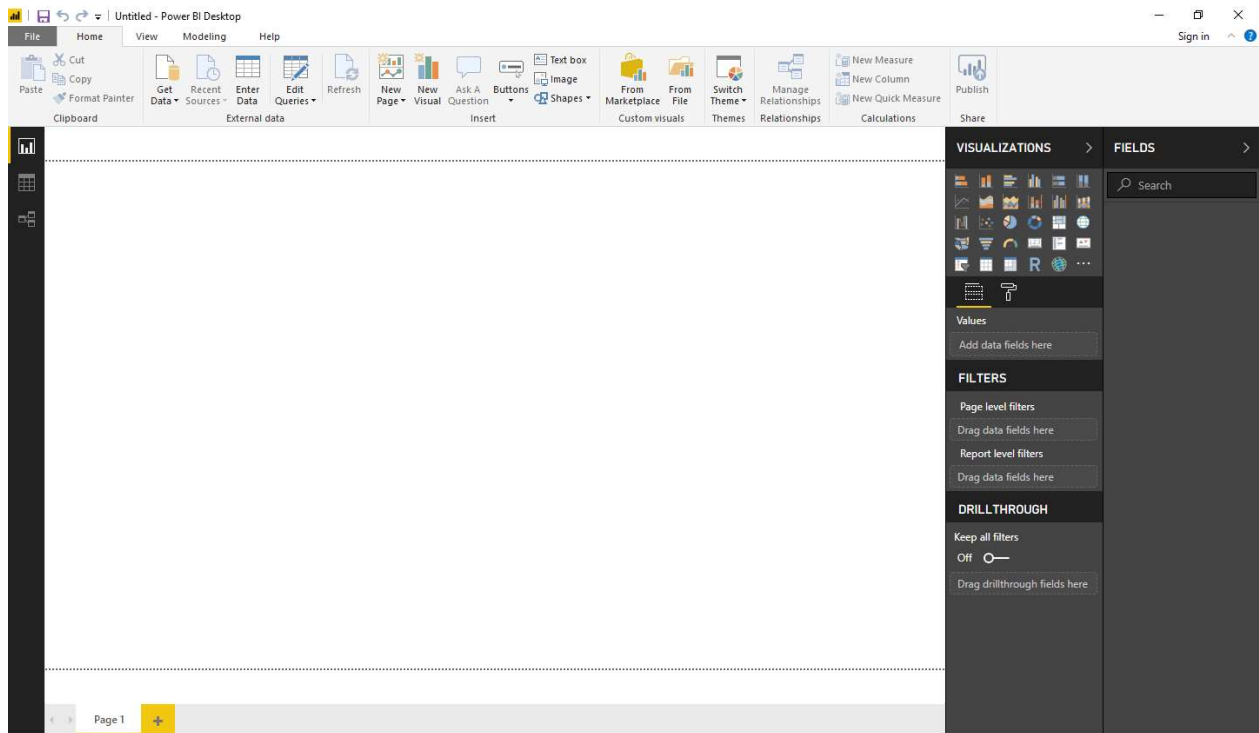
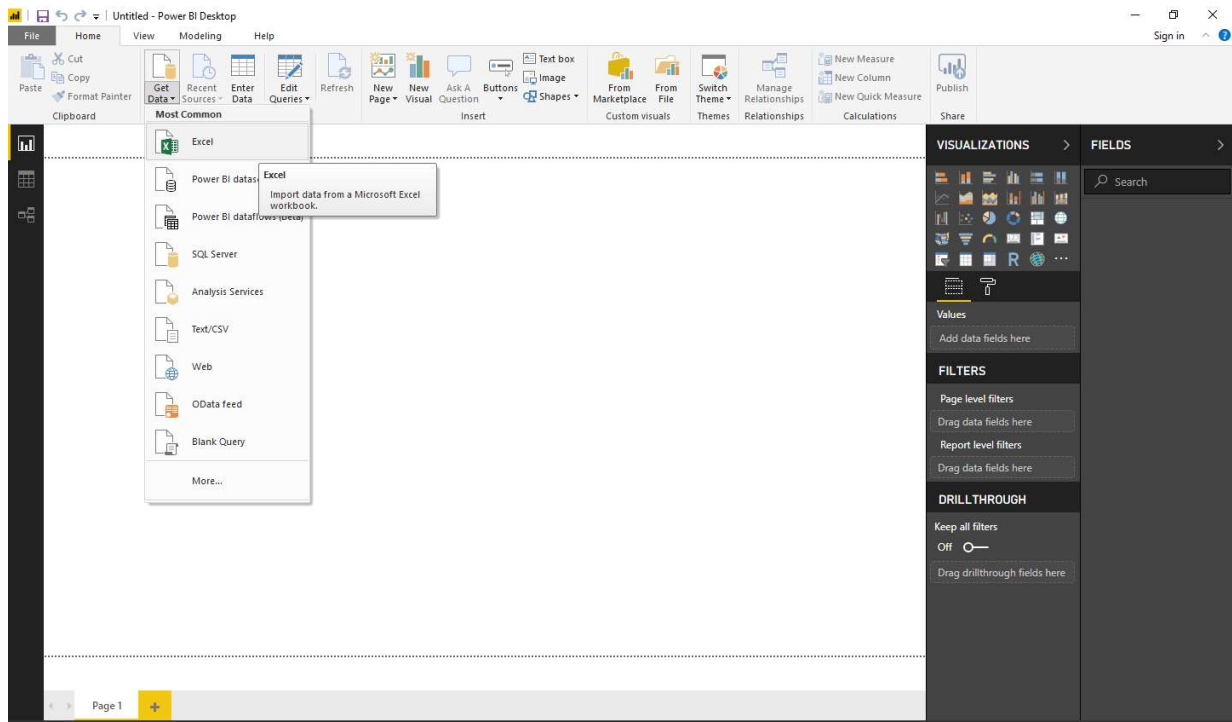


