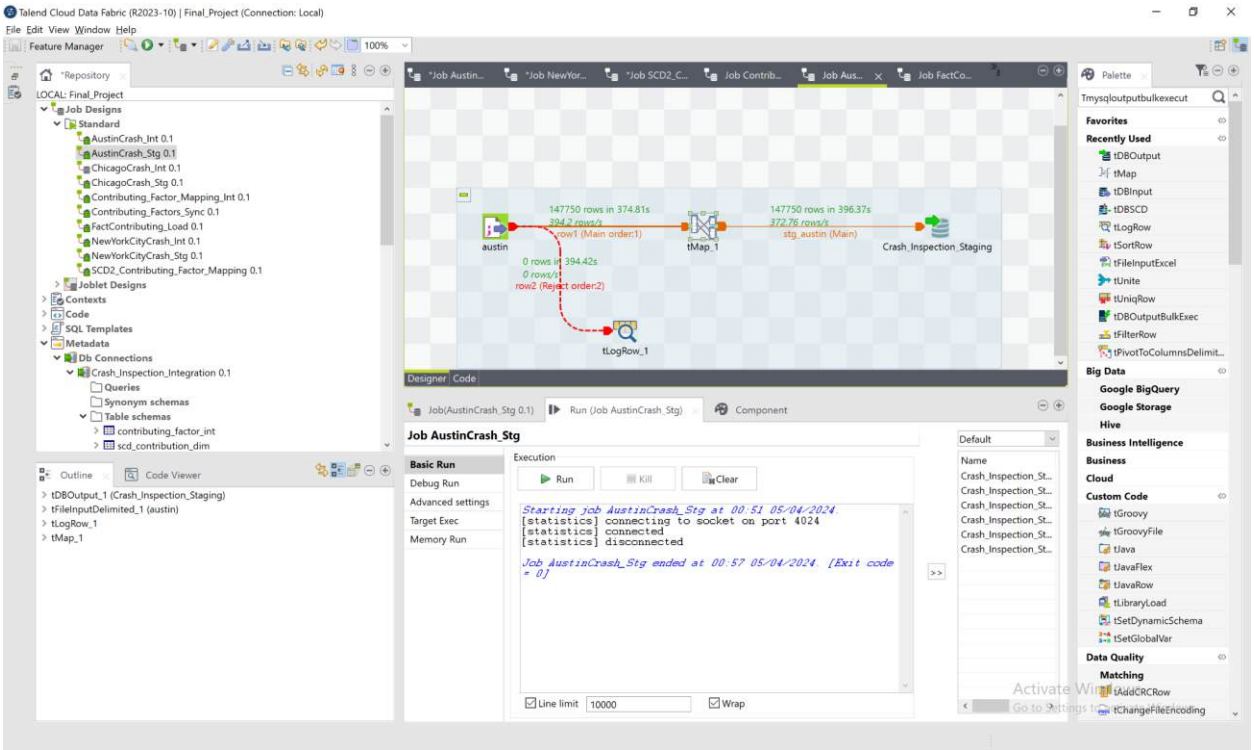
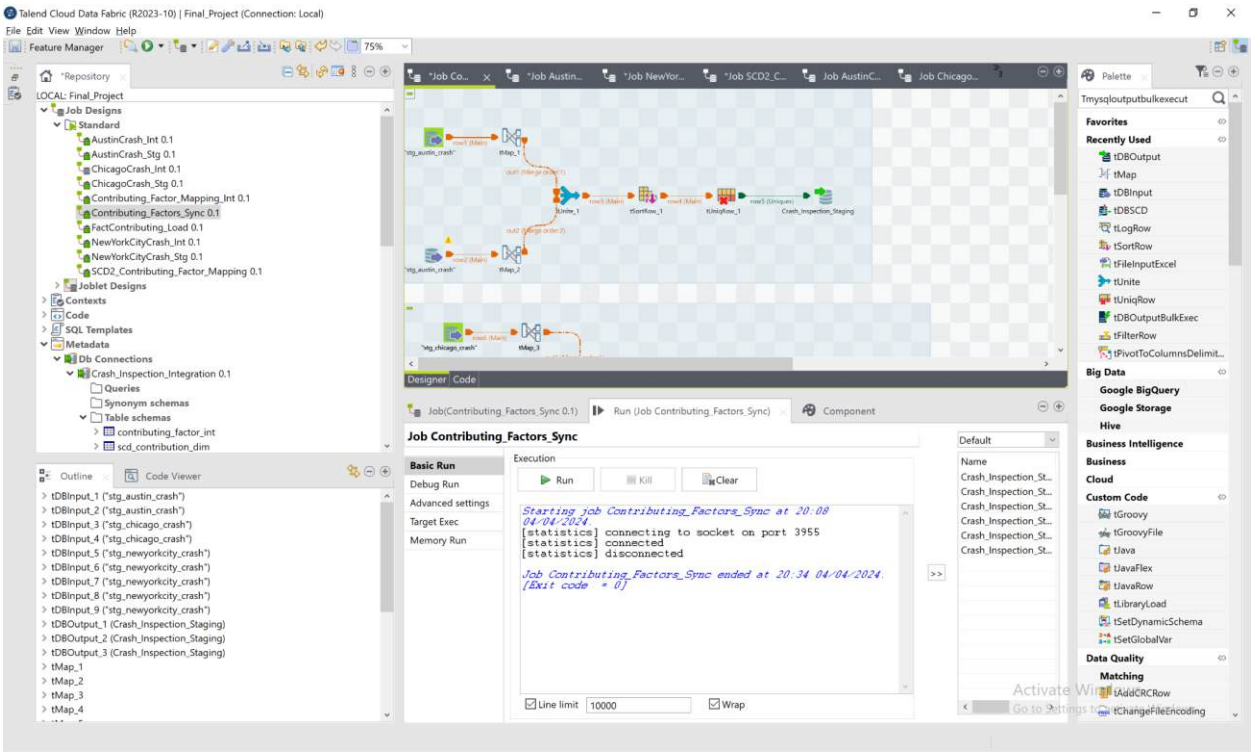
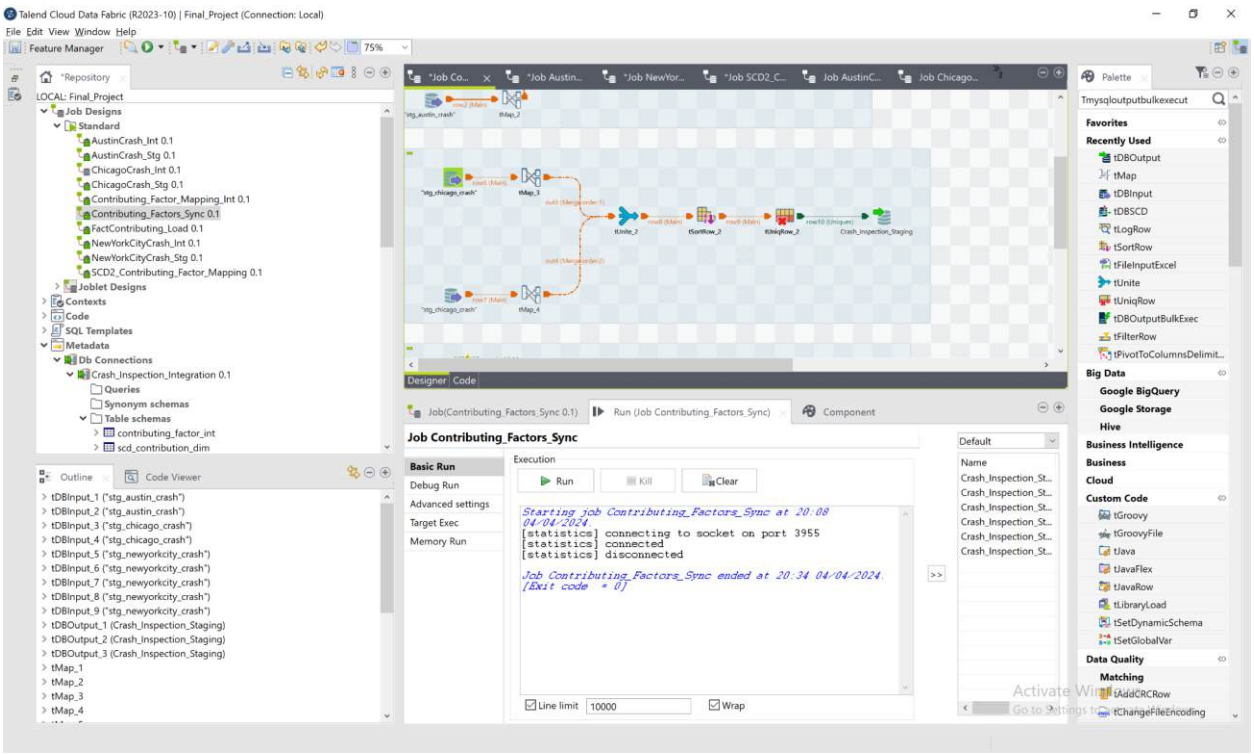
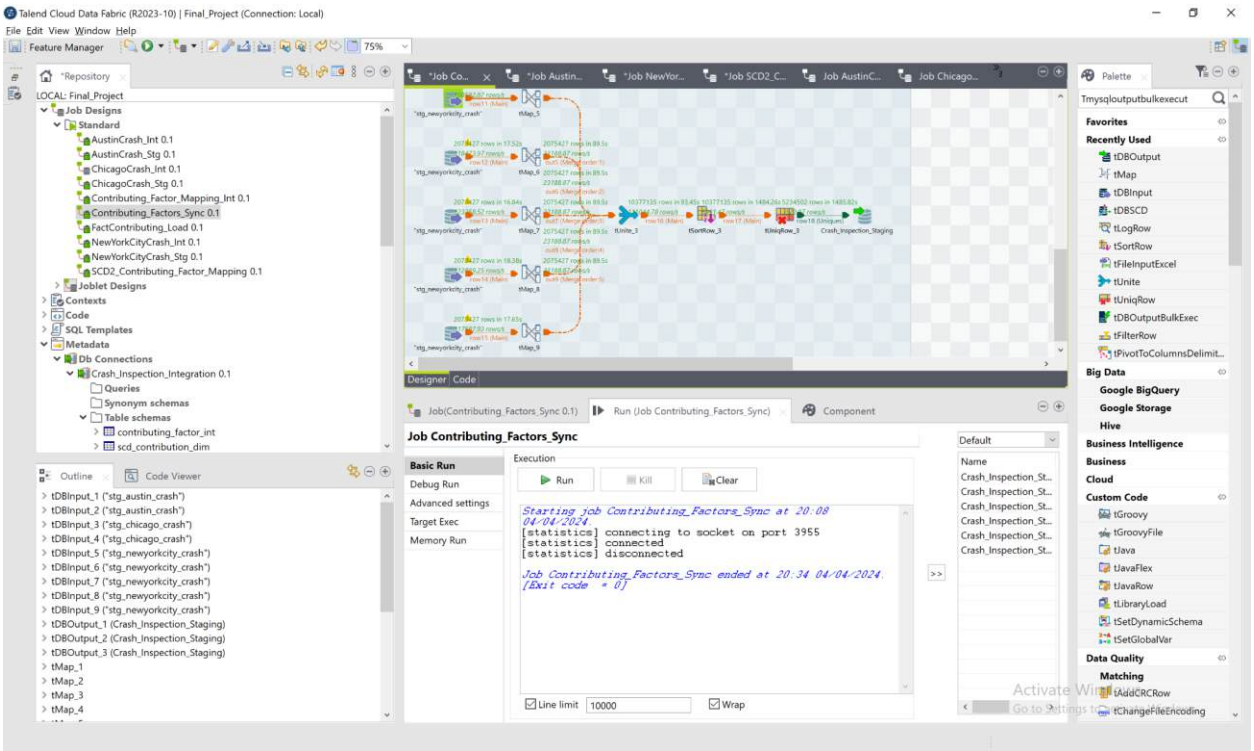
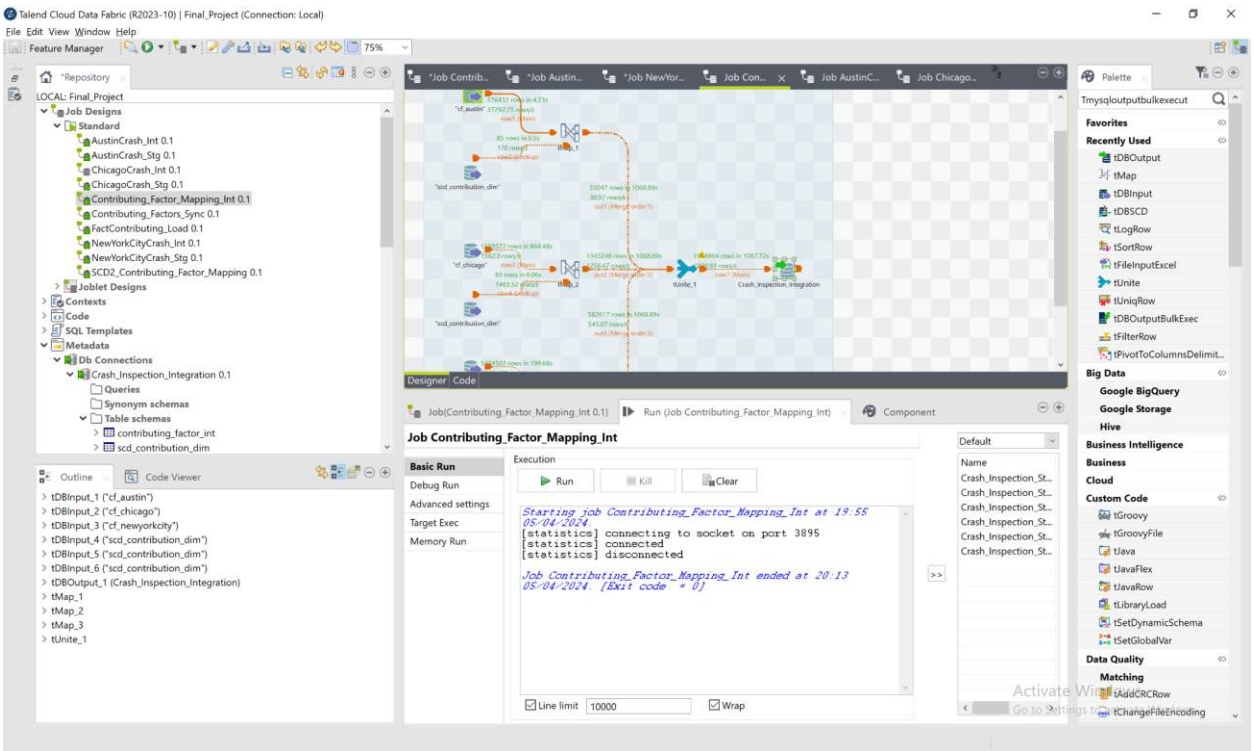
FINAL PROJECT DOCUMENTATION:

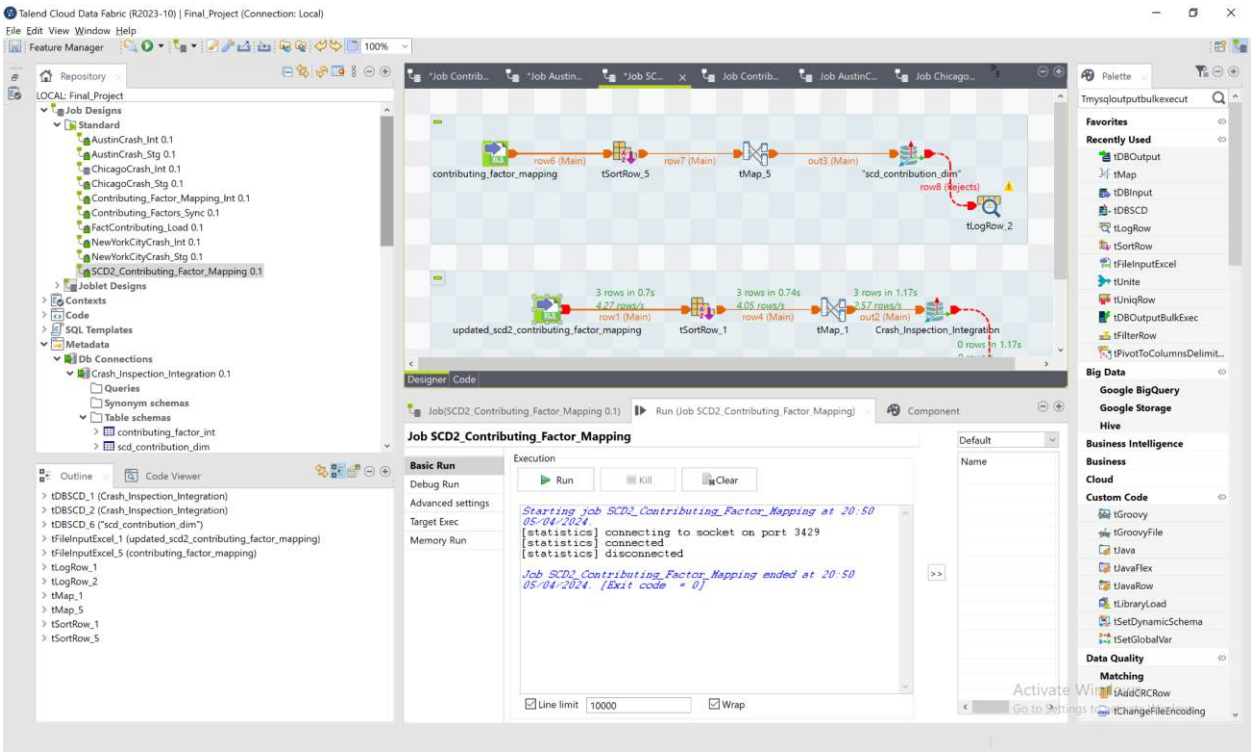


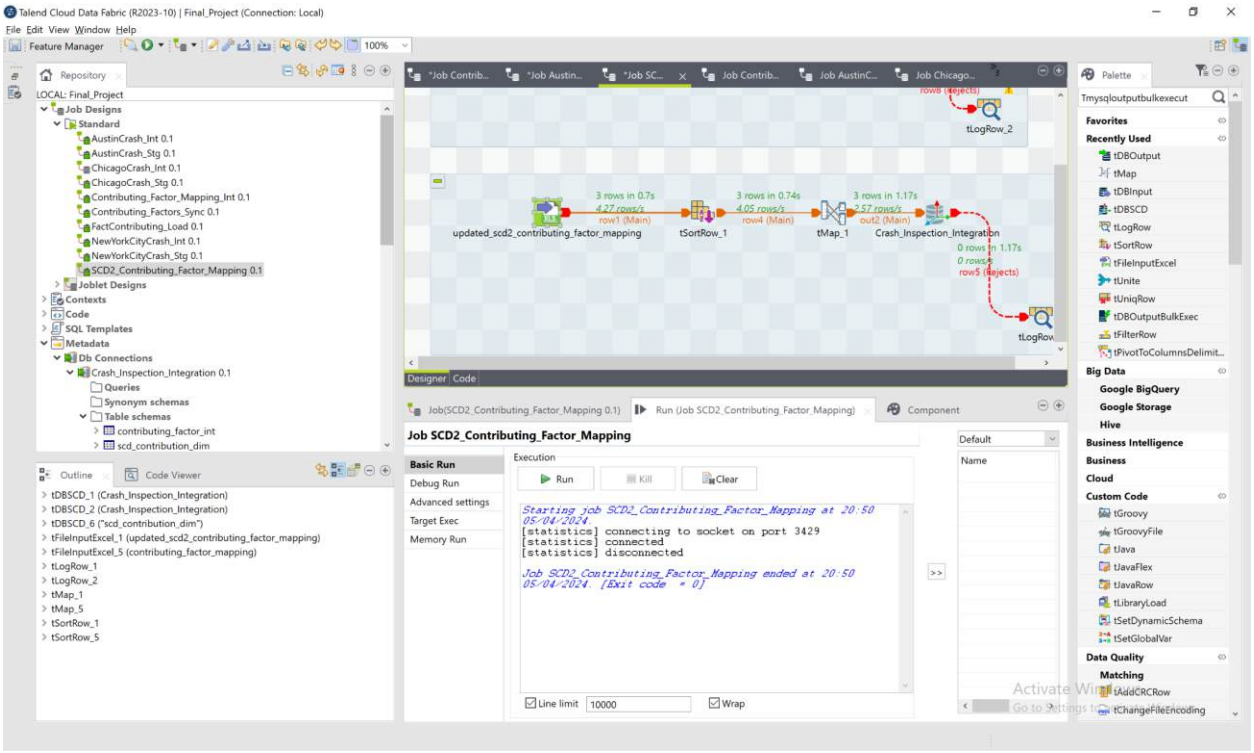

Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
75%
*Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBInput_1 ("stg_austin_crash")
tDBInput_2 ("stg_austin_crash")
tDBInput_3 ("stg_chicago_crash")
tDBInput_4 ("stg_chicago_crash")
tDBInput_5 ("stg_newyorkcity_crash")
tDBInput_6 ("stg_newyorkcity_crash")
tDBInput_7 ("stg_newyorkcity_crash")
tDBInput_8 ("stg_newyorkcity_crash")
tDBInput_9 ("stg_newyorkcity_crash")
tDBOutput_1 (Crash_Inspection_Staging)
tDBOutput_2 (Crash_Inspection_Staging)
tDBOutput_3 (Crash_Inspection_Staging)
tMap_1
tMap_2
tMap_3
tMap_4
*Job Co...
*Job Austin...
*Job NewYor...
*Job SCD2_C...
Job AustinC...
Job Chicago...
Designer
Code
Job(Contributing_Factors_Sync 0.1)
Run (Job Contributing_Factors_Sync)
Component
Job Contributing_Factors_Sync
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job Contributing_Factors_Sync at 20:08 04/04/2024.
[statistics] connecting to socket on port 3955
[statistics] connected
[statistics] disconnected
Job Contributing_Factors_Sync ended at 20:34 04/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Crash_Inspection_St...
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding

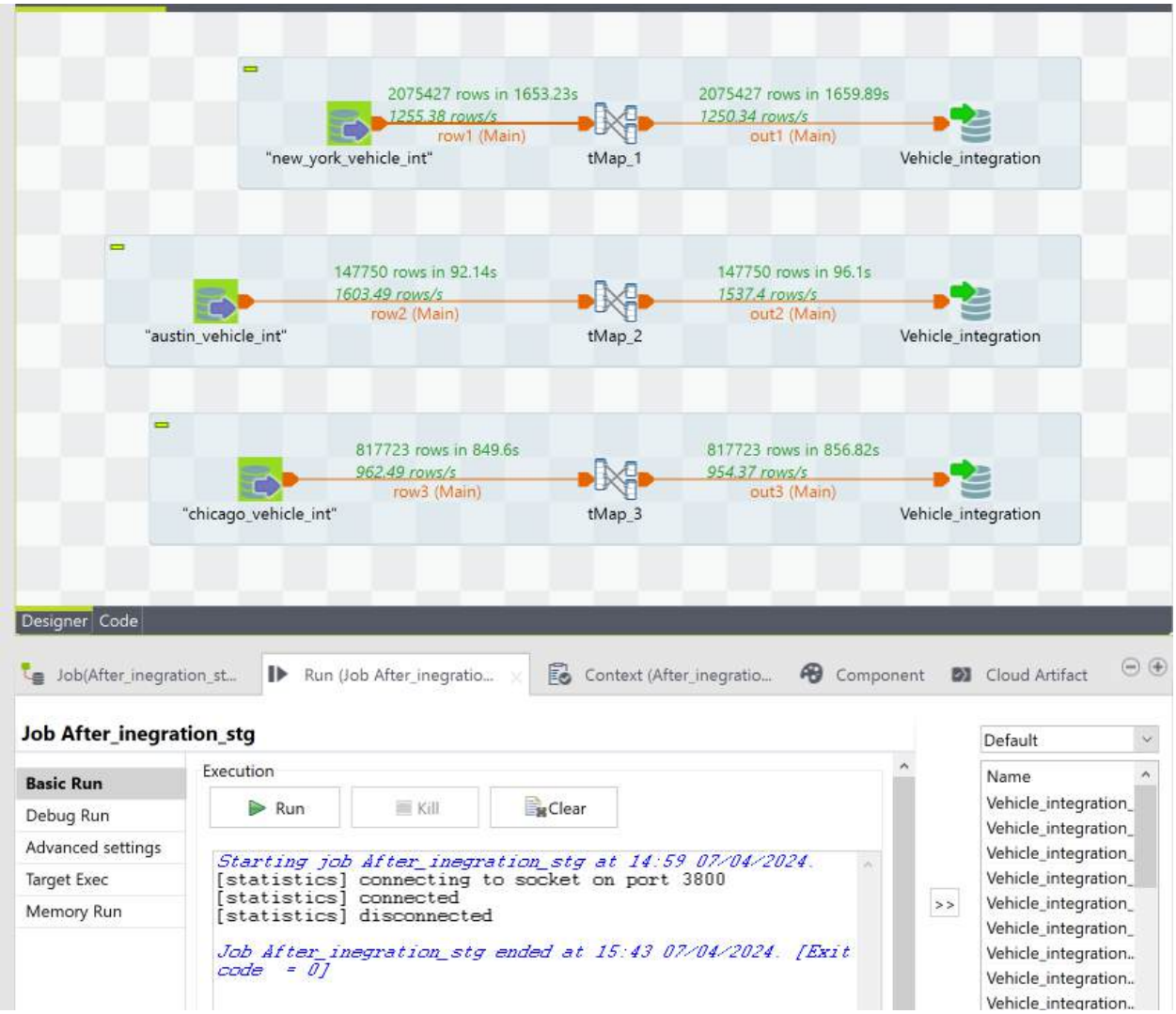

Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
75%
*Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBInput_1 ("stg_austin_crash")
tDBInput_2 ("stg_austin_crash")
tDBInput_3 ("stg_chicago_crash")
tDBInput_4 ("stg_chicago_crash")
tDBInput_5 ("stg_newyorkcity_crash")
tDBInput_6 ("stg_newyorkcity_crash")
tDBInput_7 ("stg_newyorkcity_crash")
tDBInput_8 ("stg_newyorkcity_crash")
tDBInput_9 ("stg_newyorkcity_crash")
tDBOutput_1 (Crash_Inspection_Staging)
tDBOutput_2 (Crash_Inspection_Staging)
tDBOutput_3 (Crash_Inspection_Staging)
tMap_1
tMap_2
tMap_3
tMap_4
*Job Co...
*Job Austin...
*Job NewYor...
*Job SCD2_C...
Job AustinC...
Job Chicago...
Designer
Code
Job(Contributing_Factors_Sync 0.1)
Run (Job Contributing_Factors_Sync)
Component
Job Contributing_Factors_Sync
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job Contributing_Factors_Sync at 20:08 04/04/2024.
[statistics] connecting to socket on port 3955
[statistics] connected
[statistics] disconnected
Job Contributing_Factors_Sync ended at 20:34 04/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Crash_Inspection_St...
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding


Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
75%
*Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBInput_1 ("stg_austin_crash")
tDBInput_2 ("stg_austin_crash")
tDBInput_3 ("stg_chicago_crash")
tDBInput_4 ("stg_chicago_crash")
tDBInput_5 ("stg_newyorkcity_crash")
tDBInput_6 ("stg_newyorkcity_crash")
tDBInput_7 ("stg_newyorkcity_crash")
tDBInput_8 ("stg_newyorkcity_crash")
tDBInput_9 ("stg_newyorkcity_crash")
tDBOutput_1 (Crash_Inspection_Staging)
tDBOutput_2 (Crash_Inspection_Staging)
tDBOutput_3 (Crash_Inspection_Staging)
tMap_1
tMap_2
tMap_3
tMap_4
*Job Co...
*Job Austin...
*Job NewYor...
*Job SCD2_C...
Job AustinC...
Job Chicago...
Designer
Code
Job(Contributing_Factors_Sync 0.1)
Run (Job Contributing_Factors_Sync)
Component
Job Contributing_Factors_Sync
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job Contributing_Factors_Sync at 20:08 04/04/2024.
[statistics] connecting to socket on port 3955
[statistics] connected
[statistics] disconnected
Job Contributing_Factors_Sync ended at 20:34 04/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding

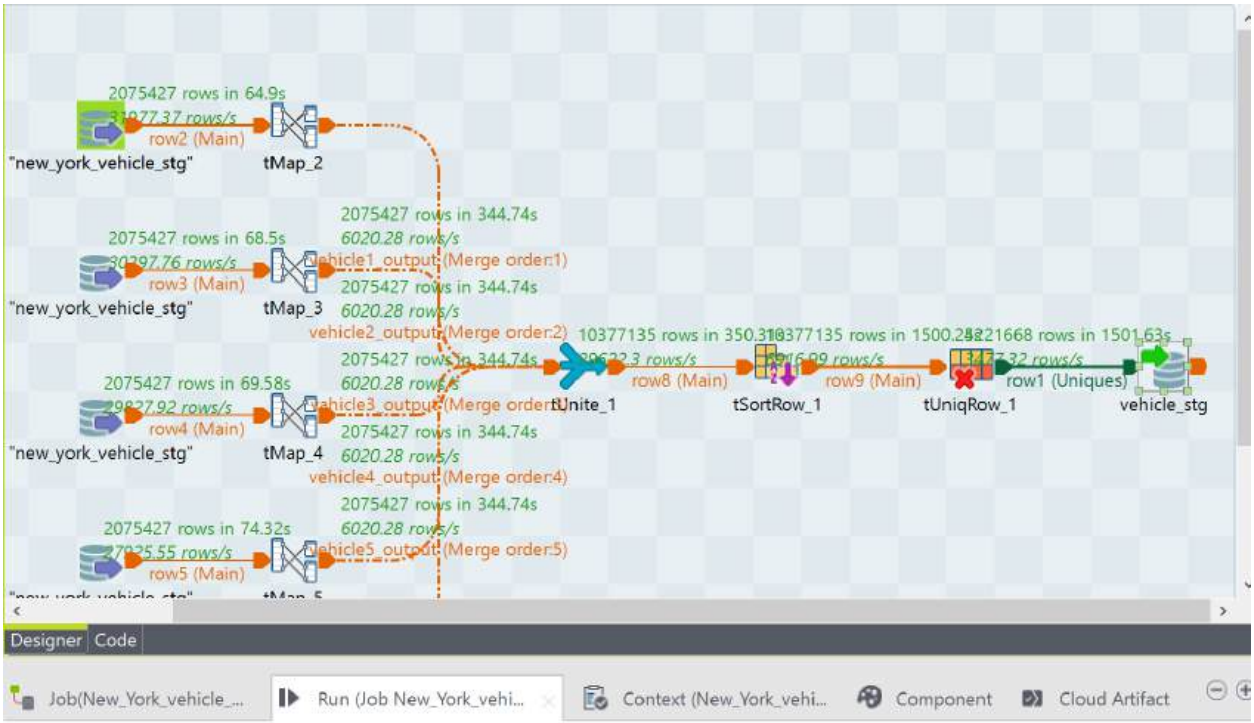

Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
75%
*Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBInput_1 ("cf_austin")
tDBInput_2 ("cf_chicago")
tDBInput_3 ("cf_newyorkcity")
tDBInput_4 ("scd_contribution_dim")
tDBInput_5 ("scd_contribution_dim")
tDBInput_6 ("scd_contribution_dim")
tDBOutput_1 (Crash_Inspection_Integration)
tMap_1
tMap_2
tMap_3
tUnite_1
*Job Contrib...
*Job Austin...
*Job NewYor...
Job Con...
Job AustinC...
Job Chicago...
Designer
Code
Job(Contributing_Factor_Mapping_Int 0.1)
Run (Job Contributing_Factor_Mapping_Int)
Component
Job Contributing_Factor_Mapping_Int
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job Contributing_Factor_Mapping_Int at 19:55 05/04/2024.
[statistics] connecting to socket on port 3895
[statistics] connected
[statistics] disconnected
Job Contributing_Factor_Mapping_Int ended at 20:13 05/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding

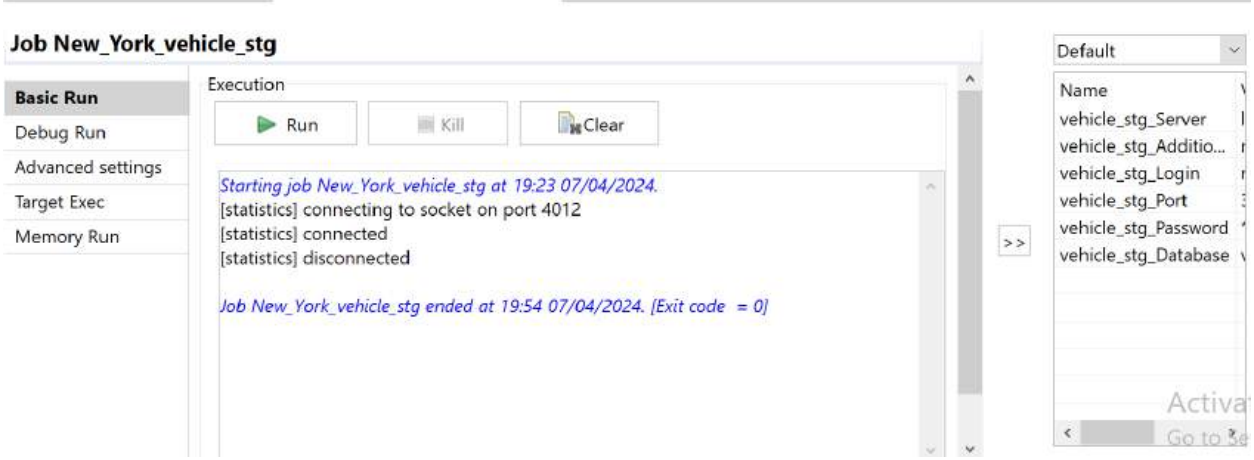

Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBSCD_1 (Crash_Inspection_Integration)
tDBSCD_2 (Crash_Inspection_Integration)
tDBSCD_6 ("scd_contribution_dim")
tFileInputExcel_1 (updated_scd2_contributing_factor_mapping)
tFileInputExcel_5 (contributing_factor_mapping)
tLogRow_1
tLogRow_2
tMap_1
tMap_5
tSortRow_1
tSortRow_5
*Job Contrib...
*Job Austin...
*Job SC...
Job Contrib...
Job AustinC...
Job Chicago...
contributing_factor_mapping
row6 (Main)
tSortRow_5
row7 (Main)
tMap_5
out3 (Main)
"scd_contribution_dim"
row8 (Rejects)
tLogRow_2
updated_scd2_contributing_factor_mapping
3 rows in 0.7s
4.27 rows/s
row1 (Main)
tSortRow_1
3 rows in 0.74s
4.05 rows/s
row4 (Main)
tMap_1
3 rows in 1.17s
2.57 rows/s
out2 (Main)
Crash_Inspection_Integration
0 rows in 1.17s
Designer
Code
Job(SCD2_Contributing_Factor_Mapping 0.1)
Run (Job SCD2_Contributing_Factor_Mapping)
Component
Job SCD2_Contributing_Factor_Mapping
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job SCD2_Contributing_Factor_Mapping at 20:50 05/04/2024.
[statistics] connecting to socket on port 3429
[statistics] connected
[statistics] disconnected
Job SCD2_Contributing_Factor_Mapping ended at 20:50 05/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding


Talend Cloud Data Fabric (R2023-10) | Final_Project (Connection: Local)
File Edit View Window Help
Feature Manager
Repository
LOCAL: Final_Project
Job Designs
Standard
AustinCrash_Int 0.1
AustinCrash_Stg 0.1
ChicagoCrash_Int 0.1
ChicagoCrash_Stg 0.1
Contributing_Factor_Mapping_Int 0.1
Contributing_Factors_Sync 0.1
FactContributing_Load 0.1
NewYorkCityCrash_Int 0.1
NewYorkCityCrash_Stg 0.1
SCD2_Contributing_Factor_Mapping 0.1
Joblet Designs
Contexts
Code
SQL Templates
Metadata
Db Connections
Crash_Inspection_Integration 0.1
Queries
Synonym schemas
Table schemas
contributing_factor_int
scd_contribution_dim
Outline
Code Viewer
tDBSCD_1 (Crash_Inspection_Integration)
tDBSCD_2 (Crash_Inspection_Integration)
tDBSCD_6 ("scd_contribution_dim")
tFileInputExcel_1 (updated_scd2_contributing_factor_mapping)
tFileInputExcel_5 (contributing_factor_mapping)
tLogRow_1
tLogRow_2
tMap_1
tMap_5
tSortRow_1
tSortRow_5
updated_scd2_contributing_factor_mapping
3 rows in 0.7s
4.27 rows/s
row1 (Main)
tSortRow_1
3 rows in 0.74s
4.05 rows/s
row4 (Main)
tMap_1
3 rows in 1.17s
2.57 rows/s
out2 (Main)
Crash_Inspection_Integration
0 rows in 1.17s
0 rows/s
row5 (Rejects)
tLogRow
Designer
Code
Job(SCD2_Contributing_Factor_Mapping 0.1)
Run (Job SCD2_Contributing_Factor_Mapping)
Component
Job SCD2_Contributing_Factor_Mapping
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job SCD2_Contributing_Factor_Mapping at 20:50 05/04/2024.
[statistics] connecting to socket on port 3429
[statistics] connected
[statistics] disconnected
Job SCD2_Contributing_Factor_Mapping ended at 20:50 05/04/2024. [Exit code = 0]
Line limit
10000
Wrap
Default
Name
Palette
Tmysqloutputbulkexecut
Favorites
Recently Used
tDBOutput
tMap
tDBInput
tDBSCD
tLogRow
tSortRow
tFileInputExcel
tUnite
tUniqRow
tDBOutputBulkExec
tFilterRow
tPivotToColumnsDelimit...
Big Data
Google BigQuery
Google Storage
Hive
Business Intelligence
Business
Cloud
Custom Code
tGroovy
tGroovyFile
tJava
tJavaFlex
tJavaRow
tLibraryLoad
tSetDynamicSchema
tSetGlobalVar
Data Quality
Matching
tAddCRCRow
tChangeFileEncoding

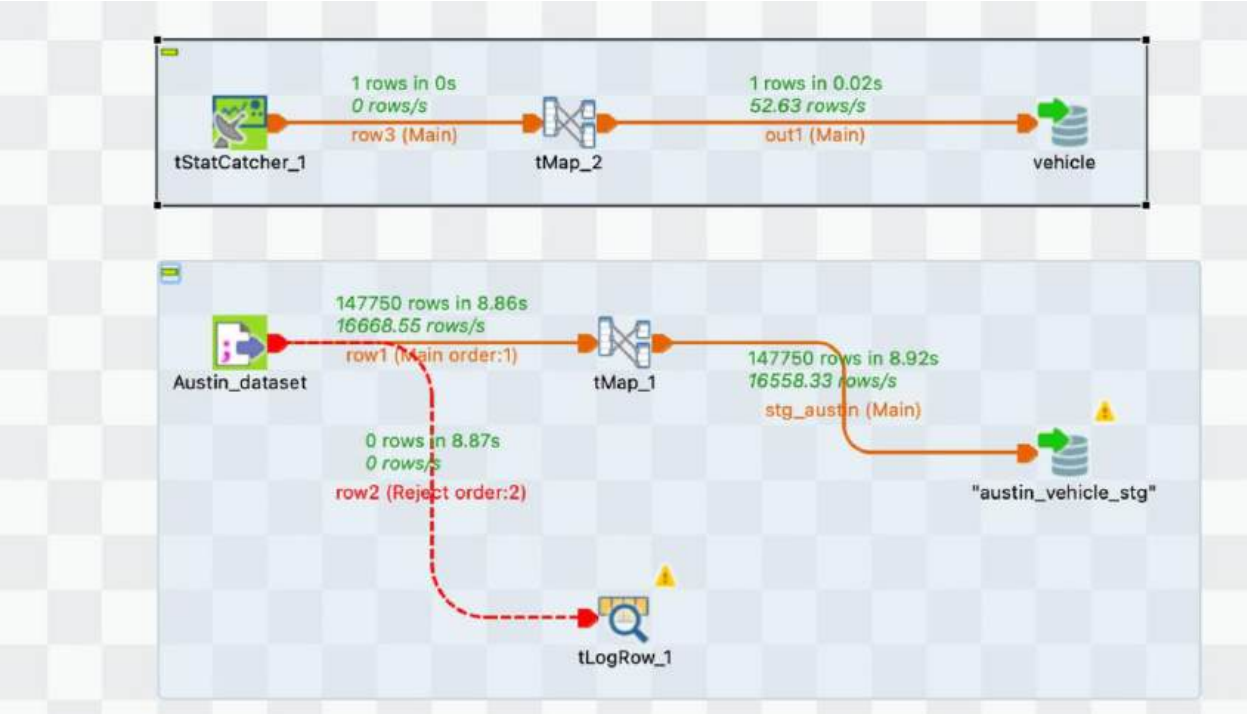

2075427 rows in 1653.23s
1255.38 rows/s
row1 (Main)
"new_york_vehicle_int"
tMap_1
2075427 rows in 1659.89s
1250.34 rows/s
out1 (Main)
Vehicle_integration
147750 rows in 92.14s
1603.49 rows/s
row2 (Main)
"austin_vehicle_int"
tMap_2
147750 rows in 96.1s
1537.4 rows/s
out2 (Main)
Vehicle_integration
817723 rows in 849.6s
962.49 rows/s
row3 (Main)
"chicago_vehicle_int"
tMap_3
817723 rows in 856.82s
954.37 rows/s
out3 (Main)
Vehicle_integration
Designer
Code

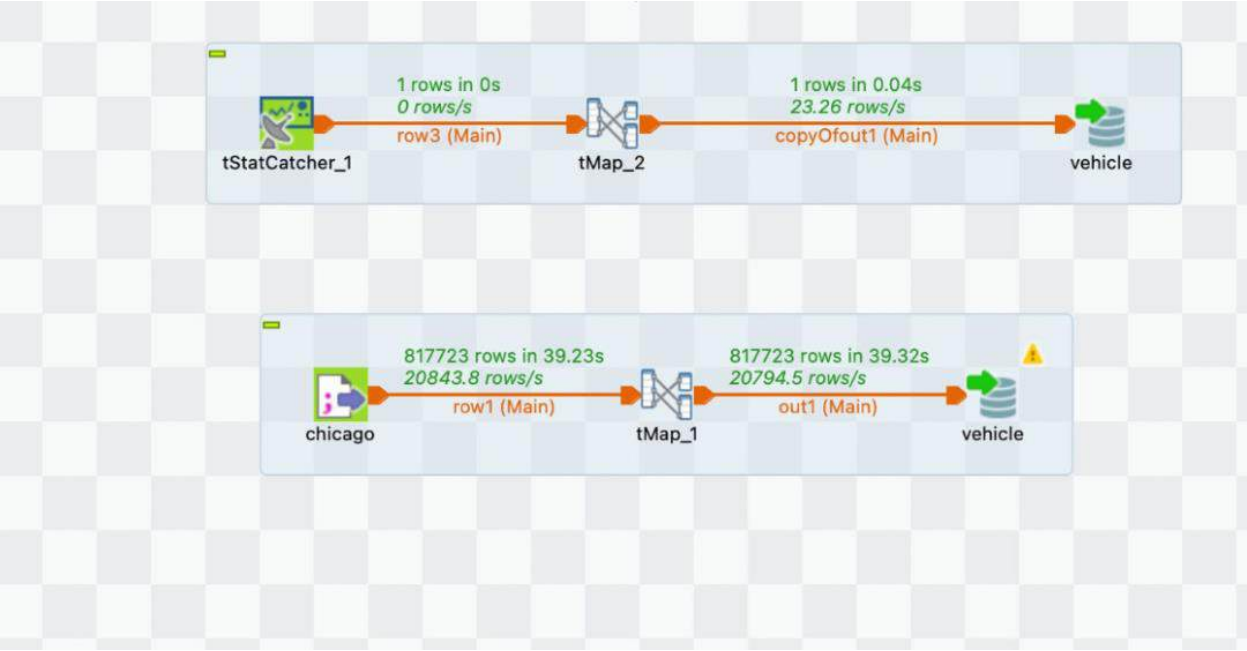
Job(After_inegration_st...
Run (Job After_inegratio...
Context (After_inegratio...
Component
Cloud Artifact
Job After_inegration_stg
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job After_inegration_stg at 14:59 07/04/2024.
[statistics] connecting to socket on port 3800
[statistics] connected
[statistics] disconnected
Job After_inegration_stg ended at 15:43 07/04/2024. [Exit code = 0]
Default
Name
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration..
Vehicle_integration..
Vehicle_integration..

