Medicare Provider Utilization And Payment Data

Database Warehouse And Business Intelligence

Prepared By:
Shaishav Shah

Objective

Analyze cost and treatment KPIs of different types of health service provider and also, analyze the nursing home facilities condition and Inpatient healthcare conditions.

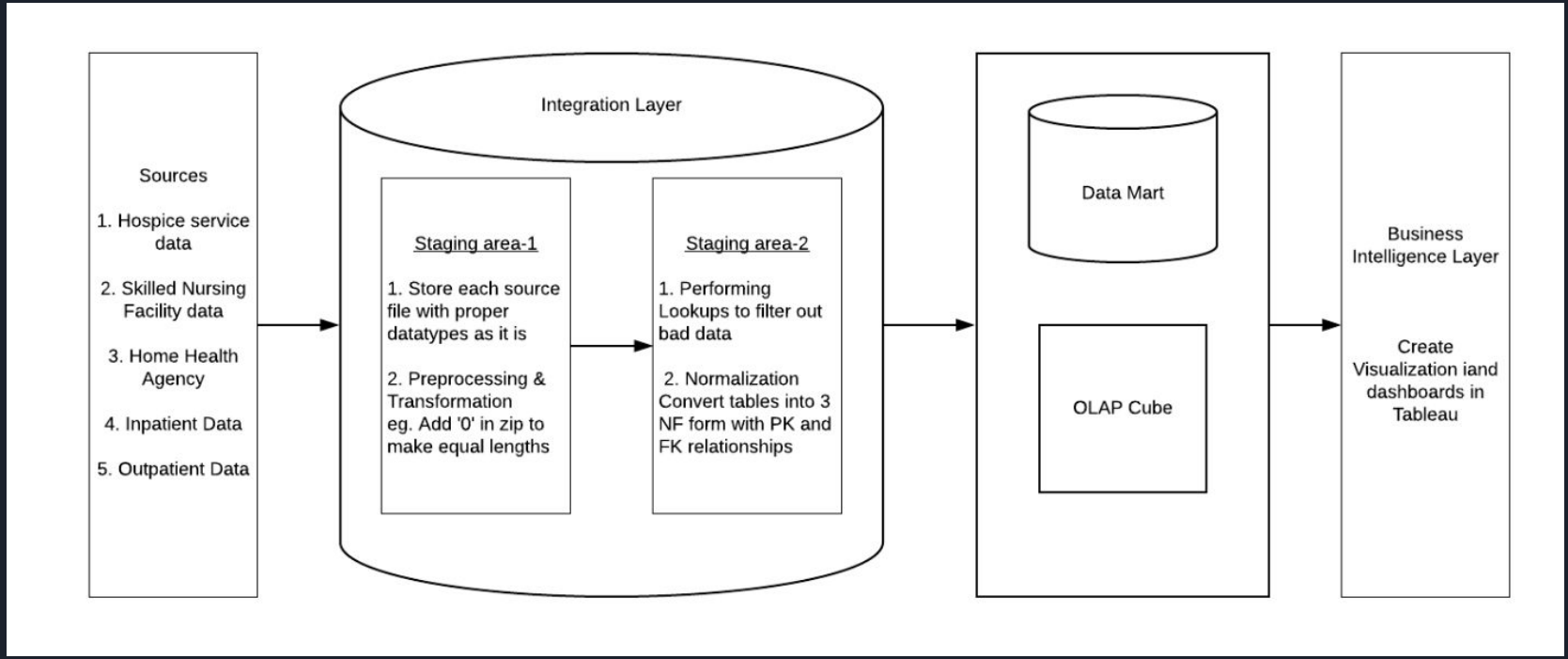
Data Source

Data was collected from from CMS website. The data set consist of five (5) files which were linked with each other based on their Provider ID over 2015 and 2016 year.

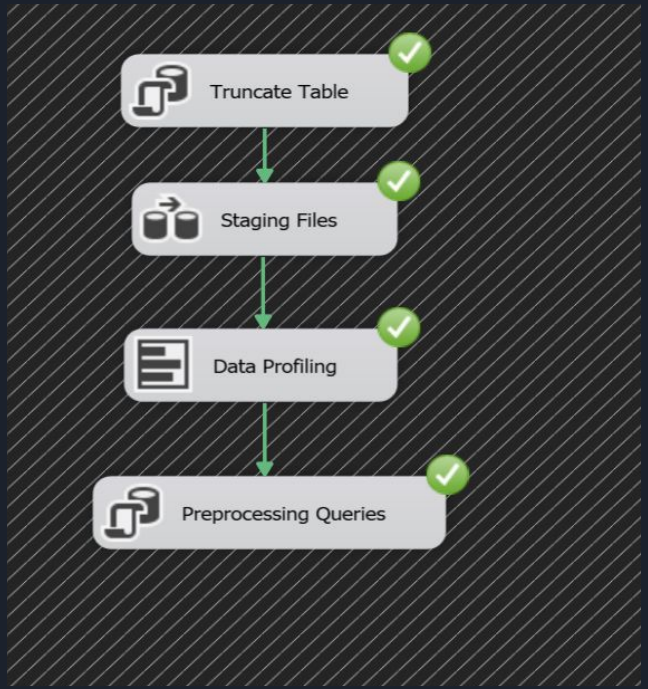
1. Medicare Provider Utilization and Payment Data: Skilled Nursing Facilities
2. Medicare Provider Utilization and Payment Data: Hospice Providers
3. Medicare Provider Utilization and Payment Data: Home Health Agencies
4. Medicare Provider Utilization and Payment Data: Outpatient
5. Medicare Provider Utilization and Payment Data: Inpatient

For all datasets key column is : Provider ID

Data and Process Flow

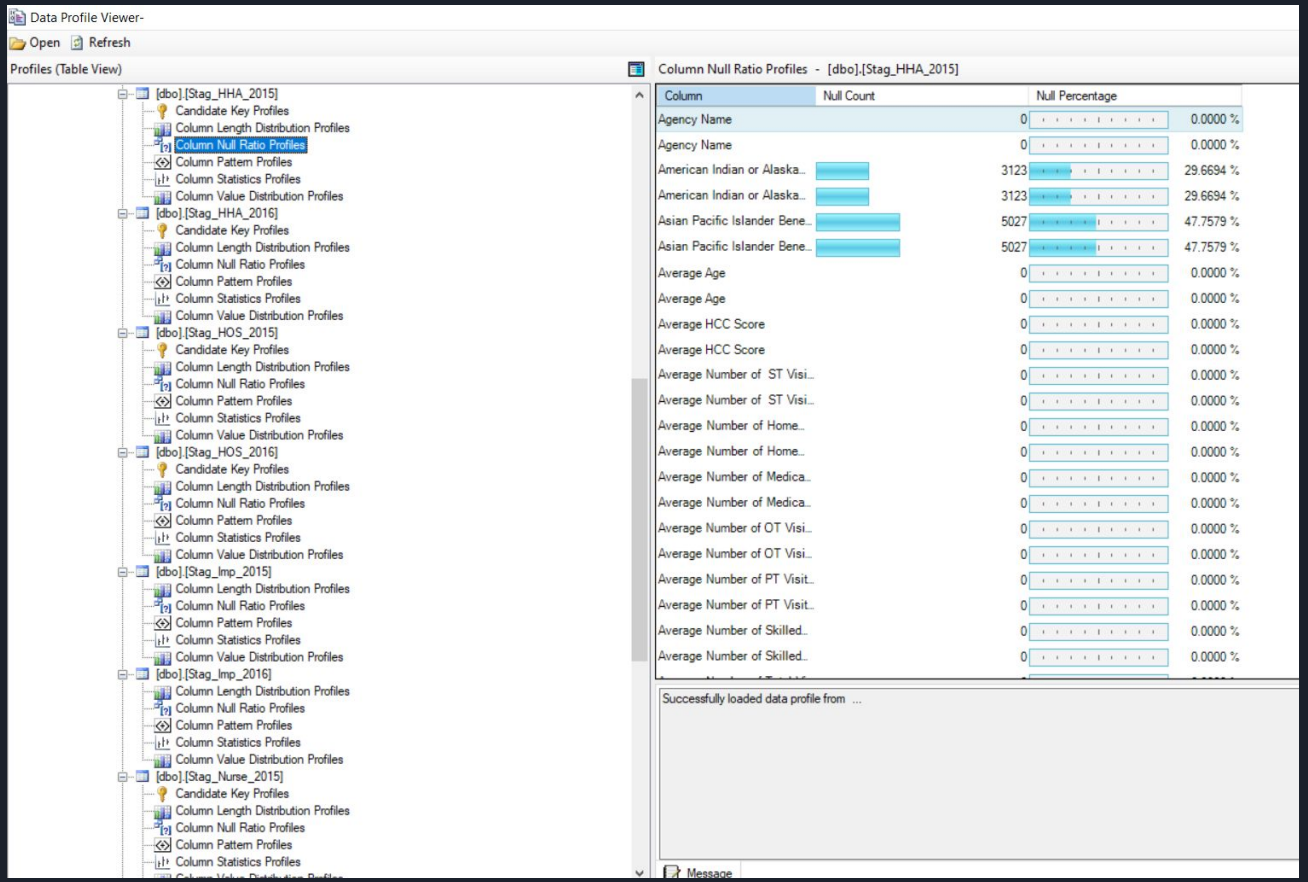


Data Profiling and Preprocessing



- Automated Data Profiling using in-built tool and SSIS Package
- Customised Pre-processing queries according to the report generated

Data Profiling and Preprocessing



Summary of Data Profiling

Some of our observations were:

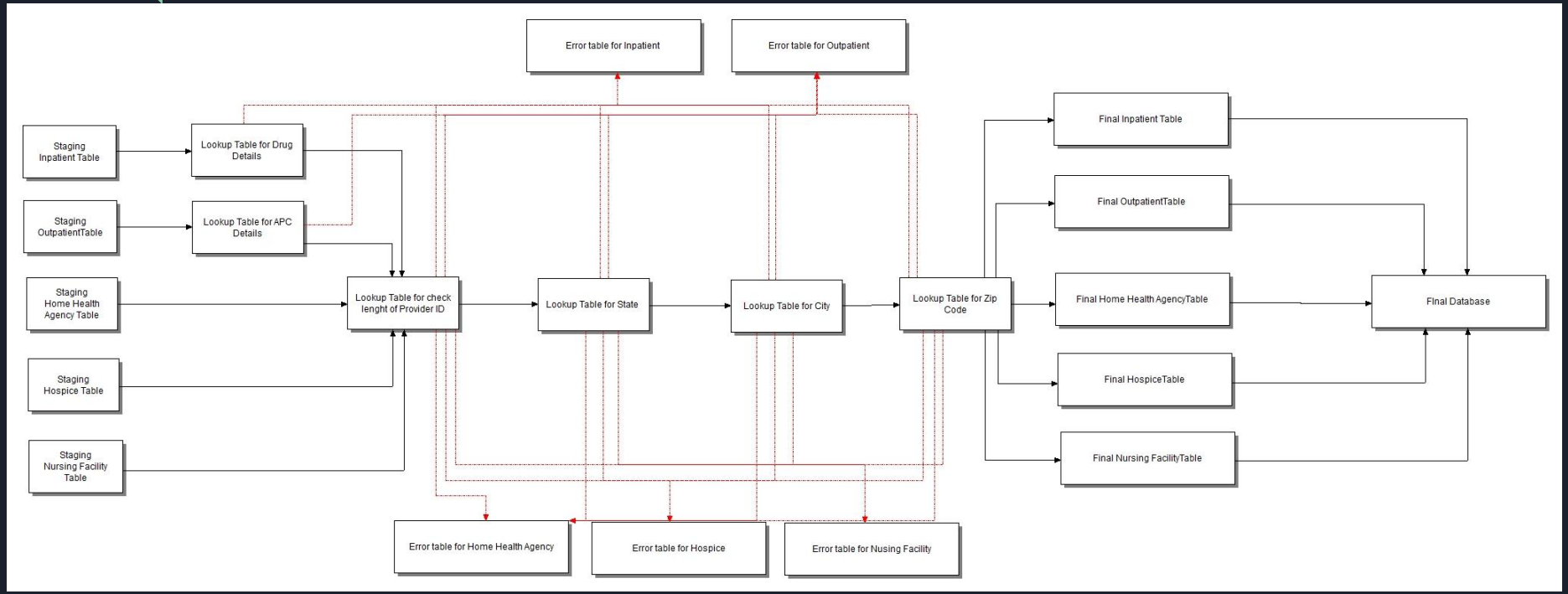
1. Zip Code were of 4 digits and 3 digits instead of 5 digits
   Example: 02215 becomes 2215
2. According to the definition, Provider ID was an unique 6 digit number. However, there were many Provider IDs which having only 5 digits.
3. There were few NULL values in the important columns like Zip Code, State Etc.
4. Data source was not consistent with Name. For example, few Provider Names were in Upper Case while Few were in lower case.

