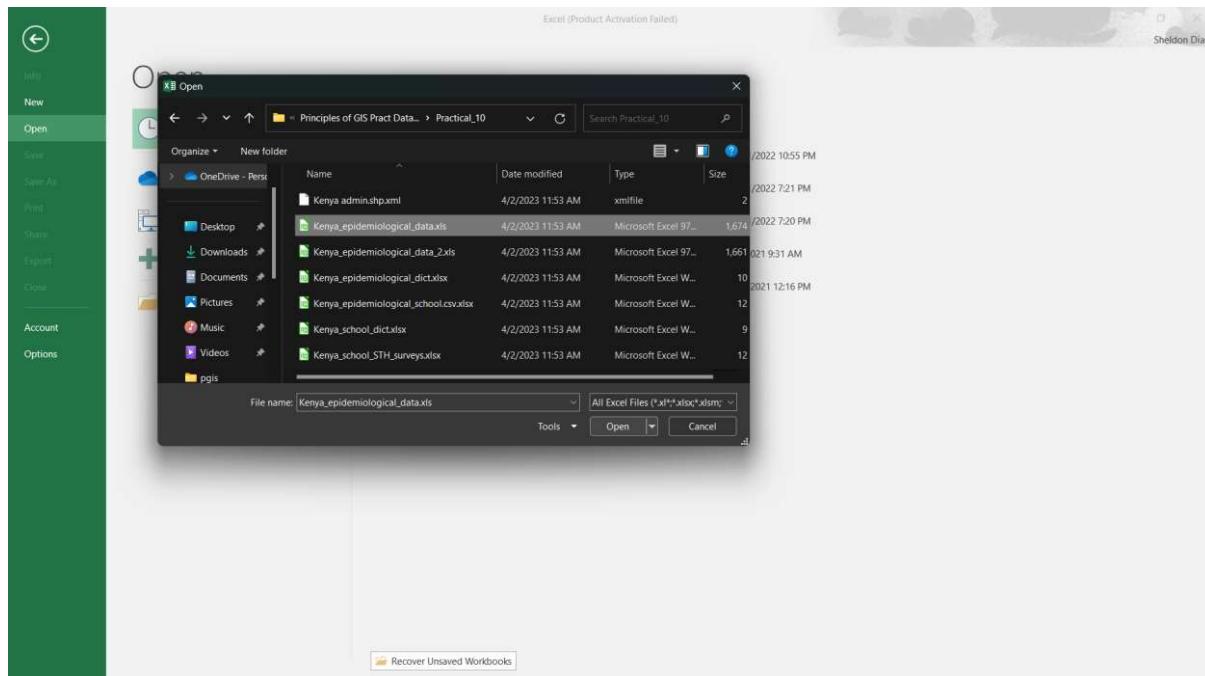


PRACTICAL – 10

Aim: Validating Map data.

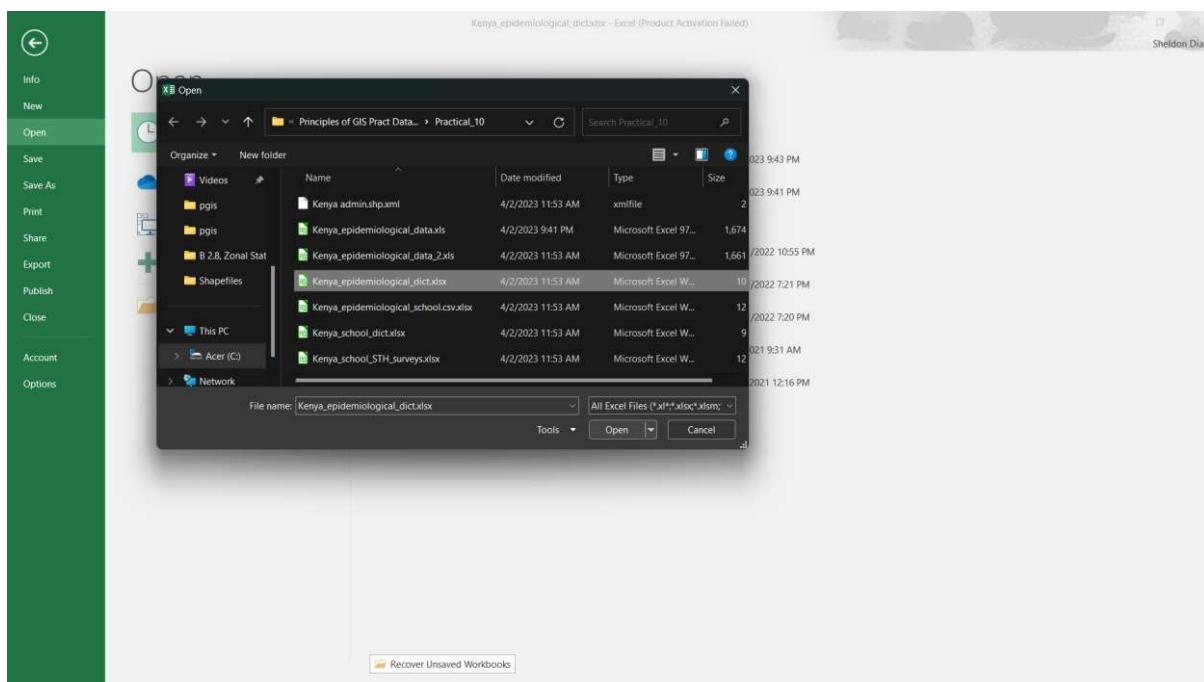
Steps:

Step 1 – Open MS Excel and open the following file.



A	B	C	D	E	F	G	H	K	L	M	N	O	P	Q	R	S	T	U	V	W
141	child_id	school	district	district	age	sex	stool	hookworm	ascaris	trichuris	smansoni	hookworm	ascaris	trichuris	anysth	smansoni	deworm	Chew		
142	70105079	NGINYANK	1203	BARINGO	701	82	2	1	0	0	0	0	0	0	0	0	0	0	0	
143	70105091	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0	
144	70105046	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	
145	70105078	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0	
146	70105011	NGINYANK	1203	BARINGO	701	7	1	1	0	0	0	0	0	0	0	0	0	0	0	
147	70105068	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0	
148	70105047	TEREBEN	1190	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0	
149	70105004	NGINYANK	1203	BARINGO	701	11	1	1	0	0	0	0	0	0	0	0	0	0	0	
150	70105070	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	
151	70105090	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0	
152	70105071	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0	
153	70105058	NGINYANK	1203	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	
154	70105099	NGINYANK	1203	BARINGO	701	14	1	1	0	0	0	0	0	0	0	0	0	0	0	
155	70105047	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0	
156	70105027	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0	
157	70105016	NGINYANK	1203	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	
158	70105074	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	
159	70105106	NGINYANK	1203	BARINGO	701	15	1	1	0	0	0	0	0	0	0	0	0	0	0	
160	70105007	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0	
161	70105093	NGINYANK	1203	BARINGO	701	12	2	1	0	0	0	0	0	0	0	0	0	0	0	
162	70105035	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0	
163	70105046	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0	
164	70105037	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0	
165	70105100	NGINYANK	1203	BARINGO	701	16	1	1	0	0	0	0	0	0	0	0	0	0	0	
166	70105003	NGINYANK	1203	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	
167	70105028	NGINYANK	1203	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	
168	70105051	NGINYANK	1203	BARINGO	701	14	1	1	0	0	0	0	0	0	0	0	0	0	0	
169	70105075	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	
170	70105038	NGINYANK	1203	BARINGO	701	12	2	1	0	0	0	0	0	0	0	0	0	0	0	
171	70105011	TEREBEN	1190	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0	0	

Step 2 – Open the following file.



