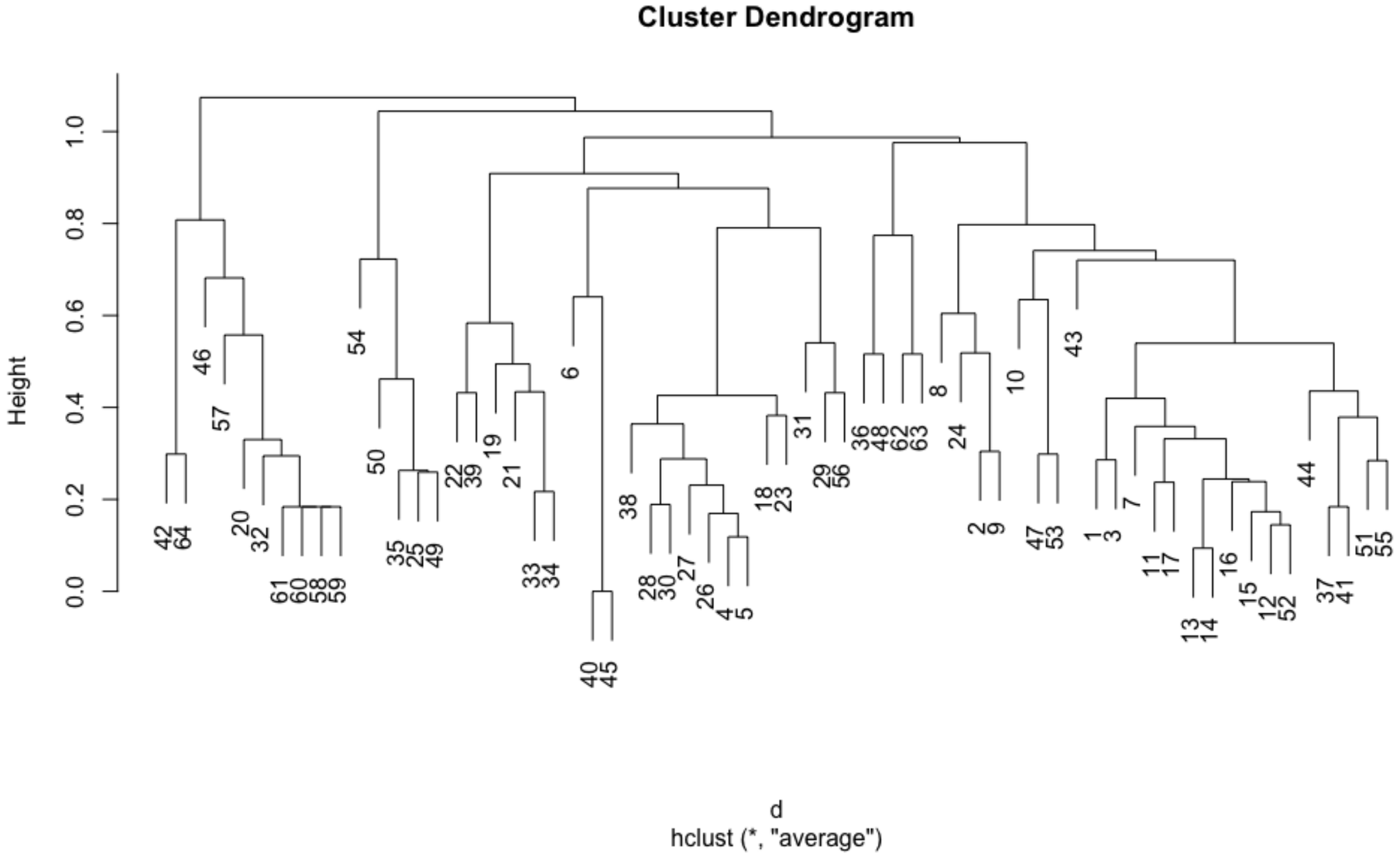
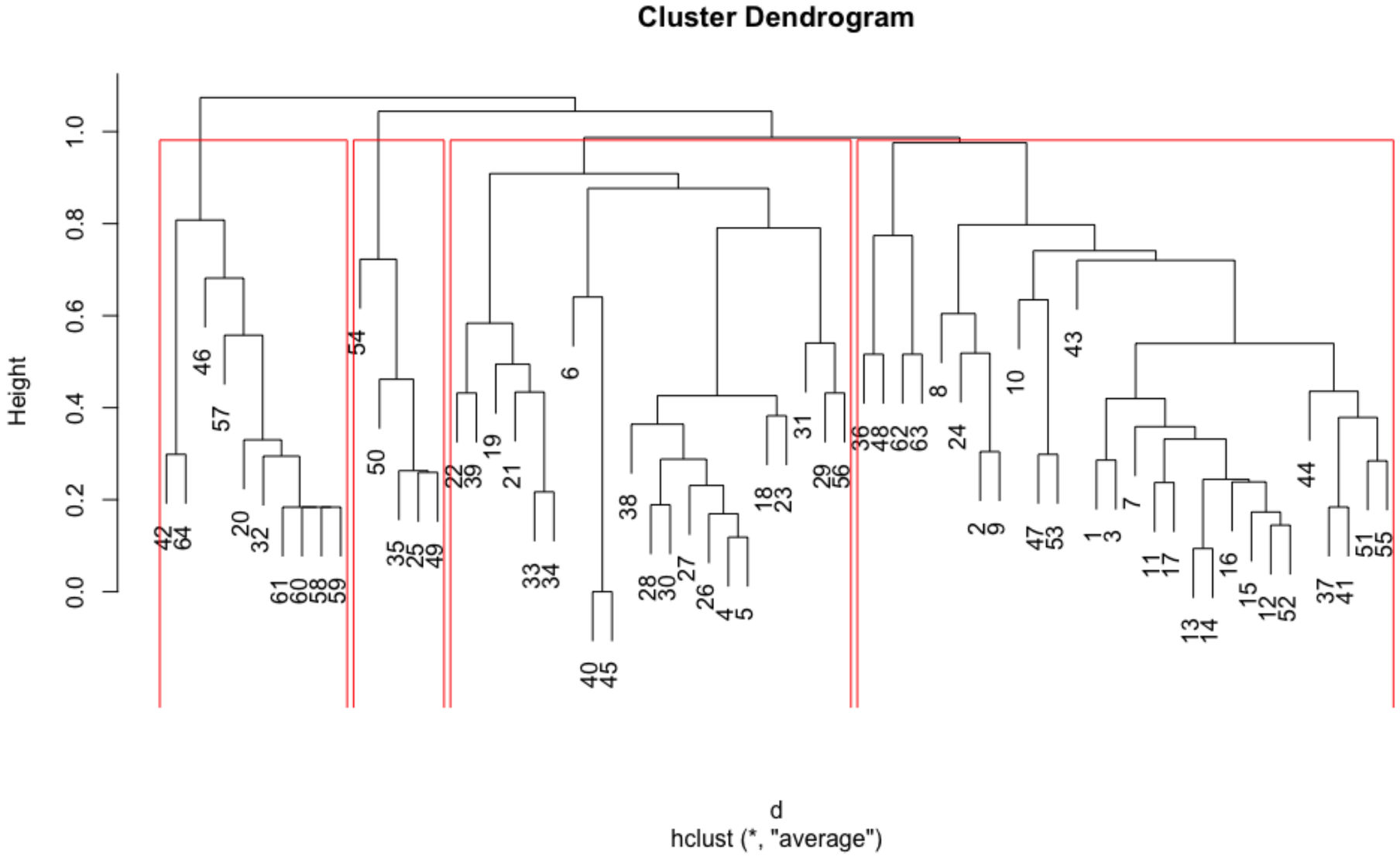
Hierarchical clustering

