GIS Techniques for Monitoring and Evaluation of HIV/AIDS and Related Programs

Exercise 3.1

Join Attribute Data to a Shapefile



*This training was developed as part of a joint effort between MEASURE Evaluation and MEASURE DHS, with funding provided by USAID and PEPFAR.

Exercise 3.1 Introduction

The data joining step is often a prerequisite for mapping and spatial analysis. Creating a color-shaded map (also known as a choropleth map), is a good first step in geographic data visualization and analysis.

Most demographic and health data is stored in tables. In order to make a map of these data, they need to be associated with (joined to) spatial data files.

It is important to consider some basic database rules. Since the shapefile contains multiple, unique records, the data to be linked must be in the same form. If a name field is used for linking, ensure that the spelling is exactly the same. Preferably, linking will be done using codes, such as "District ID", which are stored in both the data table and the shapefile.

Objectives:

- Open and examine a data file containing HIV prevalence data by district for Rwanda.
- Open and examine a shapefile that contains the administrative boundaries for the districts of Rwanda.
- Join the attribute table containing HIV prevalence data to the shapefile, and save the result as a new shapefile.
- Examine the new shapefile and revise the column names to be more meaningful using the Table Manager plug-in.

Requirements:

To complete these exercises, you will need to have QGIS and several plugins installed on your machine, as instructed in exercise 2.1. You will also need to have downloaded and unzipped the data files associated with the exercise. Remember, these files can reside anywhere on your computer, but must be kept together, with the same original file structure.

Exercise 3.1: Join Attribute Data to a Shapefile

In this exercise, you will learn how to join a text file containing HIV prevalence data from the Rwanda DHS 2010 final report to a shapefile representing the districts of Rwanda. You will then learn how to create a color-shaded map to show HIV prevalence by district.

Task Overview: Join an attribute table containing HIV prevalence data to the geometric shapes in a shapefile.

Background:

At some point in time, the majority of indicator data are stored in stand-alone tables. In order to make a map using these data, the data must first be associated with spatial features, such as administrative division boundaries or health facility locations.

As you may recall, a common format for storing spatial data is the shapefile. Before joining attribute data to a shapefile, it is important to take into account certain characteristics of shapefiles:

- The geographic entities described in a shapefile are of type vector-based data, which means they will be collections of points, lines, or polygons.
- Before attribute data, such as HIV indicators, are joined to a shapefile, the shapefile will contain only the physical properties (such as the latitude, longitude, altitude, length, or area) of the geographic entities of interest, along with one or more geographic identifiers, such as administrative name and/or code.
- Before using common geographic identifiers to join attribute data to a shapefile, the attribute data and shapefile data must be referred to at the same geographic level. For example, attribute data for second-level administrative divisions, such as districts, will need to be aggregated prior to joining them to first-level administrative divisions, such as regions or provinces.
- "Shapefiles" are actually a collection of at least three files:
 - Main file, which contains the geographic coordinates that make up the features described (e.g., administrative divisions or health facility locations).
 - Example: regions.shp
 - Index file, which identifies the byte number offset of each main file record from the beginning of the main file.
 - Example: regions.shx

- dBASE table, which contains attributes for features represented in the main file.
 - Example: regions.dbf

Source: http://www.esri.com/library/whitepapers/pdfs/shapefile.pdf

- If a field containing names is used to join attribute data to a shapefile, it is necessary to make sure that the spelling of the names is the exact same in the stand-alone attribute table and the dBASE table that is part of the shapefile.
- To avoid having to identify and correct spelling differences in names when joining attribute tables to shapefiles (for example, "Nyarugenge" vs. "NYARUGENGE" vs. "Nyrau Genge"), it is recommended to base the join on unique administrative codes (for example, district ID = RW0101).

Step 1: Open and examine a file (a data or attribute table) containing HIV prevalence data.

□ Launch QGIS with a blank (new) project document.



You should have a shortcut for QGIS on your desktop. Desktop2011 If yes, click on it to launch QGIS.

If not, on the Windows desktop click on Start > All programs > QGIS Dufour > QGIS Desktop 2.0.1.