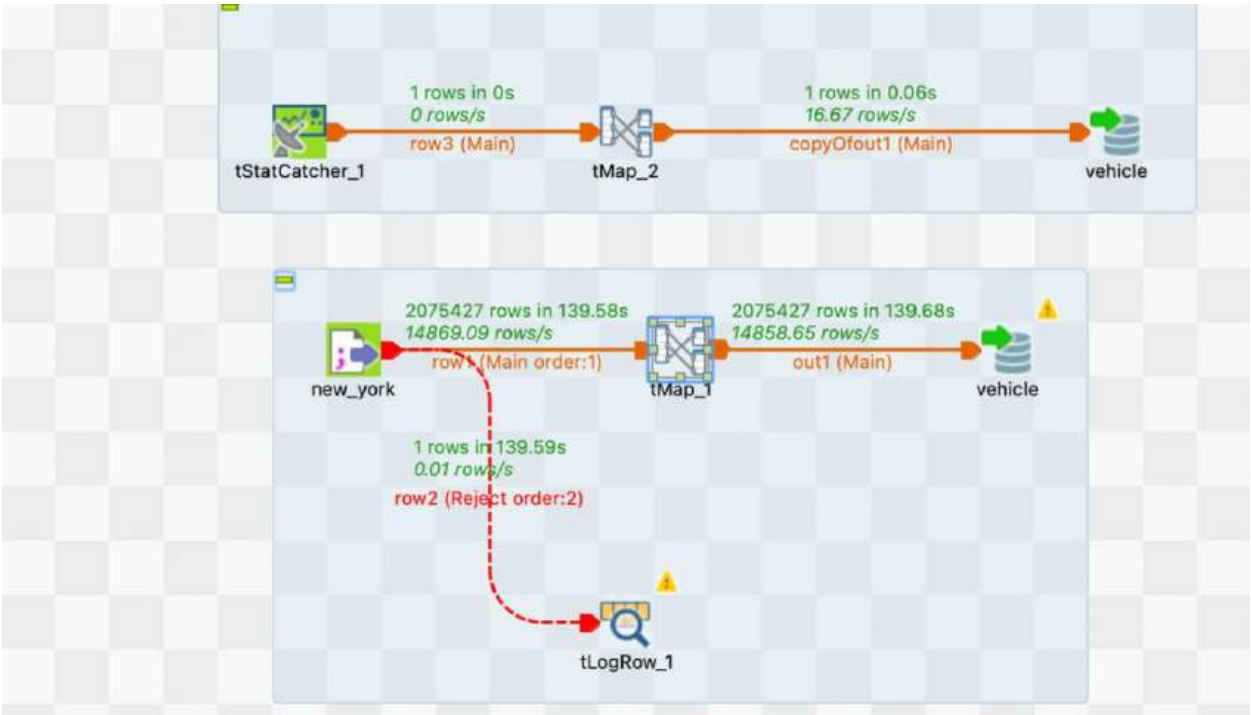

2075427 rows in 64.9s
row2 (Main)
"new_york_vehicle_stg"
tMap_2
2075427 rows in 68.5s
row3 (Main)
"new_york_vehicle_stg"
tMap_3
2075427 rows in 69.58s
row4 (Main)
"new_york_vehicle_stg"
tMap_4
2075427 rows in 74.32s
row5 (Main)
2075427 rows in 344.74s
6020.28 rows/s
vehicle1_output (Merge order:1)
vehicle2_output (Merge order:2)
vehicle3_output (Merge order:3)
vehicle4_output (Merge order:4)
vehicle5_output (Merge order:5)
10377135 rows in 350.31s
10377135 rows in 1500.28s
8221668 rows in 1501.63s
tUnite_1
row8 (Main)
tSortRow_1
row9 (Main)
tUniqRow_1
row1 (Uniques)
vehicle_stg
Designer
Code


Job(New_York_vehicle_...
Run (Job New_York_vehi...
Context (New_York_vehi...
Component
Cloud Artifact
Job New_York_vehicle_stg
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job New_York_vehicle_stg at 19:23 07/04/2024.
[statistics] connecting to socket on port 4012
[statistics] connected
[statistics] disconnected
Job New_York_vehicle_stg ended at 19:54 07/04/2024. [Exit code = 0]
Default
Name
vehicle_stg_Server
vehicle_stg_Additio...
vehicle_stg_Login
vehicle_stg_Port
vehicle_stg_Password
vehicle_stg_Database

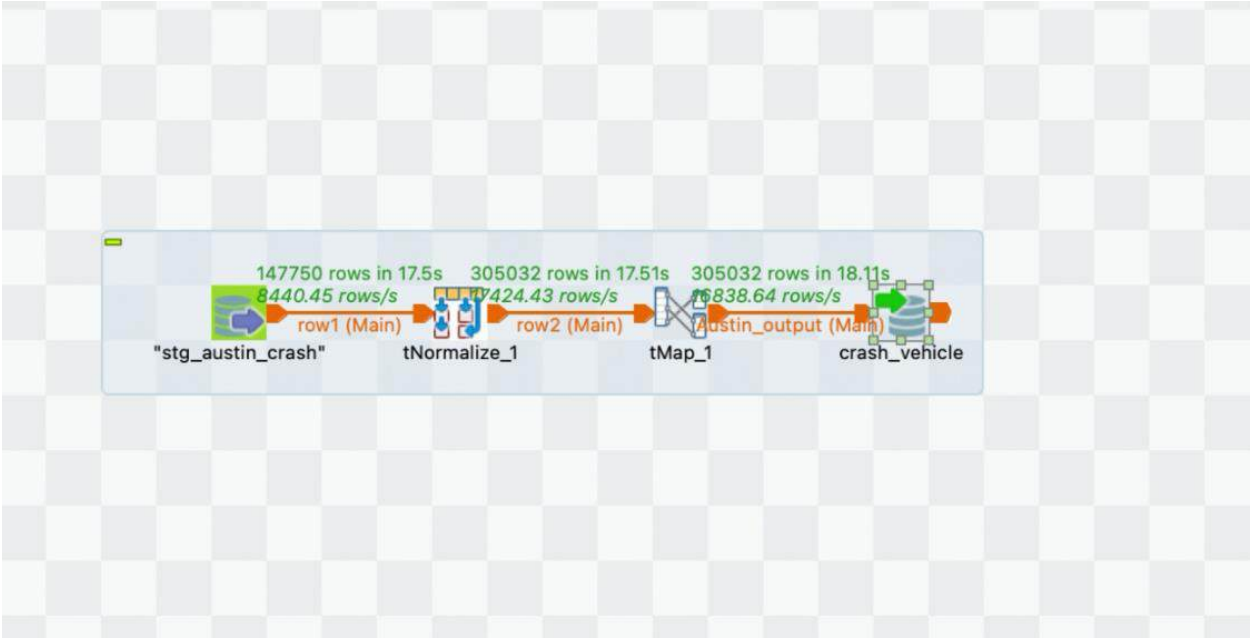
STAGING LAYER:

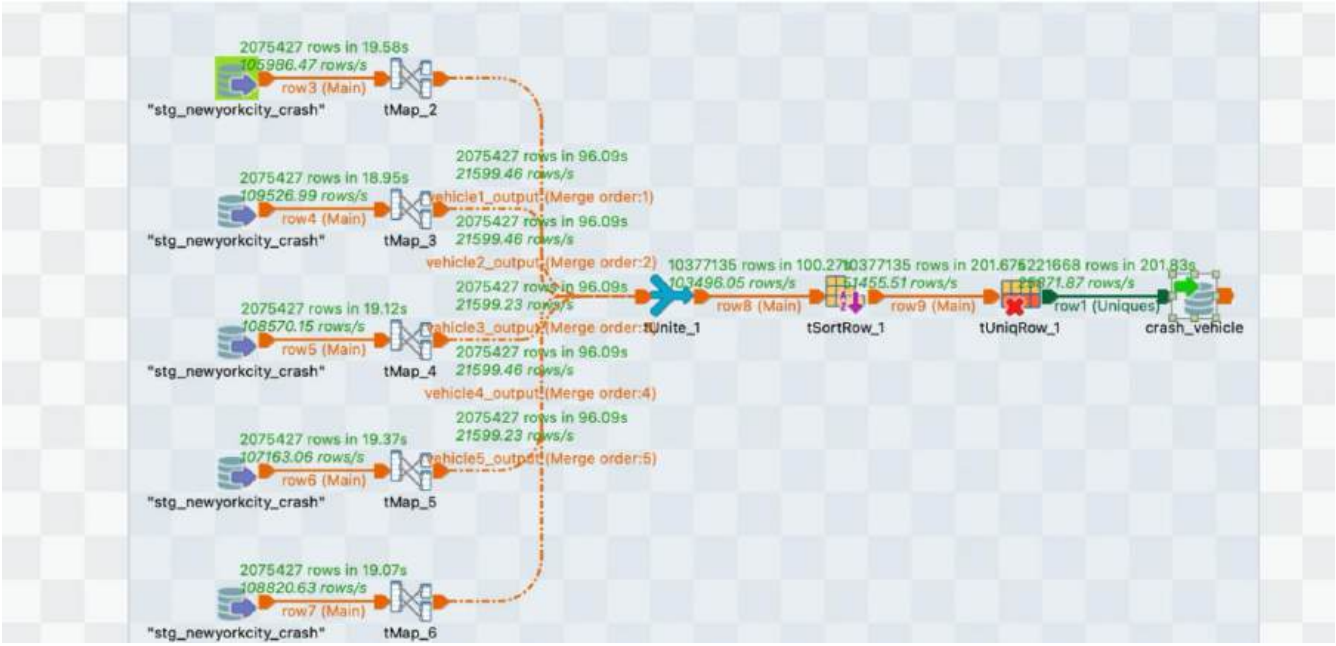


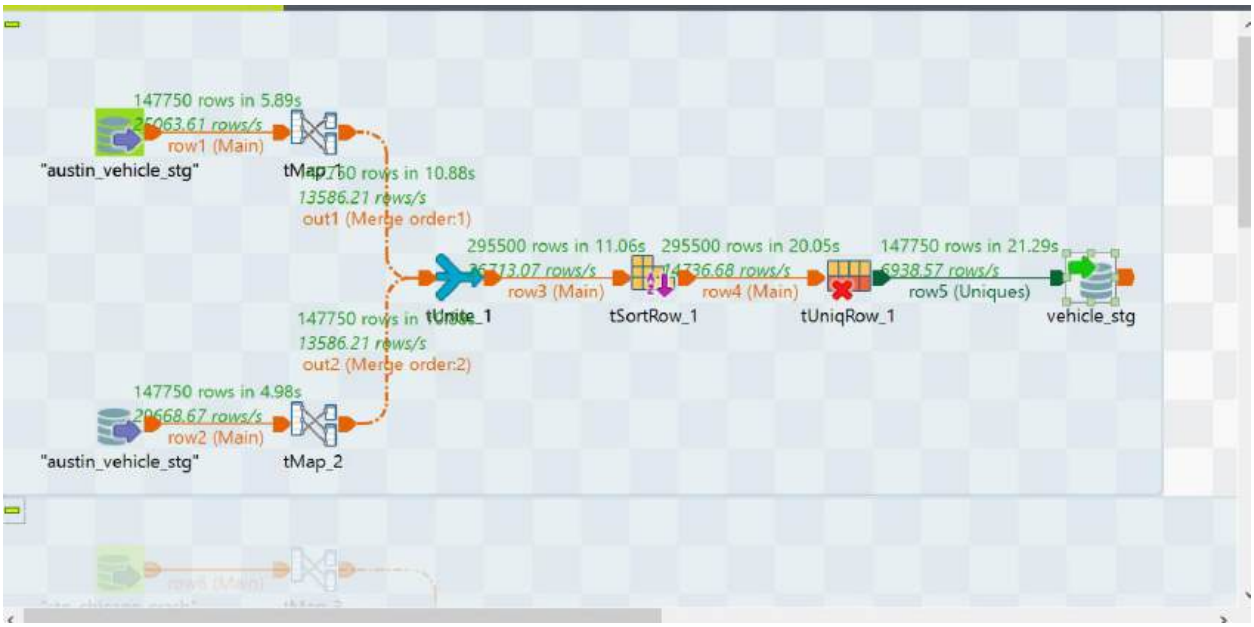


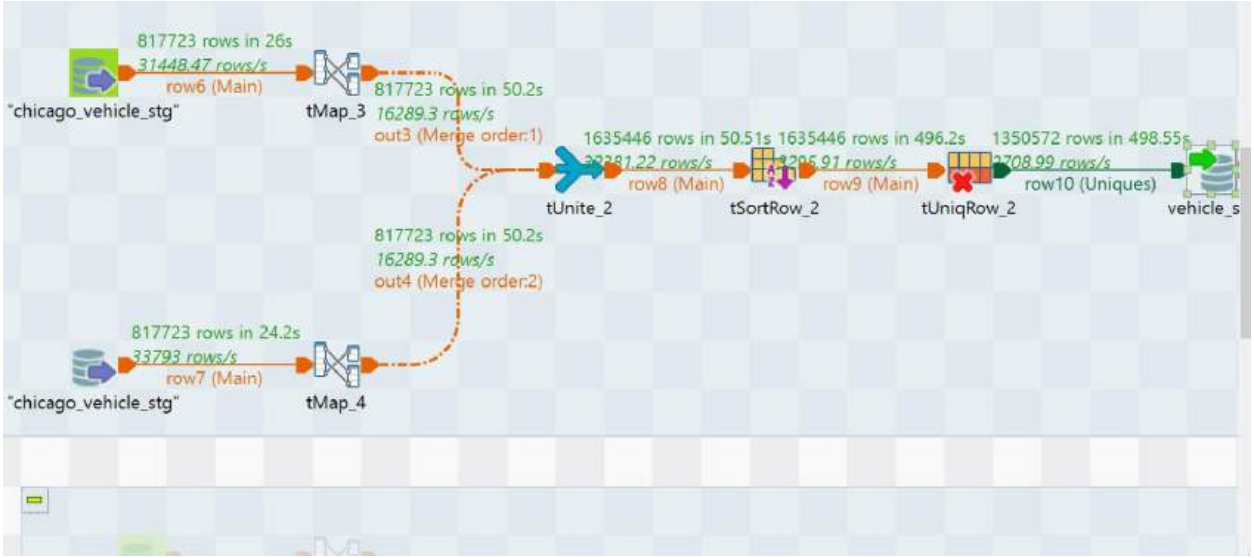


1 row was rejected as the column 'NUMBER_OF_PERSONS_INJURED' is of Datatype 'INT' and the value encountered is 'Unspecified' of datatype VARCHAR, hence that row is rejected.

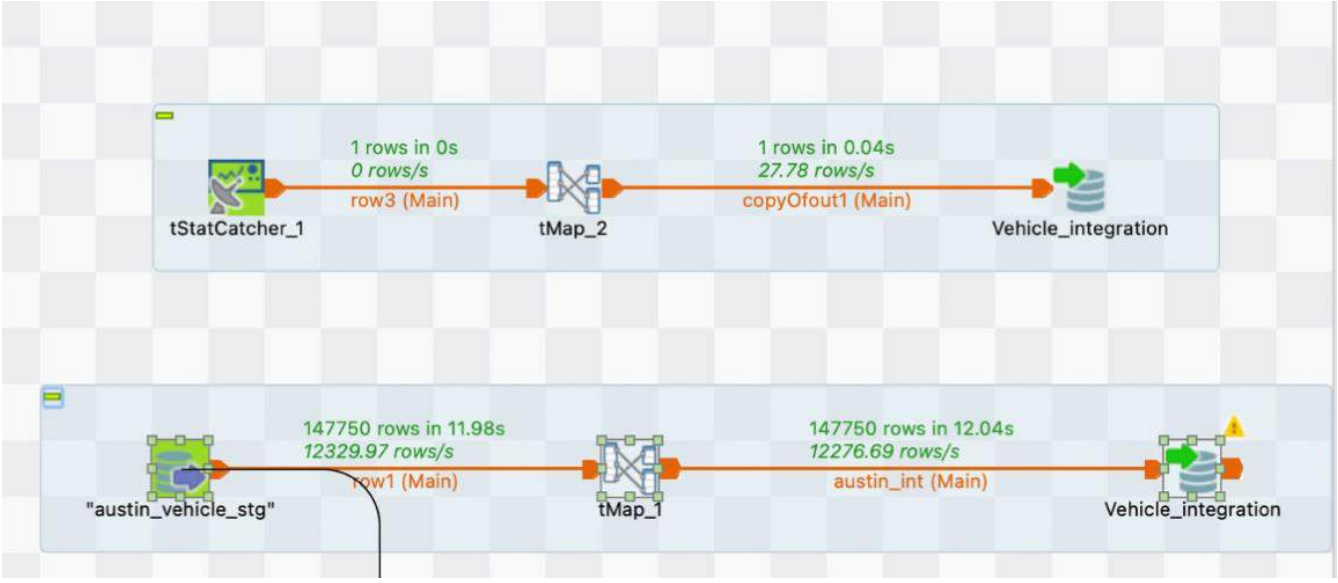


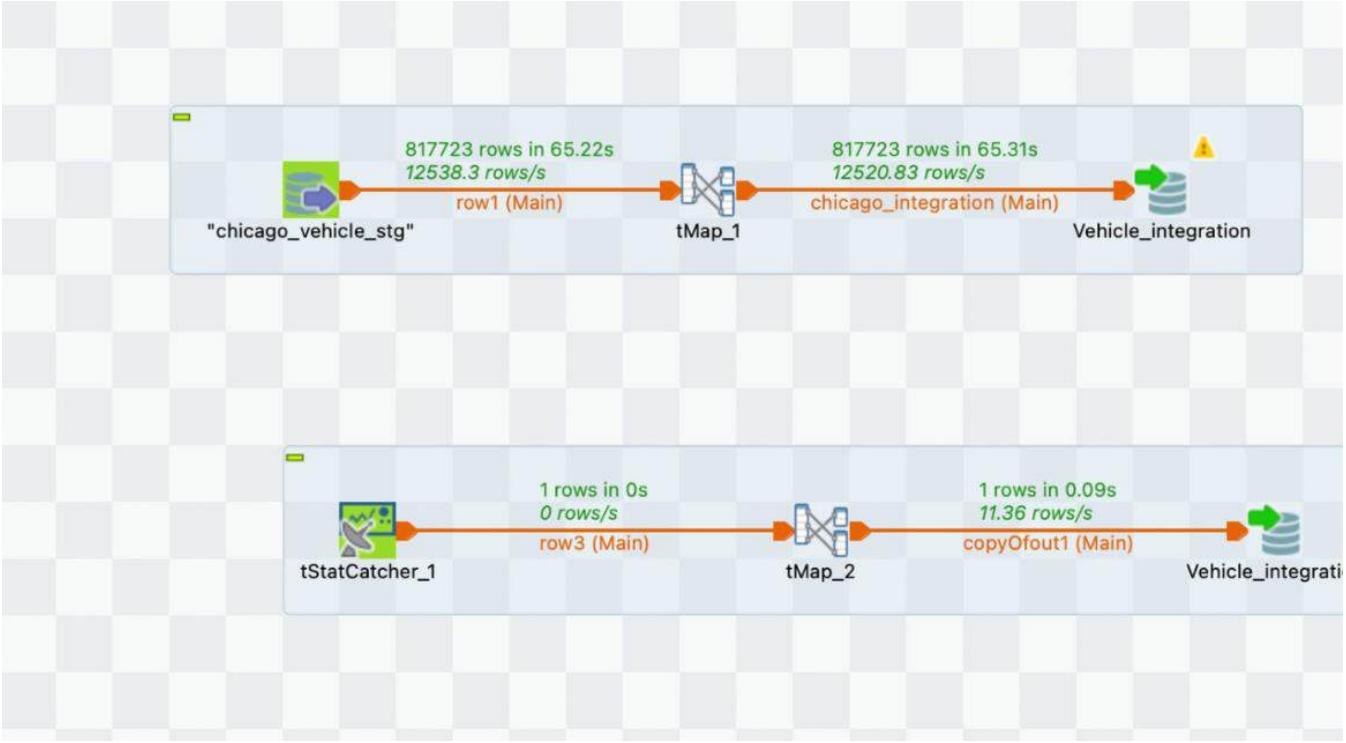

2075427 rows in 19.58s
105986.47 rows/s
row3 (Main)
"stg_newyorkcity_crash"
tMap_2
2075427 rows in 18.95s
109526.99 rows/s
row4 (Main)
"stg_newyorkcity_crash"
tMap_3
2075427 rows in 19.12s
108570.15 rows/s
row5 (Main)
"stg_newyorkcity_crash"
tMap_4
2075427 rows in 19.37s
107163.06 rows/s
row6 (Main)
"stg_newyorkcity_crash"
tMap_5
2075427 rows in 19.07s
108820.63 rows/s
row7 (Main)
"stg_newyorkcity_crash"
tMap_6
2075427 rows in 96.09s
21599.46 rows/s
vehicle1_output (Merge order:1)
2075427 rows in 96.09s
21599.46 rows/s
vehicle2_output (Merge order:2)
2075427 rows in 96.09s
21599.23 rows/s
vehicle3_output (Merge order:3)
2075427 rows in 96.09s
21599.46 rows/s
vehicle4_output (Merge order:4)
2075427 rows in 96.09s
21599.23 rows/s
vehicle5_output (Merge order:5)
tUnite_1
10377135 rows in 100.27s
103496.05 rows/s
row8 (Main)
tSortRow_1
10377135 rows in 201.67s
51455.51 rows/s
row9 (Main)
tUniqRow_1
5221668 rows in 201.83s
25871.87 rows/s
row1 (Uniques)
crash_vehicle

