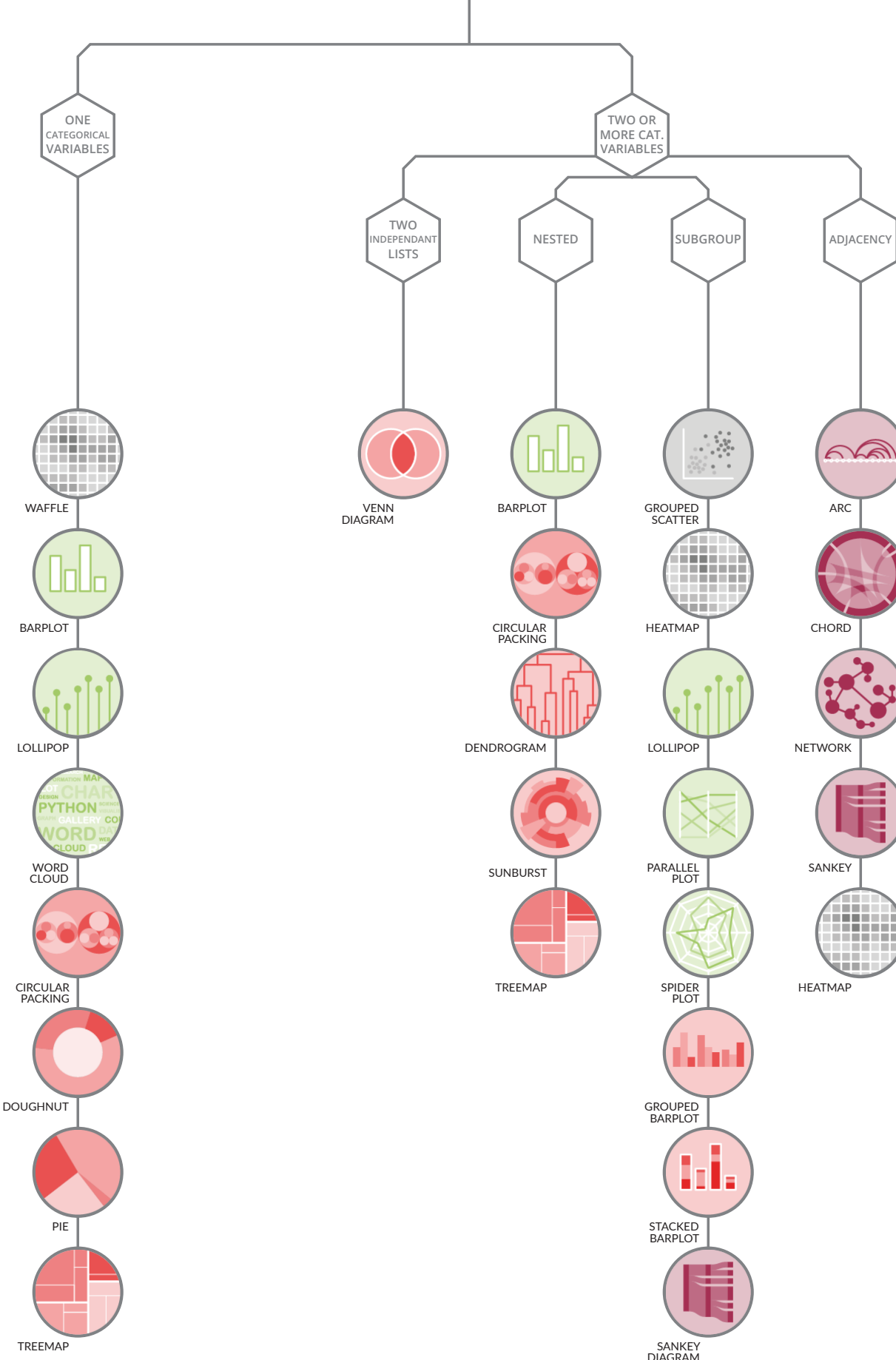# from Data to Viz

From Data to Viz is a classification of chart types based on their input data format. It can be seen as a decision tree leading to a set of potential appropriate visualizations to represent your dataset.

