

Three relational databases walk into a NoSQL bar. They left after 5 mins....

They couldn't find a table!



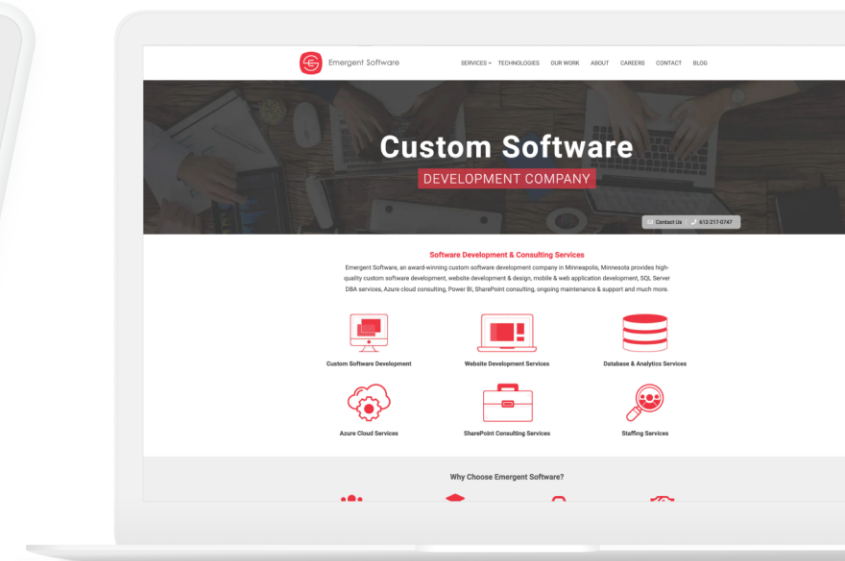
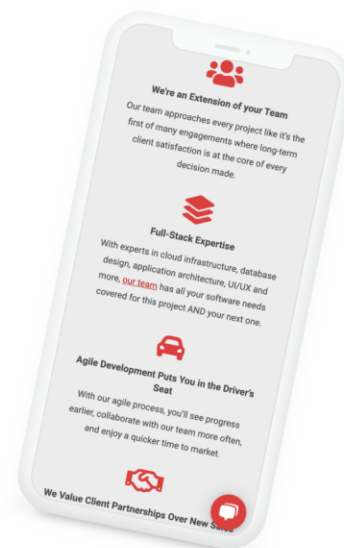
# COVID-19 Data Warehouse



An award-winning cloud, data, and software services firm.

Slide Deck:

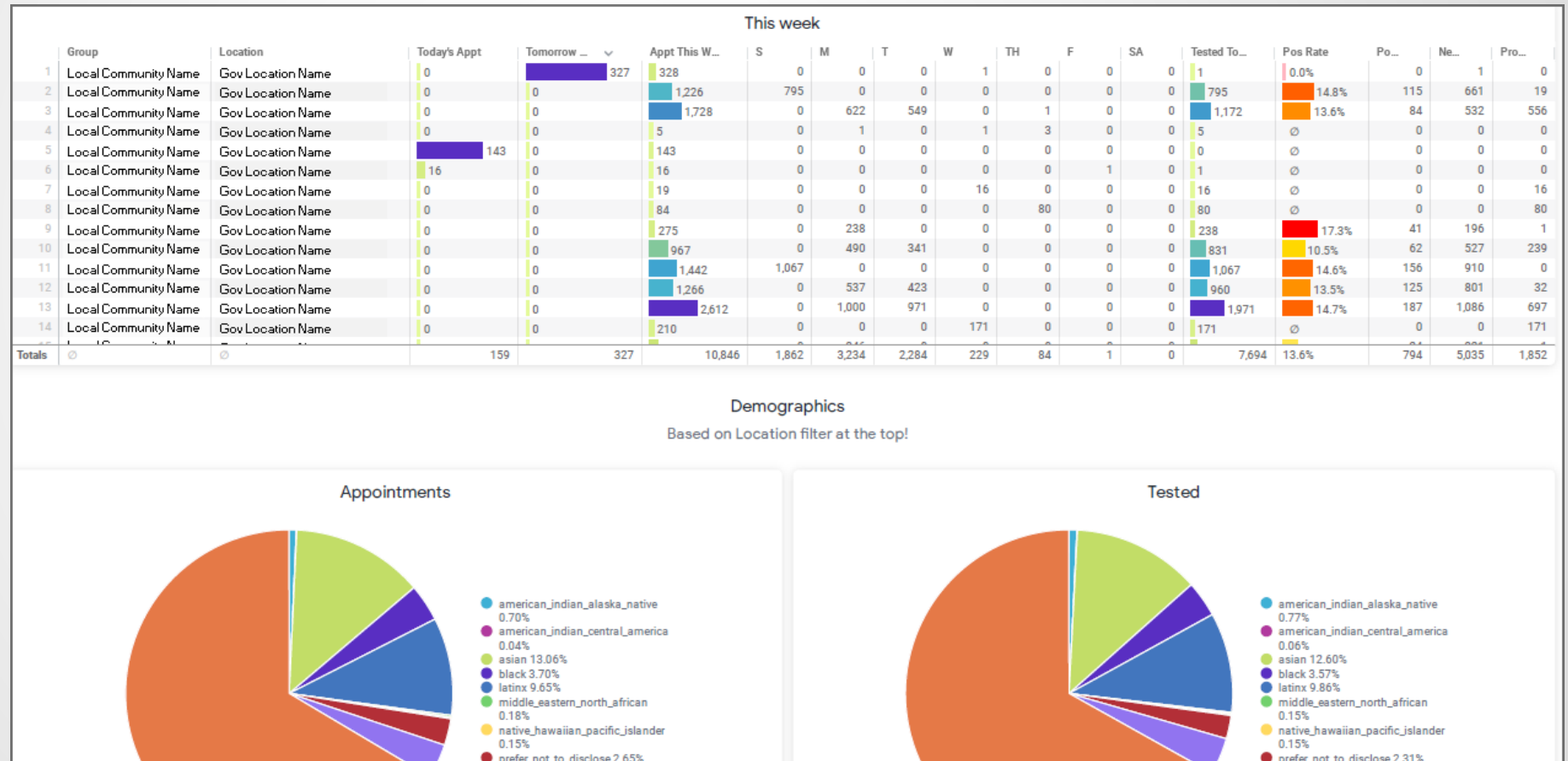
<https://kevinmartin.tech/covid-19-data-warehouse-session>



Emergent Software

# COVID-19 Data Warehouse Overview

Emergent Software gets to build cool things. I got to build the data warehouse that tracks positive COVID-19 tests for a states residents as part of a holistic data analytics solution.