Summary of Data Preprocessing

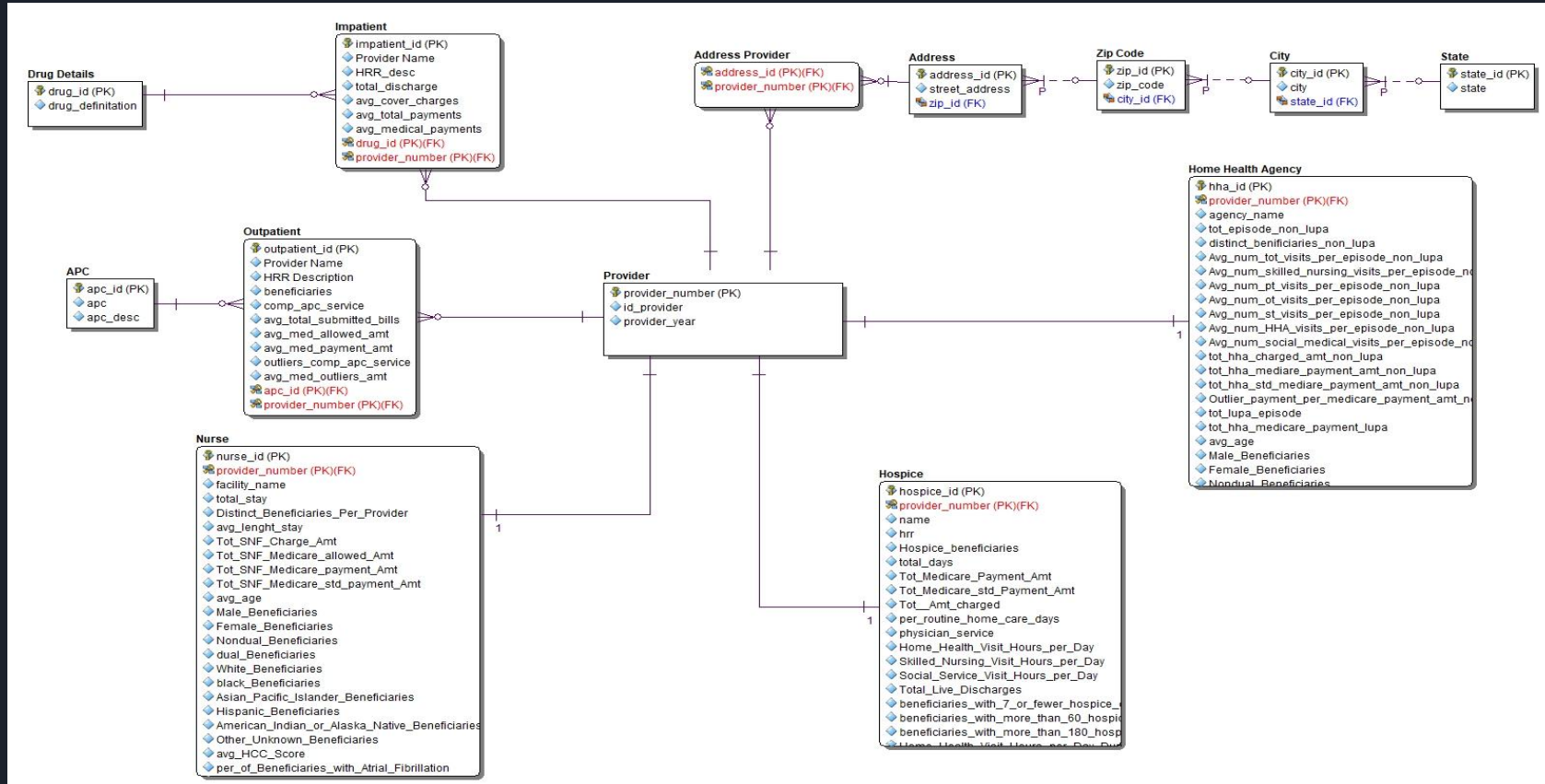
Data Preprocessing is important for data consistency. In order to make data pure and consistent, few of the techniques where used:

1. Since the Provider ID and Zip Code where of data type integers in the data source, the initial “0” were removed automatically in the data. We changed the data type to “varchar” and using the concatenating technique appended “0” in the required field
2. We removed Null values from the fields like Provider ID, State, City, APC Details, and Drug Details.
3. In order to maintain the naming consistency, we made all the rows of Provider Name, Street Address, and City lowercase and State Upper Case
4. We trimmed out white spaces and unwanted characters from important columns like Drug Definition which was leading to duplicate values.

Error Handling - Lookup Tables (Staging-2)



Data Schema - Physical Model (ERD Model)



Unique ID Creation

- We created an unique Id field using provider key and year of the data.

  E.g. For provider id ‘100010’ and year 2015, the unique id is 1000102015.

- This unique key is used as an primary key as the primary key in the destination table.

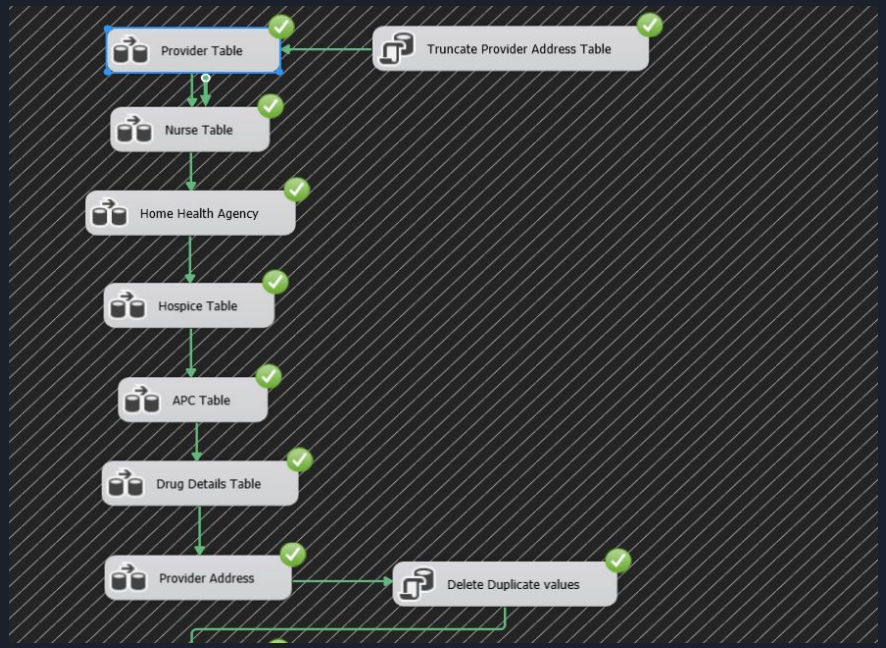
Data Loading for ERD

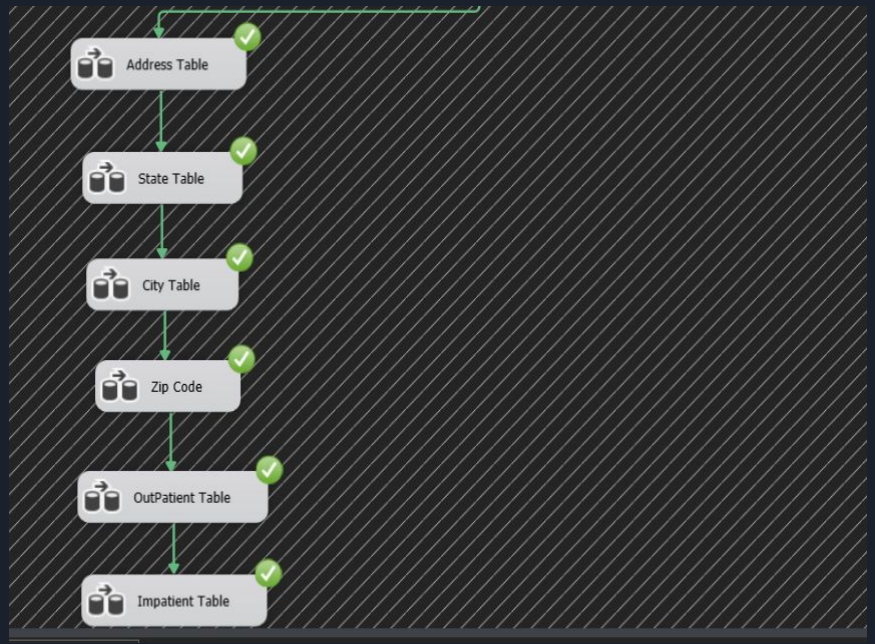
Data Flow Diagram was loading the data from source into our Data Schema.

Data Schema was designed with the help of ERStudios.

