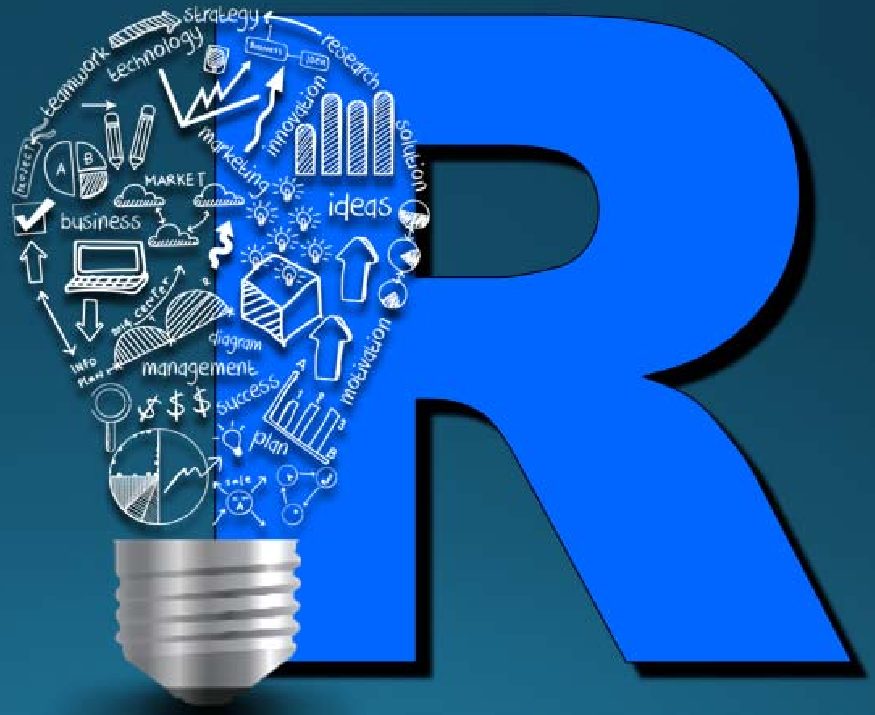
E07

TreeMap & networkD3

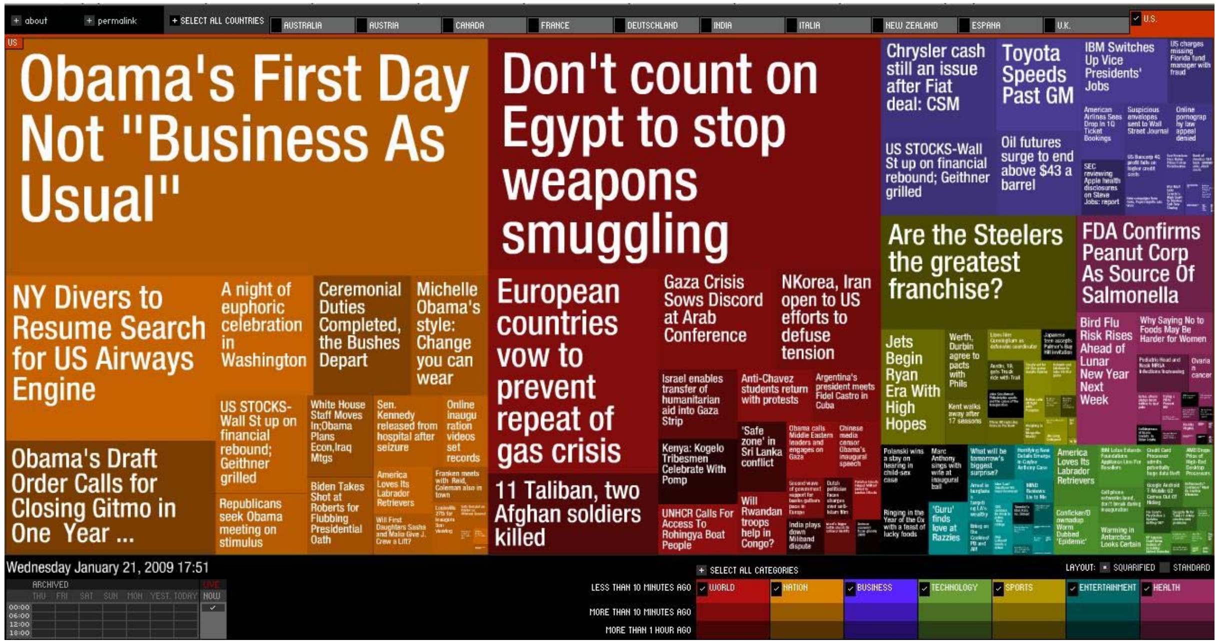


吳漢銘

國立政治大學 統計學系

https://hmwu.idv.tw

以圖為基礎的文字關係視覺化: 新聞地圖 (NewsMap)



+ about + permalink + SELECT ALL COUNTRIES AUSTRALIA AUSTRIA CANADA FRANCE DEUTSCHLAND INDIA ITALIA NEW ZEALAND ESPANA U.K. ✓ U.S.

US

Obama's First Day Not "Business As Usual"

Don't count on Egypt to stop weapons smuggling

Chrysler cash still an issue after Fiat deal: CSM

Toyota Speeds Past GM

IBM Switches Up Vice Presidents' Jobs

US charges missing Florida fund manager with fraud

American Airlines Sees Drop in 1Q Ticket Bookings

Suspicious envelopes sent to Wall Street Journal

Online pornography law appeal denied

US STOCKS-Wall St up on financial rebound; Geithner grilled

Oil futures surge to end above $43 a barrel

SEC reviewing Apple health disclosures on Steve Jobs: report

US Bancorp 4Q profit falls on higher credit costs

Are the Steelers the greatest franchise?

FDA Confirms Peanut Corp As Source Of Salmonella

NY Divers to Resume Search for US Airways Engine

A night of euphoric celebration in Washington

Ceremonial Duties Completed, the Bushes Depart

Michelle Obama's style: Change you can wear

European countries vow to prevent repeat of gas crisis

Gaza Crisis Sows Discord at Arab Conference

NKorea, Iran open to US efforts to defuse tension

Jets Begin Ryan Era With High Hopes

Werth, Durbin agree to pacts with Phils

Kent walks away after 17 seasons

Bird Flu Risk Rises Ahead of Lunar New Year Next Week

Why Saying No to Foods May Be Harder for Women

Pediatric Head and Neck MRSA Infections Increasing

Ovarian cancer

US STOCKS-Wall St up on financial rebound; Geithner grilled

White House Staff Moves In;Obama Plans Econ,Iraq Mtgs

Sen. Kennedy released from hospital after seizure

Online inauguration videos set records

Israel enables transfer of humanitarian aid into Gaza Strip

Anti-Chavez students return with protests

Argentina's president meets Fidel Castro in Cuba

Obama's Draft Order Calls for Closing Gitmo in One Year ...

Republicans seek Obama meeting on stimulus

Biden Takes Shot at Roberts for Flubbing Presidential Oath

America Loves Its Labrador Retrievers

Franken meets with Reid, Coleman also in town

Will First Daughters Sasha and Malia Give J. Crew a Lift?

11 Taliban, two Afghan soldiers killed

Kenya: Kogelo Tribesmen Celebrate With Pomp

'Safe zone' in Sri Lanka conflict

Obama calls Middle Eastern leaders and engages on Gaza

Chinese media censor Obama's inaugural speech

UNHCR Calls For Access To Rohingya Boat People

Will Rwandan troops help in Congo?

India plays down Miliband dispute

Polanski wins a stay on hearing in child-sex case

Marc Anthony sings with wife at inaugural ball

Ringing in the Year of the Ox with a feast of lucky foods

'Guru' finds love at Razzies

What will be tomorrow's biggest surprise?

America Loves Its Labrador Retrievers

Conficker/D ownadup Worm Dubbed 'Epidemic'

Warming in Antarctica Looks Certain

Credit Card Processor admits potentially huge data theft

Google Android T-Mobile G2 Comes Out Of Hiding

Wednesday January 21, 2009 17:51

ARCHIVED LIVE

THU FRI SAT SUN MON YEST. TODAY NOW

00:00 06:00 12:00 18:00 ✓

+ SELECT ALL CATEGORIES

LAYOUT: ■ SQUARIFIED □ STANDARD

LESS THAN 10 MINUTES AGO ✓ WORLD ✓ NATION ✓ BUSINESS ✓ TECHNOLOGY ✓ SPORTS ✓ ENTERTAINMENT ✓ HEALTH

MORE THAN 10 MINUTES AGO

MORE THAN 1 HOUR AGO

The rectangles in a treemap are characterized by two aesthetics: **size and color**. The sizes are derived from the proportions of the main variable. The colors can be used in several ways.

Maramushi's Newsmap takes news feeds from Google News and creates a visualisation based on the **popularity of stories**. It updates on the fly as news breaks with the tiles linking to the unabridged stories. Great for a snap-shot of what is going on in the world, even better for those far from home who want a quick over view of what is happening…about as much as ever it would appear.

- TreeMap 也是一種經典的視覺化關係佈局。
- NewsMap 就是基於 TreeMap 展示新聞，顏色用於區分新聞型別。

http://swei-industries.tumblr.com/post/387546517/maramushi-newsmap

treemap: Treemap Visualization

https://cran.r-project.org/web/packages/treemap/

Usage

```
treemap(dtf, index, vSize, vColor = NULL, stdErr = NULL, type = "index",
  fun.aggregate = "sum", title = NA, title.legend = NA,
  algorithm = "pivotSize", sortID = "-size", mirror.x = FALSE,
  mirror.y = FALSE, palette = NA, palette.HCL.options = NULL,
  range = NA, mapping = NA, n = 7, fontsize.title = 14,
  fontsize.labels = 11, fontsize.legend = 12, fontcolor.labels = NULL,
  fontface.labels = c("bold", rep("plain", length(index) - 1)),
  fontfamily.title = "sans", fontfamily.labels = "sans",
  fontfamily.legend = "sans", border.col = "black",
  border.lwds = c(length(index) + 1, (length(index) - 1):1),
  lowerbound.cex.labels = 0.4, inflate.labels = FALSE, bg.labels = NULL,
  force.print.labels = FALSE, overlap.labels = 0.5,
  align.labels = c("center", "center"), xmod.labels = 0, ymod.labels = 0,
  eval.labels = FALSE, position.legend = NULL, format.legend = NULL,
  drop.unused.levels = TRUE, aspRatio = NA, vp = NULL, draw = TRUE, ...)
```

- `dtf`: a data.frame.
- `index`: vector of column names in dtf that specify the aggregation indices. It could contain only one column name, which results in a treemap without hierarchy. If multiple column names are provided, the first name is the highest aggregation level, the second name the second-highest aggregation level, and so on.
- `vSize`: name of the column in dtf that specifies the sizes of the rectangles.
- `type`: type of the treemap, which determines how the rectangles are colored: "index", "value", "comp", "dens", "depth", "categorical", "color", "manual".

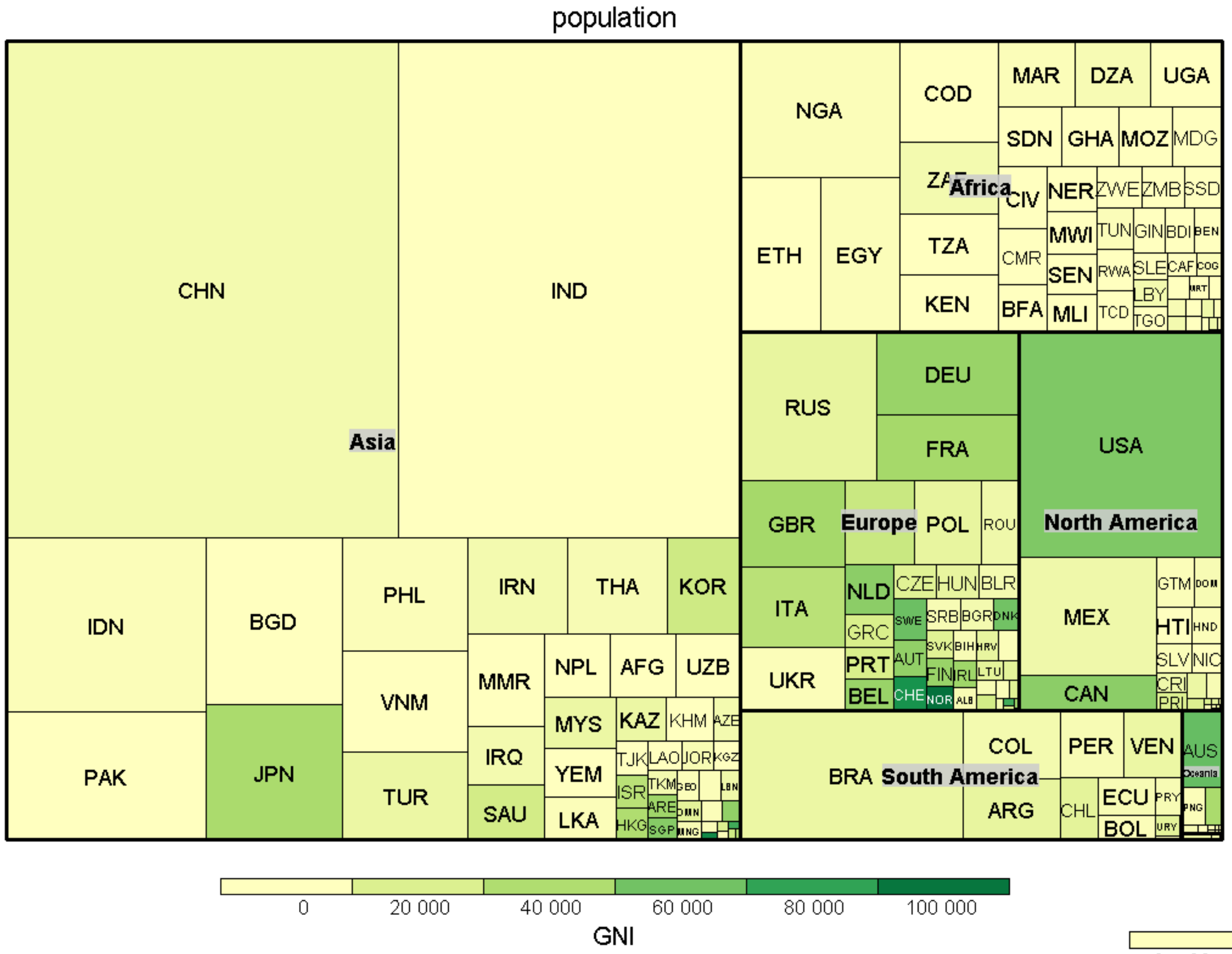
GNI2014 {treemap}

GNI 2014 Data: Gross national income (本地居民總收入) (per capita, 人均) in dollars and population totals per country in 2014.

```
> install.packages("treemap")
> library(treemap)
> data(GNI2014)
> ?GNI2014
> head(GNI2014)
  iso3           country     continent population    GNI
3  BMU           Bermuda North America      67837 106140
4  NOR            Norway        Europe    4676305 103630
5  QAT             Qatar          Asia     833285  92200
6  CHE       Switzerland        Europe    7604467  88120
7  MAC Macao SAR, China          Asia     559846  76270
8  LUX        Luxembourg        Europe     491775  75990
> str(GNI2014)
'data.frame':   188 obs. of  5 variables:
 $ iso3      : chr  "BMU" "NOR" "QAT" "CHE" ...
 $ country   : chr  "Bermuda" "Norway" "Qatar" "Switzerland" ...
 $ continent : Factor w/ 8 levels "Africa","Antarctica",..: 5 4 3 4 3 4 6 4 4 5 ...
 $ population: num  67837 4676305 833285 7604467 559846 ...
 $ GNI       : int  106140 103630 92200 88120 76270 75990 64540 61610 61310 55200 ...
```

treemap

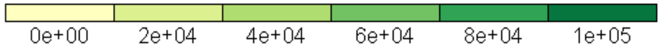
```
treemap(GNI2014,
        index=c("continent", "iso3"), vSize="population", vColor="GNI",
        type="value", format.legend = list(scientific = FALSE, big.mark = " "))
```



`type="value"`: the numeric vColor-column is directly mapped to a color palette.