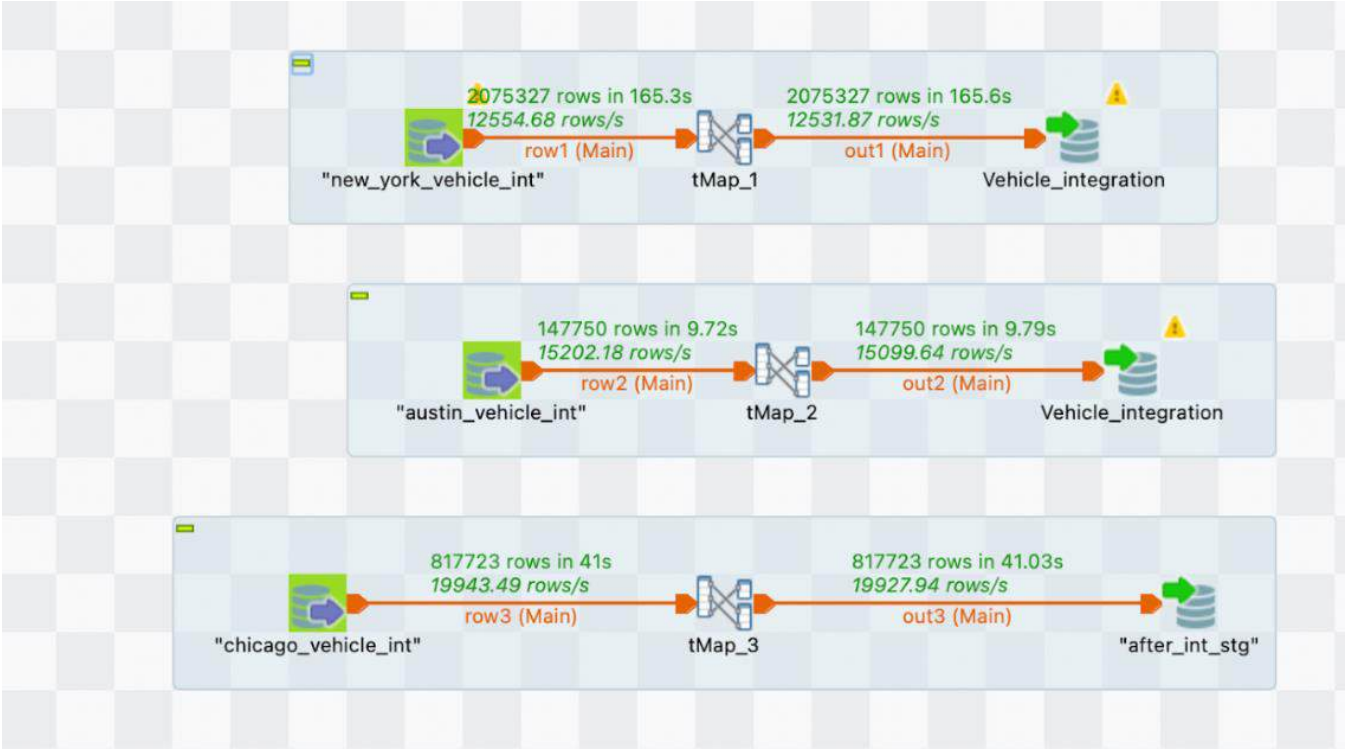


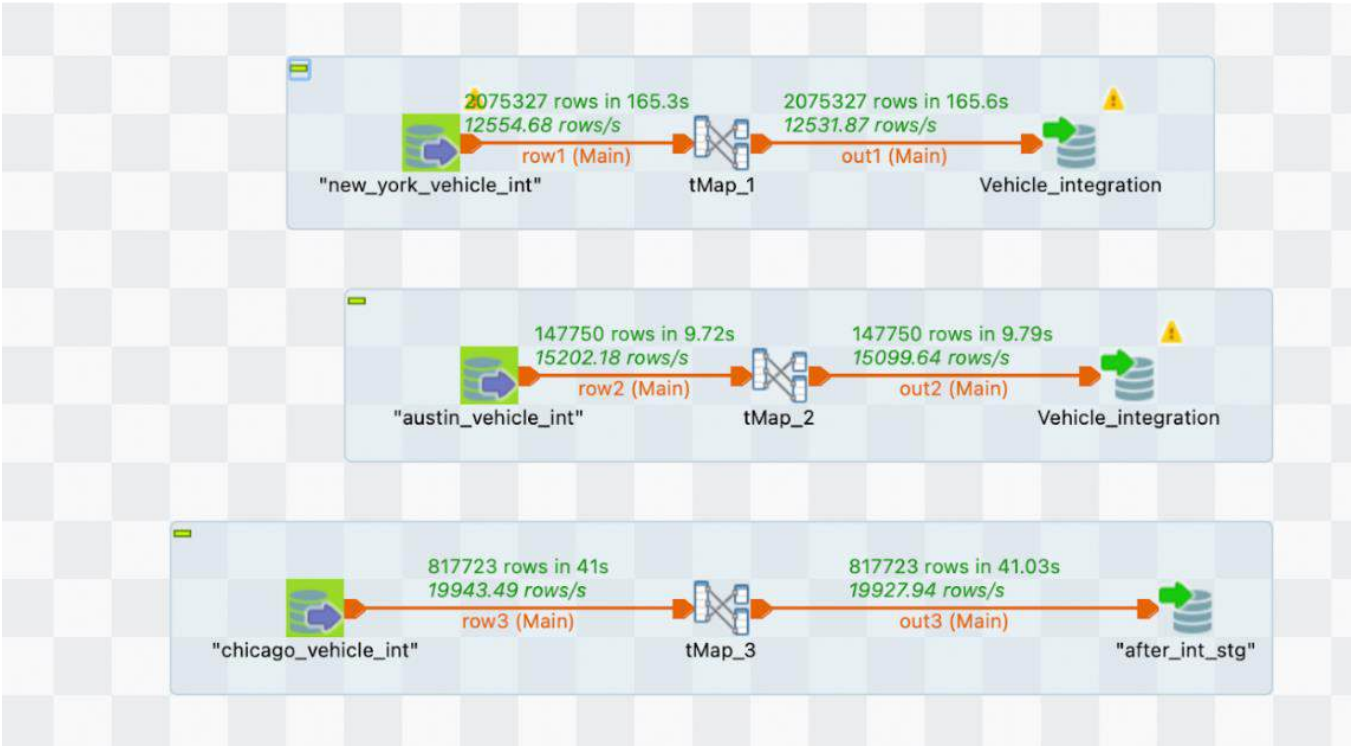


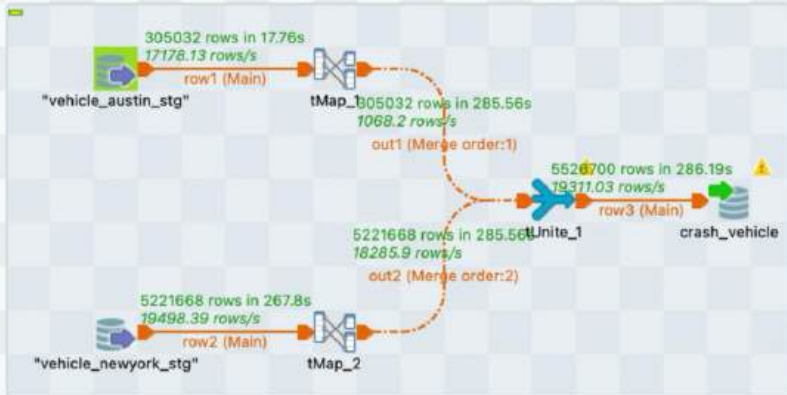
INTEGRATION LAYER:


1 rows in 0s
0 rows/s
row3 (Main)
tStatCatcher_1
tMap_2
1 rows in 0.04s
27.78 rows/s
copyOfout1 (Main)
Vehicle_integration
147750 rows in 11.98s
12329.97 rows/s
row1 (Main)
"austin_vehicle_stg"
tMap_1
147750 rows in 12.04s
12276.69 rows/s
austin_int (Main)
Vehicle_integration

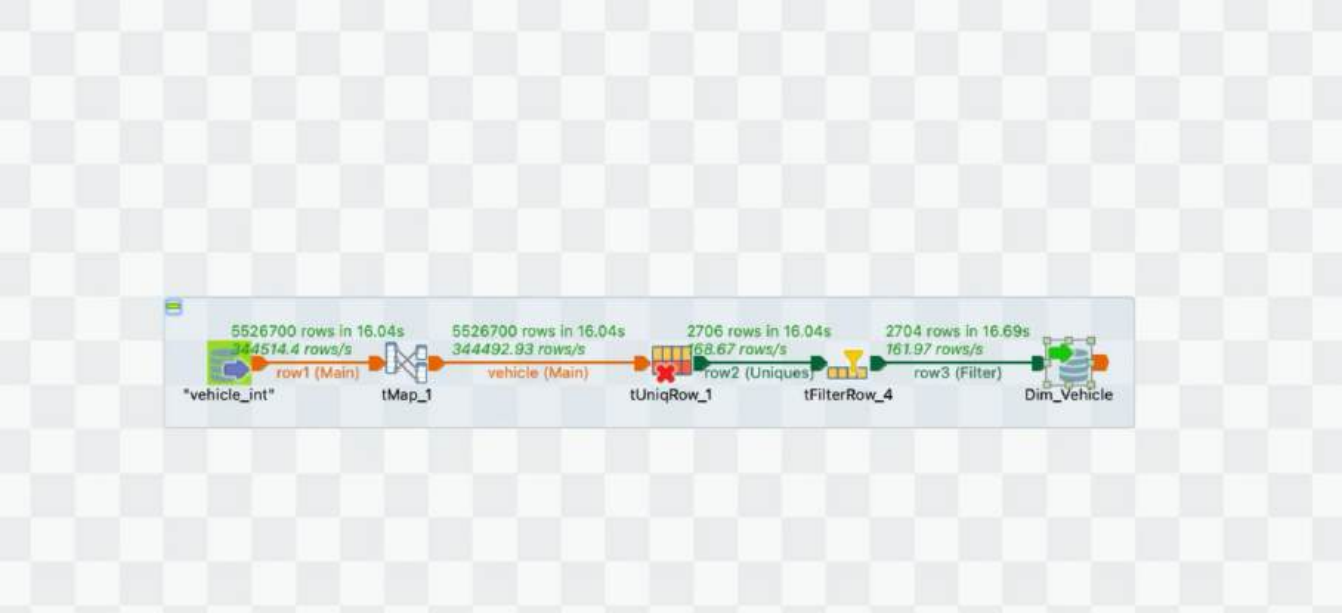

817723 rows in 65.22s
12538.3 rows/s
row1 (Main)
"chicago_vehicle_stg"
tMap_1
817723 rows in 65.31s
12520.83 rows/s
chicago_integration (Main)
Vehicle_integration
1 rows in 0s
0 rows/s
row3 (Main)
tStatCatcher_1
tMap_2
1 rows in 0.09s
11.36 rows/s
copyOfout1 (Main)
Vehicle_integrati

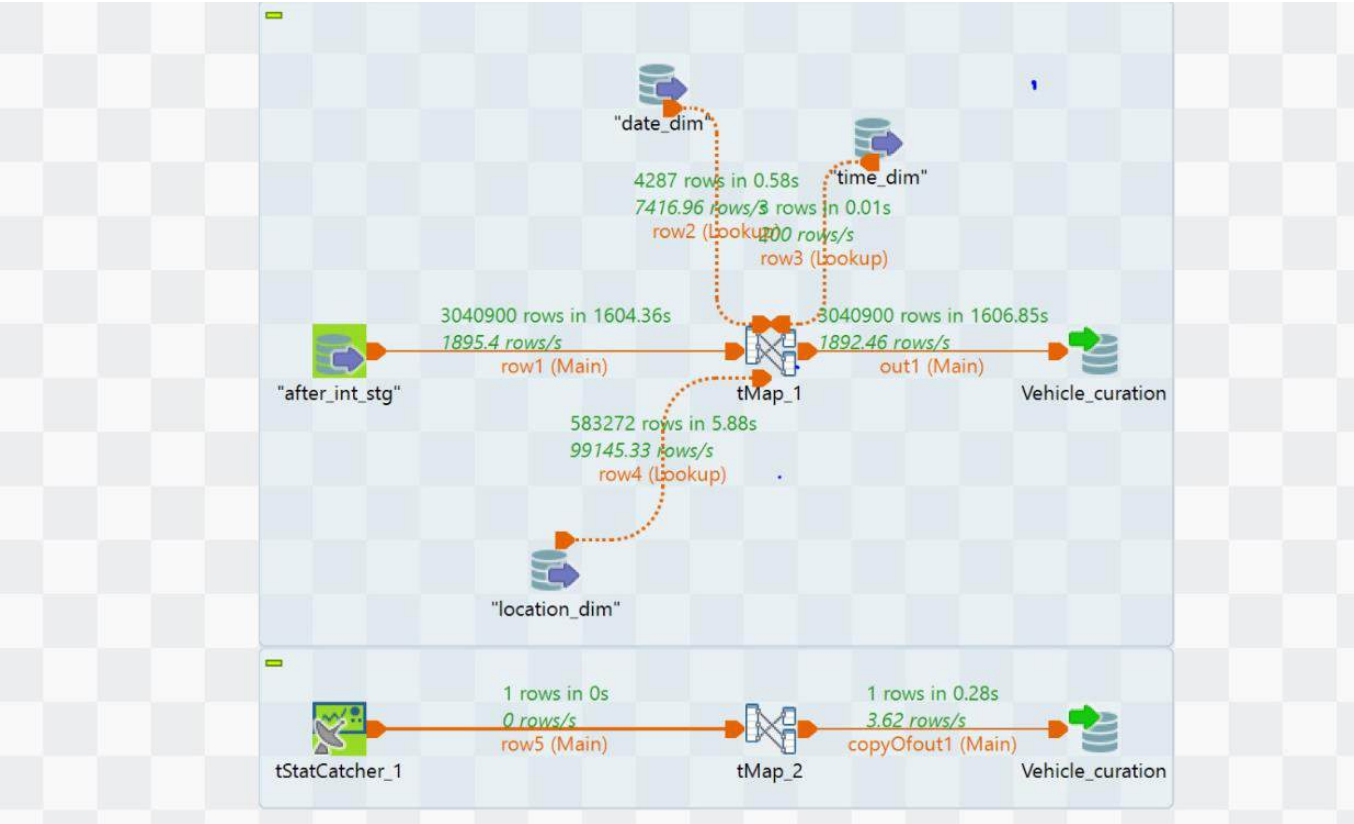

2075327 rows in 165.3s
12554.68 rows/s
row1 (Main)
"new_york_vehicle_int"
tMap_1
2075327 rows in 165.6s
12531.87 rows/s
out1 (Main)
Vehicle_integration
147750 rows in 9.72s
15202.18 rows/s
row2 (Main)
"austin_vehicle_int"
tMap_2
147750 rows in 9.79s
15099.64 rows/s
out2 (Main)
Vehicle_integration
817723 rows in 41s
19943.49 rows/s
row3 (Main)
"chicago_vehicle_int"
tMap_3
817723 rows in 41.03s
19927.94 rows/s
out3 (Main)
"after_int_stg"


2075327 rows in 165.3s
12554.68 rows/s
row1 (Main)
"new_york_vehicle_int"
tMap_1
2075327 rows in 165.6s
12531.87 rows/s
out1 (Main)
Vehicle_integration
147750 rows in 9.72s
15202.18 rows/s
row2 (Main)
"austin_vehicle_int"
tMap_2
147750 rows in 9.79s
15099.64 rows/s
out2 (Main)
Vehicle_integration
817723 rows in 41s
19943.49 rows/s
row3 (Main)
"chicago_vehicle_int"
tMap_3
817723 rows in 41.03s
19927.94 rows/s
out3 (Main)
"after_int_stg"


305032 rows in 17.76s
17178.13 rows/s
row1 (Main)
"vehicle_austin_stg"
tMap_1
305032 rows in 285.56s
1068.2 rows/s
out1 (Merge order:1)
5526700 rows in 286.19s
19311.03 rows/s
row3 (Main)
tUnite_1
crash_vehicle
5221668 rows in 285.56s
18285.9 rows/s
out2 (Merge order:2)
5221668 rows in 267.8s
19498.39 rows/s
row2 (Main)
"vehicle_newyork_stg"
tMap_2

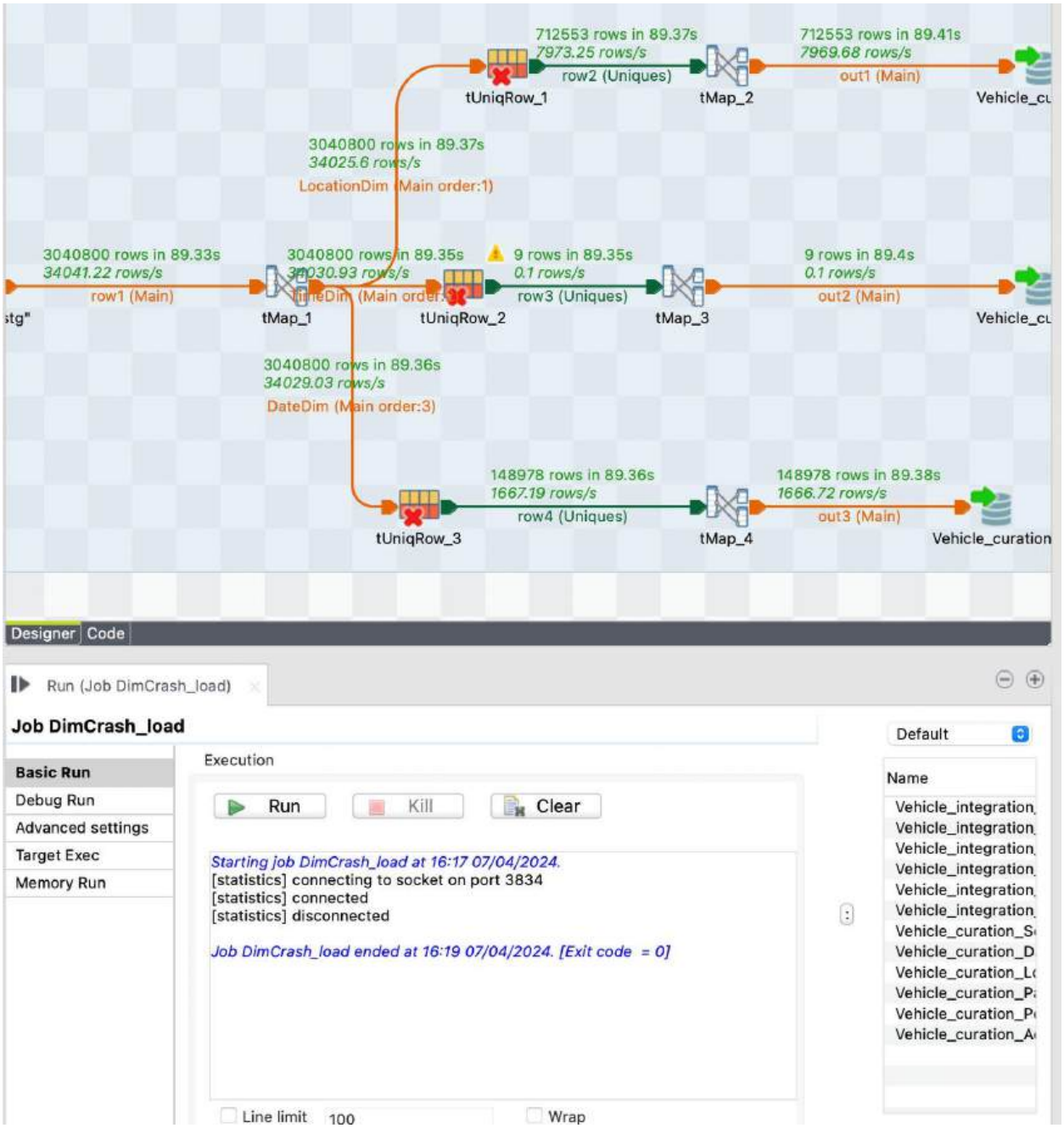
CURATION LAYER:

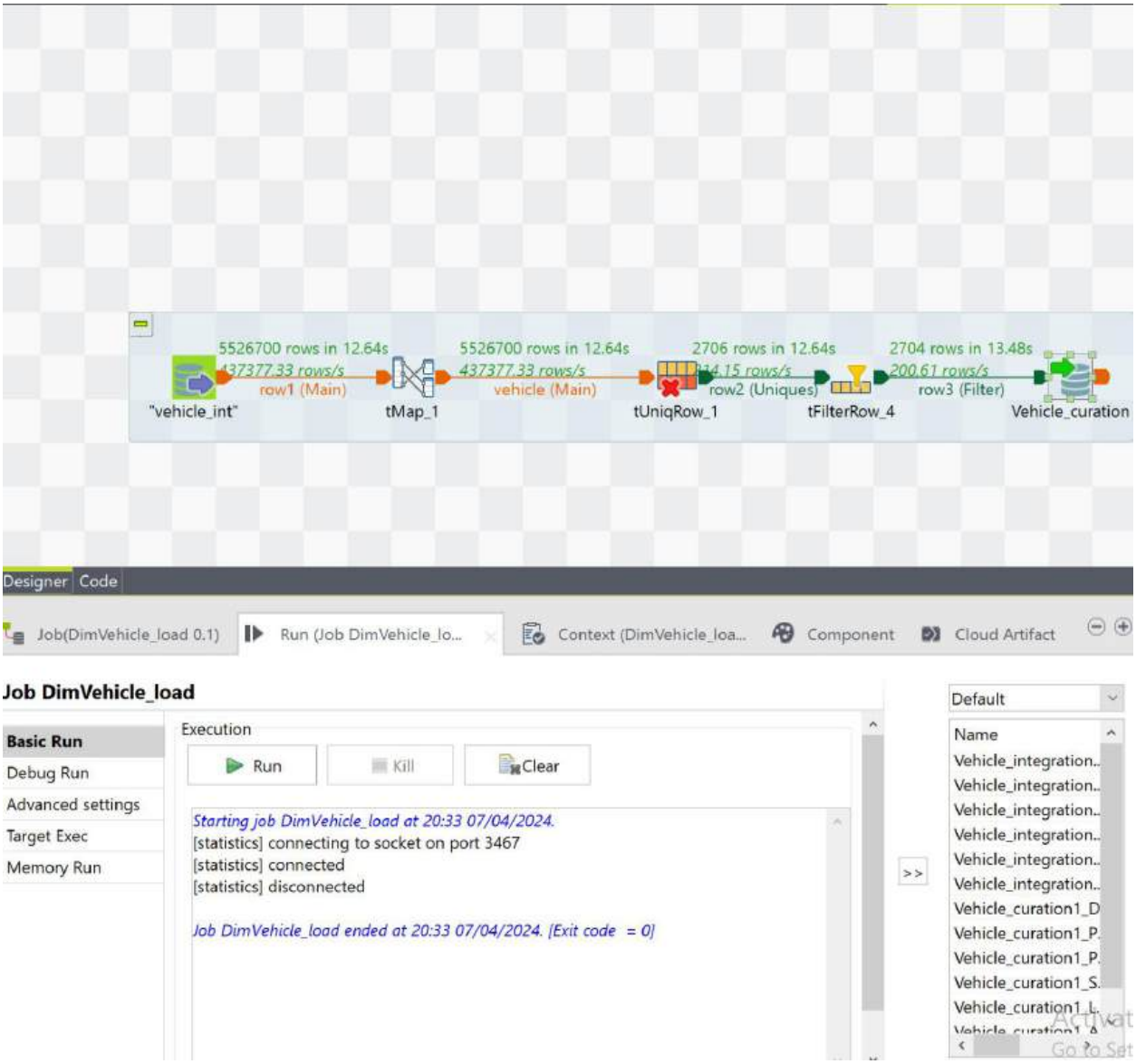


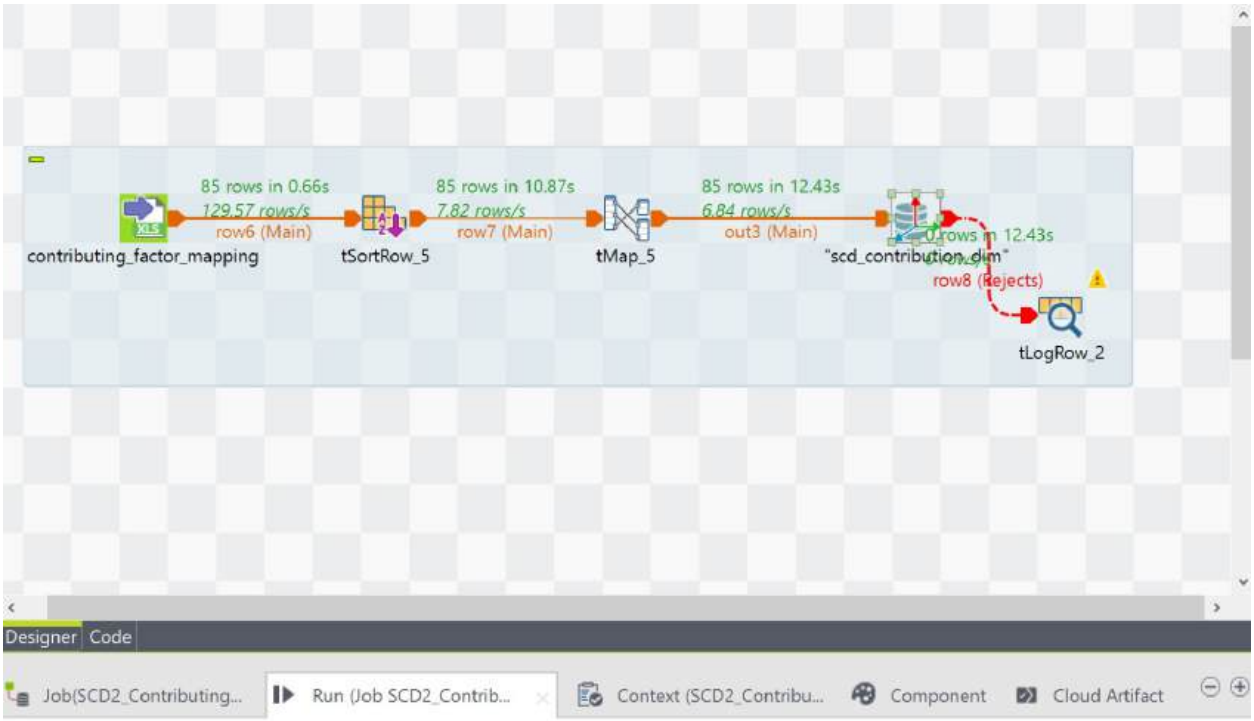


CURATION LAYER:

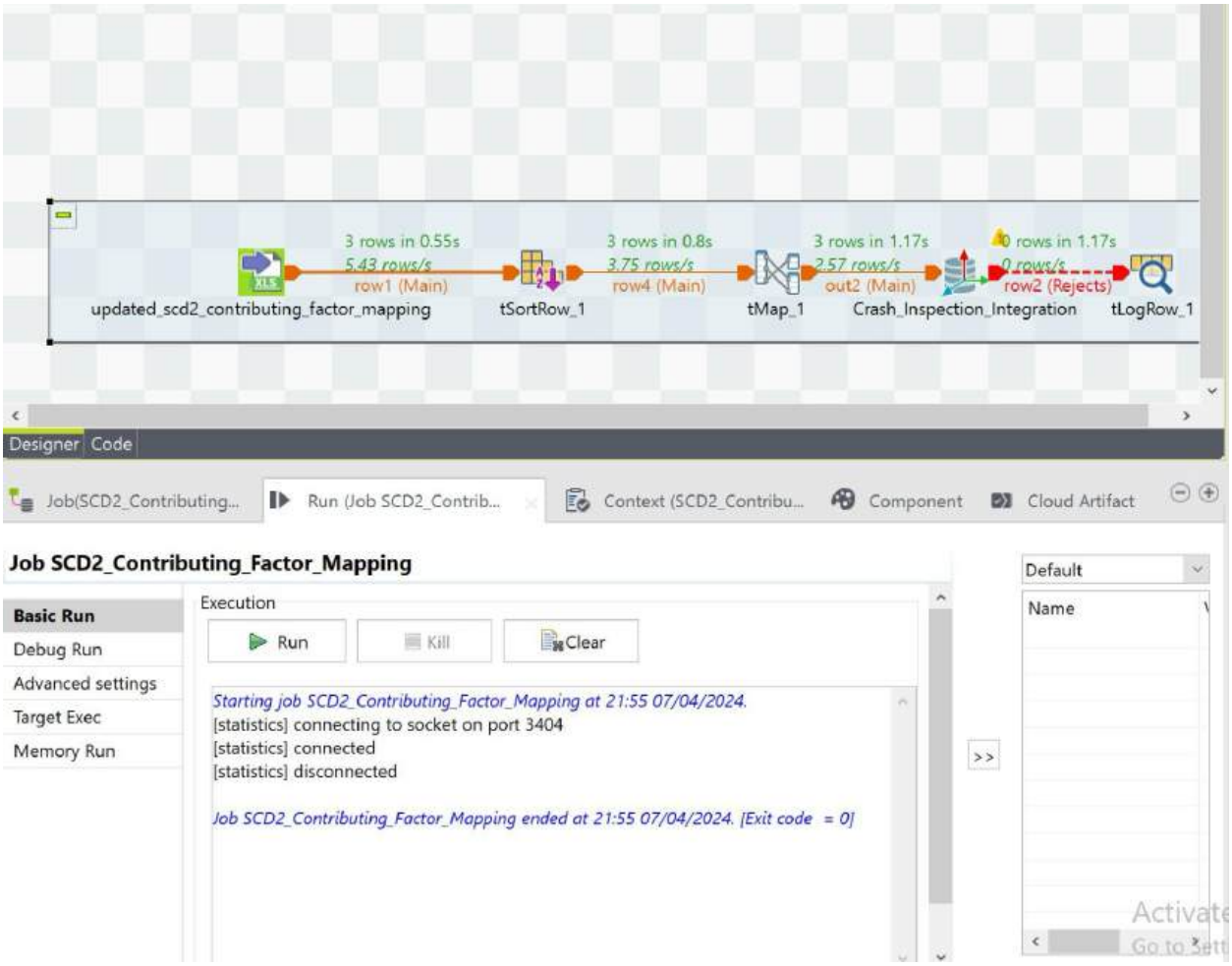
DIMS:


712553 rows in 89.37s
7973.25 rows/s
row2 (Uniques)
tUniqRow_1
tMap_2
712553 rows in 89.41s
7969.68 rows/s
out1 (Main)
3040800 rows in 89.37s
34025.6 rows/s
LocationDim (Main order:1)
3040800 rows in 89.33s
34041.22 rows/s
row1 (Main)
tMap_1
3040800 rows in 89.35s
9 rows in 89.35s
0.1 rows/s
row3 (Uniques)
tUniqRow_2
tMap_3
9 rows in 89.4s
0.1 rows/s
out2 (Main)
3040800 rows in 89.36s
34029.03 rows/s
DateDim (Main order:3)
148978 rows in 89.36s
1667.19 rows/s
row4 (Uniques)
tUniqRow_3
tMap_4
148978 rows in 89.38s
1666.72 rows/s
out3 (Main)
Vehicle_curation
Designer
Code
Run (Job DimCrash_load)
Job DimCrash_load
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job DimCrash_load at 16:17 07/04/2024.
[statistics] connecting to socket on port 3834
[statistics] connected
[statistics] disconnected
Job DimCrash_load ended at 16:19 07/04/2024. [Exit code = 0]
Line limit
100
Wrap
Default
Name
Vehicle_integration
Vehicle_integration
Vehicle_integration
Vehicle_integration
Vehicle_integration
Vehicle_integration
Vehicle_curation_S
Vehicle_curation_D
Vehicle_curation_L
Vehicle_curation_P
Vehicle_curation_P
Vehicle_curation_A

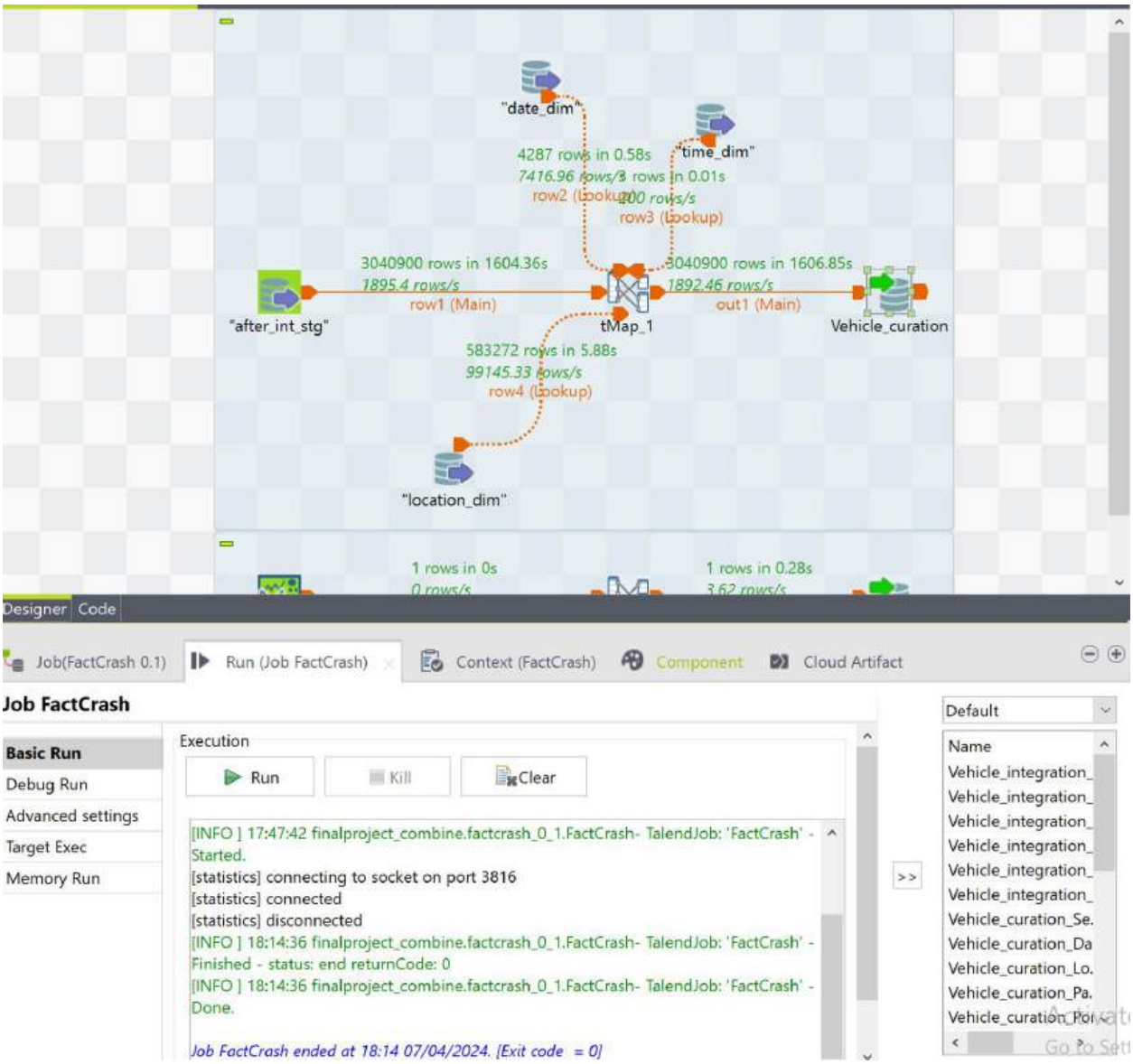

5526700 rows in 12.64s
437377.33 rows/s
row1 (Main)
"vehicle_int"
tMap_1
5526700 rows in 12.64s
437377.33 rows/s
vehicle (Main)
tUniqRow_1
2706 rows in 12.64s
214.15 rows/s
row2 (Uniques)
tFilterRow_4
2704 rows in 13.48s
200.61 rows/s
row3 (Filter)
Vehicle_curation
Designer
Code
Job(DimVehicle_load 0.1)
Run (Job DimVehicle_lo...
Context (DimVehicle_loa...
Component
Cloud Artifact
Job DimVehicle_load
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job DimVehicle_load at 20:33 07/04/2024.
[statistics] connecting to socket on port 3467
[statistics] connected
[statistics] disconnected
Job DimVehicle_load ended at 20:33 07/04/2024. [Exit code = 0]
Default
Name
Vehicle_integration..
Vehicle_integration..
Vehicle_integration..
Vehicle_integration..
Vehicle_integration..
Vehicle_integration..
Vehicle_curation1_D
Vehicle_curation1_P.
Vehicle_curation1_P.
Vehicle_curation1_S.

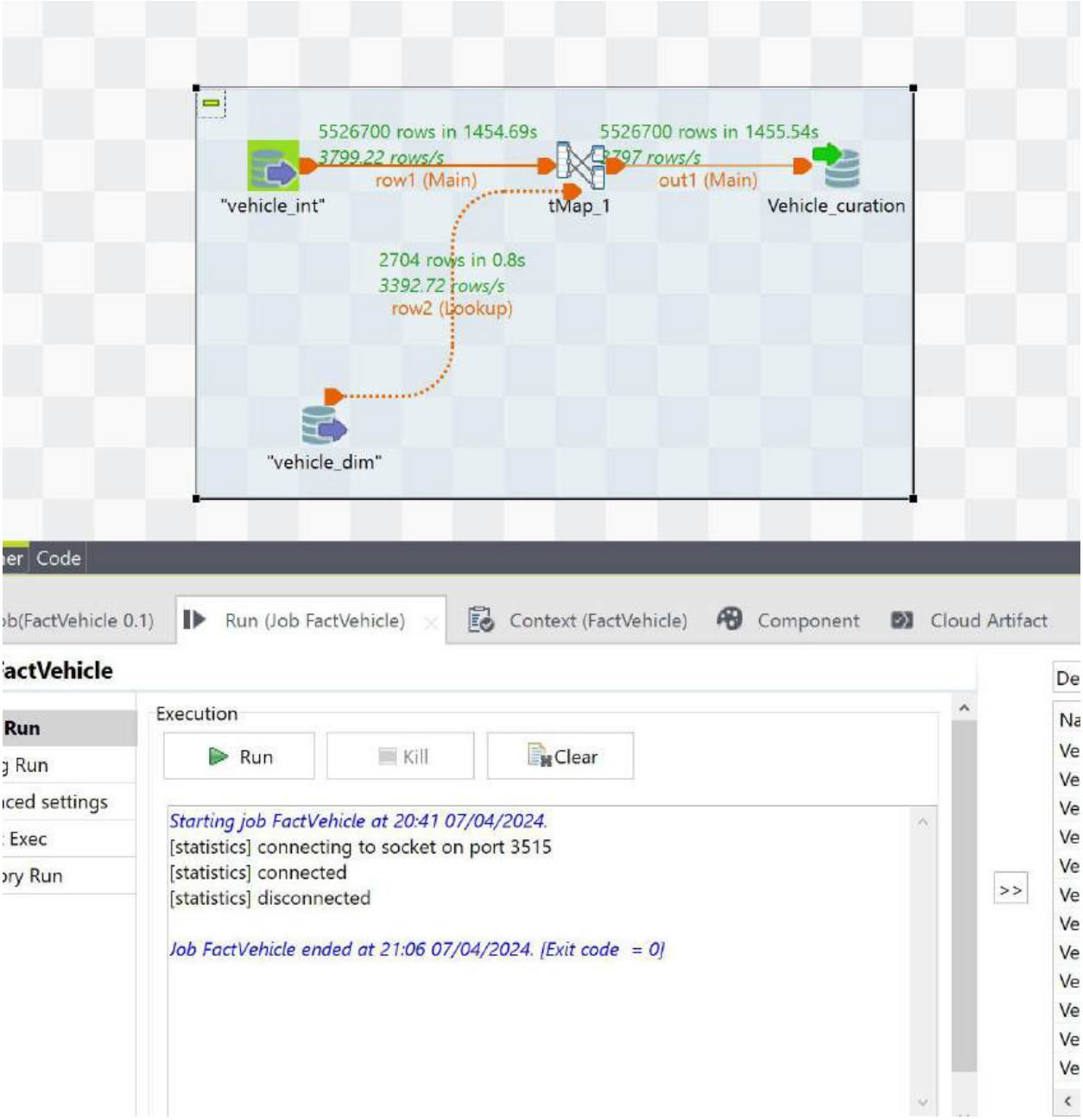

85 rows in 0.66s
129.57 rows/s
row6 (Main)
contributing_factor_mapping
tSortRow_5
85 rows in 10.87s
7.82 rows/s
row7 (Main)
tMap_5
85 rows in 12.43s
6.84 rows/s
out3 (Main)
"scd_contribution_dim"
0 rows in 12.43s
row8 (Rejects)
tLogRow_2
Designer
Code

