Final Column	Austin		New York		Chicago	
	Columns	Transformation	Columns	Transformation	Columns	Transformation
CRASH_ID	crash_id	No Transformation	COLLISION_ID	Derived CRASH_ID from the COLLISION_ID column	CRASH_RECORD_ID	Derived CRASH_ID from the CRASH_RECORD_ID column
CRASH_DATE	crash_date	Extracted Date in 'yyyy-MM-dd'	CRASH_DATE	Changed the format to 'yyyy-MM-dd'	CRASH_DATE	Derived from CRASH_DATE and changed the format to 'yyyy-MM-dd'
CRASH_TIME	crash_time	Extracted Time in 'HH:mm:ss'	CRASH_TIME	Changed the data type to Date by keeping format as 'HH:mm:ss'	CRASH_DATE	Derived from CRASH_DATE and changed the data type to Date by keeping format as 'HH-mm-ss'
LATITUDE	latitude	No Transformation	LATITUDE	No Transformation	LATITUDE	No Transformation
LONGITUDE	longitude	No Transformation	LONGITUDE	No Transformation	LONGITUDE	No Transformation
STREET_NAME	street_name	No Transformation	ON_STREET_NAME OFF_STREET_NAME	Replaced the null values for ON_STREET_NAME column with OFF_STREET_NAME	STREET_NAME	No Transformation
TOTAL_INJURY_COUNT	tot_injry_cnt	Derived TOTAL_INJURY_COUNT from the tot_injry_cnt column	NUMBER_OF_PERSONS_INJURED	Derived TOTAL_INJURY_COUNT from the NUMBER_OF_PERSONS_INJURED column	INJURIES_TOTAL	Derived TOTAL_INJURY_COUNT from the INJURIES_TOTAL column
TOTAL_FATAL_COUNT	death_cnt	Derived TOTAL_FATAL_COUNT from the death_cnt column	NUMBER_OF_PERSONS_KILLED	Derived TOTAL_FATAL_COUNT from the NUMBER_OF_PERSONS_KILLED column	INJURIES_FATAL	Derived TOTAL_FATAL_COUNT from the INJURIES_FATAL column
PEDESTRIANS_INJURED	pedestrian_serious_injury_count	Derived PEDESTRIANS_INJURED from the pedestrian_serious_injury_count column	NUMBER_OF_PEDESTRIANS_INJURED	Derived PEDESTRIANS_INJURED from the NUMBER_OF_PEDESTRIANS_INJURED column	FIRST_CRASH_TYPE INJURIES_TOTAL	Used tFilterRow to filter 'pedestrians' from FIRST_CRASH_TYPE column and mapped with INJURIES_TOTAL column to retrieve the count
PEDESTRIANS_KILLED	pedestrian_death_count	Derived PEDESTRIANS_KILLED from the pedestrian_death_count column	NUMBER_OF_PEDESTRIANS_KILLED	Derived PEDESTRIANS_KILLED from the NUMBER_OF_PEDESTRIANS_KILLED column	FIRST_CRASH_TYPE INJURIES_FATAL	Used tFilterRow to filter 'pedestrians' from FIRST_CRASH_TYPE column and mapped with INJURIES_FATAL column to retrieve the count
PEDESTRIANS_INVOLVED	pedestrian_serious_injury_count pedestrian_death_count	Calculated total sum of pedestrian_serious_injury_count and pedestrian_death_count columns	NUMBER_OF_PEDESTRIANS_INJURED NUMBER_OF_PEDESTRIANS_KILLED	Calculated total sum of NUMBER_OF_PEDESTRIANS_INJURED and NUMBER_OF_PEDESTRIANS_KILLED columns	PEDESTRIANS_INJURED PEDESTRIANS_KILLED	Calculated total sum of PEDESTRIANS_INJURED and PEDESTRIANS_KILLED
CRASH_HOUR	crash_time	Extracted CRASH_HOUR from crash_time column	CRASH_TIME	Extracted CRASH_HOUR from CRASH_TIME column	CRASH_HOUR	No Transformation
CRASH_DAY_OF_WEEK	crash_date	Extracted the day of week from crash_date column	CRASH_DATE	Extracted the day of week from CRASH_DATE column	CRASH_DAY_OF_WEEK	No Transformation
CRASH_MONTH	crash_date	Extracted the day of week from crash_date column using getMonth() function	CRASH_DATE	Extracted the day of week from CRASH_DATE column using getMonth() function	CRASH_MONTH	No Transformation
