

R riverplot package

By

Luigi Penalosa

Riverplot creates basic Sankey diagrams (Ribbon/River Plots)

Used to illustrate flows of information

Separates and joins like a river splits and merges

i.e. Comparing clusterings

This is done by:

- Creating a list of nodes
- Each item in the list, is a list of partner nodes
- Need to know at which position from left to right each node resides

Installing package:

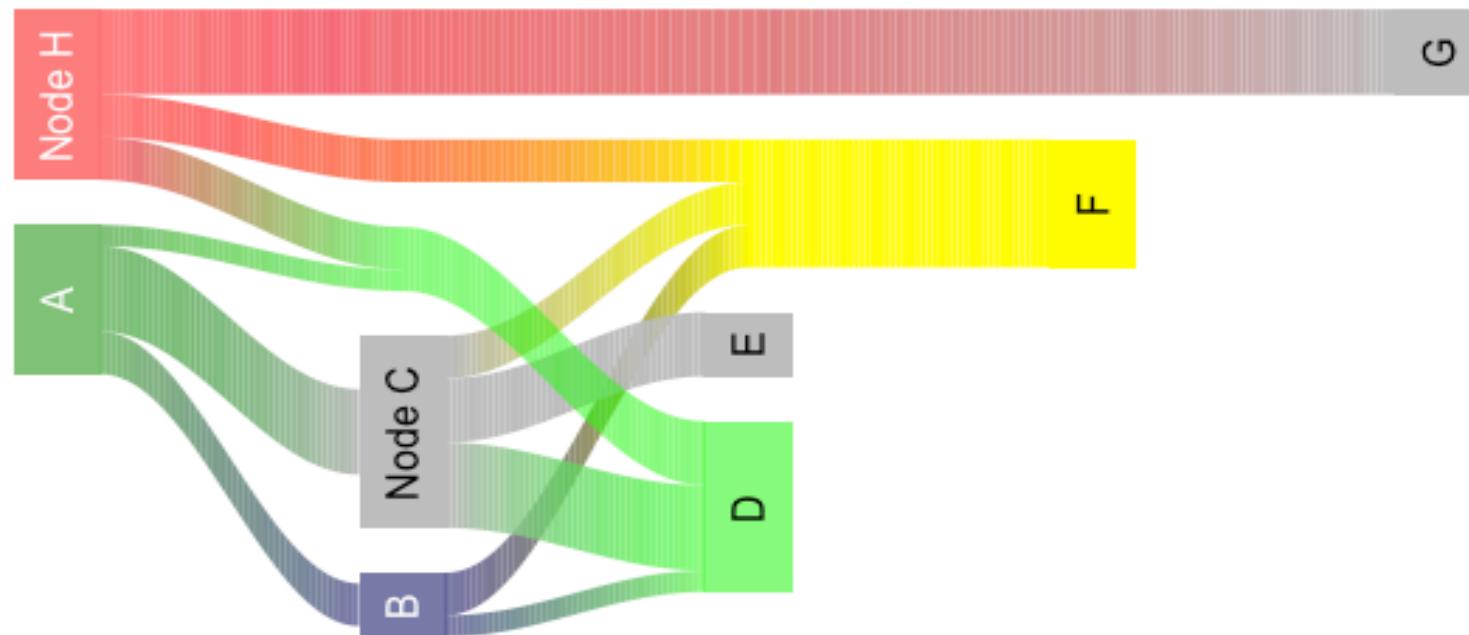
```
install.packages("riverplot")
library(riverplot)
```

Plotting: Once riverplot object is created, functions below can be used.