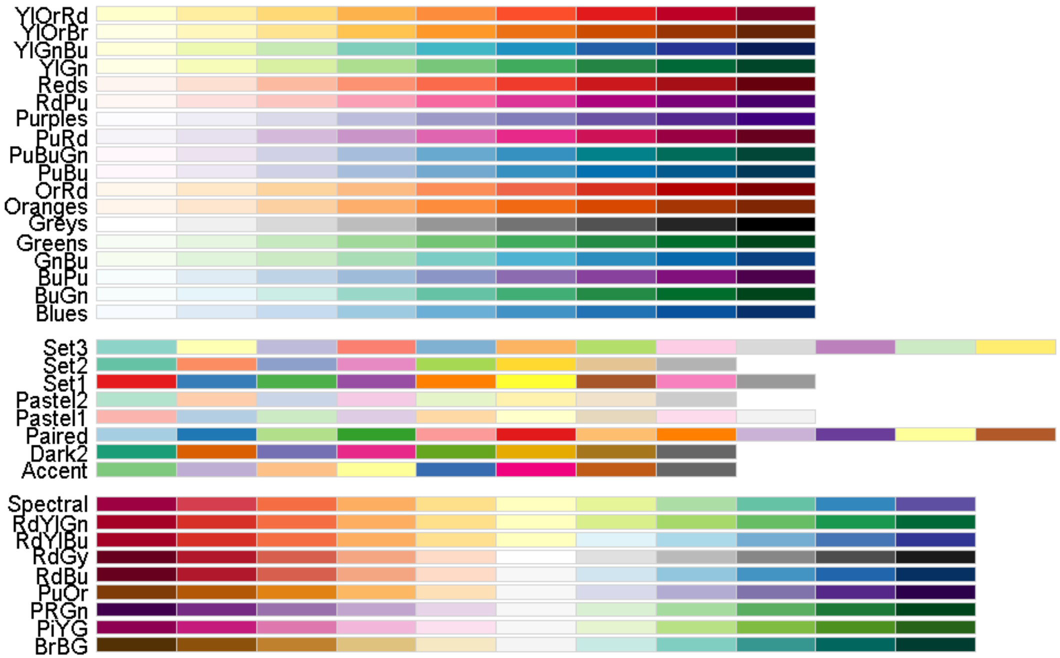
`format.legend`: a list of additional arguments for the formatting of numbers in the legend to pass to **`format()`**; only applies if type is "value", "dens" or "manual".

`scientific = TRUE`



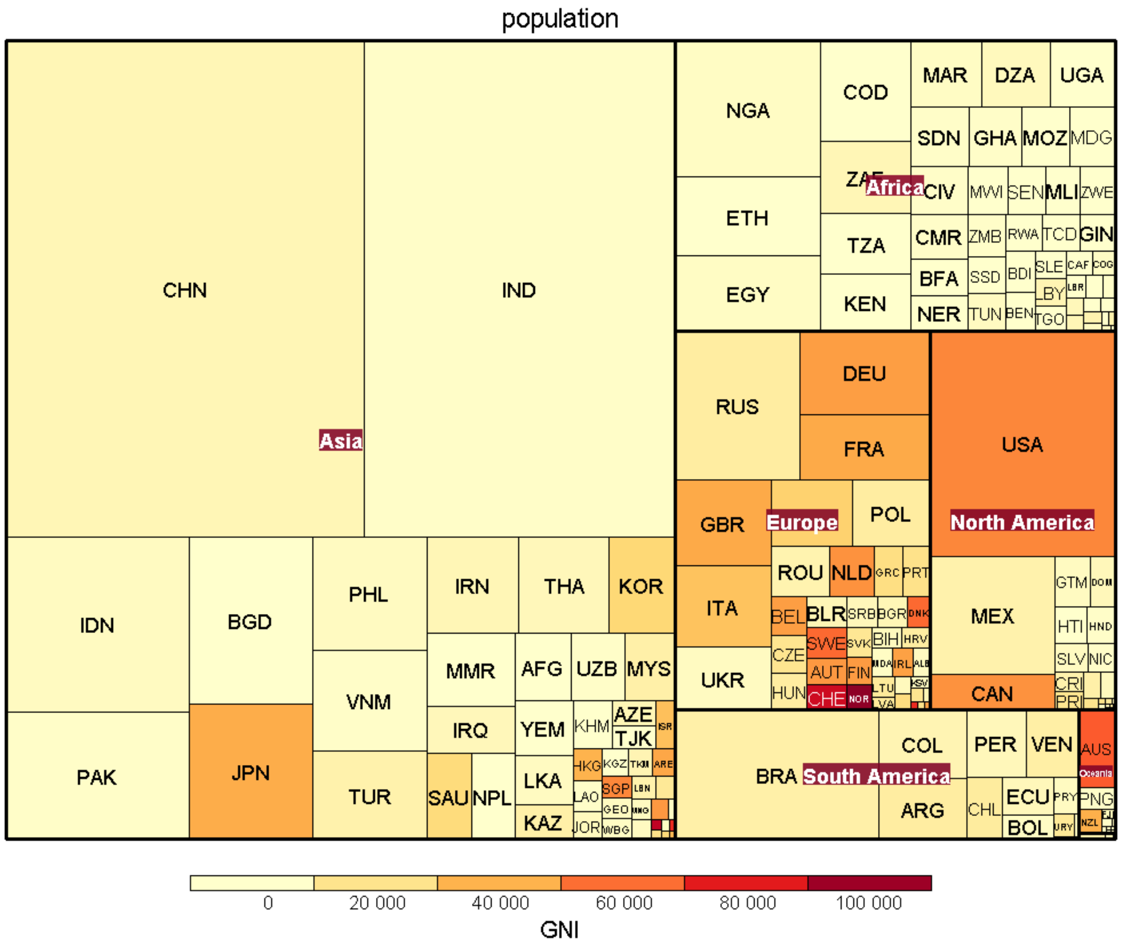
RColorBrewer::display.brewer.all()

```
> RColorBrewer::display.brewer.all()
> ?RColorBrewer::display.brewer.all
> RColorBrewer::brewer.pal.info
         maxcolors category colorblind
BrBG            11      div       TRUE
PiYG            11      div       TRUE
PRGn            11      div       TRUE
PuOr            11      div       TRUE
RdBu            11      div       TRUE
RdGy            11      div      FALSE
RdYlBu          11      div       TRUE
RdYlGn          11      div      FALSE
Spectral        11      div      FALSE
Accent           8     qual      FALSE
....
Oranges          9      seq       TRUE
OrRd             9      seq       TRUE
PuBu             9      seq       TRUE
PuBuGn           9      seq       TRUE
PuRd             9      seq       TRUE
Purples          9      seq       TRUE
RdPu             9      seq       TRUE
Reds             9      seq       TRUE
YlGn             9      seq       TRUE
YlGnBu           9      seq       TRUE
YlOrBr           9      seq       TRUE
YlOrRd           9      seq       TRUE
```



treemap

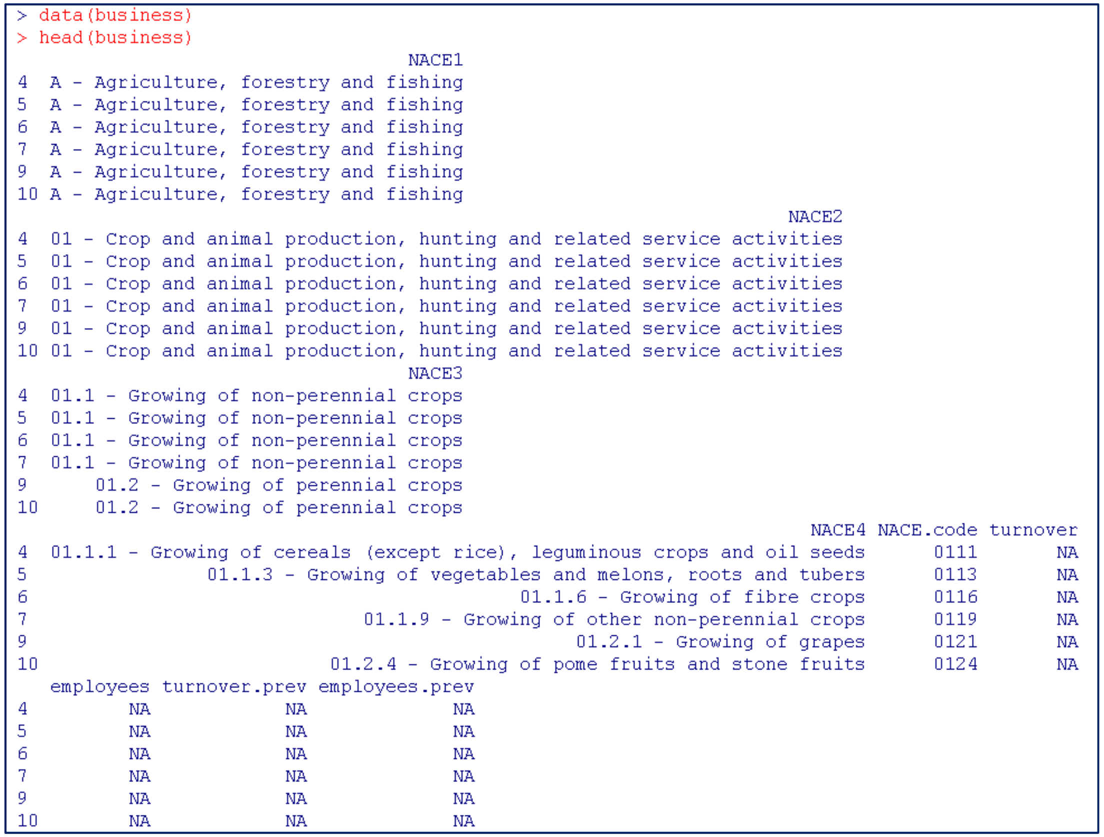
```
treemap(GNI2014,
        index=c("continent", "iso3"), vSize="population", vColor="GNI",
        type="manual", format.legend = list(scientific = FALSE, big.mark = " "),
        palette="YlOrRd")
```



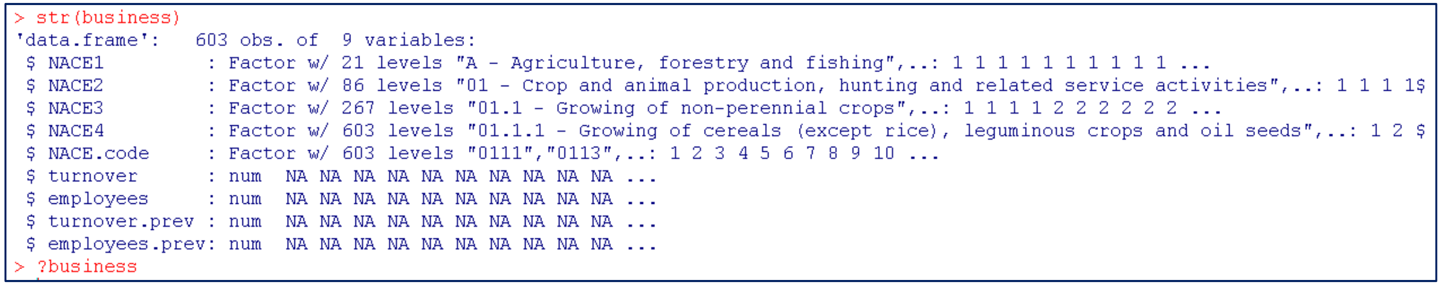
business {treemap}

Fictitious (aggregated) business statistics data. The index variables (NACE1 to NACE4) are derived from the Statistical Classification of Economic Activities in the European Community (NACE). The variables turnover(.prev) and employees(.prev) have values for NACE codes in the business economy domain only.

```
> data(business)
> head(business)
> str(business)
```



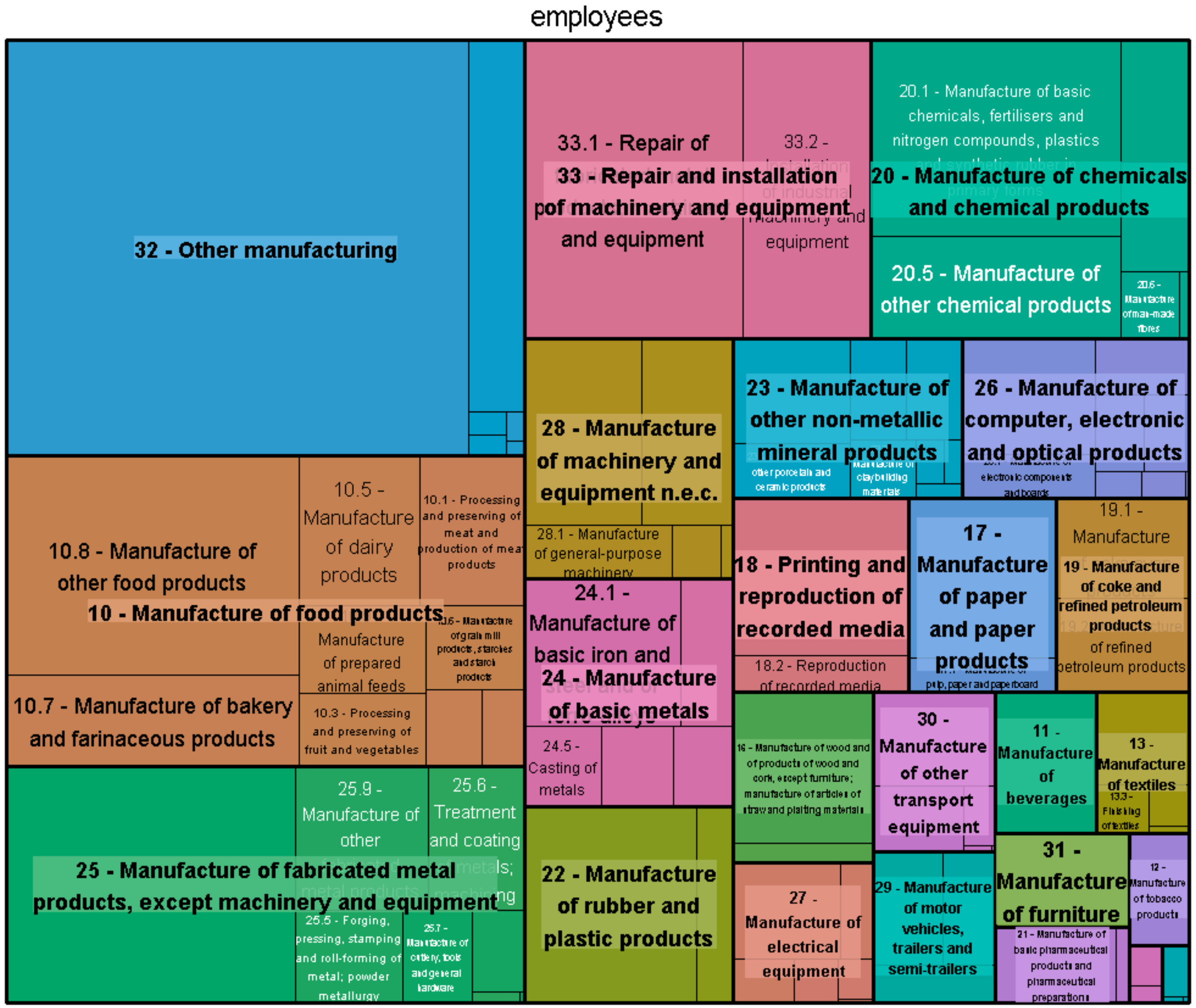
```
> data(business)
> head(business)
                                   NACE1
4  A - Agriculture, forestry and fishing
5  A - Agriculture, forestry and fishing
6  A - Agriculture, forestry and fishing
7  A - Agriculture, forestry and fishing
9  A - Agriculture, forestry and fishing
10 A - Agriculture, forestry and fishing
                                                                     NACE2
4  01 - Crop and animal production, hunting and related service activities
5  01 - Crop and animal production, hunting and related service activities
6  01 - Crop and animal production, hunting and related service activities
7  01 - Crop and animal production, hunting and related service activities
9  01 - Crop and animal production, hunting and related service activities
10 01 - Crop and animal production, hunting and related service activities
                                   NACE3
4  01.1 - Growing of non-perennial crops
5  01.1 - Growing of non-perennial crops
6  01.1 - Growing of non-perennial crops
7  01.1 - Growing of non-perennial crops
9      01.2 - Growing of perennial crops
10     01.2 - Growing of perennial crops
                                                                       NACE4 NACE.code turnover
4  01.1.1 - Growing of cereals (except rice), leguminous crops and oil seeds      0111       NA
5                 01.1.3 - Growing of vegetables and melons, roots and tubers      0113       NA
6                                            01.1.6 - Growing of fibre crops      0116       NA
7                              01.1.9 - Growing of other non-perennial crops      0119       NA
9                                                 01.2.1 - Growing of grapes      0121       NA
10                          01.2.4 - Growing of pome fruits and stone fruits      0124       NA
   employees turnover.prev employees.prev
4         NA            NA             NA
5         NA            NA             NA
6         NA            NA             NA
7         NA            NA             NA
9         NA            NA             NA
10        NA            NA             NA
```



```
> str(business)
'data.frame':   603 obs. of  9 variables:
 $ NACE1        : Factor w/ 21 levels "A - Agriculture, forestry and fishing",..: 1 1 1 1 1 1 1 1 1 1 ...
 $ NACE2        : Factor w/ 86 levels "01 - Crop and animal production, hunting and related service activities",..: 1 1 1 1$
 $ NACE3        : Factor w/ 267 levels "01.1 - Growing of non-perennial crops",..: 1 1 1 1 2 2 2 2 2 2 ...
 $ NACE4        : Factor w/ 603 levels "01.1.1 - Growing of cereals (except rice), leguminous crops and oil seeds",..: 1 2 $
 $ NACE.code    : Factor w/ 603 levels "0111","0113",..: 1 2 3 4 5 6 7 8 9 10 ...
 $ turnover     : num  NA NA NA NA NA NA NA NA NA NA ...
 $ employees    : num  NA NA NA NA NA NA NA NA NA NA ...
 $ turnover.prev : num  NA NA NA NA NA NA NA NA NA NA ...
 $ employees.prev: num  NA NA NA NA NA NA NA NA NA NA ...
> ?business
```

