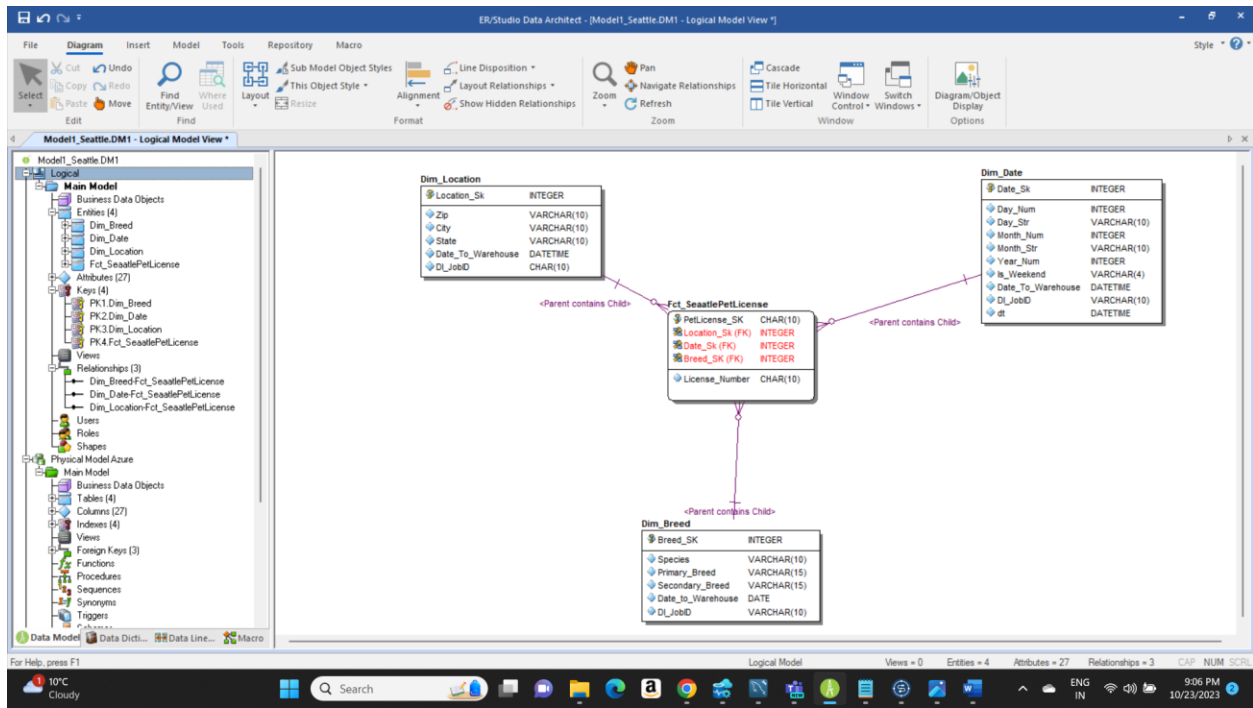
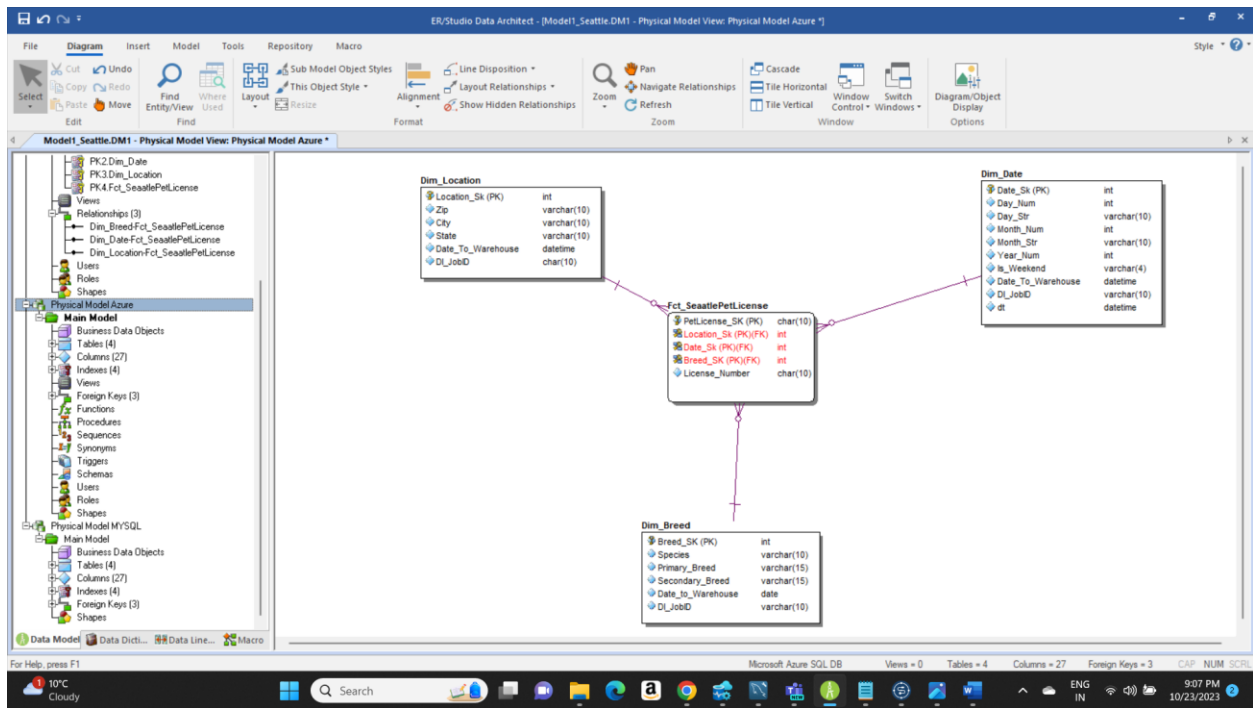