The screenshot shows an Excel spreadsheet titled 'Kenya_epidemiological_dict.xlsx'. The spreadsheet has a standard ribbon menu at the top with tabs for File, Home, Insert, Page Layout, Formulas, Data, Review, View, and Tell me what you want to do... . The Home tab is selected. The spreadsheet contains a table with three columns:

A	B	C
1 Variable	Format	Description
2 child_id	numeric	unique child identification number
3 school	string	name of school
4 school_id	numeric	unique school identification number
5 district	string	name of district
6 district_id	numeric	unique district identification number
7 age	numeric	age in years
8 sex	categorical	F=female; M = male
9 stool	categorical	0= no stool sample provided, 1= stool sample provided
10 hookworm_egp	numeric	mean hookworm egg count in eggs per gram faeces
11 ascaris_egp	numeric	mean ascaris egg count in eggs per gram faeces
12 trichuris_egp	numeric	mean trichuris egg count in eggs per gram faeces
13 smansoni_egp	numeric	mean S. mansoni egg count in eggs per gram faeces
14 hookworm_inf	categorical	1=infected with hookworm
15 ascaris_inf	categorical	1=infected with ascaris
16 trichuris_inf	categorical	1=infected with hookworm
17 smansoni_inf	categorical	1=infected with S. mansoni
18 anysth_inf	categorical	1=infected with any STH
19 dewormed	categorical	=reports receiving deworming medication in past 6 months
20		
21		
22		
23		
24		
25		
26		
27		
28		
29		
30		
31		
32		

Step 3 – View > Freeze Panes > Freeze Top Row.

The screenshot shows the Microsoft Excel interface with the 'View' tab selected. Under the 'Freeze' section, 'Freeze Top Row' is checked. The worksheet displays a dataset with columns for child_id, school, sex, age, stool, hookworm, trichuriasis, deworming, and checkup. The first row is frozen, and the rest of the data scrolls below it.

Step 4 – Select the A Column. Home > Conditional Formatting > Highlight Cell Rules > Duplicate Values.

The screenshot shows the Microsoft Excel interface with the 'Home' tab selected. In the 'Conditional Formatting' section, the 'Highlight Cells Rules' dropdown is open, and 'Duplicate Values...' is selected. The worksheet shows the same dataset as the previous step, with the A column selected for conditional formatting.

Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)

child_id	school	school	district	district	age	sex	stool	hookw.	ascaris	trichuris	smansi	hookw.	ascaris	trichuris	anysth.	smansi	deworm	Chec	
64	70105087	NGINYANK	1203	BARINGO	701	12	2	1	0	0	0	0	0	0	0	0	0	1	0
65	70105108	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0
66	70107069	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0
67	70105060	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0
68	70105015	NGINYANK	1203	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0	0
69	70107043	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0
70	70107021	TEREBEN	1190	BARINGO	701	6	1	1	0	0	0	0	0	0	0	0	0	0	0
71	70107014	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0
72	70105041	NGINYANK	1203	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0
73	70105031	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0
74	70105039	NGINYANK	1203	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0
75	70107049	TEREBEN	1190	BARINGO	701	11	1	1	0	0	0	0	0	0	0	0	0	0	0
76	70105083	NGINYANK	1203	BARINGO	701	16	1	1	0	0	0	0	0	0	0	0	0	0	0
77	70105065	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0
78	70107031	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0
79	70107052	TEREBEN	1190	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0
80	70105080	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0
81	70105084	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0
82	70105050	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0
83	70105013	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0	0
84	70105077	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0	0
85	70105069	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0	0
86	70105049	NGINYANK	1203	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0
87	70107044	TEREBEN	1190	BARINGO	701	11	2	1	0	0	0	12	0	0	0	1	1	0	0
88	70105014	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0	0
89	70105043	NGINYANK	1203	BARINGO	701	11	2	1	0	0	0	0	0	0	0	0	0	0	0
90	70105021	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0	0
91	70105052	NGINYANK	1203	BARINGO	701	14	1	1	0	0	0	0	0	0	0	0	0	0	0
92	70107030	TEREBEN	1190	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0	0
93	70105053	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0
94	70107026	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0

Step 5 – Select the entire data. Data > Remove Duplicates.

Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)

child_id	school	school	district	district	age	sex	stool	hookw.	ascaris	trichuris	smansi	hookw.	ascaris	trichuris	anysth.	smansi	deworm	Chec
64	70105087	NGINYANK	1203	BARINGO	701	12	2	1	0	0	0	0	0	0	0	0	1	0
65	70105108	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0
66	70107069	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0
67	70105060	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0
68	70105015	NGINYANK	1203	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0
69	70107043	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0
70	70107021	TEREBEN	1190	BARINGO	701	6	1	1	0	0	0	0	0	0	0	0	0	0
71	70107014	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0
72	70105041	NGINYANK	1203	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0
73	70105031	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0
74	70105039	NGINYANK	1203	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0
75	70107049	TEREBEN	1190	BARINGO	701	11	1	1	0	0	0	0	0	0	0	0	0	0
76	70105083	NGINYANK	1203	BARINGO	701	16	1	1	0	0	0	0	0	0	0	0	0	0
77	70105065	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0
78	70107031	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0
79	70107052	TEREBEN	1190	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0
80	70105080	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0
81	70105084	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0
82	70105005	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0
83	70105013	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0
84	70105077	NGINYANK	1203	BARINGO	701	13	2	1	0	0	0	0	0	0	0	0	0	0
85	70105069	NGINYANK	1203	BARINGO	701	13	1	1	0	0	0	0	0	0	0	0	0	0
86	70105049	NGINYANK	1203	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0
87	70107044	TEREBEN	1190	BARINGO	701	11	2	1	0	0	0	12	0	0	0	1	1	0
88	70105014	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0
89	70105043	NGINYANK	1203	BARINGO	701	11	2	1	0	0	0	0	0	0	0	0	0	0
90	70105021	NGINYANK	1203	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0
91	70105052	NGINYANK	1203	BARINGO	701	14	1	1	0	0	0	0	0	0	0	0	0	0
92	70107030	TEREBEN	1190	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0
93	70105053	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0
94	70107026	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0

Kenya_epidemiological_data.xlsx [Compatibility Mode] - Excel (Product Activation Failed)

Average: 4211593.936 Count: 72832 Sum: 2.74011E+11

Ready

Kenya_epidemiological_data.xlsx [Compatibility Mode] - Excel (Product Activation Failed)

Average: 4211593.936 Count: 72832 Sum: 2.74011E+11

Ready

Step 6 – Select the G column and press **ctrl + F**.

The screenshot shows a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls". The data consists of 32 rows of information, mostly from row 2 to row 33. The columns are labeled A through W. Row 1 contains column headers such as child_id, school, sex, stool, hookworm, ascaris, trichuris, smanso, hookworm, ascaris, trichuris, anysth, smanso, deworm, and Chec. A 'Find and Replace' dialog box is overlaid on the spreadsheet. In the 'Find what' field, the letter 'M' is entered. In the 'Replace with' field, the digit '1' is entered. The 'Replace All' button at the bottom left of the dialog box is currently selected.

This screenshot shows the same Excel spreadsheet and 'Find and Replace' dialog box as the previous one. However, the 'Replace All' button has been clicked, and a confirmation dialog box titled "Microsoft Excel" appears in the center. The message in the dialog box reads "All done. We made 2001 replacements." with an "OK" button at the bottom.

Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)

Average: 1 Count: 3876 Sum: 2001 100%

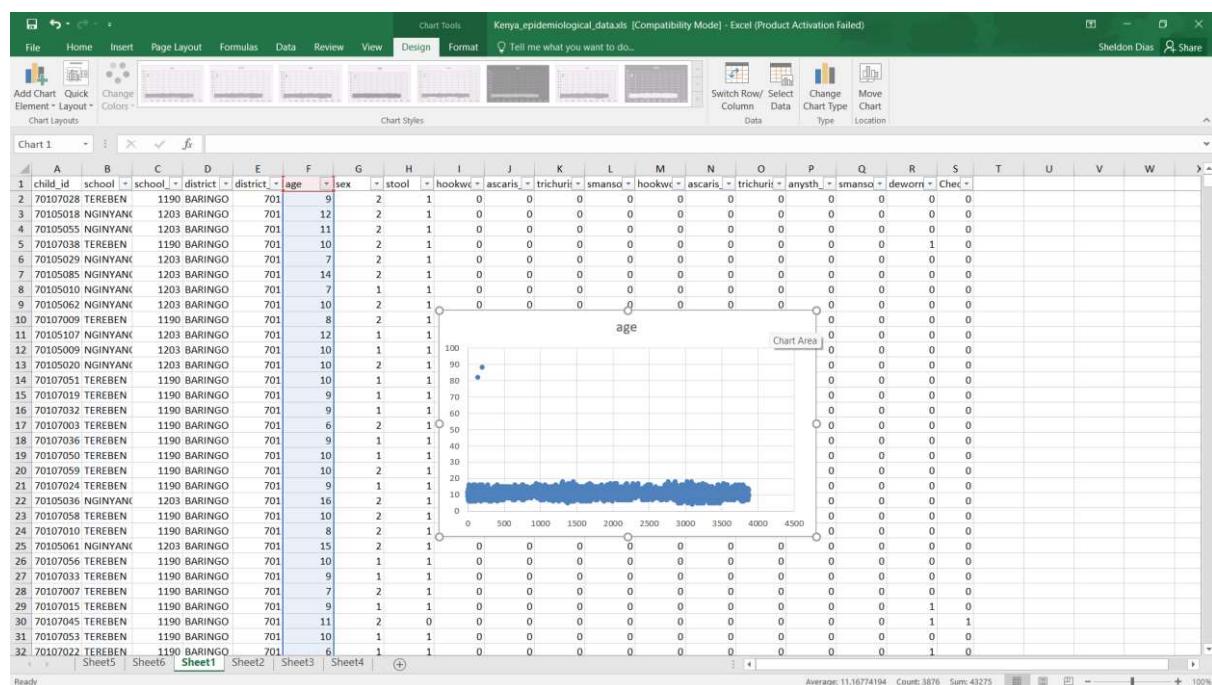
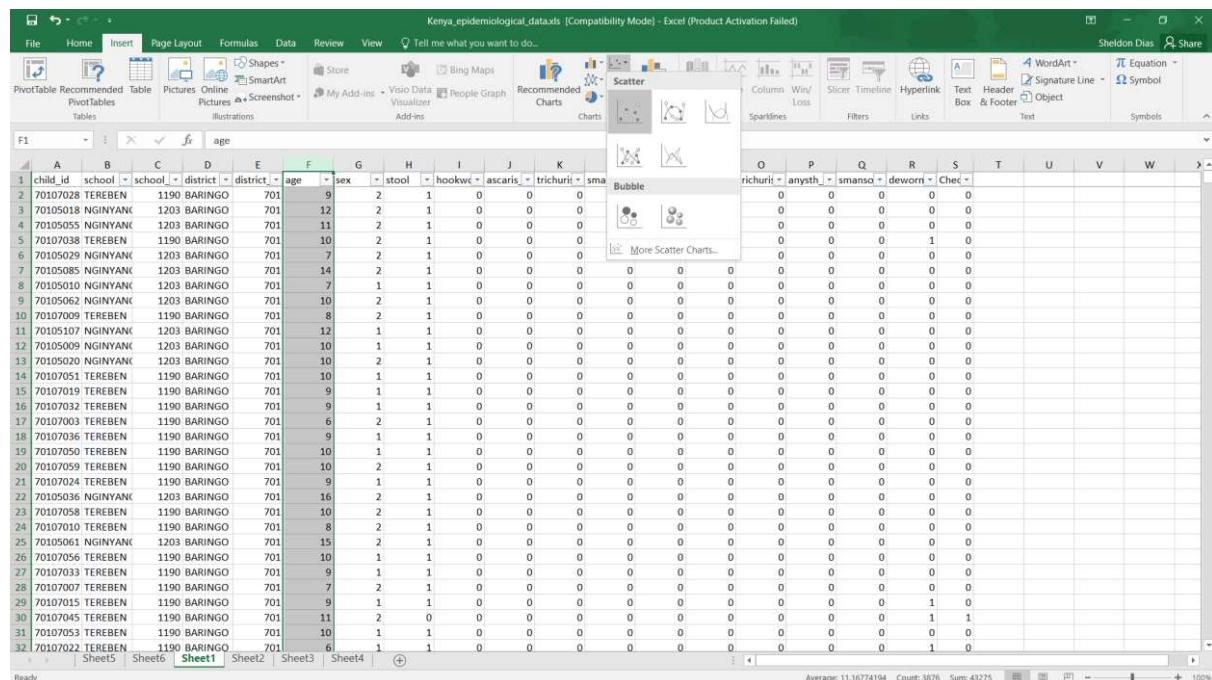
child_id	school	sex	stool	hookwt	ascaris	trichuris	smanso	hookwt	ascaris	trichuris	anysth	smanso	deworm	Ched
70107028	TEREBEN	1190 BARINGO	701	9	F	1	0	0	0	0	0	0	0	0
70105018	NGINYANK	1203 BARINGO	701	12	F	1	0	0	0	0	0	0	0	0
4	70105055	NGINYANK	1203 BARINGO	701	11	F	1	0	0	0	0	0	0	0
5	70107038	TEREBEN	1190 BARINGO	701	10	F	1	0	0	0	0	0	0	1
6	70105029	NGINYANK	1203 BARINGO	701	7	F	1	0	0	0	0	0	0	0
7	70105085	NGINYANK	1203 BARINGO	701	14	F	1	0	0	0	0	0	0	0
8	70105010	NGINYANK	1203 BARINGO	701	7	1	1	0	0	0	0	0	0	0
9	70105062	NGINYANK	1203 BARINGO	701	10	F	1	0	0	0	0	0	0	0
10	70107009	TEREBEN	1190 BARINGO	701	8	F	1	0	0	0	0	0	0	0
11	70105017	NGINYANK	1203 BARINGO	701	12	1	1	0	0	0	0	0	0	0
12	70105009	NGINYANK	1203 BARINGO	701	10	1	1	0	0	0	0	0	0	0
13	70105020	NGINYANK	1203 BARINGO	701	10	F	1	0	0	0	0	0	0	0
14	70107051	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
15	70107019	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
16	70107032	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
17	70107003	TEREBEN	1190 BARINGO	701	6	F	1	0	0	0	0	0	0	0
18	70107036	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
19	70107050	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
20	70107059	TEREBEN	1190 BARINGO	701	10	2	1	0	0	0	0	0	0	0
21	70107024	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
22	70105036	NGINYANK	1203 BARINGO	701	16	2	1	1	0	0	0	0	0	0
23	70107058	TEREBEN	1190 BARINGO	701	10	2	1	0	0	0	0	0	0	0
24	70107010	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
25	70105062	NGINYANK	1203 BARINGO	701	10	2	1	0	0	0	0	0	0	0
26	70107009	TEREBEN	1190 BARINGO	701	8	2	1	0	0	0	0	0	0	0
27	70107045	TEREBEN	1190 BARINGO	701	11	2	0	0	0	0	0	0	0	1
28	70107053	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
29	70107051	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	1
30	70107045	TEREBEN	1190 BARINGO	701	11	2	0	0	0	0	0	0	0	1
31	70107053	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
32	70107022	TEREBEN	1190 BARINGO	701	6	1	1	0	0	0	0	0	0	1

Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)

Average: 1 Count: 3876 Sum: 2001 100%

child_id	school	sex	stool	hookwt	ascaris	trichuris	smanso	hookwt	ascaris	trichuris	anysth	smanso	deworm	Ched
70107028	TEREBEN	1190 BARINGO	701	9	2	1	0	0	0	0	0	0	0	0
3	70105018	NGINYANK	1203 BARINGO	701	12	2	1	0	0	0	0	0	0	0
4	70105055	NGINYANK	1203 BARINGO	701	11	2	1	0	0	0	0	0	0	0
5	70107038	TEREBEN	1190 BARINGO	701	10	2	1	0	0	0	0	0	0	1
6	70105029	NGINYANK	1203 BARINGO	701	7	2	1	0	0	0	0	0	0	0
7	70105085	NGINYANK	1203 BARINGO	701	14	2	1	0	0	0	0	0	0	0
8	70105010	NGINYANK	1203 BARINGO	701	7	1	1	0	0	0	0	0	0	0
9	70105062	NGINYANK	1203 BARINGO	701	10	2	1	0	0	0	0	0	0	0
10	70107009	TEREBEN	1190 BARINGO	701	8	2	1	0	0	0	0	0	0	0
11	70105017	NGINYANK	1203 BARINGO	701	12	1	1	0	0	0	0	0	0	0
12	70105009	NGINYANK	1203 BARINGO	701	10	1	1	0	0	0	0	0	0	0
13	70105020	NGINYANK	1203 BARINGO	701	10	2	1	0	0	0	0	0	0	0
14	70107051	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
15	70107019	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
16	70107032	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
17	70107003	TEREBEN	1190 BARINGO	701	6	2	1	0	0	0	0	0	0	0
18	70107036	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
19	70107050	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
20	70107059	TEREBEN	1190 BARINGO	701	10	2	0	0	0	0	0	0	0	0
21	70107024	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
22	70105036	NGINYANK	1203 BARINGO	701	16	2	1	0	0	0	0	0	0	0
23	70107058	TEREBEN	1190 BARINGO	701	10	2	1	0	0	0	0	0	0	0
24	70107010	TEREBEN	1190 BARINGO	701	8	2	1	0	0	0	0	0	0	0
25	70105062	NGINYANK	1203 BARINGO	701	15	2	1	0	0	0	0	0	0	0
26	70107056	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
27	70107033	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	0
28	70107007	TEREBEN	1190 BARINGO	701	7	2	1	0	0	0	0	0	0	0
29	70107015	TEREBEN	1190 BARINGO	701	9	1	1	0	0	0	0	0	0	1
30	70107045	TEREBEN	1190 BARINGO	701	11	2	0	0	0	0	0	0	0	1
31	70107053	TEREBEN	1190 BARINGO	701	10	1	1	0	0	0	0	0	0	0
32	70107022	TEREBEN	1190 BARINGO	701	6	1	1	0	0	0	0	0	0	1

Step 7 – Select the F column. Insert > Scatter.



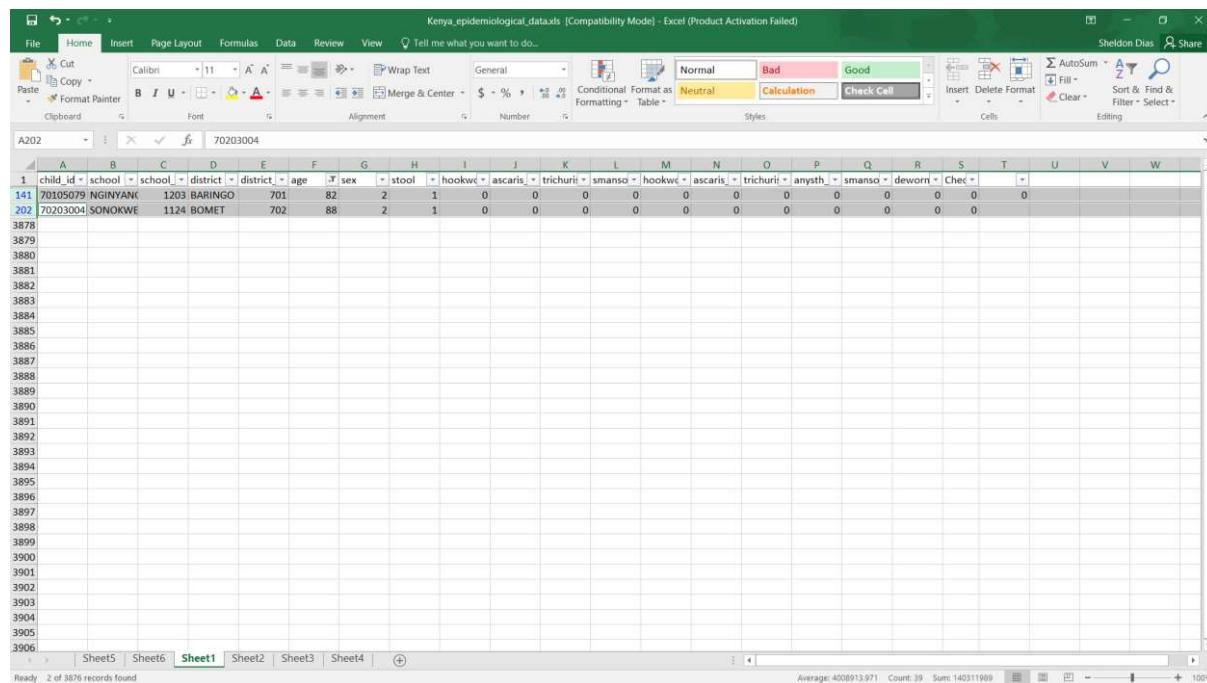
Step 8 – Select the first row. Home > Sort & Filter > Filter.

The screenshot shows a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls". The data consists of 32 rows of child information, including columns for child_id, school, sex, age, and various health metrics. The "Home" tab is selected, and the "Sort & Filter" button in the ribbon is highlighted. A context menu is open at the top-left of the data area, specifically at the first row. The menu path "Sort & Filter -> Filter" is visible. Other options in the menu include "Sort A to Z", "Sort Z to A", "Custom Sort...", "Clear", and "Reapply". The status bar at the bottom right indicates "Count: 19" and "100%".

Step 9 – Click the arrow beside age. Only select 82 and 88 and then click OK.

