Analyzing Hospital Acquired Infections and Other Healthcare Measures

Michael Marks

W205 Final Project

Spring 2016

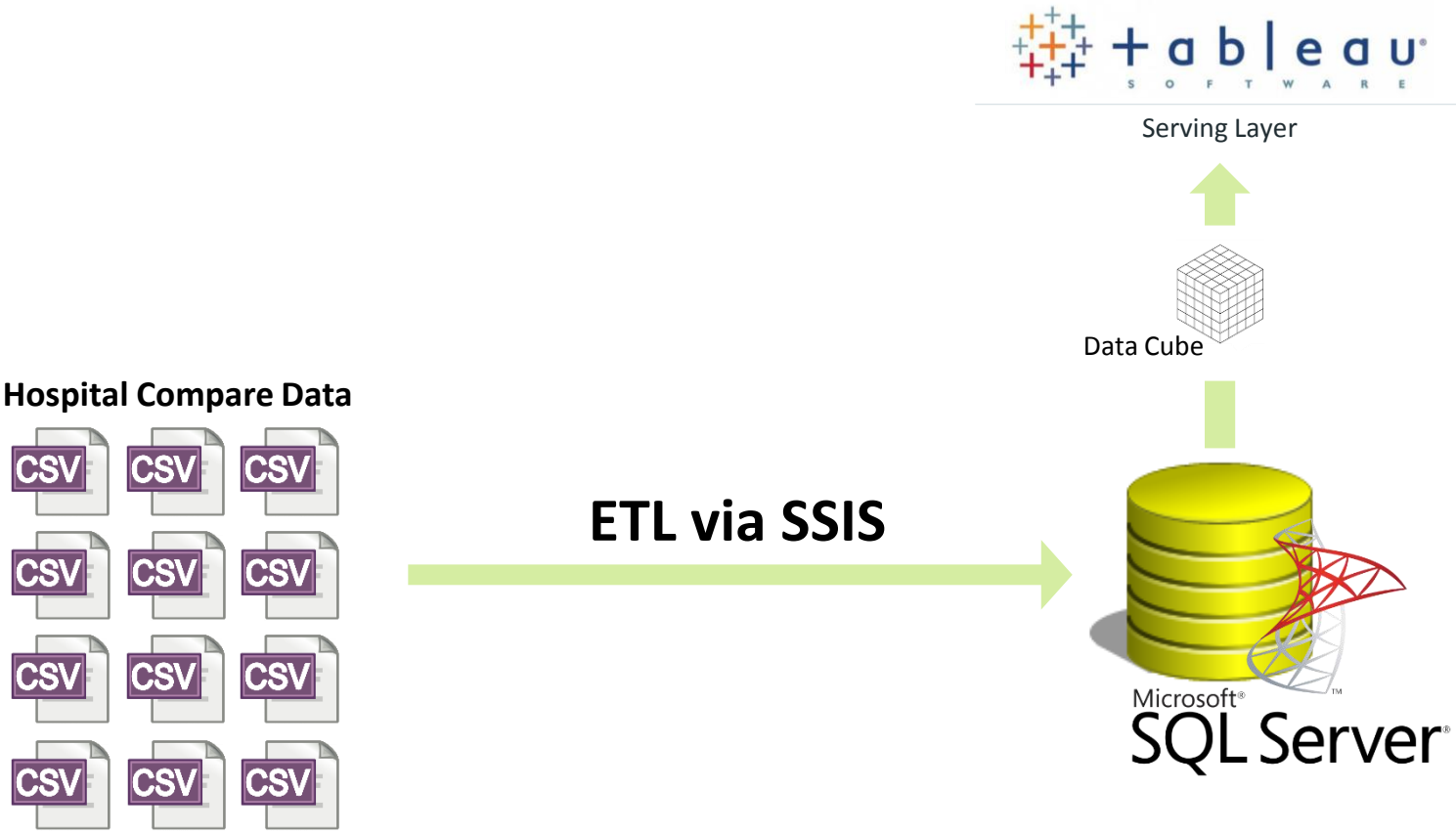
The Problems

1. Many hospitals have little insight into how their quality measures stack up against others.
2. Analysts within a company could utilize a source of hospital outcome, quality and safety measures.
   - The company's existing architecture is Microsoft based.
   - The company's main client runs highly secure Microsoft servers to which the company's software is deployed.
3. Patients have no insight into outcome, quality and safety measures when choosing a hospital.