Practical 2: Perform the Extraction Transformation and Loading (ETL) process to construct the database in SQL server.

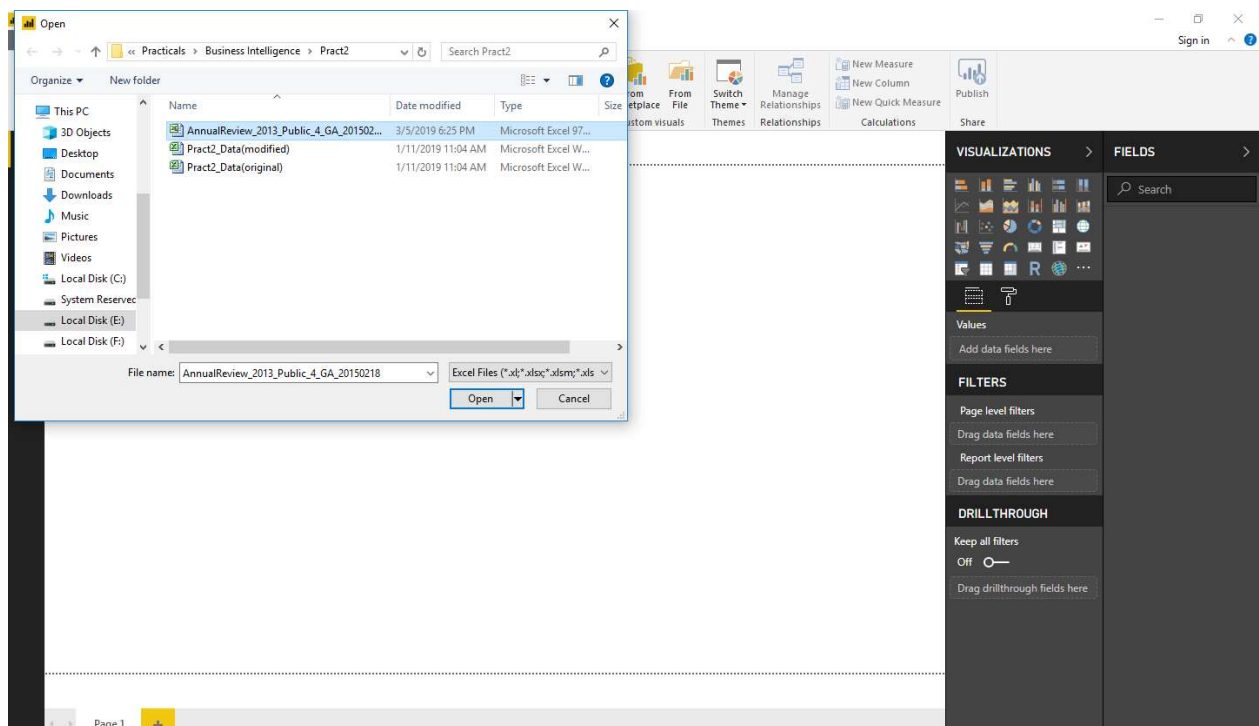
Step 1: Open Power BI.