NETWORK ANALYSIS IN THE TIDYVERSE

Massimo Franceschet

Prof. of Data Science, University of Udine (Italy)


Cluster Dendrogram
Height
0.0
0.2
0.4
0.6
0.8
1.0
d
hclust (*, "average")


Cluster Dendrogram
Height
0.0
0.2
0.4
0.6
0.8
1.0
d
hclust (*, "average")

The similarity measure

- **Single-linkage**: the similarity between two groups is the maximum of the similarities between nodes of different groups.
- **Complete-linkage**: the similarity between two groups is the minimum of the similarities between nodes of different groups.
- **Average-linkage**: the similarity between two groups is the average of the similarities between nodes of different groups.

The clustering algorithm

1. Evaluate the similarity measures for all node pairs.
2. Assign each node to a group of its own.
3. Find the pair of groups with the highest similarity and join them together into a single group.
4. Calculate the similarity between the new composite group and all others.
5. Repeat steps 3 and 4 until all nodes have been joined into a single group.

Hierarchical clustering in R

```
# distance matrix from similarity matrix
D <- 1-S
# distance object from distance matrix
d <- as.dist(D)
# average-linkage clustering method
cc <- hclust(d, method = "average")
# cut dendrogram at 4 clusters
hclust(d, method = "average")
```

```
[1] 1 1 1 2 2 2 1 1 1 1 1 1 1 1 1 1 2 2 3 2 2 2 1 4 2 2 2
[29] 2 2 2 3 2 2 4 1 1 2 2 2 1 3 1 1 2 3 1 1 4 4 1 1 1 4 1 2
[57] 3 3 3 3 3 1 1 3
```

Let's cluster our network!

NETWORK ANALYSIS IN THE TIDYVERSE

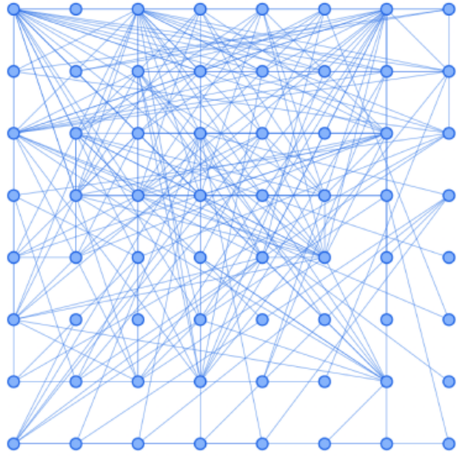
Interactive visualizations with visNetwork