Solution

- Create a functional data warehouse of the Center for Medicare and Medicare Services' (CMS) hospital compare data.
- Create a dashboard from the data warehouse that hospitals, analysts and patients can use to compare hospitals and analyze data.
- Utilize Microsoft products that can be deployed on the client's servers.

Overview


Serving Layer
Data Cube
Hospital Compare Data
CSV
ETL via SSIS
Microsoft
SQL Server

Pros and Cons of this Solution

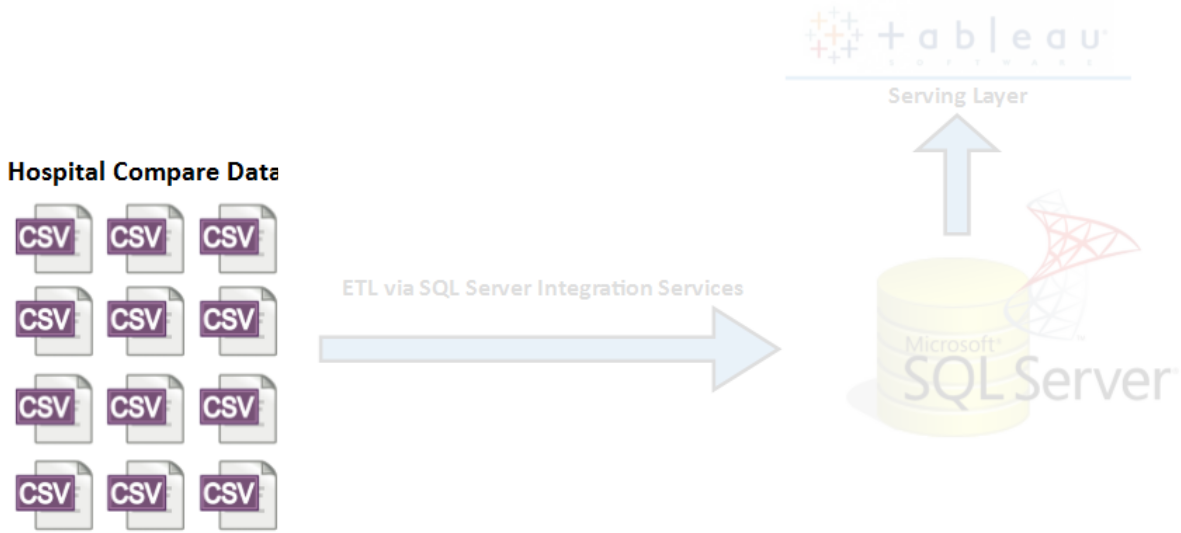
Pros

- MS SQL Server and Tableau are robust a proven tools
- Simple user interfaces
- Fits within company's existing infrastructure.
- Can be put on client's servers
- Many companies run on Microsoft

Cons

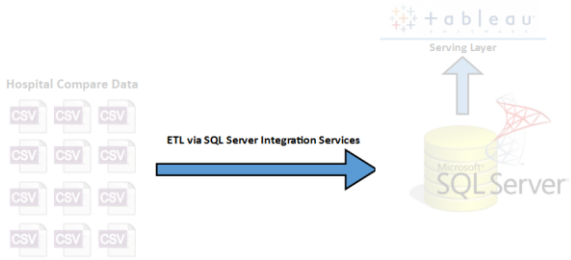
- Not open source
- Doesn't scale as well as other solutions.
- Only will work in a Windows environment
- Not easily reproducible
- Local

