built functions like GETDATE () and SYSTEM_USER VARIABLE
3) We use **LEFT** and **INNER joins** to pull data from multiple source/dimension tables.
4) We use a **temporary tables (# table)** to store the transformed source data.
5) We sync source data stored in the # table with the target table using MERGE functionality.

The Merge functionality has the following scenarios:

a) If data is present in source and not in target then: - **INSERT RECORD.**
b) If data is present in source and present in target :- **UPDATE RECORD**
c) If data is not present in source and present in target: - **SOFT DELETE** THE RECORD IN TARGET (i.e. Mark the IsActive flag in the destination as 0).

STORE PROCEDURES	DESCRIPTION
sp_LoadFactOnlineSales	Transforms and Loads data into FactOnlineSales table from Staging
sp_LoadDimCountry	Transforms and Loads data into DimCountry table from Staging
sp_LoadDimProduct	Transforms and Loads data into DimProduct table from Staging
sp_LoadDimLocation	Transforms and Loads data into DimLocation table from Staging
sp_LoadDimCustomer	Transforms and Loads data into DimCustomer table from Staging
sp_LoadFactStoreSales	Transforms and Loads data into FactStoreSales table from Staging

We use SSIS package to execute all these store procedures:

PACKAGE	DESCRIPTION
EDW-Load-Facts_Dimensions.dtsx	This package executes all the store procedure that loads facts and dimensions in the prescribed order of execution

3.6 Views:

We have created de-normalized views on top of our tables that can readily be used by Reports to pull in the data.

VIEW NAME	DESCRIPTION
vwFactOnlineSales	Projects the Online Sales data in a de-normalized format. This view is built on top of FactOnlineSales
vwFactStoreSales	Projects the Store Sales data in a de-normalized format. This view is built on top of FactStoreSales

3.7 Reports:

Reports are created to project the data to users in dashboard, charts or drill down format across various subject areas.

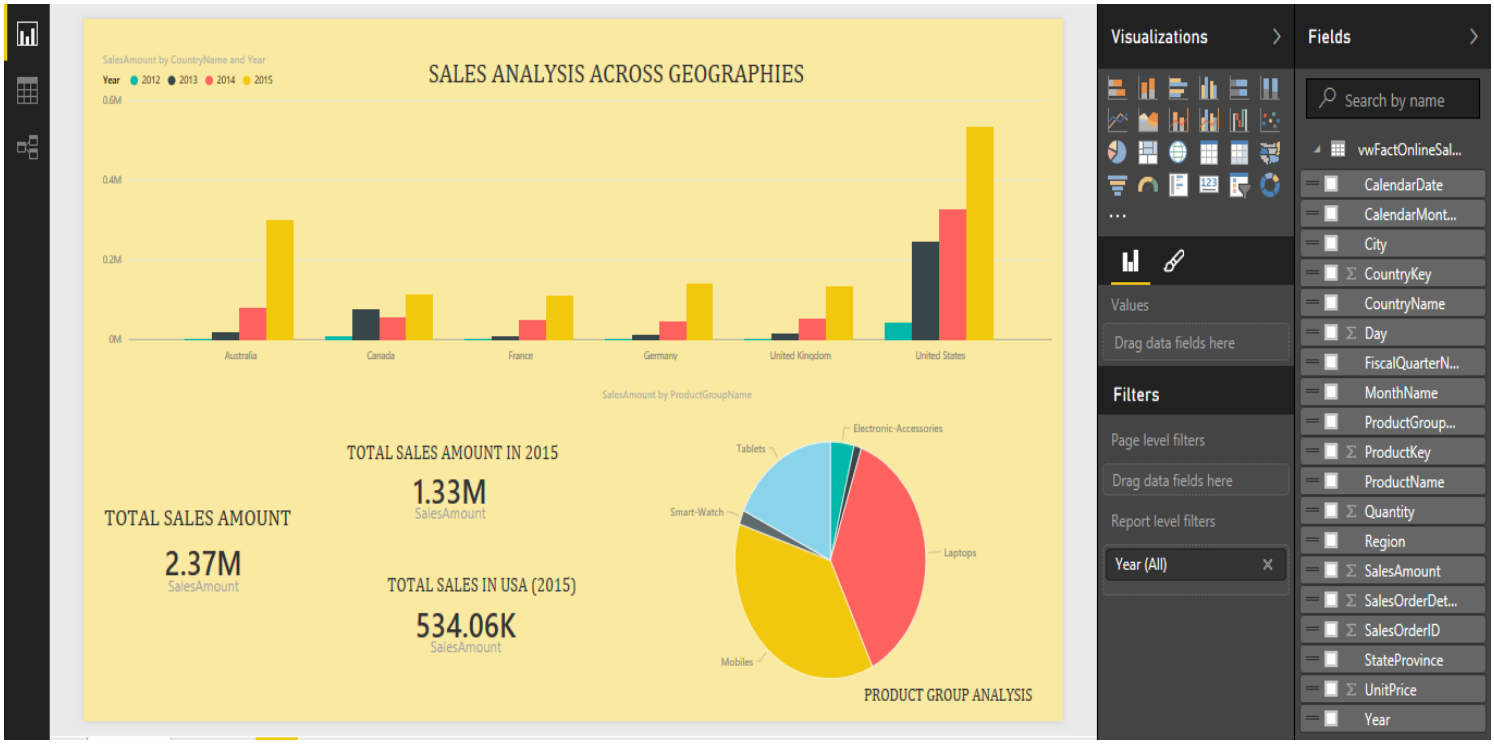
Following are the reports that are generated on top of FACTS and DIMENSIONS of datawarehouse

