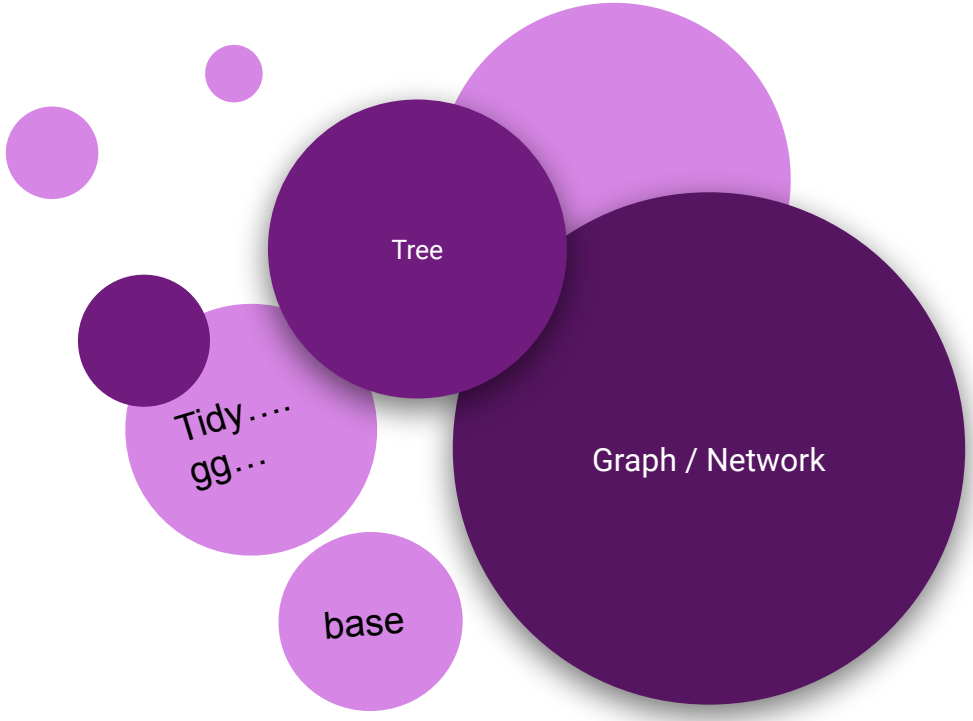
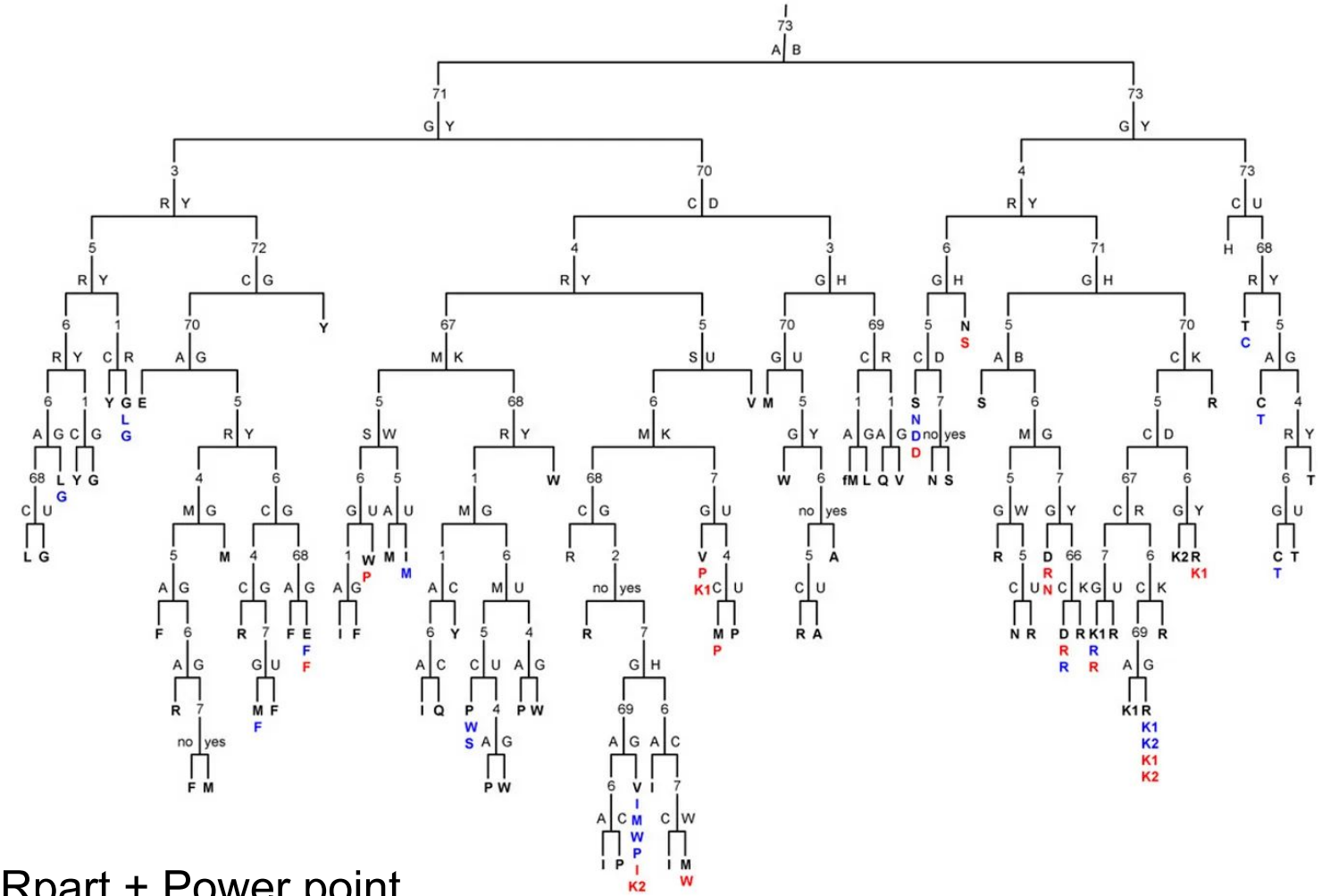
The evolution of specialized R packages for trees and graphs/networks