The Data



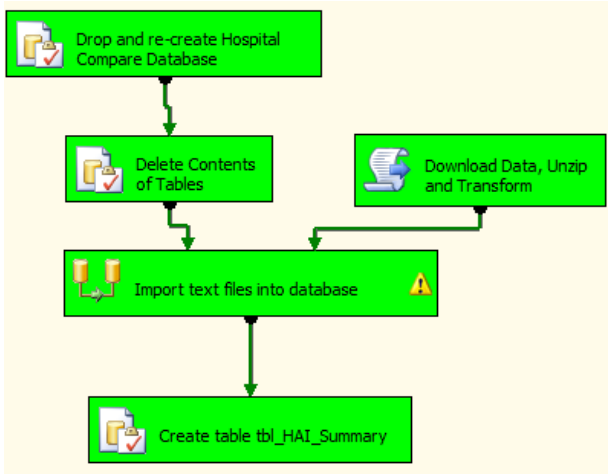
- Hospital Compare data from CMS
- Zip file contains 55 .csv files
- 284 megabytes in total

Extract, Transform, Load



- Used SQL Server Integration Services (SSIS)
  - Microsoft's GUI to create an ETL package.
- Wrote .NET code to handle many of the transformations.
- Final result is an ETL package written in XML.

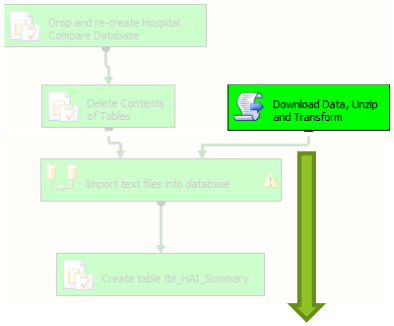
GUI



XML

```
<DTS:FlatFileColumn>
<DTS:Property DTS:Name="ColumnType">Delimited</DTS:Property>
<DTS:Property DTS:Name="ColumnDelimiter" xml:space="preserve">_x002C_</DTS:Property>
<DTS:Property DTS:Name="ColumnWidth">0</DTS:Property>
<DTS:Property DTS:Name="MaximumWidth">0</DTS:Property>
<DTS:Property DTS:Name="DataType">20</DTS:Property>
<DTS:Property DTS:Name="DataPrecision">0</DTS:Property>
<DTS:Property DTS:Name="DataScale">0</DTS:Property>
<DTS:Property DTS:Name="TextQualified">-1</DTS:Property>
<DTS:Property DTS:Name="ObjectName">Phone Number</DTS:Property>
<DTS:Property DTS:Name="DTSID">{E346820B-5952-46A5-8B4C-C934545658B6}</DTS:Property>
<DTS:Property DTS:Name="Description"></DTS:Property>
<DTS:Property DTS:Name="CreationName"></DTS:Property></DTS:FlatFileColumn>
```