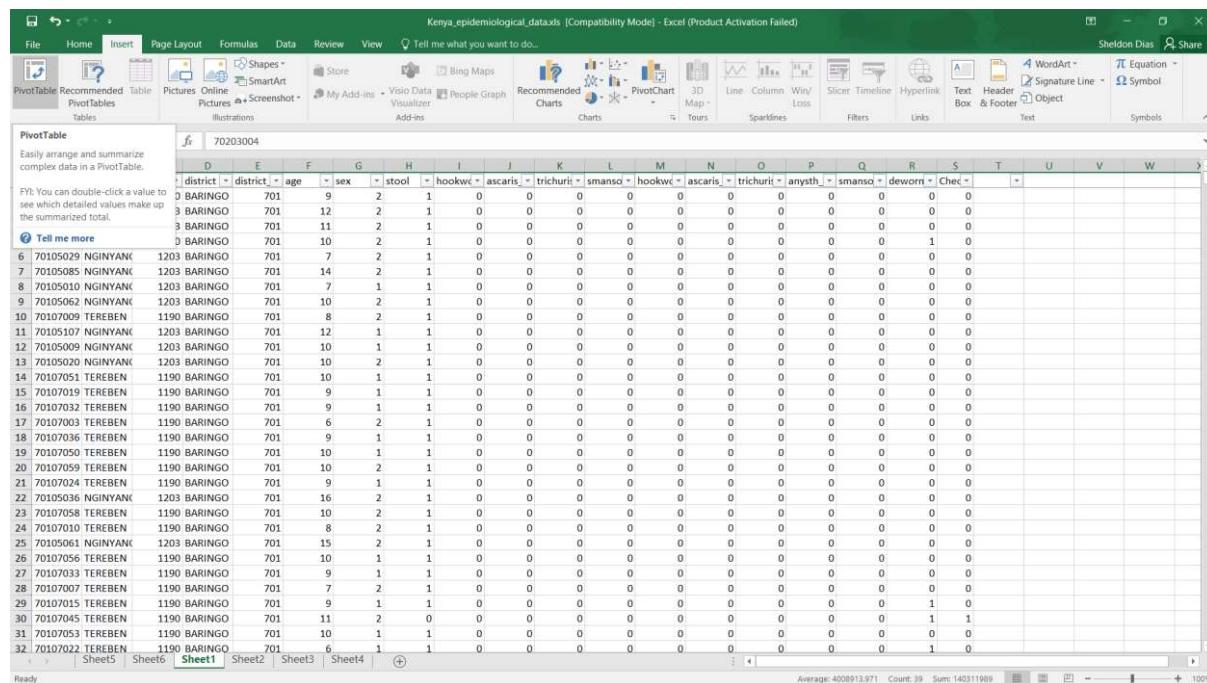
The screenshot shows the same Excel spreadsheet as above, but now the "age" column is being filtered. The "Filter" dialog box is open over the "age" column header. It displays several filter options: "Sort Smallest to Largest", "Sort Largest to Smallest", "Sort by Color", "Clear Filter From 'age'", "Filter by Color", and "Number Filters". Under "Number Filters", there is a search bar and a list of numerical values from 1 to 18. Two values are checked: 82 and 88. Below the list are "OK" and "Cancel" buttons, with "OK" being highlighted. The status bar at the bottom right indicates "Count: 19" and "100%".

Step 10 – Select both the columns and click delete.



A screenshot of a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls". The spreadsheet contains a single sheet with data from row 1 to 3878. The columns are labeled A through W, representing various epidemiological parameters. Row 141 shows data for a child from NGINYANK, aged 701, with sex 82, age 2, stool 1, hookworm 0, ascariasis 0, trichuriasis 0, smanso 0, dewormi 0, and Chec 0. Row 202 shows data for a child from SONOKWE, aged 702, with sex 88, age 2, stool 1, hookworm 0, ascariasis 0, trichuriasis 0, smanso 0, dewormi 0, and Chec 0. The rest of the rows contain similar data for other children across different locations and ages.

Step 11 – Insert > Pivot Table. Set the following attributes and click OK.



A screenshot of the same Microsoft Excel spreadsheet after inserting a Pivot Table. The Pivot Table ribbon tab is selected, and the PivotTable tool is active in the ribbon. The data table remains the same, showing 3878 rows of epidemiological data. The Pivot Table interface is visible on the left side of the screen, indicating the setup for summarizing the data.

The screenshot shows a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls". A "Create PivotTable" dialog box is open, prompting the user to select a table or range. The range "Sheet1!\$A\$1:\$S\$32" is selected. The dialog box includes options for choosing a data source, placing the PivotTable, and analyzing multiple tables.

The screenshot shows the same Microsoft Excel spreadsheet with the PivotTable Tools ribbon tab selected. The "PivotTable Fields" pane on the right lists fields such as child_id, school, sex, age, stool, hookworm_egp, and ascari_epp. The "ROWS" section of the PivotTable Fields pane is currently empty, and the "VALUES" section is also empty.

Step 12 – Drag and Drop fields as shown below.

The screenshot shows a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)". A PivotTable is being used to analyze the data. The PivotTable Fields pane on the right lists various fields: child_id, school, district, age, sex, stool, hookworm_egp, and ascaris_egp. The PivotTable itself is located in the center of the screen, showing a count of district for each child_id. The PivotTable Fields pane also shows that "age" is currently selected as a column field, and "child_id" is selected as a row field. The PivotTable structure is as follows:

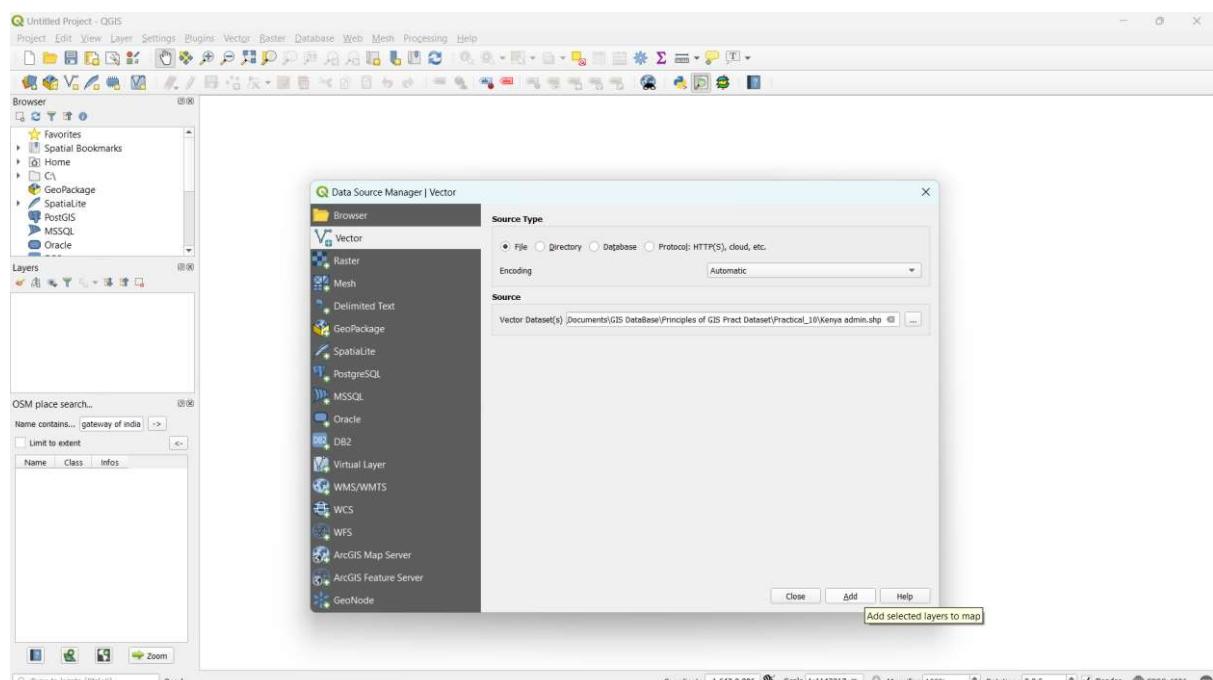
child_id	Count of district
60904001	1
60904002	1
60904003	1
60904004	1
60904005	1
60904006	1
60904007	1
60904008	1
60904009	1
60904010	1
60904011	1
60904012	1
60904013	1
60904014	1
60904015	1
60904016	1
60904017	1
60904018	1
60904019	1
60904020	1
60904021	1
60904022	1
60904023	1
60904024	1
60904025	1
60904026	1
60904027	1
60904028	1

