Optimizing Data Quality and Use

Free Analysis, Visualization and Reporting (AVR) Software Selection Tool

This tool helps planners select a free Analysis, Visualization, and Reporting (AVR) software product appropriate for the type of data being used. It also outlines some of the necessary skill sets of the data analyst and additional system requirements to use these products.

Note: Several software products requiring licensing fees (e.g., ESRI ArcGIS, Pentaho, SAS, STATA, SUDAAN) were intentionally excluded from this tool. This list is by no means an exhaustive collection of all open source AVR products relevant to epidemiological data analysis. PHII welcomes further suggestions to the development of this tool. Please send feedback to info@phii.org.

Product Name	Product Summary	Pros	Cons	System/Other Software Requirements and Data Formats	Skills Needed	Source
Epi Info 7	Different modules allow for survey creation and data capture. The Analysis module can be used with data imported from 24 different formats.	Available in desktop (Microsoft Windows), mobile and web/cloud versions; Epi Map and Epi Report allow for additional visualization and formatting; available in 16 different languages	Not available on Mac or Linux operating systems (although these projects are underway); limitations in ability to support various ANOVA and regression methods	Microsoft Windows XP or newer; Microsoft .NET 4.0 Accepted data formats include: ASCII text file, Microsoft Access, Microsoft Excel, Microsoft SQL Server Database	Basic to intermediate knowledge of statistics; familiarity with Boolean expressions and some knowledge of Microsoft Access helpful; mapping, map projections and GIS basics if using Epi Map	Centers for Disease Control and Prevention. 2016. "Epi Info." https://www.cdc.gov/epiinfo/index.html.
GRASS GIS	Geographic Resources Analysis Support System, commonly referred to as GRASS, is a software suite used for geospatial data management and analysis, image processing, graphics and maps production, spatial modeling, and visualization.	Can be used as either a stand-alone application or a backend for other software packages (e.g., R); interfaces allow for transfer of GRASS data into R and for results to be returned to GRASS. GRASS GIS 7 offers a new Python interface for application creation; Handles raster, vector, image processing, and	Support subject to the constraints of volunteer community support (commercial contractor options do exist in many countries)	Available for Microsoft Windows XP (requires Microsoft Visual C++ Redistributable Packages) or newer; Mac OS X 10.8 or newer, or Linux operating systems Accepted data formats include: Microsoft Access, MySQL, ODBC, Oracle, PostgreSQL, SQLite	Mapping, map projections, GIS basics, and spatial analysis	GRASS Development Team. 2016. "GRASS GIS: Bringing Advanced Geospatial Technologies to the World." https://grass.osgeo.org/.

Product Name	Product Summary	Pros	Cons	System/Other Software Requirements and Data Formats	Skills Needed	Source
		graphic data.				
QGIS	An open source geographic information system that provides data viewing, editing, and analysis	Desktop, spatial file browser, server application, and web applications exist currently; Android Mobile Application is in development; available in multiple languages; R and GRASS plugins available for advanced spatial statistics	Lack of 3D integration; much of the symbology palette must be downloaded opposed to pre-existing options; support subject to the constraints of volunteer community support (commercial contractor options do exist in many countries)	Available for Microsoft Windows, Mac OS X, and Linux operating systems Accepted data formats include: AutoCAD DXF, ENC, geodatabase, MapInfo, Microstation, MSSQL Spatial databases, Oracle Spatial, Shapefiles, SpatiaLite, WellKnownText (WKT)	Mapping, map projections, GIS basics, and spatial analysis	QGIS Community. 2016. "QGIS." http://www.qgis.org/en/site/about/index.html.
R	A statistical computing language and environment for data manipulations, calculation and graphical display	Ease of design for publication-quality plots; strong user community and package library	Memory management (data must be stored in physical memory), speed and efficiency are shortcomings; cannot be directly embedded into a web browser (see R-Shiny)	Available for Microsoft Windows, Mac OS X, and Linux operating systems Accepted data formats include: ASCII text file, CSV, Microsoft Excel, Minitab, SPSS	Intermediate knowledge of statistics; the syntax-based analyses will present a greater learning curve for users familiar with a menu-driven interface without knowledge of programming fundamentals; advanced users can write C code to manipulate R objects directly	The R Foundation. 2016. "R: The R Project for Statistical Computing." https://www.r-project.org/.
R-Shiny	A platform as a service (PaaS) from RStudio that allows for building web applications using R	Created for R users with no web development experience; can also be written in HTML, CSS, or JavaScript for additional flexibility	Depending on use case/business need, may incur a fee	Installation requires R and a web connection; web deployment requires servers or a hosting service	Requires functional knowledge of the R statistical language	RStudio. 2016. "Shiny." http://shiny.rstudio.com/. RStudio. 2016. "Shiny - Tutorial." http://shiny.rstudio.com/tutorial/. RStudio. 2016. "Cheatsheets – RStudio." https://www.rstudio.com/r