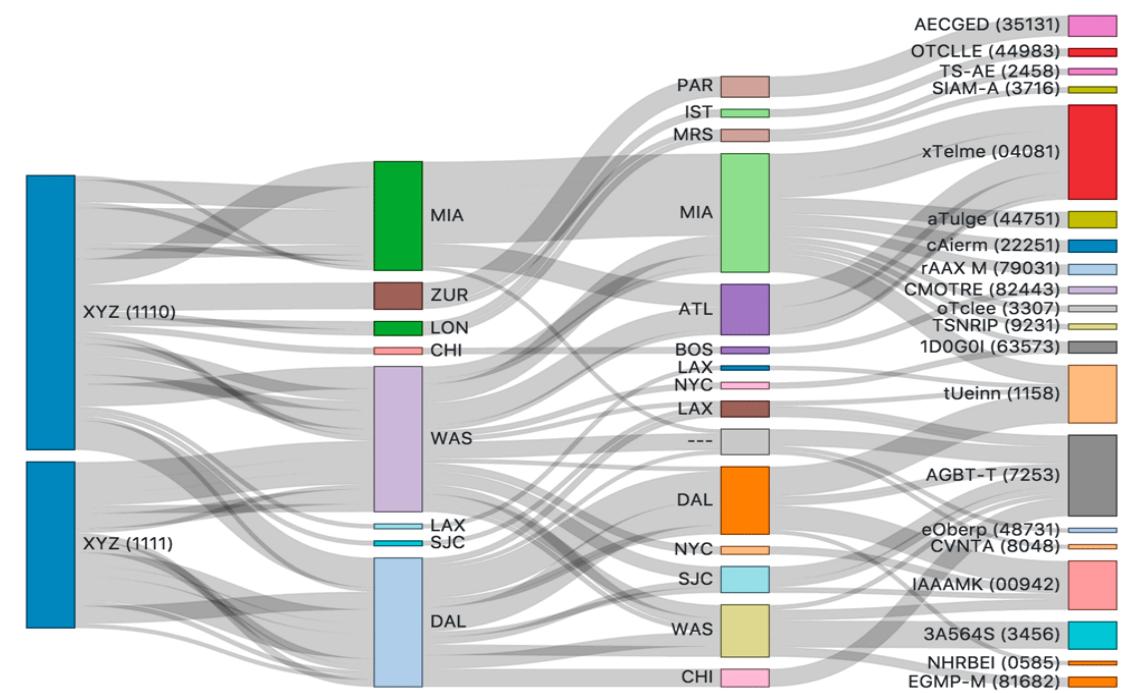
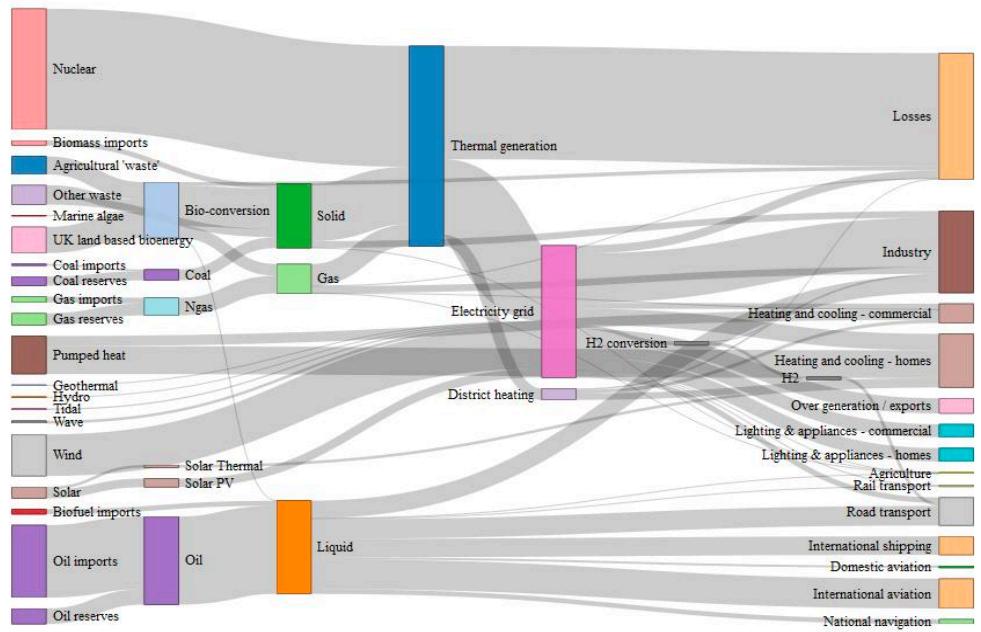
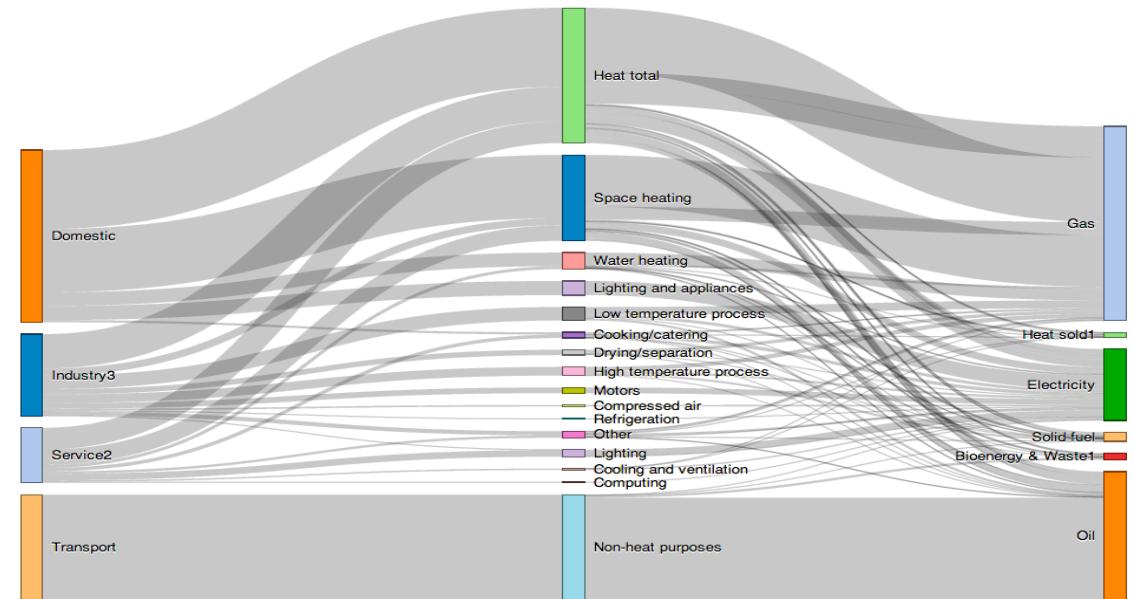
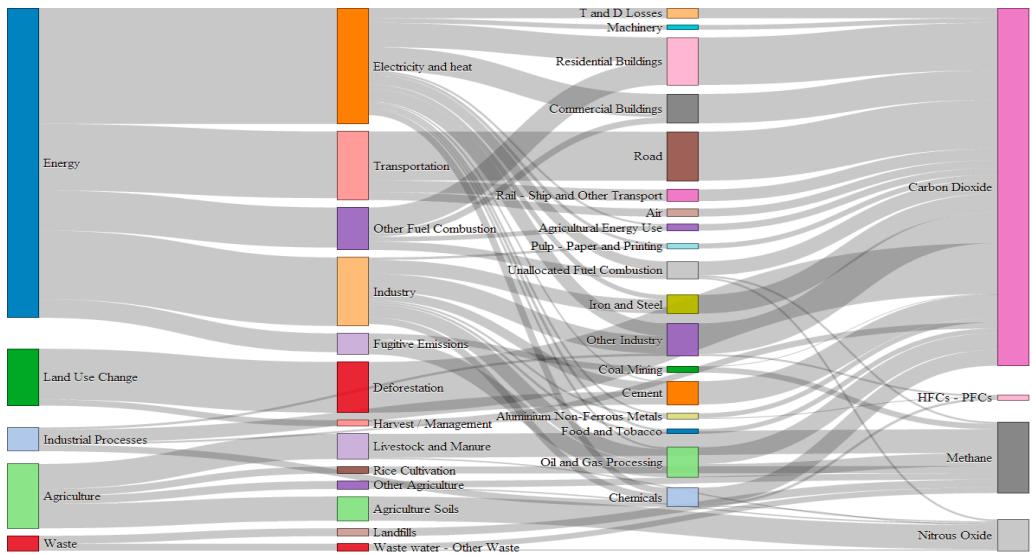
```
plot(x)
riverplot(x)
```

