



Dedup



Members:

Pratik Chavan

Pratik Kadam

Priyanka Donta

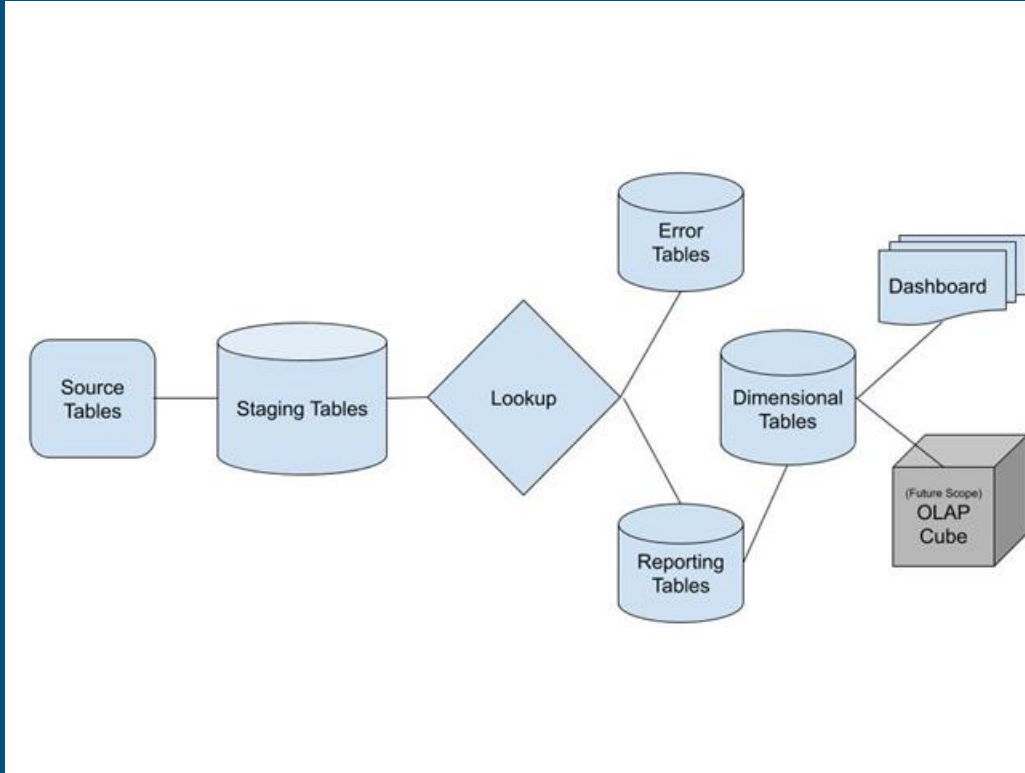
Terance Barretto



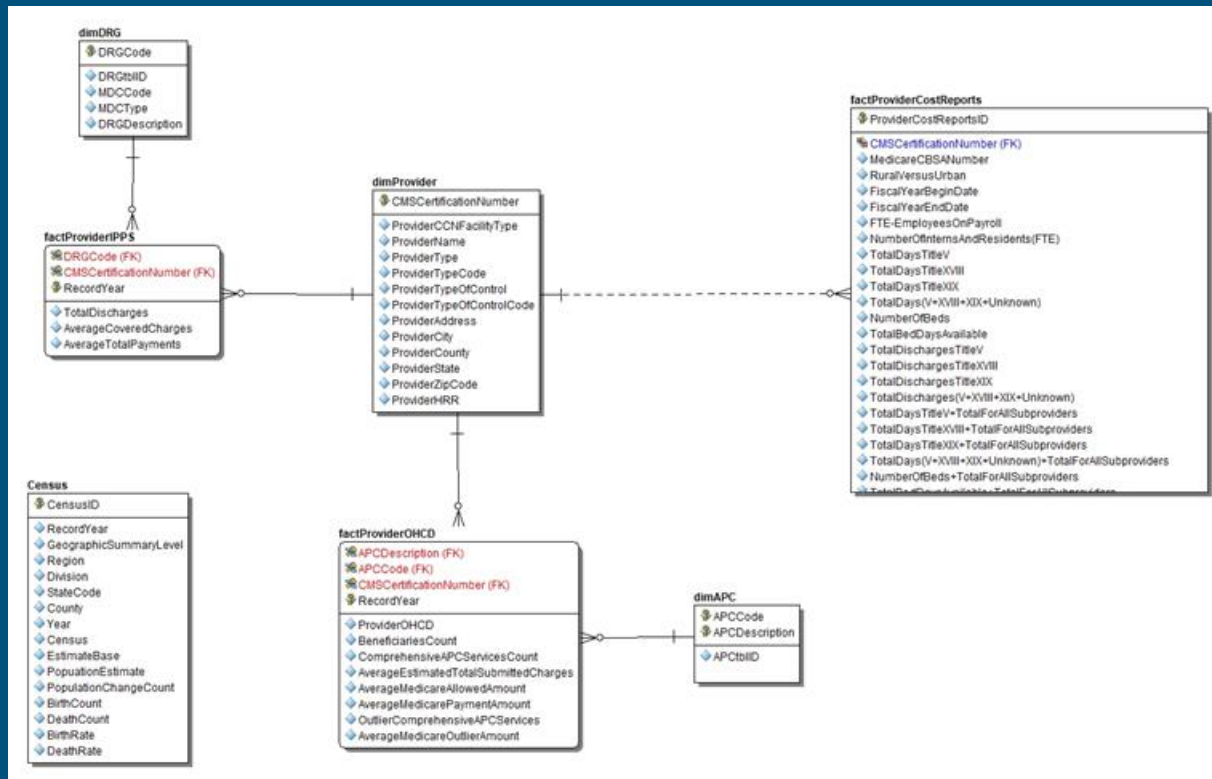
Objective

- Understanding the structure of the cost for different healthcare services for all the states in United States of America
- Working remotely and collaborating and create a data warehouse using Cloud Infrastructure
- Check the number of providers per state
- Comparing Inpatient and Outpatient revenue for Providers

Data Model



ER Diagram



Tables created

- + archive.Census_20200422
- + archive.ProviderCostReports_20200422
- + archive.ProviderIPPS_20200422
- + archive.ProviderOHCD_20200422
- + cube.Census
- + cube.dimAPC
- + cube.dimDRG
- + cube.dimProvider