Product Name	Product Summary	Pros	Cons	System/Other Software Requirements and Data Formats	Skills Needed	Source
						esources/cheatsheets/.
SaTScan	Software for the spatial, temporal, and space-time scan statistics, or SaTScan, allows users to perform geographical surveillance and evaluate statistical significance of disease clusters	Possible to scan multiple datasets simultaneously; allows for Poisson-based, Bernoulli, space-time permutation, ordinal, exponential and normal models	Initial data formatting may be cumbersome for first-time users; tutorials exist, but there does not apu group	Microsoft Windows Requirements (also available for Mac X 10.6 or newer and Linux): Vista or newer, Intel or a compatible processor, 20 MB hard drive space and Oracle Java Runtime Environment 6.0 or higher. Accepted data formats include: CSV, dBase, Shapefiles, SaTScan ASCII text file	Mapping, map projections, GIS basics, and spatial analysis	Kulldorff, Martin, and Information Management Services Inc. 2005. "SaTScan - Software for the Spatial, Temporal, and Space-Time Scan Statistics." http://www.satscan.org/. Kulldorff, Martin. 2015. "SaTScan TM User Guide." http://www.satscan.org/.
Tableau Public 9.3	Business Intelligence software with a menu-driven interface designed to share interactive visualizations on the web	"Show Me" feature offers one-click options when user knows what data they need for analysis but need assistance creating an effective view; utilizes OpenStreetMap	Limited data source options compared to paid products; files may only be saved to your Tableau Public profile, not your computer/network; as results must be posted to the web to be shared, not appropriate for confidential data	Server-side requirements: (Microsoft Windows Requirements, also available for Mac) Microsoft Windows Vista SP2 or newer, Microsoft Windows Server 2008 R2 or newer, minimum of Intel Pentium 4 or AMD Opteron processor, Internet Explorer 8 or newer Client-side requirements: Android Browser 3.2 or newer, Internet Explorer 8 or newer, Mozilla Firefox 3.x or later, Google Chrome, Tableau mobile apps	Excel proficiency highly desired; online trainings sufficient with working knowledge of Excel graphics options	Tableau Software Incorporated. 2016. "Tableau Public." https://public.tableau.com/s/. Tableau Software Incorporated. 2016. "Tableau Training and Tutorials." http://www.tableau.com/learn/training.

Product Name	Product Summary	Pros	Cons	System/Other Software Requirements and Data Formats	Skills Needed	Source
				Accepted data formats include: ASCII text file, Microsoft Access, Microsoft Excel, Microsoft SQL Server, Google Analytics, MySQL, Oracle, Tableau Server		
WEAVE	The Web-based Analysis and Visualization Environment, or Weave, is designed for users of any skill level to integrate, analyze and visualize data with the option to disseminate results in a web page	Multiple examples available from local governments currently using the platform; utilizes OpenStreetMap	Installation requires multiple software downloads which may present challenges to users without IT administrative rights	Available for Microsoft Windows, Mac OS X, and Linux operating systems Server-side requirements: MySQL, PostGreSQL, Oracle, or SQL Server; Java Development Kit (includes Java Runtime Environment); Tomcat or GlassFish; any web browser with the Adobe Flash Player plugin version 10.0 or newer Client-side requirements: any web browser with the Adobe Flash Player plugin version 10.0 or newer Accepted data formats include: CVS, .dbf component of an ESRI Shapefile, Microsoft Excel, native WEAVE datasets, WFS	Mapping, map projections and GIS basics if using GIS features	University of Massachusetts Lowell. 2016. "Weave (Web-Based Analysis and Visualization Environment)." https://oicweave.org/index.php. University of Massachusetts Lowell. 2015. "Weave User Guide." http://info.iweave.com/projects/weave/wiki/Weave_User_Guide.