EXCEL REPORTS:

REPORT NAME	DESCRIPTION
Sales Trend Analysis	Provides a breakdown of Sales data across geographies and fiscal year
Product Sale Analysis	Provides detailed analysis of Sales across product lines
Apple Watch Trend	Gives the performance trend analysis of Apple watch over the last 3 months

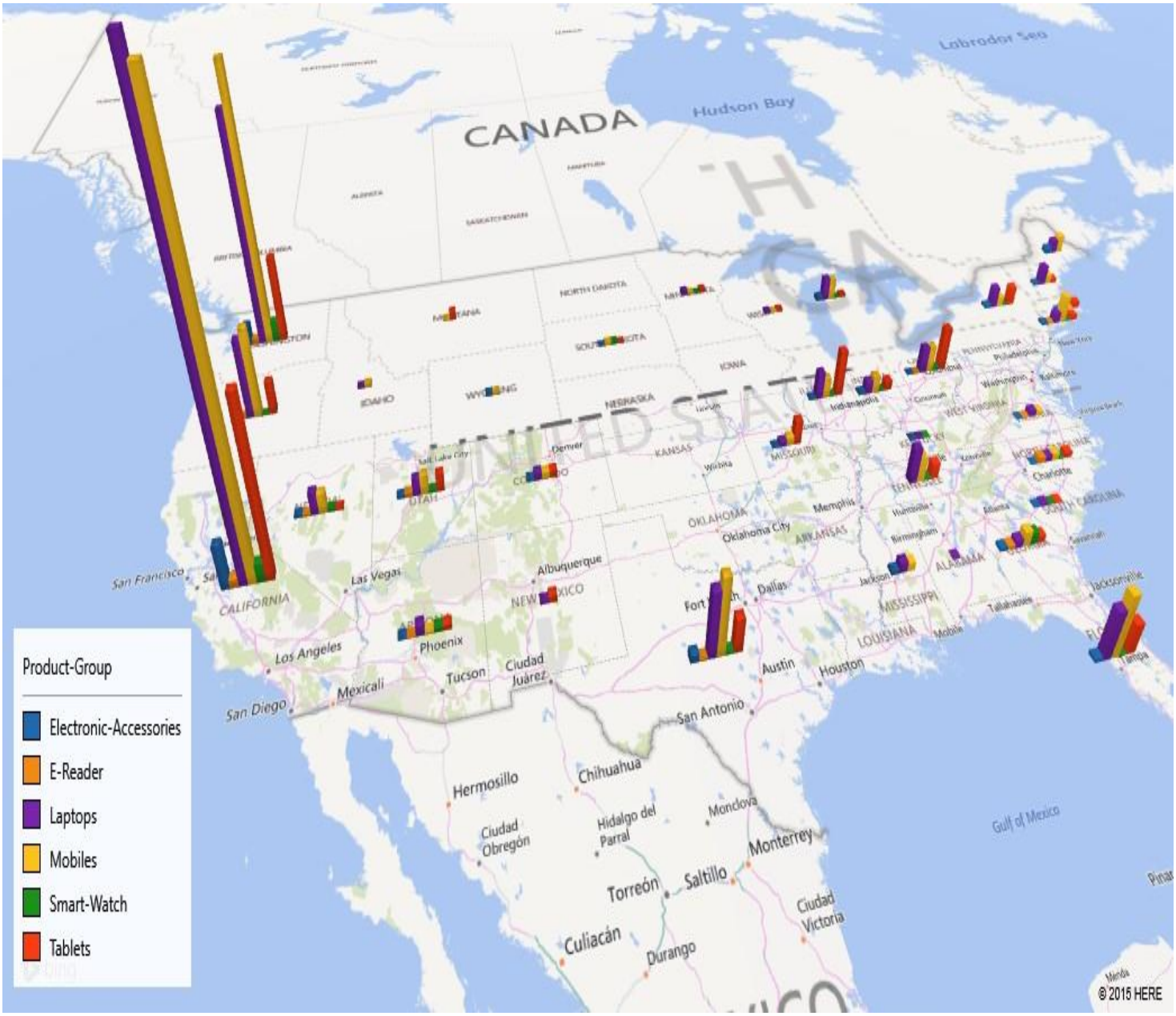
POWER BI DASHBOARD