Dror Berel
dror.berel@gmail.com


Tree
Tidy….
gg…
Graph / Network
base



Rpart + Power point

Tree Functionality

Traversal

Prune

Aggregate

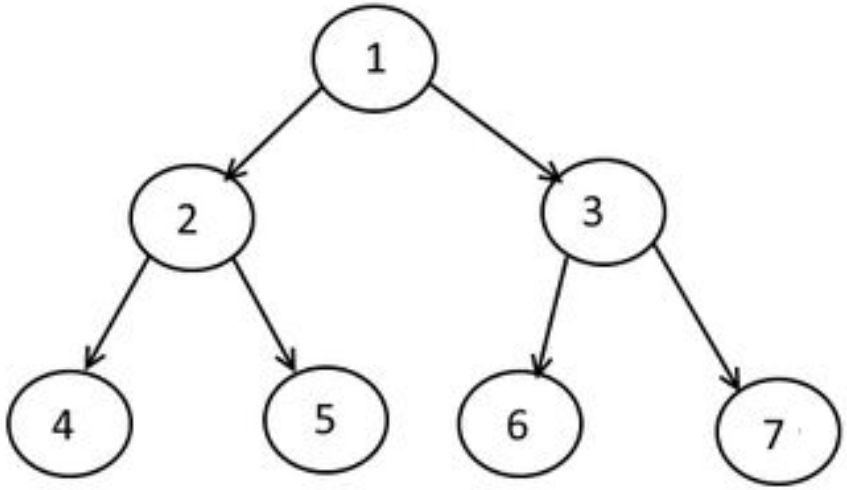
Cumulate

Sort

Plot

Print

(*graphviz, DOT)*



Inorder Traversal: 4 2 5 1 6 3 7
Preorder Traversal: 1 2 4 5 3 6 7
Postorder Traversal: 7 6 3 5 4 2 1
Breadth-First Search: 1 2 3 4 5 6 7
Depth-First Search: 1 2 4 5 3 6 7

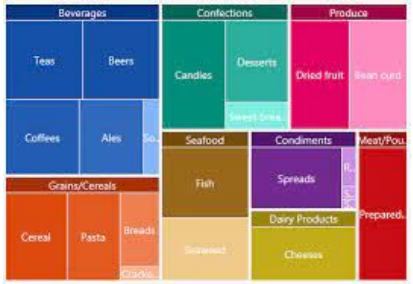
‘Tree’ Use cases:

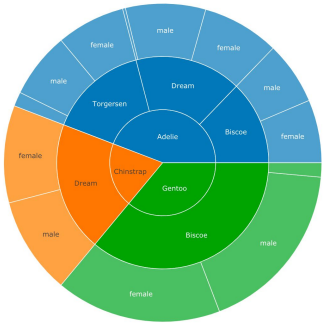
Transform data frame for treemap/sunburst/Sankey format

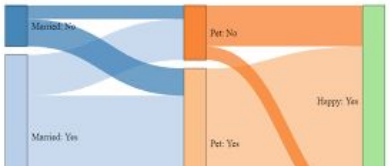
Deep nested list of list (of list) e.g. JSON / YMAL format

Folders / File system

Org plotly







Tree R packages

Structure	Back in the day	Intermediate	Modern (now)	Specialized	Base
Tree	stats: hclust, dendrogram …	ape phylobase	*treedataverse*: treeio, ggtree, tidytree	dendExtend data.tree	rraply

Trees are also:

- Bi-directional, ordered graphs.
- Undirected graph in which any two vertices are connected by exactly one path.
- Connected acyclic undirected graph

Graph / Network R packages

Structure	Back in the day	Intermediate	Modern (now)	Specialized	Base
Graph / Network	sna, statnet, network…	igraph	*TidyGraph* + ggraph		

‘Graph’ use cases

Shortest path

Best split

Centrality (hub)

Communication

Dynamic networks

Goal: Seamless conversion across data structures, with minimum information loss

Nesting lists

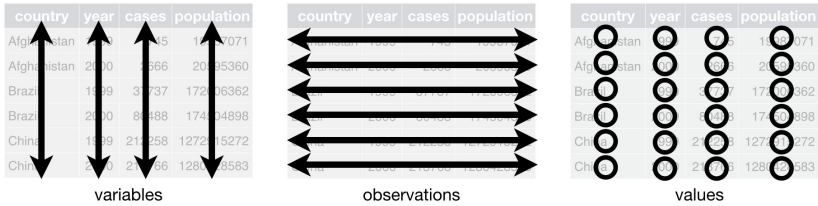
Nested list

i. item 1
   - sub item 1
   - sub item 2
     a. sub item 1
     b. sub item 2

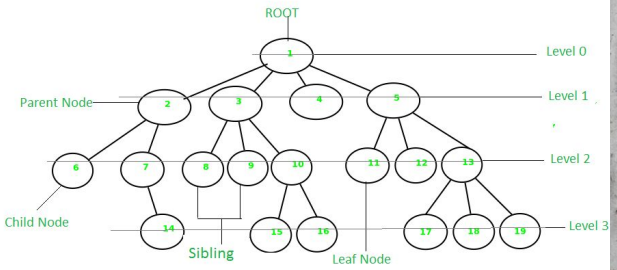
ii. item 2
   1. sub item 1
      - sub item 1
      - sub item 2
   2. sub item 2

iii. item 3

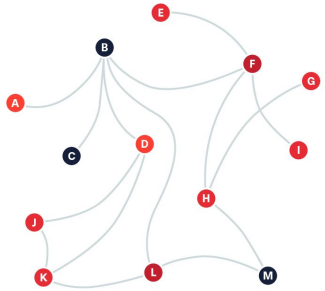
Table (data.frame, tibble, data.table, matrix,...)