- □ Open a comma-delimited file containing HIV prevalence data.
 - A comma-delimited text file (extension .csv) has been created for this exercise, as the comma-delimited format is a common file exchange format that can easily be produced using Excel or other spreadsheet or database programs.
 - When opening a comma-delimited file with the .csv extension, QGIS assumes that all fields (columns) contain strings (text) unless the **field types*** are specified in an accompanying CSVT (CSV template) file.
 - A CSVT file name must have the same file name and path as the CSV file it describes, except the file extension is .csvt instead of .csv. For this exercise, the CSV and CSVT file names are as follows:
 - CSV file: gis_exercises\data_tables\ RWA_DHS2010_TableD91.csv
 - CSVT file: gis_exercises\data_tables\ RWA_DHS2010_TableD91.csvt
 - A CSVT file contains a single line of information that specifies the types of columns contained in the CSV file as well as the maximum size of the values contained in those columns. The contents of the CSVT file for this exercise are as follows:

"string(10)","string(4)","real(4.1)","real(4.1)","real(4.1)"

This information indicates that the CSV file has five columns. The first two columns contain strings (text), with a length of 10 characters and 4 characters, respectively. The last three columns contain real (floating point) numbers with a total length of 4—including the decimal—and a precision (number of digits after the decimal) of 1 (for example, a value of 10.1).

*A "field type" refers to the way a database expects to see data stored in a particular field. The most common field types are "text" (also known as "string") or integer (for numeric whole numbers) or float (for numeric decimal values).



• To view the CSV file in QGIS, on the toolbar for managing layers click on the Add

	V	
Vector Layer button.		

 In the Add vector layer file dialog box that appears (see below), make sure the source type is set to File and click on Browse.

🧭 Add vect	or layer		? ×
Source typ File Encoding	Directory	🔿 Database	○ Protocol
Source			
Dataset		Open	Cancel Help

- In the file dialog box that appears (see below):
 - Use the pull-down menu in the lower-right corner to set the file type to "Comma Separated Value [OGR] (*.csv *.CSV)."
 - Navigate to the folder "gis_exercises\data_tables" and find the commadelimited text file **RWA_DHS2010_TableD91.csv**.
 - Click on the file name RWA_DHS2010_TableD91.csv so that it appears in the file name box.

*	Name	Date modified	Туре
	RWA_DHS2010_TableD91.csv	10/21/2013 1:41 PM	Microso
E			
);) re (D:) —	•		+
File <u>n</u> ame: RWA_DHS2010_Table	D91.csv • Com	ma Separated Value [O	GR] 🔻
		<u>Open</u> Canc	ei



 Click on Open. When the Add vector layer file dialog box reappears with the full path to the CSV file indicated in the Dataset field (see below), click once again on Open.

🌠 Add vect	or layer			? ×
-Source typ	e			
• File	O Directory	O Database		
Encoding	UTF-8			•
Source				
Dataset	V Data Sources\exe	ercises\data\RWA_DHS2010_Ta	ableD91.csv	Browse
		Open	Cancel	Help

□ After clicking on Open for the second time, you should see a new layer in the Layers window corresponding to the comma-delimited text file you just added.

🔏 QGIS	2.0.1-	Dufour						
P <u>r</u> oject	<u>E</u> dit	<u>V</u> iew	<u>L</u> ayer	<u>S</u> et	tings	<u>Plugins</u>	Vect <u>o</u> r	<u>R</u> aster
			8	8	Ą	P	0	🤻 <i>)</i>
1.	Ø	₿	° 🖸	P,	1%		~	
	1000	ananan.		Lay	ers 🛷	ananan ar		ð
V_{\bullet}	·	RW	A DHS2	2010	Table	<u>D91</u>		
œ.								
Po								

• To view the contents of the new attribute layer, right-click on the layer name and select Open Attribute Table (see below).





An alternative way to view the contents of the new data layer is to highlight the data layer name by single-clicking with the left mouse button and then going to the QGIS main menu and selecting Layer > Open Attribute Table (see below).