index treemap: colors are determined by the index argument

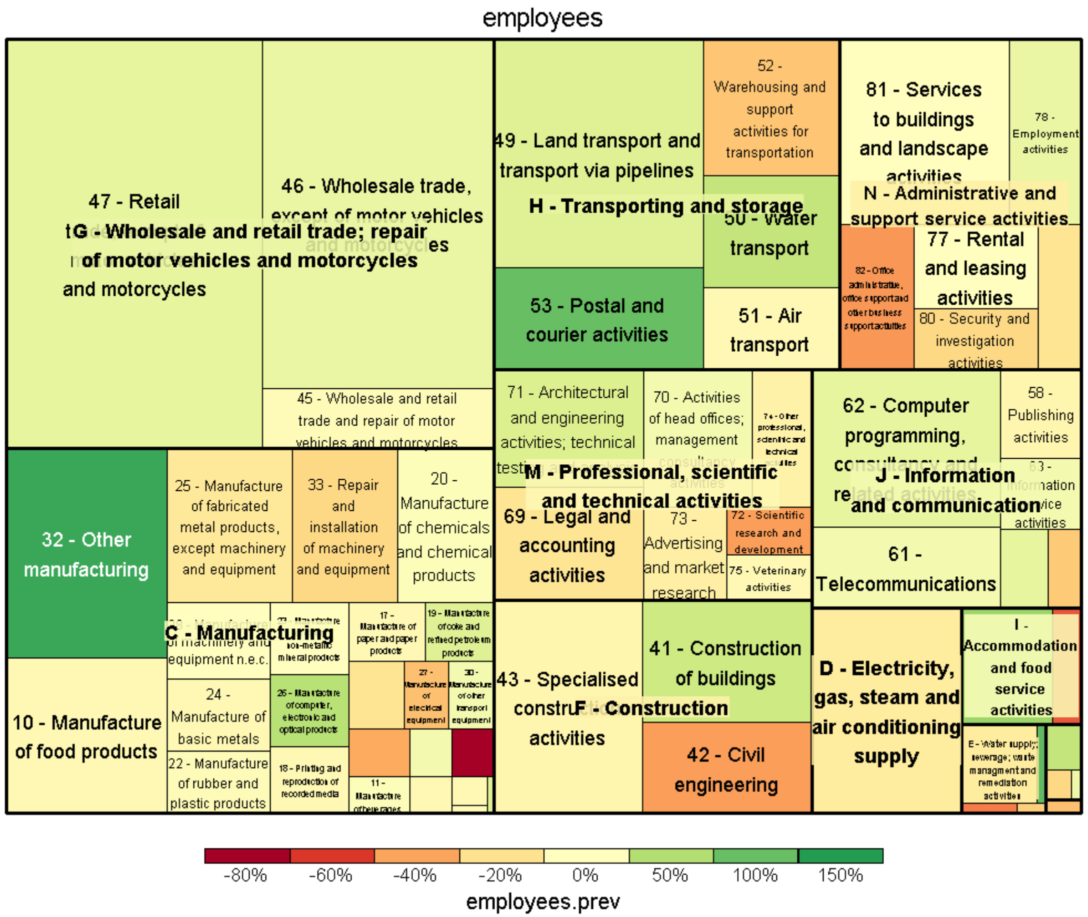
```
treemap(business[business$NACE1=="C - Manufacturing",],
        index=c("NACE2", "NACE3"), vSize=c("employees"), type="index")
```



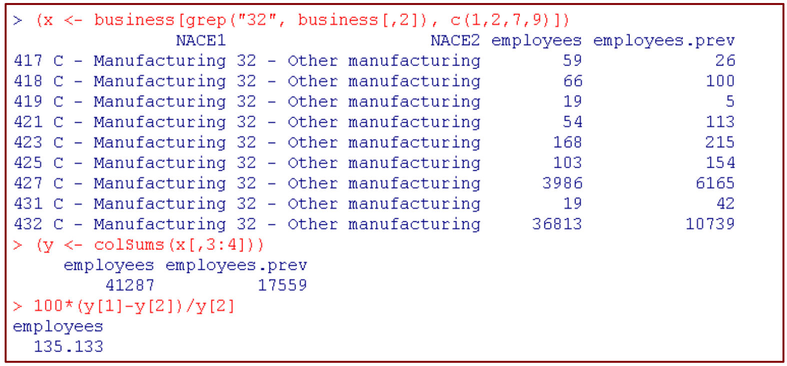
comparisson treemaps:

colors indicate change of vSize with respect to vColor

```
treemap(business,
        index=c("NACE1", "NACE2"), vSize="employees",
        vColor="employees.prev", type="comp")
```



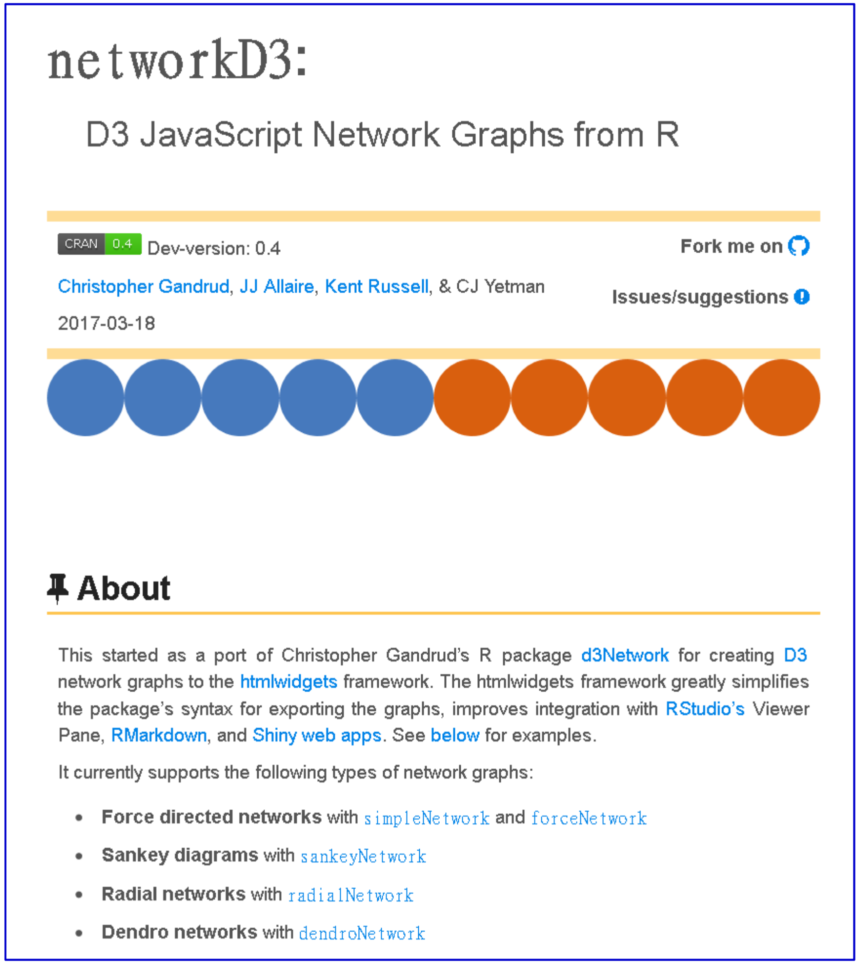
type="comp": colors indicate change of the vSize-column with respect to the numeric vColor-column in percentages. Note: the negative scale may be different from the positive scale in order to compensate for the ratio distribution.



```
> (x <- business[grep("32", business[,2]), c(1,2,7,9)])
              NACE1                     NACE2 employees employees.prev
417 C - Manufacturing 32 - Other manufacturing        59             26
418 C - Manufacturing 32 - Other manufacturing        66            100
419 C - Manufacturing 32 - Other manufacturing        19              5
421 C - Manufacturing 32 - Other manufacturing        54            113
423 C - Manufacturing 32 - Other manufacturing       168            215
425 C - Manufacturing 32 - Other manufacturing       103            154
427 C - Manufacturing 32 - Other manufacturing      3986           6165
431 C - Manufacturing 32 - Other manufacturing        19             42
432 C - Manufacturing 32 - Other manufacturing     36813          10739
> (y <- colSums(x[,3:4]))
     employees employees.prev
         41287          17559
> 100*(y[1]-y[2])/y[2]
employees
  135.133
```

networkD3: D3 JavaScript Network Graphs from R

https://christophergandrud.github.io/networkD3/



networkD3:

D3 JavaScript Network Graphs from R

CRAN 0.4 Dev-version: 0.4

Fork me on

Christopher Gandrud, JJ Allaire, Kent Russell, & CJ Yetman

Issues/suggestions

2017-03-18

About

This started as a port of Christopher Gandrud's R package d3Network for creating D3 network graphs to the htmlwidgets framework. The htmlwidgets framework greatly simplifies the package's syntax for exporting the graphs, improves integration with RStudio's Viewer Pane, RMarkdown, and Shiny web apps. See below for examples.

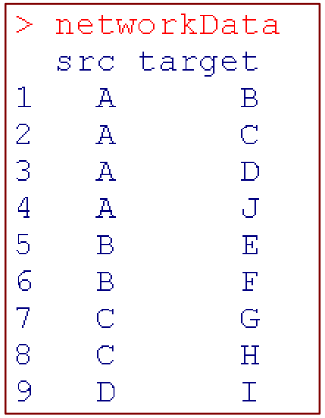
It currently supports the following types of network graphs:

- **Force directed networks** with `simpleNetwork` and `forceNetwork`
- **Sankey diagrams** with `sankeyNetwork`
- **Radial networks** with `radialNetwork`
- **Dendro networks** with `dendroNetwork`

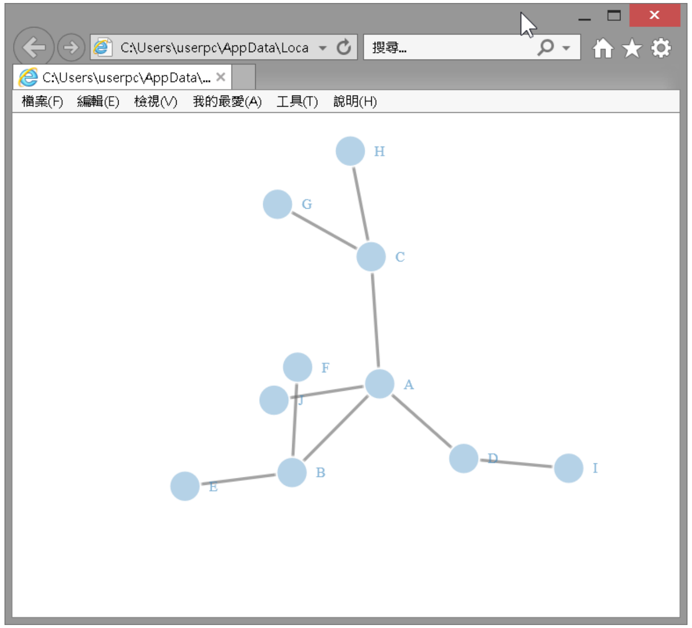
- **networkD3** supports the following types of network graphs:
  - Force directed networks with **simpleNetwork** and **forceNetwork**
  - Sankey diagrams with **sankeyNetwork**
  - Radial networks with **radialNetwork**
  - Dendro networks with **dendroNetwork**

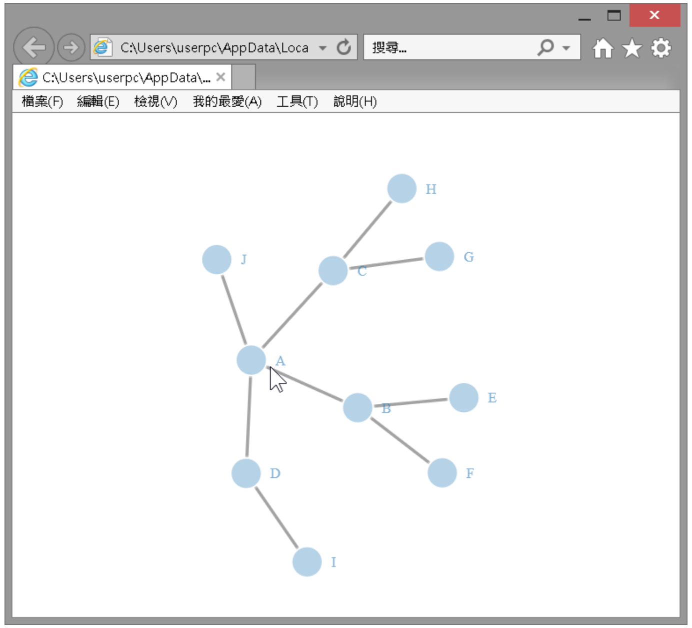
simpleNetwork

```
> src <- c("A", "A", "A", "A", "B", "B", "C", "C", "D")
> target <- c("B", "C", "D", "J", "E", "F", "G", "H", "I")
> networkData <- data.frame(src, target)
> simpleNetwork(networkData)
```



```
> networkData
  src target
1   A      B
2   A      C
3   A      D
4   A      J
5   B      E
6   B      F
7   C      G
8   C      H
9   D      I
```


C:\Users\userpc\AppData\Loca
搜尋...
C:\Users\userpc\AppData\...
檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)
H
G
C
F
J
A
D
I
E
B


C:\Users\userpc\AppData\Loca
搜尋...
C:\Users\userpc\AppData\...
檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)
H
J
C
G
A
B
E
D
F
I

https://christophergandrud.github.io/networkD3/

forceNetwork {networkD3}: Create a D3 JavaScript force directed network graph

Usage

```
forceNetwork(Links, Nodes, Source, Target, Value, NodeID, Nodesize, Group,
  height = NULL, width = NULL,
  colourScale = JS("d3.scaleOrdinal(d3.schemeCategory20);"), fontSize = 7,
  fontFamily = "serif", linkDistance = 50,
  linkWidth = JS("function(d) { return Math.sqrt(d.value); }"),
  radiusCalculation = JS(" Math.sqrt(d.nodesize)+6"), charge = -30,
  linkColour = "#666", opacity = 0.6, zoom = FALSE, legend = FALSE,
  arrows = FALSE, bounded = FALSE, opacityNoHover = 0, clickAction = NULL)
```

Arguments

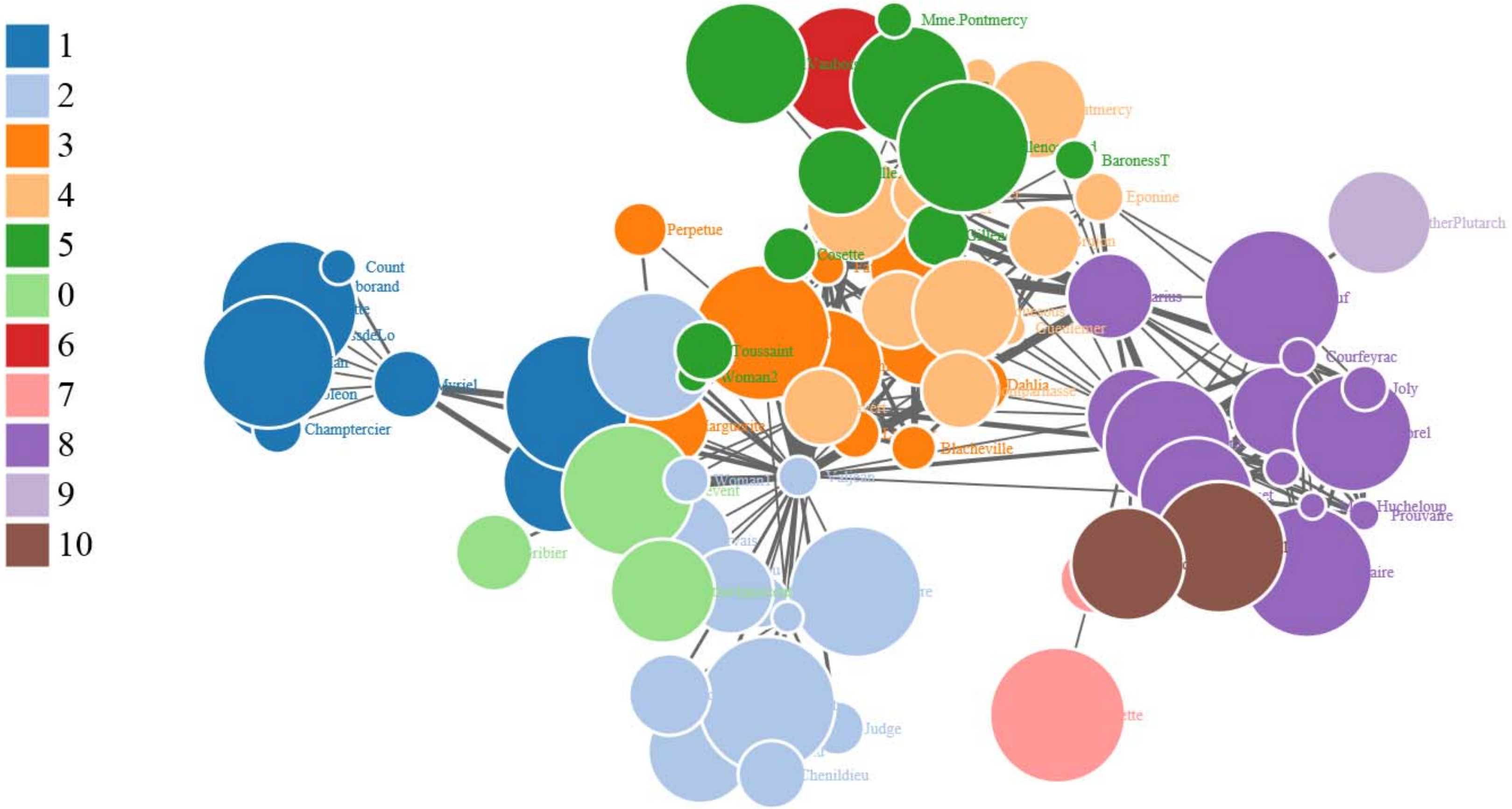
- `Links`: a data frame object with the links between the nodes. It should include the Source and Target for each link. These should be numbered starting from 0. An optional Value variable can be included to specify how close the nodes are to one another.
- `Nodes`: a data frame containing the node id and properties of the nodes. If no ID is specified then the nodes must be in the same order as the Source variable column in the Links data frame. Currently only a grouping variable is allowed.
- `Source`: character string naming the network source variable in the Links data frame.
- `Target`: character string naming the network target variable in the Links data frame.
- `Value`: character string naming the variable in the Links data frame for how wide the links are.
- `NodeID`: character string specifying the node IDs in the Nodes data frame.
- `Nodesize`: character string specifying the a column in the Nodes data frame with some value to vary the node radius's with. See also radiusCalculation.
- `Group`: character string specifying the group of each node in the Nodes data frame.