Tree



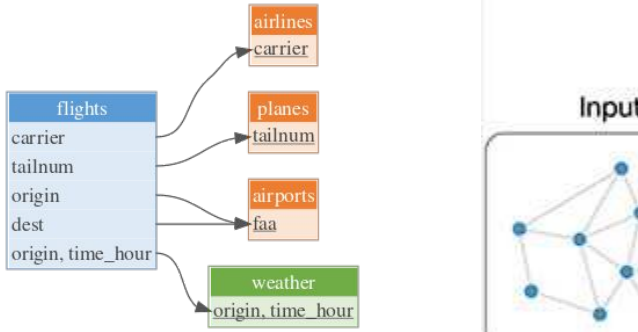
Graph / Network

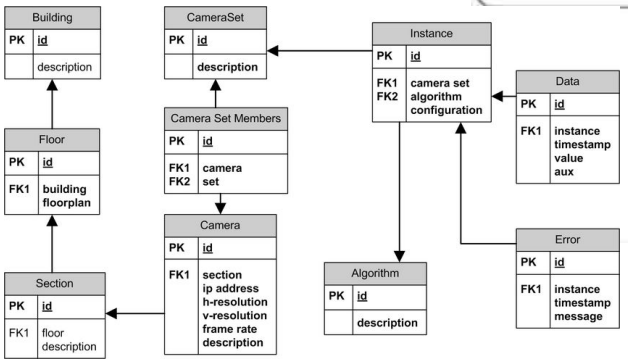


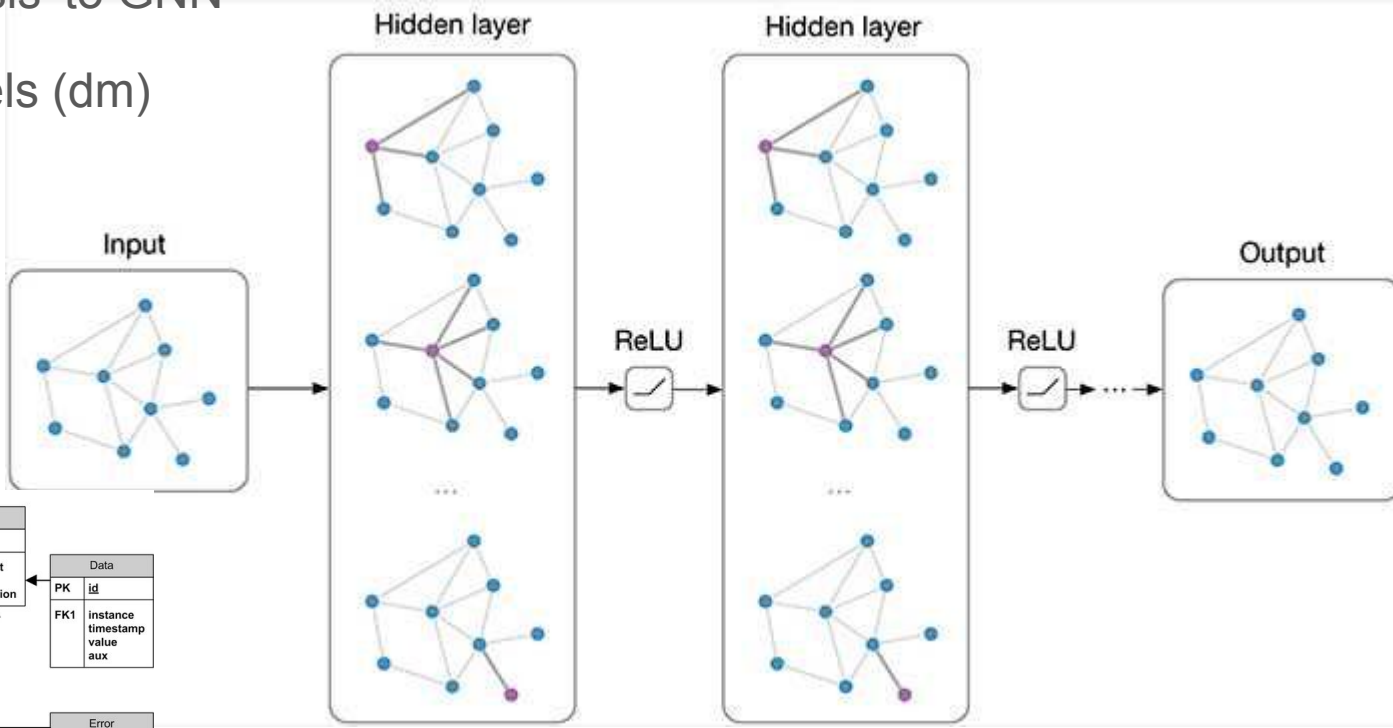
Wish list

From ‘Network analysis’ to GNN

Relational Data Models (dm)







Conclusion

Structure	Back in the day	Intermediate	Modern (now)	Specialized	Base
Tree	stats: hclust, dendrogram …	ape phylobase	*treedataverse*: treeio, ggtree, tidytree	dendExtend data.tree	rraply
Graph / Network	sna, statnet, network…	igraph	*TidyGraph* + ggraph		

https://**drorberel.medium.com**/the-evolution-of-specialized-r-packages-for-trees-and-graphs-networks-9ddb7c75b38c

Take home messages

1. ‘Tidy’ makes sense
2. Do not hesitate to peek under the hood, to find good old familiar face
3. Understand the package scope, and relationship to other packages in the ecosystem
4. ‘Base’ will always be there for you when you need it
5. Use Case motivation is fun, but learn when it is not worth going down the rabbit hole

dror.berel@gmail.com