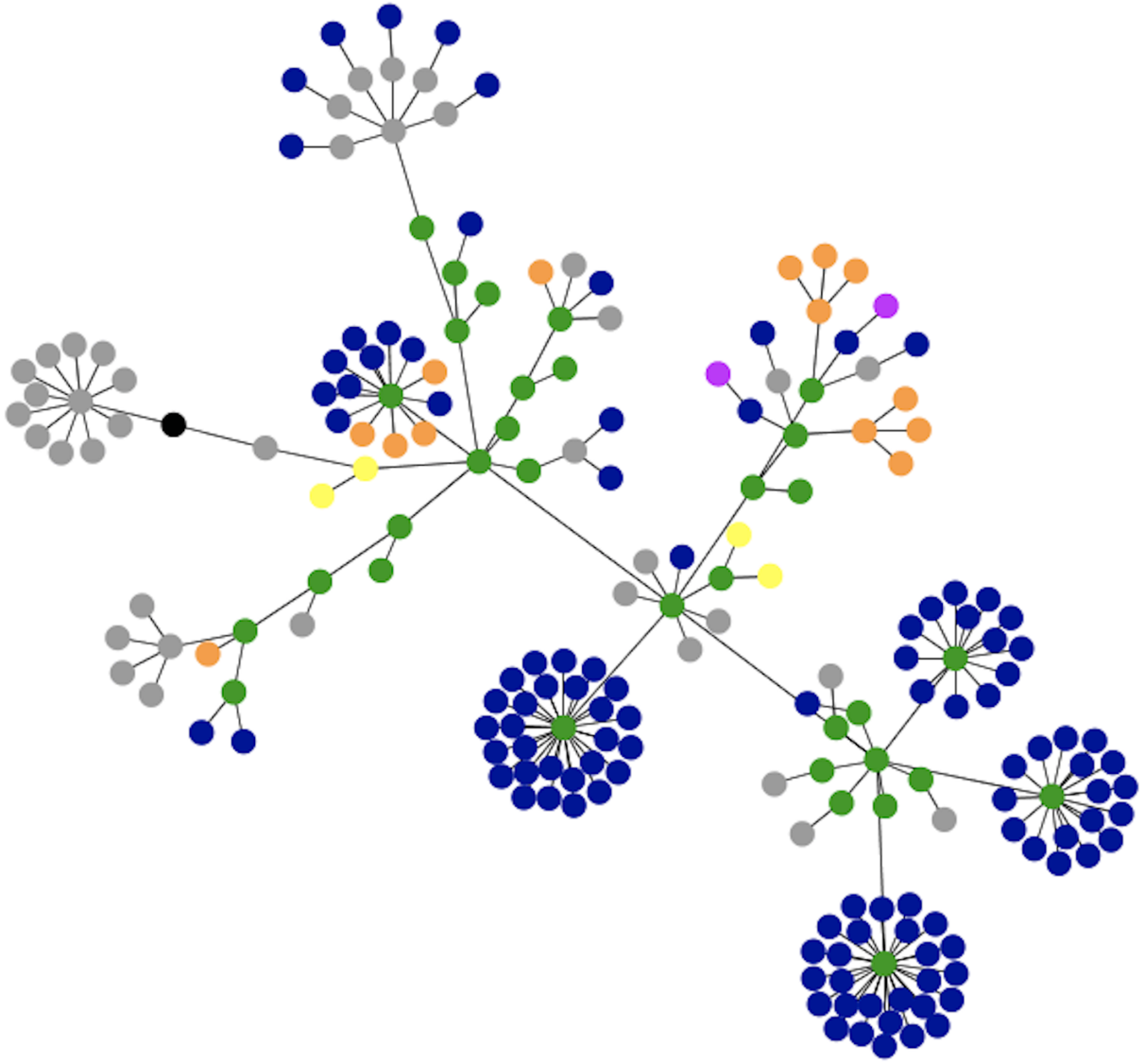
Network analysis in R: A tidy approach

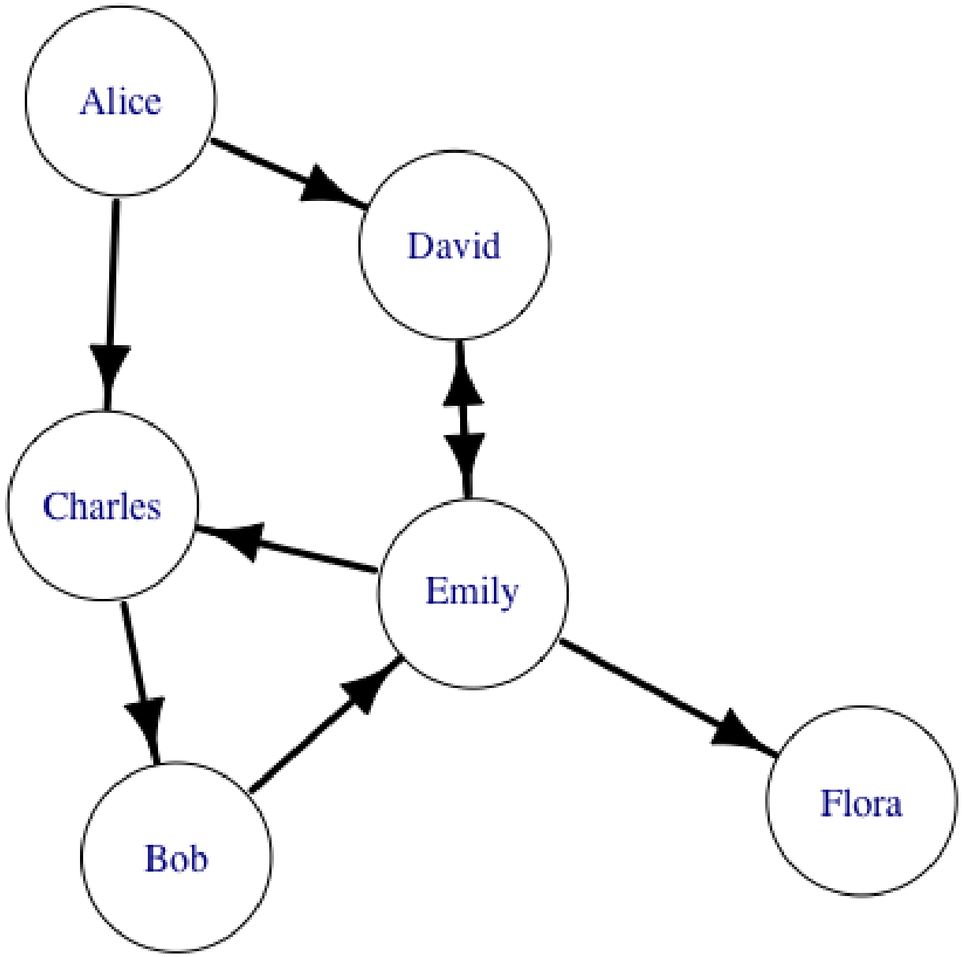
NETWORK ANALYSIS IN THE TIDYVERSE

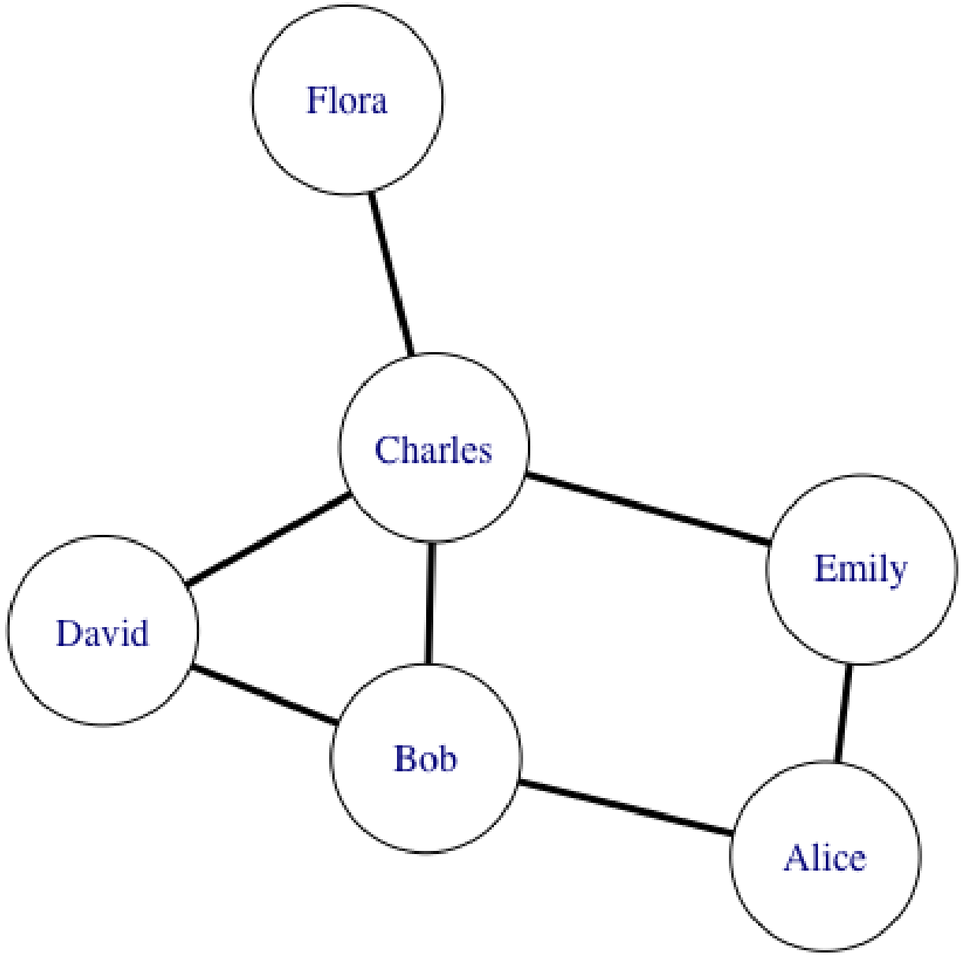
Massimo Franceschet

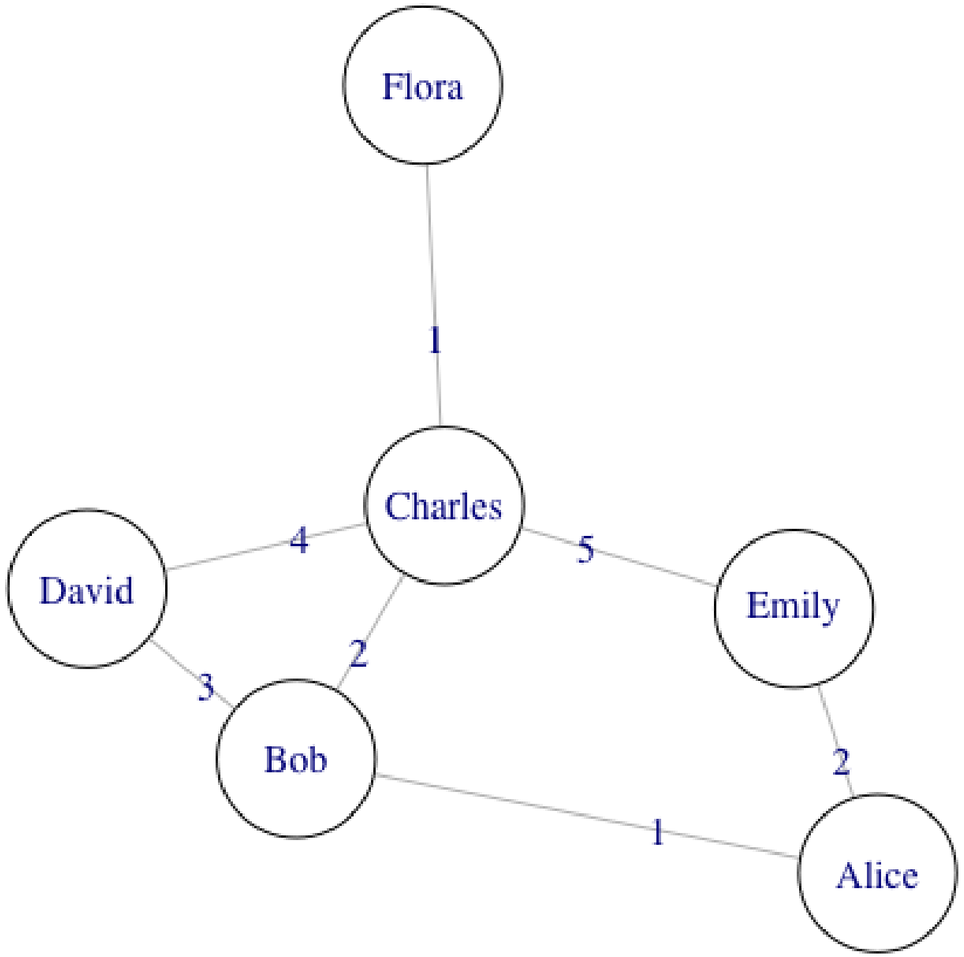
Prof. of Data Science, University of Udine (Italy)

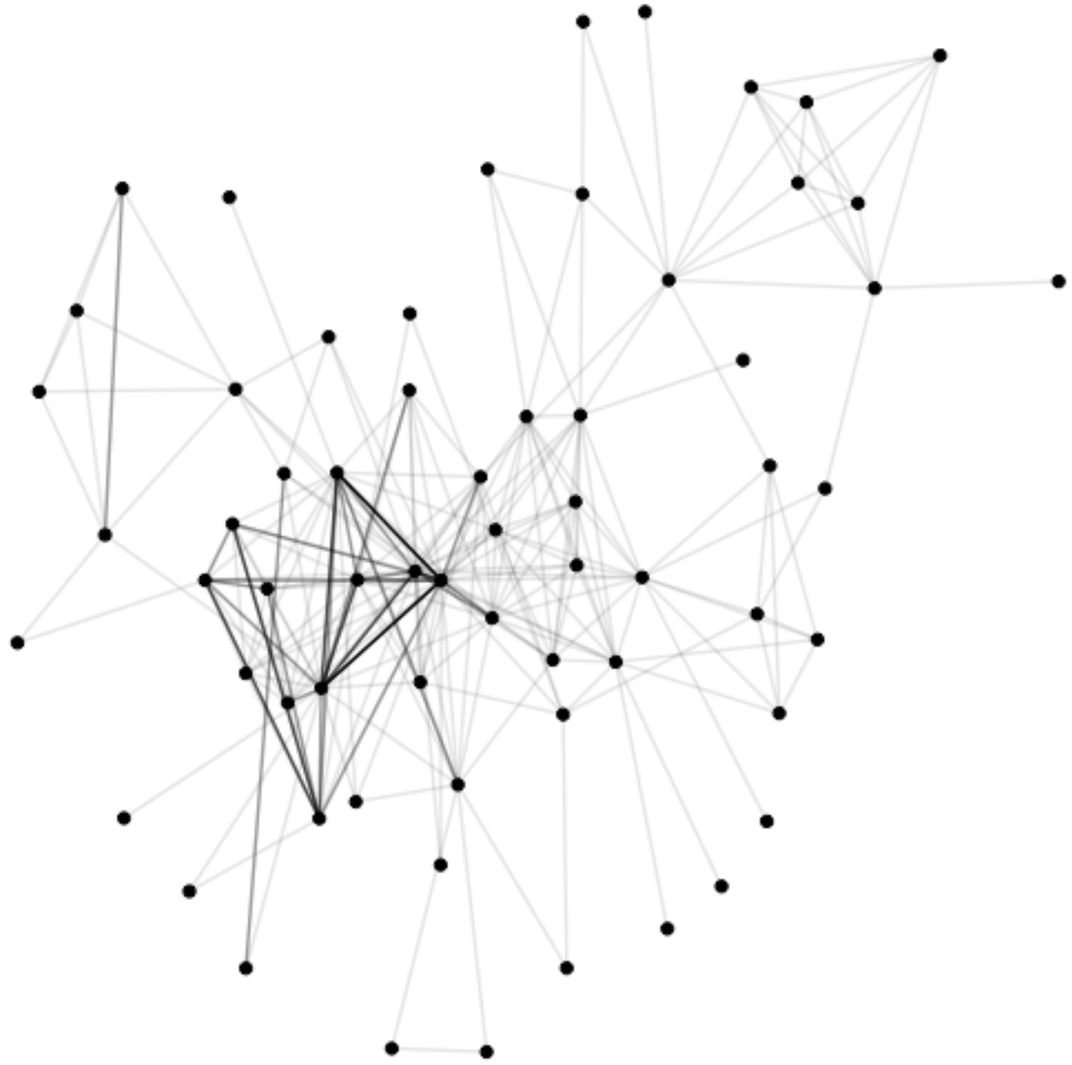
datacamp




Alice
David
Charles
Emily
Bob
Flora


Flora
Charles
Emily
David
Bob
Alice


Flora
1
Charles
4
5
David
Emily
2
3
Bob
2
1
Alice



Building the network

```
# load packages for network exploration
library(readr)
library(igraph)
# read nodes and ties data into variables
nodes <- read_csv("nodes.csv")
ties <- read_csv("ties.csv")

# build a network from data frames
g <- graph_from_data_frame(d = ties,
                           directed = FALSE,
                           vertices = nodes)
```

Exploring the network

```
# explore the set of nodes and print the number of nodes
V(g)
vcount(g)

# explore the set of ties and print the number of ties
E(g)
ecount(g)
# add the name attribute "Madrid network" to the network and print it
g$name <- "Madrid network"
g$name
# add node attribute id and print the node `id` attribute
V(g)$id <- 1:vcount(g)

# print the tie `weight` attribute
E(g)$weight
```

Let's start the investigation!

NETWORK ANALYSIS IN THE TIDYVERSE

Visualizing networks

NETWORK ANALYSIS IN THE TIDYVERSE

Massimo Franceschet

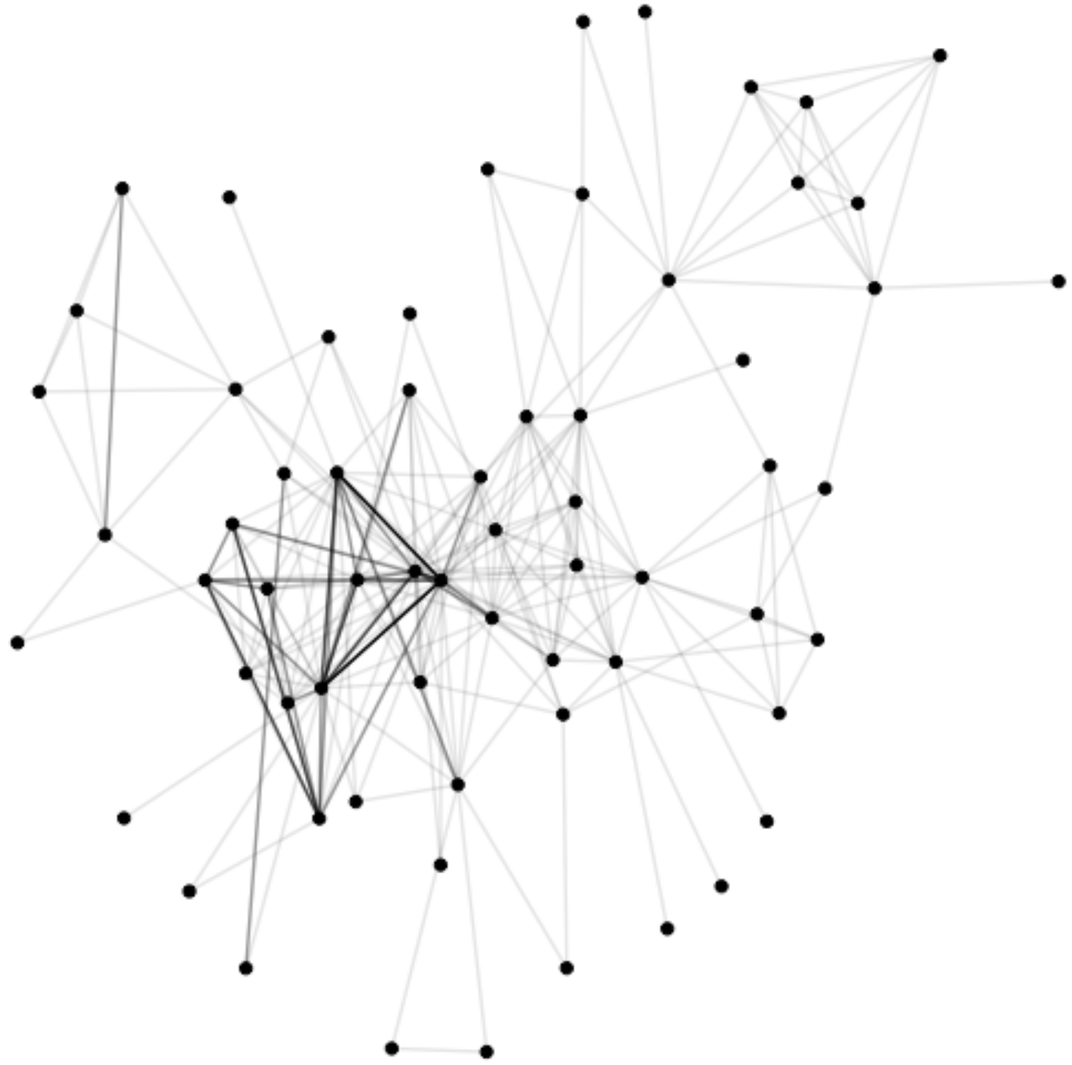
Prof. of Data Science, University of Udine (Italy)

datacamp

ggraph()

```
# load packages for data manipulation and visualization
library{igraph)
library(dplyr)
library(ggplot2)
library(ggraph)
```

```
# visualize the network
ggraph(g, layout = "with_kk") +
    geom_edge_link(aes(alpha = weight)) +
    geom_node_point()
```



Let's practice!

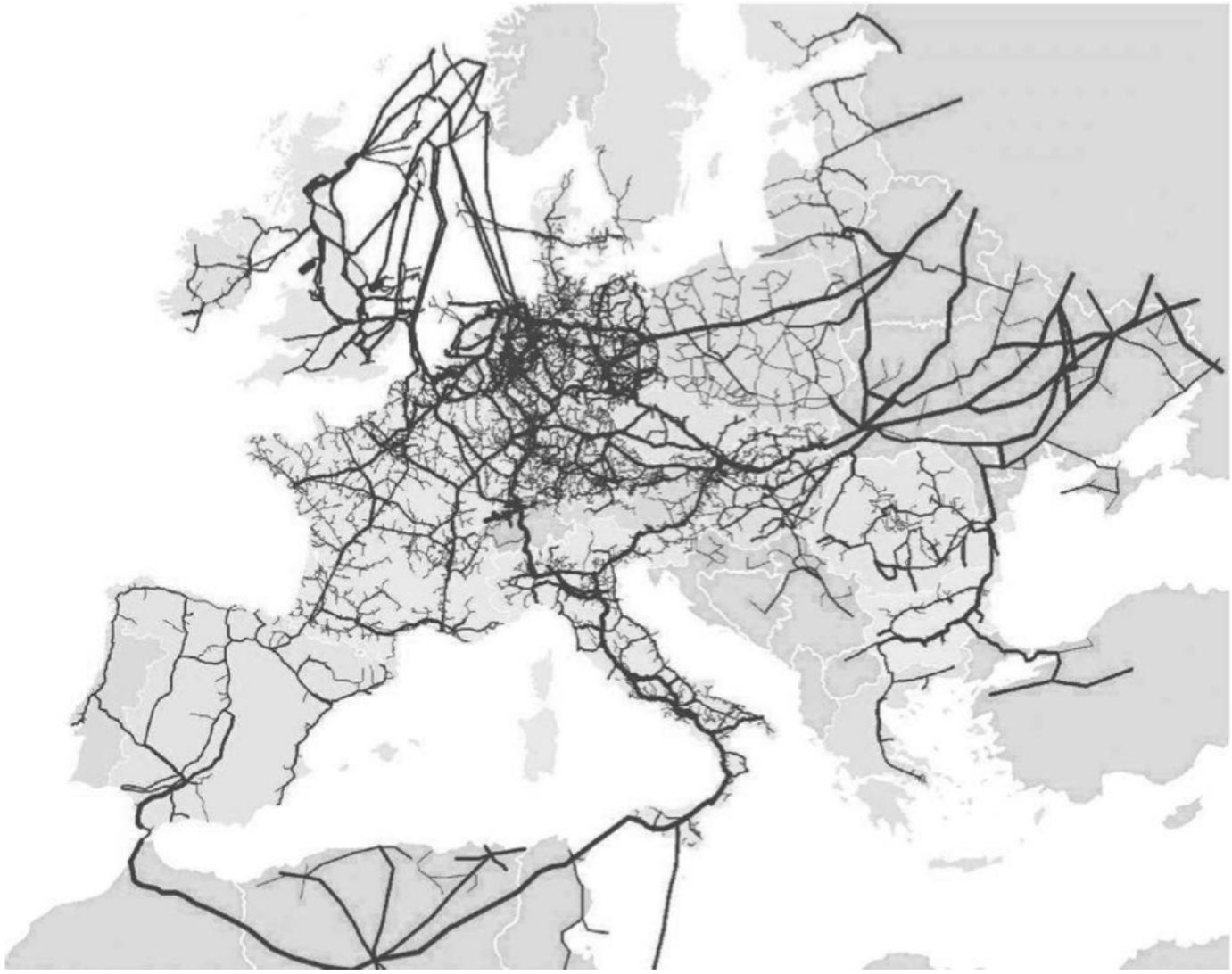
NETWORK ANALYSIS IN THE TIDYVERSE

Centrality measures

NETWORK ANALYSIS IN THE TIDYVERSE

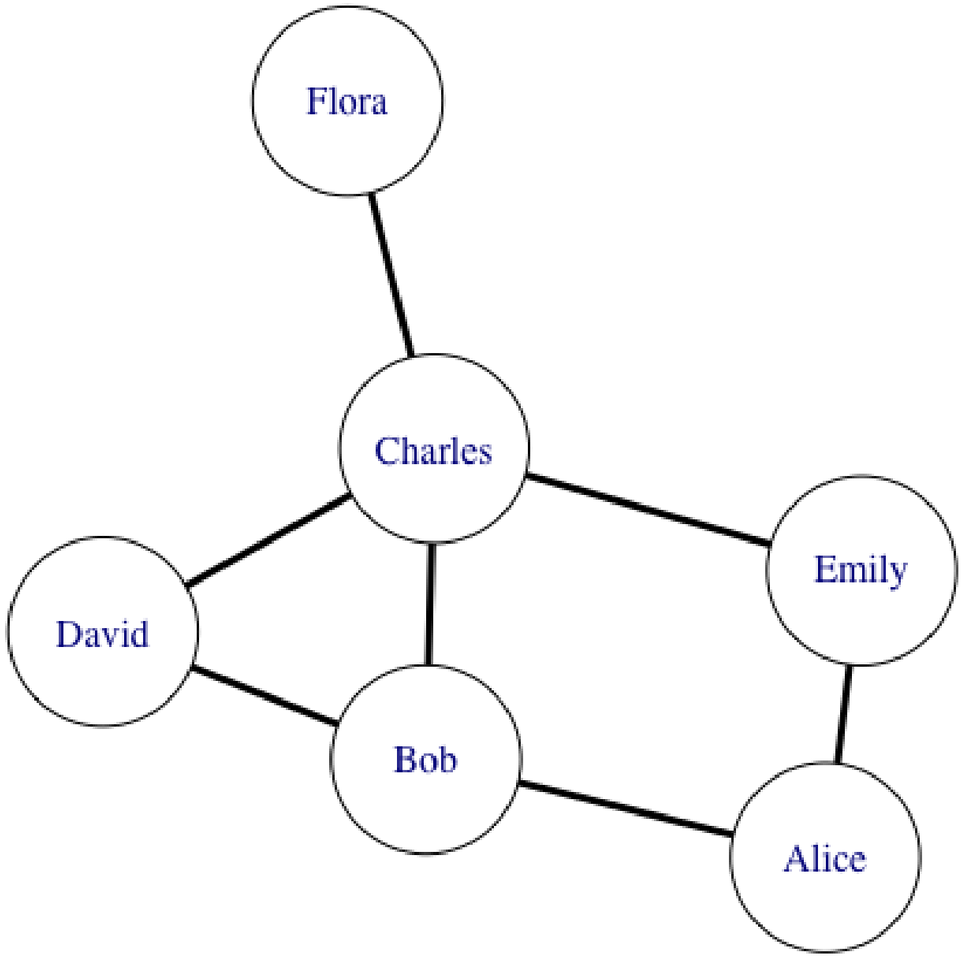
Massimo Franceschet

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



Node centrality

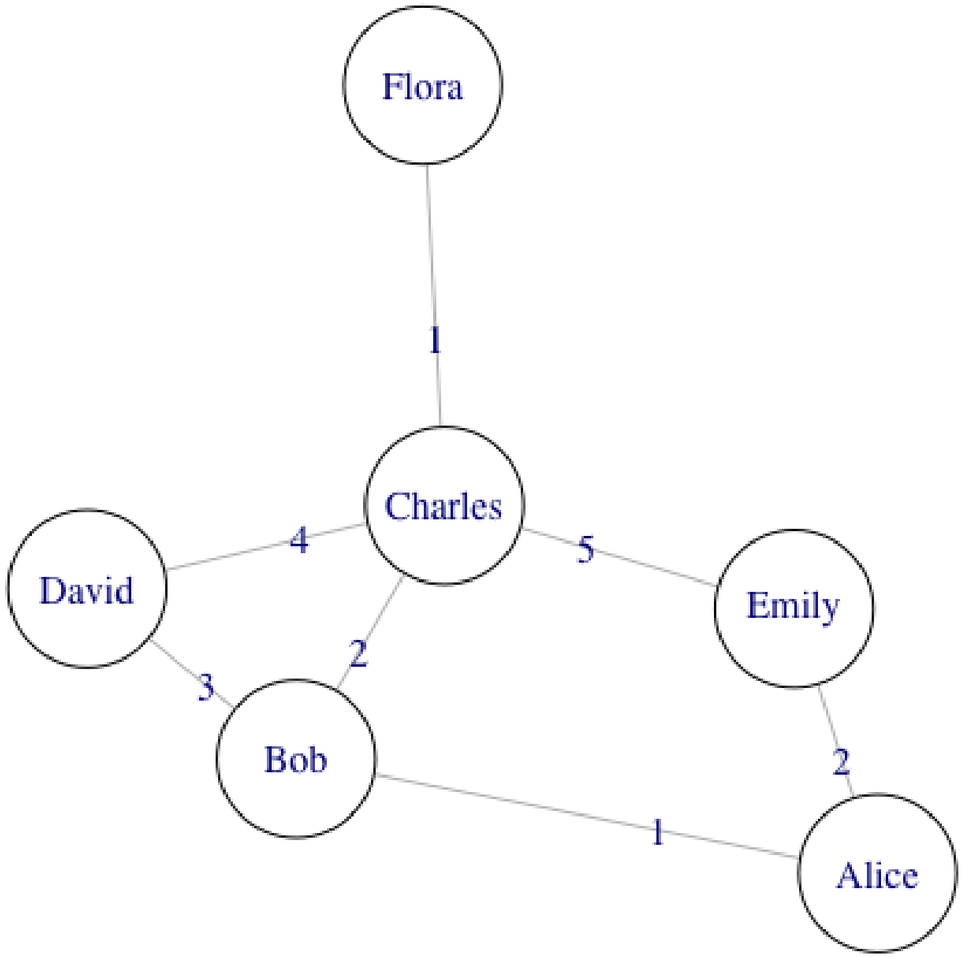
- Which are the most important nodes in a network?
  - Important web pages about a certain topic
  - Influential academic papers covering a given issue
  - Internet routers whose failure would greatly affect network connectivity


Flora
Charles
Emily
David
Bob
Alice

Computing degree

```
# compute node degrees
degree(g)
```

```
        Jamal Zougam       Mohamed Bekkali        Mohamed Chaoui
                  29                     2                    27
         Vinay Kholy         Suresh Kumar        Mohamed Chedadi
                  10                    10                     7
Imad Eddin Barakat   Abdelaziz Benyaich        Abu Abderrahame
                  22                     6                     4
     Omar Dhegayes          Amer Azizi     Abu Musad Alsakaoui
                   2                    18                    10
        Mohamed Atta      Ramzi Binalshibh        Mohamed Belfatmi
                  10                    10                    11
         Said Bahaji         Galeb Kalaje        Abderrahim Zbakh
                  11                    16                    15
```


Flora
1
Charles
4
5
David
Emily
2
3
Bob
2
1
Alice

Computing strength

```
# compute node strengths
strength(g)
```

```
      Jamal Zougam       Mohamed Bekkali        Mohamed Chaoui
                43                     2                    34
       Vinay Kholy          Suresh Kumar       Mohamed Chedadi
                10                    10                     7
Imad Eddin Barakat    Abdelaziz Benyaich       Abu Abderrahame
                35                     7                     4
     Omar Dhegayes           Amer Azizi   Abu Musad Alsakaoui
                 3                    27                    10
      Mohamed Atta       Ramzi Binalshibh      Mohamed Belfatmi
                12                    14                    19
       Said Bahaji          Galeb Kalaje      Abderrahim Zbakh
                17                    21                    15
```

Let's find the most central terrorists in the network!

NETWORK ANALYSIS IN THE TIDYVERSE