# Seattle Pet License dimensional modeling and Loading via Talend – Workshop

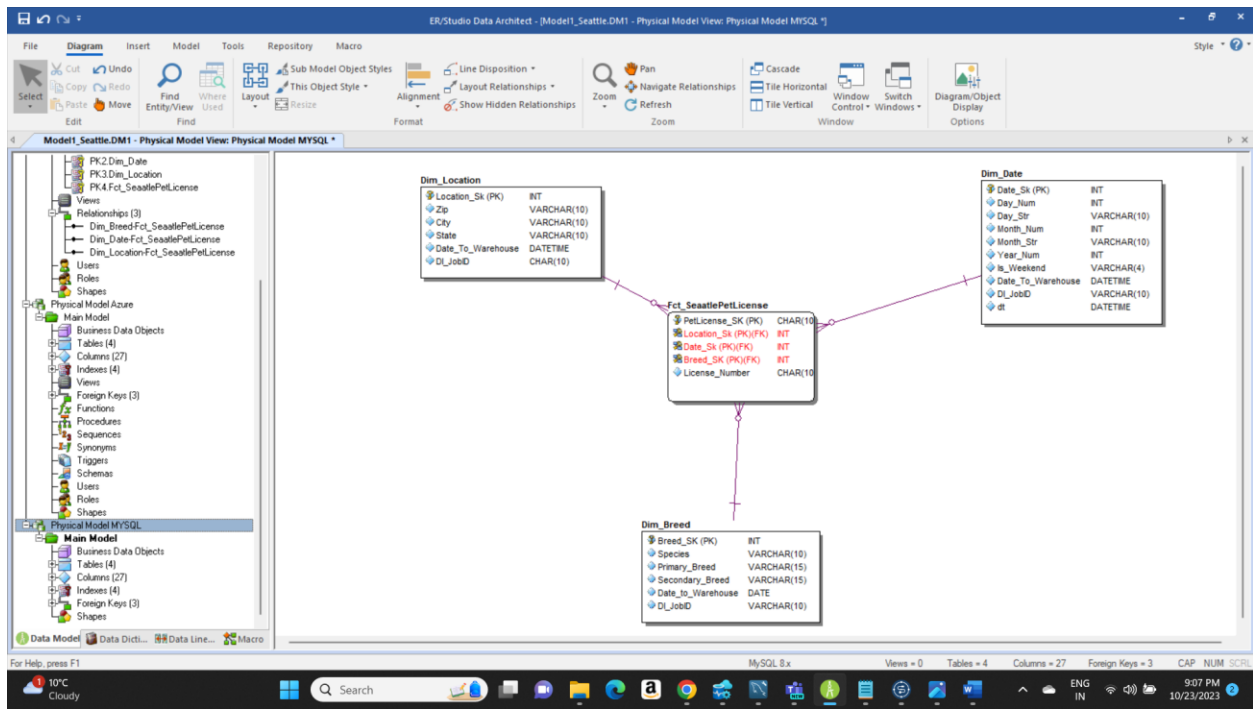
## Logical Model:



## Physical Model Azure:



## Physical Model MYSQL:



## Alteryx Workflow:

Alteryx Designer v64 - Alteryx Workflow - Seattle.yxmd

Search for tools, help, and resources...

Workflow - Configuration

Canvas Options

Layout Direction: Horizontal

Annotations: Show

Connection Progress: Show Only When Running

Finished running Alteryx Workflow - Seattle.yxmd in 0.7 seconds with 14 field conversion errors

[Learn about AMP Engine](#)

Don't show this message again

Results - Workflow - Messages

0 Errors 14 Conv Errors 0 Warnings 5 Info 5 Files

Field Summary (4) Tool #275: Tool #122: Values: "Mixed Breed, Medium (up to 44 lbs fully ...)" was truncated to 44 characters

Field Summary (4) Tool #275: Tool #122: Values: "Mixed Breed, Medium (up to 44 lbs fully ...)" was truncated to 44 characters

Field Summary (4) Tool #275: Tool #122: Values: "Mixed Breed, Medium (up to 44 lbs fully ...)" was truncated to 44 characters

Field Summary (4) Tool #275: Tool #122: Values: Field Conversion Error Limit Reached.

Designer v64 The Designer v64 reported: Beginning to compact waiting packets to reduce memory usage

Field Summary (4) Tool #275: Tool #32: Values: "Cascade MountainApexis Out of a Dream [...]" was truncated to 44 characters

Field Summary (4) Tool #275: Tool #32: Values: "Agnes "Aggie" Showball Pineapp..." was truncated to 44 characters