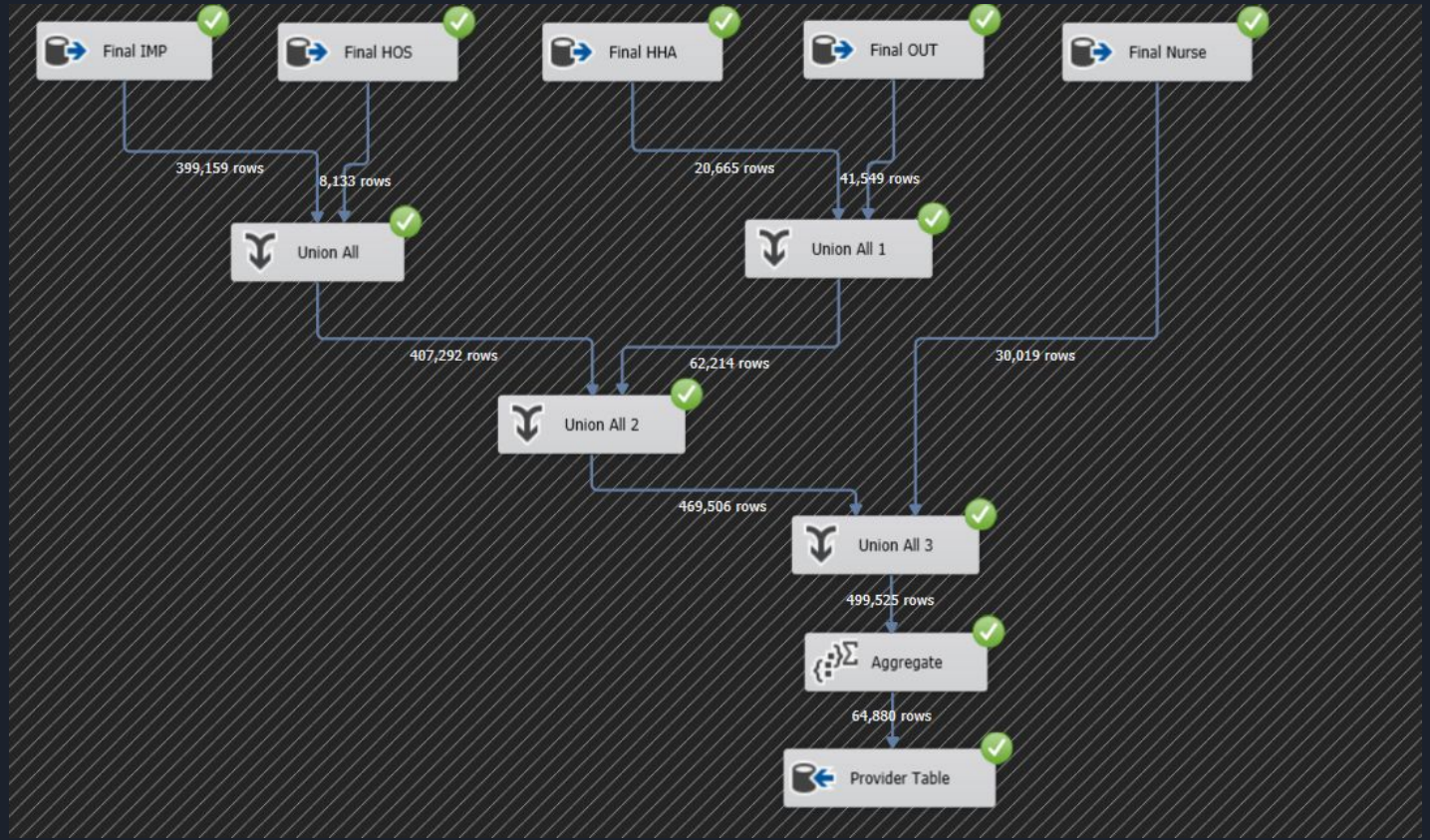
We used Visual Studio Integration Package (SSDT) to data integration.

Visual Studio’s SSIS View

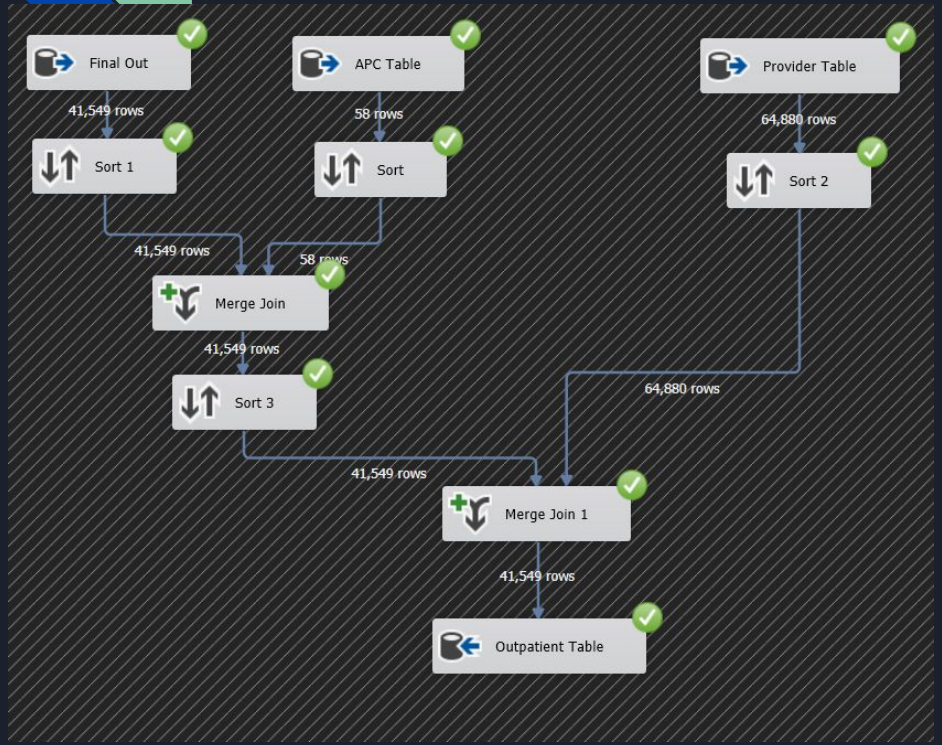


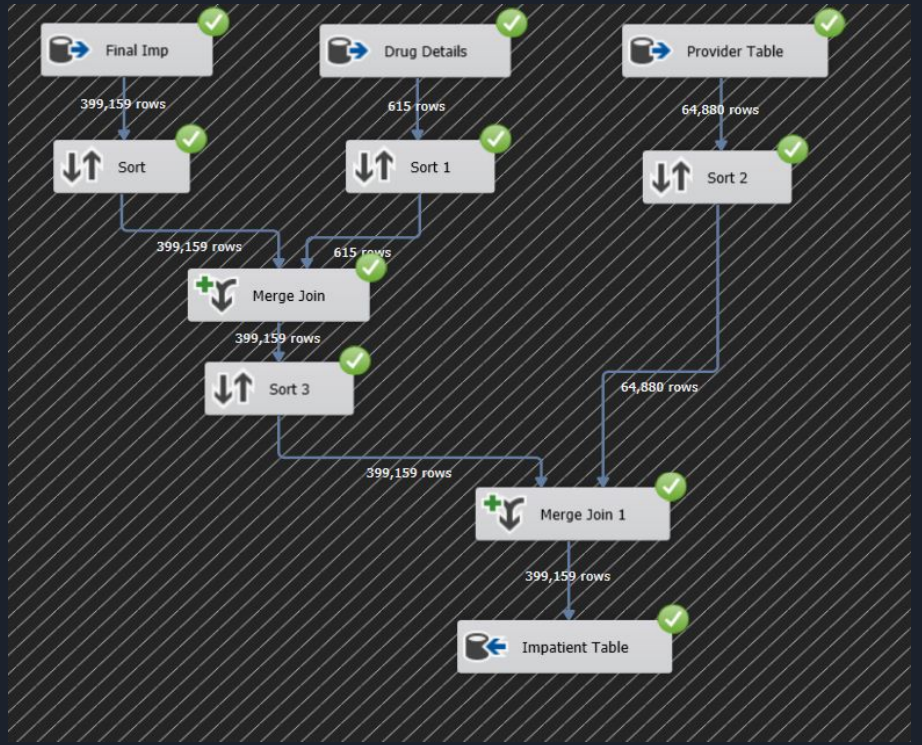


Data Flow for Provider Table

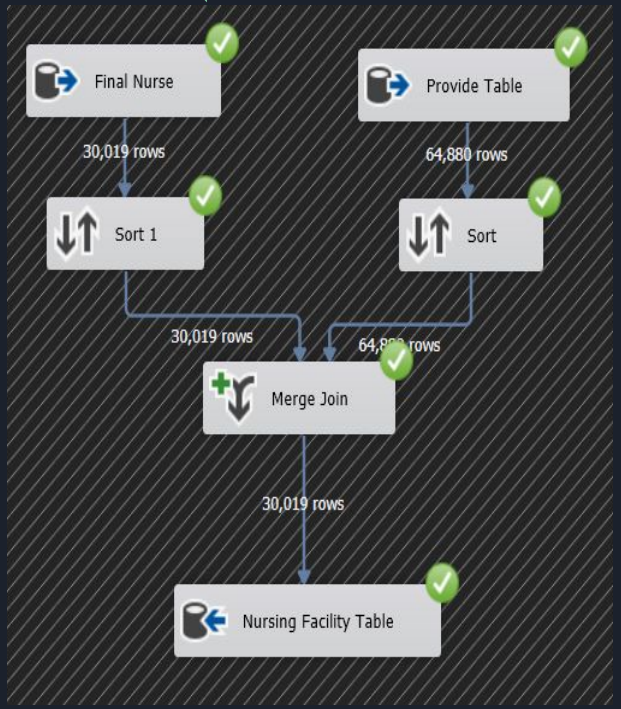


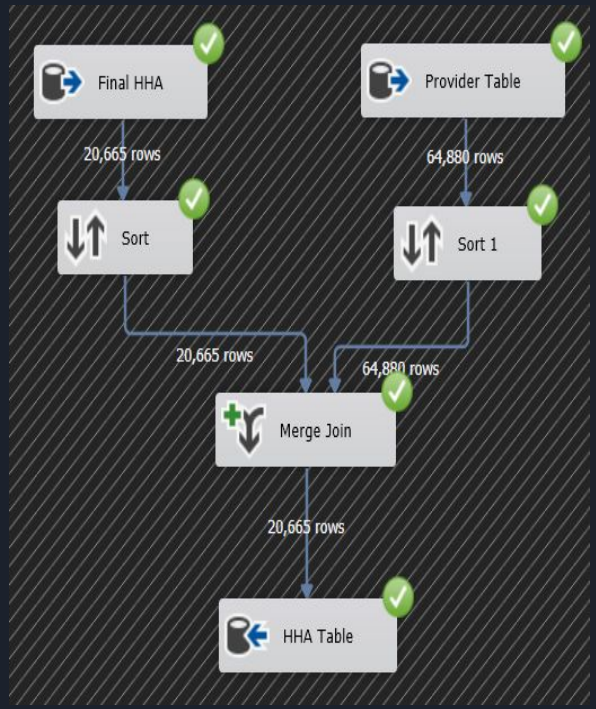
Data Flow for Inpatient and Outpatient Table