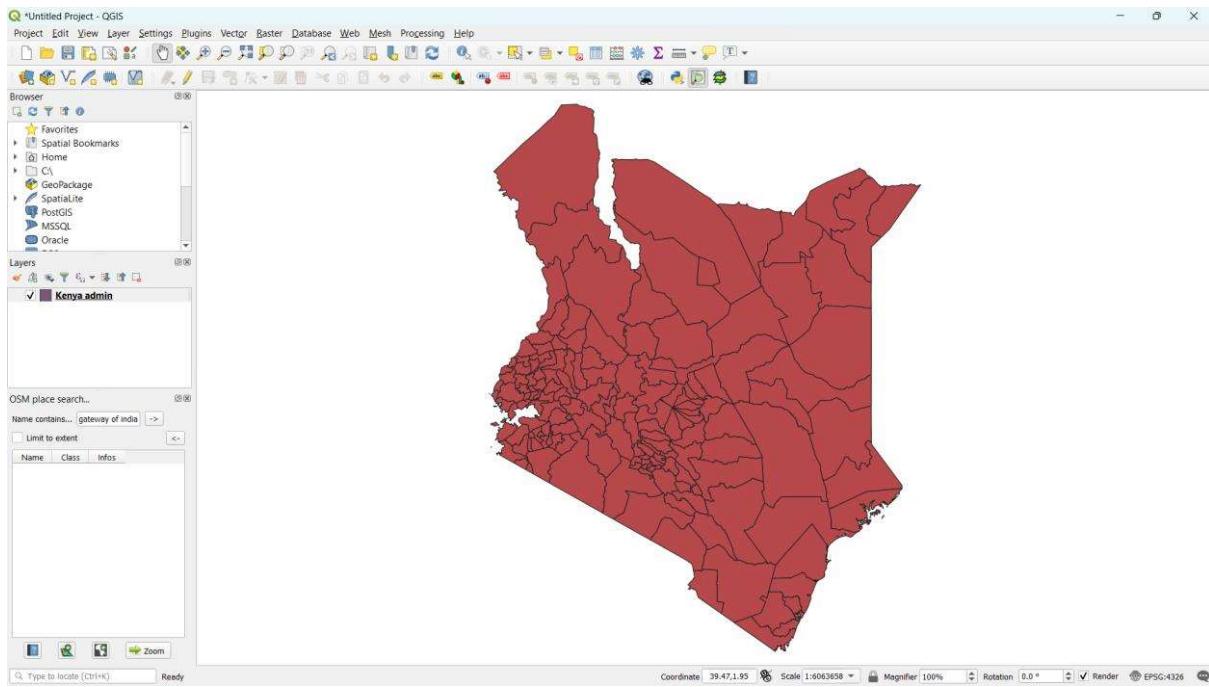
Step 13 – Select T2, type the following function and click enter.

The screenshot shows a Microsoft Excel spreadsheet titled "Kenya_epidemiological_data.xls [Compatibility Mode] - Excel (Product Activation Failed)". A formula is being entered into cell T2. The formula is =IF(AND(H2=0, NOT(P2="")), 1, 0). The formula is highlighted in green, indicating it is valid. The rest of the spreadsheet contains data for various children across different districts and schools, with various health metrics recorded.

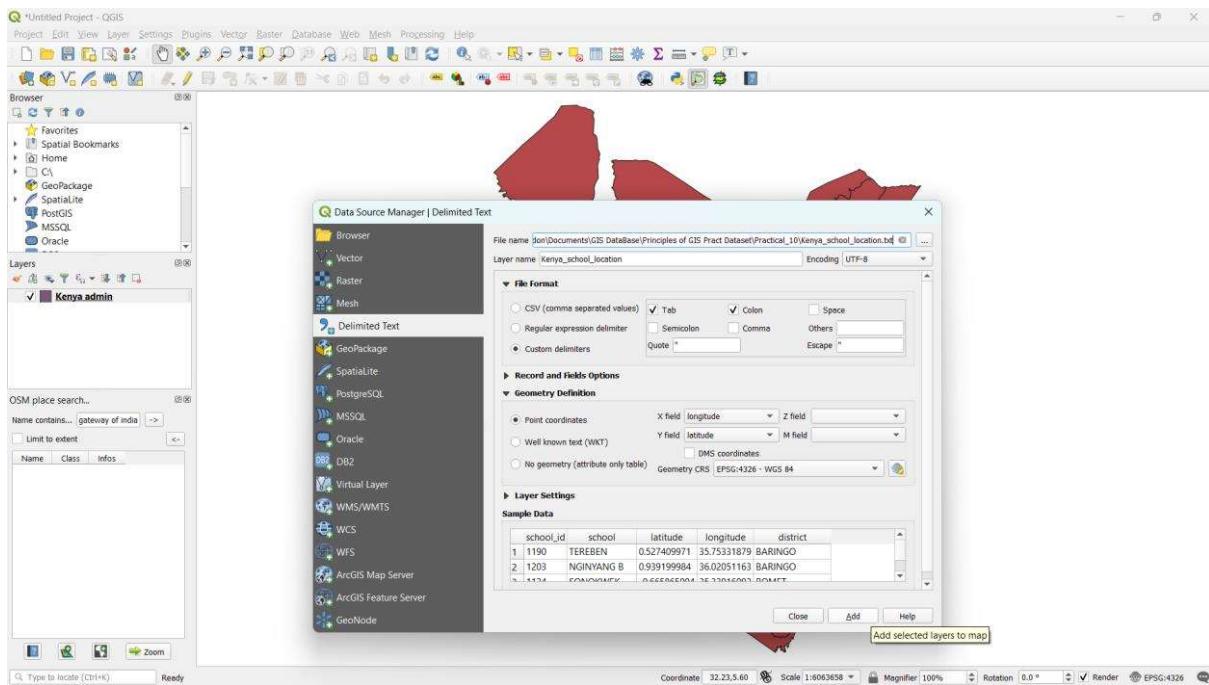
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	child_id	school	school	district	district	age	sex	stool	hookw.	ascaris	trichuris	smansi	hookwi	ascaris	trichuri	anysth.	smansi	dewormi	Ched				
2	70107028	TEREBEN	1190	BARINGO	701	9	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
3	70105018	NGINYANK	1203	BARINGO	701	12	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
4	70105055	NGINYANK	1203	BARINGO	701	11	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
5	70107038	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	1			
6	70105029	NGINYANK	1203	BARINGO	701	7	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
7	70105085	NGINYANK	1203	BARINGO	701	14	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
8	70105010	NGINYANK	1203	BARINGO	701	7	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
9	70105062	NGINYANK	1203	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
10	70107009	TEREBEN	1190	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
11	70105017	NGINYANK	1203	BARINGO	701	12	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
12	70105009	NGINYANK	1203	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
13	70105020	NGINYANK	1203	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
14	70107051	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
15	70107019	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
16	70107032	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
17	70107003	TEREBEN	1190	BARINGO	701	6	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
18	70107036	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
19	70107050	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
20	70107059	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
21	70107024	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
22	70105036	NGINYANK	1203	BARINGO	701	16	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
23	70107058	TEREBEN	1190	BARINGO	701	10	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
24	70107010	TEREBEN	1190	BARINGO	701	8	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
25	70105061	NGINYANK	1203	BARINGO	701	15	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
26	70107056	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
27	70107033	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
28	70107007	TEREBEN	1190	BARINGO	701	7	2	1	0	0	0	0	0	0	0	0	0	0	0	0			
29	70107015	TEREBEN	1190	BARINGO	701	9	1	1	0	0	0	0	0	0	0	0	0	0	1	0			
30	70107045	TEREBEN	1190	BARINGO	701	11	2	0	0	0	0	0	0	0	0	0	0	0	1	1			
31	70107053	TEREBEN	1190	BARINGO	701	10	1	1	0	0	0	0	0	0	0	0	0	0	0	0			
32	70107022	TEREBEN	1190	BARINGO	701	6	1	1	0	0	0	0	0	0	0	0	0	0	0	1			