# Who is Emergent Software

20 Years

Of consulting across the Microsoft tech stack

Nationally recognized for quality, sustained client partnerships, and breadth of services

Flexible engagement options

Senior-level consulting talent

6

Gold Microsoft Partner Competencies



# What we do



# Software Development

- Custom Web Applications & SaaS Products
- Complex CMS-Driven Websites
- System Integrations & API's
- Application Modernization
- Security and Accessibility Audits & Remediation
- Continuous Development & Production Support



# Cloud Transformation

- Application Rationalization & Refactoring
- Azure Cloud Migrations
- Cloud Backups & Disaster Recovery
- Office 365 Consulting & Implementation
- Azure Landing Zones
- Windows Virtual Desktop



# Data Engineering

- Data Warehousing & ETL Development
- Business Intelligence & Reporting
- Database Development / Modeling Services
- SQL Server Cloud Migration
- Production & Development DBA Support
- SQL Server Managed Services



# Clients we work with

100+

Actively engaged clients nationwide

317

Projects delivered in 2021





# We Are Hiring

---

- Open Positions

- Mid-Level SQL Server DBA & Developer
- Senior Software Architect (C# and .NET)
- Senior Software Engineer (C# and .NET)
- Senior Sales Executive

- Why Work Here

- **Variety of Work** – you are not performing the same 1 year of work, every year
- **Flexibility** – we are fully remote (indefinitely)
- **Growth** – surrounded by teammates who enjoy this type of work and share their knowledge

- Hiring Process

- Application (5 minutes)
- Online Assessment (40-50 minutes)
- Initial Phone Interview (30-45 minutes)
- At-Home SQL Server Challenge (3-4 hours)
- Technical Interview (60-90 minutes)
- Job Offer!

<https://www.emergentsoftware.net/careers>



RunAll.dtsx [Design] - Control Flow - Data Flow - Parameters - Event Handlers - Package Explorer

Solution Explorer

Search Solution Explorer (Ctrl+;)

- Solution
  - Database
    - CCTD
  - ETL (SQL Server 2017)
    - Project.params
    - Connection Managers
      - DW.conmgr
    - SSIS Packages
      - RunAll.dtsx
    - Package Parts
      - Control Flow
      - Miscellaneous
    - Linked Azure Resources
      - Azure-SSIS Integration Runtime
      - Azure Storage

Solution Explorer | Git Changes

Properties

RunAll Package

PackagePassword	
ProtectionLevel	DontSaveSensitive
<b>Transactions</b>	
IsolationLevel	Serializable
TransactionOption	Supported
<b>Version</b>	
VersionBuild	191
VersionComments	
VersionGUID	{AAB5AE50-6F06-400D-9811-231C}
VersionMajor	1
VersionMinor	0

**Name**  
Specifies the name of the object.

Connection Managers

Dimension Mappings Export | Dimension Mappings Import | LogFile | Source1 Data Source | Source2 Data Source | (project) DW

*I'm planning to make a film series on databases. I've got the first part ready, but I can't think of a SQL. 😊*

# Data Source File

- ▼ ETL
  - ▼ CCTD
    - ▼ archive
      - Source 1
      - Source 2
    - ▼ dimensions
    - ▼ process
      - archive
      - templates
      - inbound**
      - logs
    - ▼ misc
      - SSISPackages

_GOV_Statistics_Export_(Imported_On_20210209_0931_26).csv							
File	Home	Insert	Page Layout	Formulas	Data	Review	View
A1	Gov Statistics Organization Name						
A	B	C	D	E	F	G	H
1	Gov Statistics Organization Name	Gov Statistics Test Group Name	Gov Statistics Location Name	Gov Statistics Test ID	Gov Statistics Administered	Gov Statistics Administered At Time	Gov Statistics Appointment At Time
2	State Agency	Local Community Name	Gov Location Name	666193			12/5/2020 23:45
3	State Agency	Local Community Name	Gov Location Name	232851			10/6/2020 19:00
4	State Agency	Local Community Name	Gov Location Name	626172	FALSE		12/3/2020 20:30
5	State Agency	Local Community Name	Gov Location Name	237986			10/8/2020 19:45
6	State Agency	Local Community Name	Gov Location Name	562493	FALSE		11/23/2020 21:30
7	State Agency	Local Community Name	Gov Location Name	229508			10/7/2020 18:00
8	State Agency	Local Community Name	Gov Location Name	543914			11/20/2020 18:15
9	State Agency	Local Community Name	Gov Location Name	231438			10/7/2020 21:30
10	State Agency	Local Community Name	Gov Location Name	562478	FALSE		11/23/2020 21:30
11	State Agency	Local Community Name	Gov Location Name	228043			10/6/2020 18:00
12	State Agency	Local Community Name	Gov Location Name	562936	FALSE		11/23/2020 23:15
13	State Agency	Local Community Name	Gov Location Name	231705			10/7/2020 17:45