S.	2.0.1-	Dufour						
:	Edit	View	Layer	Settings	Plugins	Vector	Raster	Data
	_		Ne	w				•
			Em	bed Layers	and Group	os		
	1		V _a Ad	d Vector La	yer	C	trl+Shift+	۰V
			📕 Ad	d Raster La	yer	C	trl+Shift+	HR .
μ.	******		🖫 Ad	d PostGIS L	ayers	C	trl+Shift+	нD
	- L- [🖥 <u>RW</u> A	R Ad	d SpatiaLite	Layer	C	trl+Shift+	HL
			MR 🔍	d MSSQL Sp	atial Laye	r C	trl+Shift+	м
			🔍 Ad	d Oracle Sp	atial Layer	c	trl+Shift+	ю
			🚱 Ad	d WMS/WM	TS Layer	. с	trl+Shift+	w
			🕘 Ad	d WCS Laye	er			
			🕼 Ad	d WFS Laye	er			
			9. Ad	d Delimited	Text Laye	r		
			Co	py style				
			🖹 Pa	ste style				
				en Attribute	e Table			
			/ To	aale Editina				

After opening the attribute table, you should see a table like the following.

\$	🚀 Attribute table - RWA_DHS2010_TableD91 :: Features total: 30, filtered: 30, sel 📼 📼 🗮 🎫							
	? 📰 📰 📰 📢 💱 🎯 🛅 🖕 🤹 ?							
	DHSDIST 🗸	GOVTCODE	HIVPREVF	HIVPREVM	HIVPREVT			
0	Nyarugenge	0101	9.8	6.8	8.3			
1	Gasabo	0102	8.7	4.1	6.4			
2	Kicukiro	0103	10.1	5.5	7.9			
3	Nyanza	0201	2.1	2.2	2.1			
4	Gisagara	0202	1.4	0.9	1.1			
5	Nyaruguru	0203	1.3	0.5	0.9			
6	Huye	0204	4.2	2.7	3.5			
7	Nyamagabe	0205	2.9	2.8	2.8			
8	Ruhango	0206	3.4	1.6	2.5			
9	Muhanga	0207	3.9	1.6	2.9			
10	Kamonyi	0208	4.4	1.7	3.1			
11	Karongi	0301	3.4	3.3	3.3			
10	Ruteiro	0202	27	50	2 /			
	Show All Features	.]						

The attribute table you added contains five fields (columns):

- DHSDIST: The name of the district for which data are reported in Table D.91 in the Rwanda DHS 2010 final report. Data type is string. Maximum length is 10. (refer back to p.5 for a review of the CSVT file which, as you may recall, contains this data type information.)
- GOVTCODE: A unique administrative code assigned to each district by the Rwandan government. Data type is string, NOTE; even though these are numbers, the first value is actually the numeral "0". Sometimes numerals used as IDs are stored as a "string" data type. No numeric calculations can be performed on these data. Maximum length is 4.
- HIVPREVF: HIV prevalence (percent HIV positive) among females of reproductive age (age 15-49) in 2010. Data type is real (floating point). Maximum length is 4. Maximum number of decimal places is 1.
- HIVPREVM: HIV prevalence among males of reproductive age in 2010. Data type is real (floating point). Maximum length is 4. Maximum number of decimal places is 1.
- HIVPREVT: HIV prevalence among the total adult population (females and males) of reproductive age in 2010. Data type is real (floating point). Maximum length is 4. Maximum number of decimal places is 1.

Note that the DHS district names are spelled using initial capital letters. For the government codes, the first two digits represent the province and the last two



digits represent the district. For example, districts 0101, 0102, and 0103 (Nyarugenge, Gasabo, and Kicukiro, respectively), all fall into the province that contains the City of Kigali (province ID = 01).

Step 2: Open and examine a shapefile that contains the administrative boundaries for the districts of Rwanda.

Once again, in QGIS click on the Add Vector Layer button. You should see the Add vector layer dialog box.