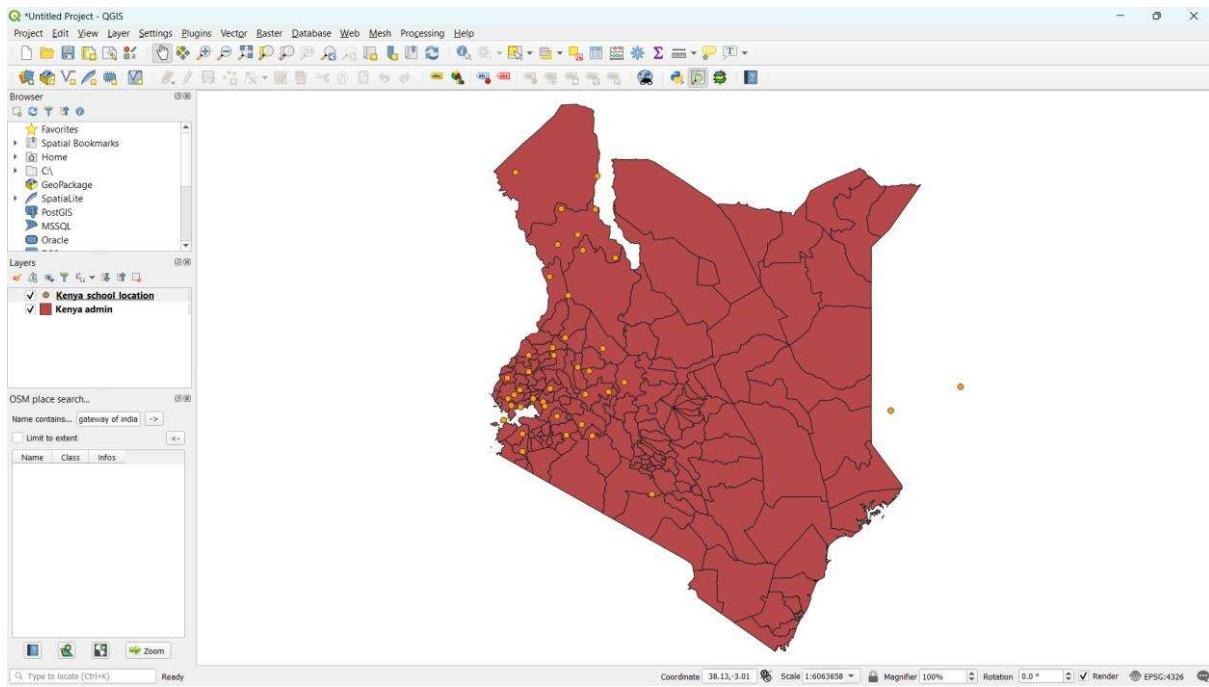
Step 14 – Open a new project in QGIS Desktop. Add the following vector layer.