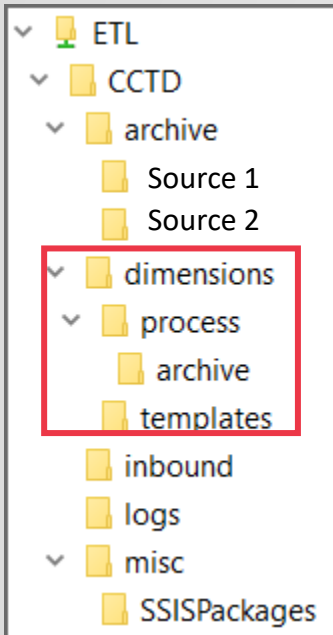
  

_Demographic_Export_02082021_(Imported_On_20210209_0928_37).csv													
File	Home	Insert	Page Layout	Formulas	Data	Review	View	Developer	Help	Team			
A1	reference_id												
A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	reference_id	restricted_id	Patient_race	Patient_ethnicity	Patient_gender	Patient_zip_code	Testing_site	Symptom_onset	Days_experiencing_symptoms	Symptoms_experienced	Result	tested_at	at_home_or_bulk
2	fr000105144228	95d44026-0488-4351-972d-9eb4e7bbf52	White	Non-Hispanic/Latino	FEMALE	55811	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 21:29	bulk
3	fr000386841840	1e42dce2-109a-4ac8-8ea5-f8c16eb07bb0	White	Unknown	MALE	55427	Gov Agency - City Name	2/6/2021	2	[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 20:31	bulk
4	fr000685143238	02394bf7-7721-48f6-baff-233a80be6145	White	Non-Hispanic/Latino	MALE	55416	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 18:52	bulk
5	fr000822347957	ff199b0f-5540-4c46-b163-31cc432187fc	White	Non-Hispanic/Latino	MALE	55374	Gov Agency - City Name			[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 16:25	bulk
6	fr001276120514	9a128fc5-2b74-41b7-b50c-c7eb550b0438	White	Unknown	MALE	55812	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 19:10	bulk
7	fr001584167195	22ead9b6-d982-4ee0-a7a8-3bfb9b0892e5	White	Non-Hispanic/Latino	MALE	56560	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 18:41	bulk
8	fr001610477641	5f822461-3698-484b-9c73-76f705200f0b	White	Non-Hispanic/Latino	FEMALE	55436	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 23:55	bulk
9	fr001821818456	456fd05e-ca1c-407e-b051-dcf2a514717d	White	Non-Hispanic/Latino	MALE	56560	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 17:22	bulk
10	fr002797957864	07fb2f9c-e685-452c-9f77-b4ad5a5af9da	White	Non-Hispanic/Latino	FEMALE	55604	Gov Agency - City Name	2/7/2021	1	[MALAISE]	Not Detected	2/7/2021 20:56	bulk
11	fr003298013969	6076a472-21c1-4e4d-8f2e-3b5c930e818e	White	Unknown	MALE	56377	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 21:26	bulk
12	fr003299313534	04c85859-3ccc-4dc8-b3b1-240b0c9ce2bd	White	Non-Hispanic/Latino	FEMALE	55362	Gov Agency - City Name			[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 17:53	bulk
13	fr004035238314	bbbfee1a-2a36-43b3-9646-b41d3f4ba74b	White	Non-Hispanic/Latino	MALE	55406	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 13:28	bulk
14	fr004412914573	73fc67e3-3283-4762-ba58-c73863e9fc7e	White	Non-Hispanic/Latino	MALE	56057	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 17:09	bulk
15	fr004465377326	9e0aca8d-2fdf-446f-9e36-fd910d936d37	Asian	Non-Hispanic/Latino	FEMALE	55109	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/2021 16:41	bulk
16	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
17	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
18	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
19	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
20	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
21	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
22	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
23	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
24	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
25	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
26	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
27	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
28	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
29	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
30	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
31	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
32	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk
33	fr004500532005	447629b0-10ed-fc0b0b0b0b0b0b0b	Black	Non-Hispanic/Latino	MALE	55660	Gov Agency - City Name	2/8/2021		[NO_SYMPTOMS_OF_COVID]	Not Detected	2/7/23/2021 14:50	bulk



Emergent Software 1 data source is “truncate & dump”, 1 is incremental. No PII data!

# Dimension Mapping Excel



ETL > CCTD > dimensions

Name	Date modified	Type	Size
process	2/9/2021 9:28 AM	File folder	
templates	1/6/2021 12:26 PM	File folder	
DimensionMappings.xls	2/9/2021 9:37 AM	Microsoft Exce...	
<input checked="" type="checkbox"/> INSTRUCTIONS.txt	12/11/2020 1:02 PM	Text Document	

ETL > CCTD > dimensions > templates

Name	Date modified	Type	Size
DoNotEditOrMoveThisDimensionMappingsTemplate - Copy.xls	1/5/2021 9:15 PM	Microsoft Exce...	41 KB
DoNotEditOrMoveThisDimensionMappingsTemplate.xls	1/6/2021 12:26 PM	Microsoft Exce...	41 KB
<input checked="" type="checkbox"/> README.txt	12/8/2020 9:34 PM	Text Document	1 KB

INSTRUCTIONS.txt - Notepad

File Edit Format View Help

These instructions explain how to map the original dimension value to the conformed value. With conformed dimensions they can be used across many fact tables which will allow reporting consistency.

1. Open the DimensionsMapping.xls file in Excel
  - a. The folder location is a variable and can be changed
  - b. The account that runs the SSIS package must have permissions to the folder location
2. Only edit the 'Conformed Value' column
  - a. The 'Facilty Name' column has been included to decern the data source original dimension value
3. Repeat for each dimension worksheet (tab)
4. Close the DimensionsMapping.xls file in Excel, the scheduled import will fail if the file is kept open in Excel
5. Move (cut and paste or drag and drop the DimensionsMapping.xls file into the 'process' folder
6. The updated dimension mappings will be processed on the next scheduled import
7. A new DimensionsMapping.xls file will be created on the next scheduled import