The Power BI dashboard provides instant snapshot of the sales data across different dimensions and also provides important numbers for the Sales team.



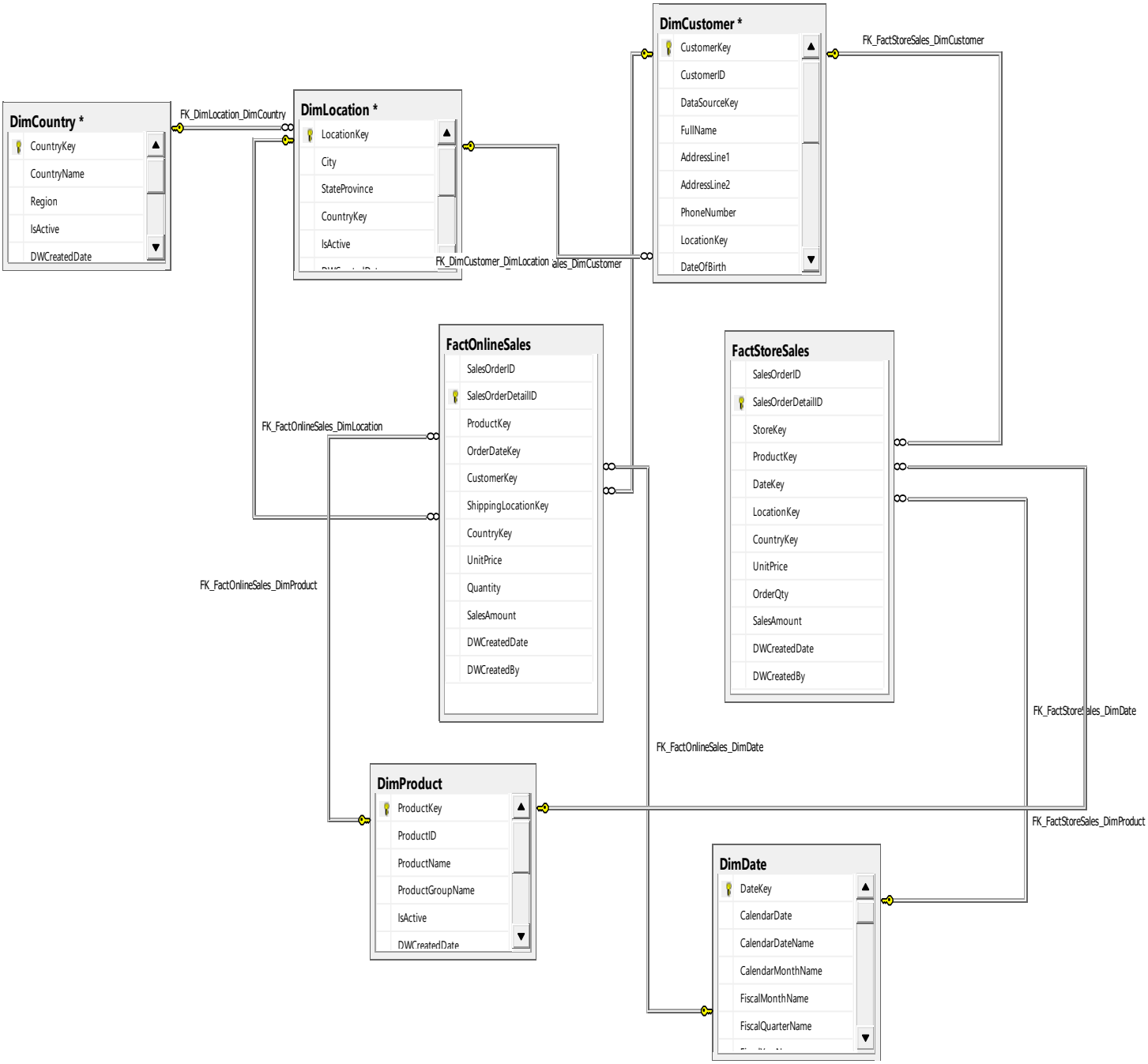
Power Map

Power Map provides a 3D data visualization of the data on a geographical map. Following is the screen shot which has been taken after the DW data was projected using Power Map.