MOTORISTS_INJURED	motorcycle_serious_injury_count motor_vehicle_serious_injury_count	Calculated total sum of motorcycle_serious_injury_count and motor_vehicle_serious_injury_count columns	NUMBER_OF_MOTORIST_INJURED	Derived MOTORISTS_INJURED from the NUMBER_OF_MOTORIST_INJURED column	FIRST_CRASH_TYPE INJURIES_TOTAL	Manually extracted motorists from FIRST_CRASH_TYPE from values such as 'PEDALCYCLIST' , 'PARKED MOTOR VEHICLE' Used tFilterRow to filter 'motorists' from FIRST_CRASH_TYPE column and mapped with INJURIES_TOTAL column to retrieve the count
MOTORISTS_KILLED	motorcycle_death_count motor_vehicle_death_count	Calculated total sum of motorcycle_death_count and motor_vehicle_death_count columns	NUMBER_OF_MOTORIST_KILLED	Derived MOTORISTS_KILLED from the NUMBER_OF_MOTORIST_KILLED column	FIRST_CRASH_TYPE INJURIES_FATAL	Manually extracted motorists from FIRST_CRASH_TYPE from values such as 'PEDALCYCLIST' , 'PARKED MOTOR VEHICLE' Used tFilterRow to filter 'motorists' from FIRST_CRASH_TYPE column and mapped with INJURIES_FATAL column to retrieve the count
CRASH_SEASON	CRASH_MONTH	We are deriving SEASON from the CRASH_MONTH column. Below is the pseudo code for implementation of the logic: if crash_month is either 1 or 12 then 'Winter' else if crash_month is between 3 and 5 then 'Spring' else if crash_month is between 6 and 8 then 'Summer' else if crash_month is between 9 and 11 then 'Fall'	CRASH_MONTH	We are deriving SEASON from the CRASH_MONTH column. Below is the pseudo code for implementation of the logic: if crash_month is either 1 or 12 then 'Winter' else if crash_month is between 3 and 5 then 'Spring' else if crash_month is between 6 and 8 then 'Summer' else if crash_month is between 9 and 11 then 'Fall'	CRASH_MONTH	We are deriving SEASON from the CRASH_MONTH column. Below is the pseudo code for implementation of the logic: if crash_month is either 1 or 12 then 'Winter' else if crash_month is between 3 and 5 then 'Spring' else if crash_month is between 6 and 8 then 'Summer' else if crash_month is between 9 and 11 then 'Fall'
CONTRIBUTING_FACTORS_D	Mapping Document contrib_factr_p1_id contrib_factr_p2_id	Mapped Contributing Factors and merged using '$' delimiter	CONTRIBUTING_FACTOR_VEHICLE_1 CONTRIBUTING_FACTOR_VEHICLE_2 CONTRIBUTING_FACTOR_VEHICLE_3 CONTRIBUTING_FACTOR_VEHICLE_4 CONTRIBUTING_FACTOR_VEHICLE_5	We have merged all the 5 contributing factor columns using '$' delimiter	PRIM_CONTRIBUTORY_CAUSE SEC_CONTRIBUTORY_CAUSE	We have merged the 2 contributing factor columns using '$' delimiter
CONTRIBUTING_FACTORS_I	contrib_factr_p1_id contrib_factr_p2_id	We have merged the 2 contributing factor id columns using '$' delimiter	Mapping Document CONTRIBUTING_FACTOR_VEHICLE_1 CONTRIBUTING_FACTOR_VEHICLE_2 CONTRIBUTING_FACTOR_VEHICLE_3 CONTRIBUTING_FACTOR_VEHICLE_4 CONTRIBUTING_FACTOR_VEHICLE_5	Mapped Contributing Factors and merged using '$ delimiter'	Mapping Document PRIM_CONTRIBUTORY_CAUSE SEC_CONTRIBUTORY_CAUSE	Mapped Contributing Factors and merged using '$ delimiter'
VEHICLE_TYPE	units_involved	Derived VEHICLE_TYPE from the units_involved column	VEHICLE_TYPE_CODE_1 VEHICLE_TYPE_CODE_2 VEHICLE_TYPE_CODE_3 VEHICLE_TYPE_CODE_4 VEHICLE_TYPE_CODE_5	We have merged all the 5 vehicle types columns using a delimiter	-	column has been added for Chicago to maintain the granularity and null values have been handled
VEHICLE_COUNT	units_involved	Calculated the total count of vehicle types present using '&' delimiter	VEHICLE_TYPE_CODE_1 VEHICLE_TYPE_CODE_2 VEHICLE_TYPE_CODE_3 VEHICLE_TYPE_CODE_4 VEHICLE_TYPE_CODE_5	Calculated the total count of vehicle types present	-	column has been added for Chicago to maintain the granularity and null values have been handled