# Dimension Mapping Excel Document

AutoSave  Off DimensionMappings.xls - Compatibility Mode

File Home Insert Page Layout Formulas Data Review View Developer Help Team

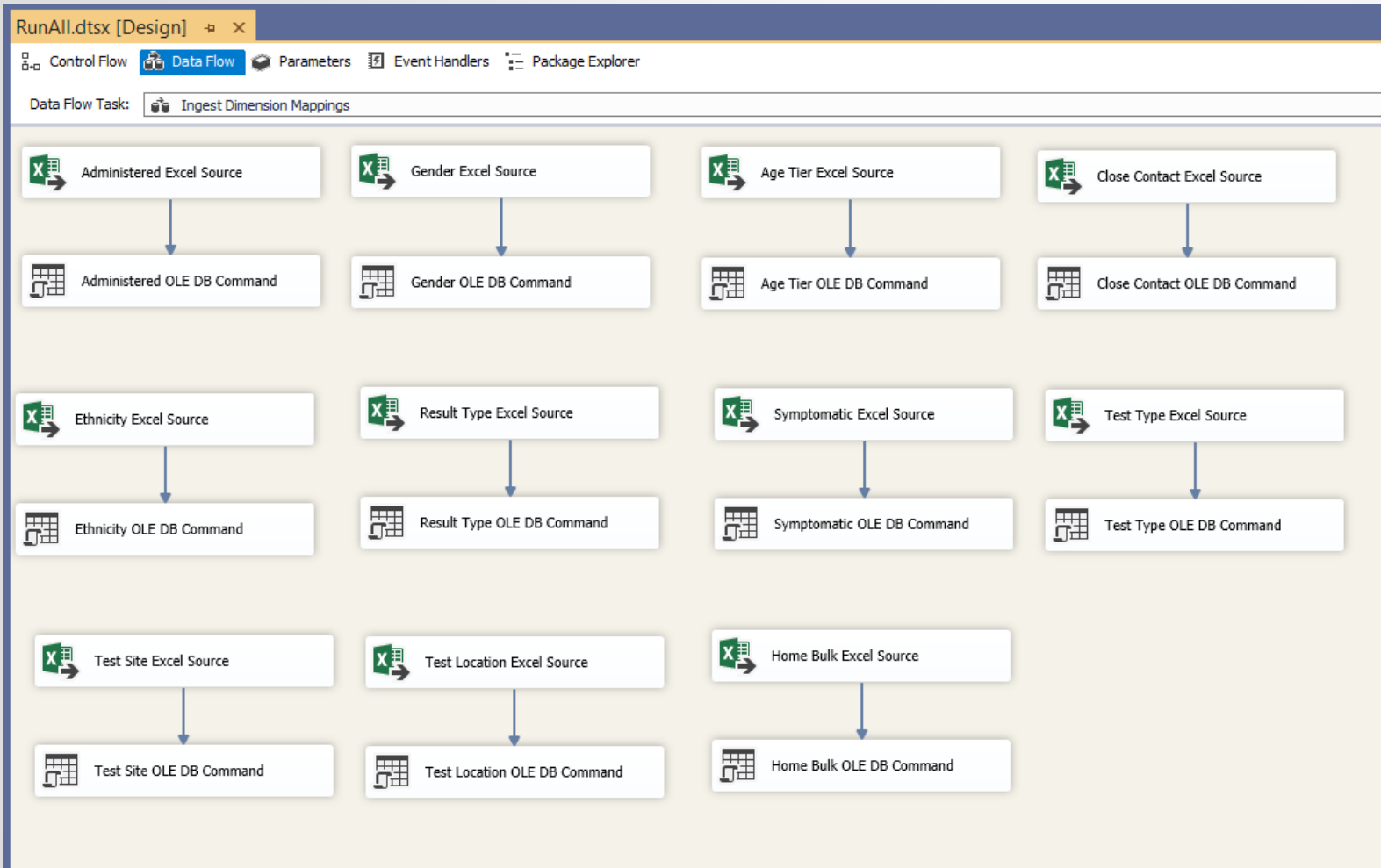
C5    'Source 1

	A	B	C	D	E	F	G	H	I	J	K
1	<b>Gender Map Key</b>	<b>Facility Key</b>	<b>Facility Name</b>	<b>Original Value</b>	<b>Conformed Value</b>						
2	1	1	Source 1	cisgender	Cisgender						
3	2	1	Source 1	female	Female						
4	3	1	Source 1	gender_queer	Queer						
5	4	1	Source 1	male	Male						
6	5	1	Source 1	non_binary	Non-Binary						
7	6	1	Source 1	other	Other						
8	7	1	Source 1	prefer_not_to_disclose	Prefer Not To Disclose						
9	8	1	Source 1	prefer_to_self_describe	Prefer Not To Disclose						
10	9	1	Source 1	transgender	Transgender						
11	10	1	Source 1	unknown	Unknown						
12	11	2	Source 2		Unknown						
13	12	2	Source 2	DECLINE_TO_IDENTIFY	Prefer Not To Disclose						
14	13	2	Source 2	FEMALE	Female						
15	14	2	Source 2	MALE	Male						
16											

Administered **Gender** Age Tier Close Contact Ethnicity Result Type Symptomatic Test Type Test Site Test Location Home Bulk ...

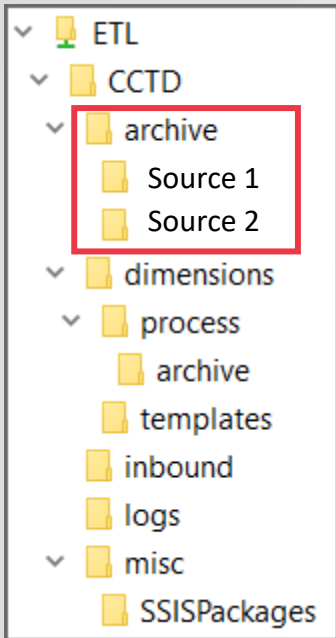
Ready

# Dimension Mapping from Excel Process



1. If a new Excel file is found
2. Then each Excel dimension sheet is imported into a mapping table
3. Then the Excel file is archived with a new "imported on" file name
4. Then the mapping tables are processed to make conformed dimensions
5. Then the Excel template is copied to "Dimension/Process" folder for any new dimension mappings

# Archive Folder for Ingested Source Files



File Explorer window showing the contents of the 'Source 1' folder:

Path: ETL > CCTD > archive > Source 1

Name	Date modified	Type	Size	Date created
_GOV_Statistics_Export_(Imported_On_20210209_0931_26).csv	2/9/2021 4:10 AM	Microsoft Exce...	56,274 KB	2/9/2021 4:10 AM
_GOV_Statistics_Export_(Imported_On_20210208_0503_08).csv	2/8/2021 4:10 AM	Microsoft Exce...	56,251 KB	2/8/2021 4:10 AM
_GOV_Statistics_Export_(Imported_On_20210208_0902_07).csv	2/8/2021 4:10 AM	Microsoft Exce...	56,251 KB	2/8/2021 4:10 AM
_GOV_Statistics_Export_(Imported_On_20210207_0502_48).csv	2/7/2021 4:40 AM	Microsoft Exce...	56,250 KB	2/7/2021 4:40 AM
_GOV_Statistics_Export_(Imported_On_20210207_0902_12).csv	2/7/2021 4:40 AM	Microsoft Exce...	56,250 KB	2/7/2021 4:40 AM
_GOV_Statistics_Export_(Imported_On_20210206_0502_38).csv	2/6/2021 4:10 AM	Microsoft Exce...	56,247 KB	2/6/2021 4:10 AM