- + cube.factProviderCostReports
- + cube.factProviderIPPS
- + cube.factProviderOHCD
- + error.Census
- + error.ProviderIPPS
- + error.ProviderOHCD

- + lookup.DRGCodes
- + lookup.ProviderCCN
- + lookup.ZipCodesUS
- + reporting.Census
- + reporting.ProviderCostReports
- + reporting.ProviderIPPS
- + reporting.ProviderOHCD
- + source.Census
- + source.ProviderCostReports
- + source.ProviderIPPS
- + source.ProviderOHCD
- + staging.Census
- + staging.ProviderCostReports
- + staging.ProviderIPPS
- + staging.ProviderOHCD

Schemas:

Archive:

TableName + Date of Archive

Cube:

Dimensional Tables

Error:

Store errors generated during Lookups

Staging:

To move data from source and store with proper data types

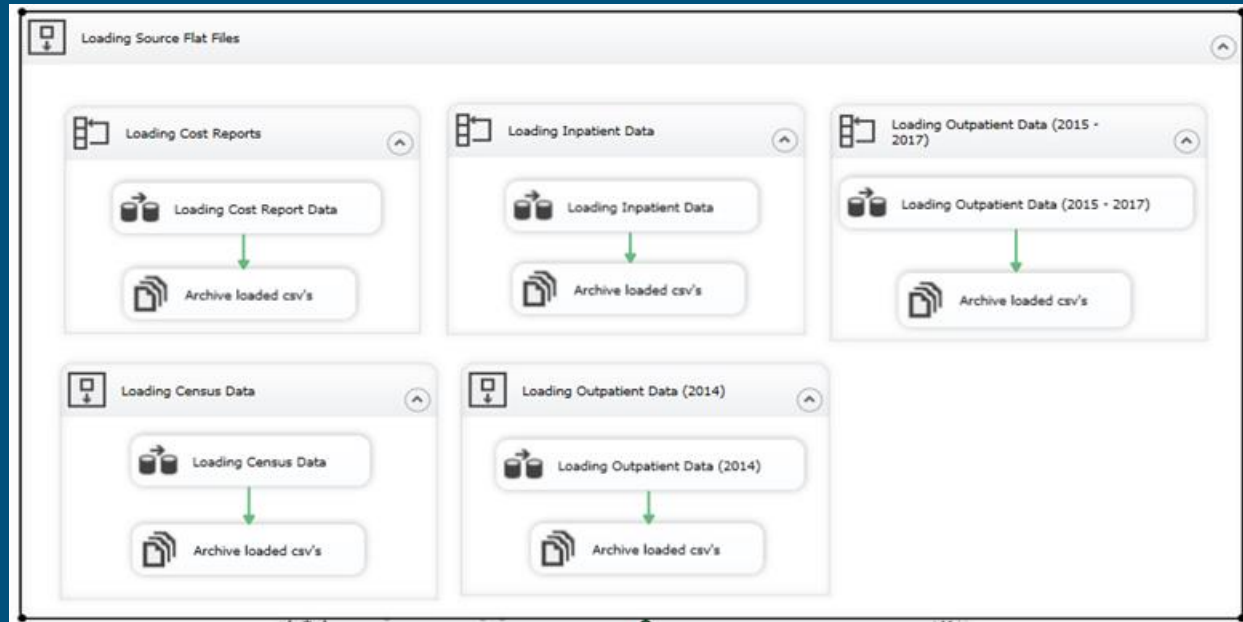
Source:

To load source flat files

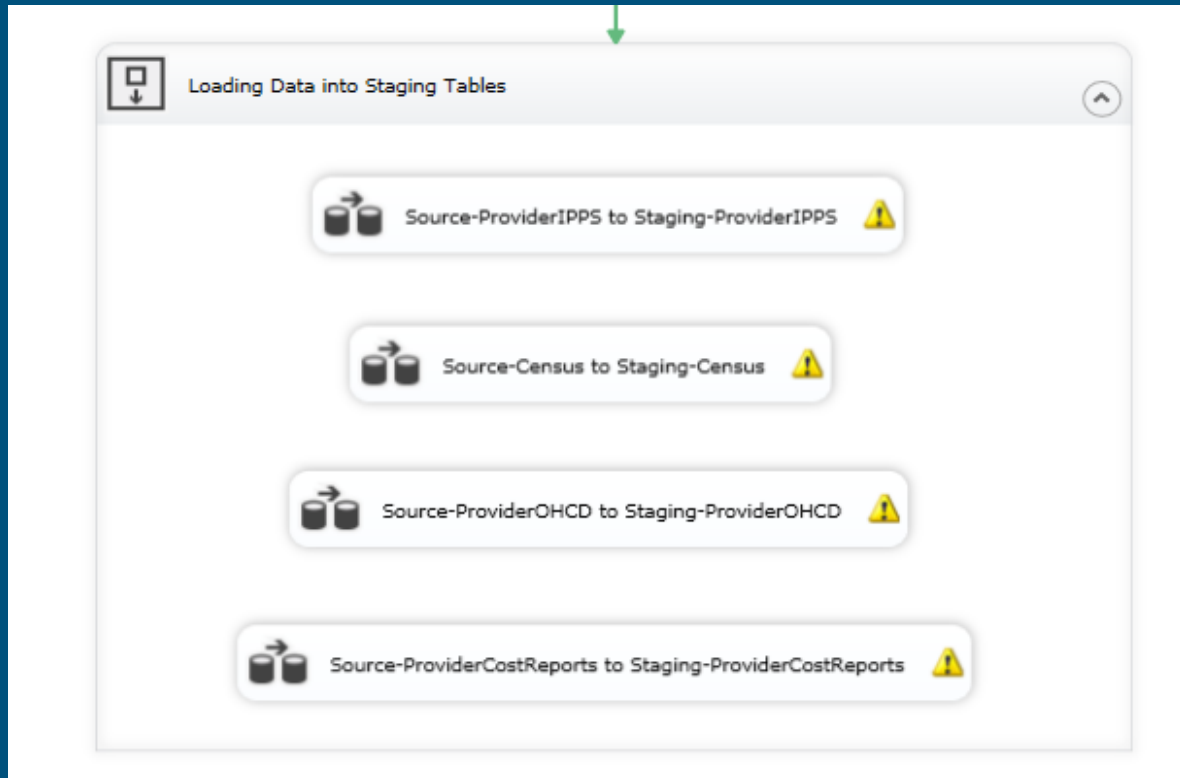
Reporting:

To have reporting tables for SQL Queries

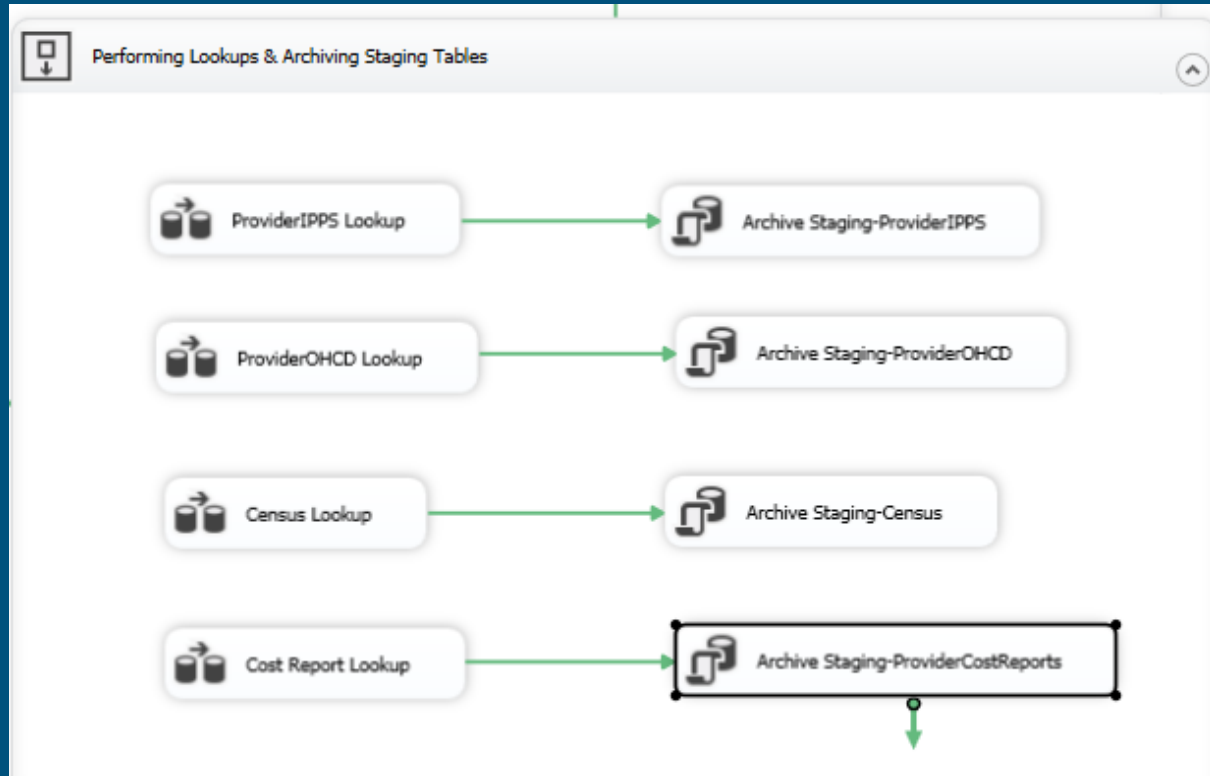
SSIS Control Flow



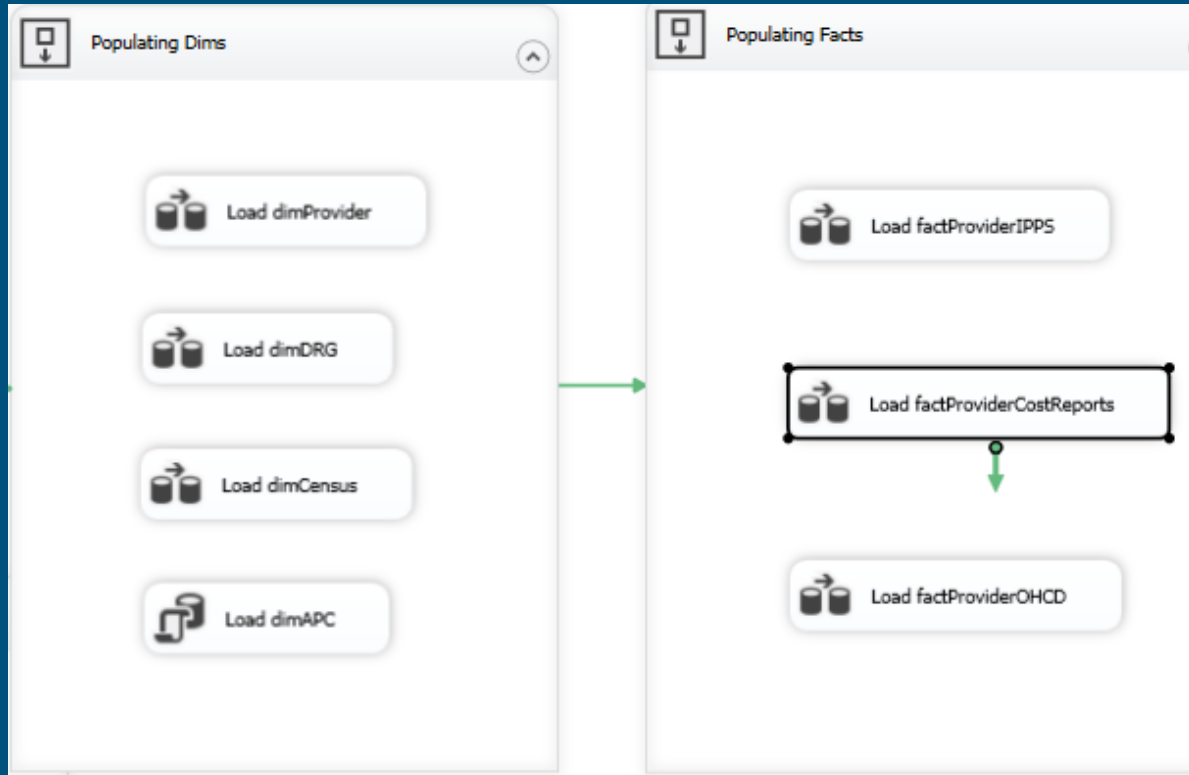
SSIS Control Flow



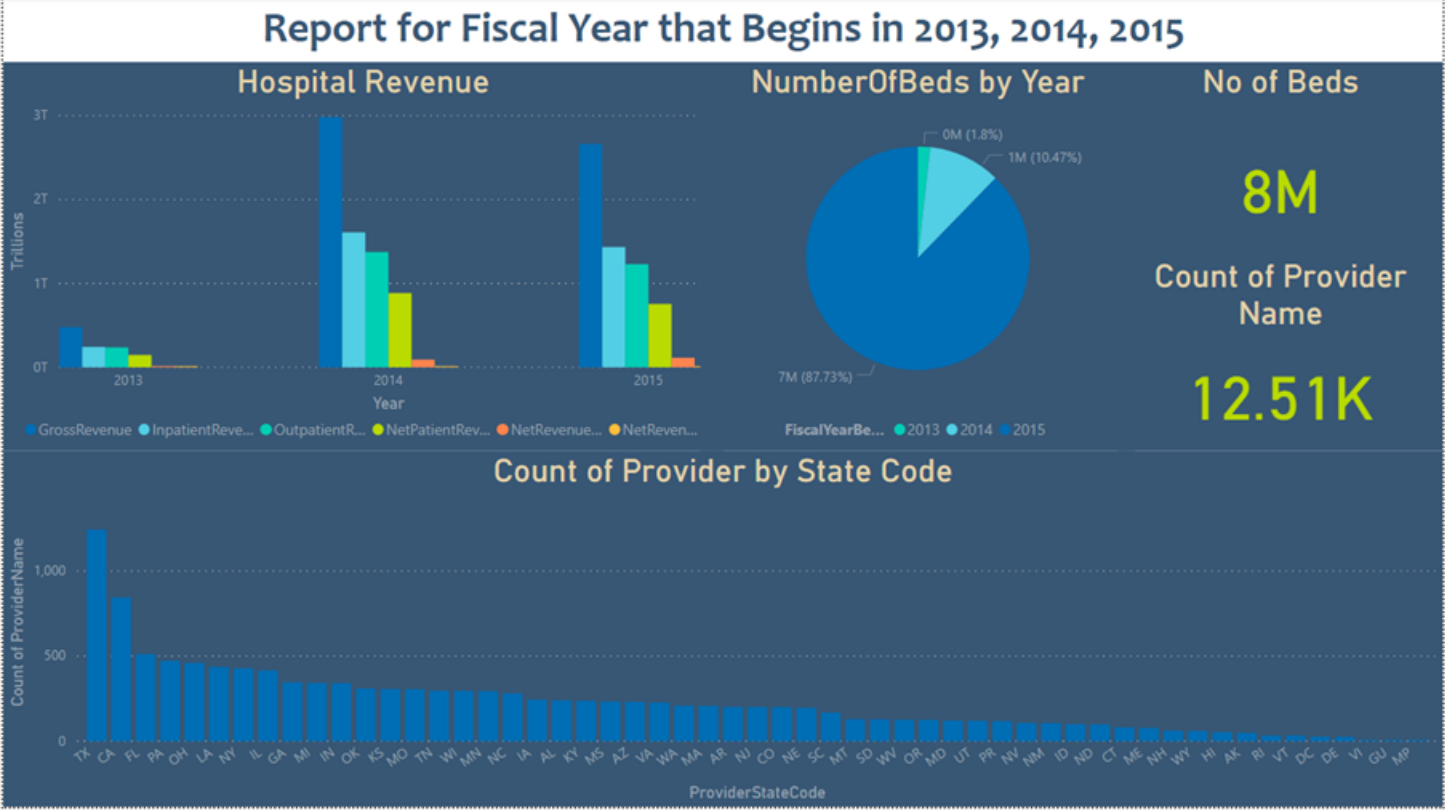
SSIS Control Flow