Job(SCD2_Contributing...
Run (Job SCD2_Contrib...
Context (SCD2_Contribu...
Component
Cloud Artifact
Job SCD2_Contributing_Factor_Mapping
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job SCD2_Contributing_Factor_Mapping at 21:47 07/04/2024.
[statistics] connecting to socket on port 3966
[statistics] connected
[statistics] disconnected
Job SCD2_Contributing_Factor_Mapping ended at 21:47 07/04/2024. [Exit code = 0]
Default
Name


3 rows in 0.55s
5.43 rows/s
row1 (Main)
updated_scd2_contributing_factor_mapping
tSortRow_1
3 rows in 0.8s
3.75 rows/s
row4 (Main)
tMap_1
3 rows in 1.17s
2.57 rows/s
out2 (Main)
Crash_Inspection_Integration
0 rows in 1.17s
0 rows/s
row2 (Rejects)
tLogRow_1
Designer
Code
Job(SCD2_Contributing...
Run (Job SCD2_Contrib...
Context (SCD2_Contribu...
Component
Cloud Artifact
Job SCD2_Contributing_Factor_Mapping
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
Starting job SCD2_Contributing_Factor_Mapping at 21:55 07/04/2024.
[statistics] connecting to socket on port 3404
[statistics] connected
[statistics] disconnected
Job SCD2_Contributing_Factor_Mapping ended at 21:55 07/04/2024. [Exit code = 0]
Default
Name

FACTS:


"date_dim"
"time_dim"
4287 rows in 0.58s
7416.96 rows/s
row2 (Lookup)
3 rows in 0.01s
300 rows/s
row3 (Lookup)
3040900 rows in 1604.36s
1895.4 rows/s
row1 (Main)
3040900 rows in 1606.85s
1892.46 rows/s
out1 (Main)
"after_int_stg"
tMap_1
Vehicle_curation
583272 rows in 5.88s
99145.33 rows/s
row4 (Lookup)
"location_dim"
1 rows in 0s
1 rows in 0.28s
Designer
Code
Job(FactCrash 0.1)
Run (Job FactCrash)
Context (FactCrash)
Component
Cloud Artifact
Job FactCrash
Basic Run
Debug Run
Advanced settings
Target Exec
Memory Run
Execution
Run
Kill
Clear
[INFO] 17:47:42 finalproject_combine.factcrash_0_1.FactCrash- TalendJob: 'FactCrash' - Started.
[statistics] connecting to socket on port 3816
[statistics] connected
[statistics] disconnected
[INFO] 18:14:36 finalproject_combine.factcrash_0_1.FactCrash- TalendJob: 'FactCrash' - Finished - status: end returnCode: 0
[INFO] 18:14:36 finalproject_combine.factcrash_0_1.FactCrash- TalendJob: 'FactCrash' - Done.
Job FactCrash ended at 18:14 07/04/2024. [Exit code = 0]
Default
Name
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_integration_
Vehicle_curation_Se.
Vehicle_curation_Da
Vehicle_curation_Lo.
Vehicle_curation_Pa.
Vehicle_curation_Por


5526700 rows in 1454.69s
3799.22 rows/s
row1 (Main)
"vehicle_int"
tMap_1
5526700 rows in 1455.54s
3797 rows/s
out1 (Main)
Vehicle_curation
2704 rows in 0.8s
3392.72 rows/s
row2 (Lookup)
"vehicle_dim"
Code
Run (Job FactVehicle)
Context (FactVehicle)
Component
Cloud Artifact
Run
Execution
Run
Kill
Clear
Starting job FactVehicle at 20:41 07/04/2024.
[statistics] connecting to socket on port 3515
[statistics] connected
[statistics] disconnected
Job FactVehicle ended at 21:06 07/04/2024. [Exit code = 0]

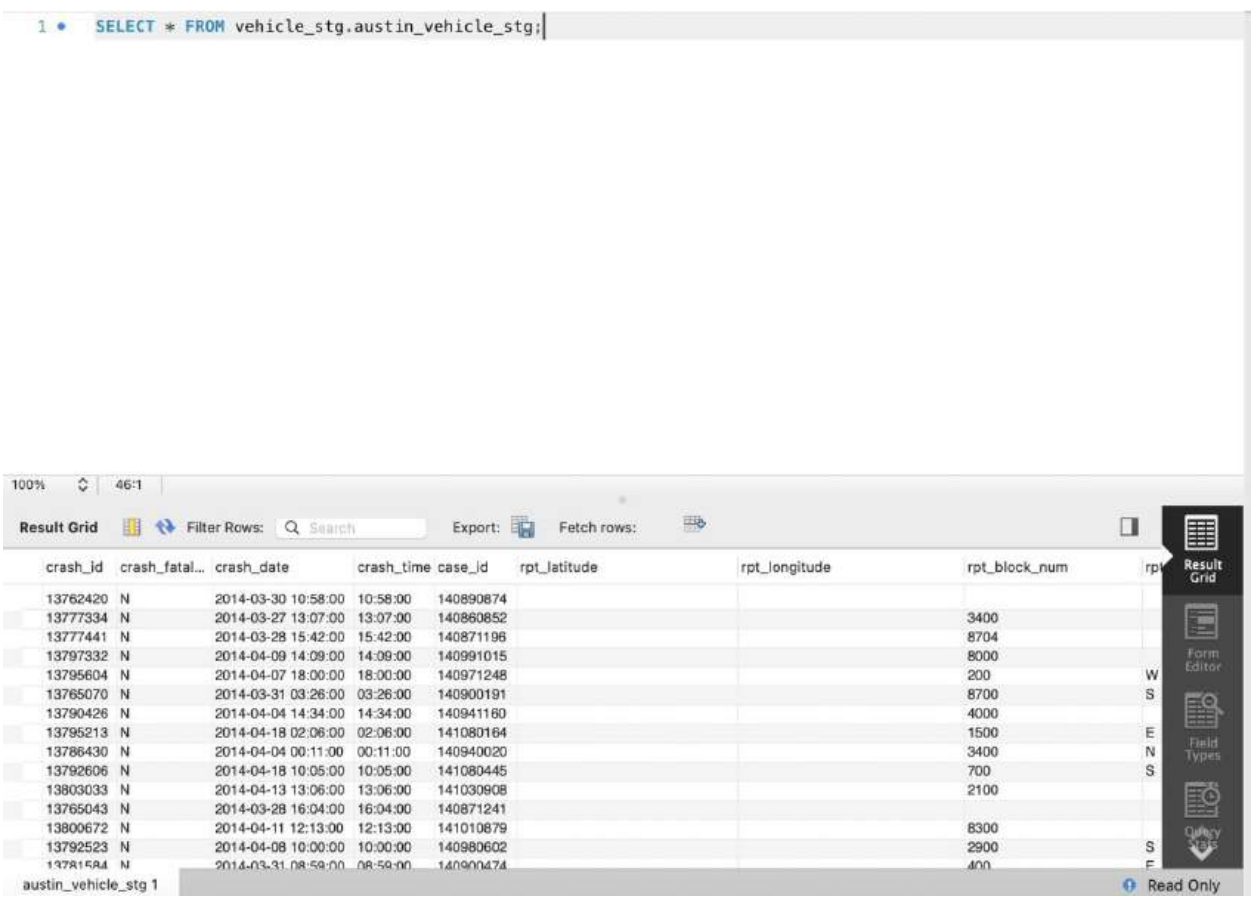
DDL Scripts:

STAGING LAYER:

1. austin_vehicle_stg

```
CREATE TABLE `austin_vehicle_stg` (
  `crash_id` varchar(100) DEFAULT NULL,
  `crash_fatal_fl` char(1) DEFAULT NULL,
  `crash_date` datetime DEFAULT NULL,
  `crash_time` varchar(8) DEFAULT NULL,
  `case_id` varchar(500) DEFAULT NULL,
  `rpt_latitude` varchar(50) DEFAULT NULL,
  `rpt_longitude` varchar(50) DEFAULT NULL,
  `rpt_block_num` varchar(100) DEFAULT NULL,
  `rpt_street_pfx` varchar(100) DEFAULT NULL,
  `rpt_street_name` varchar(200) DEFAULT NULL,
  `rpt_street_sfx` varchar(4) DEFAULT NULL,
  `crash_speed_limit` varchar(2) DEFAULT NULL,
  `road_constr_zone_fl` varchar(1) DEFAULT NULL,
  `latitude` decimal(18,9) DEFAULT NULL,
  `longitude` decimal(18,9) DEFAULT NULL,
  `street_name` varchar(100) DEFAULT NULL,
  `street_nbr` varchar(5) DEFAULT NULL,
  `street_name_2` varchar(100) DEFAULT NULL,
  `street_nbr_2` varchar(100) DEFAULT NULL,
  `crash_sev_id` int DEFAULT NULL,
  `sus_serious_injry_cnt` int DEFAULT NULL,
  `nonincap_injry_cnt` int DEFAULT NULL,
  `poss_injry_cnt` int DEFAULT NULL,
  `non_injry_cnt` int DEFAULT NULL,
  `unkn_injry_cnt` int DEFAULT NULL,
  `tot_injry_cnt` int DEFAULT NULL,
  `death_cnt` int DEFAULT NULL,
  `contrib_factr_p1_id` varchar(2) DEFAULT NULL,
  `contrib_factr_p2_id` varchar(2) DEFAULT NULL,
  `units_involved` varchar(500) DEFAULT NULL,
  `atd_mode_category_metadata` varchar(5000) DEFAULT NULL,
  `pedestrian_fl` varchar(1) DEFAULT NULL,
  `motor_vehicle_fl` char(1) DEFAULT NULL,
  `motorcycle_fl` varchar(1) DEFAULT NULL,
  `bicycle_fl` varchar(1) DEFAULT NULL,
  `other_fl` varchar(1) DEFAULT NULL,
  `point` varchar(45) DEFAULT NULL,
```

```
  `apd_confirmed_fatality` char(1) DEFAULT NULL,
  `apd_confirmed_death_count` int DEFAULT NULL,
  `motor_vehicle_death_count` int DEFAULT NULL,
  `motor_vehicle_serious_injury_count` int DEFAULT NULL,
  `bicycle_death_count` int DEFAULT NULL,
  `bicycle_serious_injury_count` int DEFAULT NULL,
  `pedestrian_death_count` int DEFAULT NULL,
  `pedestrian_serious_injury_count` int DEFAULT NULL,
  `motorcycle_death_count` int DEFAULT NULL,
  `motorcycle_serious_injury_count` int DEFAULT NULL,
  `other_death_count` int DEFAULT NULL,
  `other_serious_injury_count` int DEFAULT NULL,
  `onsys_fl` char(1) DEFAULT NULL,
  `private_dr_fl` char(1) DEFAULT NULL,
  `micromobility_serious_injury_count` int DEFAULT NULL,
  `micromobility_death_count` int DEFAULT NULL,
  `micromobility_fl` varchar(1) DEFAULT NULL,
  `source_system` varchar(20) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```



crash_id	crash_fatal...	crash_date	crash_time	case_id	rpt_latitude	rpt_longitude	rpt_block_num	rp
13762420	N	2014-03-30 10:58:00	10:58:00	140890874				
13777334	N	2014-03-27 13:07:00	13:07:00	140860852			3400	
13777441	N	2014-03-28 15:42:00	15:42:00	140871196			8704	
13797332	N	2014-04-09 14:09:00	14:09:00	140991015			8000	
13795604	N	2014-04-07 18:00:00	18:00:00	140971248			200	W
13765070	N	2014-03-31 03:26:00	03:26:00	140900191			8700	S
13790426	N	2014-04-04 14:34:00	14:34:00	140941160			4000	
13795213	N	2014-04-18 02:06:00	02:06:00	141080164			1500	E
13786430	N	2014-04-04 00:11:00	00:11:00	140940020			3400	N
13792606	N	2014-04-18 10:05:00	10:05:00	141080445			700	S
13803033	N	2014-04-13 13:06:00	13:06:00	141030908			2100	
13765043	N	2014-03-28 16:04:00	16:04:00	140871241				
13800672	N	2014-04-11 12:13:00	12:13:00	141010879			8300	
13792523	N	2014-04-08 10:00:00	10:00:00	140980602			2900	S
13781584	N	2014-03-31 08:59:00	08:59:00	140900474			400	E

```
1 •  SELECT count(*) FROM vehicle_stg.austin_vehicle_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(*)
147750

2. chicago_vehicle_stg

```
CREATE TABLE `chicago_vehicle_stg` (
  `CRASH_RECORD_ID` varchar(300) DEFAULT NULL,
  `CRASH_DATE_EST_I` varchar(1) DEFAULT NULL,
  `CRASH_DATE` datetime DEFAULT NULL,
  `POSTED_SPEED_LIMIT` int DEFAULT NULL,
  `TRAFFIC_CONTROL_DEVICE` varchar(200) DEFAULT NULL,
  `DEVICE_CONDITION` varchar(200) DEFAULT NULL,
  `WEATHER_CONDITION` varchar(100) DEFAULT NULL,
  `LIGHTING_CONDITION` varchar(22) DEFAULT NULL,
  `FIRST_CRASH_TYPE` varchar(28) DEFAULT NULL,
  `TRAFFICWAY_TYPE` varchar(31) DEFAULT NULL,
  `LANE_CNT` varchar(60) DEFAULT NULL,
  `ALIGNMENT` varchar(100) DEFAULT NULL,
  `ROADWAY_SURFACE_COND` varchar(200) DEFAULT NULL,
  `ROAD_DEFECT` varchar(17) DEFAULT NULL,
  `REPORT_TYPE` varchar(26) DEFAULT NULL,
  `CRASH_TYPE` varchar(32) DEFAULT NULL,
  `INTERSECTION_RELATED_I` varchar(1) DEFAULT NULL,
  `NOT_RIGHT_OF_WAY_I` varchar(1) DEFAULT NULL,
  `HIT_AND_RUN_I` varchar(1) DEFAULT NULL,
  `DAMAGE` varchar(13) DEFAULT NULL,
  `DATE_POLICE_NOTIFIED` varchar(22) DEFAULT NULL,
  `PRIM_CONTRIBUTORY_CAUSE` varchar(200) DEFAULT NULL,
  `SEC_CONTRIBUTORY_CAUSE` varchar(80) DEFAULT NULL,
  `STREET_NO` int DEFAULT NULL,
  `STREET_DIRECTION` char(1) DEFAULT NULL,
  `STREET_NAME` varchar(100) DEFAULT NULL,
```

```
  `BEAT_OF_OCCURRENCE` int DEFAULT NULL,
  `PHOTOS_TAKEN_I` varchar(1) DEFAULT NULL,
  `STATEMENTS_TAKEN_I` varchar(1) DEFAULT NULL,
  `DOORING_I` varchar(1) DEFAULT NULL,
  `WORK_ZONE_I` varchar(1) DEFAULT NULL,
  `WORK_ZONE_TYPE` varchar(50) DEFAULT NULL,
  `WORKERS_PRESENT_I` varchar(1) DEFAULT NULL,
  `NUM_UNITS` int DEFAULT NULL,
  `MOST_SEVERE_INJURY` varchar(24) DEFAULT NULL,
  `INJURIES_TOTAL` int DEFAULT NULL,
  `INJURIES_FATAL` int DEFAULT NULL,
  `INJURIES_INCAPACITATING` int DEFAULT NULL,
  `INJURIES_NON_INCAPACITATING` int DEFAULT NULL,
  `INJURIES_REPORTED_NOT_EVIDENT` int DEFAULT NULL,
  `INJURIES_NO_INDICATION` int DEFAULT NULL,
  `INJURIES_UNKNOWN` int DEFAULT NULL,
  `CRASH_HOUR` int DEFAULT NULL,
  `CRASH_DAY_OF_WEEK` int DEFAULT NULL,
  `CRASH_MONTH` int DEFAULT NULL,
  `LATITUDE` decimal(13,9) DEFAULT NULL,
  `LONGITUDE` decimal(13,9) DEFAULT NULL,
  `LOCATION` varchar(40) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_stg.chicago_vehicle_stg;
```

100% 47:1

Result Grid | Filter Rows: Search | Export: | Fetch rows:

CRASH_RECORD_ID	CRASH_DATE_EST_I	CRASH_DATE	POSTED_SPEED_LIMIT	TRAFFIC_CONTROL_DEVICE	DEVICE_C
61fcb8c1eb522a6469b460e2134df3d15f82e81f...		2023-08-18 17:58:00	30	NO CONTROLS	NO CONTR
004cd14d0303a9163aad69a2d7f341b7da2a857...		2019-11-26 08:38:00	25	NO CONTROLS	NO CONTR
a1d5f0ea90897745365a4cbb06cc60329a120d8...		2023-08-18 10:45:00	20	NO CONTROLS	NO CONTR
b236c77d59e32b7b469a6e2f17f438b7457e1bd...		2023-07-29 13:00:00	30	TRAFFIC SIGNAL	FUNCTION
35156ce97cab22747495e92e8bbb16c57e0e60...		2023-02-06 17:30:00	30	NO CONTROLS	NO CONTR
0e208d23344f0d1b3a9fcd4bb07676a750ddb73...		2023-08-13 13:30:00	35	NO CONTROLS	FUNCTION
14386daec6219c6032b71612b28f0e4cd38e289...		2023-08-13 00:11:00	30	TRAFFIC SIGNAL	FUNCTION
359bf9f5872d646bb63576e55b1e0b480dc93c2...		2022-01-31 19:45:00	25	NO CONTROLS	NO CONTR
36360857c079418cba1b1d70cf653595bbfb456...	Y	2022-01-01 16:32:00	10	NO CONTROLS	NO CONTR
9e405b05680e0cb67ca12f910d962fcba7937ee...		2023-09-20 11:00:00	25	NO CONTROLS	NO CONTR
2cb686b41c122f94972f317da3d0f8e41bb56ee4...		2023-10-09 07:15:00	15	NO CONTROLS	NO CONTR
fd05285e9d273fe20cbbebf84794045828a2ba58...		2023-07-29 14:30:00	10	NO CONTROLS	NO CONTR
fda2491d33ac819033f4aaa7ed901120f2f6785b...		2023-07-29 00:50:00	30	TRAFFIC SIGNAL	FUNCTION
436cf5c475779879e882928e0576f67481478de...		2023-09-20 12:35:00	20	NO CONTROLS	NO CONTR

chicago_vehicle_stg 1 | Read Only

Result Grid | Form Editor | Field Types | Query Stats

```
1 •  SELECT count(*) FROM vehicle_stg.chicago_vehicle_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(*)
817723

3. new_york_vehicle_stg

CREATE TABLE `new_york_vehicle_stg` (

```
  `CRASH_DATE` datetime DEFAULT NULL,
  `CRASH_TIME` datetime DEFAULT NULL,
  `BOROUGH` varchar(13) DEFAULT NULL,
  `ZIP_CODE` int DEFAULT NULL,
  `LATITUDE` decimal(30,9) DEFAULT NULL,
  `LONGITUDE` decimal(30,9) DEFAULT NULL,
  `LOCATION` varchar(50) DEFAULT NULL,
  `ON_STREET_NAME` varchar(32) DEFAULT NULL,
  `CROSS_STREET_NAME` varchar(500) DEFAULT NULL,
  `OFF_STREET_NAME` varchar(500) DEFAULT NULL,
  `NUMBER_OF_PERSONS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PERSONS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_KILLED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_KILLED` int DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_1` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_2` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_3` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_4` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_5` varchar(100) DEFAULT NULL,
  `COLLISION_ID` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_1` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_2` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_3` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_4` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_5` varchar(100) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle.new_york_vehicle_stg;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

CRASH_DATE	CRASH_TIME	BOROUGH	ZIP_CODE	LATITUDE	LONGITUDE	LOCATION	ON_STREET_NAME	C
2021-09-11 00:00:00	1970-01-01 02:00:00		-99	NULL	NULL		WHITESTONE EXPRESSWAY	2
2022-03-26 00:00:00	1970-01-01 11:00:00		-99	NULL	NULL		QUEENSBORO BRIDGE UPPER	
2022-06-29 00:00:00	1970-01-01 06:00:00		-99	NULL	NULL		THROGS NECK BRIDGE	
2021-09-11 00:00:00	1970-01-01 09:00:00	BROOKLYN	11208	40.667202000	-73.866500000	(40.667202, -73.8665)		
2021-12-14 00:00:00	1970-01-01 08:00:00	BROOKLYN	11233	40.683304000	-73.917274000	(40.683304, -73.917274)	SARATOGA AVENUE	D
2021-04-14 00:00:00	1970-01-01 12:00:00		-99	NULL	NULL		MAJOR DEEGAN EXPRESSWAY RAMP	
2021-12-14 00:00:00	1970-01-01 17:00:00		-99	40.709183000	-73.956825000	(40.709183, -73.956825)	BROOKLYN QUEENS EXPRESSWAY	
2021-12-14 00:00:00	1970-01-01 08:00:00	BRONX	10475	40.868160000	-73.831480000	(40.86816, -73.83148)		
2021-12-14 00:00:00	1970-01-01 21:00:00	BROOKLYN	11207	40.671720000	-73.897100000	(40.67172, -73.8971)		
2021-12-14 00:00:00	1970-01-01 14:00:00	MANHATTAN	10017	40.751440000	-73.973970000	(40.75144, -73.97397)	3 AVENUE	E
2021-12-13 00:00:00	1970-01-01 00:00:00		-99	40.701275000	-73.888870000	(40.701275, -73.88887)	MYRTLE AVENUE	
2021-12-14 00:00:00	1970-01-01 16:00:00	QUEENS	11413	40.675884000	-73.755770000	(40.675884, -73.75577)	SPRINGFIELD BOULEVARD	E
2021-12-14 00:00:00	1970-01-01 08:00:00		-99	NULL	NULL		broadway	w
2021-12-14 00:00:00	1970-01-01 00:00:00		-99	40.596620000	-74.002310000	(40.59662, -74.00231)	BELT PARKWAY	
2021-12-14 00:00:00	1970-01-01 23:00:00	QUEENS	11434	40.666840000	-73.789410000	(40.66684, -73.78941)	NORTH CONDUIT AVENUE	1

new_york_vehicle_stg 1 Read Only

Result Grid Form Editor Field Types Query Stats

```
1 •  SELECT count(*) FROM vehicle_stg.new_york_vehicle_stg;
```