This project is built on two underlying philosophies. First, that most of the data analysis can be summarized in about twenty different dataset formats. Second, that it is the data and the context that decide the chart. Thus the suggested method consists of identifying and trying all the feasible chart types to find out which suits your data and idea best.
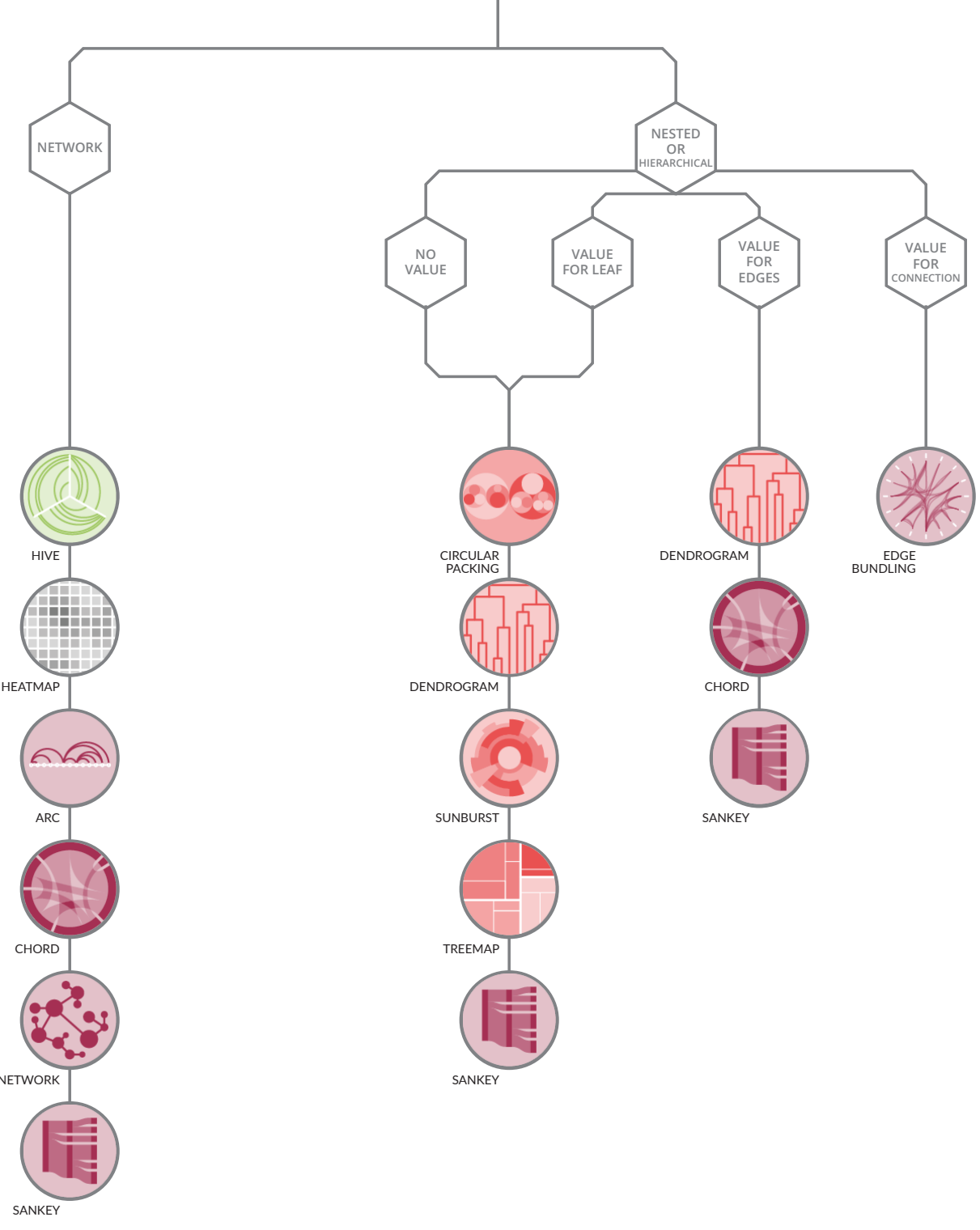
Dataviz is a world with endless possibilities and this project does not claim to be exhaustive. However it should lead you to a good starting point. An interactive version of this classification is available online. It provides more information on each chart type and highlights a few pitfalls to avoid. For each dataset format, an example of real data analysis is offered to illustrate the different chart types with their pro and con.
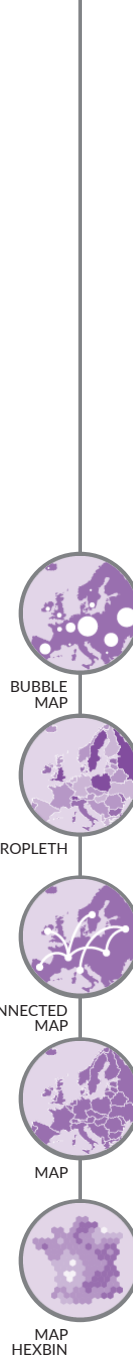
www.data-to-viz.com

