# 1 Graphs Networks and Trees

Mostly from: Visual Analysis of Large Graphs

...recall Schneiderman Data Types.

A **graph** is a set of nodes with a set of edges (relations) between them. We have directed graphs and undirected graphs. Nodes and edges can also be characterized differently using edge/node types. Colors, shapes, etc. The nodes and edges can also have different weights, size is a good indicator of this. A **tree** is a special case of a graph in which any two of the nodes are connected by exactly one path. No cycles also.

In graph theory, a network is a directed graph with weighted edge. In info viz, a network is any graph with attributes associated to nodes and edges. A minimum spanning tree is a subset of a graph connecting all vertices of a graph. **compound graphs** are a way of grouping nodes together to create a tree structure.

**Dynamic Graphs**. These can evolve over time, with changes occuring in atttributes of nodes, edges, structure, or combinations of these. The *data* is changing.

**Topological properties:** graph size - the number of nodes. Graph density - number of edges. Sparse graphs have few edges, complete graphs have everything connected, cliques are a subset of a graph which is fully connected.

So what is a large graph? In algorithmic analysis, this means long computation times or a large memory footprint. In info viz, this means anything leading to cluttered displays. A question we ask is *how can we reduce the size of our graph?*.

We look at the degree of the notes. Graph filtering: stochastic sampling - randomly remove nodes or edges from graph. Geodesic clustering: looks at structure of graph and tries to keep connection information and throwing away duplicate information. Structure based filtering: remove any edges so long as elements remain connected.