forceNetwork {networkD3}: Create a D3 JavaScript force directed network graph.

```
> data(MisLinks)
> data(MisNodes)
> head(MisLinks)
  source target value
1      1      0     1
2      2      0     8
3      3      0    10
4      3      2     6
5      4      0     1
6      5      0     1
> str(MisLinks)
'data.frame':   254 obs. of  3 variables:
 $ source: int  1 2 3 3 4 5 6 7 8 9 ...
 $ target: int  0 0 0 2 0 0 0 0 0 0 ...
 $ value : int  1 8 10 6 1 1 1 1 2 1 ...
> head(MisNodes)
             name group size
1          Myriel     1   15
2        Napoleon     1   20
3 Mlle.Baptistine     1   23
4    Mme.Magloire     1   30
5    CountessdeLo     1   11
6        Geborand     1    9
> str(MisNodes)
'data.frame':   77 obs. of  3 variables:
 $ name : Factor w/ 77 levels "Anzelma","Babet",..: 63 64 51 57 21 33 12 23 20 65 ...
 $ group: int  1 1 1 1 1 1 1 1 1 1 ...
 $ size : int  15 20 23 30 11 9 11 30 8 29 ...
> ?forceNetwork
```

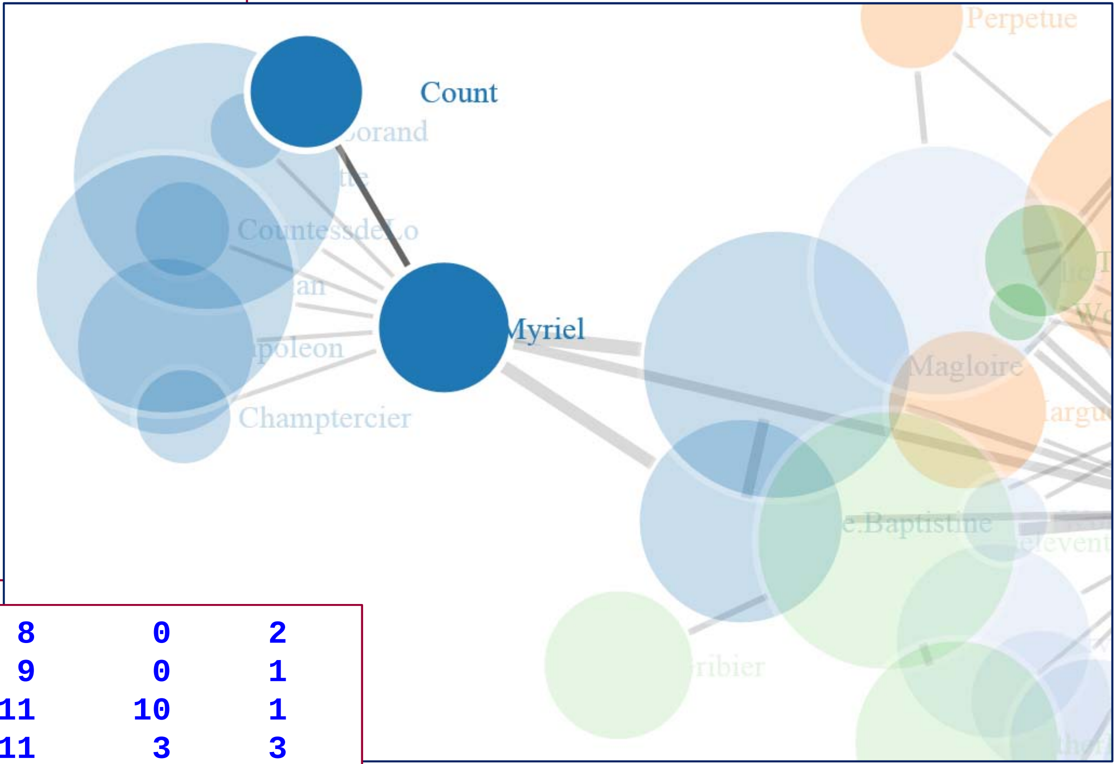
forceNetwork

```
forceNetwork(Links = MisLinks, Nodes = MisNodes, Source = "source",
             Target = "target", Value = "value", NodeID = "name",
             Nodesize = 'size', radiusCalculation = "d.nodesize",
             Group = "group", opacity = 1, legend = T, zoom = T, opacityNoHover = 1)
```



forceNetwork

```
> data.frame(id=0:(nrow(MisNodes)-1), MisNodes)
   id               name group size
1   0             Myriel     1   15
2   1           Napoleon     1   20
3   2    Mlle.Baptistine     1   23
4   3       Mme.Magloire     1   30
5   4       CountessdeLo     1   11
6   5           Geborand     1    9
7   6       Champtercier     1   11
8   7           Cravatte     1   30
9   8              Count     1    8
10  9             OldMan     1   29
11 10            Labarre     2   29
12 11            Valjean     2    9
13 12         Marguerite     3   18
...
76 75             Brujon     4   16
77 76       Mme.Hucheloup     8    6
```

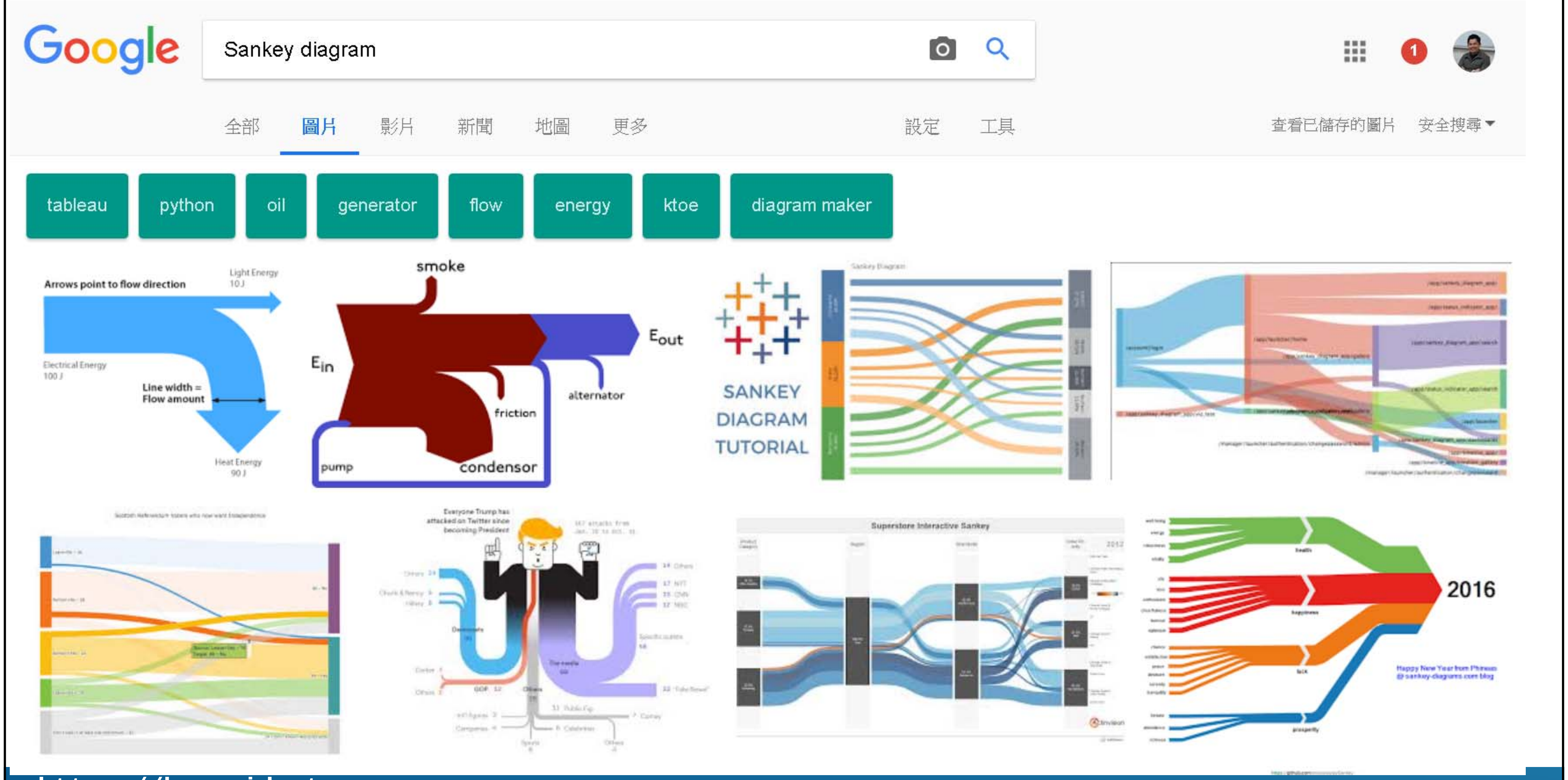


```
> MisLinks
    source target value
1        1      0     1
2        2      0     8
3        3      0    10
4        3      2     6
5        4      0     1
6        5      0     1
7        6      0     1
8        7      0     1
```

```
9          8      0     2
10         9      0     1
11        11     10     1
12        11      3     3
13        11      2     3
14        11      0     5
15        12     11     1
...
253       76     48     1
254       76     58     1
```

Sankey Diagram

http://www.sankey-diagrams.com/



Sankey diagram

sankeyNetwork {networkD3}: Create a D3 JavaScript Sankey diagram

Usage

```
sankeyNetwork(Links, Nodes, Source, Target, Value, NodeID, NodeGroup = NodeID,
  LinkGroup = NULL, units = "",
  colourScale = JS("d3.scaleOrdinal(d3.schemeCategory20);"), fontSize = 7,
  fontFamily = NULL, nodeWidth = 15, nodePadding = 10, margin = NULL,
  height = NULL, width = NULL, iterations = 32, sinksRight = TRUE)
```

Arguments

- `Links`: a data frame object with the links between the nodes. It should have include the Source and Target for each link. An optional Value variable can be included to specify how close the nodes are to one another.
- `Nodes`: a data frame containing the node id and properties of the nodes. If no ID is specified then the nodes must be in the same order as the Source variable column in the Links data frame. Currently only grouping variable is allowed.
- `Source`: character string naming the network source variable in the Links data frame.
- `Target`: character string naming the network target variable in the Links data frame.
- `Value`: character string naming the variable in the Links data frame for how far away the nodes are from one another.
- `NodeID`: character string specifying the node IDs in the Nodes. data frame. Must be 0-indexed.
- `NodeGroup`: character string specifying the node groups in the Nodes. Used to color the nodes in the network.
- `LinkGroup`: character string specifying the groups in the Links. Used to color the links in the network.

sankeyNetwork {networkD3}: Create a D3 JavaScript Sankey diagram

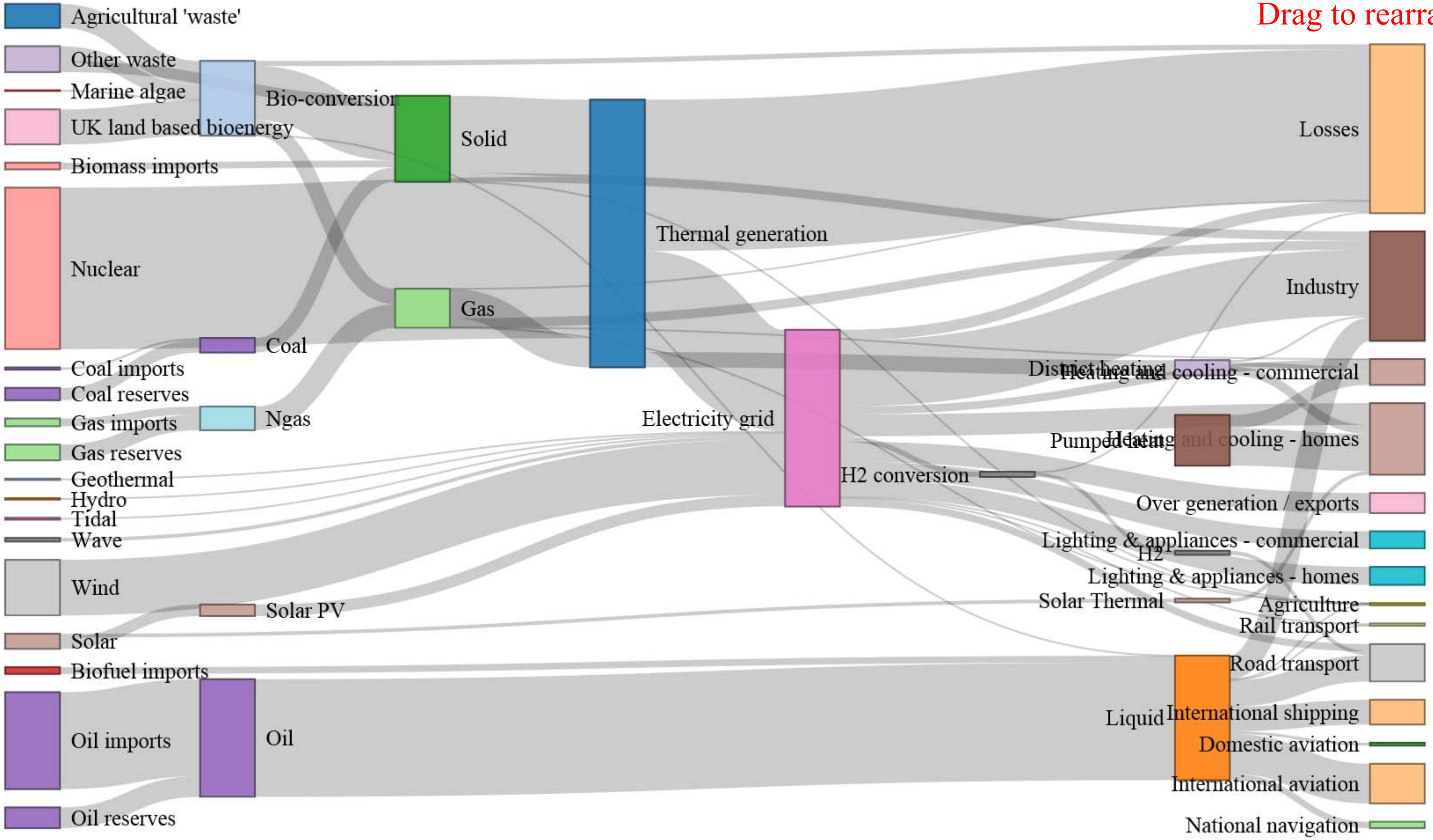
```
> URL <- paste0(
+         "https://cdn.rawgit.com/christophergandrud/networkD3/",
+         "master/JSONdata/energy.json")
> Energy <- jsonlite::fromJSON(URL)
> str(Energy)
List of 2
 $ nodes:'data.frame':  48 obs. of  1 variable:
  ..$ name: chr [1:48] "Agricultural 'waste'" "Bio-conversion" "Liquid" "Losses" ...
 $ links:'data.frame':  68 obs. of  3 variables:
  ..$ source: int [1:68] 0 1 1 1 1 6 7 8 10 9 ...
  ..$ target: int [1:68] 1 2 3 4 5 2 4 9 9 4 ...
  ..$ value : num [1:68] 124.729 0.597 26.862 280.322 81.144 ...
> lapply(Energy, head)
$nodes
                  name
1 Agricultural 'waste'
2       Bio-conversion
3               Liquid
4               Losses
5                Solid
6                  Gas

$links
  source target   value
1      0      1 124.729
2      1      2   0.597
3      1      3  26.862
4      1      4 280.322
5      1      5  81.144
6      6      2  35.000
```

```
> sankeyNetwork(Links = Energy$links, Nodes = Energy$nodes,
+               Source = "source", Target = "target",
+               Value = "value", NodeID = "name",
+               fontSize = 12, nodeWidth = 30)
```

sankeyNetwork

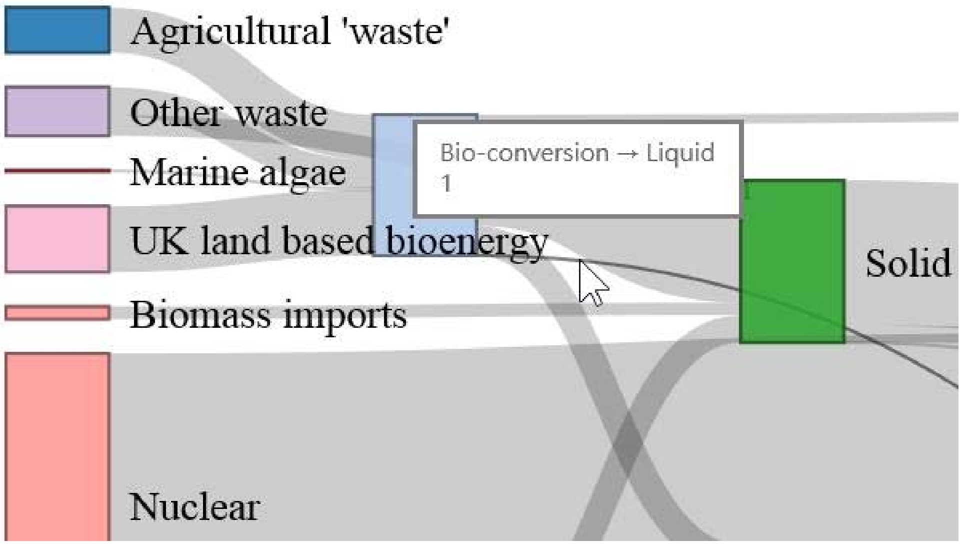
```
> sankeyNetwork(Links = Energy$links, Nodes = Energy$nodes,
+               Source = "source", Target = "target",
+               Value = "value", NodeID = "name",
+               fontSize = 12, nodeWidth = 30)
```