🌠 Add vect	or layer		? ×
Source typ	pe		
• File		 Database 	O Protocol
Encoding	UTF-8		•
Source			
Dataset			Browse
		Open	Cancel Help

 Click on the Browse button, and in the file dialog window that appears, use the pull-down menu in the lower-right corner to set the file type to "ESRI Shapefiles [OGR] (*.shp *.SHP)." Navigate to the folder "gis_exercises\shapefiles" and find the shapefile RWA_districts.shp. Click on the file name RWA_districts.shp so that it appears in the file name box. The file dialog box should resemble the one below.

	•	Name	Date m	odified	Туре
	II	RWA_districts.shp	2/14/20	012 10:15 AM	SHP File
5)	< III			•
File <u>n</u> ame: RWA_districts.shp			ESRI Shapefile: <u>O</u> pen	s [OGR] (*.shp Canc	*.S ▼ el



- Click on Open. When the Add vector layer file dialog box reappears with the full path to the shapefile indicated in the Dataset field, click once again on Open.
- You should see a new layer in the Layers window corresponding to the shapefile you just added.



□ To view the contents of the attribute table for the shapefile just added, right-click on the layer name and select Open Attribute Table (see below).

%	🔏 Attribute table - RWA_districts :: Features total: 30, filtered: 30, selected: 0								
		E 🛃) 🚳 🐯 🎾						?
	OBJECTID 🗸	GovtDistri	PROVINCE	DISTRICT	AREA_SQ_KM	PERIMETER_	POPULATION	Shape_Leng	Shape_Area
0	1	0404	NORD	BURERA	647.6230000000	162.8100000000	324653.0000000	1.464418119560	0.052395556233
1	2	0403	NORD	MUSANZE	532.8750000000	116.48199999999	299683.0000000	1.047734626490	0.043112619032
2	3	0103	VILLE DE KIGALI	KICUKIRO	167.4950000000	82.34210000000	204962.0000000	0.740730185697	0.013554958738
3	4	0301	OUEST	KARONGI	997.6259999999	192.7410000000	265487.0000000	1.734042146808	0.080742629718
4	5	0205	SUD	NYAMAGABE	1095.430000000	204.050000000	284852.0000000	1.835887446403	0.088674623992
5	6	0306	OUEST	RUSIZI	962.8750000000	257.4370000000	331950.0000000	2.316306846381	0.077953079939
6	7	0502	EST	NYAGATARE	1931.049999999	280.0910000000	255104.0000000	2.519292301606	0.156221970480
7	8	0503	EST	GATSIBO	1718.490000000	247.4319999999	240566.0000000	2.225708711195	0.139043684719
8	9	0304	OUEST	NYABIHU	533.9829999999	127.8580000000	268367.0000000	1.150092335458	0.043205296775
9	10	0303	OUEST	RUBAVU	390.1309999999	115.2920000000	287579.0000000	1.037060276863	0.031566271862
10	11	0402	NORD	GAKENKE	707.3940000000	146.6860000000	322043.0000000	1.319447255202	0.057237784654
11	12	0305	OUEST	NGORORERO	682.1699999999	171.7059999999	282249.0000000	1.544612002896	0.055202171853
12	13	0302	OUEST	RUTSIRO	1162.619999999	184.8650000000	269434.0000000	1.663042445640	0.094082074204
13	14	0101	VILLE DE KIGALI	NYARUGENGE	134.5860000000	106.3130000000	236990.0000000	0.956362847325	0.010891596903
14	15	0505	EST	KIREHE	1190.279999999	191.4290000000	229468.0000000	1.722201652990	0.096340838955
15	16	0507	EST	BUGESERA	1296.670000000	204.8180000000	266775.0000000	1.842670327904	0.104952104625
16	17	0506	EST	NGOMA	871.801000000	163.0329999999	232165.0000000	1.466785713926	0.070560658520
17	18	0307	OUEST	NYAMASHEKE	1179.299999999	181.4960000000	325032.0000000	1.632919480310	0.095459734200
18	19	0204	SUD	HUYE	584.2649999999	130.46999999999	264295.0000000	1.173874747922	0.047300079320
19	20	0202	SUD	GISAGARA	682.4059999999	149.103000000	266424.0000000	1.341473705963	0.055249328463
20	21	0501	EST	RWAMAGANA	685.1720000000	135.5790000000	209423.0000000	1.219651495015	0.055448275666
21	22	0504	EST	KAYONZA	1813.210000000	196.9290000000	220802.0000000	1.771475399563	0.146726570834
22	23	0201	SUD	NYANZA	675.31799999999	175.7210000000	176456.0000000	1.580989569265	0.054663730824
23	24	0206	SUD	RUHANGO	629.7380000000	163.6839999999	210000.0000000	1.472662258286	0.050969261878
24	25	0207	SUD	MUHANGA	650.77499999999	179.2119999999	340369.0000000	1.612143453993	0.052664092036
25	26	0208	SUD	KAMONYI	658.639000000	169.800000000	292772.0000000	1.527499154389	0.053302180684
26	27	0405	NORD	GICUMBI	833.47699999999	199.853000000	362668.0000000	1.797638086922	0.067436969841
27	28	0102	VILLE DE KIGALI	GASABO	431.2409999999	110.15999999999	318569.0000000	0.990987147546	0.034896897795
28	29	0401	NORD	RULINDO	569.6770000000	159.1949999999	278310.0000000	1.431975750659	0.046095584951
29	30	0203	SUD	NYARUGURU	1014.970000000	188.627000000	233815.0000000	1.697339066385	0.082179775426
	Show All Features]							