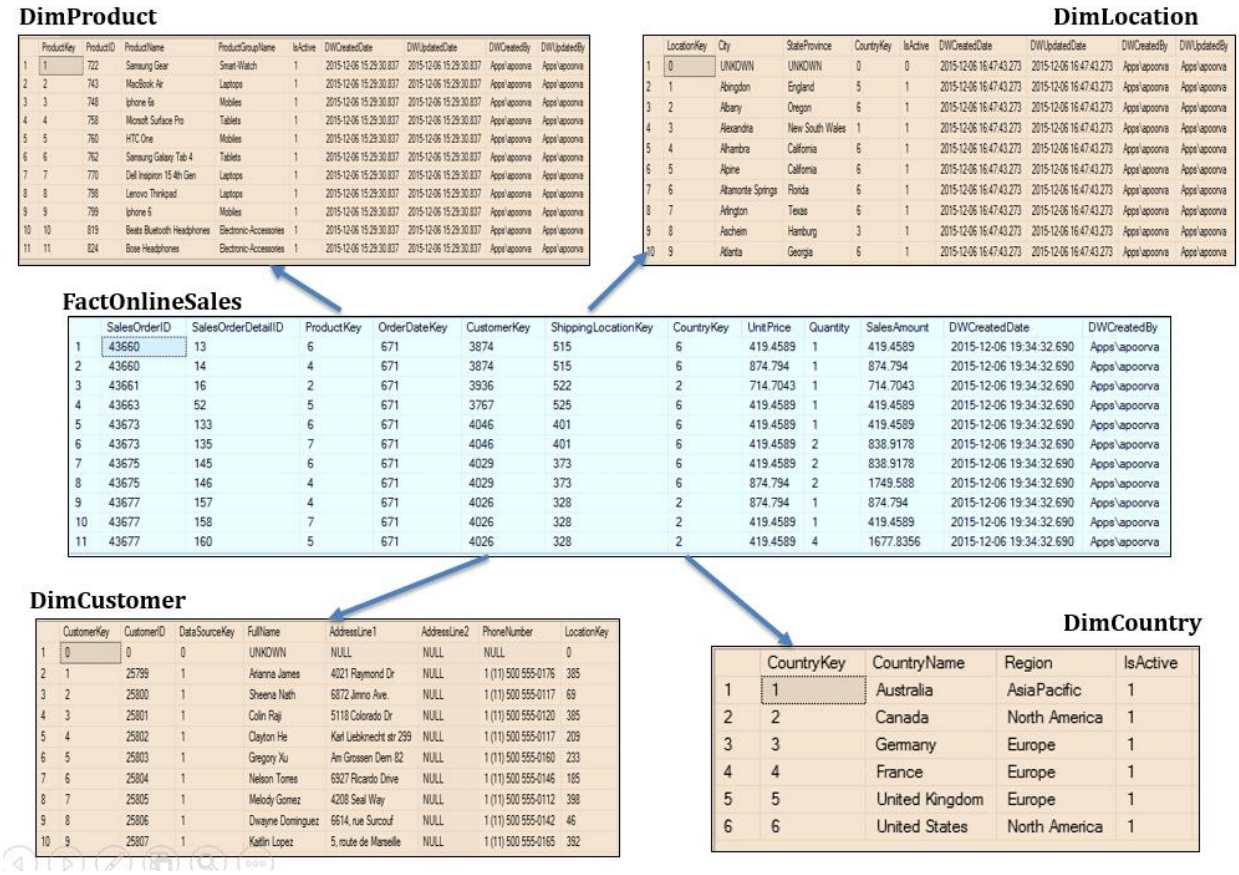


4 Database Design and Sample Data:

The following is the database design for this project:



Sample Data:



DimProduct

	ProductKey	ProductID	ProductName	ProductGroupName	IsActive	DWCreatedDate	DWUpdatedDate	DWCreatedBy	DWUpdatedBy
1	1	722	Samsung Gear	Smart-Watch	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
2	2	743	MacBook Air	Laptops	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
3	3	748	Iphone 6s	Mobiles	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
4	4	758	Microsoft Surface Pro	Tablets	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
5	5	760	HTC One	Mobiles	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
6	6	762	Samsung Galaxy Tab 4	Tablets	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
7	7	770	Dell Inspiron 15 4th Gen	Laptops	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
8	8	798	Lenovo Thinkpad	Laptops	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
9	9	799	Iphone 6	Mobiles	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
10	10	819	Beats Bluetooth Headphones	Electronic-Accessories	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva
11	11	824	Bose Headphones	Electronic-Accessories	1	2015-12-06 15:29:30.837	2015-12-06 15:29:30.837	Apps\apoorva	Apps\apoorva