CATEGORIC

ONE CATEGORICAL VARIABLES
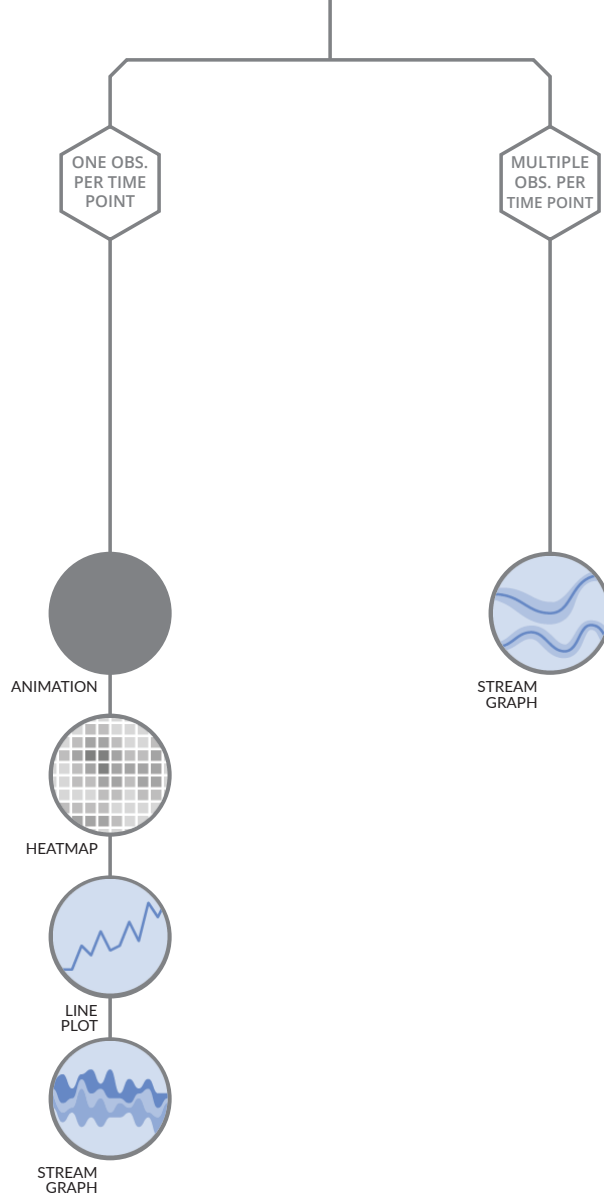- WAFFLE
- BARPLOT
- LOLLIPOP
- WORD CLOUD
- CIRCULAR PACKING
- DOUGHNUT
- PIE
- TREEMAP