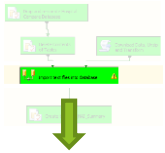
Extract, Transform, Load



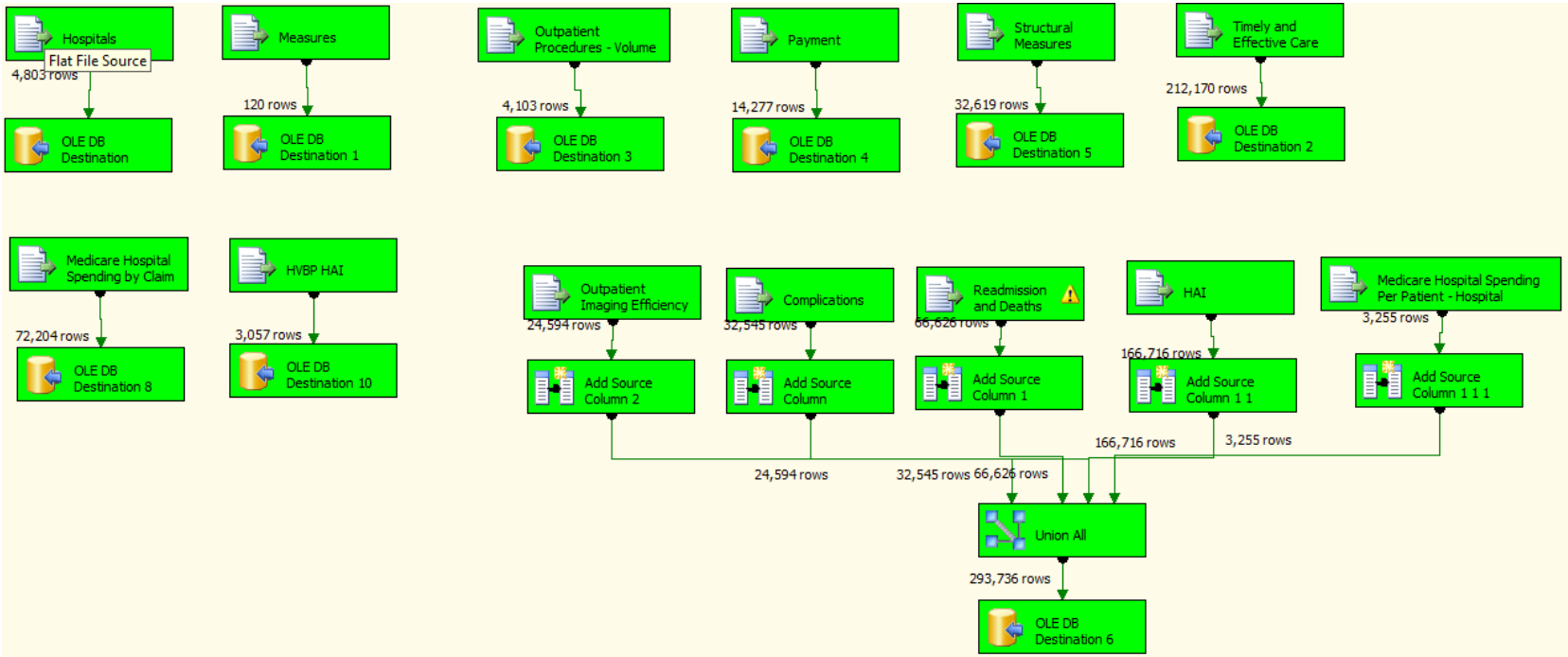
Download data, Unzip and Transform Process

- Uses .NET code to:
  - Download the zip file from CMS
  - Unzip it
  - Clean all text files
    - Remove leading and trailing spaces
    - Replace ‘N/A’ and ‘Not Available’ values with empty strings.

Extract, Transform, Load

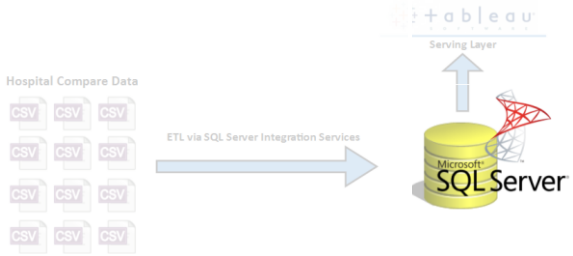


Import text files into database process



- 13 individual .csv files were loaded into 9 tables.
- Overall process (including preceding steps) takes about 5 minutes
  - SQL Server automatically uses all cores and multi-threading