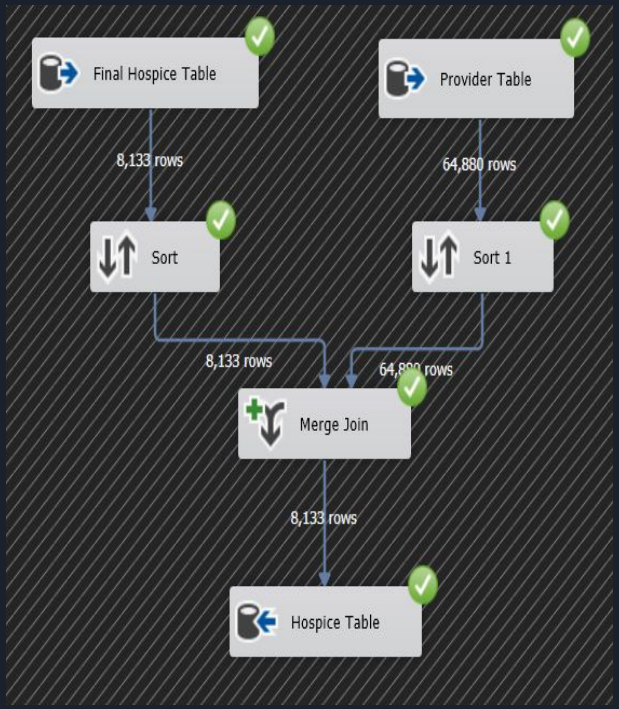




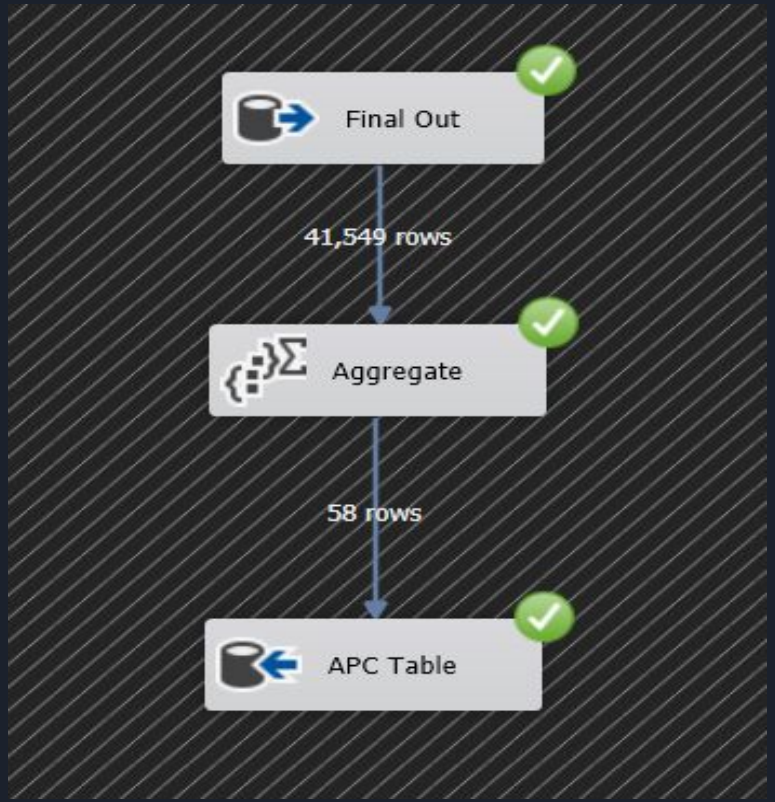
Data Flow for Nursing, Hospice and Home Health Agency Table



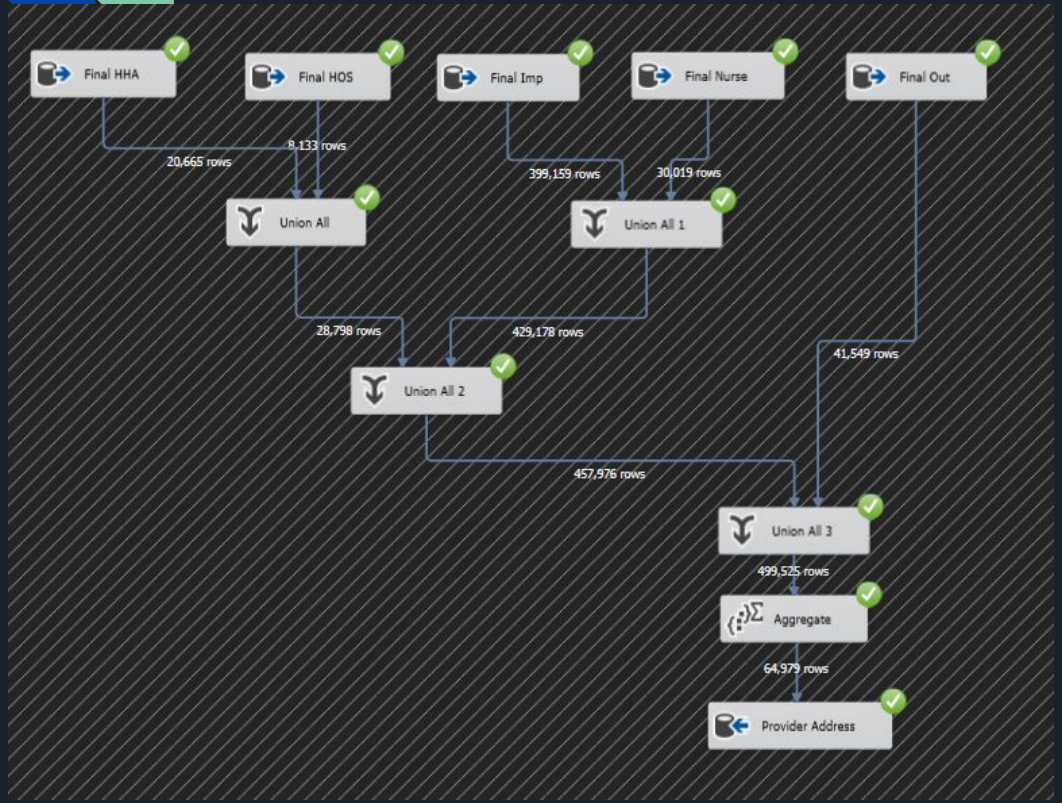


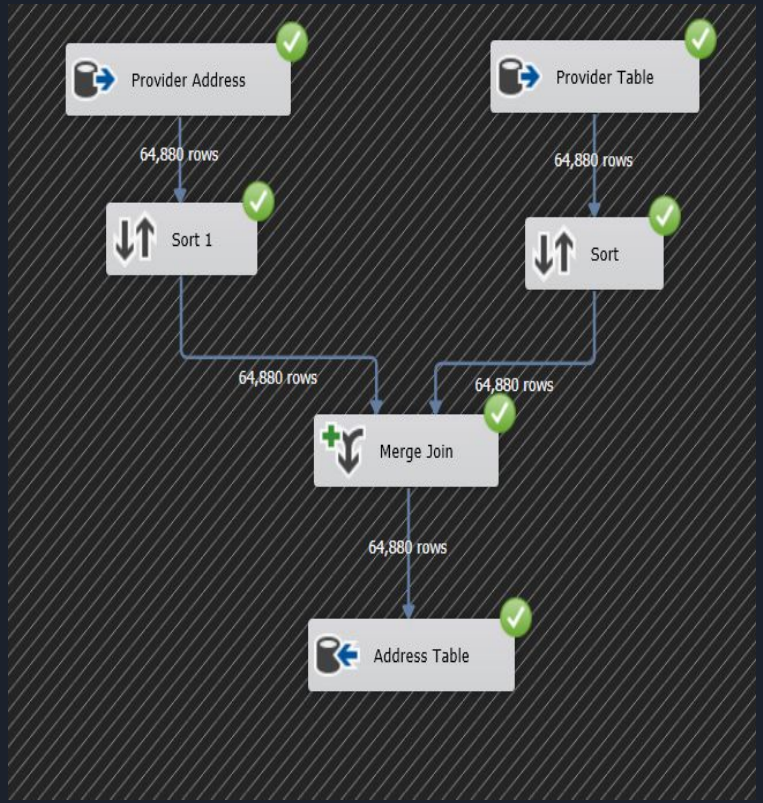


Data Flow for APC Details and Drug Details Table

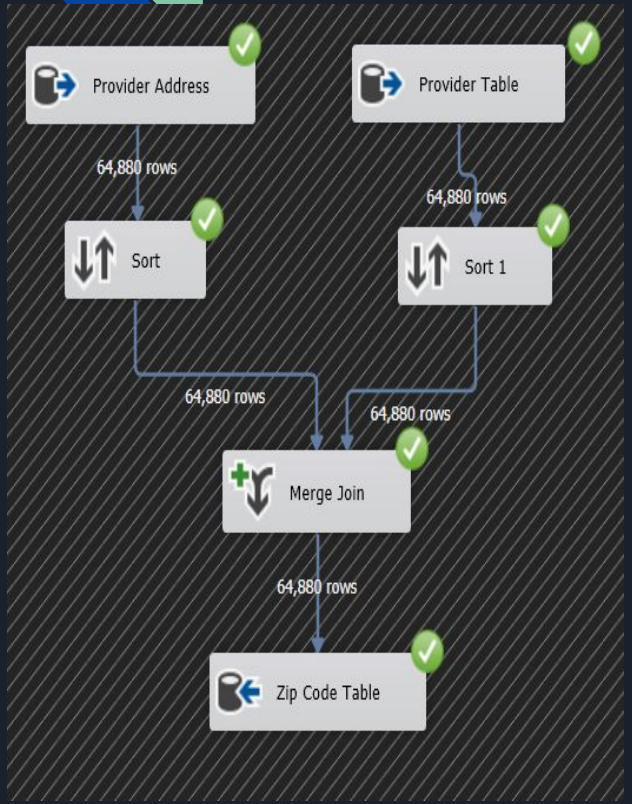


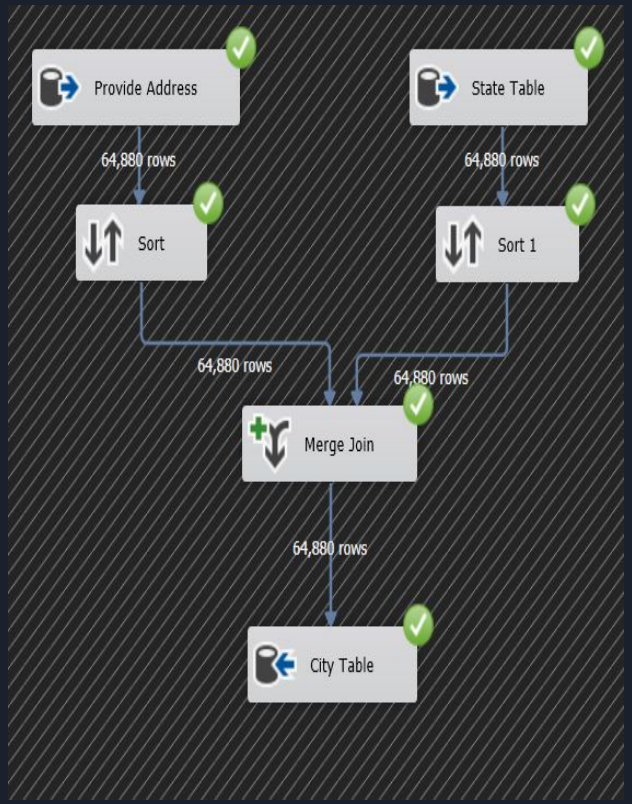
Data Flow for Provider Address and Address Table