The attribute table for the shapefile you added contains nine fields:

- OBJECTID: This is a unique object identifier for each row in the attribute table, which is assigned automatically by ArcGIS when an ESRI feature class is created in a geodatabase. This field exists in the shapefile because the shapefile was derived from an ESRI geodatabase dated April 2011, which was downloaded from the Web site for the Rwanda Ministry of Health (www.moh.gov.rw).
- GovtDistri: A unique administrative code assigned to each district by the Rwandan government. The values in this field are the same as those in the field GOVTCODE contained in the comma-delimited text file added to the QGIS project earlier in this exercise.
- PROVINCE: French province names from the MOH geodatabase.
- DISTRICT: District names in Kinyarwanda, the native language of Rwanda. These are the same district names used in the Rwanda DHS 2010.
- AREA_SQ_KM: Surface area of each district calculated in square kilometers.
- PERIMETER_: The length of the perimeter of each district in kilometers.
- POPULATION: Total district population from the 2000 census.

- Shape_Leng: The length of the perimeter of each district in decimal degrees. Note that the decimal degree, which represents an unprojected coordinate system, is an inconsistent and imprecise unit of measurement for calculating distances. The Shape Length field is created and updated automatically in ArcGIS feature classes, but not in shapefiles.
- Shape_Area: Surface area of each district in decimal degrees. The Shape Area field is created and updated automatically in ArcGIS feature classes, but not in shapefiles.

NOTE: The DISTRICT column in the shapefile table contains the same administrative names as the DHSDIST column in the attribute table containing HIV prevalence data from the Rwanda DHS 2010. The district names in the shapefile table, however, are spelled in all upper case letters. This is different from the district names in the attribute table containing the HIV prevalence data you want to join to the shapefile. As a result, you cannot use the district names as they are currently formatted, to join the HIV prevalence data to the shapefile.

To resolve the problem you have two primary options: (1) create a new column in one of the tables to contain district names with the same capitalization as in the other table or (2) use an administrative code the two tables have in common. Fortunately, you have a common administrative code that can be used to join the two tables, GOVTCODE in the HIV prevalence table and GovtDistri in the shapefile table.

Step 3: Join the attribute table containing HIV prevalence data to the shapefile.

- □ First, open the Properties dialog box for the shapefile layer.
 - To do that, you have two options: (1) double-click on the layer name for the shapefile (RWA_districts) in the Layers panel, or (2) right-click on the shapefile layer name in the Layers panel and on the popup menu select "Properties."
 - When the Properties window opens, highlight the "Joins" icon in the left-hand navigation panel by left-clicking on it (see below).

<u>R</u>	Layer	Properties - RWA_di	strie	:ts		
	$\mathbf{\hat{x}}$	General	Jo	ins		
	~	Style		Join layer	Join field	Target field
	abc	Labels				
		Fields				
	\bigcirc	Display				
I,	Ô	Actions				
	•	Joins				
17		Diagrams				
	i	Metadata				

□ Then, on the Joins screen within the Layer Properties window, click on the green plus sign (see below).