# Logs

The screenshot shows a file explorer window with the following folder structure:

- ETL
  - CCTD
    - archive
      - Source 1
      - Source 2
    - dimensions
    - process
      - archive
      - templates
      - inbound
      - logs**
      - misc
    - SSISPackages

The main pane shows the 'logs' folder containing three log files:

Name	Date modified	Type	Size	Date created
<input checked="" type="checkbox"/> RunAll_2021-02-09_09-28-.log	2/9/2021 9:37 AM	LOG File	54 KB	2/9/2021 9:28 AM
<input type="checkbox"/> RunAll_2021-02-09_09-27-.log	2/9/2021 9:28 AM	LOG File	3 KB	2/9/2021 9:27 AM
<input type="checkbox"/> RunAll_2021-02-09_09-01-.log	2/9/2021 9:01 AM	LOG File	10 KB	2/9/2021 9:01 AM

The Notepad window displays the content of the selected log file:

```
*RunAll_2021-02-09_09-28-.log - Notepad
File Edit Format View Help
#Fields: event,computer,operator,source,sourceid,executionid,starttime,endtime,datacode,databytes,message
PackageStart,DOMAIN,Domain\Kevin,RunAll,{DEE2D654-9C62-41EC-8B3A-56D37F789366},{84079DDC-59B7-44A6-97F4-48DB13F0585D},2/9/2021 9:28:20 AM,2/9/2021 9:28:20 AM,0,0x,Beginning of package execution.

OnWarning,DOMAIN,Domain\Kevin,Check for Dimension File,{1484388F-8079-4219-AF81-91B8175A9452},{84079DDC-59B7-44A6-97F4-48DB13F0585D},2/9/2021 9:28:21 AM,2/9/2021 9:28:21 AM,-2147368956,0x,The For Each File enumerator is empty. The For Each File enumerator did not find any files that matched the file pattern, or the specified directory was empty.

OnWarning,DOMAIN,Domain\Kevin,RunAll,{DEE2D654-9C62-41EC-8B3A-56D37F789366},{84079DDC-59B7-44A6-97F4-48DB13F0585D},2/9/2021 9:28:21 AM,2/9/2021 9:28:21 AM,-2147368956,0x,The For Each File enumerator is empty. The For Each File enumerator did not find any files that matched the file pattern, or the specified directory was empty.

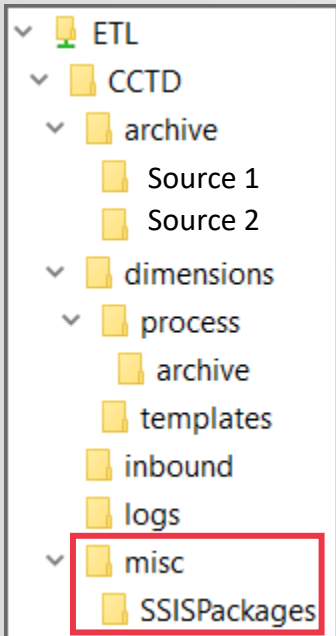
OnInformation,DOMAIN,Domain\Kevin,Import Source Files,{FBF7D62C-DA61-484C-BBB1-DA50445BBB25},{84079DDC-59B7-44A6-97F4-48DB13F0585D},2/9/2021 9:28:22 AM,2/9/2021 9:28:22 AM,1074016266,0x,Validation phase is beginning.
```

@[\$Project::LoggingFolder] + @[System::PackageName] + "\_" + REPLACE(REPLACE(SUBSTRING((DT\_STR, 50, 1252) GETDATE(), 1, 17), ":", "-"), " ", "\_") + ".log"





# SSIS Deployment Files



> ETL > CCTD > misc > SSISPackages

<input type="checkbox"/> Name ^	Date modified	Type	Size
ETL v1.ispac	12/4/2020 10:57 A...	Integration Services Project Deployment File	15 KB
ETL v2.ispac	12/10/2020 5:55 P...	Integration Services Project Deployment File	31 KB
ETL v3.ispac	12/11/2020 1:04 P...	Integration Services Project Deployment File	31 KB
ETL v4.ispac	1/4/2021 5:31 PM	Integration Services Project Deployment File	31 KB
ETL v5.ispac	1/6/2021 1:46 AM	Integration Services Project Deployment File	36 KB
ETL v6.ispac	1/8/2021 11:24 PM	Integration Services Project Deployment File	37 KB
ETL v7.ispac	1/20/2021 9:58 AM	Integration Services Project Deployment File	37 KB



- [-] CCTD
  - [+] Database Diagrams
  - [-] Tables
    - [+] System Tables
    - [+] FileTables
    - [+] External Tables
    - [+] Graph Tables
    - [+] Dimension.Administered
    - [+] Dimension.AgeTier
    - [+] Dimension.CloseContact
    - [+] Dimension.Date
    - [+] Dimension.Ethnicity
    - [+] Dimension.Facility
    - [+] Dimension.Gender
    - [+] Dimension.HomeBulk
    - [+] Dimension.Number
    - [+] Dimension.Person
    - [+] Dimension.ResultType
    - [+] Dimension.StateProvince
    - [+] Dimension.Symptomatic
    - [+] Dimension.TestLocation
    - [+] Dimension.TestSite
    - [+] Dimension.TestType
    - [+] Dimension.Time
    - [+] Dimension.ZipCode
    - [+] Fact.Result
    - [+] Integration.AdministeredMap
    - [+] Integration.AgeTierMap
    - [+] Integration.CloseContactMap
    - [+] Integration.DevelopCheckToSkip
    - [+] Integration.EthnicityMap
    - [+] Integration.GenderMap
    - [+] Integration.HomeBulkMap
    - [+] Integration.Popular Demographics in the United States - County
    - [+] Integration.Popular Demographics in the United States - State
    - [+] Integration.Popular Demographics in the United States - Tract
    - [+] Integration.Popular Demographics in the United States - Zip Code

