

Information Technology Infrastructure and Set Up

IT infrastructure

The Kaiser Permanente Southern California (KPSC) Information Technology team developed a virtual Linux server (RedHat Enterprise Linux operating system). This Linux server runs the R shiny server, software and code that allows to create the mapping application. R shiny and the software used to develop the Heart Disease Data Portal (HDDP) is explained in detail below. The HDDP Linux server is public facing but located behind a firewall.

Dataset format

Hypertension prevalence estimates are calculated as described in the document [How hypertension prevalence was estimated](#). Any prevalence estimates that are based on numbers that are too small to be released are deleted from the final dataset. More details on that process are available in the document [How we ensure data privacy and safety](#). The dataset itself is structured in long format, in other words, each prevalence estimate for each health district and age-sex-race/ethnicity-year combination has its own row. Our program requires that the data is saved as comma separate values (csv) file. However, any file type that R can read could be used.

Open-Source Software

The KPSC-RAND Heart Disease Data Portal was created using a powerful suite of open-source software tools all of which can be accessed by anyone. Below we describe these tools and where to locate them. We also share computer code that was used to create our mapping portal (below and on our webpage).

The mapping application was developed using the R Shiny web framework.¹ R is a programming language for statistical computing² that also provides access to high-quality tools for creating statistical graphics, including maps. Many such tools are not available in the core R language, but rather in extensions of R provided in external code libraries, also referred to as packages. The primary repository of R libraries is the Comprehensive R Archive Network (CRAN);³ although many developers also post their own R libraries on GitHub. GitHub is a code hosting platform that allows users to store and share code.

R Shiny is a web framework that allows programmers to create websites that run R in the background for preparing data and generating output such as graphics and tables that are then presented on a website. The Heart Disease Data Portal was created using R Shiny and RStudio. We use an R-Shiny Server from RStudio.⁴ Shiny Server is a software for hosting web applications that were developed using R Shiny. The Shiny server that creates the web application is also hosted on the Linux server along with all other software and the prevalence estimates. The map itself was created using the leaflet library for R,⁵ which provides an R interface to the Leaflet JavaScript library. Extensions to the map were custom-built for the HDDP using JavaScript and the shinyjs⁶

¹ <https://shiny.rstudio.com/>

² R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

³ <https://cran.r-project.org/>

⁴ <https://www.rstudio.com/products/shiny/shiny-server/>

⁵ <https://rstudio.github.io/leaflet/>

⁶ Dean Attali (2020). shinyjs: Easily Improve the User Experience of Your Shiny Apps in Seconds. R package version 2.0.0. <https://CRAN.R-project.org/package=shinyjs>

R library. Other R libraries used in the app include dplyr,⁷ ggplot2,⁸ glue,⁹ htmltools,¹⁰ htmlwidgets,¹¹ magrittr,¹² readxl,¹³ sf,¹⁴ shinydashboard,¹⁵ stringr,¹⁶ and tidy.¹⁷

The maps used in the HDDP app are from OpenStreetMap¹⁸ and Carto.¹⁹ Geographic shapefiles for Los Angeles County Health Districts were downloaded from the Los Angeles County website.²⁰ Shapefiles for freeways in LA County were downloaded from the City of Los Angeles GeoHub.²¹

The R-Code for the mapping application is available at

<https://github.com/RANDCorporation/KPSC-RAND-Heart-Disease-Data-Portal>

Website

The KPSC Research Communication team developed the HDDP homepage using WordPress. The HDDP web page is located on the KPSC Department of Research & Evaluation's public website. It provides access to the mapping portal, all technical documentation and the code to generate the portal.

⁷ Wickham H, François R, Henry L, Müller K (2022). dplyr: A Grammar of Data Manipulation.

<https://dplyr.tidyverse.org>, <https://github.com/tidyverse/dplyr>.

⁸ H. Wickham. ggplot2: Elegant Graphics for Data Analysis.

Springer-Verlag New York, 2016.

⁹ Jim Hester (2021). glue: Interpreted String Literals. R package version 1.5.0. <https://CRAN.R-project.org/package=glue>

¹⁰ Joe Cheng, Carson Sievert, Barret Schloerke, Winston Chang, Yihui Xie and Jeff Allen (2021). htmltools: Tools for HTML. R package version 0.5.2. <https://CRAN.R-project.org/package=htmltools>

¹¹ Ramnath Vaidyanathan, Yihui Xie, JJ Allaire, Joe Cheng, Carson Sievert and Kenton Russell (2021). htmlwidgets: HTML Widgets for R. R package version 1.5.4. <https://CRAN.R-project.org/package=htmlwidgets>

¹² Stefan Milton Bache and Hadley Wickham (2020). magrittr: A Forward-Pipe Operator for R. R package version 2.0.1. <https://CRAN.R-project.org/package=magrittr>

¹³ Hadley Wickham and Jennifer Bryan (2019). readxl: Read Excel Files. R package version 1.3.1. <https://CRAN.R-project.org/package=readxl>

¹⁴ Pebesma, E., 2018. Simple Features for R: Standardized Support for Spatial Vector Data. The R Journal 10 (1), 439-446, <https://doi.org/10.32614/RJ-2018-009>

¹⁵ Winston Chang and Barbara Borges Ribeiro (2021). shinydashboard: Create Dashboards with 'Shiny'. R package version 0.7.2. <https://CRAN.R-project.org/package=shinydashboard>

¹⁶ Hadley Wickham (2019). stringr: Simple, Consistent Wrappers for Common String Operations. R package version 1.4.0. <https://CRAN.R-project.org/package=stringr>

¹⁷ Hadley Wickham (2021). tidy: Tidy Messy Data. R package version 1.1.4. <https://CRAN.R-project.org/package=tidy>

¹⁸ <https://www.openstreetmap.org>

¹⁹ <https://carto.com/>

²⁰ <https://data.lacounty.gov/dataset/Health-Districts-2012-/7yts-2b49>

²¹ <https://geohub.lacity.org/datasets/a1543cfa466b45aab01d5ee75152ccb0/explore>