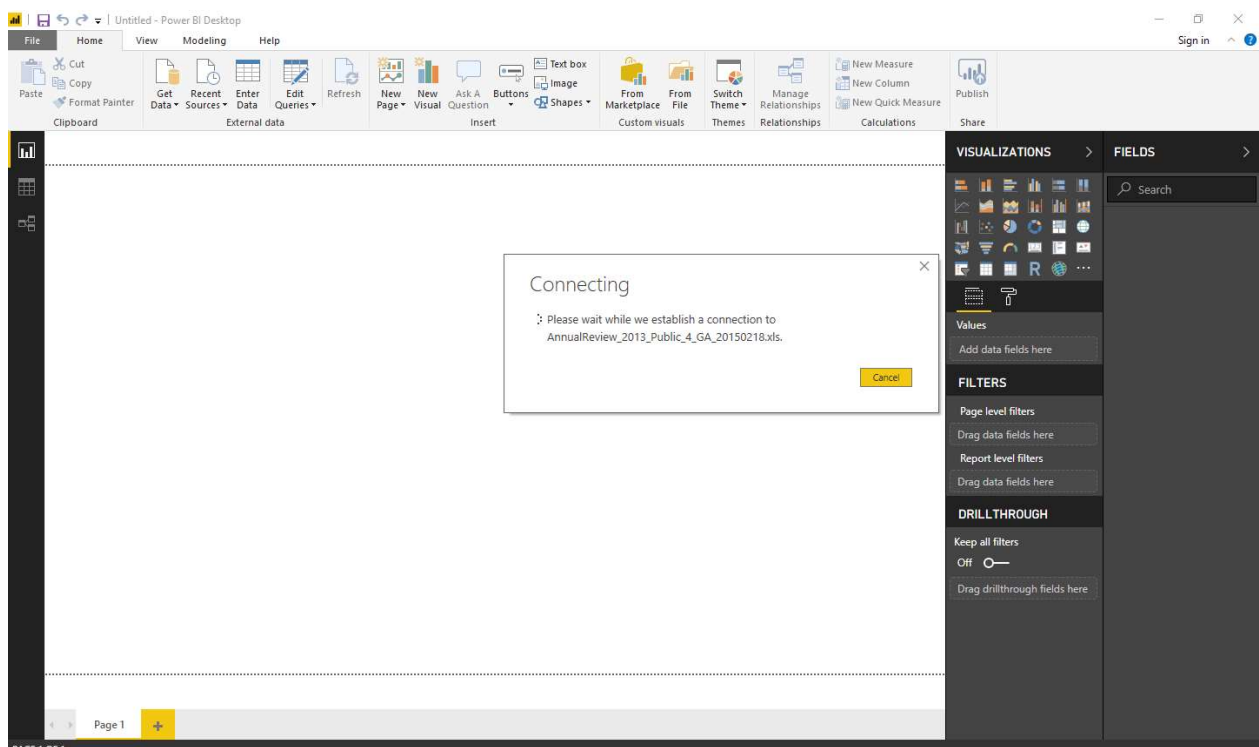
Step 2: Get Data → Excel



Step 3: Choose you Excel data.