- [+] Integration.Source1
- [+] Integration.ResultTypeMap
- [+] Integration.SymptomaticMap
- [+] Integration.TestLocationMap
- [+] Integration.TestSiteMap
- [+] Integration.TestTypeMap
- [+] Integration.Source2
- [+] Integration.Source2History

# This data warehouse has:

- 1 Fact Table
- 18 Dimension Tables
- 11 Dimension Mapping Tables
- ? Vaccination Data, Developed Later

# Test Results Fact Table

SELECT \* FROM Fact.Result;

	Result Key	Source Test Id	Facility Key	Test Site Key	Test Location Key	State Key	Zip Code Key	Test Type Key	Administered Key	Result Type Key	Result Date Key
1	17421301	-1	1	126	-1	1	-1	8	1	9	-1
2	17421302	-1	1	126	-1	1	-1	8	1	9	-1
3	17421303	-1	1	126	-1	1	-1	8	1	9	-1
4	17421304	-1	1	126	-1	1	-1	8	1	9	-1

	Result Time Key	Administered Date Key	Administered Time Key	Appointment Date Key	Appointment Time Key	Check In Date Key	Check In Time Key	Gender Key
1	-1	-1	-1	20210119	153000	-1	-1	10
2	-1	-1	-1	20210119	181500	-1	-1	10
3	-1	-1	-1	20210120	104500	-1	-1	10
4	-1	-1	-1	20210120	161500	-1	-1	10

	Ethnicity K...	Age	Age Tier K...	Close Contact K...	Symptomatic Onset Date Key	Days Experiencing Symptoms Number K...	Symptomatic Key	Home Bulk Key	Person Key
1	11	0	12	4	-1	-1	1	4	-1
2	11	0	12	4	-1	-1	1	4	-1
3	11	0	12	4	-1	-1	1	4	-1
4	11	0	12	4	-1	-1	1	4	-1

# Date Dimension

Connect 100% `SELECT * FROM Dimension.Date AS D;`

Dimension.Date

- Columns
  - Date Key (PK, int, not null)
  - Full Date Time (datetime2(7), not null)
  - Date Name (varchar(10), not null)
  - Date Name US (varchar(10), not null)
  - Date Name EU (varchar(10), not null)
  - Day Of Week (tinyint, not null)
  - Day Name Of Week (varchar(9), not null)
  - Day Of Month (tinyint, not null)
  - Day Of Year (smallint, not null)
  - Weekday Weekend (char(7), not null)
  - Week Of Year (tinyint, not null)
  - Month Name (varchar(9), not null)
  - Month Of Year (tinyint, not null)
  - Last Day Of Month (varchar(3), not null)
  - Work Day (varchar(3), not null)
  - Calendar Quarter (tinyint, not null)
  - Calendar Year (smallint, not null)
  - Calendar Year Month (char(7), not null)
  - Calendar Year Quarter (char(6), not null)
  - Fiscal Month Of Year (tinyint, not null)
  - Fiscal Quarter (tinyint, not null)
  - Fiscal Year (int, not null)
  - Fiscal Year Month (char(7), not null)
  - Fiscal Year Quarter (char(6), not null)
- Keys
- Constraints
- Triggers
- Indexes
- Statistics

Results

	Date Key	Full Date Time	Date Name	Date Name US	Date Name EU	Day Of Week	Day Name Of Week	Day Of Mo
1	-1	0001-01-01 00:00:00.0000000	UNKNOWN	UNKNOWN	UNKNOWN	2	UNKNOWN	1
2	10101	0001-01-01 00:00:00.0000000	0001/01/01	01/01/0001	01/01/0001	2	Monday	1
3	10102	0001-01-02 00:00:00.0000000	0001/01/02	01/02/0001	02/01/0001	3	Tuesday	2
4	10103	0001-01-03 00:00:00.0000000	0001/01/03	01/03/0001	03/01/0001	4	Wednesday	3
5	10104	0001-01-04 00:00:00.0000000	0001/01/04	01/04/0001	04/01/0001	5	Thursday	4
6	10105	0001-01-05 00:00:00.0000000	0001/01/05	01/05/0001	05/01/0001	6	Friday	5
7	10106	0001-01-06 00:00:00.0000000	0001/01/06	01/06/0001	06/01/0001	7	Saturday	6
8	10107	0001-01-07 00:00:00.0000000	0001/01/07	01/07/0001	07/01/0001	1	Sunday	7
9	10108	0001-01-08 00:00:00.0000000	0001/01/08	01/08/0001	08/01/0001	2	Monday	8
10	10109	0001-01-09 00:00:00.0000000	0001/01/09	01/09/0001	09/01/0001	3	Tuesday	9
11	10110	0001-01-10 00:00:00.0000000	0001/01/10	01/10/0001	10/01/0001	4	Wednesday	10
12	10111	0001-01-11 00:00:00.0000000	0001/01/11	01/11/0001	11/01/0001	5	Thursday	11
13	10112	0001-01-12 00:00:00.0000000	0001/01/12	01/12/0001	12/01/0001	6	Friday	12
14	10113	0001-01-13 00:00:00.0000000	0001/01/13	01/13/0001	13/01/0001	7	Saturday	13
15	10114	0001-01-14 00:00:00.0000000	0001/01/14	01/14/0001	14/01/0001	1	Sunday	14
16	10115	0001-01-15 00:00:00.0000000	0001/01/15	01/15/0001	15/01/0001	2	Monday	15
17	10116	0001-01-16 00:00:00.0000000	0001/01/16	01/16/0001	16/01/0001	3	Tuesday	16
18	10117	0001-01-17 00:00:00.0000000	0001/01/17	01/17/0001	17/01/0001	4	Wednesday	17
19	10118	0001-01-18 00:00:00.0000000	0001/01/18	01/18/0001	18/01/0001	5	Thursday	18
20	10119	0001-01-19 00:00:00.0000000	0001/01/19	01/19/0001	19/01/0001	6	Friday	19