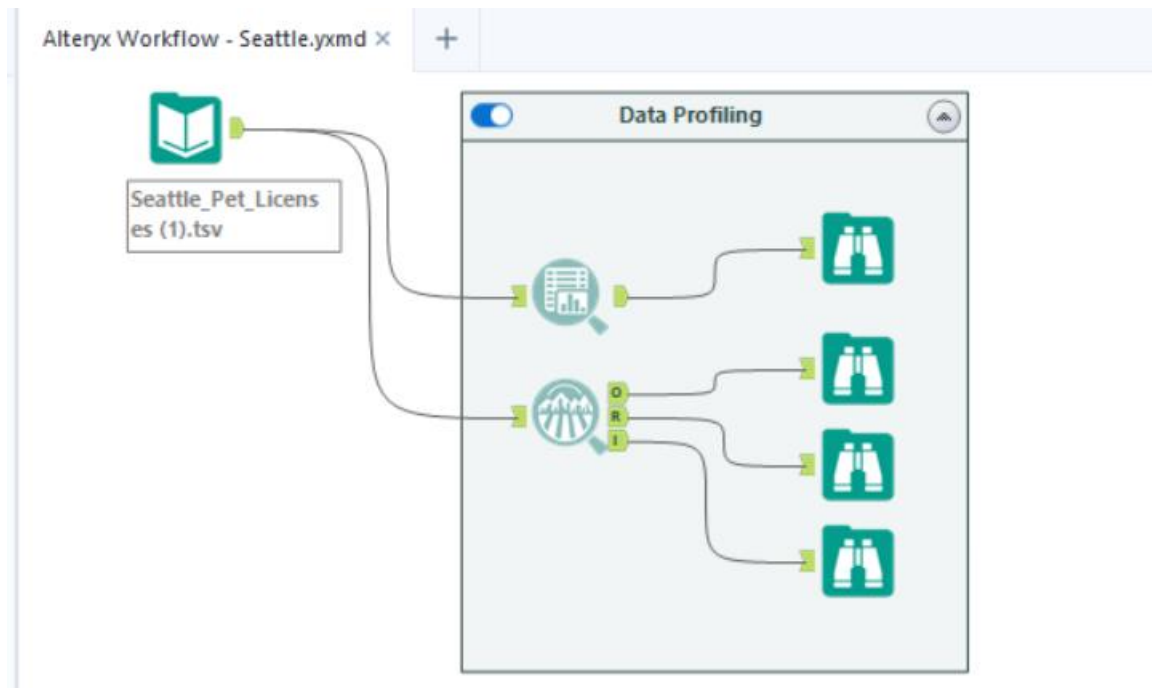
Field Summary (4) Tool #275: Tool #32: Values: "Houlene "Hula" Elizabeth Honey..." was truncated to 44 characters

Browse (5) 7 records

Browse (6) 1 records

Browse (12) 1 records

Designer v64 Finished running Alteryx Workflow - Seattle.yxmd in 0.7 seconds with 14 field conversion errors



Summary of the Values:

Summary			^
Type	Records	Data Type Size	
V_WString	147	1,073,741,823	
● Ok	133		90.48%
Unique	64		43.54%
● Null	13		8.84%
● Not Ok	1		0.68%
● Empty	0		0.00%

Length Statistics		^
Min		1
Max		59
Average		8.40
Shortest Value		0
Longest Value	File: C:\Users\muska\Downloads\Seattle_Pet_Licenses (1).tsv	
First Alphanumeric Value	"Zippy"	
Last Alphanumeric Value	"Dr." Noah Pickles	
Blanks		0
Values with Leading Whitespace		0
Values with Trailing Whitespace		1

There are 90% Okay values. There are around 8.84% Null values, and 43% unique values.

The min length of all values is 1 and max is 54.

License Issue Date	String	0 Null Values
License Number	String	0 Null Values, Avg length: 6.8
Animal's Name	String	32 Null Values, Longest length: 50
Species	String	0 Null Values, Longest Length: 4, Uniques: 4
Primary Breed	String	0 Null Values, Longest Length: 46
Secondary Breed	String	15661 Null Values, Longest length: 46
ZIP Code	String	129 Null Values, 42957 OK values

Alteryx Workflow - Seattle.yrmd x + Run

Results - Browse (5) - Input

22 of 22 Fields ☒ Cell Viewer ☒ 7 records displayed, 2,681 bytes Search Data Metadata Actions 000

Record #3, Field Remarks (V\_WString)  
Some values of this field have a small number of value counts. If Appropriate, consider combining some value levels together.

Record	Name	Field Categ...	M	Percent Missing	Unique Values	Remarks	Shortest_V...	Longest_Value	MinValueCount	MaxValueCount
1	ZIP Code	String	[ ]	0.299401	208	Some values of this field have a small number of...	981	98117	1	4,038
2	Primary Breed	String	[ ]	0	341	Some values of this field have a small number of...	Mix	Mixed Breed, Medium (up to 44 lbs fu...	1	7,307
3	License Issue Date	String	[ ]	0	907	Some values of this field have a small number of...	May 19 2020	September 01 2019	1	181
4	Animal's Name	String	[ ]	0.07427	12,781	Some values of this field have a small number of...	Z	Nuit Athoor Hecate Sappho Zebe...	1	395
5	License Number	String	[ ]	0	43,083	Some values of this field have a small number of...	9268	S107948	1	2
6	Species	String	[ ]	0	4	Some values of this field have a small number of...	Cat	Goat	4	29,917
7	Secondary Breed	String	[ ]	36.348234	263	This field has over 10% missing values. Consider L...	Mix	Mixed Breed, Medium (up to 44 lbs fu...	1	17,458

1 record displayed, 1 field, 34 KB Report Profile

1 of 1 Fields ☒ Records 1 to 1

### Frequent Values

98115	4038
98103	3689
98117	3100
98125	2416
98122	2151
98118	2084

### Summary

Count	43086
Valid	42957
Unique	207
Blank	129
Longest	98117
Shortest	981

### ZIP Code

The histogram shows the frequency of ZIP codes. The x-axis lists ZIP codes: 98115, 98027, 98057, 98060, 95621, 33029, 98095. The y-axis represents frequency from 0 to 4,000. The bar for 98115 is the tallest, reaching approximately 4,000.

### Frequent Values

98115	4038
98103	3689
98117	3100

## MYSQL Staging Table

The screenshot displays a data integration workflow in a tool like Talend. The workflow starts with a **tsvFile** component, which feeds into a **tMap\_1** component. From **tMap\_1**, the data is split into two paths: one leading to **tLogRow\_1** and another to a **mySQL** component. Performance metrics are shown for each component: **tsvFile** (43086 rows in 1.7s, 25344.71 rows/s), **tMap\_1** (43086 rows in 2.31s, 18684.3 rows/s), **tLogRow\_1** (43086 rows in 2.31s, 18684.3 rows/s), and **mySQL** (43086 rows in 2.31s, 18651.95 rows/s).