Step 4: Loading might take time.



Step 5: Select Data_GA.

The screenshot shows the Power BI Desktop interface. The Navigator pane on the left lists various data sources under 'AnnualReview_2013_Public_4_GA_20150218...'. The 'Data_GA' source is selected. The Data_GA table preview is shown on the right, displaying columns: Column1, Column2, Column3, and Column4. The table contains data for General Aviation Accident Aircraft, 2013, with columns: ntsb_no, aircraft_key, ev_date, and latitude.

Column1	Column2	Column3	Column4
General Aviation Accident Aircraft, 2013	nts_b_no	aircraft_key	ev_date
WPR13CA079	1	1/1/2013	423428N
WPR13LA082	1	1/2/2013	361238N
WPR13FA083	1	1/2/2013	350358N
WPR13FA080	1	1/2/2013	354337N
CEN13FA122	1	1/2/2013	430834N
ERA13FA101	1	1/1/2013	335310N
CEN13FA121	1	1/2/2013	351826N
WPR13CA084	1	1/3/2013	381727N
ERA13LA104	1	1/2/2013	275456N
CEN13CA124	1	1/2/2013	305556N
ERA13FA105	1	1/4/2013	292732N
ANC13CA019	1	1/4/2013	615419N
WPR13LA085	1	1/5/2013	330009N
WPR13FA086	1	1/6/2013	354237N
ERA13CA107	1	1/7/2013	371014N
WPR13LA087	1	1/8/2013	361959N
WPR13CA088	1	1/8/2013	340745N
WPR13CA089	1	1/8/2013	385515N
CEN13FA130	1	1/11/2013	411206N
CEN13LA127	1	1/2/2013	413925N
CEN13FA131	1	1/12/2013	333945N

Step 6: Click on Edit button

The screenshot shows the Power BI Desktop interface. The Navigator pane on the left lists various data sources under 'AnnualReview_2013_Public_4_GA_20150218...'. The 'Data_GA' source is selected. The Data_GA table preview is shown on the right, displaying columns: Column1, Column2, Column3, and Column4. The table contains data for General Aviation Accident Aircraft, 2013, with columns: ntsb_no, aircraft_key, ev_date, and latitude. The 'Edit' button is highlighted in the bottom right corner of the Data_GA preview pane.

Column1	Column2	Column3	Column4
General Aviation Accident Aircraft, 2013	nts_b_no	aircraft_key	ev_date
WPR13CA079	1	1/1/2013	423428N
WPR13LA082	1	1/2/2013	361238N
WPR13FA083	1	1/2/2013	350358N
WPR13FA080	1	1/2/2013	354337N
CEN13FA122	1	1/2/2013	430834N
ERA13FA101	1	1/1/2013	335310N
CEN13FA121	1	1/2/2013	351826N
WPR13CA084	1	1/3/2013	381727N
ERA13LA104	1	1/2/2013	275456N
CEN13CA124	1	1/2/2013	305556N
ERA13FA105	1	1/4/2013	292732N
ANC13CA019	1	1/4/2013	615419N
WPR13LA085	1	1/5/2013	330009N
WPR13FA086	1	1/6/2013	354237N
ERA13CA107	1	1/7/2013	371014N
WPR13LA087	1	1/8/2013	361959N
WPR13CA088	1	1/8/2013	340745N
WPR13CA089	1	1/8/2013	385515N
CEN13FA130	1	1/11/2013	411206N
CEN13LA127	1	1/2/2013	413925N
CEN13FA131	1	1/12/2013	333945N

Step 7: Power Query Editor window will open.