Drag to rearrange nodes.



sankeyNetwork

```
> Energy
$nodes
                               name
1              Agricultural 'waste'
2                    Bio-conversion
3                            Liquid
4                            Losses
5                             Solid
6                               Gas
7                   Biofuel imports
8                   Biomass imports
...
46           UK land based bioenergy
47                             Wave
48                             Wind

$links
   source target   value
1       0      1 124.729
2       1      2   0.597
3       1      3  26.862
4       1      4 280.322
5       1      5  81.144
6       6      2  35.000
7       7      4  35.000
...
67     46     15  19.013
68     47     15 289.366
```



Sankey diagrams are closely related to alluvial diagrams, which show how network structure changes over time.
https://en.wikipedia.org/wiki/Alluvial_diagram

radialNetwork {networkD3}

Create Reingold-Tilford T:ree network diagrams

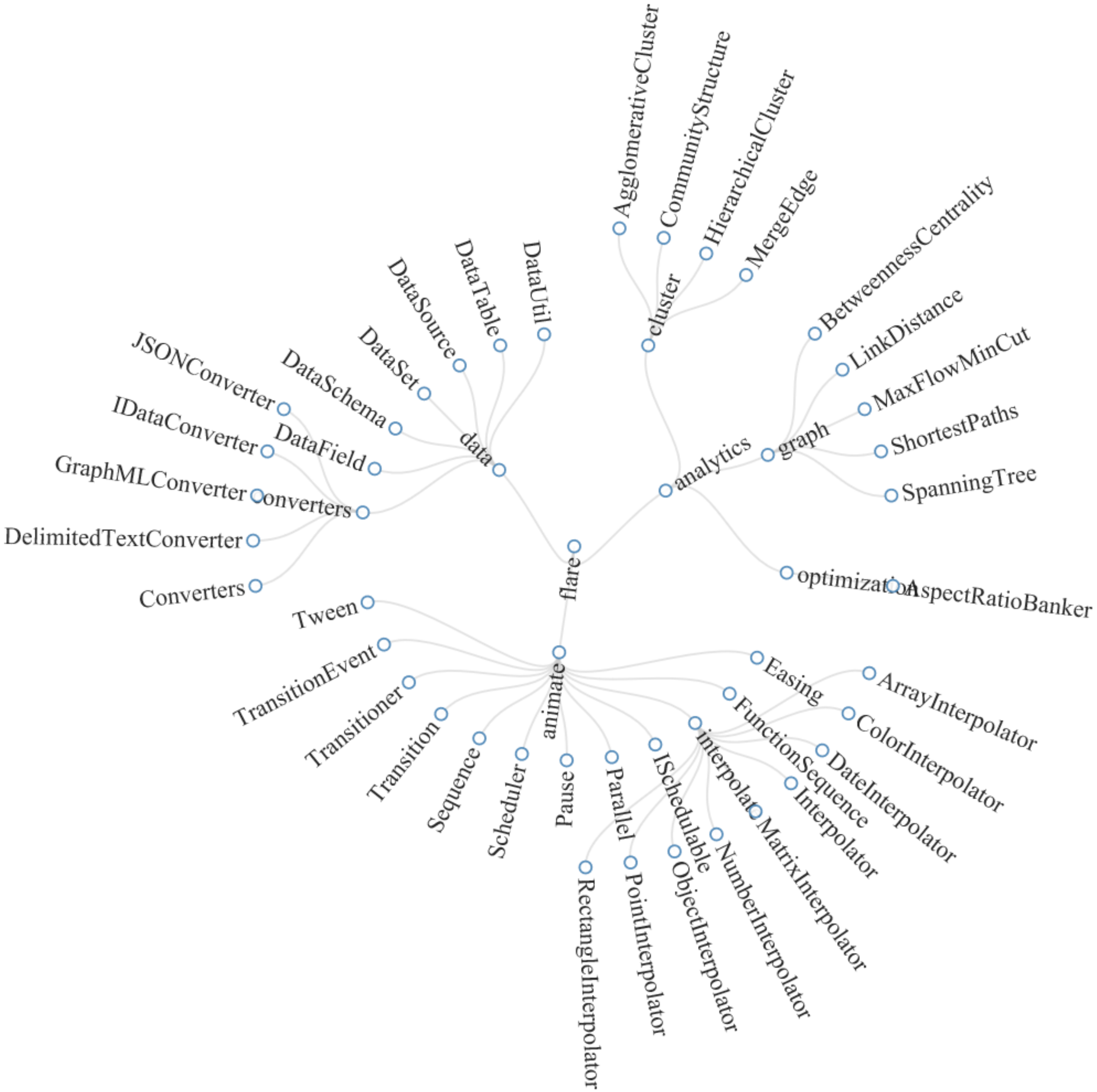
Usage

```
radialNetwork(List, height = NULL, width = NULL, fontSize = 10,
  fontFamily = "serif", linkColour = "#ccc", nodeColour = "#fff",
  nodeStroke = "steelblue", textColour = "#111", opacity = 0.9,
  margin = NULL)
```

```
> URL <- paste0("https://cdn.rawgit.com/christophergandrud/networkD3/",
+               "master/JSONdata//flare.json")
> Flare <- jsonlite::fromJSON(URL, simplifyDataFrame = FALSE)
> Flare$children = Flare$children[1:3]
> str(Flare)
```

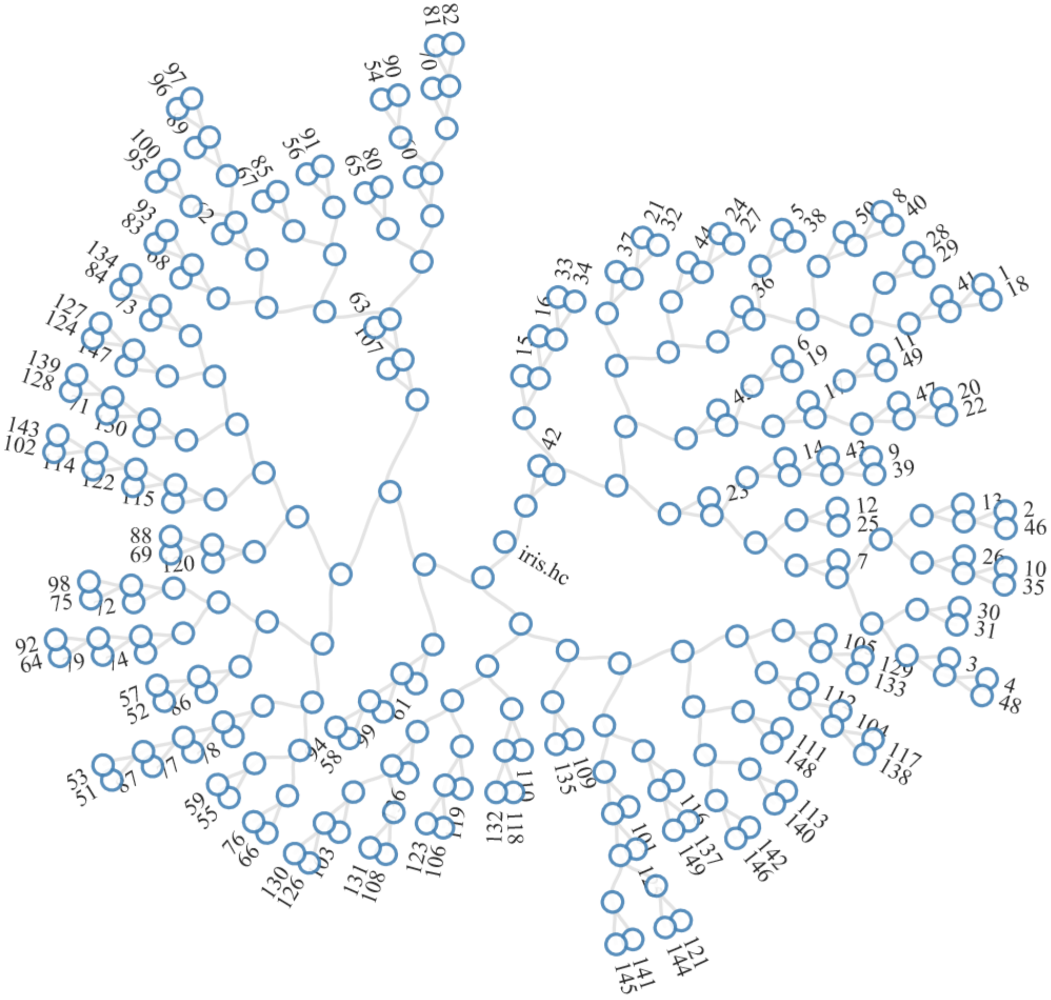
> radialNetwork(List = Flare, fontSize = 14)

```
> str(Flare)
List of 2
 $ name    : chr "flare"
 $ children:List of 3
  ..$ :List of 2
  .. ..$ name    : chr "analytics"
  .. ..$ children:List of 3
  .. .. ..$ :List of 2
  .. .. .. ..$ name    : chr "clus
  .. .. .. ..$ children:List of 4
  .. .. .. .. ..$ :List of 2
...
  ..$ :List of 2
  .. ..$ name    : chr "animate"
  .. ..$ children:List of 12
  .. .. ..$ :List of 2
  .. .. .. ..$ name: chr "Easing"
  .. .. .. ..$ size: int 17010
  .. .. ..$ :List of 2
...
  ..$ :List of 2
  .. ..$ name    : chr "data"
  .. ..$ children:List of 7
  .. .. ..$ :List of 2
  .. .. .. ..$ name    : chr "conv
  .. .. .. ..$ children:List of 5
  .. .. .. .. ..$ :List of 2
  .. .. .. .. .. ..$ name: chr "Cc
```

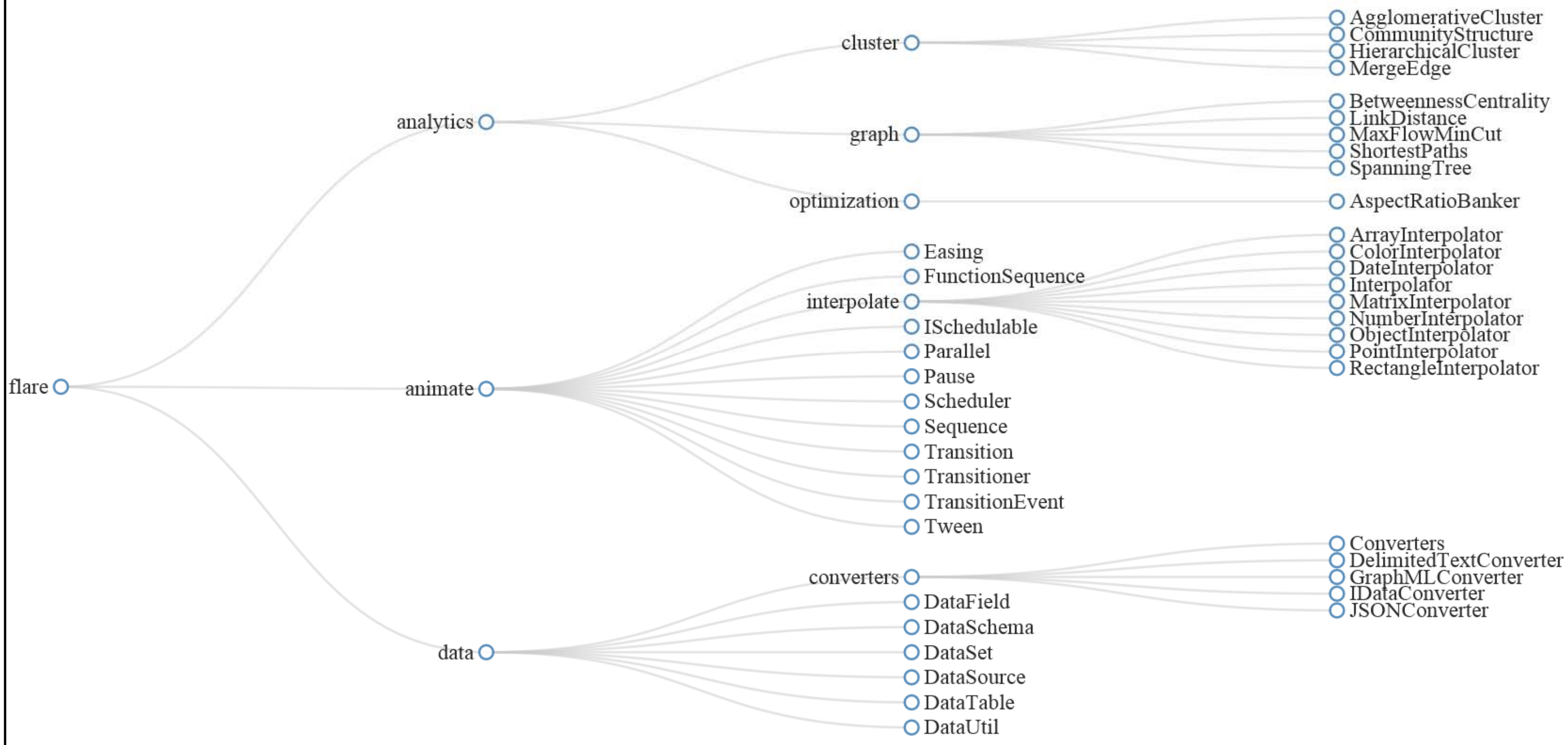


radialNetwork for a tree dendrogram

```
> iris.hc <- hclust(dist(iris[,1:4]), "ave")
> radialNetwork(as.radialNetwork(iris.hc))
```



diagonalNetwork(List = Flare, fontSize = 14)

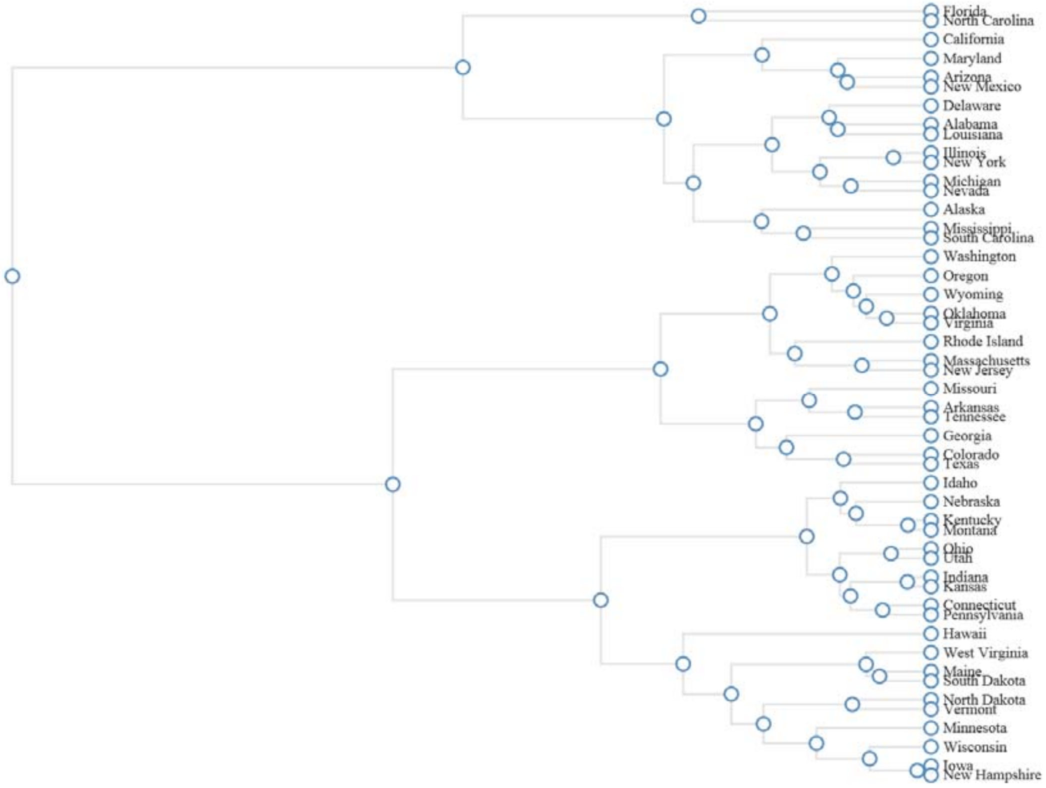


dendroNetwork {networkD3}: Create hierarchical cluster network diagrams

Usage

```
dendroNetwork(hc, height = 500, width = 800, fontSize = 10,
  linkColour = "#ccc", nodeColour = "#fff", nodeStroke = "steelblue",
  textColour = "#111", textOpacity = 0.9, textRotate = NULL,
  opacity = 0.9, margins = NULL, linkType = c("elbow", "diagonal"),
  treeOrientation = c("horizontal", "vertical"), zoom = FALSE)
```

```
> head(USArrests)
           Murder Assault UrbanPop Rape
Alabama      13.2     236       58 21.2
Alaska       10.0     263       48 44.5
Arizona       8.1     294       80 31.0
Arkansas      8.8     190       50 19.5
California    9.0     276       91 40.6
Colorado      7.9     204       78 38.7
> str(USArrests)
'data.frame':   50 obs. of  4 variables:
 $ Murder  : num  13.2 10 8.1 8.8 9 7.9
 $ Assault : int  236 263 294 190 276 20
 $ UrbanPop: int  58 48 80 50 91 78 77 7
 $ Rape    : num  21.2 44.5 31 19.5 40.6
> hc <- hclust(dist(USArrests), "ave")
> dendroNetwork(hc, height = 600)
```