Result Grid Filter Rows: Export: Wrap Cell Content:

count(*)
2075427

4. vehicle_austin_stg

CREATE TABLE `vehicle_austin_stg` (

```
  `crash_id` varchar(100) NOT NULL,
  `units_involved` varchar(500) DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```sql
SELECT * FROM vehicle_stg.vehicle_austin_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell C

crash_id	units_involved
13762420	Passenger car
13777334	Large passenger vehicle
13777441	Large passenger vehicle
13797332	Motor vehicle – other
13795604	Passenger car
13765070	Passenger car
13790426	Passenger car
13795213	Large passenger vehicle
13786430	Passenger car
13792606	Passenger car
13792606	Passenger car
13803033	Large passenger vehicle
13803033	Large passenger vehicle
13765043	Large passenger vehicle
13765043	Passenger car
13800672	Passenger car
13800672	Large passenger vehicle
13800672	Passenger car
13792523	Passenger car
13792523	Large passenger vehicle
13791584	Large passenger vehicle

```
SELECT count(*) FROM vehicle_stg.vehicle_austin_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(*)
305032

5. **vehicle_new_york_stg**

```
CREATE TABLE `new_york_vehicle_stg` (
  `CRASH_DATE` datetime DEFAULT NULL,
  `CRASH_TIME` datetime DEFAULT NULL,
  `BOROUGH` varchar(13) DEFAULT NULL,
  `ZIP_CODE` int DEFAULT NULL,
```

```
  `LATITUDE` decimal(30,9) DEFAULT NULL,
  `LONGITUDE` decimal(30,9) DEFAULT NULL,
  `LOCATION` varchar(50) DEFAULT NULL,
  `ON_STREET_NAME` varchar(32) DEFAULT NULL,
  `CROSS_STREET_NAME` varchar(500) DEFAULT NULL,
  `OFF_STREET_NAME` varchar(500) DEFAULT NULL,
  `NUMBER_OF_PERSONS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PERSONS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_KILLED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_KILLED` int DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_1` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_2` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_3` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_4` varchar(100) DEFAULT NULL,
  `CONTRIBUTING_FACTOR_VEHICLE_5` varchar(100) DEFAULT NULL,
  `COLLISION_ID` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_1` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_2` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_3` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_4` varchar(100) DEFAULT NULL,
  `VEHICLE_TYPE_CODE_5` varchar(100) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_stg.vehicle_new_york_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fet

COLLISION_ID	VEHICLE_TYPE_CODE
	NULL
100	PASSENGER VEHICLE
100	NULL
1000	PASSENGER VEHICLE
1000	UNKNOWN
1000	NULL
10000	TAXI
10000	NULL
100000	PASSENGER VEHICLE
100000	LARGE COM VEH(6 OR MORE TIRES)
100000	NULL
1000000	PASSENGER VEHICLE
1000000	NULL
1000001	PASSENGER VEHICLE
1000001	NULL
1000002	PASSENGER VEHICLE
1000002	NULL
1000003	SPORT UTILITY / STATION WAGON
1000003	NULL
1000004	SPORT UTILITY / STATION WAGON

```
1 •  SELECT count(*) FROM vehicle_stg.vehicle_new_york_stg;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	count(*)
▶	5221668

INTEGRATION LAYER:

1. austin_vehicle_int

```
CREATE TABLE `austin_vehicle_int` (
  `crash_id` varchar(100) NOT NULL,
  `latitude` decimal(18,0) DEFAULT NULL,
  `longitude` decimal(18,0) DEFAULT NULL,
  `JUST_INJURED_FL` varchar(18) DEFAULT NULL,
  `PEDESTRIAN_INVOLVED_FL` varchar(18) DEFAULT NULL,
  `SEASON` varchar(200) DEFAULT NULL,
  `MOTORIST_INVOLVED_COUNT` int DEFAULT NULL,
  `TIME_OF_DAY` varchar(20) DEFAULT NULL,
  `DAY_OF_WEEK` varchar(20) DEFAULT NULL,
  `WEEKDAY_OR_WEEKEND` varchar(20) DEFAULT NULL,
  `CRASH_FATAL_FL` char(1) DEFAULT NULL,
  `PEDESTRIAN_KILLED_FL` varchar(5) DEFAULT NULL,
  `street_nbr` varchar(5) DEFAULT NULL,
  `street_name` varchar(100) DEFAULT NULL,
  `crash_date` datetime DEFAULT NULL,
  `crash_time` varchar(8) DEFAULT NULL,
  `contrib_factr_p1_id` varchar(2) DEFAULT NULL,
  `contrib_factr_p2_id` varchar(2) DEFAULT NULL,
  `units_involved` varchar(500) DEFAULT NULL,
  `source_system` varchar(20) DEFAULT NULL,
  `Pedestrian_most_killed` int DEFAULT NULL,
  `Total_death_count` int DEFAULT NULL,
  `Total_injury_count` int DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_integration.austin_vehicle_int;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

crash_id	latitude	longitude	JUST_INJURED_FL	PEDESTRIAN_INVOLVED_FL	SEASON	MOTORIST_INVOLVED_COUNT	TIME_OF_DAY	DAY_OF_WEEK	WEEKDA
13762420	NULL	NULL	N	N	Winter	0	Morning	Sunday	Weekend
13777334	30	-98	N	N	Winter	0	Afternoon	Thursday	Weekday
13777441	30	-98	N	N	Winter	0	Afternoon	Friday	Weekday
13797332	30	-98	N	N	Spring	0	Afternoon	Wednesday	Weekday
13795604	30	-98	N	N	Spring	0	Night	Monday	Weekday
13765070	30	-98	Y	N	Winter	0	Night	Monday	Weekday
13790426	30	-98	N	N	Spring	0	Afternoon	Friday	Weekday
13795213	30	-98	Y	N	Spring	0	Night	Friday	Weekday
13786430	30	-98	N	N	Spring	0	Night	Friday	Weekday
13792606	30	-98	N	N	Spring	0	Morning	Friday	Weekday
13803033	30	-98	Y	N	Spring	0	Afternoon	Sunday	Weekend
13765043	30	-98	N	N	Winter	0	Afternoon	Friday	Weekday
13800672	30	-98	Y	N	Spring	0	Afternoon	Friday	Weekday
13792523	30	-98	Y	N	Spring	0	Morning	Tuesday	Weekday
13781584	30	-98	Y	N	Winter	0	Morning	Monday	Weekday

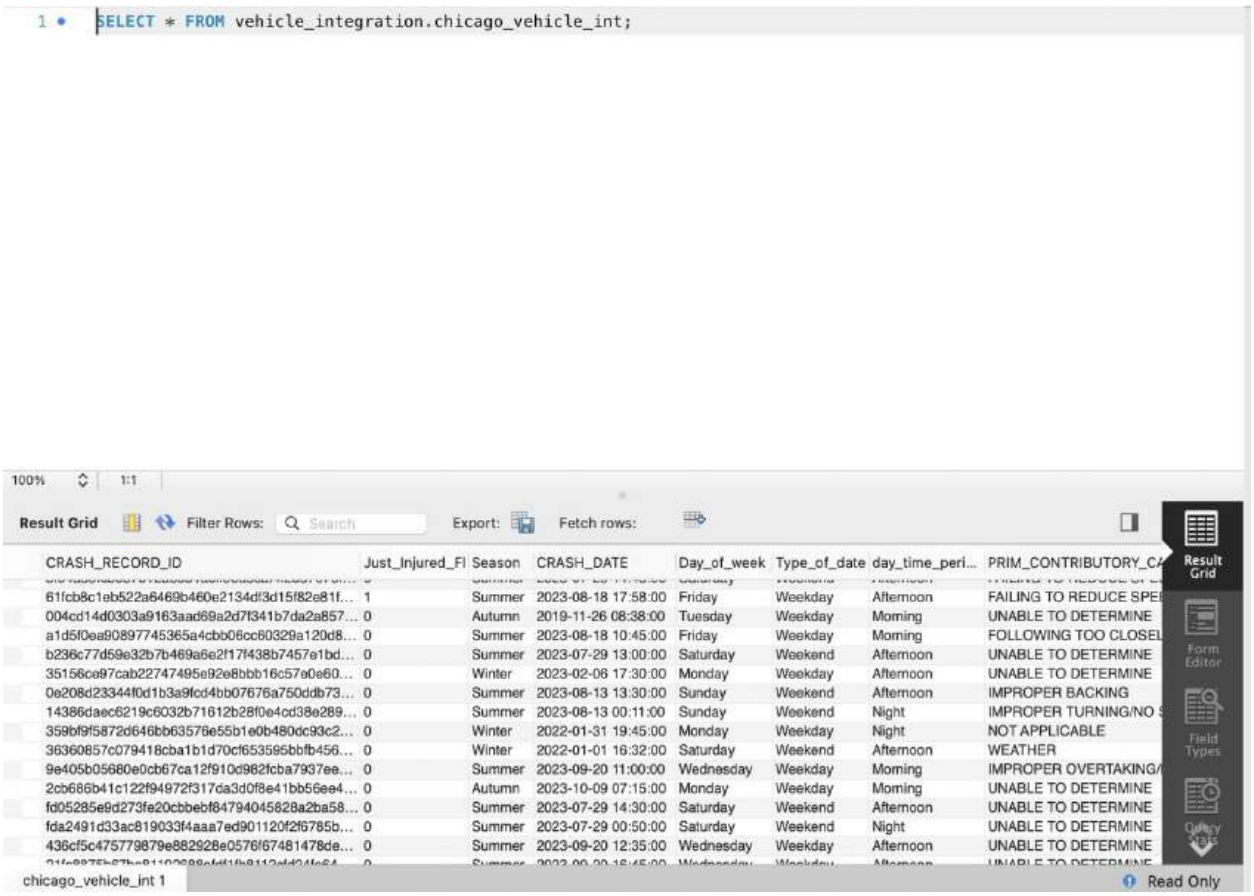
Result Grid Form Editor Field Types Query Stats

austin_vehicle_int 1 Read Only

2. chicago_vehicle_int

```
CREATE TABLE `austin_vehicle_int` (
  `crash_id` varchar(100) NOT NULL,
  `latitude` decimal(18,0) DEFAULT NULL,
  `longitude` decimal(18,0) DEFAULT NULL,
  `JUST_INJURED_FL` varchar(18) DEFAULT NULL,
  `PEDESTRIAN_INVOLVED_FL` varchar(18) DEFAULT NULL,
  `SEASON` varchar(200) DEFAULT NULL,
  `MOTORIST_INVOLVED_COUNT` int DEFAULT NULL,
  `TIME_OF_DAY` varchar(20) DEFAULT NULL,
  `DAY_OF_WEEK` varchar(20) DEFAULT NULL,
  `WEEKDAY_OR_WEEKEND` varchar(20) DEFAULT NULL,
  `CRASH_FATAL_FL` char(1) DEFAULT NULL,
  `PEDESTRIAN_KILLED_FL` varchar(5) DEFAULT NULL,
  `street_nbr` varchar(5) DEFAULT NULL,
  `street_name` varchar(100) DEFAULT NULL,
  `crash_date` datetime DEFAULT NULL,
```

```
  `crash_time` varchar(8) DEFAULT NULL,
  `contrib_factr_p1_id` varchar(2) DEFAULT NULL,
  `contrib_factr_p2_id` varchar(2) DEFAULT NULL,
  `units_involved` varchar(500) DEFAULT NULL,
  `source_system` varchar(20) DEFAULT NULL,
  `Pedestrian_most_killed` int DEFAULT NULL,
  `Total_death_count` int DEFAULT NULL,
  `Total_injury_count` int DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

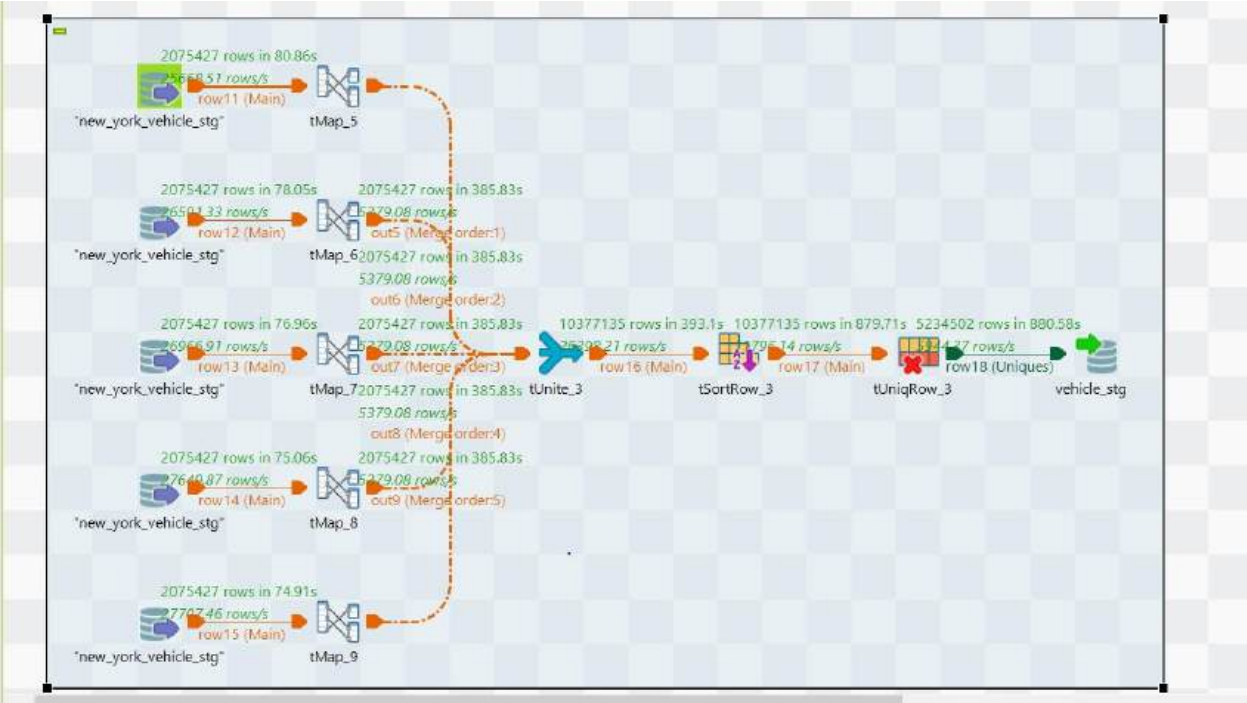


CRASH_RECORD_ID	Just_Injured_Fl	Season	CRASH_DATE	Day_of_week	Type_of_date	day_time_peri...	PRIM_CONTRIBUTORY_CA
61fcb8c1eb522a6469b460e2134df3d15f82e81f...	1	Summer	2023-08-18 17:58:00	Friday	Weekday	Afternoon	FAILING TO REDUCE SPE
004cd14d0303a9163aad69a2d7f341b7da2a857...	0	Autumn	2019-11-26 08:38:00	Tuesday	Weekday	Morning	UNABLE TO DETERMINE
a1d5f0ea90897745365a4cbb06cc60329a120d8...	0	Summer	2023-08-18 10:45:00	Friday	Weekday	Morning	FOLLOWING TOO CLOSEL
b236c77d59e32b7b469a6e2f17f438b7457e1bd...	0	Summer	2023-07-29 13:00:00	Saturday	Weekend	Afternoon	UNABLE TO DETERMINE
35156ce97cab22747495e92e8bbb16c57e0e60...	0	Winter	2023-02-06 17:30:00	Monday	Weekday	Afternoon	UNABLE TO DETERMINE
0e208d23344f0d1b3a9fcd4bb07676a750ddb73...	0	Summer	2023-08-13 13:30:00	Sunday	Weekend	Afternoon	IMPROPER BACKING
14386daec6219c6032b71612b28f0e4cd38e289...	0	Summer	2023-08-13 00:11:00	Sunday	Weekend	Night	IMPROPER TURNING/NO S
359bf9f5872d646bb63576e55b1e0b480dc93c2...	0	Winter	2022-01-31 19:45:00	Monday	Weekday	Night	NOT APPLICABLE
36360857c079418cba1b1d70cf653595bbfb456...	0	Winter	2022-01-01 16:32:00	Saturday	Weekend	Afternoon	WEATHER
9e405b05680e0cb67ca12f910d982fcba7937ee...	0	Summer	2023-09-20 11:00:00	Wednesday	Weekday	Morning	IMPROPER OVERTAKING/
2cb686b41c122f94972f317da3d0f8e41bb56ee4...	0	Autumn	2023-10-09 07:15:00	Monday	Weekday	Morning	UNABLE TO DETERMINE
fd05285e9d273fe20cbbebf84794045828a2ba58...	0	Summer	2023-07-29 14:30:00	Saturday	Weekend	Afternoon	UNABLE TO DETERMINE
fda2491d33ac819033f4aaa7ed901120f2f6785b...	0	Summer	2023-07-29 00:50:00	Saturday	Weekend	Night	UNABLE TO DETERMINE
436cf5c475779879e882928e0576f67481478de...	0	Summer	2023-09-20 12:35:00	Wednesday	Weekday	Afternoon	UNABLE TO DETERMINE

3. new_york_vehicle_int

```
 CREATE TABLE `new_york_vehicle_int` (
  `CRASH_ID` varchar(20) DEFAULT NULL,
  `LATITUDE` decimal(30,9) DEFAULT NULL,
  `LONGITUDE` decimal(30,9) DEFAULT NULL,
  `ON_STREET_NAME` varchar(32) DEFAULT NULL,
  `ZIP_CODE` int DEFAULT NULL,
```

```
  `NUMBER_OF_PERSONS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PERSONS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_INJURED` int DEFAULT NULL,
  `NUMBER_OF_PEDESTRIANS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_CYCLIST_KILLED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_INJURED` int DEFAULT NULL,
  `NUMBER_OF_MOTORIST_KILLED` int DEFAULT NULL,
  `Season` varchar(19) DEFAULT NULL,
  `day_in_time` varchar(19) DEFAULT NULL,
  `Type_of_date` varchar(19) DEFAULT NULL,
  `Day_of_week` varchar(19) DEFAULT NULL,
  `CRASH_DATE` datetime DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```



```
1 • SELECT * FROM vehicle_integration.new_york_vehicle_int;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

CRASH_ID	LATITUDE	LONGITUDE	ON_STREET_NAME	ZIP_CODE	NUMBER_OF_PERSONS_INJURED	NUMBER_OF_PERSONS_KILLED
4455765	NULL	NULL	WHITESTONE EXPRESSWAY	-99	2	0
4513547	NULL	NULL	QUEENSBORO BRIDGE UPPER	-99	1	0
4541903	NULL	NULL	THROGS NECK BRIDGE	-99	0	0
4456314	40.667202000	-73.866500000		11208	0	0
4486609	40.683304000	-73.917274000	SARATOGA AVENUE	11233	0	0
4407458	NULL	NULL	MAJOR DEEGAN EXPRESSWAY RAMP	-99	0	0
4486555	40.709183000	-73.956825000	BROOKLYN QUEENS EXPRESSWAY	-99	0	0
4486660	40.868160000	-73.831480000		10475	2	0
4487074	40.671720000	-73.897100000		11207	0	0
4486519	40.751440000	-73.973970000	3 AVENUE	10017	0	0
4486934	40.701275000	-73.888870000	MYRTLE AVENUE	-99	0	0
4487127	40.675884000	-73.755770000	SPRINGFIELD BOULEVARD	11413	0	0
4486634	NULL	NULL	broadway	-99	0	0
4486564	40.596620000	-74.002310000	BELT PARKWAY	-99	0	0
4486635	40.666840000	-73.789410000	NORTH CONDUIT AVENUE	11434	2	0