The screenshot shows the Power Query Editor window with a table of aircraft accident data. The table has 21 columns and 999 rows. The columns are: General Aviation Accident Aircraft, 2013, aircraft_key, ev_date, latitude, longitude, ev_city, ev_state, and ev_count. The data is sorted by aircraft_key. The table contains 31 rows of data, starting with 'General Aviation Accident Aircraft, 2013' and ending with 'Moriarty'.

General Aviation Accident Aircraft, 2013	aircraft_key	ev_date	latitude	longitude	ev_city	ev_state	ev_count
WPR13CA079	1	1/1/2013	423428N	1215226W	Chiloquin	OR	USA
WPR13LA082	1	1/2/2013	361238N	1151140W	North Las Vegas	NV	USA
WPR13FA083	1	1/2/2013	350358N	1203706W	Oceano	CA	USA
WPR13FA080	1	1/2/2013	354337N	1190945W	Delano	CA	USA
CEN13FA122	1	1/2/2013	430834N	0932858W	Clear Lake	IA	USA
ERA13FA101	1	1/1/2013	355310N	0871900W	Jasper	AL	USA
CEN13FA121	1	1/2/2013	351826N	0963450W	Seminole	OK	USA
WPR13CA084	1	1/3/2013	381727N	1213840W	Five Points	CA	USA
ERA13LA104	1	1/2/2013	275456N	0822658W	Tampa	FL	USA
CEN13CA124	1	1/2/2013	305556N	0994833W	Menard	TX	USA
ERA13FA105	1	1/4/2013	292732N	0811108W	Palm Coast	FL	USA
ANC13CA019	1	1/4/2013	615419N	1485525W	Palmer	AK	USA
WPR13LA085	1	1/5/2013	330009N	1164305W	Julian	CA	USA
WPR13FA086	1	1/6/2013	354237N	1185414W	Woody	CA	USA
ERA13CA107	1	1/7/2013	371014N	0833455W	Bear Branch	KY	USA
WPR13LA087	1	1/8/2013	361959N	1194930W	Riverdale	CA	USA
WPR13CA088	1	1/8/2013	340745N	1172405W	Rialto	CA	USA
WPR13CA089	1	1/8/2013	385515N	1205153W	Georgetown	CA	USA
CEN13FA130	1	1/11/2013	411206N	1002813W	Maxwell	NE	USA
CEN13LA127	1	1/2/2013	413925N	0860205W	Elkhart	IN	USA
CEN13FA131	1	1/12/2013	333945N	0953252W	Paris	TX	USA
ERA13FA109	1	1/12/2013	272310N	0823343W	Sarasota	FL	USA
ERA13LA110	1	1/12/2013	335858N	0834003W	Winder	GA	USA
ERA13LA113	1	1/13/2013	355404N	0754205W	Manteo	NC	USA
CEN13LA132	1	1/7/2013	341352N	0953715W	Antlers	OK	USA
CEN13CA133	1	1/11/2013	414128N	0933359W	Ankeny	IA	USA
ERA13LA111	1	1/13/2013	391021N	0752925W	Dover	DE	USA
ERA13LA112	1	1/11/2013	275616N	0821604W	Brandon	FL	USA
CEN13CA134	1	1/10/2013	345856N	1060018W	Moriarty	NM	USA

Step 8: Reduce→Remove rows→Remove top rows.

The screenshot shows the Power Query Editor window with the 'Remove Top Rows' dialog box open. The dialog box has two options: 'Remove the top N rows from this table.' and 'Remove the bottom N rows from this table.' The first option is selected. The table data is visible in the background, showing the same aircraft accident data as in the previous screenshot.