The Database

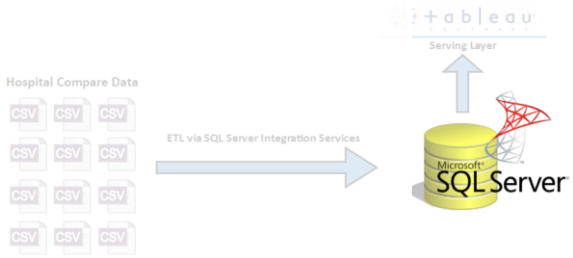


- Hospital_Compare
  - Database Diagrams
  - Tables
    - System Tables
    - FileTables
    - dbo.tbl_All_Measures_By_Hospital
      - Columns
        - Provider ID (varchar(50), null)
        - Measure ID (varchar(150), null)
        - Score (float, null)
        - Measure Start Date (datetime, null)
        - Measure End Date (datetime, null)
        - Source (nvarchar(29), null)
        - Denominator (varchar(50), null)
        - Lower Estimate (float, null)
        - Higher Estimate (float, null)
        - Measure Name (varchar(660), null)
      - Keys
      - Constraints
      - Triggers
      - Indexes
        - IX1 (Non-Unique, Non-Clustered)
        - IX2 (Non-Unique, Non-Clustered)
        - IX3 (Non-Unique, Non-Clustered)
      - Statistics
    - dbo.tbl_HAI_Summary
    - dbo.tbl_Hospitals
    - dbo.tbl_HVBP_HAI
    - dbo.tbl_Measures
    - dbo.tbl_Medicare_Spending_By_Claim
    - dbo.tbl_Outpatient_Procedure_Volume
    - dbo.tbl_Payment
    - dbo.tbl_Structural_Measures
    - dbo.tbl_Timely_and_Effective_Care

Results | Messages

	Provider ID	Measure ID	Score	Measure Start Date	Measure End Date	Source	Denominator	Lower Estimate	Higher Estimate	Measure Name
57	410004	MSPB_1	0.96	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
58	420069	MSPB_1	0.75	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
59	440001	MSPB_1	0.94	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
60	440081	MSPB_1	0.95	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
61	440218	MSPB_1	0.87	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
62	450090	MSPB_1	1.09	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
63	450221	MSPB_1	0.81	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
64	450460	MSPB_1	0.92	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
65	450661	MSPB_1	0.99	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
66	450808	MSPB_1	1.07	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
67	460006	MSPB_1	0.98	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
68	490018	MSPB_1	0.93	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
69	490111	MSPB_1	0.98	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
70	500044	MSPB_1	1.02	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
71	510055	MSPB_1	0.96	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
72	520088	MSPB_1	0.9	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
73	670008	MSPB_1	1.04	2014-01-01 00:00:00.000	2014-12-31 00:00:00.000	Medicare Spending	NULL	NULL	NULL	Medicare hospital spending per patient (Medicare...
74	010001	OP_10	6.5	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
75	010016	OP_14	4	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
76	010032	OP_10	2.1	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
77	010040	OP_14	2.3	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
78	010055	OP_10	5.3	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
79	010069	OP_14	4.8	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
80	010089	OP_10	9.9	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
81	010100	OP_14	1.9	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
82	010112	OP_10	12.9	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
83	010128	OP_14	0	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
84	010148	OP_10	50	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
85	010169	OP_14	0	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...
86	020017	OP_10	6.2	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Abdomen CT Use of Contrast Material
87	030007	OP_14	2.8	2013-07-01 00:00:00.000	2014-06-30 00:00:00.000	Outpatient Imagin...	NULL	NULL	NULL	Outpatients with brain CT scans who got a sinus ...

The Database



tbl_Hospitals
- PK **Provider_ID**
- Hospital Name
- Address
- City
- State
- ZIP Code
- County Name
- Phone Number
- Hospital Type
- Hospital Ownership
- Emergency Services

tbl_All_Measures_By_Hospital
- Provider_ID
- Measure_ID
- Score
- Measure Start Date
- Measure End Date
- Source
- Denominator
- Lower Estimate
- Higher Estimate
- Measure Name

tbl_Medicare_Spending_By_Claim
- Provider_ID
- Period
- Claim Type
- Average Spending Per Episode (Hospital)
- Average Spending Per Episode (State)
- Average Spending Per Episode (Nation)
- Percent of Spending (Hospital)
- Percent of Spending (State)
- Percent of Spending (Nation)
- Measure Start Date
- Measure End Date

