ceylon: An R package for plotting the maps of Sri Lanka

26 December 2023

Summary

The rapid evolution in the fields of computer science, data science and Artificial Intelligence has significantly transformed the utilization of data for decision making. Data visualisation plays a critical role in any work that involves data. Visualizing data on maps is frequently encountered in many fields. Visualizing data on maps not only transforms raw data into visually comprehensible representations but also convert complex spatial information into simple understandable form. Locating data files necessary for map creation can be a challenging task. Establishing a centralized repository can alleviate the challenging task of finding shape files, providing users to efficiently discover and access geographic data. The ceylon R package is designed to make simple features data related to Sri Lanka's administrative boundaries and rivers and streams accessible for a diverse range of R (R Core Team 2023) users. With straightforward functionalities, this package allows users to quickly plot and explore administrative boundaries and rivers and streams in Sri Lanka.

Statement of Need

The `ceylon` R package conveniently packages shape files corresponding to the geographic features of Sri Lanka, enhancing user-friendliness for seamless integration and analysis. This allows minimizing the time spent on data searching and cleaning efforts and Hence, the package ceylon stands out as a catalyst for research efficiency.Furthermore, the package support for research reproducibility allowing others to independently verify and build upon the work that utilize the data in this package.

Datasets available in the package

dataset	description	data source
sf_sl_0	country boundary	https://data.humdata.org/
province	province boundaries	https://data.humdata.org/
district	district boundaries	https://data.humdata.org/
sf_sl_3	divisional secretariat boundaries	https://data.humdata.org/
rivers	Sri Lanka rivers and streams shapefiles	https://data.humdata.org/

Usage

`ceylon` is available on GitHub, and can be installed and loaded into the R session using:

```
install.packages("devtools")
devtools::install_github("thiyangt/ceylon")
library(ceylon)
```

Additional packages require for plotting are as follows:

```
library(ggplot2)
library(sp)
library(sf)
library(viridis)
```

The package ggplot2(Wickham 2016) is used for data visualization. The sp(E. J. Pebesma and Bivand 2005) provides tools for handling spatial data. The sf simple features(E. Pebesma 2018) builds upon the strengths of the sp package, introducing efficient approach to handling spatial data.The viridis (Garnier et al. 2023) package provides a collection of color palettes that are color blind friendly.

```
data(sf_sl_0)
a <- ggplot(sf_sl_0) + geom_sf() + ggtitle ("a: Country")
data(province)
b <- ggplot(province) + geom_sf() + ggtitle("b: Province")
data(district)
c <- ggplot(district) + geom_sf() + ggtitle("c: District")
data(sf_sl_3)
d <- ggplot(sf_sl_3) + geom_sf() + ggtitle("d: Divisional Secretariat")
```

```
library(patchwork)
(a|b|c|d)
```