Step 9: Enter the number of rows you want to reduce (here it is just one row from top)

The screenshot shows the Power Query Editor interface. The 'Transform' tab is active, and the 'Remove Rows' button is highlighted. A dialog box titled 'Remove Top Rows' is open, prompting the user to specify how many rows to remove from the top. The 'Number of rows' field is set to 1. The background data table shows columns: aircraft_key, ev_date, latitude, longitude, ev_city, ev_state, and ev_county. The 'QUERY SETTINGS' pane on the right shows the 'APPLIED STEPS' list with 'Changed Type' selected.

Step 10: Transform → Use first row as headers.

The screenshot shows the Power Query Editor interface. The 'Transform' tab is active, and the 'Use First Row as Headers' button is highlighted. The background data table shows columns: aircraft_key, ev_date, latitude, longitude, ev_city, ev_state, and ev_county. The 'QUERY SETTINGS' pane on the right shows the 'APPLIED STEPS' list with 'Removed Top Rows' selected.

Step 11: Your table has appropriate headers now.

The screenshot shows the Power Query Editor interface. The main area displays a table with 21 columns and 199 rows. The columns are: ntssb_no, aircraft_key, ev_date, latitude, longitude, ev_city, ev_state, and ev_country. The table contains data for various aircraft events across different states and countries. The 'ev_date' column shows dates from 1/1/2013 to 1/15/2013. The 'ev_state' column shows various US states like OR, NV, CA, IA, AL, OK, FL, TX, AK, IN, KY, NE, PA, DE, GA, NC, and OH. The 'ev_country' column shows 'USA' for all entries.

Step 12: Right click on the columns which you don't want to use → Remove Columns.

The screenshot shows the Power Query Editor interface with a right-click context menu open over the 'longitude' column. The menu options include: Copy, Remove Columns, Remove Other Columns, Add Column From Examples..., Remove Duplicates, Remove Errors, Replace Values..., Fill, Change Type, Merge Columns, Group By..., Unpivot Columns, Unpivot Other Columns, Unpivot Only Selected Columns, and Move. The 'Remove Columns' option is highlighted.

Step 13: Replace values by selecting that cell→right click→replace values...

The screenshot shows the Power Query Editor interface. A right-click context menu is open over a cell in the 'inj_tot_f' column, with the 'Replace Values...' option selected. The background table contains data with columns: ntstb_no, ev_date, ev_city, ev_state, ev_country, inj_tot_f, inj_tot_s, and ev_highest_injury. The 'Data_GA' query is selected in the left pane. The right pane shows the 'QUERY SETTINGS' for 'Data_GA', including 'PROPERTIES' and 'APPLIED STEPS'.

ntstb_no	ev_date	ev_city	ev_state	ev_country	inj_tot_f	inj_tot_s	ev_highest_injury
1	1/1/2013	Onioquin	OR	USA		null	NONE
2	1/2/2013	North Las Vegas	NV	USA			
3	1/2/2013	Oceano	CA	USA			
4	1/2/2013	Delano	CA	USA			
5	1/2/2013	Clear Lake	IA	USA			
6	1/1/2013	Jasper	AL	USA			
7	1/2/2013	Seminole	OK	USA			
8	1/3/2013	Five Points	CA	USA		null	MINR
9	1/2/2013	Tampa	FL	USA		null	NONE
10	1/2/2013	Menard	TX	USA		null	NONE
11	1/4/2013	Palm Coast	FL	USA	3	null	FATL
12	1/4/2013	Palmer	AK	USA	null	null	NONE
13	1/5/2013	Julian	CA	USA	null	null	MINR
14	1/6/2013	Woody	CA	USA	2	null	FATL
15	1/7/2013	Bear Branch	KY	USA	null	null	NONE
16	1/8/2013	Riverdale	CA	USA	null	1	SERS
17	1/8/2013	Rialto	CA	USA	null	null	MINR
18	1/8/2013	Georgetown	CA	USA	null	null	MINR
19	1/11/2013	Maxwell	NE	USA	4	null	FATL
20	1/2/2013	Elkhart	IN	USA	null	null	NONE
21	1/2/2013	Paris	TX	USA	3	null	FATL
22	1/12/2013	Sarasota	FL	USA	2	null	FATL
23	1/12/2013	Winder	GA	USA	null	null	NONE
24	1/13/2013	Manteo	NC	USA	1	null	FATL
25	1/7/2013	Anders	OK	USA	null	null	NONE
26							

