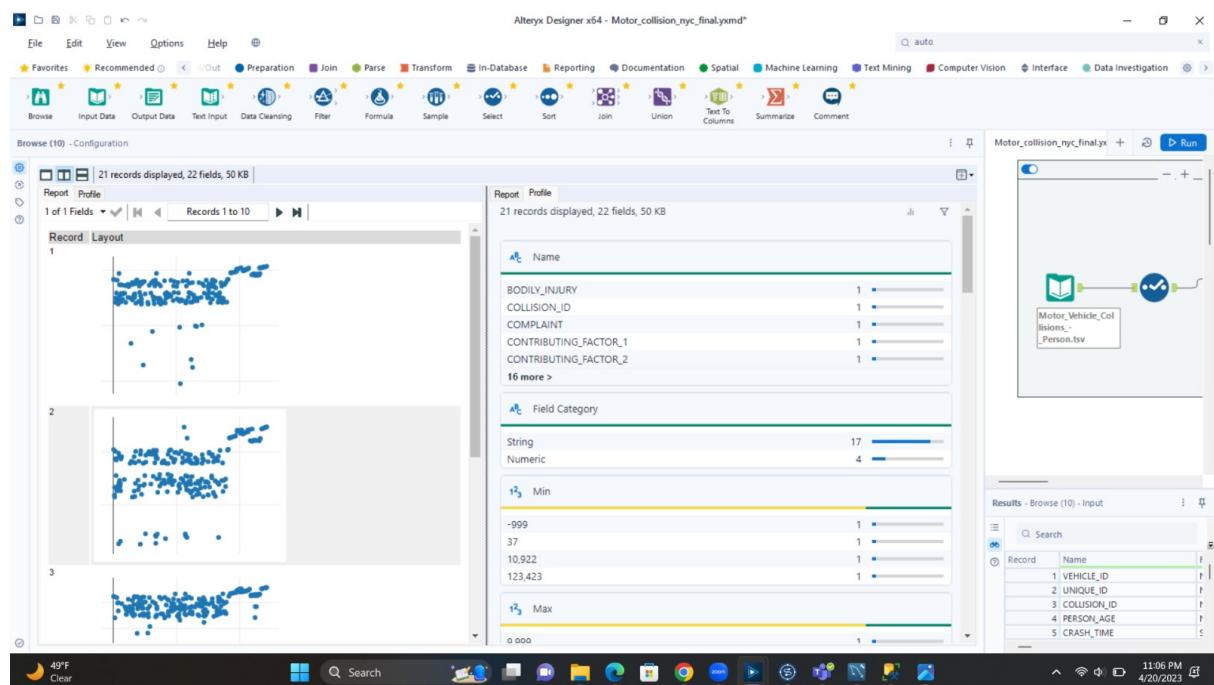
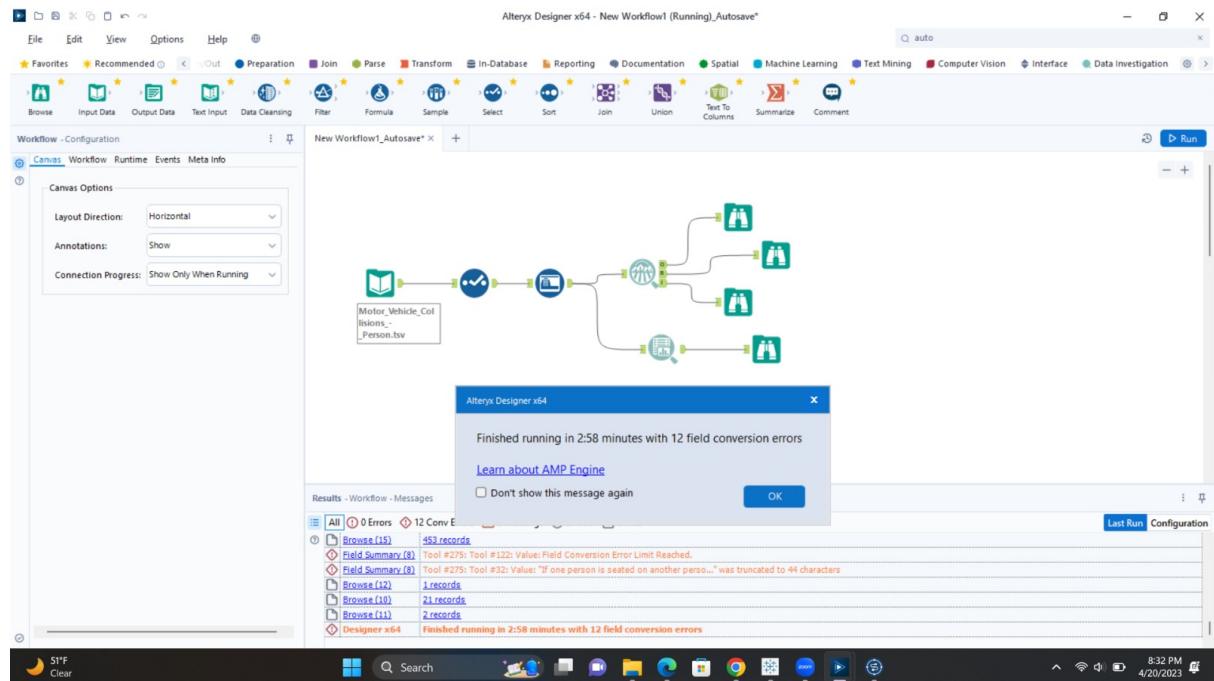