Result Grid Form Editor Field Types Query Stats

new_york_vehicle_int 1 Read Only

4. after_int_stg

```
CREATE TABLE `after_int_stg` (
`CRASH_ID` varchar(400) DEFAULT NULL,
`LATITUDE` decimal(30,9) DEFAULT NULL,
`LONGITUDE` decimal(30,9) DEFAULT NULL,
`ON_STREET_NAME` varchar(32) DEFAULT NULL,
`ZIP_CODE` int DEFAULT NULL,
`Total_injured` int DEFAULT NULL,
`Total_killed` int DEFAULT NULL,
`PEDESTRIAN_INVOLVED` int DEFAULT NULL,
`MOTORISTS_KILLED_INJURED` int DEFAULT NULL,
`Season` varchar(19) DEFAULT NULL,
`day_in_time` varchar(19) DEFAULT NULL,
`Type_of_date` varchar(19) DEFAULT NULL,
`Day_of_week` varchar(10) DEFAULT NULL,
`CRASH_DATE` datetime DEFAULT NULL,
`NUMBER_OF_PEDESTRIANS_INJURED` int DEFAULT NULL,
`NUMBER_OF_PEDESTRIANS_KILLED` int DEFAULT NULL,
```

```
  `NUMBER_OF_PERSONS_KILLED` int DEFAULT NULL,
  `NUMBER_OF_PERSONS_INJURED` int DEFAULT NULL,
  `PEDESTRIAN_KILLED_MOST` int DEFAULT NULL,
  `Just_Injured_Fl` int DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_integration.new_york_vehicle_int;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

CRASH_ID	LATITUDE	LONGITUDE	ON_STREET_NAME	ZIP_CODE	NUMBER_OF_PERSONS_INJURED	NUMBER_OF_PERSONS_KILLED
4455765	NULL	NULL	WHITESTONE EXPRESSWAY	-99	2	0
4513547	NULL	NULL	QUEENSBORO BRIDGE UPPER	-99	1	0
4541903	NULL	NULL	THROGS NECK BRIDGE	-99	0	0
4456314	40.667202000	-73.866500000		11208	0	0
4486609	40.683304000	-73.917274000	SARATOGA AVENUE	11233	0	0
4407458	NULL	NULL	MAJOR DEEGAN EXPRESSWAY RAMP	-99	0	0
4486555	40.709183000	-73.956825000	BROOKLYN QUEENS EXPRESSWAY	-99	0	0
4486660	40.868160000	-73.831480000		10475	2	0
4487074	40.671720000	-73.897100000		11207	0	0
4486519	40.751440000	-73.973970000	3 AVENUE	10017	0	0
4486934	40.701275000	-73.888870000	MYRTLE AVENUE	-99	0	0
4487127	40.675884000	-73.755770000	SPRINGFIELD BOULEVARD	11413	0	0
4486634	NULL	NULL	broadway	-99	0	0
4486564	40.596620000	-74.002310000	BELT PARKWAY	-99	0	0
4486635	40.666840000	-73.789410000	NORTH CONDUIT AVENUE	11434	2	0

Result Grid Form Editor Field Types Query Stats

new_york_vehicle_int 1 Read Only

CURATION LAYER:

DIMENSION

1. DimDate

```
CREATE TABLE `DimDate` (
  `Date_id` int NOT NULL,
  `CRASH_DATE` datetime DEFAULT NULL,
  `Day_of_week` varchar(10) DEFAULT NULL,
  `Month` varchar(20) DEFAULT NULL,
  `Year` varchar(20) DEFAULT NULL,
  `Weekday_Or_Weekend` varchar(19) DEFAULT NULL,
  `Season` varchar(19) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_curation.DimDate;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

Date_id	CRASH_DATE	Day_of_week	Month	Year	Weekday_Or_Weekend	Season	source_system	DI_CreateDate
1	2021-09-11 00:00:00	Saturday	September	2021	Weekend	Summer	New_York	2024-04-07 16:17:59
2	2022-03-26 00:00:00	Saturday	March	2022	Weekend	Winter	New_York	2024-04-07 16:17:59
3	2022-06-29 00:00:00	Wednesday	June	2022	Weekday	Spring	New_York	2024-04-07 16:17:59
4	2021-12-14 00:00:00	Tuesday	December	2021	Weekday	Autumn	New_York	2024-04-07 16:17:59
5	2021-04-14 00:00:00	Wednesday	April	2021	Weekday	Spring	New_York	2024-04-07 16:17:59
6	2021-12-13 00:00:00	Monday	December	2021	Weekday	Autumn	New_York	2024-04-07 16:17:59
7	2021-12-11 00:00:00	Saturday	December	2021	Weekend	Autumn	New_York	2024-04-07 16:17:59
8	2021-12-12 00:00:00	Sunday	December	2021	Weekend	Autumn	New_York	2024-04-07 16:17:59
9	2021-12-16 00:00:00	Thursday	December	2021	Weekday	Autumn	New_York	2024-04-07 16:17:59
10	2021-04-15 00:00:00	Thursday	April	2021	Weekday	Spring	New_York	2024-04-07 16:17:59
11	2021-07-07 00:00:00	Wednesday	July	2021	Weekday	Summer	New_York	2024-04-07 16:17:59
12	2022-07-12 00:00:00	Tuesday	July	2022	Weekday	Summer	New_York	2024-04-07 16:17:59
13	2022-03-23 00:00:00	Wednesday	March	2022	Weekday	Winter	New_York	2024-04-07 16:17:59
14	2021-07-09 00:00:00	Friday	July	2021	Weekday	Summer	New_York	2024-04-07 16:17:59
15	2022-04-24 00:00:00	Sunday	April	2022	Weekend	Spring	New_York	2024-04-07 16:17:59

Result Grid Form Editor Field Types Query Stats

DimDate 1 Read Only

2. DimLocation

```
CREATE TABLE `DimLocation` (
  `Loc_id` int NOT NULL,
  `LATITUDE` decimal(30,9) DEFAULT NULL,
  `LONGITUDE` decimal(30,9) DEFAULT NULL,
  `STREET_NAME` varchar(32) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
SELECT * FROM vehicle_curation.DimLocation;
```

100% 1:1

Result Grid Filter Rows: Search Export: Fetch rows:

Loc_id	LATITUDE	LONGITUDE	STREET_NAME	source_system	DI_CreateDate
1	-200.000000000	-200.000000000	WHITESTONE EXPRESSWAY	New_York	2024-04-07 16:17:59
2	-200.000000000	-200.000000000	QUEENSBORO BRIDGE UPPER	New_York	2024-04-07 16:17:59
3	-200.000000000	-200.000000000	THROGS NECK BRIDGE	New_York	2024-04-07 16:17:59
4	40.667202000	-73.866500000		New_York	2024-04-07 16:17:59
5	40.683304000	-73.917274000	SARATOGA AVENUE	New_York	2024-04-07 16:17:59
6	-200.000000000	-200.000000000	MAJOR DEEGAN EXPRESSWAY RAMP	New_York	2024-04-07 16:17:59
7	40.709183000	-73.956825000	BROOKLYN QUEENS EXPRESSWAY	New_York	2024-04-07 16:17:59
8	40.868160000	-73.831480000		New_York	2024-04-07 16:17:59
9	40.671720000	-73.897100000		New_York	2024-04-07 16:17:59
10	40.751440000	-73.973970000	3 AVENUE	New_York	2024-04-07 16:17:59
11	40.701275000	-73.888870000	MYRTLE AVENUE	New_York	2024-04-07 16:17:59
12	40.675884000	-73.755770000	SPRINGFIELD BOULEVARD	New_York	2024-04-07 16:17:59
13	-200.000000000	-200.000000000	broadway	New_York	2024-04-07 16:17:59
14	40.596620000	-74.002310000	BELT PARKWAY	New_York	2024-04-07 16:17:59
15	40.666840000	-73.789410000	NORTH CONDUIT AVENUE	New_York	2024-04-07 16:17:59

Result Grid Form Editor Field Types Query Stats

DimLocation 1 Read Only

3. DimTime

```
CREATE TABLE `DimTime` (
  `Time_Of_Day` varchar(30) DEFAULT NULL,
  `source_system` varchar(30) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  `Time_id` int DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```sql
SELECT * FROM vehicle_curation.time_dim;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Time_Of_Day	DI_CreateDate	Time_id
Morning	2024-04-07 16:57:52	1
Afternoon	2024-04-07 16:57:52	2
Night	2024-04-07 16:57:52	3

```
1 •  SELECT COUNT(*) FROM vehicle_curation.time_dim;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	COUNT(*)
▶	3

4. vehicle_dim

```
CREATE TABLE  vehicle_dim  (
  Vehicle_Id  int NOT NULL,
  Vehicle_Type  varchar(5000) DEFAULT NULL,
  source_system  varchar(10) DEFAULT NULL,
  DI_CreateDate  datetime DEFAULT NULL,
  PRIMARY KEY ( Vehicle_Id )
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_curation.vehicle_dim;
```

Result Grid | Filter Rows: | Edit: | Export/Import:

Vehicle_Id	Vehicle_Type	source_system	DI_CreateDate
1	Passenger car	Austin	2024-04-07 20:33:39
2	Large passenger vehicle	Austin	2024-04-07 20:33:39
4	Motor vehicle – other	Austin	2024-04-07 20:33:39
46	Motorcycle	Austin	2024-04-07 20:33:39
50	Other/Unknown	Austin	2024-04-07 20:33:39
54	Pedestrian	Austin	2024-04-07 20:33:39
125	Bicycle	Austin	2024-04-07 20:33:39
2102	Train	Austin	2024-04-07 20:33:39
23307	Micromobility device	Austin	2024-04-07 20:33:39
111766	E-scooter	Austin	2024-04-07 20:33:39
305034	PASSENGER VEHICLE	New_York	2024-04-07 20:33:39
305037	UNKNOWN	New_York	2024-04-07 20:33:39
305039	TAXI	New_York	2024-04-07 20:33:39
305042	LARGE COM VEH(6 OR ...	New_York	2024-04-07 20:33:39
305050	SPORT UTILITY / STATI...	New_York	2024-04-07 20:33:39
305062	SMALL COM VEH(4 TIR...	New_York	2024-04-07 20:33:39
305075	OTHER	New_York	2024-04-07 20:33:39
305116	LIVERY VEHICLE	New_York	2024-04-07 20:33:39
305129	BICYCLE	New_York	2024-04-07 20:33:39
305150	PICK-UP TRUCK	New_York	2024-04-07 20:33:39
[illegible]	VAN	New_York	[illegible]

```
1 •  SELECT COUNT(*) FROM vehicle_curation.vehicle_dim;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	COUNT(*)
▶	2704

5. SCD_Contribution_dim

```
CREATE TABLE  scd_contribution_dim  (
   SK_Contri  int NOT NULL AUTO_INCREMENT
  Description  varchar(100) DEFAULT NULL,
  Code  varchar(5) NOT NULL,
  scd_start  datetime NOT NULL,
  scd_end  datetime DEFAULT NULL,
  scd_version  int NOT NULL,
  scd_active  bit(10) NOT NULL,
  PRIMARY KEY ( SK_Contri )
) ENGINE=InnoDB AUTO_INCREMENT=89 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM `vehicle_curation`.`scd_contribution_dim`;
2
```

Result Grid | Filter Rows: | Edit: | Export/Import: | Wrap Cell Content:

Description	Code	SK_Contri	scd_start	scd_end	scd_version	scd_active
Animal on Road - Domestic	1	1	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Animal on Road - Wild	2	2	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Backed without Safety	3	3	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Changed Lane when Unsafe	4	4	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Disabled in Traffic Lane	14	5	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Disregard Stop and Go Signal	15	6	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Disregard Stop Sign or Light	16	7	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Disregard Turn Marks at Intersection	17	8	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Disregard Warning Sign at Construction	18	9	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Distraction in Vehicle	19	10	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Driver Inattention	20	11	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Drove Without Headlights	21	12	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Control Speed	22	13	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Drive in Single Lane	23	14	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Give Half of Roadway	24	15	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Heed Warning Sign or Traffic C...	25	16	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Pass to Left Safely	26	17	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Pass to Right Safely	27	18	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Signal or Gave Wrong Signal	28	19	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Stop at Proper Place	29	20	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1
Failed to Stop for School Bus	30	21	2024-04-07 21:47:29	9999-01-01 12:00:00	1	1

scd_contribution_dim 2

```
1 •  SELECT COUNT(*) FROM vehicle_curation.scd_contribution_dim;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

COUNT(*)
88

Result 2 ×

FACTS

```
CREATE TABLE  fact_crash  (
  Accident_id  int NOT NULL,
  CRASH_ID  varchar(300) DEFAULT NULL,
  Date_id  int NOT NULL,
  Loc_id  int NOT NULL,
  Time_id  int DEFAULT NULL,
  DI_CreateDate  datetime DEFAULT NULL,
  Count_of_Motorist_involved  int DEFAULT NULL,
  Total_death  int DEFAULT NULL,
  source_system  varchar(30) DEFAULT NULL,
  PEDESTRIAN_KILLED_MOST  int DEFAULT NULL,
  Just_Injured_FI  int DEFAULT NULL,
  Total_injured  int DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehicle_curation.fact_crash;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content: | Fetch rows:

Accident_id	CRASH_ID	Date_id	Loc_id	Time_id	DI_CreateDate	Count_of_Motorist_involved	Total_death	source_system	PEDESTRIAN_KILLED
1	4455765	1	0	3	2024-04-07 17:48:00	2	0	New_York	0
2	4513547	24	0	1	2024-04-07 17:48:00	1	0	New_York	0
3	4541903	288	0	1	2024-04-07 17:48:00	0	0	New_York	0
4	4456314	1	1	1	2024-04-07 17:48:00	0	0	New_York	0
5	4486609	2	2	1	2024-04-07 17:48:00	0	0	New_York	0
6	4407458	70	0	2	2024-04-07 17:48:00	0	0	New_York	0
7	4486555	2	3	2	2024-04-07 17:48:00	0	0	New_York	0
8	4486660	2	4	1	2024-04-07 17:48:00	2	0	New_York	0
9	4487074	2	5	3	2024-04-07 17:48:00	0	0	New_York	0
10	4486519	2	6	2	2024-04-07 17:48:00	0	0	New_York	0
11	4486934	3	7	3	2024-04-07 17:48:00	0	0	New_York	0
12	4487127	2	8	2	2024-04-07 17:48:00	0	0	New_York	0
13	4486634	2	0	1	2024-04-07 17:48:00	0	0	New_York	0
14	4486564	2	9	3	2024-04-07 17:48:00	0	0	New_York	0
15	4486635	2	10	3	2024-04-07 17:48:00	2	0	New_York	0
16	4486604	2	11	2	2024-04-07 17:48:00	0	0	New_York	0
17	4486991	2	12	3	2024-04-07 17:48:00	4	0	New_York	0
18	4486284	2	0	3	2024-04-07 17:48:00	3	0	New_York	0
19	4487040	4	13	3	2024-04-07 17:48:00	1	0	New_York	0

```
1 •  SELECT COUNT(*) FROM vehicle_curation.fact_crash;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

COUNT(*)
3040900

```
CREATE TABLE `fact_vehicle` (
  `Vehicle_Id` int NOT NULL,
  `VF_Id` int NOT NULL,
  `Crash_Id` varchar(100) NOT NULL,
  `Source_System` varchar(10) DEFAULT NULL,
  `DI_CreateDate` datetime DEFAULT NULL,
  PRIMARY KEY (`VF_Id`)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```sql
SELECT * FROM vehicle_curation.fact_vehicle;
```

Result Grid | Filter Rows: | Edit: | Export/Import:

Vehicle_Id	VF_Id	Crash_Id	Source_System	DI_CreateDate
1	1	13762420	Austin	2024-04-07 20:41:56
2	2	13777334	Austin	2024-04-07 20:41:56
2	3	13777441	Austin	2024-04-07 20:41:56
4	4	13797332	Austin	2024-04-07 20:41:56
1	5	13795604	Austin	2024-04-07 20:41:56
1	6	13765070	Austin	2024-04-07 20:41:56
1	7	13790426	Austin	2024-04-07 20:41:56
2	8	13795213	Austin	2024-04-07 20:41:56
1	9	13786430	Austin	2024-04-07 20:41:56
1	10	13792606	Austin	2024-04-07 20:41:56
1	11	13792606	Austin	2024-04-07 20:41:56
2	12	13803033	Austin	2024-04-07 20:41:56
2	13	13803033	Austin	2024-04-07 20:41:56
2	14	13765043	Austin	2024-04-07 20:41:56
1	15	13765043	Austin	2024-04-07 20:41:56
1	16	13800672	Austin	2024-04-07 20:41:56
2	17	13800672	Austin	2024-04-07 20:41:56
1	18	13800672	Austin	2024-04-07 20:41:56
1	19	13792523	Austin	2024-04-07 20:41:56
2	20	13792523	Austin	2024-04-07 20:41:56
2	21	13781584	Austin	2024-04-07 20:41:56

```
1 •  SELECT COUNT(*) FROM vehicle_curation.fact_vehicle;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

	COUNT(*)
▶	5526700

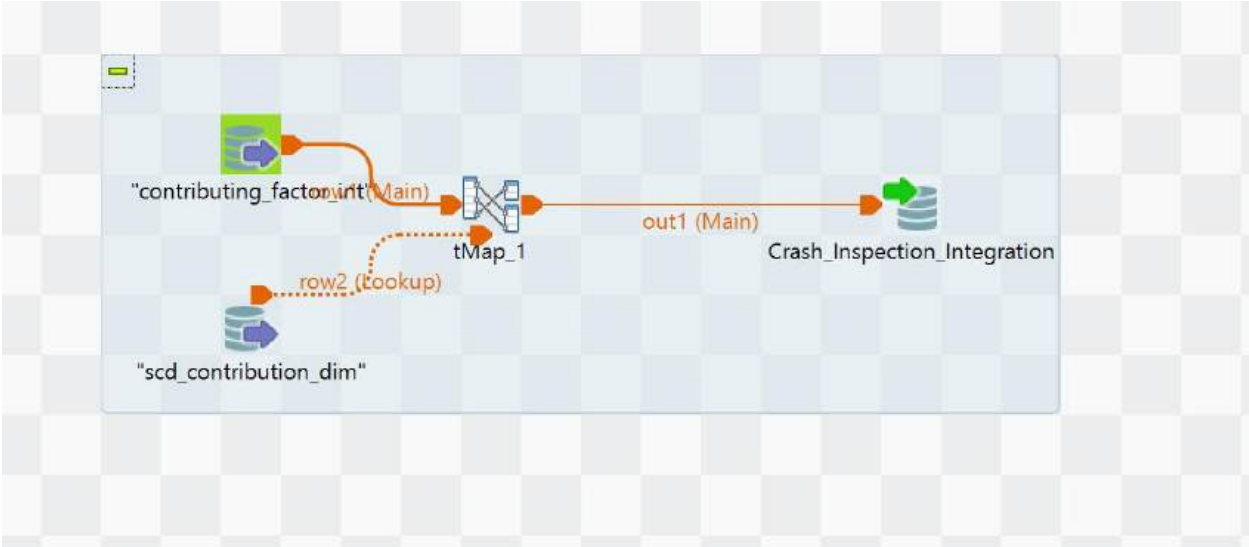
```
CREATE TABLE fact_contribution (
  Contri_Fact_Id int NOT NULL,
  crash_id varchar(300) NOT NULL,
  Code varchar(5) NOT NULL,
  source_system varchar(50) NOT NULL,
  SK_Contri int NOT NULL,
  PRIMARY KEY (Contri_Fact_Id,SK_Contri)
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 •  SELECT * FROM vehcile_curation.fact_contribution;
```

Result Grid | Filter Rows: | Edit: | Export/I

Contri_Fact_Id	crash_id	Code	source_system	SK_Contri
1	13756966	20	Austin	11
2	13762356	22	Austin	13
3	13762394	20	Austin	11
4	13765043	20	Austin	11
5	13765164	36	Austin	27
6	13766397	68	Austin	59
7	13766398	20	Austin	11
8	13767417	34	Austin	25
9	13767511	20	Austin	11
10	13772704	20	Austin	11
11	13773529	20	Austin	11
12	13774913	67	Austin	58
13	13776763	60	Austin	51
14	13776798	20	Austin	11
15	13776803	19	Austin	10
16	13776853	20	Austin	11
17	13776853	3	Austin	3
18	13776976	60	Austin	51
19	13777288	60	Austin	51
20	13777289	41	Austin	32
21	13777290	61	Austin	52

fact_contribution 3




```
1 •  SELECT count(*) FROM vehicle_curation.fact_contribution;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(*)
1958904

Audit Tables

```
CREATE TABLE log_fact (
  moment datetime DEFAULT NULL,
  pid varchar(20) DEFAULT NULL,
  father_pid varchar(20) DEFAULT NULL,
  root_pid varchar(20) DEFAULT NULL,
  system_pid bigint DEFAULT NULL,
  project varchar(50) DEFAULT NULL,
  job varchar(255) DEFAULT NULL,
  job_repository_id varchar(255) DEFAULT NULL,
  job_version varchar(255) DEFAULT NULL,
  context varchar(50) DEFAULT NULL,
  origin varchar(255) DEFAULT NULL,
  message_type varchar(255) DEFAULT NULL,
  message varchar(255) DEFAULT NULL,
  duration bigint DEFAULT NULL
) ENGINE=InnoDB DEFAULT CHARSET=utf8mb4 COLLATE=utf8mb4_0900_ai_ci
```

```
1 • SELECT * FROM vehicle_curation.log_fact;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

moment	pid	father_pid	root_pid	system_pid	project	job	job_repository_id	job_version	contex
2024-04-07 17:26:16	ABn3KK	ABn3KK	ABn3KK	23080	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 17:29:07	d1yb2p	d1yb2p	d1yb2p	9884	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 17:46:01	d1yb2p	d1yb2p	d1yb2p	9884	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 17:47:42	OCx7gM	OCx7gM	OCx7gM	13688	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 18:14:36	OCx7gM	OCx7gM	OCx7gM	13688	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 22:25:01	VscB8r	VscB8r	VscB8r	23524	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 22:36:07	XUUjBe	XUUjBe	XUUjBe	22396	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 22:37:36	xURqko	xURqko	xURqko	15240	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default
2024-04-07 22:43:17	Gr3b7X	Gr3b7X	Gr3b7X	24412	FINALPROJECT_COMBINE	FactCrash	_DTR14PQQEe68ocLG8VEQQg	0.1	Default

```
1 •  SELECT count(*) FROM vehicle_curation.log_fact;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

count(*)
9