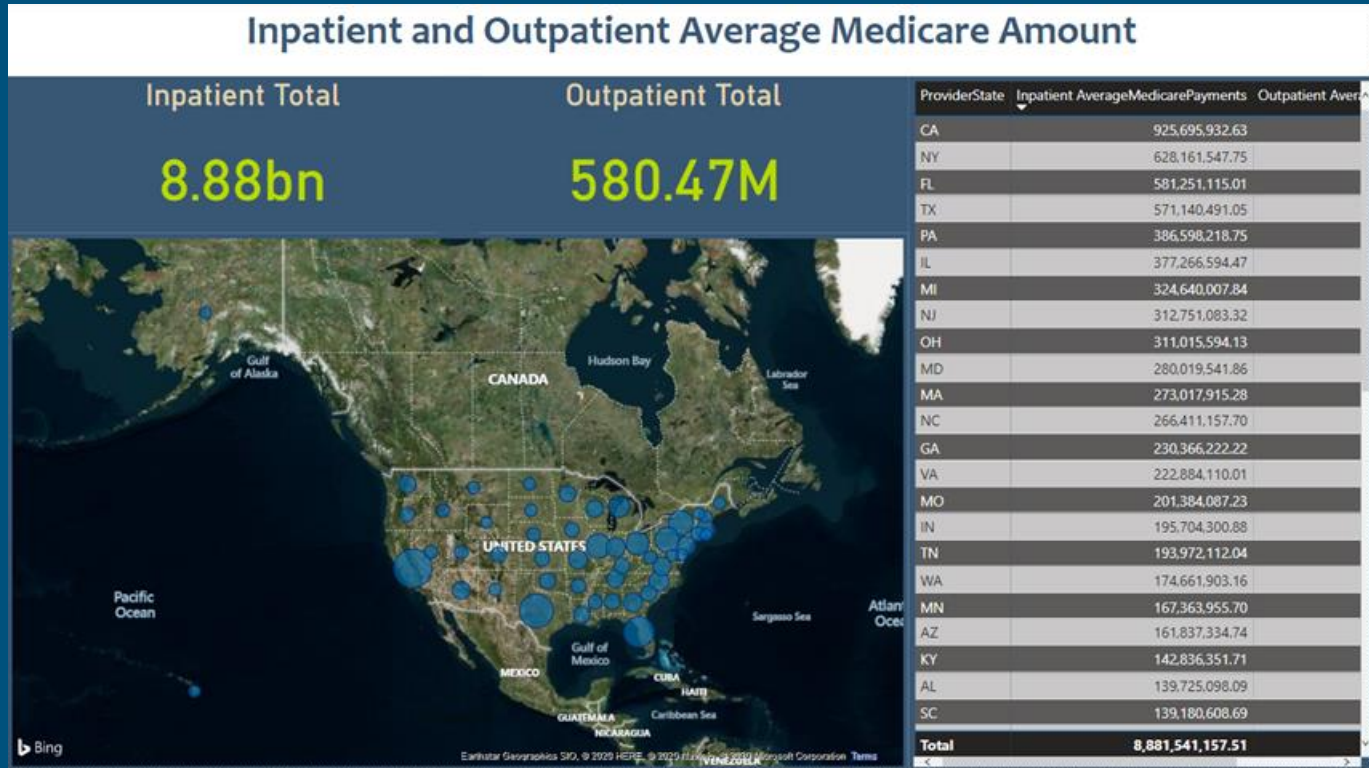
SSIS Control Flow



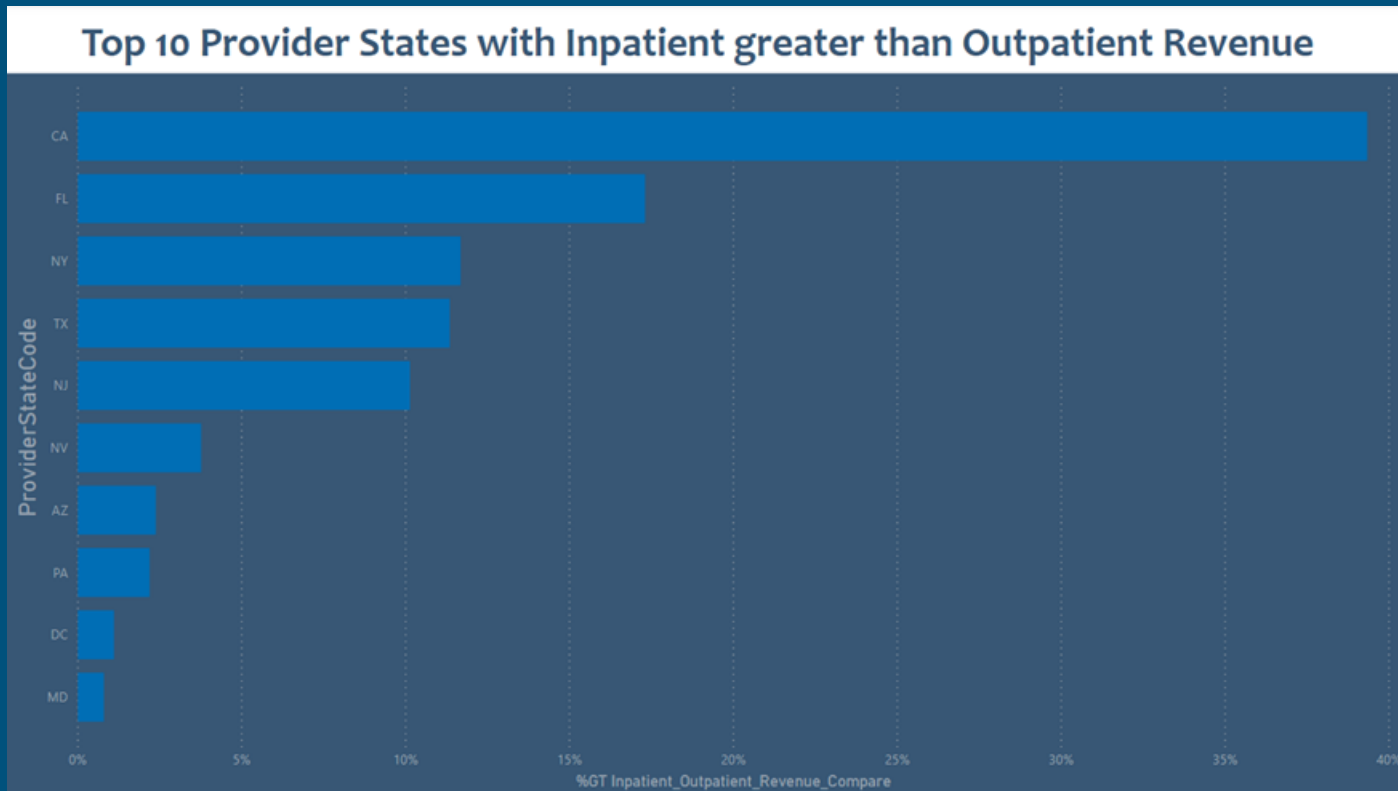
Dashboard



Dashboard



Dashboard



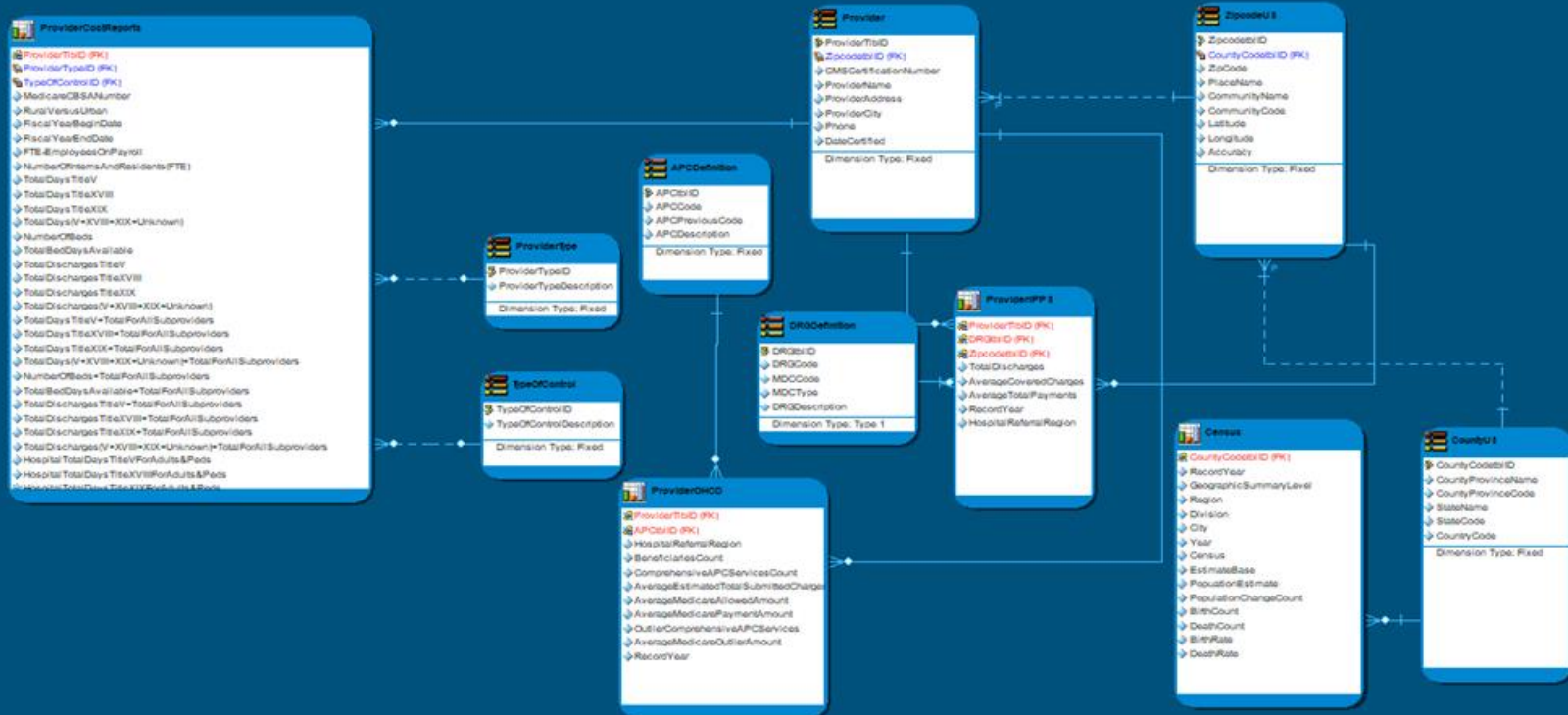
Pain Points

- Less compute resources available on Azure due to Cost Constraint
- Unavailability of source of truth lookups for APCCodes etc.
- Data structure change for Outpatient Prospective Payment System after 2014
- Aligning on the business rules and dimensional model

Future Scope

- Improve dimensional model
- Move flat file loading to Azure to enable all-cloud infrastructure
- Improve lookup tables with better data sources
- Implement a cube based on the dimensional model

Future Scope



Conclusion

- Built a data warehouse for healthcare Provider data using SSIS
- Created dashboard on Power BI to visualize statewise Provider data and Inpatient, Outpatient Revenue