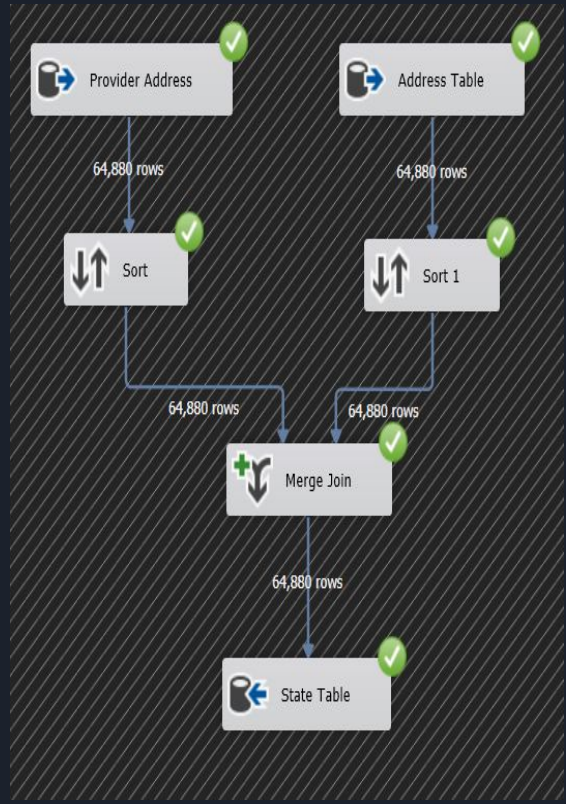




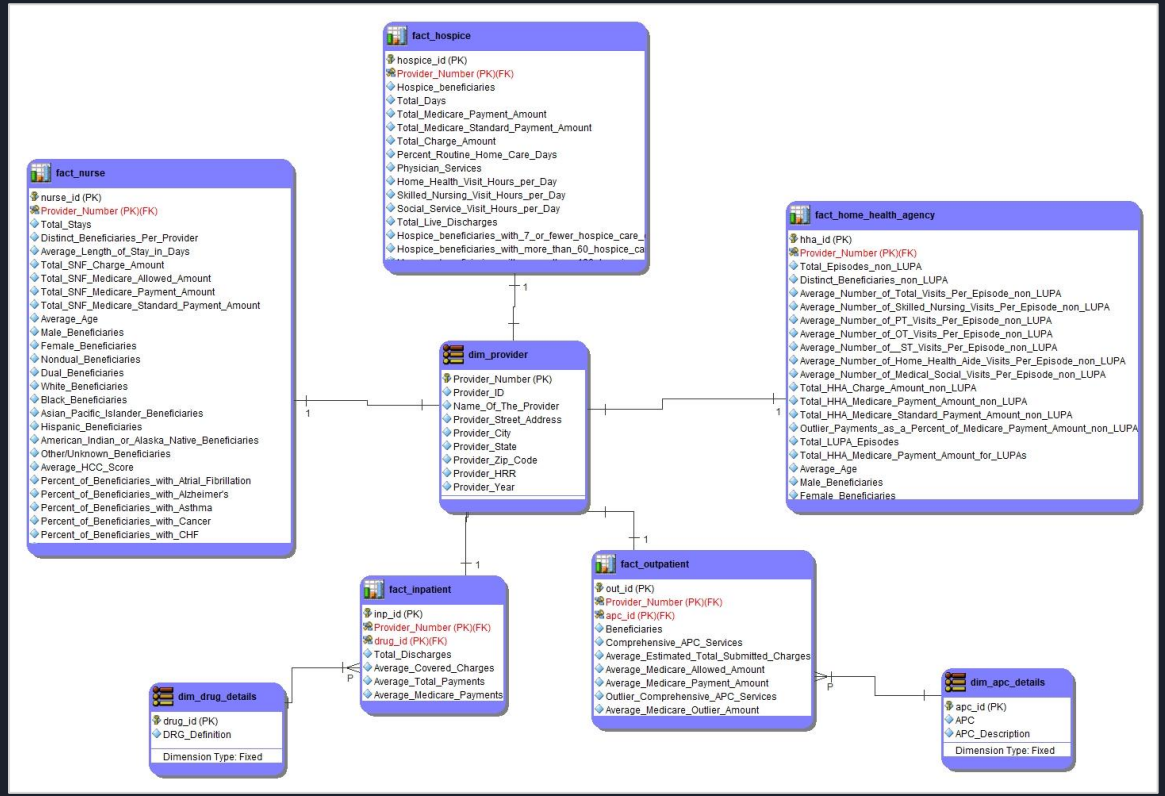
Data Flow for Zip Code, City, and State Tables





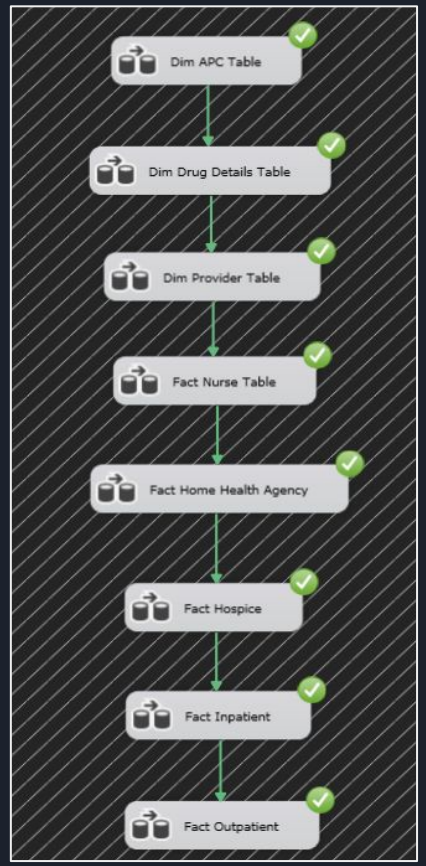


Data Mart : Dim Fact Model



- ERStudio
- 5 Fact Table
- 3 Dimensions
- Purpose:
  Dim Fact Model
  Multidimension
  Data Mart
  Visualization

ETL SSIS Package



- Software:

  Visual Studio

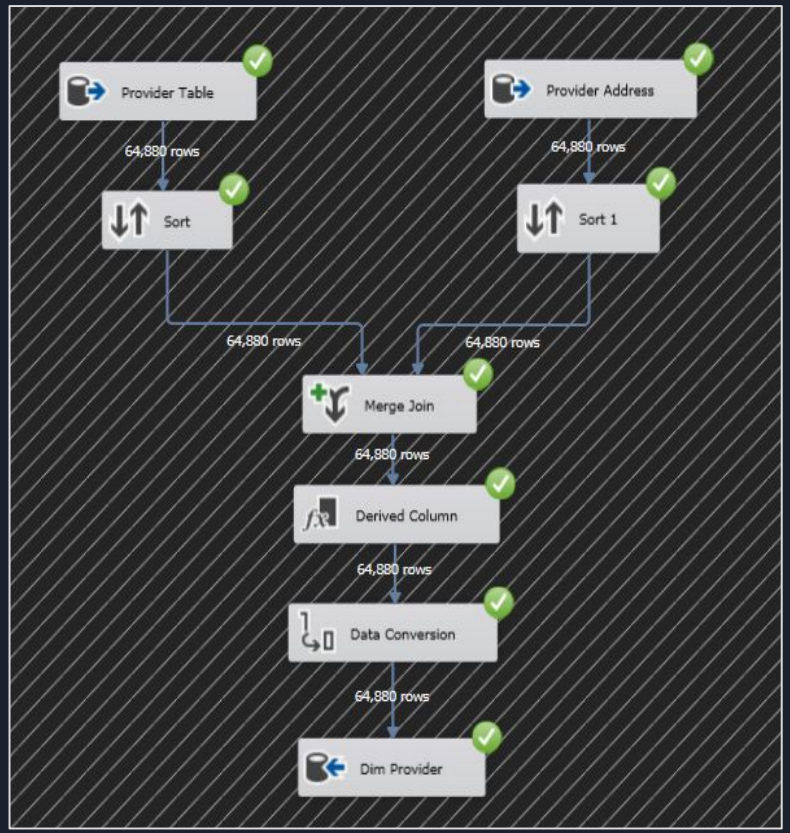
- Purpose:

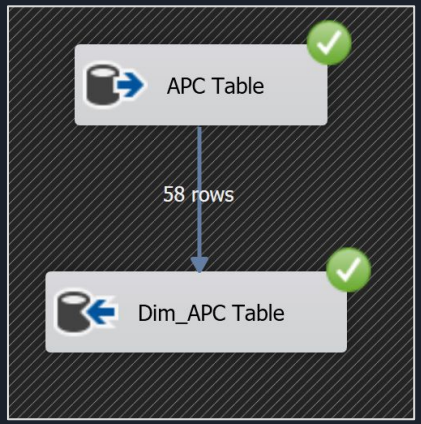
  Table Creation

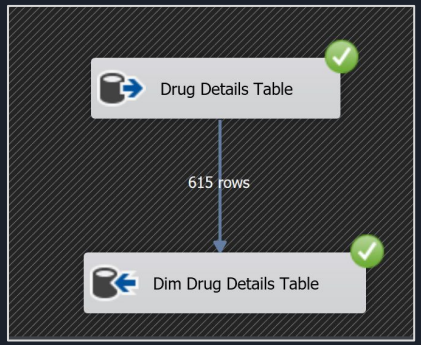
  Relationship (Parent-Child)

  Data Loading (From ERD Schema)