tbl_Payment
- Provider_ID
- Denominator
- Payment
- Lower Estimate
- Higher Estimate
- Measure Start Date
- Measure End Date
- Payment Measure Name
- Payment Measure ID
- Payment Category
- Payment Footnote
- Value of Care display name
- Value of Care display ID
- Value of Care category
- Value of care footnore

tbl_HVBP_HAI
- Provider_ID
- SCIP-INF-1 Performance Rate
- SCIP-INF-1 Achievement Points
- SCIP-INF-1 Improvement Points
- SCIP-INF-1 Measure Score
- SCIP-INF-2 Performance Rate
- SCIP-INF-2 Achievement Points
- SCIP-INF-2 Improvement Points
- SCIP-INF-2 Measure Score
- SCIP-INF-3 Performance Rate
- SCIP-INF-3 Achievement Points
- SCIP-INF-3 Improvement Points
- SCIP-INF-3 Measure Score
- SCIP-INF-4 Performance Rate
- SCIP-INF-4 Achievement Points
- SCIP-INF-4 Improvement Points
- SCIP-INF-4 Measure Score
- SCIP-INF-9 Performance Rate
- SCIP-INF-9 Achievement Points
- SCIP-INF-9 Improvement Points
- SCIP-INF-9 Measure Score
- HAI Condition Procedure Score

tbl_Measures
- PK **Measure ID**
- Measure Name
- Measure Start Quarter
- Measure Start Date
- Measure End Quarter
- Measure End Date

tbl_Outpatient Procedure_Volume
- Provider_ID
- Measure_ID
- Gastrointestinal
- Eye
- Musculoskelatal
- Skin
- Genitourinary
- Cardiovascular
- Start Date
- End Date
- Nervous System

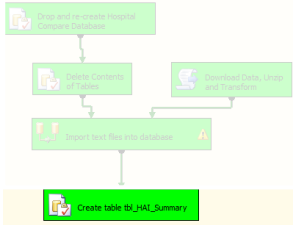
tbl_Structural_Measures
- Provider_ID
- Measure_ID
- Measure Response
- Measure Start Date
- Measure End Date

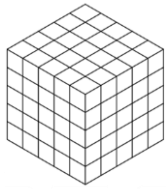
tbl_Timely_and_Effective_Care
- Provider ID
- Condition
- Measure ID
- Score
- Sample
- Measure Start Date
- Measure End Date

Data Cube

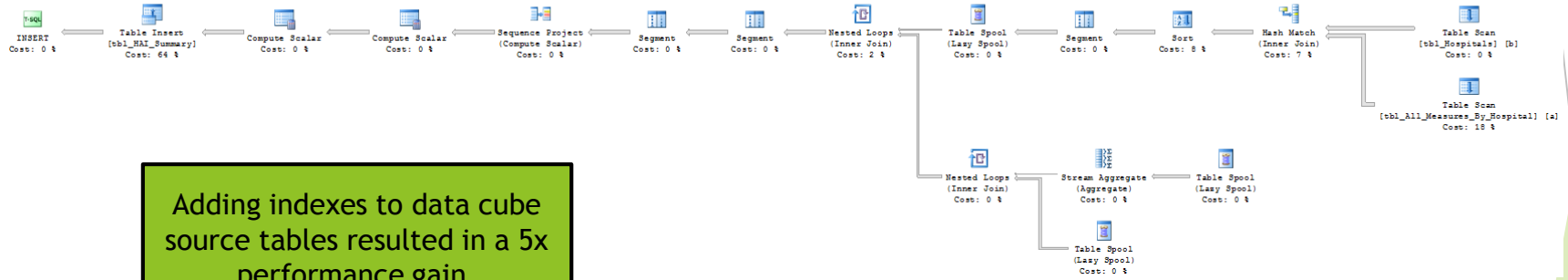
tbl_HAI Summary
- Hospital
- City
- State
- Zip_Code
- County
- Measure
- Score
- Percentile