NYC Motor Vehicle collisions Part1

Data Profiling:

NYC_MV_Persons:

This job was completed in 2:58min



Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (11) - Configuration

Report Profile 1 of 1 Fields Records 1 to 2

Record Report

Numeric Fields

Name	Plot	% Missing	Unique Values	Min	Mean	P
VEHICLE_ID		4.0%	2,241,675	123,423.000	18,324,822.168	18,571.1
UNIQUE_ID		0.0%	4,863,648	10,922.000	8,666,028.176	9,082.9

Results - Browse (11) - Input

2 of 2 Fields 2 records displayed, 52 KB

Record	Desc	Report
1	Numeric	Table - View Browse Tool Report Tab
2	String	Table - View Browse Tool Report Tab

Container 36

Motor_Vehicle_Collisions_-_Person.csv

11:07 PM 4/20/2023

Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse (12) - Configuration

Report Profile 1 of 1 Fields Records 1 to 1

Field Summary

RecordText

1 VEHICLE_ID

UNIQUE_ID

Results - Browse (12) - Input

1 of 1 Fields 1 record displayed, 56 KB

Record	Text	Report
1	Text - View Browse Tool Report Tab	

Container 36

Motor_Vehicle_Collisions_-_Person.csv

11:07 PM 4/20/2023

Alteryx Designer x64 - Motor_collision_nyc_final.yxmd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (15) - Configuration

Profile 453 records displayed, 3 fields, 9,849 bytes

FieldName

COLLISION_ID	24
PERSON_AGE	24
UNIQUE_ID	24
VEHICLE_ID	24
BODILY_INJURY	21
16 more >	

Name

Data Type	21
Description	21
Maximum	21
Minimum	21
Name	21
28 more >	

Value

0	74
File: C:\DADB\Data - NYC Motor Vehicle Collisions\Motor_Vehicle_Collisions_-_Vehicles.csv	21
V_String	13

Container 36

Results - Browse (15) - Input

Record	FieldName	Name	Value
1	UNIQUE_ID	Name	UNIQUE_ID
2	UNIQUE_ID	Data Type	Int32
3	UNIQUE_ID	Size	4
4	UNIQUE_ID	Source	File: C:\DADB\Data - NYC Motor Vehicle Collisio...
5	UNIQUE_ID	Description	[Null]

49°F Clear Search 11:07 PM 4/20/2023

NYC_MV_Vehicles:

This job was completed in 2:01min

Alteryx Designer x64 - Motor_collision_nyc_final.yxmd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Workflow - Configuration

Canvas Workflow Runtime Events Meta Info

Canvas Options

- Layout Direction: Horizontal
- Annotations: Show
- Connection Progress: Show Only When Running

Motor_collision_nyc_final.yxmd*

Alteryx Designer x64

Finished running Motor_collision_nyc_final in 2:01 minutes with 22 field conversion errors

Learn about AMP Engine

OK

Results - Workflow

- All 0 Errors
- Field Summa 0 Errors
- Field Summa 0 Errors
- Field Summa 0 Errors
- Browse (20) 1 records
- Browse (18) 25 records
- Browse (19) 2 records
- Designer x64 Finished running Motor_collision_nyc_final in 2:01 minutes with 22 field conversion errors

Last Run Configuration

51°F Clear Search 8:39 PM 4/20/2023

Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (19) - Configuration

Report Profile 1 of 1 Fields Records 1 to 2

Record Report Numeric Fields

Name	Plot	% Missing	Unique Values	Min	Mean	Median	Max	Std D
VEHICLE_OCCUPANTS		46.4%	121	0.000	1,010.392	2.000	999,999,999.000	994,574.1
VEHICLE_YEAR		48.5%	300	1,000.000	2,014.806	2,013.000	20,063.000	155.1

Results - Browse (19) - Input

Record	Desc	Report
1	Numeric	Table - View-Browse Tool Report Tab
2	String	Table - View-Browse Tool Report Tab