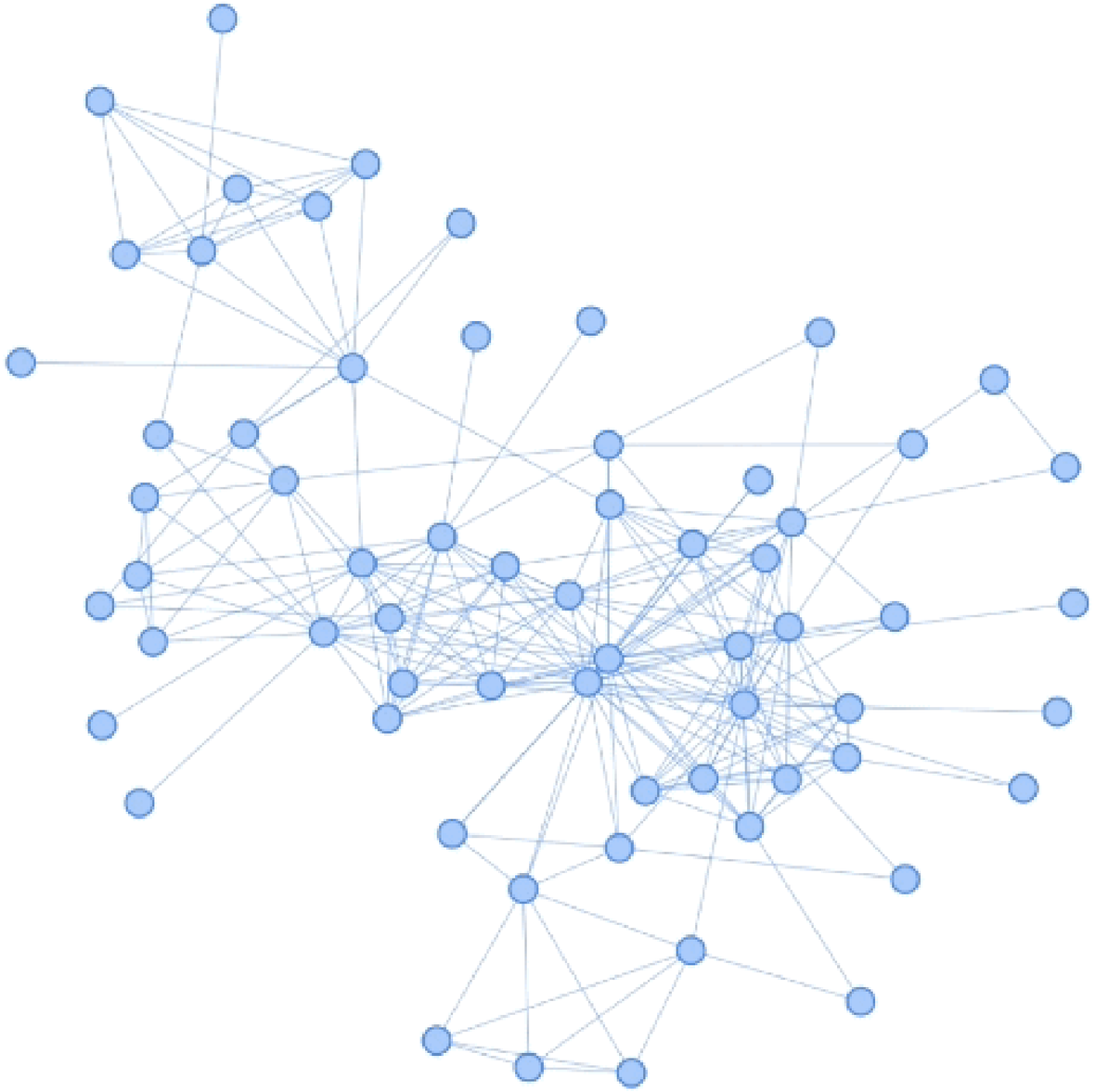
NETWORK ANALYSIS IN THE TIDYVERSE

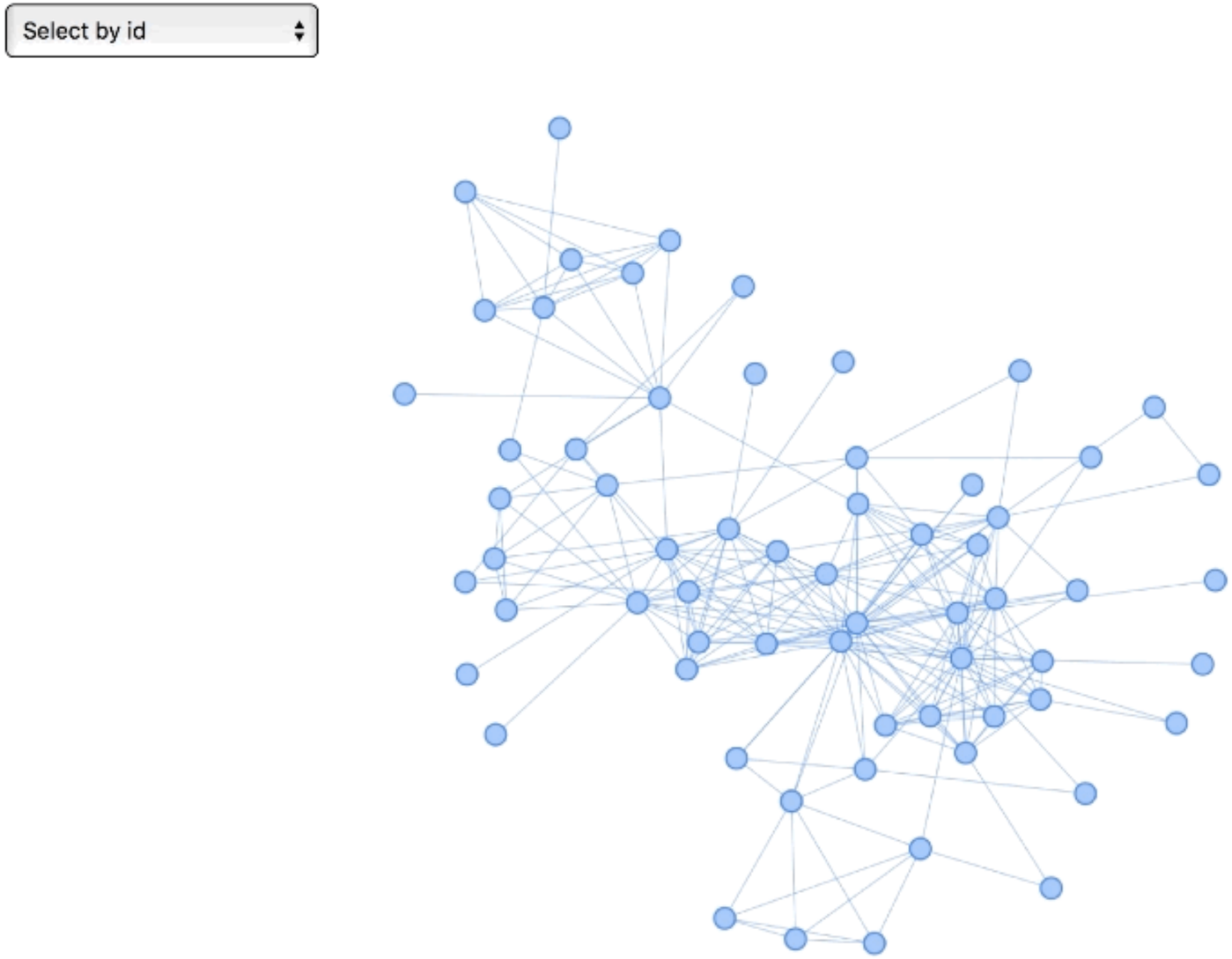
Massimo Franceschet
Prof. of Data Science, University of Udine (Italy)

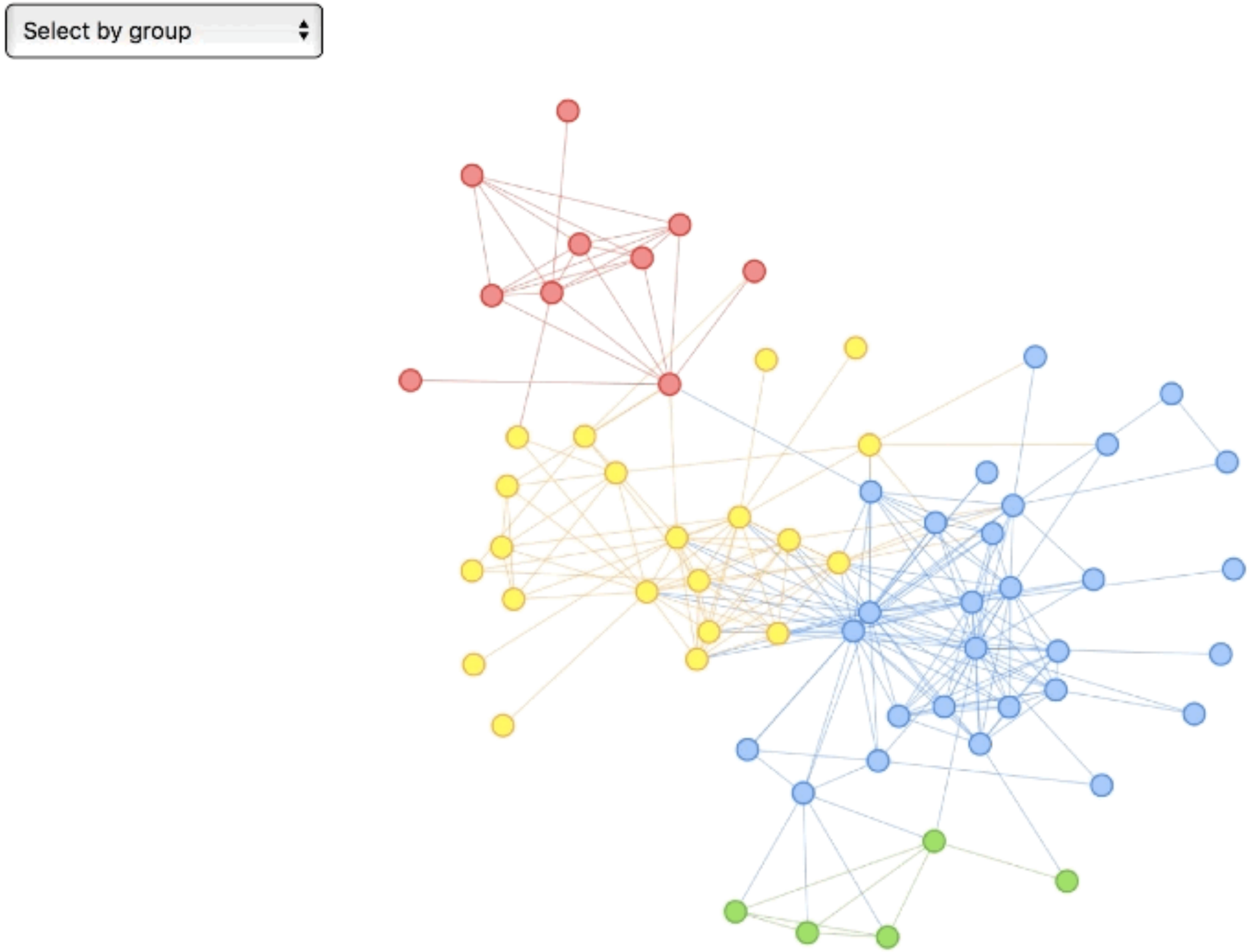
datacamp

Different layouts






Select by id


Select by group

Let's interact!

NETWORK ANALYSIS IN THE TIDYVERSE

Congratulations!

NETWORK ANALYSIS IN THE TIDYVERSE

Massimo Franceschet

Prof. of Data Science, University of Udine (Italy)

Deeper inside network science

You now know how to:

- Analyze any network with basic centrality and similarity measures
- Produce beautiful network visualizations, including interactive ones

For more information:

- **University of Udine Network Science Course**

Continue the journey!

NETWORK ANALYSIS IN THE TIDYVERSE