Data Cube for Serving Layer





```
Query 4: Query cost (relative to the batch): 100%
Insert into tbl_HAI_Summary (Hospital ,City ,State ,ZIP_Code ,County ,Measure ,Score ,Percentile) Select [Hospital Name] ,[City] ,[State] ,[ZIP Code] ,[County Name] ,[Measure Name] , [Score] ,1- PERCENT_R…
Missing Index (Impact 25.4168): CREATE NONCLUSTERED INDEX [<Name of Missing Index, sysname,>] ON [dbo].[tbl_All_Measures_By_Hospital] ([Provider ID],[Score]) INCLUDE ([Measure ID],[Measure Name])
```



Adding indexes to data cube source tables resulted in a 5x performance gain.

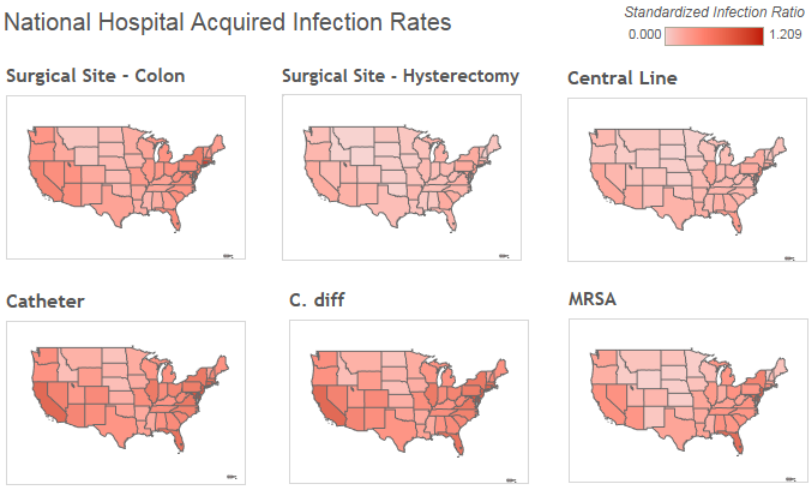
	Hospital	City	State	ZIP_Code	County	Measure	Score	Percentile
3113	TRINITY REGIONAL MEDICAL CENTER	FORT DODGE	IA	50501	WEBSTER	Central line-associated blood stream infections (CLABSI)	0	1
3114	IOWA LUTHERAN HOSPITAL	DES MOINES	IA	50316	POLK	Central line-associated blood stream infections (CLABSI)	0	1
3115	MARY GREELEY MEDICAL CENTER	AMES	IA	50010	STORY	Central line-associated blood stream infections (CLABSI)	0	1
3116	SKIFF MEDICAL CENTER	NEWTON	IA	50208	JASPER	Central line-associated blood stream infections (CLABSI)	0	1
3117	THE UNIVERSITY OF TN MEDICAL CEN...	KNOXVILLE	TN	37920	KNOX	Central line-associated blood stream infections (CLABSI)	0.03	0.32699...
3118	SENTARA NORFOLK GENERAL HOSPIT...	NORFOLK	VA	23507	NORFOLK CITY	Central line-associated blood stream infections (CLABSI)	0.034	0.32678...
3119	FROEDTERT MEMORIAL LUTHERAN H...	MILWAUKEE	WI	53226	MILWAUKEE	Central line-associated blood stream infections (CLABSI)	0.038	0.32656...
3120	MERCY SAN JUAN MEDICAL CENTER	CARMICHAEL	CA	95608	SACRAMENTO	Central line-associated blood stream infections (CLABSI)	0.041	0.32634...
3121	MAIMONIDES MEDICAL CENTER	BROOKLYN	NY	11219	KINGS	Central line-associated blood stream infections (CLABSI)	0.05	0.32613...
3122	SCOTTSDALE OSBORN MEDICAL CENT...	SCOTTSDALE	AZ	85251	MARICOPA	Central line-associated blood stream infections (CLABSI)	0.052	0.32591...
3123	WELLMONT HOLSTON VALLEY MEDICA...	KINGSPORT	TN	37662	SULLIVAN	Central line-associated blood stream infections (CLABSI)	0.058	0.32570...
3124	GROSSMONT HOSPITAL	LA MESA	CA	91942	SAN DIEGO	Central line-associated blood stream infections (CLABSI)	0.059	0.32548...
3125	METHODIST HEALTHCARE MEMPHIS H...	MEMPHIS	TN	38104	SHELBY	Central line-associated blood stream infections (CLABSI)	0.063	0.32526...
3126	BAPTIST HOSPITAL OF MIAMI	MIAMI	FL	33176	MIAMI-DADE	Central line-associated blood stream infections (CLABSI)	0.066	0.32505...