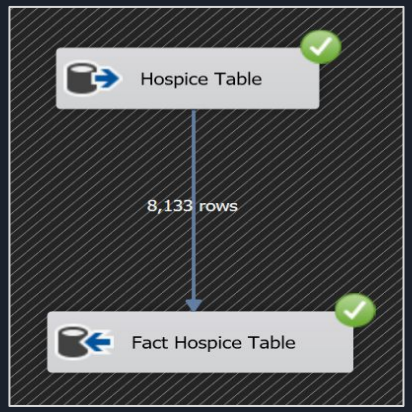
Dimension Table for Fact-Dim Model

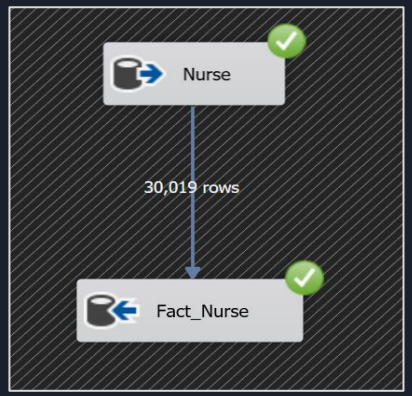


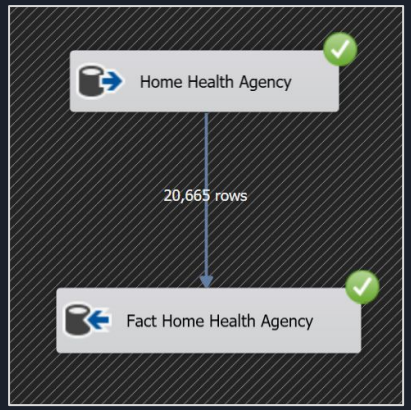


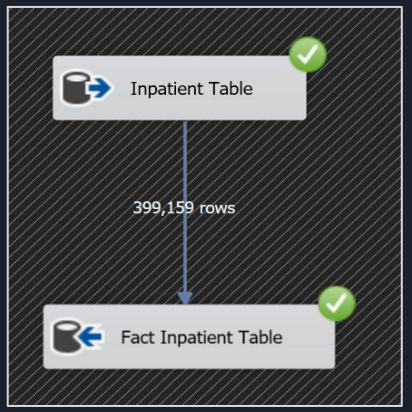


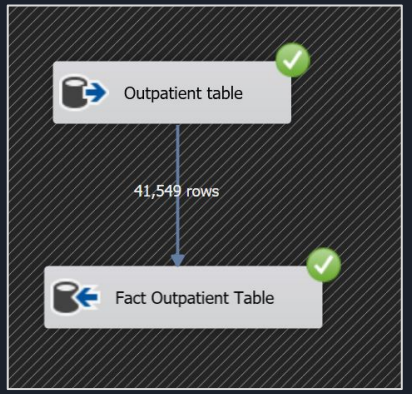
Fact Table for Fact-Dim Model



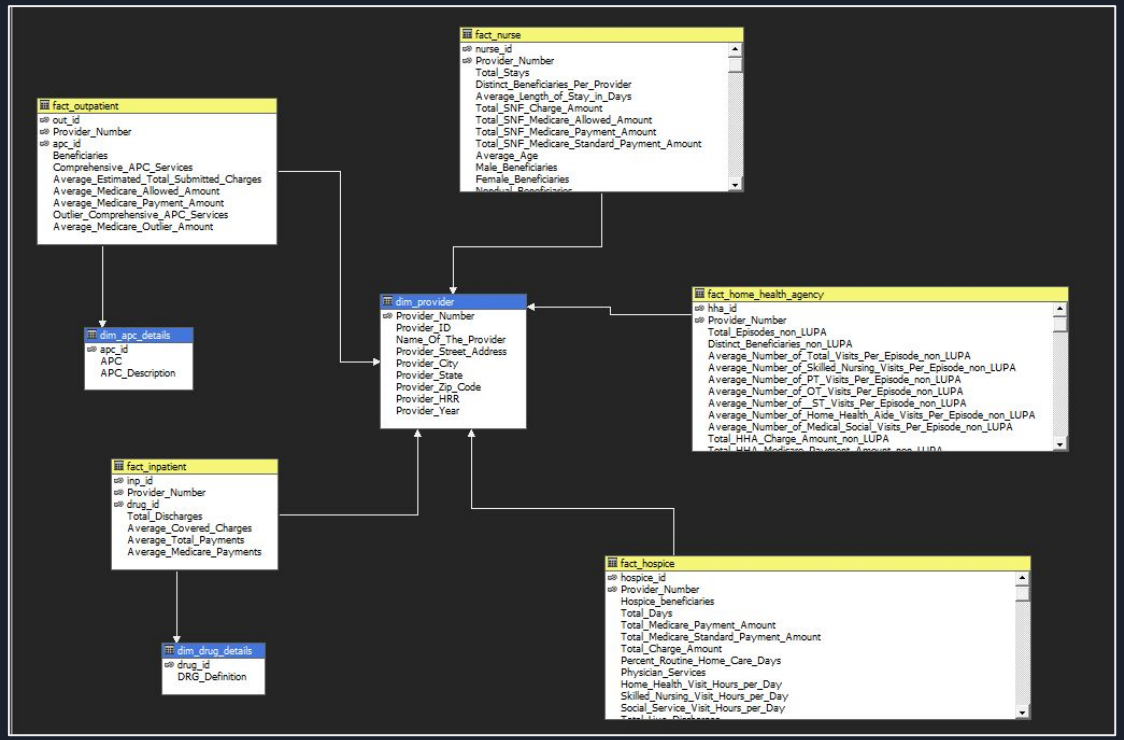






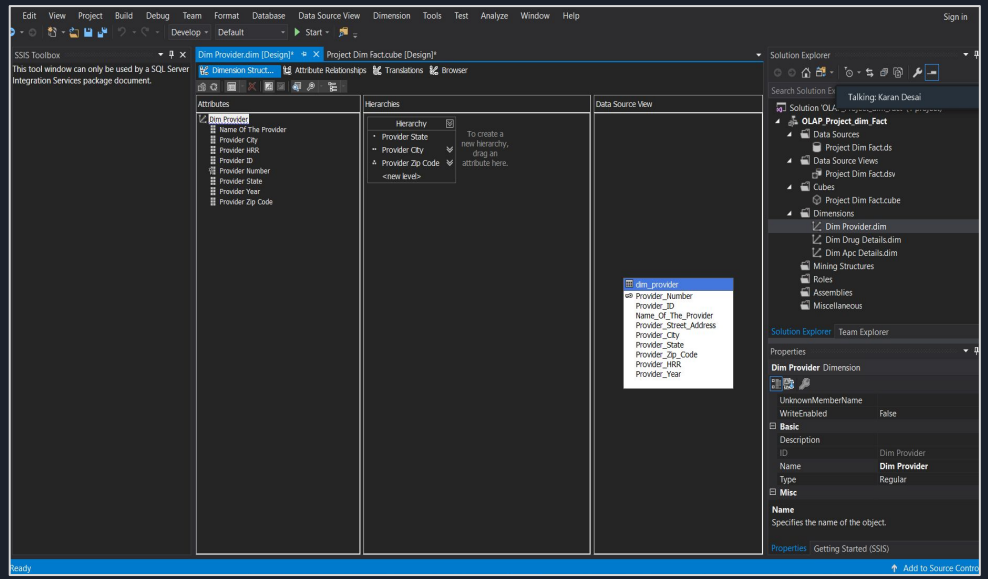


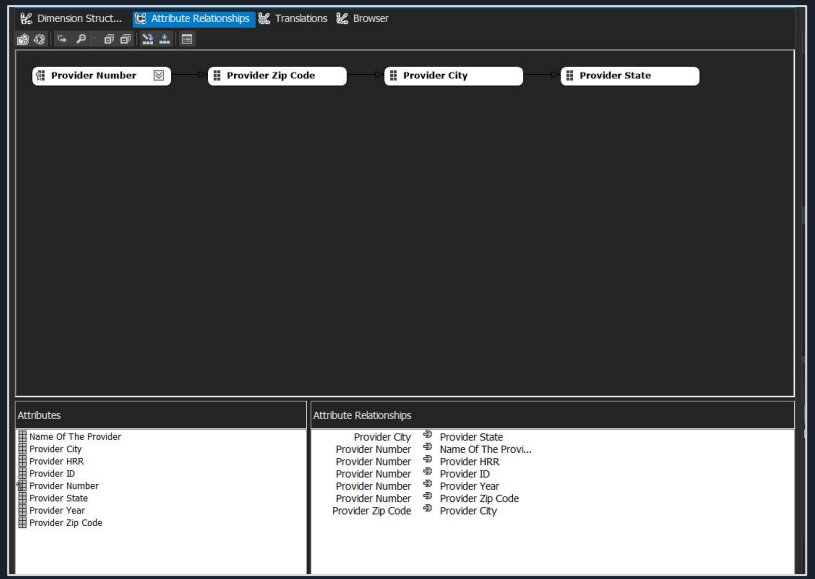
OLAP
Multidimensional Cube



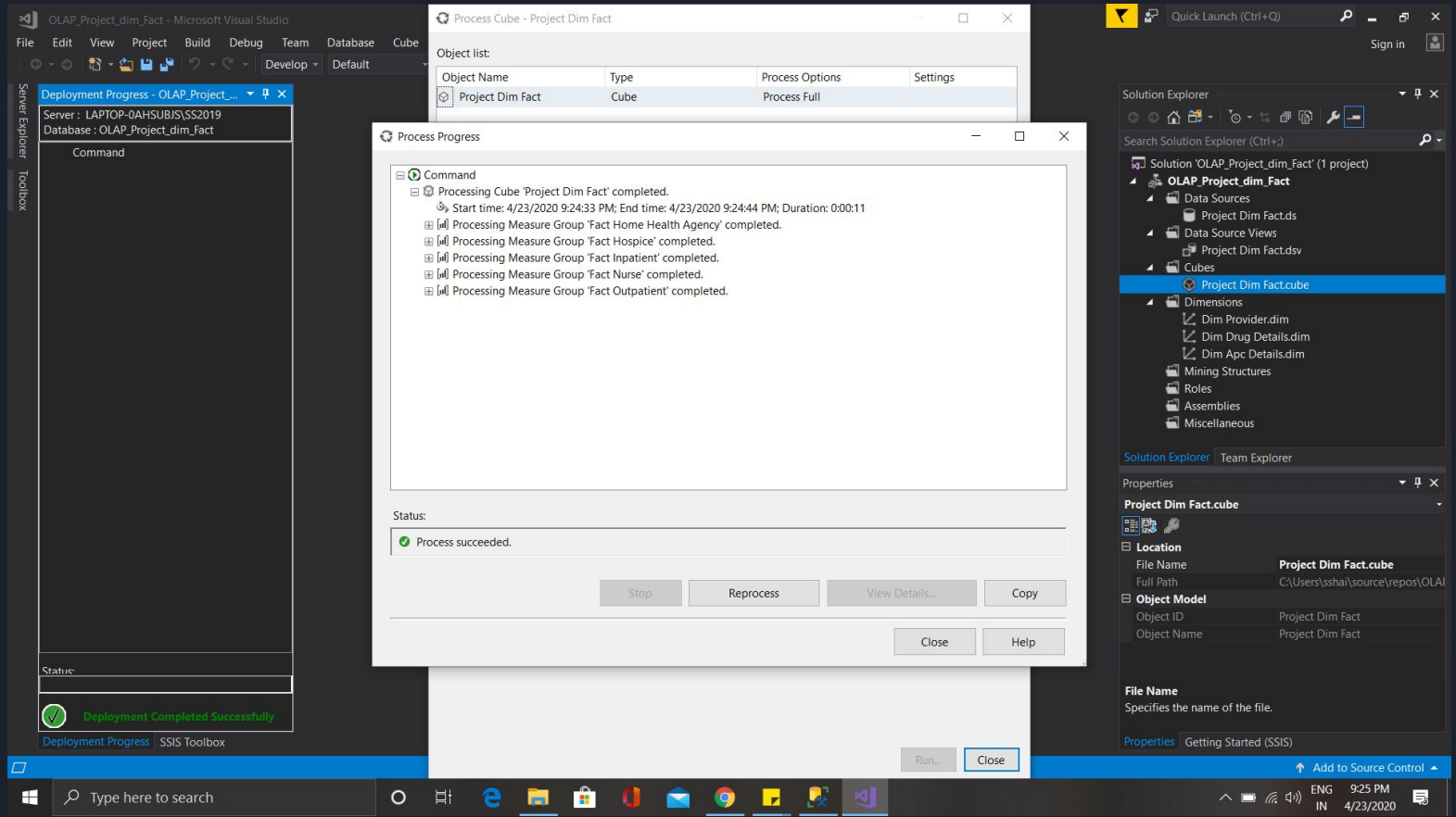
- Software:
  Visual Studio
- Mode:
  Multidimensional
- Fact Tables: 5
- Dimensions: 3
- Purpose:
  Cube Process
  Hierarchy
  Data Browse