Layer Properties - RWA_di	stricts
General	Joins
😻 Style	Join layer
(abc) Labels	
Fields	
🥟 Display	
Actions	
0 Joins	
Diagrams	
🕡 Metadata	•

• In the "Add vector join" file dialog box that appears, use the pull-down menus to specify the join layer, join field, and target field as shown below. (NOTE: you will see a box checked that says "Cache join layer in virtual memory". The joined table will only be temporary and will need to be saved separately if it is to be used for other tasks in the future. This will be done at a later step.)



🧭 Add vector join	? <mark>×</mark>	
Join layer	RWA_DHS2010_TableD91	
Join field	GOVTCODE	
Target field	GovtDistri	
🕱 Cache join layer in virtual memory		
Create attribute in	ndex on join field	
[[OK Cancel	

• Click on OK and inspect the Layer Properties window to verify that the join has been accomplished (see below).

🔏 Layer Properties - RWA_districts			
General	Joins		
😻 Style	Join layer Join field Target field		
abc Labels			
Fields			
Display			
Actions			
Joins			
Diagrams			
Metadata			

□ To view the result of the join, click on OK once again to close the Layer Properties window. Next, right-click on the name of the shapefile layer (RWA_districts) in the left-hand Layers panel and choose "Open Attribute Table" (see below).





• When the shapefile's attribute table opens, scroll to the right to view the new columns containing HIV prevalence data (see below).

	A REAL PROPERTY AND A REAL			×
				?
RWA_DHS2010_TableD91_DHSDIST	RWA_DHS2010_TableD91_HIVPREVF	RWA_DHS2010_TableD91_HIVPREVM	RWA_DHS2010_TableD91_HIVPREVT	
Burera	6.0	0.6	3.5	1
Musanze	3.3	2.1	2.7	
Kicukiro	10.1	5.5	7.9	
Karongi	3.4	3.3	3.3	
Nyamagabe	2.9	2.8	2.8	
Rusizi	2.8	2.8	2.8	
Nyagatare	2.4	1.4	1.9	
Gatsibo	1.2	0.5	0.9	
Nyabihu	2.1	3.4	2.7	
Rubavu	4.3	1.3	2.8	
Gakenke	0.5	2.5	1.4	
Ngororero	2.6	1.4	2.1	
Rutsiro	3.7	3.0	3.4	
Nvarugenge	9.8	6.8	83 	

Note that the column names are very long, as they contain a combination of the input file name (RWA_DHS2010_TableD91) and the column names in the input file (DHSDIST, HIVPREVF, etc.).

Note also that the join is temporary; the variables for HIV prevalence will not be permanently added to the shapefile table unless you save out a new shapefile using the results of this join.

Step 4: Create a new shapefile from your joined data, and add it to the QGIS project.

To make the join permanent, save a copy of the shapefile containing the temporarily joined HIV prevalence data. To do so, perform the following tasks:

- Close the shapefile attribute table by clicking on the X in the red box in the upper-right corner of the attribute table.
- Right-click on the name of the shapefile layer in the Layers panel and select "Save As."



- In the Save As dialog box, choose the following settings:
 - Format: ESRI Shapefile
 - Save as: Session gis_exercises\shapefiles\RWA_districts_HIV.shp
 - Add saved file to map: Click on this box so that an X appears in it.
- □ Leave the remaining settings as they are so that the Save As dialog box looks like the one below.

🔏 Save vector layer as		
Format	ESRI Shapefile	•
Save as	ses/shapefiles/RWA_districts_HIV.shp	Browse
Encoding	UTF-8	•
CRS	Layer CRS	•
	WGS 84	Browse
Symbology export	No symbology	•
Scale	1:50000	
OGR creation options		
Data source		
Layer		
Skip attribute creation Add saved file to map		
	OK Cancel	Help

Note: For now it is important to make sure that the choice for Encoding is "UTF-8" and the choice for CRS is "Layer CRS." This will keep the output file in an unprojected coordinate reference system (CRS), which stores features using latitude and longitude in degrees.