TWO OR MORE CAT. VARIABLES

TWO INDEPENDANT LISTS
- VENN DIAGRAM

NESTED
- BARPLOT
- CIRCULAR PACKING
- DENDROGRAM
- SUNBURST
- TREEMAP

SUBGROUP
- GROUPED SCATTER
- HEATMAP
- LOLLIPOP
- PARALLEL PLOT
- SPIDER PLOT
- GROUPED BARPLOT
- STACKED BARPLOT
- SANKEY DIAGRAM

ADJACENCY
- ARC
- CHORD
- NETWORK
- SANKEY
- HEATMAP

RELATIONAL

NETWORK
- HIVE
- HEATMAP
- ARC
- CHORD
- NETWORK
- SANKEY

NESTED OR HIERARCHICAL

NO VALUE
- CIRCULAR PACKING
- DENDROGRAM
- SUNBURST
- TREEMAP
- SANKEY

VALUE FOR LEAF

VALUE FOR EDGES
- DENDROGRAM
- CHORD
- SANKEY

VALUE FOR CONNECTION
- EDGE BUNDLING

MAP

ONE OBS. PER TIME POINT
- BUBBLE MAP
- CLOROPLETH
- CONNECTED MAP
- MAP
- MAP HEXBIN

TIME SERIES

ONE OBS. PER TIME POINT
- ANIMATION
- HEATMAP
- LINE PLOT
- STREAM GRAPH

MULTIPLE OBS. PER TIME POINT
- STREAM GRAPH

CATEGORIC AND NUMERIC

ONE NUM. ONE CAT.

ONE OBS. PER GROUP
- WAFFLE
- BARPLOT
- LOLLIPOP
- WORD CLOUD
- CIRCULAR PACKING
- DOUGHNUT
- PIE
- TREEMAP

SEVERAL OBS. PER GROUP
- BOXPLOT
- VIOLIN
- RIDGE LINE
- DENSITY
- HISTOGRAM

ONE CAT. SEVERAL NUM.

NO ORDER
- BOXPLOT
- VIOLIN
- GROUPED SCATTER
- 2D DENSITY
- PCA
- CORRELOGRAM

A NUM. IS ORDERED
- CONNECTED SCATTER
- AREA
- LINE PLOT
- STACKED AREA
- STREAM GRAPH

ONE VALUE PER GROUP
- GROUPED SCATTER
- HEATMAP
- LOLLIPOP
- PARALLEL PLOT
- SPIDER PLOT
- GROUPED BARPLOT
- STACKED BARPLOT
- SANKEY DIAGRAM

SEVERAL CAT. ONE NUM.

SUBGROUP

ONE OBS. PER GROUP
- BOXPLOT
- VIOLIN

SEVERAL OBS. PER GROUP
- BOXPLOT
- VIOLIN

NESTED

ONE OBS. PER GROUP
- BARPLOT
- CIRCULAR PACKING
- DENDROGRAM
- SUNBURT
- TREEMAP

SEVERAL OBS. PER GROUP
- BOXPLOT
- VIOLIN

ADJACENCY
- HEATMAP
- ARC
- CHORD
- NETWORK
- SANKEY

NUMERIC

ONE NUMERIC VARIABLE
- DENSITY
- HISTOGRAM

TWO NUMERIC VARIABLES

NOT ORDERED

FEW POINTS
- BOX PLOT
- HISTOGRAM
- SCATTER PLOT

MANY POINTS
- DENSITY
- VIOLIN
- 2D DENSITY
- MARGINAL DISTRIBUTION

ORDERED
- CONNECTED SCATTER
- AREA PLOT
- LINE PLOT

THREE NUMERIC VARIABLES

NOT ORDERED
- BOXPLOT
- VIOLIN
- BUBBLE PLOT
- 3D SCATTER OR SURFACE

ORDERED
- AREA
- LINE PLOT
- STACKED AREA
- STREAM GRAPH

SEVERAL NUMERIC

ORDERED

NOT ORDERED
- BOXPLOT
- RIDGE LINE
- VIOLIN
- CORRELOGRAM
- HEATMAP
- PCA
- DENDROGRAM