Hierarchy

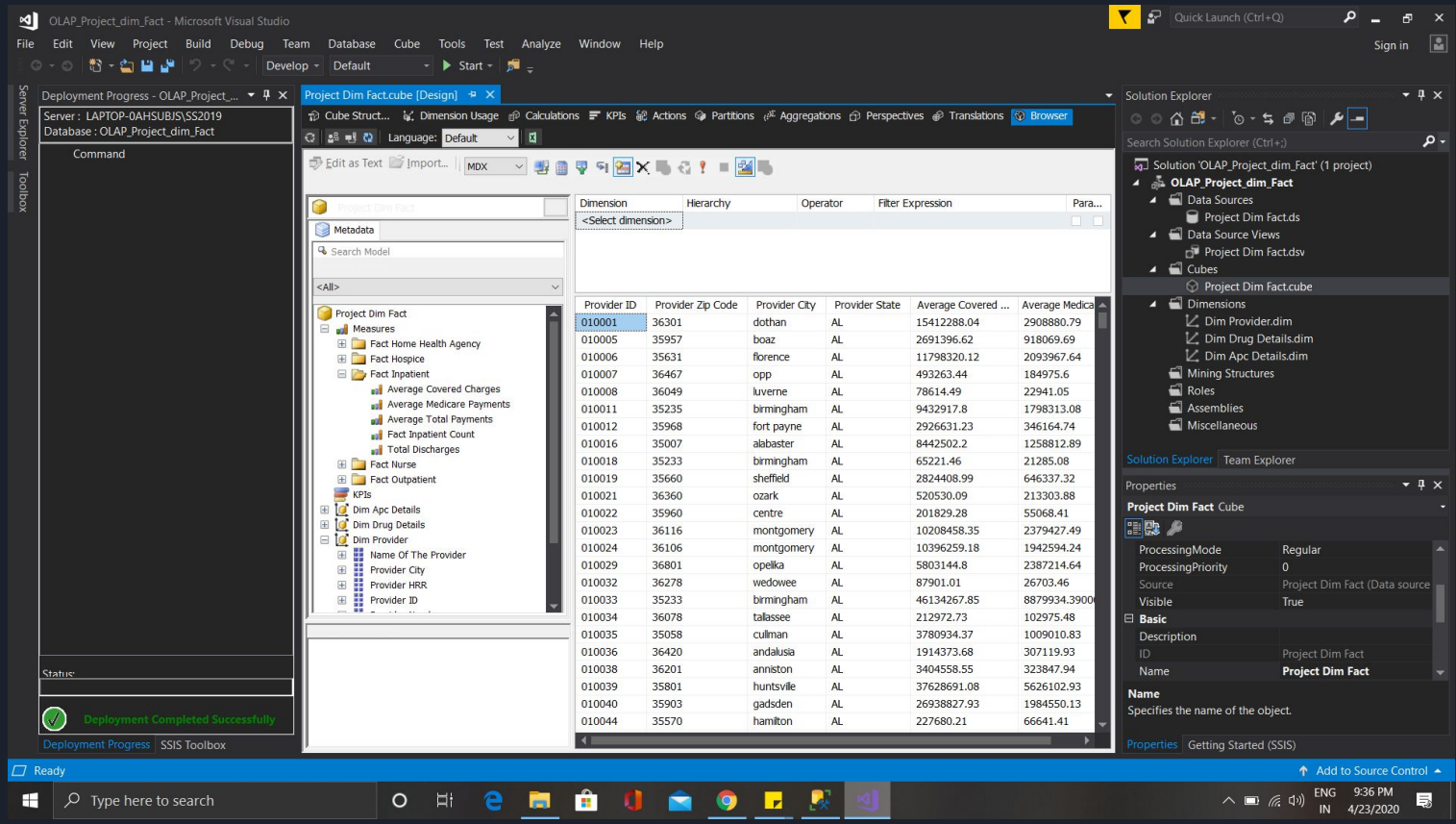

Edit View Project Build Debug Team Format Database Data Source View Dimension Tools Test Analyze Window Help
Sign in
SSIS Toolbox
This tool window can only be used by a SQL Server Integration Services package document.
Dim Provider.dim [Design]*
Project Dim Fact.cube [Design]*
Dimension Struct...
Attribute Relationships
Translations
Browser
Attributes
Dim Provider
Name Of The Provider
Provider City
Provider HRR
Provider ID
Provider Number
Provider State
Provider Year
Provider Zip Code
Hierarchies
Hierarchy
Provider State
Provider City
Provider Zip Code
<new level>
To create a new hierarchy, drag an attribute here.
Data Source View
dim_provider
Provider_Number
Provider_ID
Name_Of_The_Provider
Provider_Street_Address
Provider_City
Provider_State
Provider_Zip_Code
Provider_HRR
Provider_Year
Solution Explorer
Talking: Karan Desai
OLAP_Project_dim_Fact
Data Sources
Project Dim Fact.ds
Data Source Views
Project Dim Fact.dsv
Cubes
Project Dim Fact.cube
Dimensions
Dim Provider.dim
Dim Drug Details.dim
Dim Apc Details.dim
Mining Structures
Roles
Assemblies
Miscellaneous
Solution Explorer
Team Explorer
Properties
Dim Provider Dimension
UnknownMemberName
WriteEnabled
False
Basic
Description
ID
Dim Provider
Name
Dim Provider
Type
Regular
Misc
Name
Specifies the name of the object.
Properties
Getting Started (SSIS)
Ready
Add to Source Contro


Dimension Struct...
Attribute Relationships
Translations
Browser
Provider Number
Provider Zip Code
Provider City
Provider State
Attributes
Name Of The Provider
Provider City
Provider HRR
Provider ID
Provider Number
Provider State
Provider Year
Provider Zip Code
Attribute Relationships
Provider City
Provider State
Provider Number
Name Of The Provi...
Provider Number
Provider HRR
Provider Number
Provider ID
Provider Number
Provider Year
Provider Number
Provider Zip Code
Provider Zip Code
Provider City