 Click on OK and verify that the new shapefile has been added to the Layers panel in QGIS.

Step 5: Revise the new column names to be more meaningful.

□ If you recall, the new column names generated by QGIS automatically in the shapefile table during the join were very long. In the process of creating the new shapefile after the join, however, QGIS automatically shortened these column

names to a maximum of 10 characters. This created column names that are not very descriptive and therefore not very useful. In this step we will use Table Manager to create new, more useful names for these columns.

• To verify the new column names corresponding to the HIV prevalence indicators, right-click on the name of the new shapefile layer (RWA_districts_HIV) in the Layers panel and scroll to the right. You should see columns with abbreviated names as shown in the column headings below.

RWA_DHS201	RWA_DHS2_1	RWA_DHS2_2	RWA_DHS2_3
------------	------------	------------	------------

 To change the column names, close the shapefile table, highlight the name of the new shapefile layer in the Layers panel, and start the **Table Manager** plug-in by going to the main menu in QGIS and selecting Vector > Table Manager > Table Manager (see below).



<u>NOTE</u>: If you do not see the Table Manager plug-in on the Vector menu, please refer back to Exercise 2.1, the "installing plugins" section.

• In the Table Manager window, highlight each new variable name, click on the

Rename button Rename, and make the following changes:

- From RWA_DHS201 to DHSDIST. This is the DHS district name.
- From RWA_DHS2_1 to HIVPREVF. This is female HIV prevalence.
- From RWA_DHS2_2 to HIVPREVM. This is male HIV prevalence.
- From RWA_DHS2_3 to HIVPREVT. This is total HIV prevalence.
- Save the changes to the shapefile table by clicking on the Save button at the bottom of the Table Manager dialog box.

- A new dialog box will open, which will ask you whether you want to keep the layer style currently applied. Answer Yes and then click on the Close button at the bottom of the Table Manager dialog box.
- Verify the changes to the shapefile table by right-clicking on the shapefile layer name in the Layers panel, selecting Open Attribute Table, and scrolling to the right (see below).

DHSDIST	HIVPREVF	HIVPREVM	HIVPREVT
Burera	6.0	0.6	3.5
Musanze	3.3	2.1	2.7
Kicukiro	10.1	5.5	7.9
Karongi	3.4	3.3	3.3
Nyamagabe	2.9	2.8	2.8
Rusizi	2.8	2.8	2.8
Nyagatare	2.4	1.4	1.9
Gatsibo	1.2	0.5	0.9
Nyabihu	2.1	3.4	2.7
Rubavu	4.3	1.3	2.8
Gakenke	0.5	2.5	1.4

□ You now have a shapefile containing attributes for each district which you can use for mapping and analysis of HIV data.

Step 6: Save the QGIS project.

□ To save the QGIS project with which you have been working, go to the main QGIS

menu and select Project > Save or click on the Save button.

• In the file dialog box that appears, specify the following output name:

Session 3 HIV Data Sources/exercises/qgis_exercise_3.1.qgs

- □ Quit QGIS.
 - To close QGIS, go to the main menu and select Project > Exit QGIS or use a combination of the Ctrl and Q keys (see below).



NOTE on joining other types of data files:

QGIS can also join other types of data files to shapefiles. All that is required is that both files (the attribute data and the shapefile) have a field in common on which to perform the join (such as a geographic identifier code).

• Use Excel files (with the file name extension **.xlsx** or **.xls**), or DBF files, which are dBASE files with the file name extension **.dbf**.

To use these files without losing the intended formatting for each of the data fields, make sure the column types (cell formats) are specified in the source files before attempting to join them.

To add one of these file to QGIS, use the following steps:

- Click on the Add Vector Layer button.
- In the Add Vector Layer dialog box, click on Browse.
 Browse
- In the file dialog box that opens, change the file to "All files (*)(*.*)."

All files (*) (*.*)	•
Open	Cancel

• Specify the file name and click Open twice to open the file. Proceed exactly as you did when joining a shapefile to a CSV file.

END