DimLocation

	LocationKey	City	StateProvince	CountryKey	IsActive	DWCreatedDate	DWUpdatedDate	DWCreatedBy	DWUpdatedBy
1	0	UNKOWN	UNKOWN	0	0	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
2	1	Abingdon	England	5	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
3	2	Albany	Oregon	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
4	3	Alexandria	New South Wales	1	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
5	4	Alhambra	California	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
6	5	Alpine	California	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
7	6	Altamonte Springs	Florida	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
8	7	Arlington	Texas	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
9	8	Ascheim	Hamburg	3	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva
10	9	Atlanta	Georgia	6	1	2015-12-06 16:47:43.273	2015-12-06 16:47:43.273	Apps\apoorva	Apps\apoorva

FactOnlineSales

	SalesOrderID	SalesOrderDetailID	ProductKey	OrderDateKey	CustomerKey	ShippingLocationKey	CountryKey	UnitPrice	Quantity	SalesAmount	DWCreatedDate	DWCreatedBy
1	43660	13	6	671	3874	515	6	419.4589	1	419.4589	2015-12-06 19:34:32.690	Apps\apoorva
2	43660	14	4	671	3874	515	6	874.794	1	874.794	2015-12-06 19:34:32.690	Apps\apoorva
3	43661	16	2	671	3936	522	2	714.7043	1	714.7043	2015-12-06 19:34:32.690	Apps\apoorva
4	43663	52	5	671	3767	525	6	419.4589	1	419.4589	2015-12-06 19:34:32.690	Apps\apoorva
5	43673	133	6	671	4046	401	6	419.4589	1	419.4589	2015-12-06 19:34:32.690	Apps\apoorva
6	43673	135	7	671	4046	401	6	419.4589	2	838.9178	2015-12-06 19:34:32.690	Apps\apoorva
7	43675	145	6	671	4029	373	6	419.4589	2	838.9178	2015-12-06 19:34:32.690	Apps\apoorva
8	43675	146	4	671	4029	373	6	874.794	2	1749.588	2015-12-06 19:34:32.690	Apps\apoorva
9	43677	157	4	671	4026	328	2	874.794	1	874.794	2015-12-06 19:34:32.690	Apps\apoorva
10	43677	158	7	671	4026	328	2	419.4589	1	419.4589	2015-12-06 19:34:32.690	Apps\apoorva
11	43677	160	5	671	4026	328	2	419.4589	4	1677.8356	2015-12-06 19:34:32.690	Apps\apoorva

DimCustomer

	CustomerKey	CustomerID	DataSourceKey	FullName	AddressLine1	AddressLine2	PhoneNumber	LocationKey
1	0	0	0	UNKOWN	NULL	NULL	NULL	0
2	1	25799	1	Arianna James	4021 Raymond Dr	NULL	1 (11) 500 555-0176	385
3	2	25800	1	Sheena Nath	6872 Jimno Ave.	NULL	1 (11) 500 555-0117	69
4	3	25801	1	Colin Raji	5118 Colorado Dr	NULL	1 (11) 500 555-0120	385
5	4	25802	1	Clayton He	Karl Liebknecht str 299	NULL	1 (11) 500 555-0117	209
6	5	25803	1	Gregory Xu	Am Grossen Dem 82	NULL	1 (11) 500 555-0160	233
7	6	25804	1	Nelson Torres	6927 Ricardo Drive	NULL	1 (11) 500 555-0146	185
8	7	25805	1	Melody Gomez	4208 Seal Way	NULL	1 (11) 500 555-0112	398
9	8	25806	1	Dwayne Dominguez	6614, rue Surcouf	NULL	1 (11) 500 555-0142	46
10	9	25807	1	Kaitlin Lopez	5, route de Marseille	NULL	1 (11) 500 555-0165	392

DimCountry

	CountryKey	CountryName	Region	IsActive
1	1	Australia	AsiaPacific	1
2	2	Canada	North America	1
3	3	Germany	Europe	1
4	4	France	Europe	1
5	5	United Kingdom	Europe	1
6	6	United States	North America	1

5 Deployment:

All the scripts and deployment steps are placed in the parent folder of this document.

Please refer to the Readme file for more details.