Good and bad graphs

- Good graphs clearly show the important features of the data.
- They should always have:
  - a title;
  - labelled axes;
  - a key.
- In general they should tell a story and be memorable but also have a "low information to ink ratio" (junk kept to a minimum and no distracting features) and not mislead the viewer.
- Choice of color when designing charts and graphs is also important to allow for color blindness and black and white printing.

https://iase-web.org/islp/apps/gov_stats_graphing/GoodBad/GoodBadGraphs.pdf

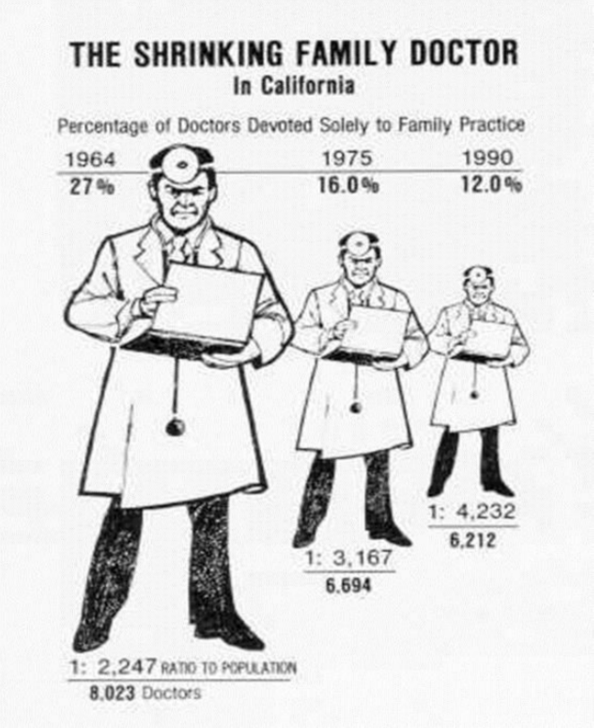
Good and bad graphs

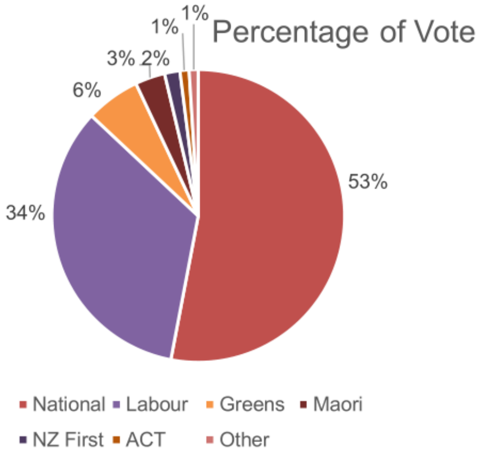
- Graphs are often made misleading for advertising or other purposes, or even just by accident, by:
  - Leaving gaps/changing the scale in vertical axes
  - Uneven shading/colours
  - Unfair emphasis on some sections
  - Distorting areas in histograms (bar widths should always be equal - if you have different widths then the bar height must be adjusted so AREAS reflect counts)
  - Use of 3-dimensions instead of two
  - Misleading use of pictograms

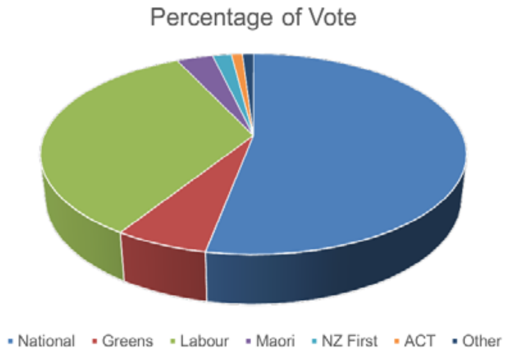
In particular, watch out for missing zero points on axes, spurious colouring and annotation, and unjustifiable extrapolation. Pictograms are often misleading as in the case of the graph as areas or volumes (instead of heights) are used to represent numbers exaggerating differences visually.

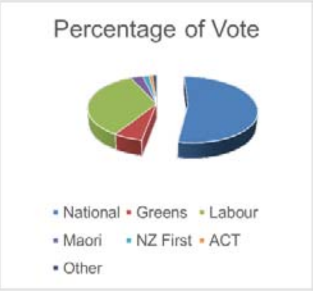
Example (from the Los Angeles Times, August 5 1979, p3)

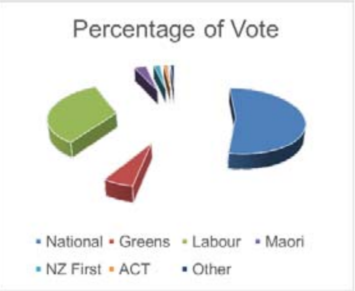
The data show a reduction in the ratio from 27% to 12%. this is represented by a change in the height of the doctor but your eye sees a change in the area.

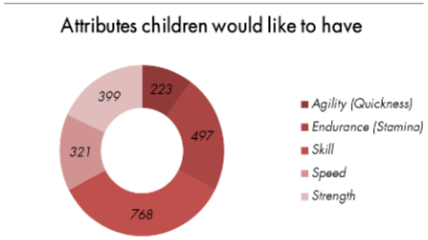




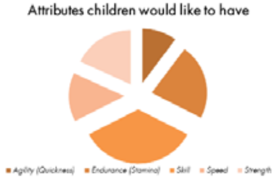




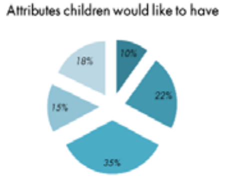




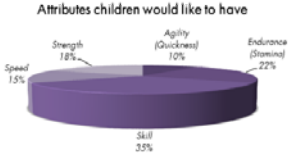
- Sectors don’t touch in the middle, (so angles can’t be compared easily)



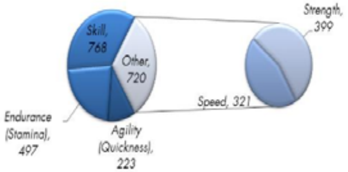
- Sectors should be touching each other
- No percentage labels
- No indication of how many children participated



- Sectors should be touching each other
- No key to say what each sector represents
- No indication as to how may children participated



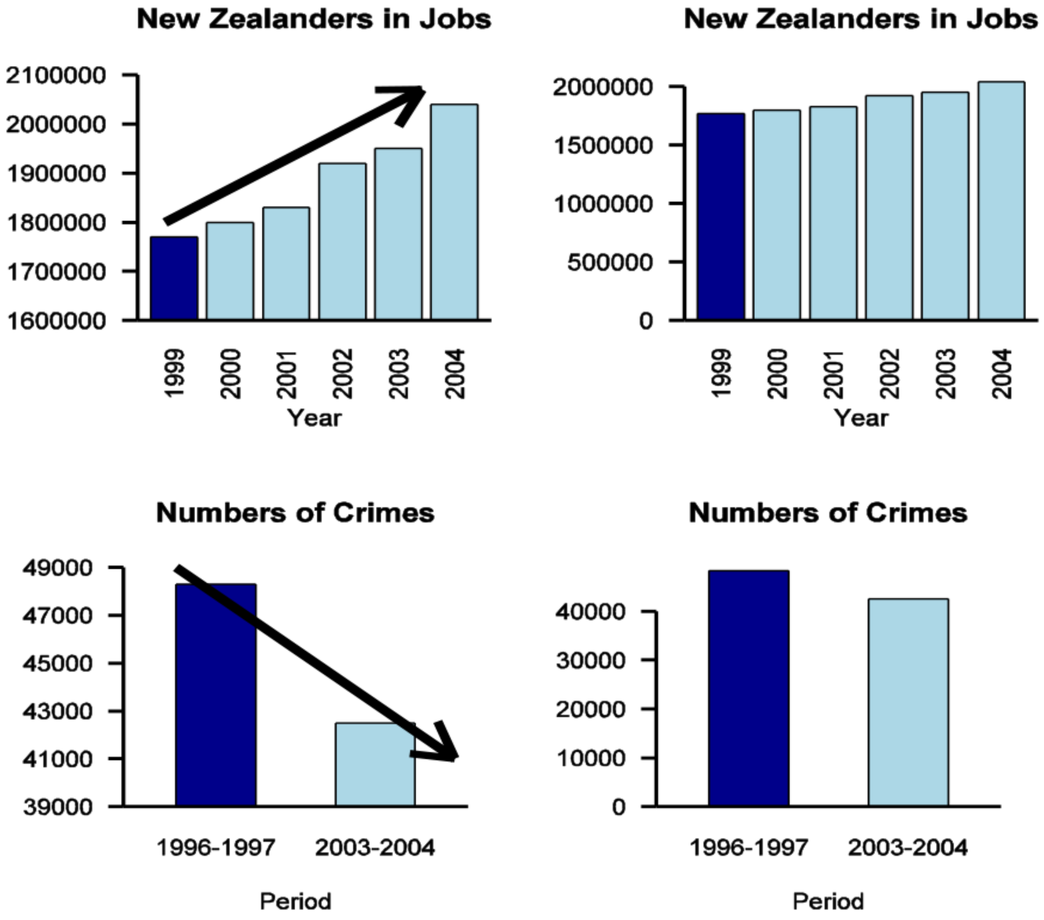
- Pie charts should always be 2-dimensional, in 3-dimensions the visual representation of the sectors is distorted
- A key should be used for the attributes



- All the sectors should be presented in the same pie
- A key should be used for the attributes
- No title

Bar charts

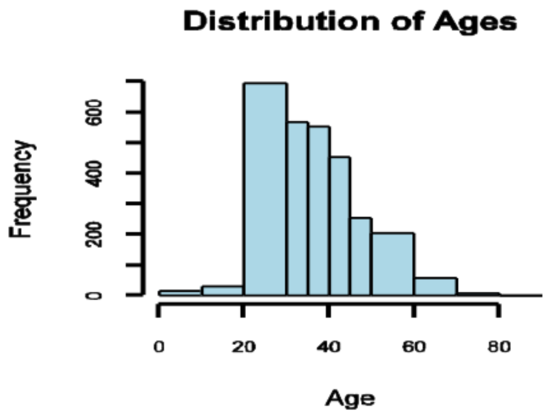
Both the (made up) graphs on the left mislead the reader by exaggerating the differences in the heights of the bars by not starting the vertical axis at zero. The correct graphs are given on the right.

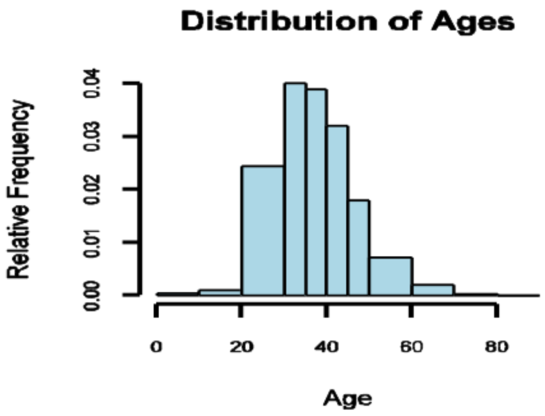


https://methodsavvy.com/good-data-bad-graphs/

Histograms

The graph on the left misleads the reader by doubling the width of some of the bars. The correct graph on the right halves the heights of these bars so that the area still represents the frequency.





The top ten worst graphs

The top ten worst graphs

https://www.biostat.wisc.edu/~kbroman/topten_worstgraphs/

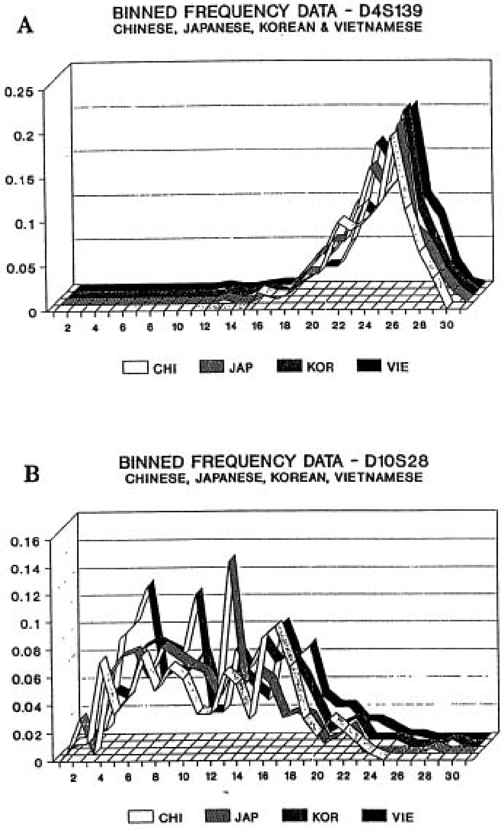
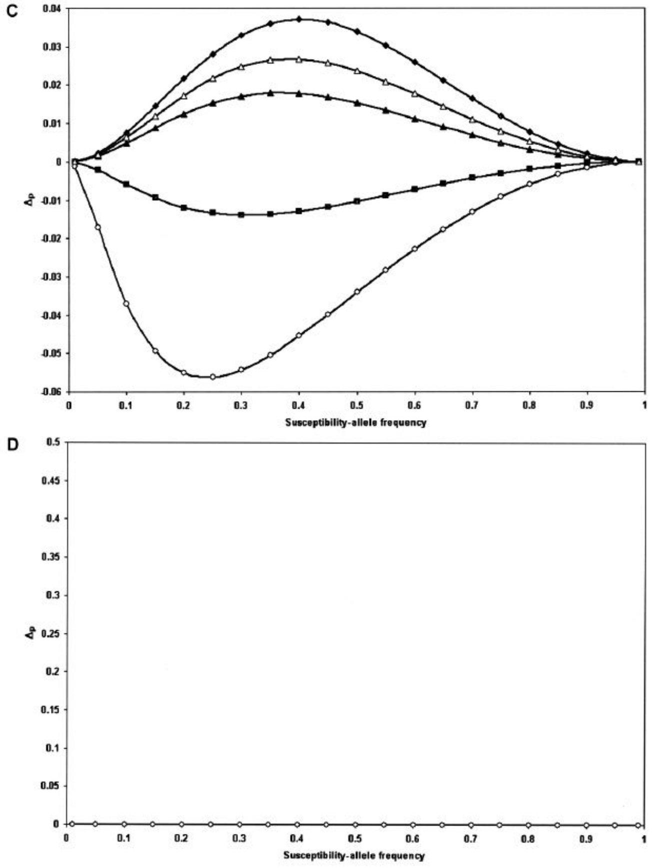


FIG. 4. *Fixed bin distribution (histogram) for two loci and four Asian subpopulations (used with permission from John Hartmann): the boundaries of the 30 bins (vertical axis) are determined by the FBI; these bins are not of equal length. Sample sizes (numbers of individuals) for Chinese, Japanese, Korean and Vietnamese are* 103, 125, 93 *and* 215 *for* D4S139 *and* 120, 137, 100 *and* 193 *for* D10S28. *The horizontal axis is the bin number; bins are not of equal length.*



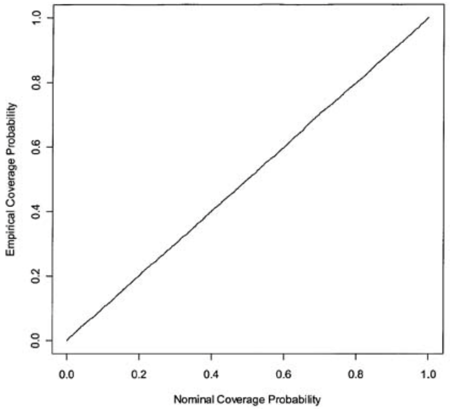


Figure 1 Empirical coverage of CIs for the relative-risk parameter β of haplotype 01100. Results are based on 10,000 simulated data sets with the same haplotype frequencies as the FUSION data. Haplotype 01100 has a multiplicative effect on disease risk, with $\beta = 0.35$.