Cube Processing



Cube Browser



Data Visualisation

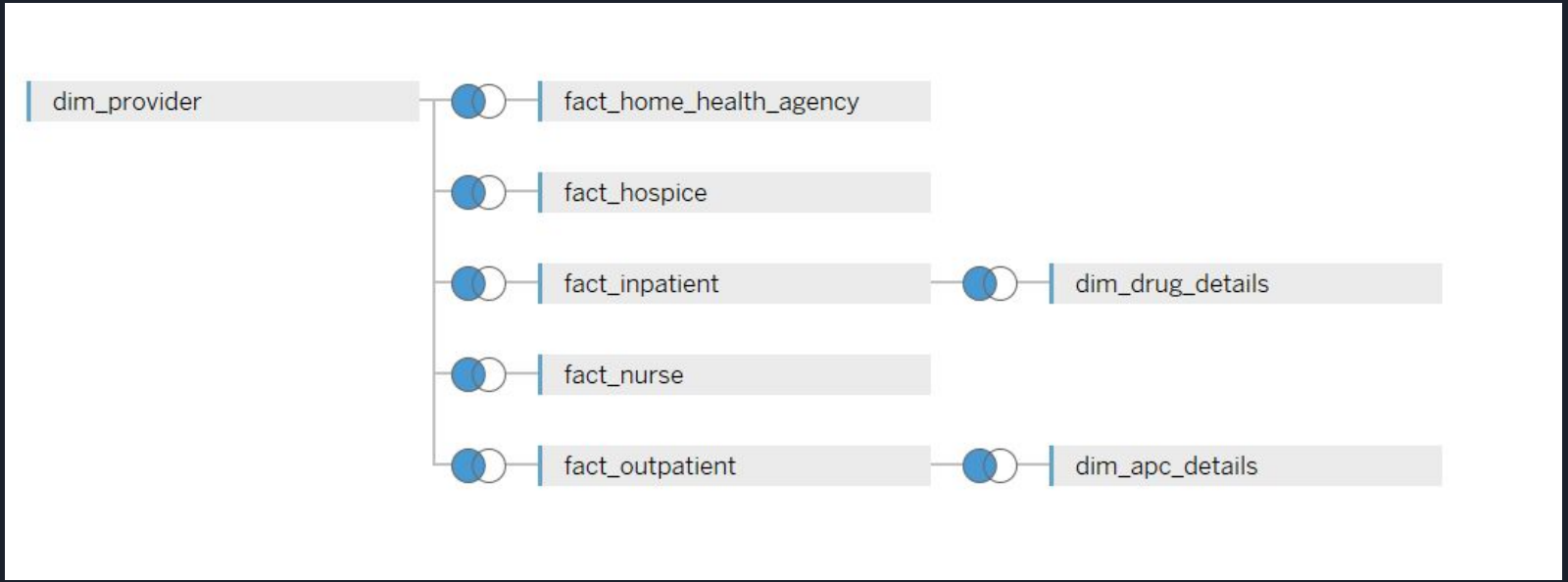


Tableau Dashboard: Health Services Comparison

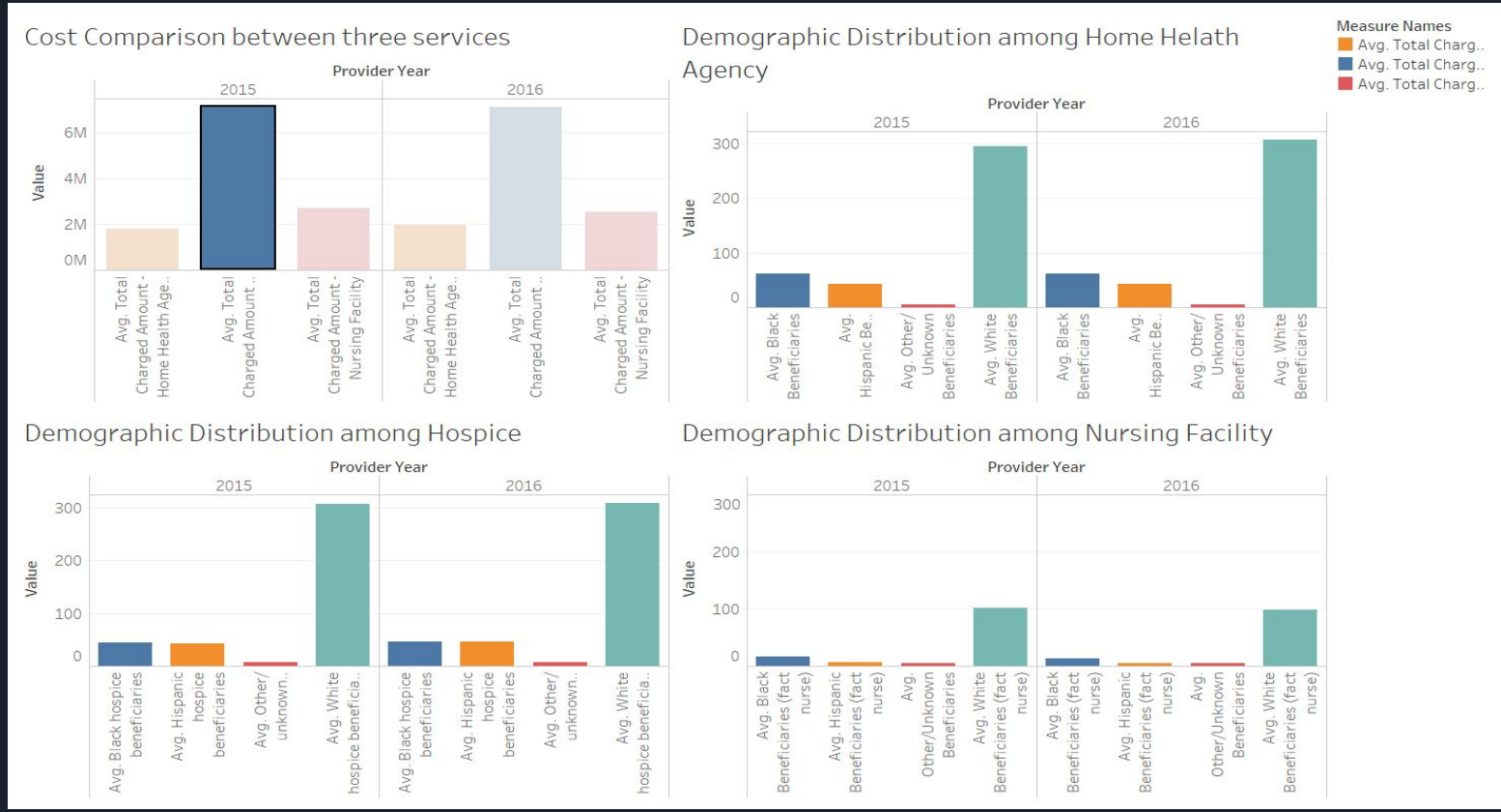


Tableau Dashboard: Disease Comparison in Nursing and Home Health Agency

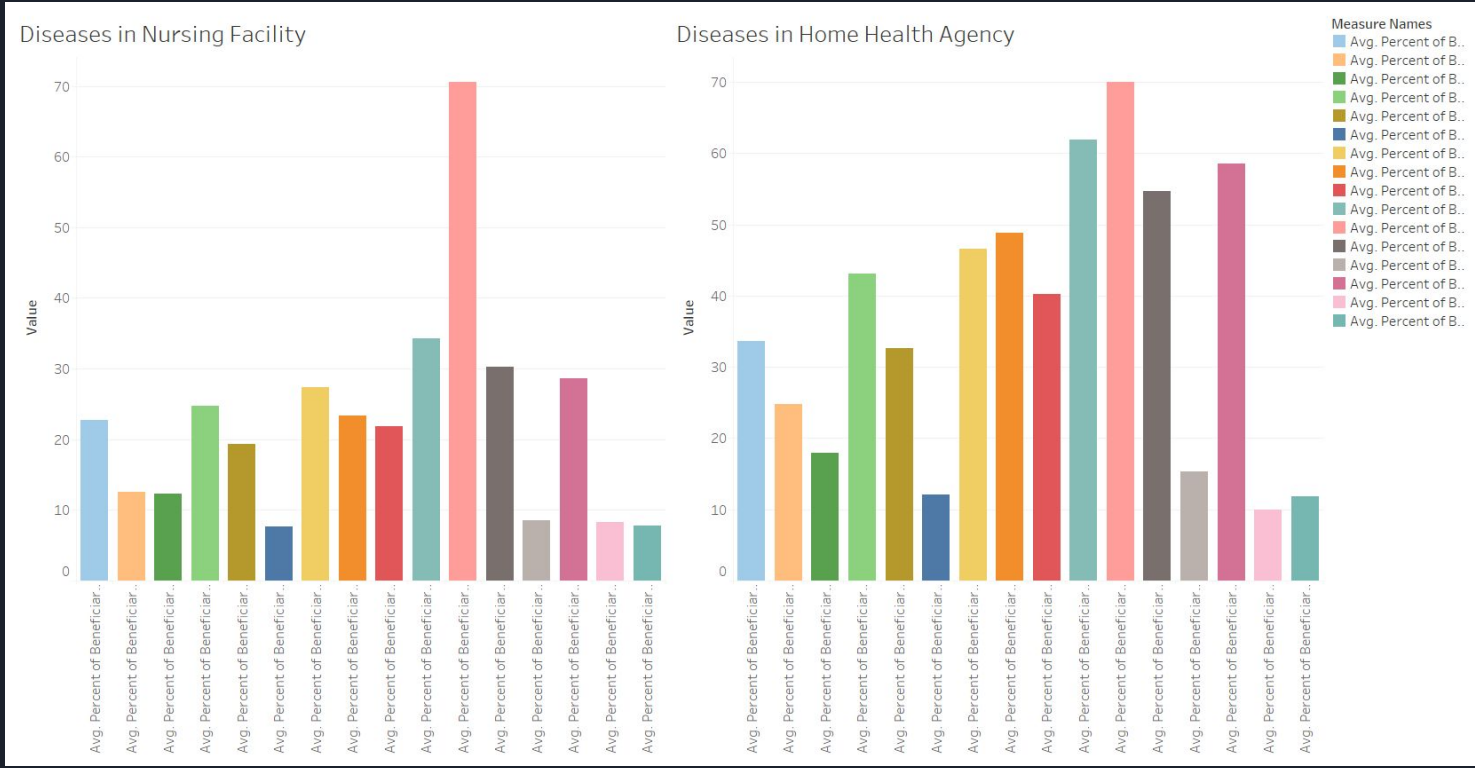


Tableau Dashboard: Beneficiary Distribution and their HCC Score

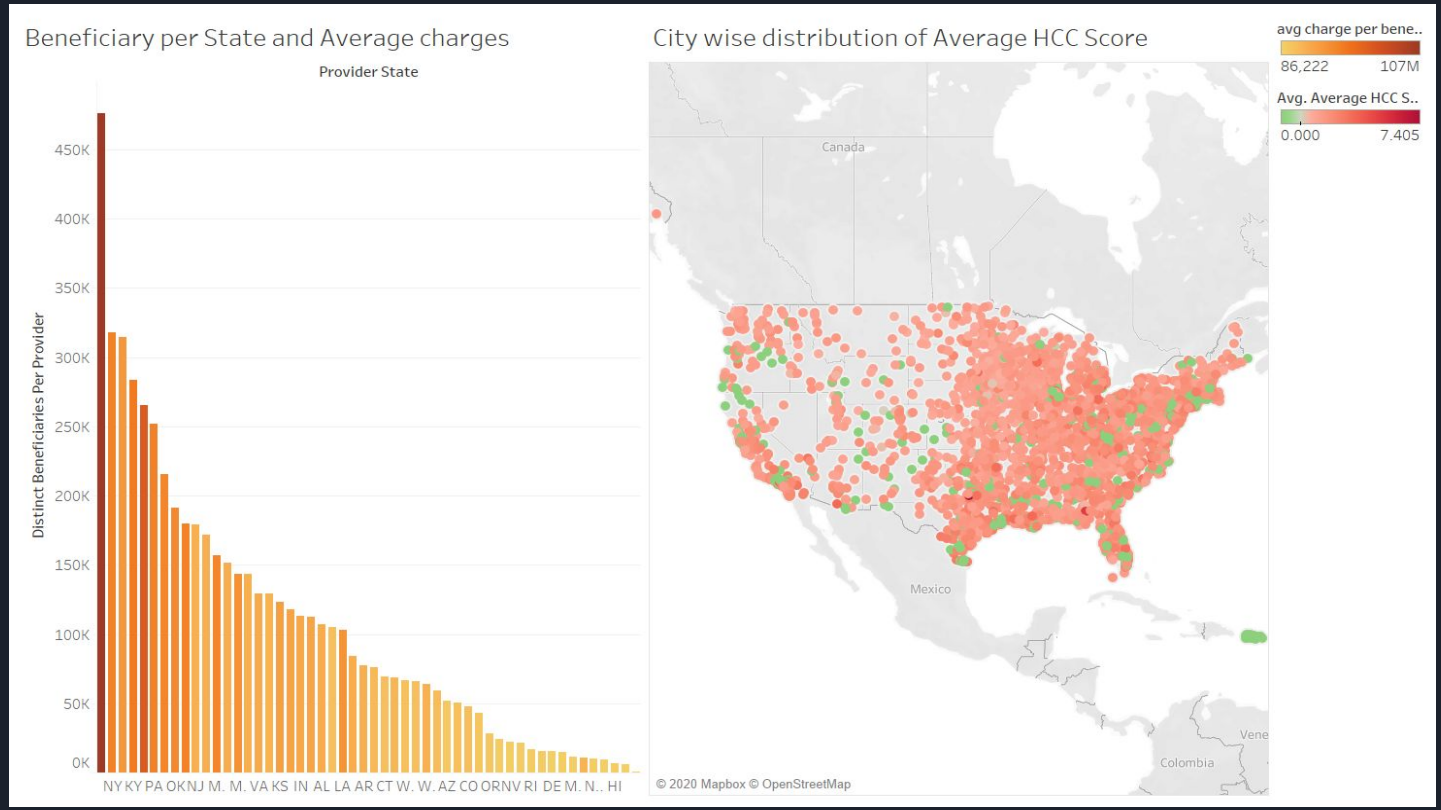
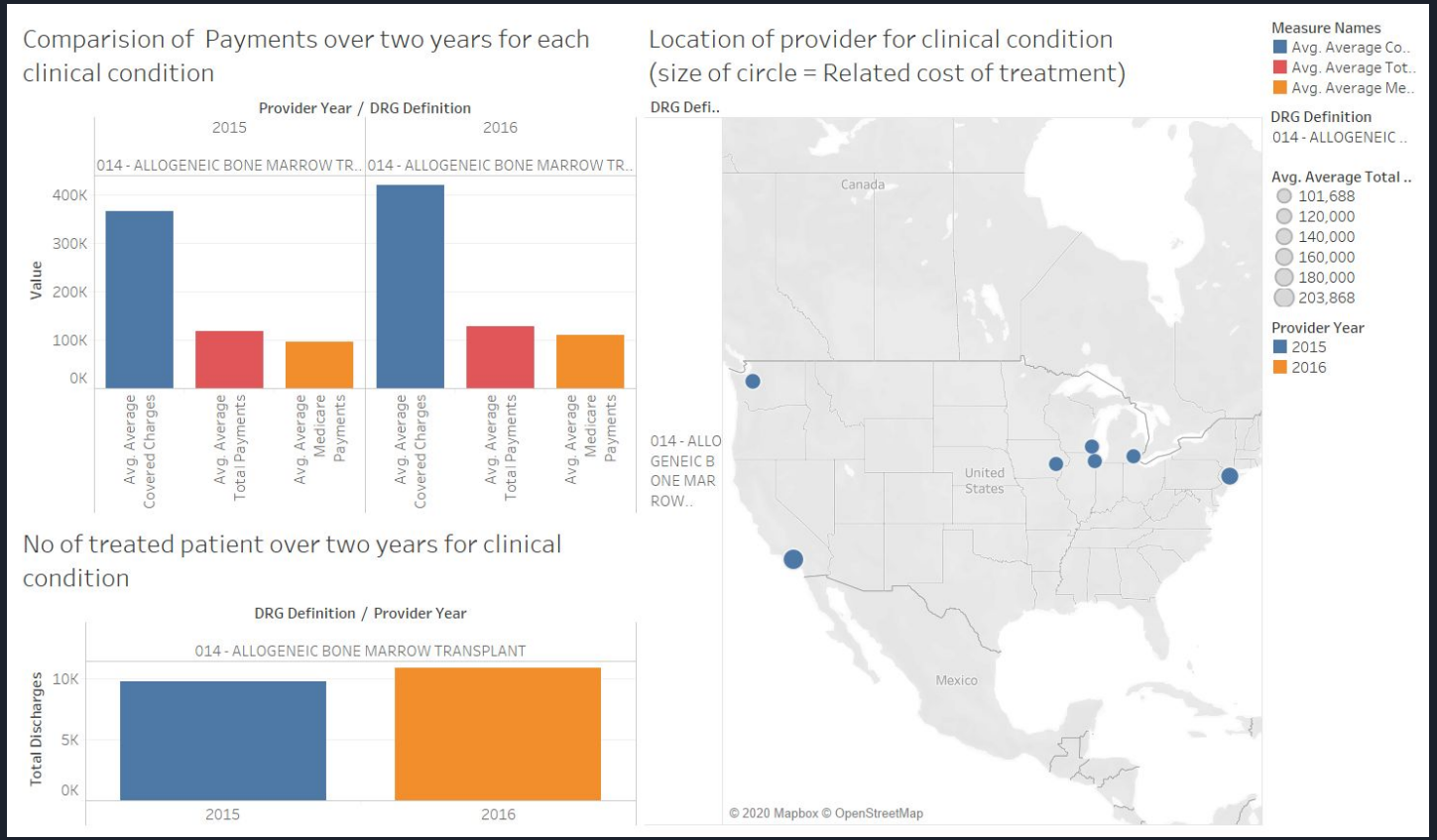


Tableau Dashboard: Drug Definition wise Payment, Discharge and Provider Distribution



Thank you!