49°F Clear 11:12 PM 4/20/2023

Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (18) - Configuration

Report Profile 1 of 1 Fields Records 1 to 10

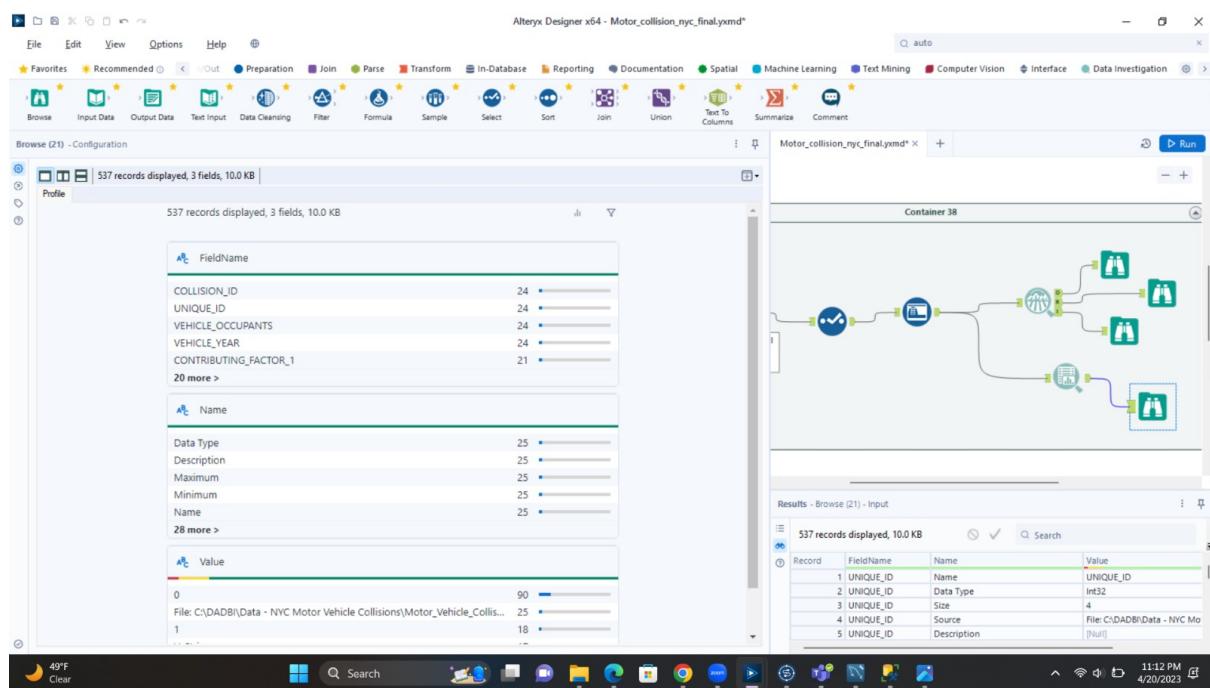
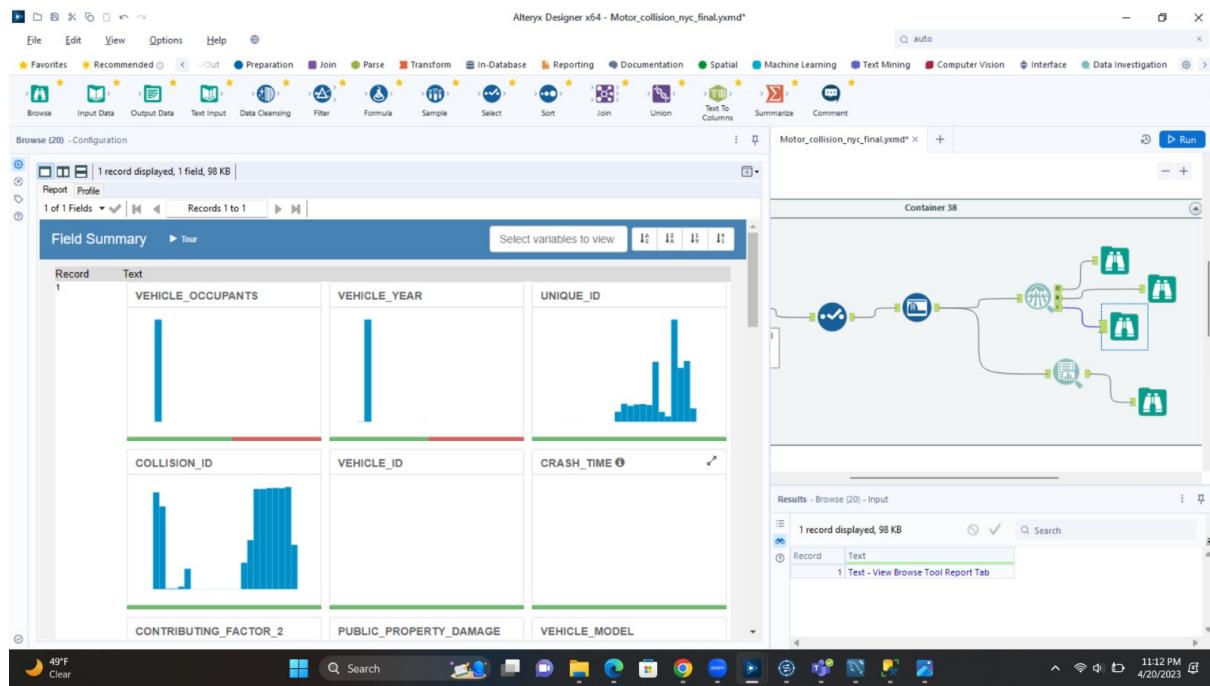
Record Layout