Sankey Diagram

```
riverplot.example()  
plot(riverplot.example())
```



Examples



`makeRiver()`

Functions to create a new object of the `riverplot` class from the provided data.

`makeRiver` creates a plot from an object which specifies the graph directly.

Nodes can be specified as a character vector (simply listing the nodes) or as a data frame.

- character vector: in this case, you also need to provide the `node_xpos` argument to specify the horizontal positions of the nodes.
- data frame: the data frame must have at least a column called "ID"; the horizontal position can be specified either with `node_xpos` argument or by column "x" in the data frame. Optionally, the data frame can include columns "labels" and "y" (vertical positions of the node). Any NA values are ignored (not entered into the riverplot project). Additionally, the data frame may contain style information.

```
makeRiver(nodes, edges, node_labels = NULL, node_xpos = NULL,  
node_ypos = NULL, node_styles = NULL, edge_styles = NULL,  
default_style = NULL)
```

Structure:

nodes- Data frame specifying nodes. Contains ID and x (Horizontal position of node). Columns “labels” and y(vertical position of node) are optional.

edges- Data frame specifying edges, and graph topology. Contains ID of the edge “ID”, parent node “N1”, child node “N2” and size of edge “Value”.

styles- Names of this list are the node or edge IDs. Values are styles specifying the style of the given node or edge.