Below the workflow is the **Job SeattlePetLicense** configuration. The **Basic Run** tab shows a sample data table with columns: License\_Issue\_Date, License\_Number, Animal\_s\_Name, Species, Primary\_Breed, Secondary\_Breed, ZIP\_Code, DI\_CreatedJob, and DI\_ProcessID. The **Default** configuration table shows settings for the job, including Name, Value, and a list of components.

At the bottom, a SQL query window shows the query `select * from stgseattle;` and a **Result Grid** displaying the data from the `stgseattle` table. The result grid includes columns for License\_Issue\_Date, License\_Number, Animal\_s\_Name, Species, Primary\_Breed, Secondary\_Breed, ZIP\_Code, DI\_CreatedJob, and DI\_ProcessID.

License_Issue_Date	License_Number	Animal_s_Name	Species	Primary_Breed	Secondary_Breed	ZIP_Code	DI_CreatedJob	DI_ProcessID
2015-12-18 00:00:00	S107948	Zen	Cat	Domestic Longhair	Mix	98117	2023-10-23 12:51:02	BHU0D6
2016-06-14 00:00:00	S116503	Misty	Cat	Siberian		98117	2023-10-23 12:51:02	BHU0D6
2016-08-04 00:00:00	S119301	Lyra	Cat	Mix		98121	2023-10-23 12:51:02	BHU0D6
2019-01-27 00:00:00	8005097	Jolene	Cat	Maine Coon	Mix	98133	2023-10-23 12:51:02	BHU0D6
2019-02-13 00:00:00	962273	Veronica	Cat	Domestic Longhair		98107	2023-10-23 12:51:02	BHU0D6
2019-06-01 00:00:00	208746	Sweetheart	Cat	Domestic Medium Hair		98116	2023-10-23 12:51:02	BHU0D6
2019-06-06 00:00:00	79347	Mr Darcy	Cat	Domestic Shorthair		98103	2023-10-23 12:51:02	BHU0D6
2019-06-25 00:00:00	8007918	Kali	Cat	Domestic Shorthair		98133	2023-10-23 12:51:02	BHU0D6
2019-07-04 00:00:00	209285	Daisy	Cat	Domestic Shorthair		98117	2023-10-23 12:51:02	BHU0D6
2019-07-30 00:00:00	S132137	Ada	Cat	American Curl	Mix	98115	2023-10-23 12:51:02	BHU0D6
2019-08-10 00:00:00	S133113	Spider	Cat	LaPerm		98115	2023-10-23 12:51:02	BHU0D6
2019-08-16 00:00:00	S130442	Fog	Cat	Domestic Medium Hair	Mix	98106	2023-10-23 12:51:02	BHU0D6
2019-08-23 00:00:00	S133417	Alicia Keys	Cat	Domestic Longhair		98133	2023-10-23 12:51:02	BHU0D6

## Dimensions:

## Breed:

The screenshot displays a data integration tool interface. At the top, a job execution flow is visualized with the following components and data:

- stg\_seattlepetlicense**: 43086 rows in 0.87s
- tAggregateRow\_1**: 49353.95 rows/s
- tMap\_2**: 3351 rows in 1.18s
- Breed\_Dim\_MySQL (Main)**: 3351 rows in 1.23s
- mysql**: 2715.56 rows/s

Below the flow, the **Job Breed\_Dimension\_MySQL** configuration is shown. The **Basic Run** tab is active, displaying the execution log:

```
Starting job Breed_Dimension_MySQL at 19:49 23/10/2023.
[statistics] connecting to socket on port 4050
[statistics] connected
[statistics] disconnected
Job Breed_Dimension_MySQL ended at 19:49 23/10/2023. [Exit code = 0]
```

On the right, the **Default** configuration for the MySQL connection is listed:

Name	Value
mysql_Passwo...	****
mysql_Databa...	Seattle
mysql_Additi...	noDate
mysql_Port	3306
mysql_Server	127.0.0.

The screenshot shows a SQL query editor with the following query:

```
select * from dim_breed;
```

The results are displayed in a table with the following columns:

Breed_SK	Species	Primary_Breed	Secondary_Breed	DI_CreatedDate	DI_ProcessID
1	Dog	Bichon Frise	Lhasa Apso	2023-10-15 19:04:57	C60kH
2	Cat	California Spangled	Mix	2023-10-15 19:04:57	C60kH
3	Dog	Poodle, Toy	Spaniel, English Cocker	2023-10-15 19:04:57	C60kH
4	Dog	Welsh Corgi, Cardigan	Bulldog, English	2023-10-15 19:04:57	C60kH
5	Dog	Shih Tzu	Mixed Breed, Small (under 24 lbs fully grown)	2023-10-15 19:04:57	C60kH
6	Dog	Spaniel, American Cocker	Border Collie	2023-10-15 19:04:57	C60kH
7	Dog	Brittany	Retriever, Labrador	2023-10-15 19:04:57	C60kH
8	Dog	Mixed Breed, Large (over 44 lbs fully grown)	Terrier, Pit Bull	2023-10-15 19:04:57	C60kH
9	Dog	Spaniel	Mix	2023-10-15 19:04:57	C60kH
10	Cat	American Shorthair	Munchkin	2023-10-15 19:04:57	C60kH
11	Dog	Terrier, American Staffordshire	Retriever	2023-10-15 19:04:57	C60kH
12	Dog	Chihuahua, Short Coat	Terrier, Fox, Toy	2023-10-15 19:04:57	C60kH
13	Dog	Terrier	Dachshund, Standard Smooth Haired	2023-10-15 19:04:57	C60kH

Date:

Job Date\_Dimension\_MySQL

Basic Run  
Debug Run  
Advanced settings  
Target Exec  
Memory Run

Execution

Run Kill Clear

Starting job Date\_Dimension\_MySQL at 19:49 23/10/2023.  
[statistics] connecting to socket on port 3485  
[statistics] connected  
[statistics] disconnected  
Job Date\_Dimension\_MySQL ended at 19:49 23/10/2023. [Exit code = 0]

Default

Name	Value
mySQL_Passwo...	****
mySQL_Databa...	Seattle
mySQL_Additi...	noDate
mySQL_Port	3306
mySQL_Server	127.0.0.
mySQL_Login	root

Query 1 SQL File 3\* SQL File 4\* SQL File 5\* x

Limit to 1000 rows

1 • select \* from dim\_date;

Result Grid Filter Rows: Edit: Export/Imports: Wrap Cell Contents: Fetch rows:

	Date_SK	Day_Num	Day_Str	Month_Num	Month_Str	Year_Num	is_weekend	DI_CreatedDate	DI_ProcessID	Dt
▶	1012015	1	Thursday	0	January	2015	N	2023-10-23 13:00:53	5Tvw0d	2015-01-01 00:00:00
	1012016	1	Friday	0	January	2016	N	2023-10-23 13:00:54	5Tvw0d	2016-01-01 00:00:00
	1012017	1	Sunday	0	January	2017	Y	2023-10-23 13:00:54	5Tvw0d	2017-01-01 00:00:00
	1012018	1	Monday	0	January	2018	N	2023-10-23 13:00:54	5Tvw0d	2018-01-01 00:00:00
	1012019	1	Tuesday	0	January	2019	N	2023-10-23 13:00:54	5Tvw0d	2019-01-01 00:00:00
	1012020	1	Wednesday	0	January	2020	N	2023-10-23 13:00:54	5Tvw0d	2020-01-01 00:00:00
	1012021	1	Friday	0	January	2021	N	2023-10-23 13:00:55	5Tvw0d	2021-01-01 00:00:00
	1012022	1	Saturday	0	January	2022	Y	2023-10-23 13:00:55	5Tvw0d	2022-01-01 00:00:00
	1012023	1	Sunday	0	January	2023	Y	2023-10-23 13:00:55	5Tvw0d	2023-01-01 00:00:00
	1012024	1	Monday	0	January	2024	N	2023-10-23 13:00:55	5Tvw0d	2024-01-01 00:00:00
	1012025	1	Wednesday	0	January	2025	N	2023-10-23 13:00:55	5Tvw0d	2025-01-01 00:00:00
	1012026	1	Thursday	0	January	2026	N	2023-10-23 13:00:55	5Tvw0d	2026-01-01 00:00:00
	1012027	1	Friday	0	January	2027	N	2023-10-23 13:00:55	5Tvw0d	2027-01-01 00:00:00

dim\_date 8 x

Apply Revert



## Location

The screenshot shows a data integration tool interface with a job named "Job Location\_Dimension\_MySQL". The job is in the "Run" state, and the execution log shows the following details:

```
Starting job Location_Dimension_MySQL at 20:01 23/10/2023.
[statistics] connecting to socket on port 3699
[statistics] connected
[statistics] disconnected
Job Location_Dimension_MySQL ended at 20:01 23/10/2023. [Exit code = 0]
```

The job configuration shows a flow from a "geo\_zip" component to a "tMap\_1" component, and then to a "mysql" component. The flow is labeled "row1 (Main)" and "Location\_dim (Main)". The flow statistics show 33103 rows in 6.73s for the first segment and 33103 rows in 7.07s for the second segment.

The screenshot shows a SQL query editor with the following query:

```
1 • select * from dim_location;
```

The result grid displays the following data:

	Location_SK	zipcode	city	state	DI_CreatedDate	DI_ProcessDate
1	35004	Acmar	Alabama	2023-10-16 01:09:07	j4jEkl	
2	35005	Adamsville	Alabama	2023-10-16 01:09:07	j4jEkl	
3	35006	Adger	Alabama	2023-10-16 01:09:07	j4jEkl	
4	35007	Keystone	Alabama	2023-10-16 01:09:07	j4jEkl	
5	35010	New site	Alabama	2023-10-16 01:09:07	j4jEkl	
6	35014	Alpine	Alabama	2023-10-16 01:09:07	j4jEkl	
7	35016	Arab	Alabama	2023-10-16 01:09:07	j4jEkl	
8	35019	Baileyton	Alabama	2023-10-16 01:09:07	j4jEkl	
9	35020	Bessemer	Alabama	2023-10-16 01:09:07	j4jEkl	
10	35022	Zcta 35022	Alabama	2023-10-16 01:09:07	j4jEkl	
11	35023	Hueytown	Alabama	2023-10-16 01:09:07	j4jEkl	
12	35031	Blountsville	Alabama	2023-10-16 01:09:07	j4jEkl	
13	35033	Bremen	Alabama	2023-10-16 01:09:07	j4jEkl	

## Fact Table MSSQL

The screenshot displays a data integration tool interface with a job configuration and its execution results.