Record	Layout
1	
2	
3	

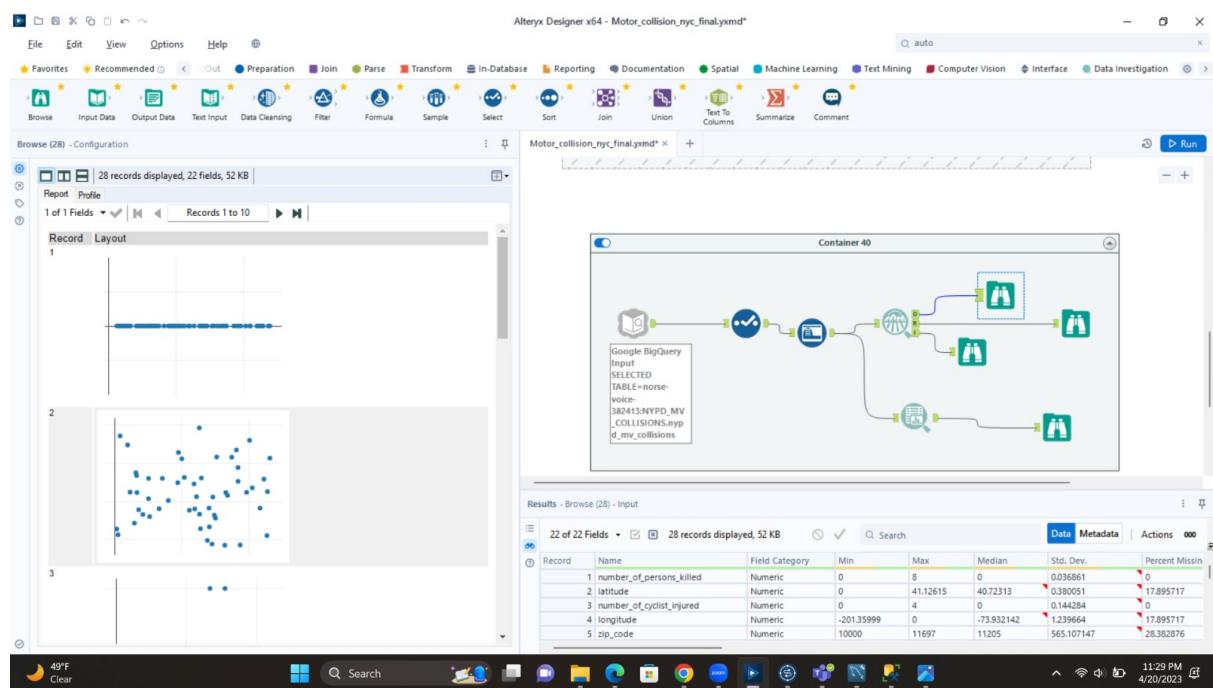
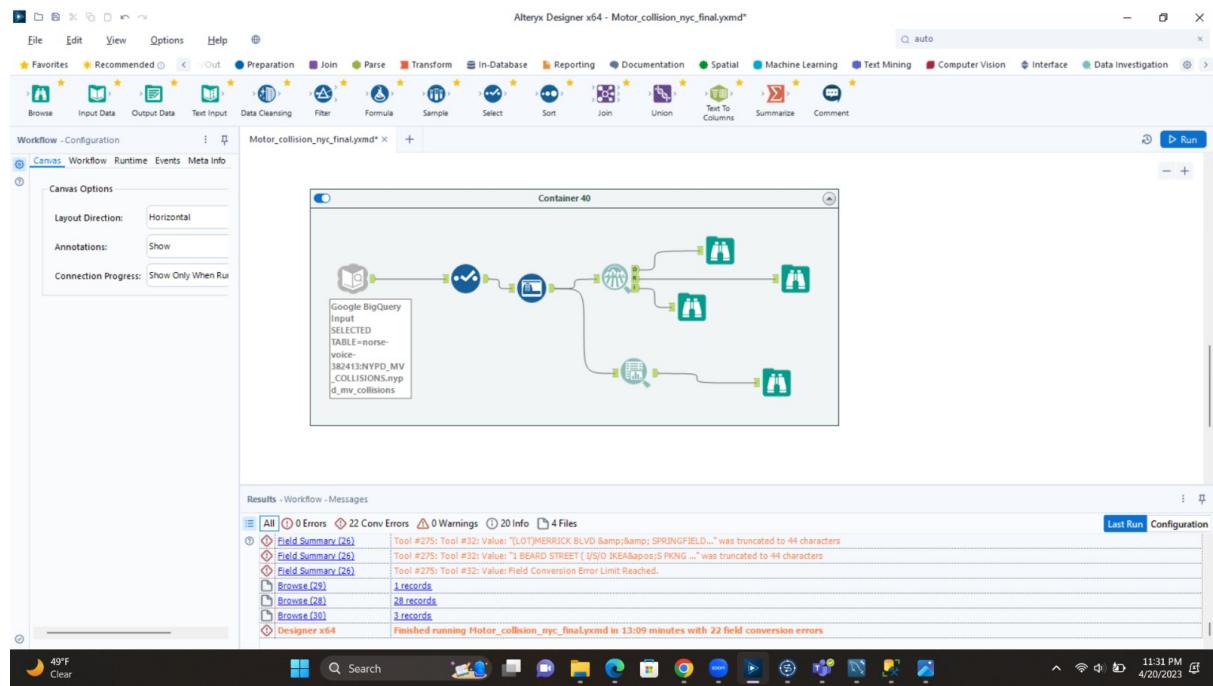
Results - Browse (18) - Input