# Time Dimension

Connect `SELECT * FROM Dimension.Time AS T;`

100 %

Results Messages

	Time Key	Hour 24	Hour 24 Short	Hour 24 Medium	Hour 24 Full	Hour 12	Hour 12 Short	Hour 12 Short Trim	Hour 12 Medium	Hour 12 Full
1	-1	0	UN	UNKWN	UNKNOWN	12	UNKWN	UNKWN	UNKNOWN	UNKWN
2	0	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
3	1	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
4	2	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
5	3	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
6	4	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
7	5	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
8	6	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
9	7	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
10	8	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
11	9	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
12	10	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
13	11	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
14	12	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
15	13	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
16	14	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
17	15	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
18	16	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
19	17	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
20	18	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
21	19	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
22	20	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM
23	21	0	00	00:00	00:00:00	12	12 AM	12 AM	12:00 AM	12:00 AM

Dimension.Time

- Columns
  - Time Key (PK, int, not null)
  - Hour 24 (int, not null)
  - Hour 24 Short (char(2), not null)
  - Hour 24 Medium (char(5), not null)
  - Hour 24 Full (char(8), not null)
  - Hour 12 (int, not null)
  - Hour 12 Short (char(5), not null)
  - Hour 12 Short Trim (varchar(5), not null)
  - Hour 12 Medium (char(8), not null)
  - Hour 12 Medium Trim (varchar(8), not null)
  - Hour 12 Full (char(11), not null)
  - Hour 12 Full Trim (varchar(11), not null)
  - AM PM Code (bit, not null)
  - AM PM (char(2), not null)
  - Minute (int, not null)
  - Minute Of Day (int, not null)
  - Minute Code (int, not null)
  - Minute Short (char(2), not null)
  - Minute 24 Full (char(8), not null)
  - Minute 12 Full (char(11), not null)
  - Minute 12 Full Trim (varchar(50), not null)
  - Half Hour Code (int, not null)
  - Half Hour (bit, not null)
  - Half Hour Short (char(2), not null)
  - Half Hour 24 Full (char(8), not null)
  - Half Hour 12 Full (char(11), not null)
  - Half Hour 12 Full Trim (varchar(50), not null)
  - Second (int, not null)
  - Second Of Day (int, not null)
  - Second Short (char(2), not null)
  - Full Time 24 (char(8), not null)
  - Full Time 12 (char(11), not null)
  - Full Time 12 Trim (varchar(50), not null)
  - Full Time (time(7), not null)



# Gender Dimension

Dimension.Gender

- Columns
  - Gender Key (PK, int, not null)
  - Gender Name (nvarchar(200), not null)
- Keys
  - Gender\_Key
- Constraints
- Triggers
- Indexes
  - Gender\_Key (Clustered)
  - Gender\_Name (Non-Unique, Non-Clustered)
- Statistics

Integration.GenderMap

- Columns
  - Gender Map Key (PK, int, not null)
  - Facility Key (int, not null)
  - Original Value (nvarchar(200), null)
  - Conformed Value (nvarchar(200), null)
- Keys
  - Gender\_Key
- Constraints
- Triggers
- Indexes
  - Conformed\_Value (Non-Unique, Non-Clustered)
  - Gender\_Key (Clustered)
  - Original\_Value (Non-Unique, Non-Clustered)
- Statistics

```
SELECT * FROM Dimension.Gender AS G;
SELECT * FROM Integration.GenderMap AS GM;
```

	Gender Key	Gender Name
1	1	Cisgender
2	2	Female
3	4	Male
4	5	Non-Binary
5	6	Other
6	7	Prefer Not To Disclose
7	3	Queer
8	9	Transgender
9	10	Unknown
10	-1	UNMAPPED

	Gender Map Key	Facility Key	Original Value	Conformed Value
1	1	1	cisgender	Cisgender
2	2	1	female	Female
3	3	1	gender_queer	Queer
4	4	1	male	Male
5	5	1	non_binary	Non-Binary
6	6	1	other	Other
7	7	1	prefer_not_to_disclose	Prefer Not To Disclose
8	8	1	prefer_to_self_describe	Prefer Not To Disclose
9	9	1	transgender	Transgender
10	10	1	unkown	Unknown
11	11	2	NULL	Unknown
12	12	2	DECLINE_TO_IDENTIFY	Prefer Not To Disclose



# Integration.DimensionsCreateDefaults

---

```
CREATE OR ALTER PROCEDURE Integration.DimensionsCreateDefaults
AS
BEGIN
    SET NOCOUNT ON;

    /*****
    SET IDENTITY_INSERT Dimension.Gender ON;

    INSERT Dimension.Gender ([Gender Key], [Gender Name])
    SELECT
        -1
        , 'UNMAPPED'
    WHERE
        NOT EXISTS (
            SELECT
                1
            FROM
                Dimension.Gender WITH (UPDLOCK, SERIALIZABLE)
            WHERE
                [Gender Key] = -1
        );

    SET IDENTITY_INSERT Dimension.Gender OFF;

    *****/
```

<https://emergentsoftware.github.io/SQL-Server-Development-Assessment/best-practices-and-potential-findings/sql-code-conventions#76>



# Integration.DimensionsMerge

---

```
CREATE OR ALTER PROCEDURE Integration.DimensionsMerge
AS
BEGIN
    SET NOCOUNT ON;

    /*****
MERGE INTO Dimension.Gender AS T
USING (
    SELECT
        [Gender Name] = GM.[Conformed Value]
    FROM
        Integration.GenderMap AS GM
    WHERE
        GM.[Conformed Value] IS NOT NULL
    GROUP BY
        GM.[Conformed Value]
    ) AS S
ON (T.[Gender Name] = S.[Gender Name])
WHEN MATCHED THEN
    UPDATE SET
        T.[Gender Name] = S.[Gender Name]
WHEN NOT MATCHED BY TARGET THEN
    INSERT ([Gender Name])
    VALUES
        (S.[Gender Name]);
*****/
```

<https://emergentsoftware.github.io/SQL-Server-Development-Assessment/best-practices-and-potential-findings/sql-code-conventions#76>