**Job Configuration (Designer View):**

- Job:** Job Fact\_SeattlePetLicense
- Contexts:** Fact\_SeattlePetLicense
- Components:** Run (Job Fact\_SeattlePetLicense)
- Cloud Artifact:** Default

**Job Fact\_SeattlePetLicense Execution Details:**

- Basic Run:** Run, Kill, Clear buttons.
- Advanced settings:** Target Exec, Memory Run.
- Execution Log:**

```
tLogRow_2
-----
License_Issue_Date|License_Number
-----
[statistics] disconnected
Job Fact_SeattlePetLicense ended at 20:02 23/10/2023. [Exit code = 0]
```

**SQL Query:**

```
1 • select * from fct_seattle;
```

**Result Grid:**

	PetLicenseSK	Date_SK	Location_SK	Breed_SK	License_Number
1	1	18122015	30842	1090	S107948
2	2	14062016	30842	2572	S116503
3	3	4082016	30845	2904	S119301
4	4	27012019	30849	1246	8005097
5	5	13022019	30835	2981	962273
6	6	1062019	30841	2638	208746
7	7	6062019	30831	406	79347
8	8	25062019	30849	406	8007918
9	9	4072019	30842	406	209285
10	10	30072019	30840	1262	S132137
11	11	10082019	30840	1587	S133113
12	12	16082019	30834	1165	S130442
13	13	23082019	30849	2981	S133417

## Azure Staging Table:

The screenshot displays the SSDT interface with a data flow diagram and a job execution window.

**Data Flow Diagram:**

- SeattlePetLicense** (Source) → **tMap\_1** (Transformation): 43086 rows in 10.75s, 4009.12 rows/s, row1 (Main).
- tMap\_1** → **LogRow\_1** (Destination): 43086 rows in 11.72s, 3675.65 rows/s, LogRow (Main order:1).
- tMap\_1** → **Azure** (Destination): 43086 rows in 12.26s, 3575.22 rows/s, AzureDS (Main order:2).

**Job SeattlePetLicense\_Staging**

**Basic Run**

License_Issue_Date	License_Number	Animal_s_Name	Species	Primary_Breed	Secondary_Breed	ZIP_Code	DI_CreatedDate	DI_JobID
07-04-2023	8017262	Theo	Cat	Domestic Longhair	Mix	98117	2023-10-23 19:46:40.257	c3huGQ
07-04-2023	8045573	Bronny	Cat	Siberian		98117	2023-10-23 19:46:40.267	c3huGQ
07-04-2023	8045574	Red	Cat	Mix		98121	2023-10-23 19:46:40.267	c3huGQ
07-04-2023	8045577	Linus	Cat	Maine Coon	Mix	98133	2023-10-23 19:46:40.267	c3huGQ

[statistics] disconnected

Job SeattlePetLicense\_Staging ended at 19:46 23/10/2023. [Exit code = 0]

The screenshot displays the SSDT interface with a query execution window.

**Query:**

```
select * from [server\datalog7370].[dbo].[SeattlePetLicense];
```

**Results:**

License_Issue_Date	License_Number	Animal_s_Name	Species	Primary_Breed	Secondary_Breed	ZIP_Code	DI_CreatedDate	DI_JobID
2015-12-18 00:00:00.000	5107948	Zen	Cat	Domestic Longhair	Mix	98117	2023-10-23 19:46:40.257	c3huGQ
2016-06-14 00:00:00.000	5116503	Misty	Cat	Siberian		98117	2023-10-23 19:46:40.267	c3huGQ
2016-08-04 00:00:00.000	5119301	Lyra	Cat	Mix		98121	2023-10-23 19:46:40.267	c3huGQ
2019-01-27 00:00:00.000	8005097	Jolene	Cat	Maine Coon	Mix	98133	2023-10-23 19:46:40.267	c3huGQ
2019-02-13 00:00:00.000	962273	Veronica	Cat	Domestic Longhair		98107	2023-10-23 19:46:40.267	c3huGQ
2019-06-01 00:00:00.000	208746	Sweetheart	Cat	Domestic Medium Hair		98116	2023-10-23 19:46:40.267	c3huGQ
2019-06-06 00:00:00.000	79347	Mr Darcy	Cat	Domestic Shorthair		98103	2023-10-23 19:46:40.267	c3huGQ
2019-06-25 00:00:00.000	8007918	Kali	Cat	Domestic Shorthair		98133	2023-10-23 19:46:40.267	c3huGQ
2019-07-04 00:00:00.000	209285	Daisy	Cat	Domestic Shorthair		98117	2023-10-23 19:46:40.267	c3huGQ
2019-07-30 00:00:00.000	5132137	Ada	Cat	American Curl	Mix	98115	2023-10-23 19:46:40.267	c3huGQ
2019-08-10 00:00:00.000	5133113	Spider	Cat	LaPerm		98115	2023-10-23 19:46:40.267	c3huGQ
2019-08-16 00:00:00.000	5130442	Fog	Cat	Domestic Medium Hair	Mix	98106	2023-10-23 19:46:40.267	c3huGQ
2019-08-23 00:00:00.000	5133417	Alicia Keys	Cat	Domestic Longhair		98133	2023-10-23 19:46:40.267	c3huGQ

## Dimensions:

## Breed

The diagram illustrates a data pipeline with the following components and performance metrics:

- stg\_SeattlePetLicense**: 43086 rows in 1.5s
- tAggregateRow\_1**: 28685.75 rows/s (row1 (Main)), 3351 rows in 225.09s
- tMap\_1**: 14.89 rows/s (row2 (Main)), 3351 rows in 225.53s
- AzureDB (Main)**: 3351 rows in 225.53s
- Azure**: Final destination

**Job Breed\_Dim**

**Basic Run**

Debug Run

Advanced settings

Target Exec

Memory Run

Starting job Breed\_Dim at 21:27 23/10/2023.  
[statistics] connecting to socket on port 3951  
[statistics] connected  
[statistics] disconnected  
Job Breed\_Dim ended at 21:30 23/10/2023. [Exit code = 0]

Line limit 100000

Default

Name	Value
Azure_Schema	
Azure_Database	damg7.
Azure_Login	srivasta
Azure_Server	serverd
Azure_Additio...	
Azure_Password	****