Record	Name	Field Category	Min	Max
1	VEHICLE_OCCUPANTS	Numeric	0	99999999
2	VEHICLE_YEAR	Numeric	1000	20063
3	UNIQUE_ID	Numeric	111711	20121717
4	COLLISION_ID	Numeric	22	4484197
5	VEHICLE_ID	String	[Null]	[Null]

49°F Clear 11:12 PM 4/20/2023



NYC_MV_Bigquery:
This job was completed in 13:09min



Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (30) - Configuration

Report Profile 1 of 1 Fields Records 1 to 3

Record Report Numeric Fields

Name	Plot	% Missing	Unique Values
number_of_persons_killed		0.0%	7
latitude		17.9%	118,107

Container 40

Results - Browse (30) - Input

Record	Desc	Report
1	Numeric	Table - View Browse Tool Report Tab
2	Date	Table - View Browse Tool Report Tab
3	String	Table - View Browse Tool Report Tab

49°F Clear 11:29 PM 4/20/2023

Alteryx Designer x64 - Motor_collision_nyc_final.lymd*

File Edit View Options Help

Favorites Recommended Out Preparation Join Parse Transform In-Database Reporting Documentation Spatial Machine Learning Text Mining Computer Vision Interface Data Investigation

Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment

Browse (29) - Configuration

Report Profile 1 of 1 Fields Records 1 to 1

Field Summary Tour Select variables to view 1 1 1 1 1 1

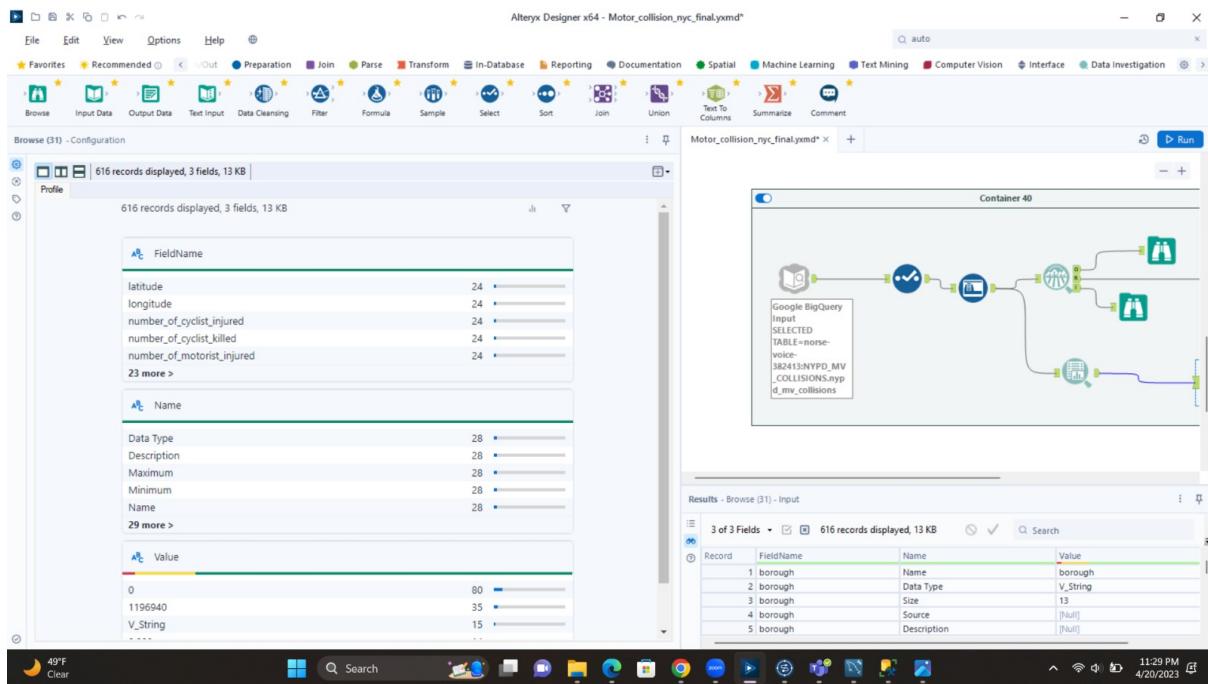
Record	Text										
1	<table border="1"> <tr> <td>number_of_pedestrians_injured</td> <td>number_of_persons_killed</td> <td>latitude</td> </tr> <tr> <td></td> <td></td> <td></td> </tr> </table> <table border="1"> <tr> <td>number_of_persons_injured</td> <td>number_of_motorist_injured</td> </tr> <tr> <td></td> <td></td> </tr> </table>	number_of_pedestrians_injured	number_of_persons_killed	latitude				number_of_persons_injured	number_of_motorist_injured		
number_of_pedestrians_injured	number_of_persons_killed	latitude									
number_of_persons_injured	number_of_motorist_injured										