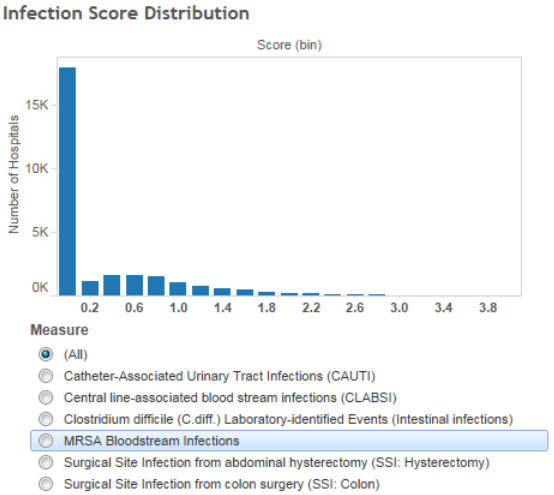
Serving Layer

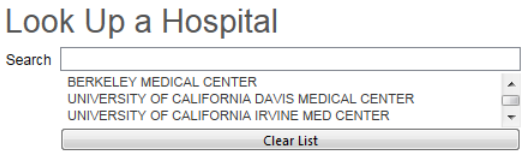
- Tableau
- Connected via ODBC
  - Uses Tableau public server for publishing and display purposes. Ideally I would have a Tableau server license. This would enable an end to end system.

Serving Layer

https://public.tableau.com/profile/publish/Dashboard_354/Dashboard1#!/publish-confirm







Percentile of Standardized Infection Ratio

Measure	BERKELEY MEDICAL CENTER	UNIVERSITY OF CALIFORNIA DAVIS MEDICAL CENTER	UNIVERSITY OF CALIFORNIA IRVINE MED CENTER	UNIVERSITY OF CALIFORNIA SAN DIEGO MEDICAL CENTER
Catheter-Associated Urinary Tract Infections (CAUTI)	0.4058	0.2918	0.1730	0.0996
Central line-associated blood stream infections (CLABSI)	1.0000	0.1732	0.3192	0.0724
Clostridium difficile (C.diff.) Laboratory-identified Events (Intestinal infections)	0.4536	0.2559	0.0819	0.1235
MRSA Bloodstream Infections	1.0000	0.0298	0.2691	0.2311
Surgical Site Infection from abdominal hysterectomy (SSI: Hysterectomy)	1.0000	0.0823	1.0000	0.0043
Surgical Site Infection from colon surgery (SSI: Colon)	0.3233	0.2050	0.1000	0.1060

Roadmap for Improvement

- Scale up using Microsoft Azure Products
  - Unfortunately there is no simple way to move the whole existing structure to Azure products. ETL would have to be re-written.
  - Azure Hybrid Cloud solution would enable it to be integrated into existing infrastructure.

Local Solution

Scale Up Cloud Solution

Azure Data Factory

Azure SQL Database

Overall my experience with Azure was good. It gets a lot done. Great out of the box. Looks good and easy to get started. I look forward to exploring its capabilities in the future.