SQLQuery\_11 - serverdamg7370.database.windows.net.damg7370 (srivastavamu) - Azure Data Studio

serverdamg7370.database.windows.net.damg7370 SQLQuery\_12 - server...tavamu SQLQuery\_11 - server...tavamu SQLQuery\_12 - server...tavamu

Run Cancel Disconnect Change Connection damg7370 Estimated Plan Enable Actual Plan Parse Enable SQLCMD

1 select \* from [srivastavamu].[Dim\_Breed1]

Results Messages

	Breed_SK	Species	Primary_Breed	Secondary_Breed	DI_CreatedDate	DI_JobID
1	1	Dog	Bichon Frise	Lhasa Apso	2023-10-23 21:27:11.903	ovKeGr
2	2	Cat	California Spangled	Mix	2023-10-23 21:27:12.027	ovKeGr
3	3	Dog	Poodle, Toy	Spaniel, English Cocker	2023-10-23 21:27:12.167	ovKeGr
4	4	Dog	Welsh Corgi, Cardigan	Bulldog, English	2023-10-23 21:27:12.317	ovKeGr
5	5	Dog	Shih Tzu	Mixed Breed, Small (under 24...	2023-10-23 21:27:12.447	ovKeGr
6	6	Dog	Spaniel, American Cocker	Border Collie	2023-10-23 21:27:12.587	ovKeGr
7	7	Dog	Brittany	Retriever, Labrador	2023-10-23 21:27:12.643	ovKeGr
8	8	Dog	Mixed Breed, Large (over 44 ...	Terrier, Pit Bull	2023-10-23 21:27:12.857	ovKeGr
9	9	Dog	Spaniel	Mix	2023-10-23 21:27:13.000	ovKeGr
10	10	Cat	American Shorthair	Munchkin	2023-10-23 21:27:13.133	ovKeGr
11	11	Dog	Terrier, American Staffordsh...	Retriever	2023-10-23 21:27:13.263	ovKeGr
12	12	Dog	Chihuahua, Short Coat	Terrier, Fox, Toy	2023-10-23 21:27:13.413	ovKeGr
13	13	Dog	Terrier	Dachshund, Standard Smooth H...	2023-10-23 21:27:13.553	ovKeGr
14	14	Dog	Dachshund, Standard Long Hai...	Pomeranian	2023-10-23 21:27:13.637	ovKeGr
15	15	Dog	Dachshund, Miniature Smooth ...	Dachshund, Standard Smooth H...	2023-10-23 21:27:13.773	ovKeGr
16	16	Dog	Spaniel, English Cocker	Retriever, Flat-Coated	2023-10-23 21:27:13.920	ovKeGr

## Location:

Job Location\_Dim

Basic Run  
Debug Run  
Advanced settings  
Target Exec  
Memory Run

Starting job Location\_Dim at 20:01 23/10/2023.  
[statistics] connecting to socket on port 3429  
[statistics] connected  
[statistics] disconnected  
Job Location\_Dim ended at 20:36 23/10/2023. [Exit code = 0]

Default

Name	Value
GeoData_Row...	"\n"
GeoData_Field...	","
GeoData_File	"D:\DD..."
GeoData_Head...	1
GeoData_Enco...	US-ASC
Azure_Database	damg7...

Run Cancel Disconnect Change Connection damg7370 Estimated Plan End

```
1 select * from [srivastavamu].[Dim_Location];
```

Results Messages

	Location_SK	Zip	City	State	Date_To_Warehouse	DI_JobID
1	1	35004	Acmar	Alabama	2023-10-23 20:02:02.450	18fcWM
2	2	35005	Adamsville	Alabama	2023-10-23 20:02:02.597	18fcWM
3	3	35006	Adger	Alabama	2023-10-23 20:02:02.737	18fcWM
4	4	35007	Keystone	Alabama	2023-10-23 20:02:02.870	18fcWM
5	5	35010	New site	Alabama	2023-10-23 20:02:03.017	18fcWM
6	6	35014	Alpine	Alabama	2023-10-23 20:02:03.153	18fcWM
7	7	35016	Arab	Alabama	2023-10-23 20:02:03.397	18fcWM
8	8	35019	Baileyton	Alabama	2023-10-23 20:02:03.543	18fcWM
9	9	35020	Bessemer	Alabama	2023-10-23 20:02:03.677	18fcWM
10	10	35022	Zcta 35022	Alabama	2023-10-23 20:02:03.813	18fcWM
11	11	35023	Hueytown	Alabama	2023-10-23 20:02:03.873	18fcWM
12	12	35031	Blountsville	Alabama	2023-10-23 20:02:03.987	18fcWM
13	13	35033	Bremen	Alabama	2023-10-23 20:02:04.163	18fcWM
14	14	35034	Brent	Alabama	2023-10-23 20:02:04.297	18fcWM
15	15	35035	Brierfield	Alabama	2023-10-23 20:02:04.407	18fcWM
16	16	35036	Brookside	Alabama	2023-10-23 20:02:04.463	18fcWM
17	17	35040	Calera	Alabama	2023-10-23 20:02:04.580	18fcWM

## Date:

The screenshot shows a data integration tool interface. The top bar displays several job names: Job Seattle..., Job Fct\_Se..., Job Breed..., Job Seattle..., Job Locatio..., Job Dat..., Job Locatio..., and Job Fact\_S... The central canvas shows a data flow diagram with the following components and statistics:

- tRowGenerator\_1: 10000 rows in 613.39s
- tMap\_1: 10000 rows in 614.14s
- AzureDB (Main): 10000 rows in 614.14s
- Azure: 10000 rows in 614.14s

The bottom panel shows the job details for 'Job Date\_Dim'. The job is running and has completed. The output shows the following statistics:

- Starting job Date\_Dim at 19:52 23/10/2023.
- [statistics] connecting to socket on port 3444
- [statistics] connected
- [statistics] disconnected
- Job Date\_Dim ended at 20:02 23/10/2023. [Exit code = 0]

The configuration table on the right shows the following values:

Name	Value
Azure_Schema	
Azure_Database	damg7.
Azure_Login	srivasta
Azure_Server	serverd
Azure_Additio...	
Azure_Password	****

The screenshot shows a SQL query editor with the following query:

```
select * from [srivastavamu].[Dim_Date];
```

The query is executed, and the results are displayed in a table. The table has the following columns: Date\_Sk, Day\_Num, Day\_Str, Month\_Num, Month\_Str, Year\_Num, Is\_Weekend, Date\_to\_Warehouse, DI\_JobID, and dt. The results show data for the years 2017 through 2029.