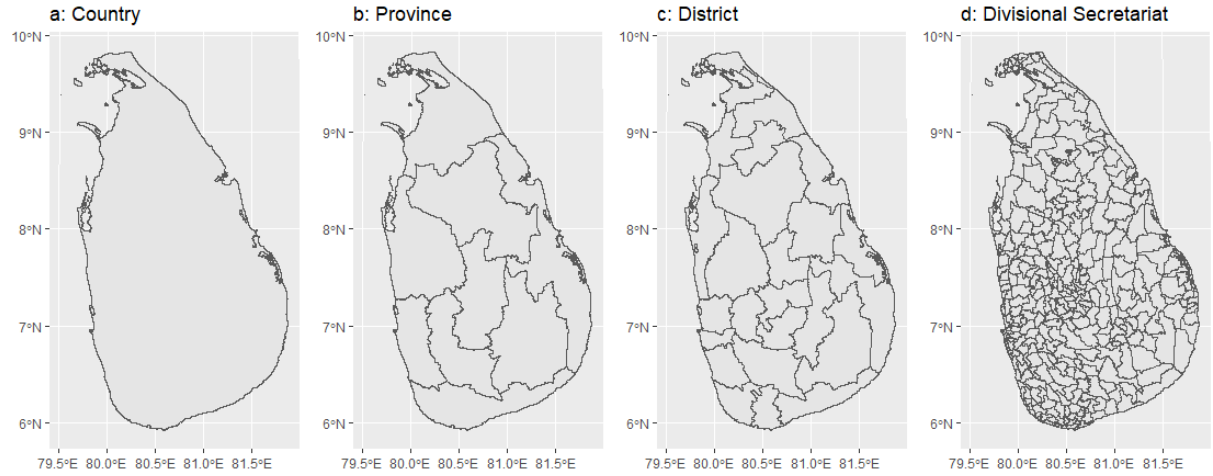

Figure 1: Maps of differnt administrative divisions in Sri Lanka

Point Map: Adding a point to the map

The Global Positioning System (GPS) coordinates of Bandaranaike International Airport, Sri Lanka is Latitude: 7.1753 Longitude: 79.8835. The goal is to plot this point along with the province boundaries. The EPSG:4326 geographic CRS system gives latitude and longitude coordinates to specify a location on the surface of the earth. Hence, before plotting first longitude and latitudes should be converted into sf object to the same coordinate reference system as the province data set. For ths we use sp(E. J. Pebesma and Bivand 2005) and sf (E. Pebesma 2018) packages in R. In the following code `st_as_sf` specify the current coordinate reference system for longitude and latitude. The function `st_transform` convert the current CRS to target CRS. The target CRS is the CRS associated with province which is defined as `crs = st_crs(province)` inside the `st_transform` function.

```
airport <- data.frame(lng = 79.8835, lat = 7.1753)
airport.new <- airport |>
  st_as_sf(coords = c("lng", "lat"), crs = 4326) |>
  st_transform(crs = st_crs(province))
```

```
point <- ggplot(province) +
  geom_sf() + geom_sf(data = airport.new, size = 2, col = "darkred") +
  ggtitle("a: Point")
```

Line Map: Plot rivers and streams in Sri Lanka

```
data("rivers")
line <- ggplot(data = sf_sl_0) +
  geom_sf(fill="#edf8b1", color="#AAAAAA") +
  geom_sf(data=rivers, colour="#253494") +
  labs(title =  "b: Line")
```

Polygon Map: Creating a choropleth map

A choropleth map shows different regions coloured according to the numerical values associated with each individual region.

```
polygon <- ggplot(province) +
  geom_sf(mapping = aes(fill = population)) + scale_fill_viridis() +
  ggtitle("c: Polygon")
```

```
(point|line|polygon)
```

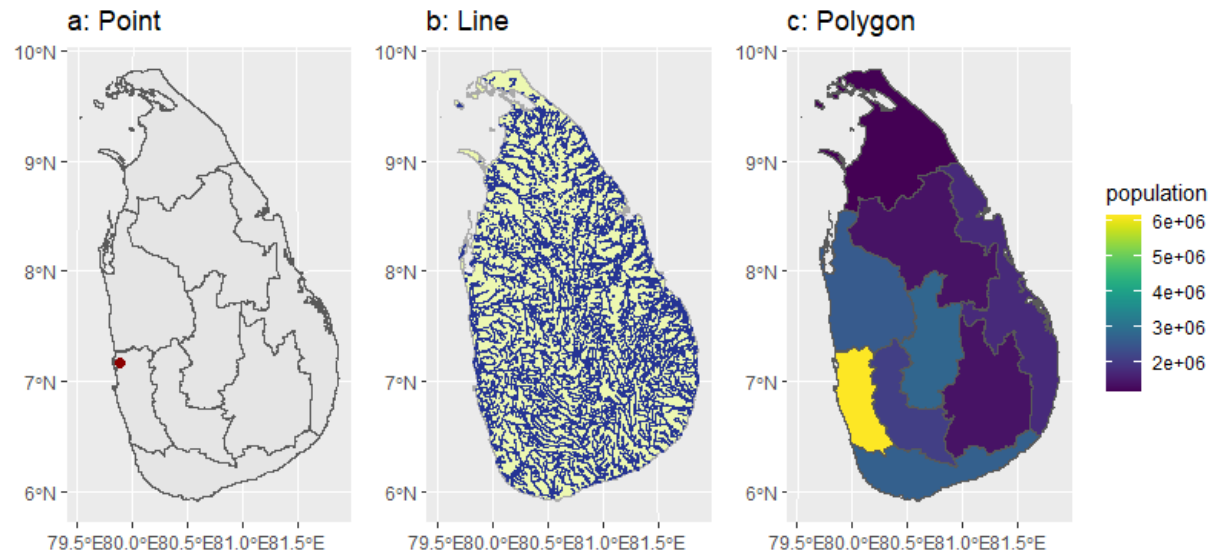

Figure 2: Illustration of point, line and polygon map

Acknowledgements

We extend our sincere gratitude to Stephanie Kobakian for her enlightening talk at the R-Ladies Colombo meetup. This package owes its inspiration to the valuable insights shared during Stephanie's presentation

References

Garnier, Simon, Ross, Noam, Rudis, Robert, Camargo, et al. 2023. "viridis(Lite) - Colorblind-Friendly Color Maps for r." https://doi.org/10.5281/zenodo.4679423.

Pebesma, Edzer. 2018. "Simple Features for R: Standardized Support for Spatial Vector Data." *The R Journal* 10 (1): 439–46. https://doi.org/10.32614/RJ-2018-009.

Pebesma, Edzer J., and Roger Bivand. 2005. "Classes and Methods for Spatial Data in R." *R News* 5 (2): 9–13. https://CRAN.R-project.org/doc/Rnews/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org.