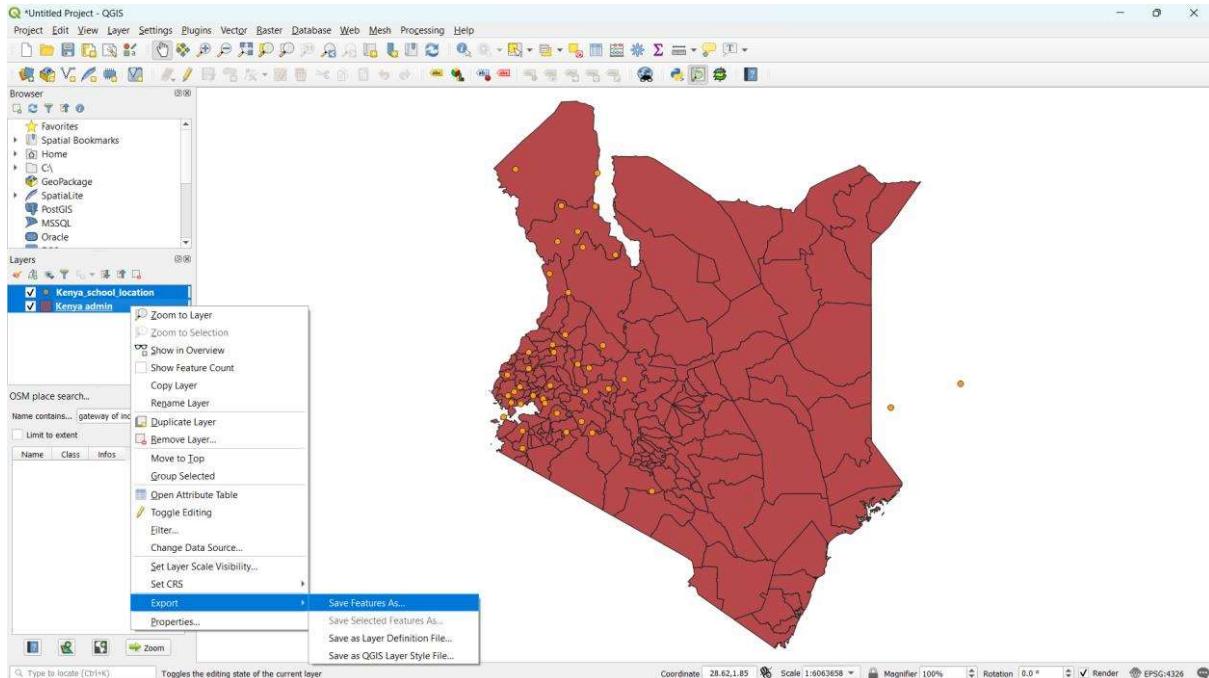


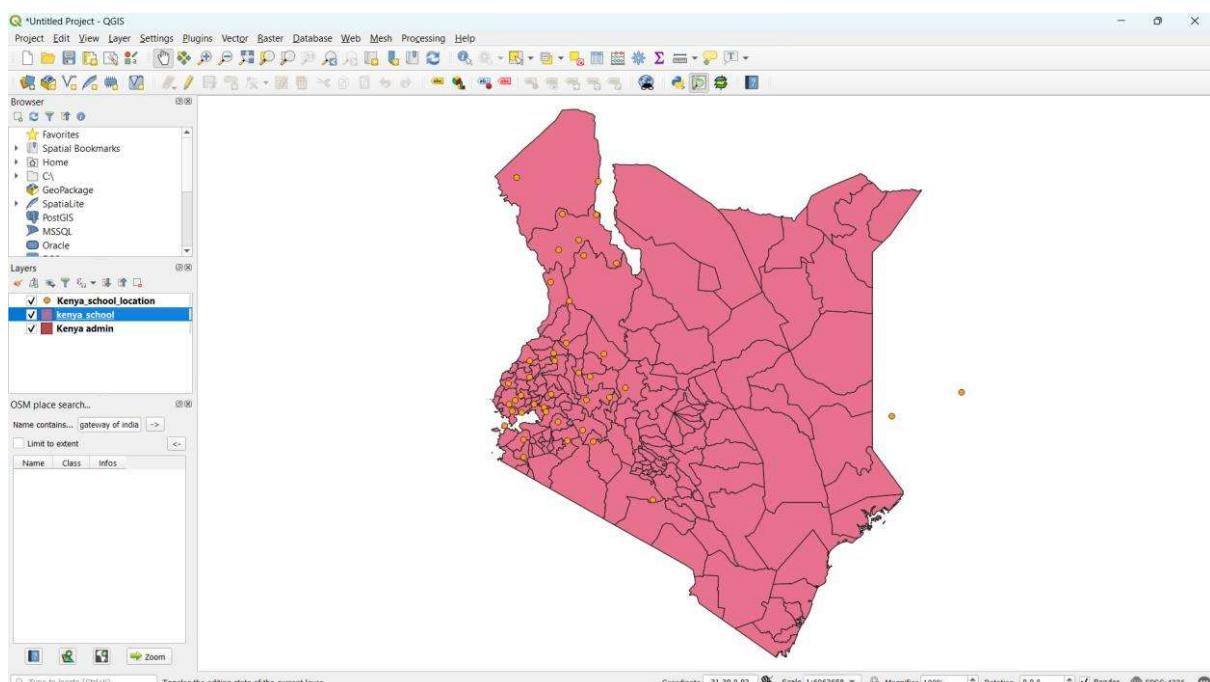
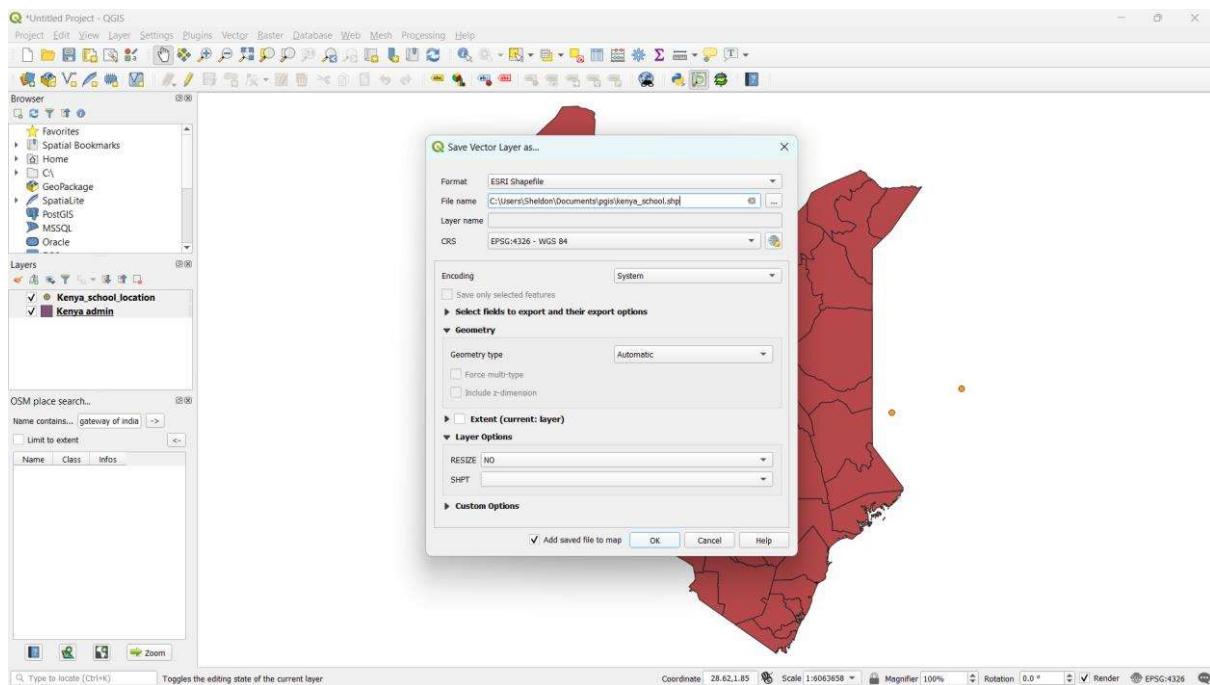
Step 15 – Add the following delimited text layer.



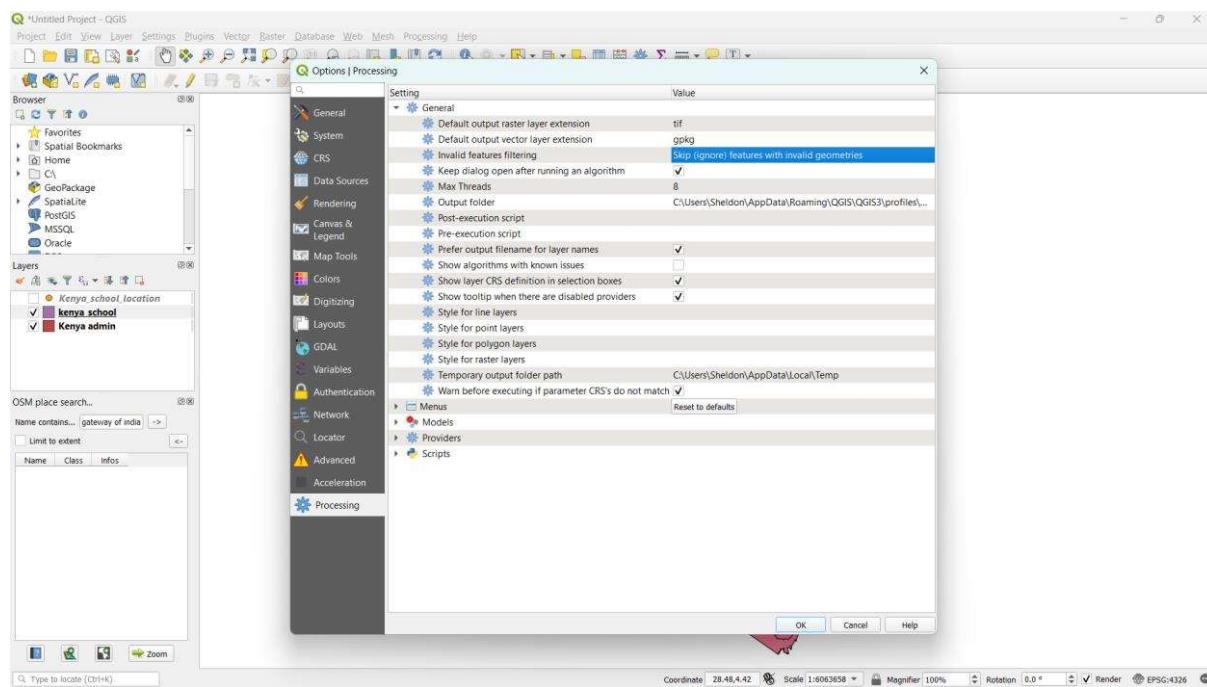


Step 16 – Select both the layers and right click. Export > Save Features as. Set the following parameters and click OK.





Step 17 – Settings > Option. Set the following parameters and then click OK.



Step 18 – Vector > Data Management Tools > Join Attributes by Location. Set the following attributes and hit run.