	Date_Sk	Day_Num	Day_Str	Month_Num	Month_Str	Year_Num	Is_Weekend	Date_to_Warehouse	DI_JobID	dt
1	1012017	1	Sunday	0	January	2017	Y	2023-10-23 19:52:45.600	ieIudn	2017-01-01 00:00:00
2	1012018	1	Monday	0	January	2018	N	2023-10-23 19:53:07.647	ieIudn	2018-01-01 00:00:00
3	1012019	1	Tuesday	0	January	2019	N	2023-10-23 19:53:30.373	ieIudn	2019-01-01 00:00:00
4	1012020	1	Wednesday	0	January	2020	N	2023-10-23 19:53:55.027	ieIudn	2020-01-01 00:00:00
5	1012021	1	Friday	0	January	2021	N	2023-10-23 19:54:16.883	ieIudn	2021-01-01 00:00:00
6	1012022	1	Saturday	0	January	2022	Y	2023-10-23 19:54:42.340	ieIudn	2022-01-01 00:00:00
7	1012023	1	Sunday	0	January	2023	Y	2023-10-23 19:55:08.617	ieIudn	2023-01-01 00:00:00
8	1012024	1	Monday	0	January	2024	N	2023-10-23 19:55:30.230	ieIudn	2024-01-01 00:00:00
9	1012025	1	Wednesday	0	January	2025	N	2023-10-23 19:55:56.137	ieIudn	2025-01-01 00:00:00
10	1012026	1	Thursday	0	January	2026	N	2023-10-23 19:56:20.340	ieIudn	2026-01-01 00:00:00
11	1012027	1	Friday	0	January	2027	N	2023-10-23 19:56:45.853	ieIudn	2027-01-01 00:00:00
12	1012028	1	Saturday	0	January	2028	Y	2023-10-23 19:57:07.240	ieIudn	2028-01-01 00:00:00
13	1012029	1	Monday	0	January	2029	N	2023-10-23 19:57:27.990	ieIudn	2029-01-01 00:00:00

## Fact Table Azure:

The screenshot displays the SQL Server Data Tools (SSDT) interface for a job named "Job Fct\_SeattlePetLicense". The top pane shows a data flow diagram with the following components and statistics:

- Dim\_Breed**: 3351 rows in 0.72s, 4667.13 rows/s, row4 (lookup)
- Dim\_Location**: 33103 rows in 0.54s, 61188.54 rows/s, row2 (lookup)
- Dim\_Date**: 28571.43 rows/s, row3 (lookup)
- stg\_SeattlePetLicense**: 43086 rows in 2722.01s, 15.83 rows/s, row1 (Main)
- tMap\_1**: 43086 rows in 2722.01s, 15.83 rows/s, row1 (Main)
- tLogRow\_2**: 0 rows in 2722.06s, 0 rows/s, OutReject (Main order:2)
- Azure**: 43086 rows in 2722.99s, 15.82 rows/s, OutputAzure (Main order:1)

The bottom pane shows the execution log for the job "Job Fct\_SeattlePetLicense 0.1". The log indicates that the job started at 21:41 on 23/10/2023 and ended at 22:27 on the same date with an exit code of 0.

**Job Fct\_SeattlePetLicense**

Execution

Run Kill Clear

Starting job Fct\_SeattlePetLicense at 21:41 23/10/2023.  
[statistics] connecting to socket on port 3536  
[statistics] connected  
[statistics] disconnected  
Job Fct\_SeattlePetLicense ended at 22:27 23/10/2023. [Exit code = 0]

Default

Name	Value
Azure_Schema	
Azure_Database	damg7.
Azure_Login	srivasta
Azure_Server	serverd
Azure_Additio...	
Azure_Password	****

serverdamg7370.database.windows.net:damg7370
SQLQuery\_12 - server...tavamu
SQLQuery\_11

Run
Cancel
Disconnect
Change Connection
damg7370
Estimated Plan
Enable Ac

1
select \* from [srivastavamu].[fct\_seattlepetlicense];

Results		Messages				
	PetLicense_Sk	Date_Sk	Location_SK	Breed_SK	License_Number	
14	14	1092019	30840	1587	5133114	
15	15	6092019	30839	406	434326	
16	16	11102019	30841	406	825362	
17	17	24102019	30833	1587	346697	
18	18	24102019	30833	1587	346695	
19	19	1112019	30839	406	832211	
20	20	1112019	30835	2638	832698	
21	21	13112019	30842	2638	833084	
22	22	21112019	30847	959	8002549	
23	23	23112019	30831	1165	5104077	
24	24	3122019	30839	1354	5130444	
25	25	11122019	30834	2638	733464	
26	26	12012020	30841	3094	5151906	
27	27	30012020	30840	406	5140122	
28	28	30012020	30840	2638	5140121	
29	29	13022020	30841	406	8008951	
30	30	23022020	30833	406	732370	
31	31	23022020	30833	406	732369	