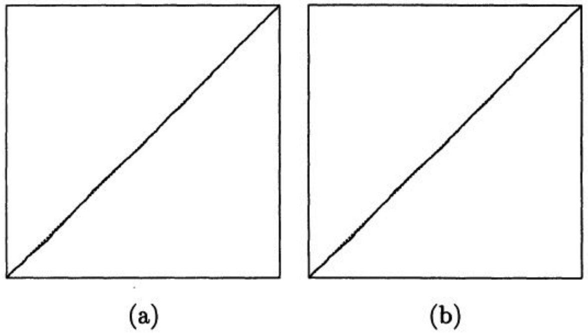


Figure 1. SRQ Plots of T_i/T_n (Vertical Axes) Against i/n (Horizontal Axes) for the Gibbs Sampler (a) and an Alternating Gibbs/Independence Sampler (b) for the Pump Failure Data Based on Runs of Length 5,000. Lines through the origin with unit slope are shown dashed; axis ranges are from 0 to 1 for all axes.

The top ten worst graphs

The top ten worst graphs

https://www.biostat.wisc.edu/~kbroman/topten_worstgraphs/

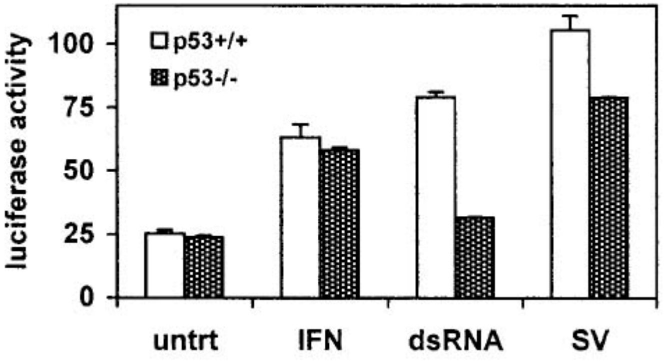


FIG. 4. ISG15 promoter activity mimics endogenous ISG15 mRNA regulation by p53, dsRNA, and virus. Cells (6×10^5 HCT 116) were seeded in 32-mm plates and allowed to attach overnight. Cells were transfected with 500 ng of pGL3/ISG15-Luc, 50 ng of pRL null (Promega), and 450 ng of pcDNA3 for carrier DNA by using Lipofectamine Plus (Life Technologies) following the manufacturer's instructions. Twenty-four hours posttransfection, the medium was aspirated and replaced with medium containing either 1,000 U of IFN-α/ml, 50 µg of dsRNA/ml, or Sendai virus (multiplicity of infection, 10). Cells were incubated for 12 h and then lysed, and luciferase assays were performed. Luciferase activity was assessed on 20 µl of each lysate as directed by the supplier (Dual Luciferase Kit, Promega) using a TD 20/20 luminometer (Turner Designs). Luciferase activity is presented as the ratio of firefly activity to renilla activity to control for differences in transfection efficiency. Each data point is the mean of triplicate samples ± the standard error; the data presented are representative of four independent experiments.

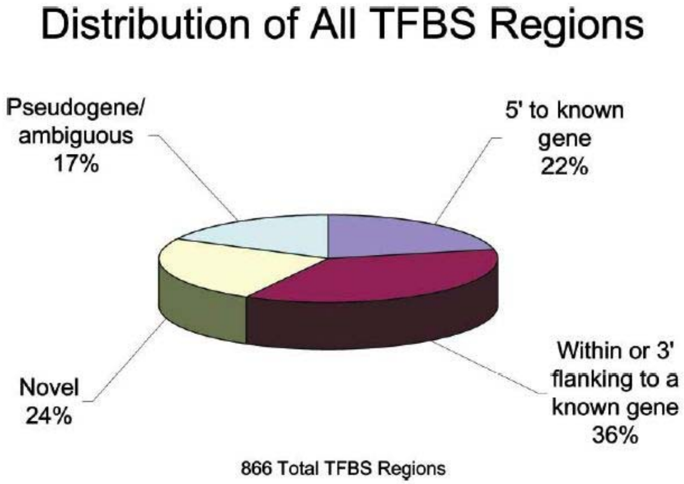


Figure 1. Classification of TFBS Regions

TFBS regions for Sp1, cMyc, and p53 were classified based upon proximity to annotations (RefSeq, Sanger hand-curated annotations, GenBank full-length mRNAs, and Ensembl predicted genes). The proximity was calculated from the center of each TFBS region. TFBS regions were classified as follows: within 5 kb of the 5′ most exon of a gene, within 5 kb of the 3′ terminal exon, or within a gene, novel or outside of any annotation, and pseudogene/ambiguous (TFBS overlapping or flanking pseudogene annotations, limited to chromosome 22, or TFBS regions falling into more than one of the above categories).

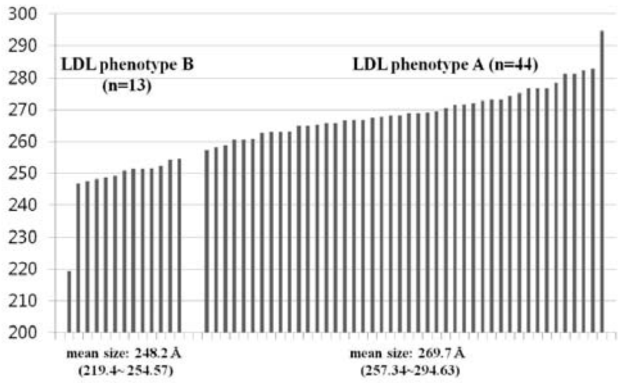


Fig. 1. Distribution of low-density lipoprotein (LDL) particle size in all study subjects (LDL phenotypes A and B). *LDL phenotype A group* (mean size: 269.7 Å, n = 44), subjects with buoyant-mode profiles [peak LDL particle diameter ≥ 264 Å] including intermediate LDL subclass pattern [256 Å ≤ peak LDL particle diameter ≤ 263 Å]; *LDL phenotype B group* (mean size: 248.2 Å, n = 13), subjects with dense-mode profiles [peak LDL particle diameter ≤ 255 Å]

The top ten worst graphs

The top ten worst graphs

https://www.biostat.wisc.edu/~kbroman/topten_worstgraphs/

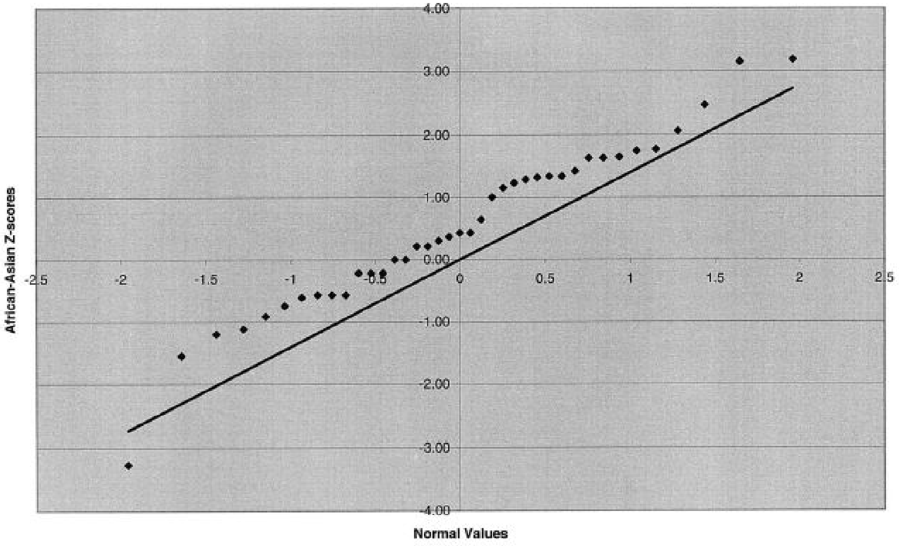
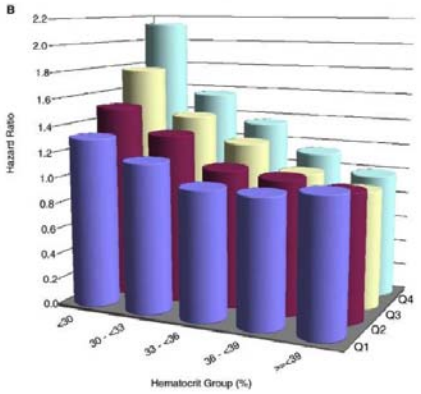


Figure 2 Q-Q plots of *Z* scores for telomeric interval-length differences. *a*, African Americans versus Asians. *b*, Whites versus Asians.



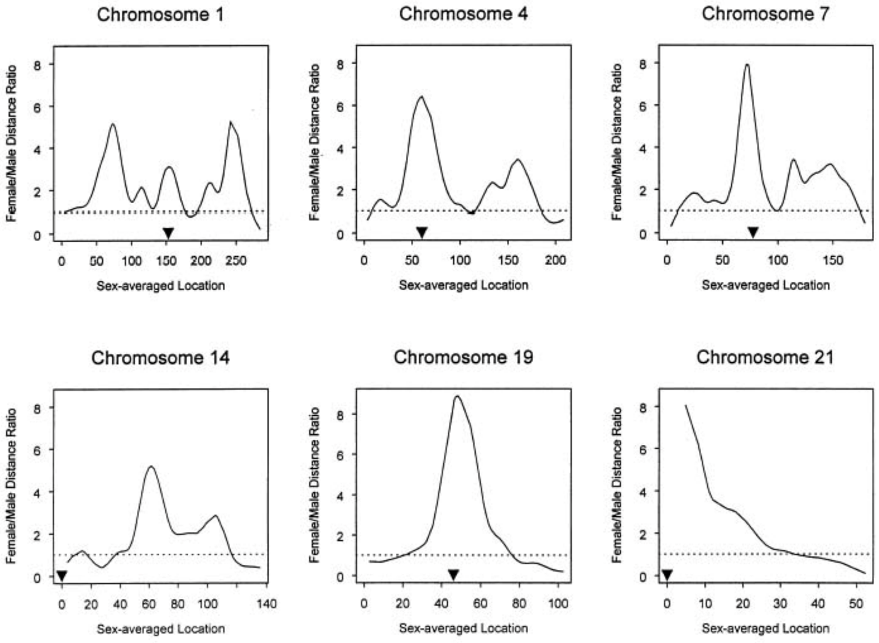
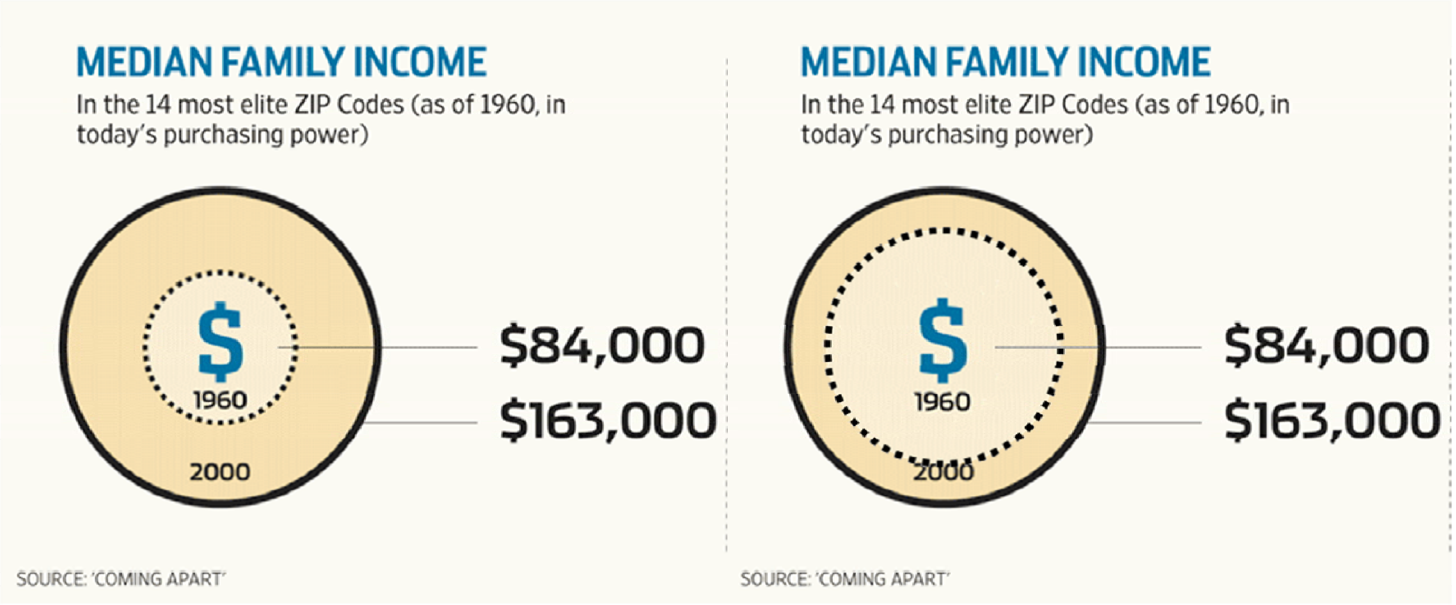


Figure 1 Plots of the female:male genetic-distance ratio against sex-averaged genetic location (in cM) along six selected chromosomes. Approximate locations of the centromeres are indicated by the triangles. The dashed lines correspond to equal female and male distances.

Tragic Graphic: The Wall Street Journal Lies with Statistics?

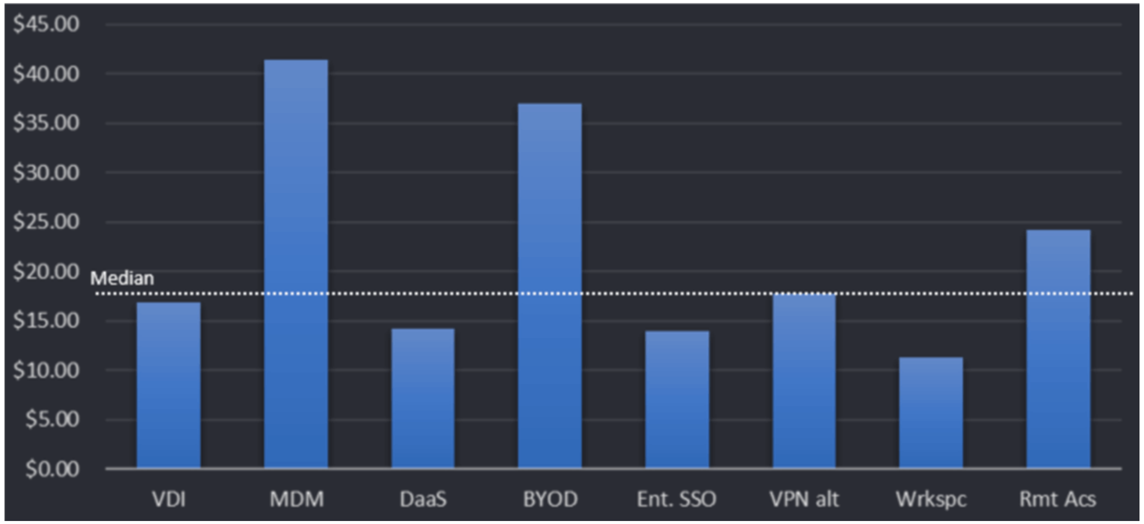


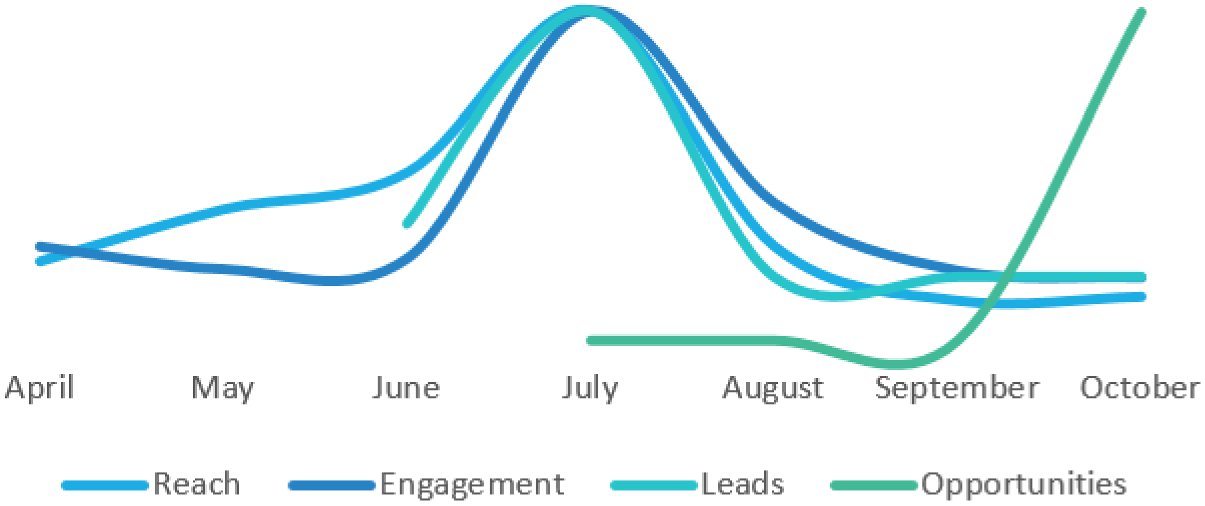
Inaccurate graph as it appeared in the *Wall Street Journal* (1/21/2012)