# Integration.Source2HistoryMerge

```
CREATE OR ALTER PROCEDURE Integration.Source2HistoryMerge
AS
BEGIN
    SET NOCOUNT ON;

    /* Import new rows from Integration.Source2 stage table into the
    Integration.Source2History table, Need to keep original
    dimension values for when the mapping changes */
    MERGE INTO Integration.Source2History AS T
    USING (
        SELECT
            V.* /* Full columns used here */
        FROM
            Integration.Source2 AS V
    ) AS S
    ON (T.reference_id = S.reference_id)
    WHEN MATCHED AND EXISTS (
        SELECT
            S.* /* Full columns used here */
        EXCEPT
        SELECT
            T.* /* Full columns used here */
    ) THEN
        UPDATE SET
            T.reference_id = S.reference_id
            /*...*/
    WHEN NOT MATCHED BY TARGET THEN
        INSERT (
            reference_id
            /*...*/
        )
        VALUES
            (S.reference_id
            /*...*/);
END;
```

<https://emergentsoftware.github.io/SQL-Server-Development-Assessment/best-practices-and-potential-findings/sql-code-conventions#76>



# Integration.DimensionMap\_Source1

```
CREATE OR ALTER PROCEDURE Integration.DimensionMap_Source1
AS
BEGIN
    SET NOCOUNT ON;

    /**
    MERGE INTO Integration.GenderMap AS T
    USING (
        SELECT
            [Original Value] = PB.Gov_Statistics_Gender
        FROM
            Integration.Source1 AS PB
        GROUP BY
            PB.Gov_Statistics_Gender
    ) AS S
    ON (
        ISNULL(T.[Original Value], -1) = ISNULL(S.[Original Value], -1)
        AND T.[Facility Key] = 1
    )
    WHEN NOT MATCHED BY TARGET THEN
        INSERT ([Original Value], [Facility Key])
        VALUES
            (S.[Original Value], 1);

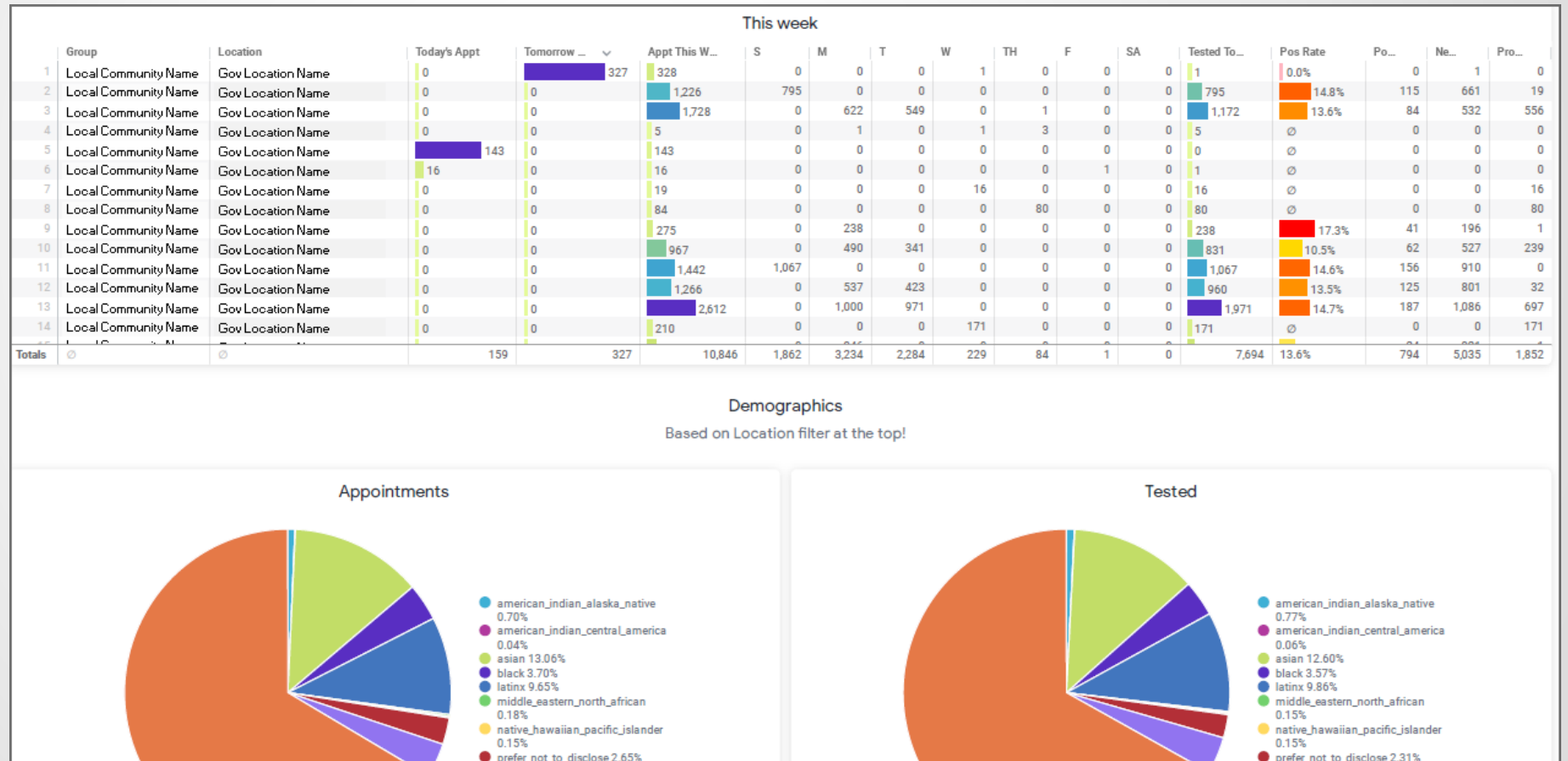
    /**
```

<https://emergentsoftware.github.io/SQL-Server-Development-Assessment/best-practices-and-potential-findings/sql-code-conventions#76>



# Results

The results are great looking visuals that allowed government officials to be informed on the COVID-19 testing results occurring all over their state. The state agency was able to make informed, quick decisions. Vaccination data was later added into the data warehouse.





END;

Do we have time for a few questions?

---

Slide Deck: <https://kevinmartin.tech/covid-19-data-warehouse-session>

*Alexa, do you have a boyfriend?*

*I'll leave the relations to the databases. 😄*