Arguments

nodes	Data frame with node ID's, positions and optionally other information
edges	A named list or a data frame specifying the edges between the nodes.
node_labels	A named character vector of labels for the nodes
node_xpos	A named vector of numeric values specifying the horizontal positions on the plot.
node_ypos	A named vector of numeric values specifying the vertical positions on the plot.
node_styles	A named list specifying the styles for the nodes
edge_styles	A named list specifying the styles for the edges
default_style	list containing style information which is applied to every node and every edge

Example Code: Recreation of Famous figure by Charles Minard

```
data( minard )
nodes <- minard$nodes
edges <- minard$edges
colnames( nodes ) <- c( "ID", "x", "y" )
colnames( edges ) <- c( "N1", "N2", "Value", "direction" )

# color the edges by troop movement direction
edges$col <- c( "#e5cbaar", "black" )[ factor( edges$direction ) ]
# color edges by their color rather than by gradient between the nodes
edges$edgecol <- "col"

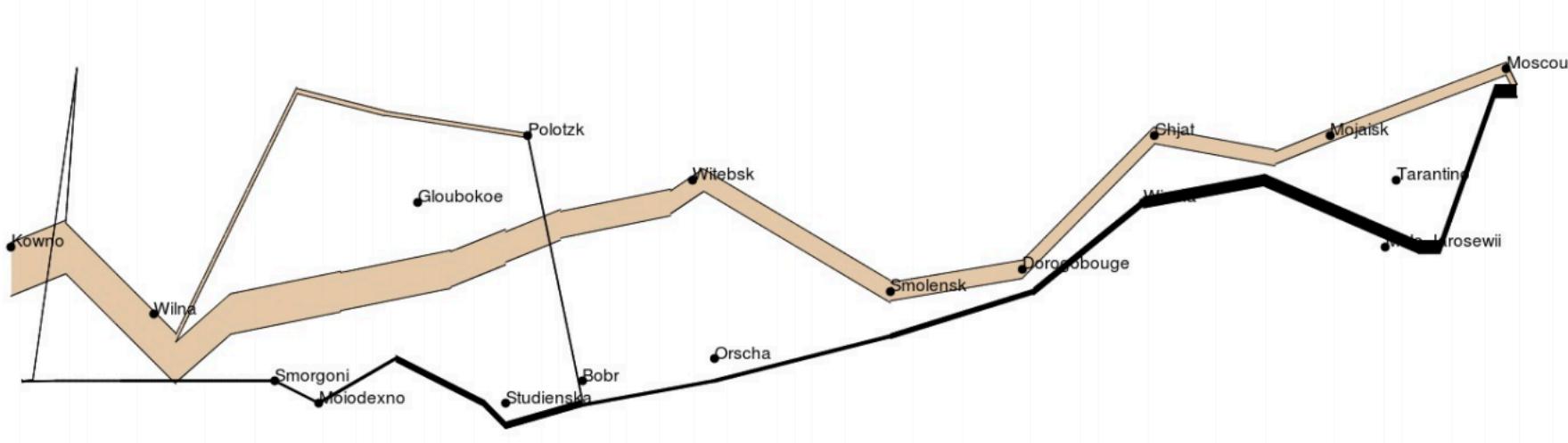
# generate the riverplot object and a style
river <- makeRiver( nodes, edges )
style <- list( edgestyle= "straight", nodestyle= "invisible" )

# plot the generated object
plot( river, lty= 1, default_style= style )
# Add cities
with( minard$cities, points( Longitude, Latitude, pch= 19 ) )
with( minard$cities, text( Longitude, Latitude, Name, adj= c( 0, 0 ) ) )
```

Recreation of Famous figure using Minard Napoleon Russian Campaign Data

nodes data frame with geographic locations of the Napoleon army (longitude and latitude) and the direction of the march

edges connections between positions



Thank You.



LUIGI PENALOZA

Sources:

- <https://cran.r-project.org>
- <https://stackoverflow.com>
- <https://github.com>