Accurate graph constructed by EvalBlog.com

https://evalblog.com/2012/01/23/tragic-graphic-the-wall-street-journal-lies-with-statistics/

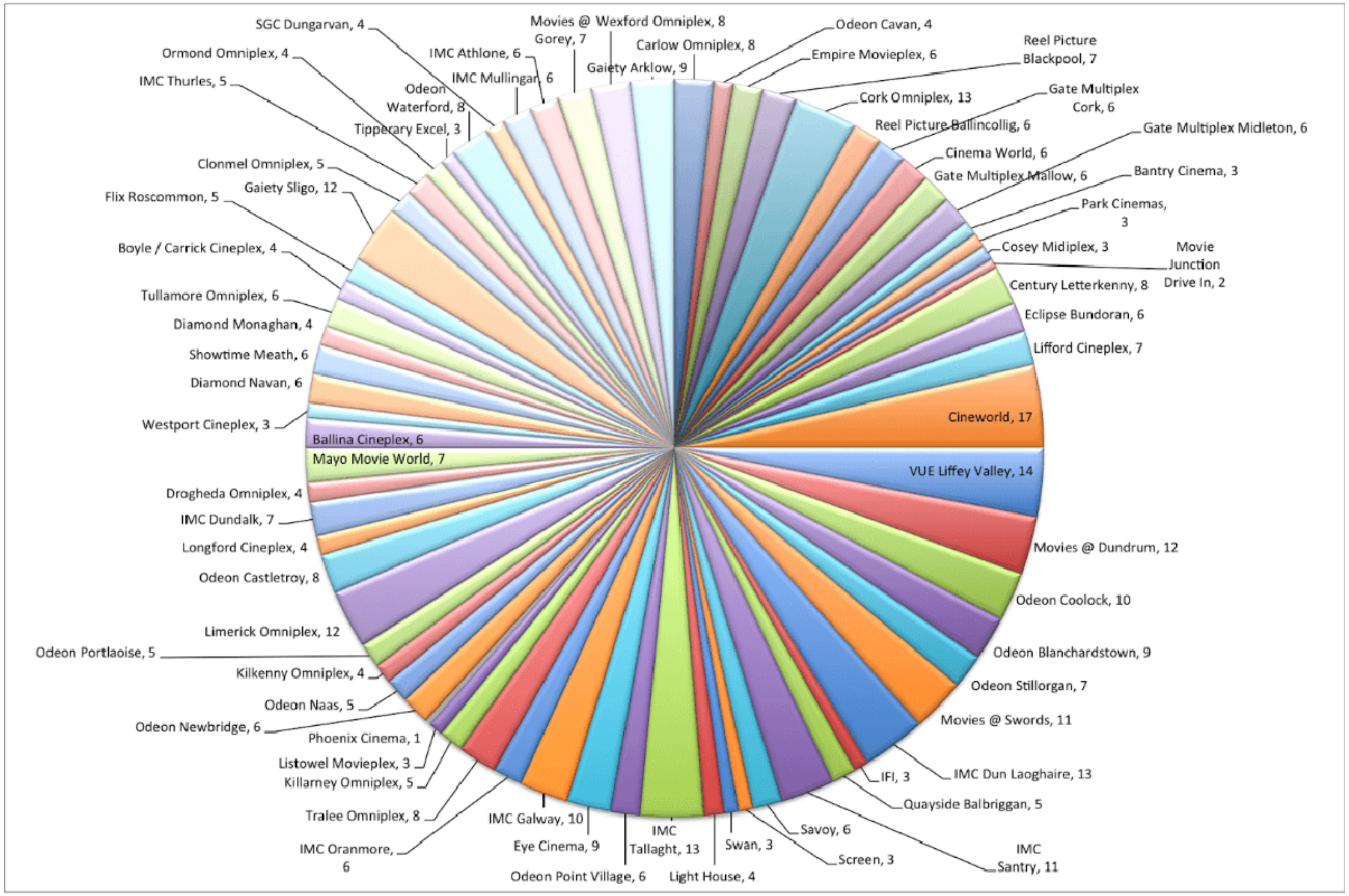
Use Labels Properly





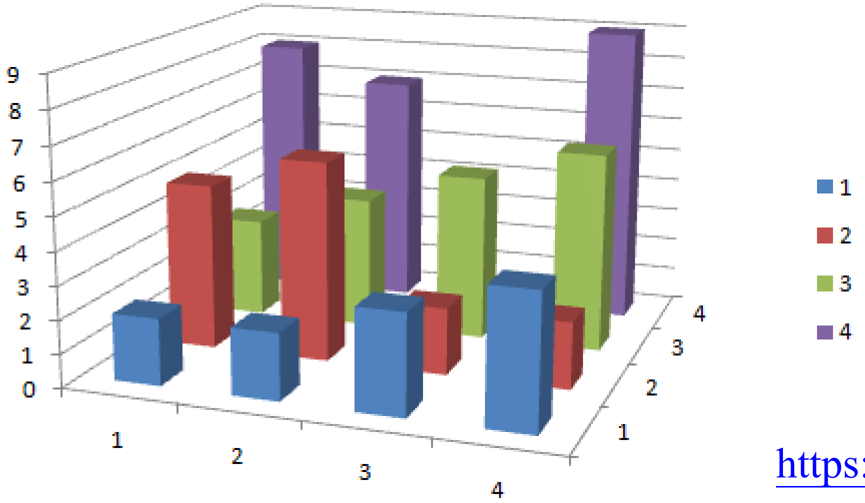
https://methodsavvy.com/good-data-bad-graphs/

Limit Dimensions



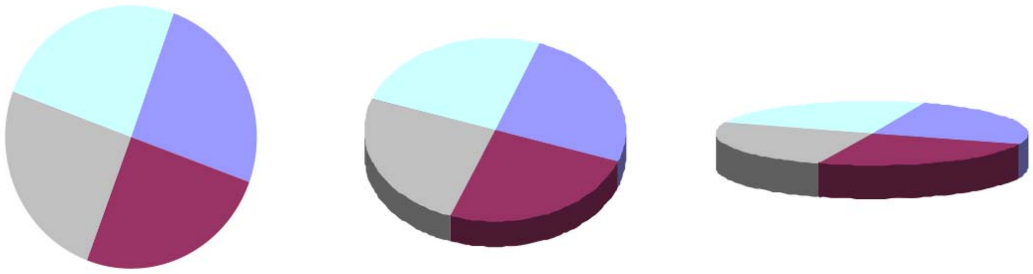
https://methodsavvy.com/good-data-bad-graphs/

3D Graphs – Just Say No

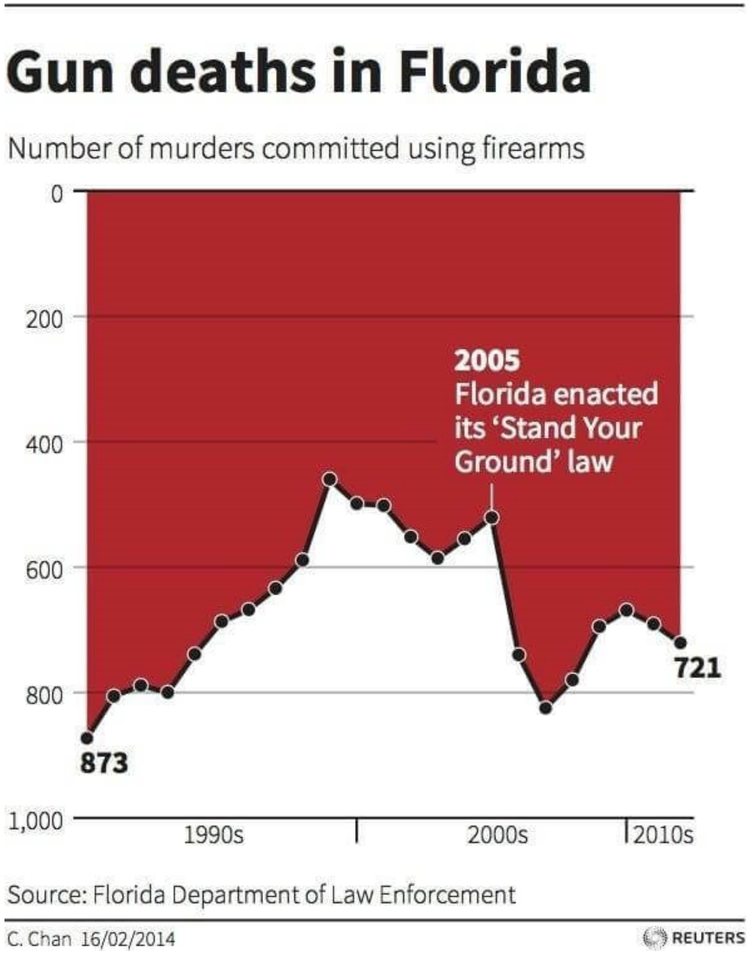


https://methodsavvy.com/good-data-bad-graphs/

Stay away from 3-D charts!



Deceptive Graphs

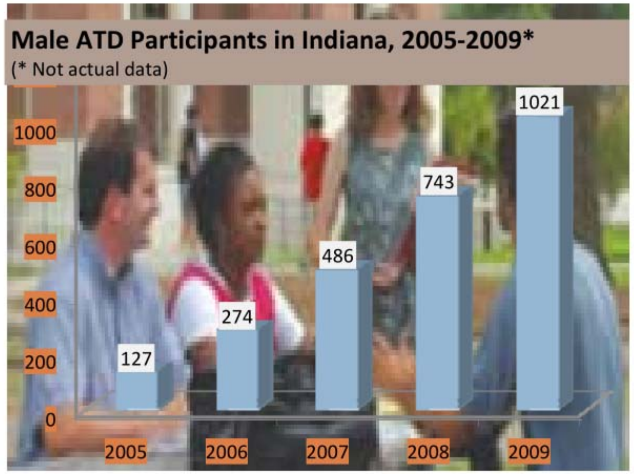


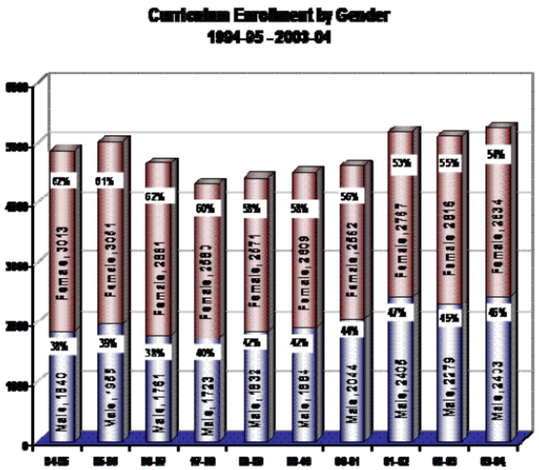
This next graph is honestly pretty deceptive to the average casual reader and goes a long way to demonstrate how data can be clearly used but easily misunderstood. At first glance, this graph would appear to indicate gun violence in Florida has substantially worsened in recent years. Then after the passage of the Stand Your Ground legislation, gun violence immediately decreased.

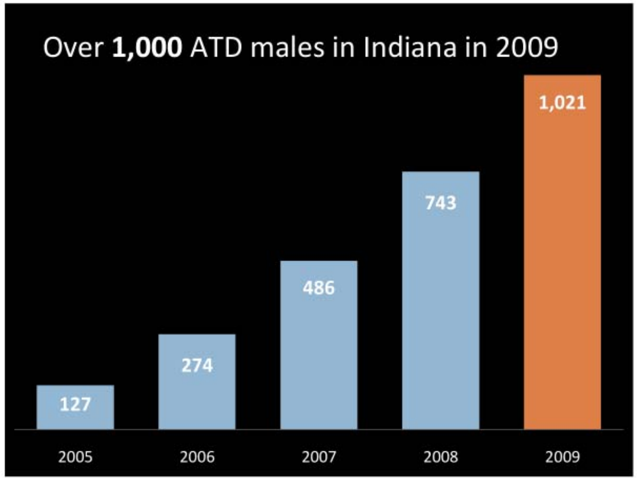
However, the careful observer will notice the Y axis is inverted. The story of gun violence decreasing is a false narrative. The data shows that gun violence was actually lower before the passage of the legislation, and actually got substantially worse after the law was passed.

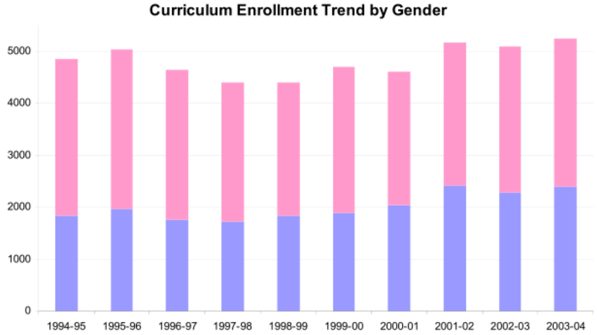
https://methodsavvy.com/good-data-bad-graphs/

Bad graphs







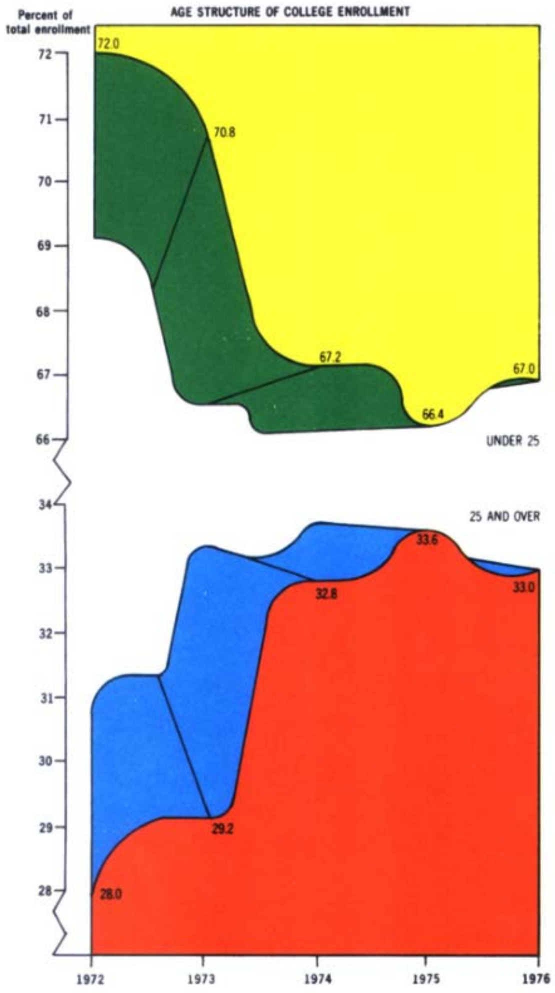


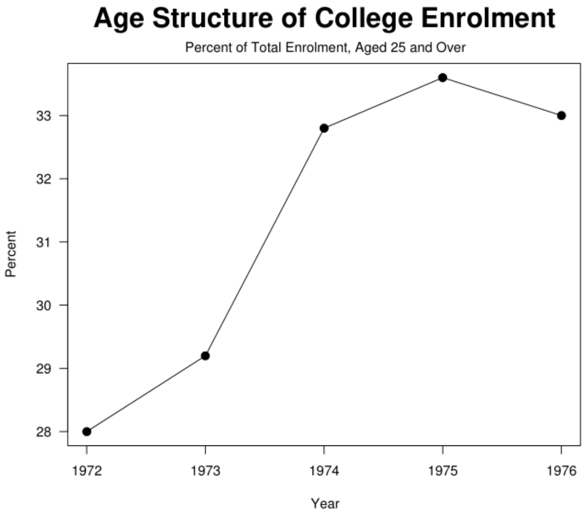
http://aaude.org/system/files/documents/public/strategy-institute-handout-final.pdf

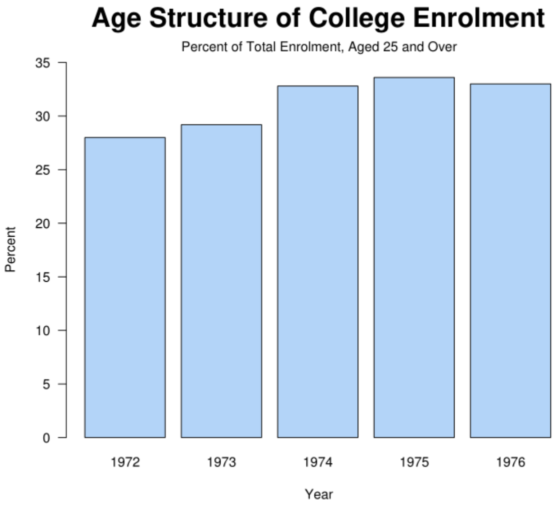
Bad graphs

Age Structure of College Enrolment (1972-1976)

This graph presents five values. It uses six colours, unnecessary perspective and a split axis to do so. American Education Magazine.


AGE STRUCTURE OF COLLEGE ENROLLMENT
Percent of total enrollment
UNDER 25
25 AND OVER


Age Structure of College Enrolment
Percent of Total Enrolment, Aged 25 and Over
Percent
Year


Age Structure of College Enrolment
Percent of Total Enrolment, Aged 25 and Over
Percent
Year