Step 14: Provide the new value and click on OK.

The screenshot shows the Power Query Editor interface with the 'Replace Values' dialog box open. The dialog box prompts the user to 'Replace one value with another in the selected columns.' It has two input fields: 'Value To Find' (containing 'null') and 'Replace With' (containing '0'). The 'OK' button is highlighted. The background table is the same as in the previous screenshot, but the 'inj_tot_f' column now contains the value '1' for row 3.

ntstb_no	ev_date	ev_city	ev_state	ev_country	inj_tot_f	inj_tot_s	ev_highest_injury
1	1/1/2013	Onioquin	OR	USA		null	NONE
2	1/2/2013	North Las Vegas	NV	USA		1	FATL
3	1/2/2013	Oceano	CA	USA	1	null	FATL
4	1/2/2013	Delano	CA	USA	1	null	FATL
5	1/2/2013	Clear Lake	IA	USA	3	null	FATL
6	1/1/2013	Jasper	AL	USA	3	null	FATL
7	1/2/2013	Seminole	OK	USA			
8	1/3/2013	Five Points	CA	USA		null	MINR
9	1/2/2013	Tampa	FL	USA		null	NONE
10	1/2/2013	Menard	TX	USA		null	NONE
11	1/4/2013	Palm Coast	FL	USA	3	null	FATL
12	1/4/2013	Palmer	AK	USA	null	null	NONE
13	1/5/2013	Julian	CA	USA	null	null	MINR
14	1/6/2013	Woody	CA	USA	2	null	FATL
15	1/7/2013	Bear Branch	KY	USA	null	null	NONE
16	1/8/2013	Riverdale	CA	USA	null	1	SERS
17	1/8/2013	Rialto	CA	USA	null	null	MINR
18	1/8/2013	Georgetown	CA	USA	null	null	MINR
19	1/11/2013	Maxwell	NE	USA	4	null	FATL
20	1/2/2013	Elkhart	IN	USA	null	null	NONE
21	1/2/2013	Paris	TX	USA	3	null	FATL
22	1/12/2013	Sarasota	FL	USA	2	null	FATL
23	1/12/2013	Winder	GA	USA	null	null	NONE
24	1/13/2013	Manteo	NC	USA	1	null	FATL
25	1/7/2013	Anders	OK	USA	null	null	NONE
26							

Step 15: All values having the values you wished to replace earlier will be replaced with new values.

The screenshot displays the Microsoft Power Query Editor interface. The main area shows a table with 18 columns and 999+ rows. The columns are: ntsb_no, ev_date, ev_city, ev_state, ev_country, inj_tot_f, inj_tot_s, and ev_highest_injury. The data includes various incident records with details like date, location, state, and injury counts.

On the right side, the 'QUERY SETTINGS' pane is open, showing the 'APPLIED STEPS' section. The steps listed are: Source, Navigation, Changed Type, Removed Top Rows, Promoted Headers, Changed Type1, Removed Columns, and 'X Replaced Value'. The 'Replaced Value' step is currently selected.

The bottom status bar indicates '18 COLUMNS, 999+ ROWS' and 'PREVIEW DOWNLOADED AT 6:31 PM'. The Windows taskbar at the very bottom shows the time as 6:31 PM on 3/5/2019.