Container 40

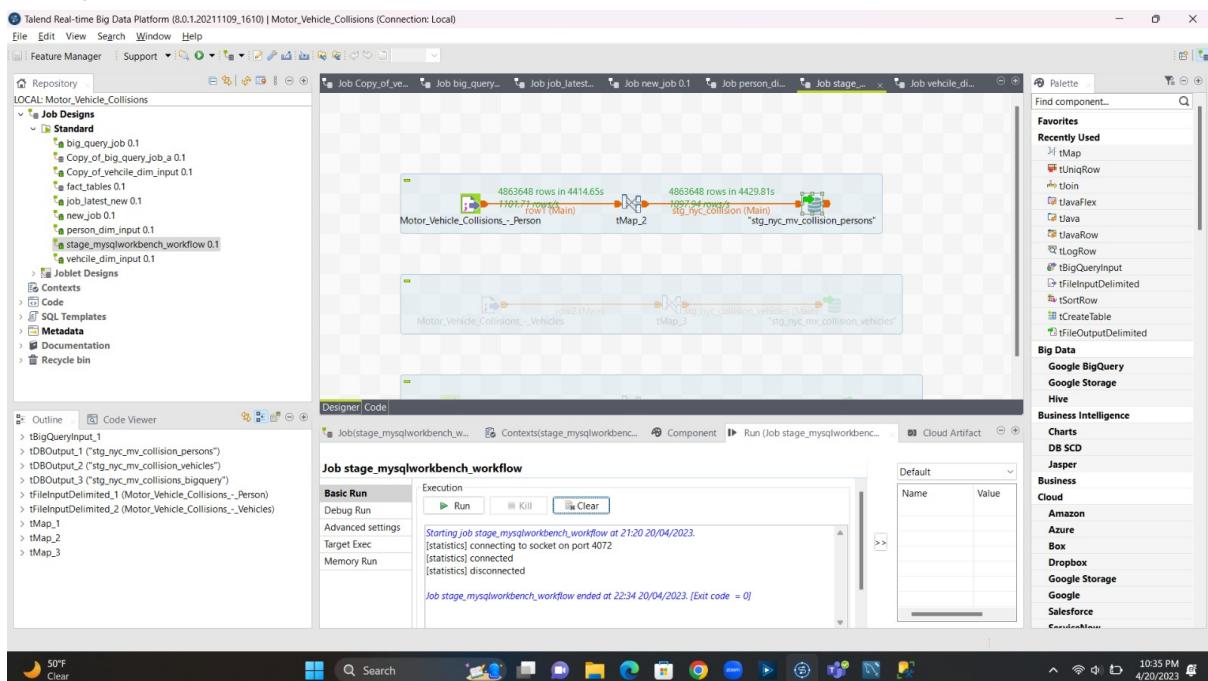
Results - Browse (29) - Input

Record	Text
1	Text - View Browse Tool Report Tab

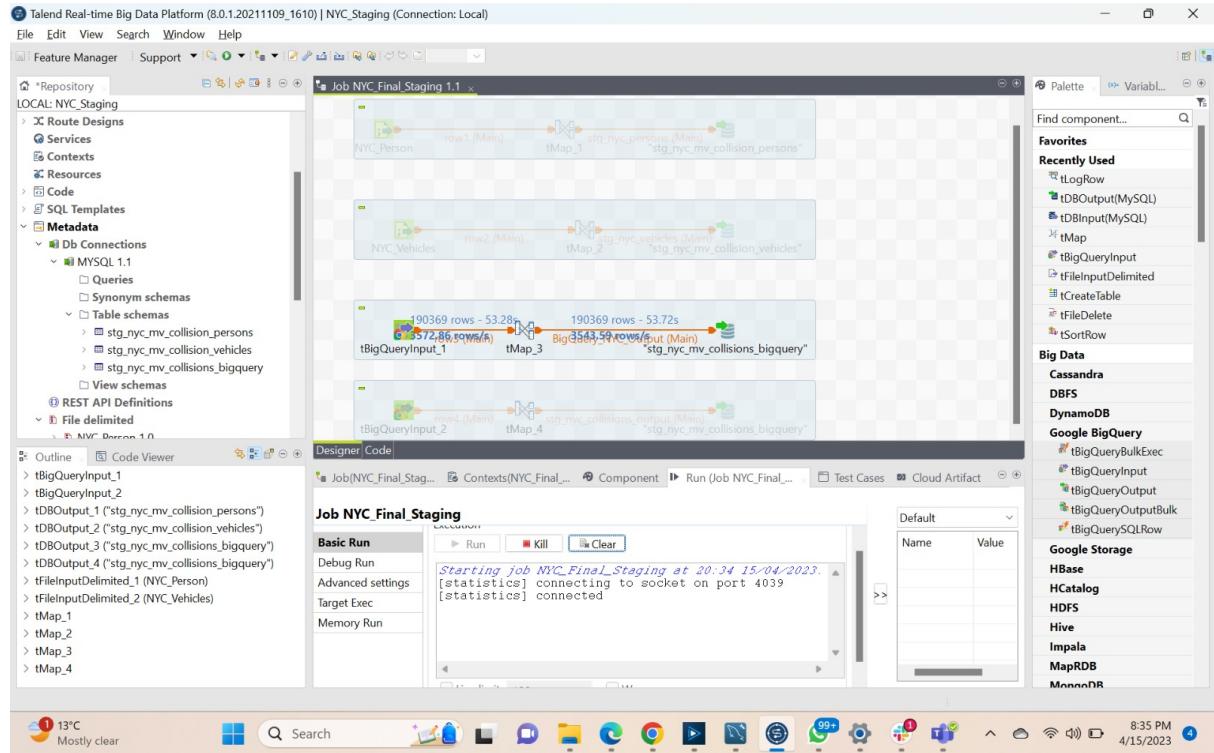
49°F Clear 11:29 PM 4/20/2023



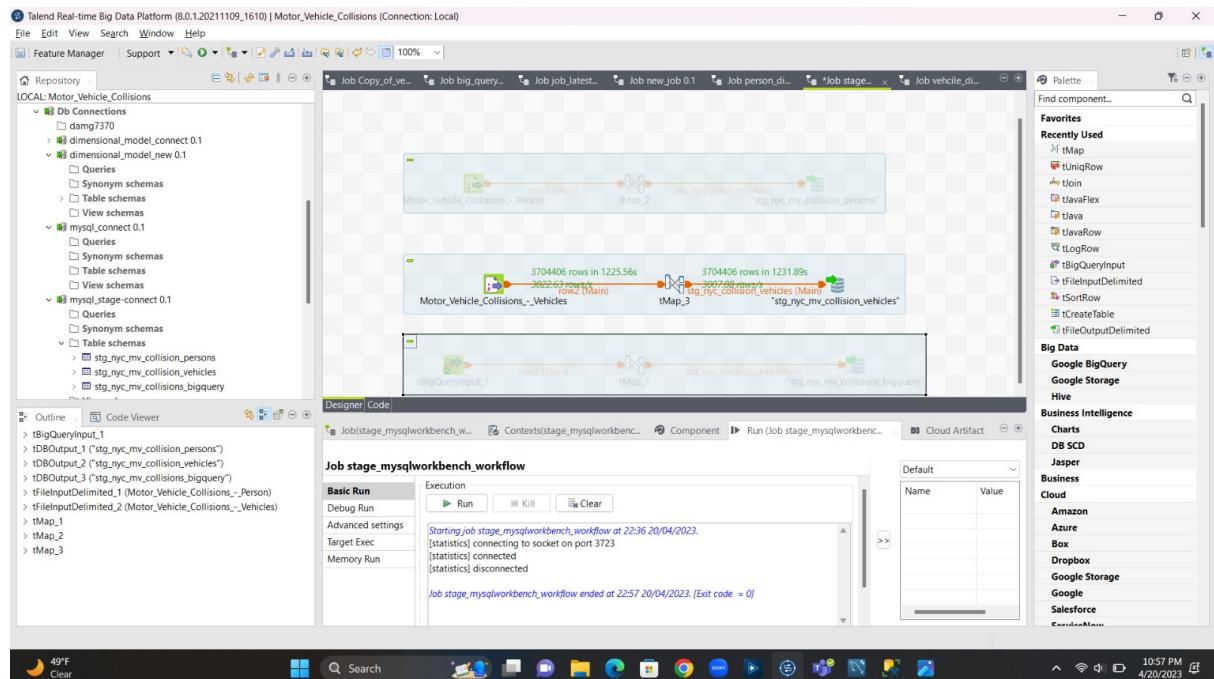
Loading data into Staging tables using Talend:
stg_MV_Collision_Persons:
This job was completed in 4429.82sec



stg_MV_Collisions_Bigquery:
This job was completed in 53.72sec



stg_MV_Collisions_Vehicles:
This job was completed in 1231.89sec



Staged Tables Row Counts:

stg_MV_Collision_Persons:

The screenshot shows the MySQL Workbench interface with the 'motor_vehicle' schema selected. The SQL editor contains the following query:

```
1 • select count(*) FROM motor_vehicle.stg_nyc_mv_collision_persons;
2 • select count(*) FROM motor_vehicle.stg_nyc_mv_collision_vehicles;
3 • SELECT count(*) FROM stg_nyc_mv_collisions_bigquery;
```

The Result Grid shows the output of the first query:

count(*)
4863648

The status bar at the bottom right indicates the session was run at 10:58 PM on 4/20/2023.

stg_MV_Collision_Vehicles:

The screenshot shows the MySQL Workbench interface with the 'motor_vehicle' schema selected. The SQL editor contains the following query:

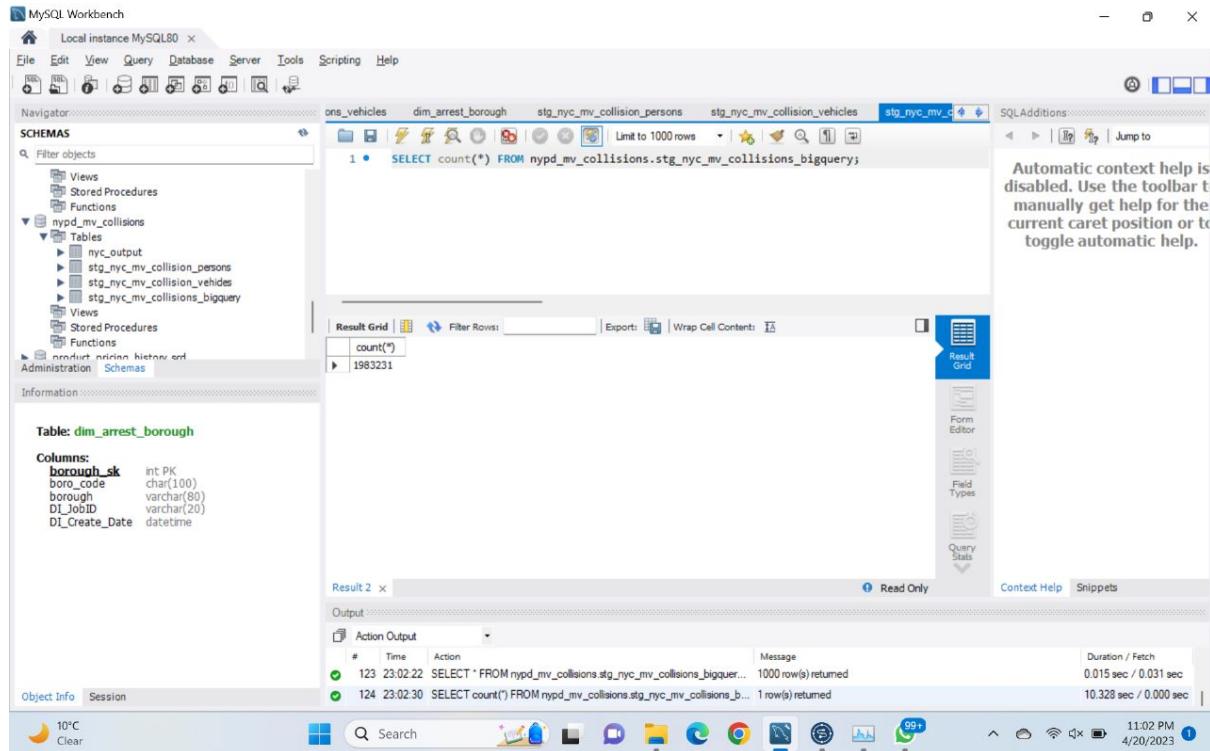
```
1 • select count(*) FROM motor_vehicle.stg_nyc_mv_collision_persons;
2 • select count(*) FROM motor_vehicle.stg_nyc_mv_collision_vehicles;
3 • SELECT count(*) FROM stg_nyc_mv_collisions_bigquery;
```

The Result Grid shows the output of the second query:

count(*)
3704406

The status bar at the bottom right indicates the session was run at 10:58 PM on 4/20/2023.

stg_MV_Collision_Bigquery:



Preliminary Dimensional model:

List of facts:

1. Fct_Vehicle_Collision
2. Fct_Collision

List of dimensions:

1. Dim_Vehicle
2. Dim_Contributing_Factor
3. Dim_Person
4. Dim_Time
5. Dim_Date
6. Dim_Borough
7. Dim_Location

Data Model:

