Lab 10 - Network data

Ariane Ducellier

2024-11-21

Load R packages

```
library(igraph)
```

```
##
## Attaching package: 'igraph'
```

```
## The following objects are masked from 'package:stats':
##
##     decompose, spectrum
```

```
## The following object is masked from 'package:base':
##
##     union
```

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x lubridate::%--%()      masks igraph::%--%()
## x dplyr::as_data_frame() masks tibble::as_data_frame(), igraph::as_data_frame()
## x purrr::compose()       masks igraph::compose()
## x tidyr::crossing()      masks igraph::crossing()
## x dplyr::filter()        masks stats::filter()
## x dplyr::lag()           masks stats::lag()
## x purrr::simplify()      masks igraph::simplify()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

Load the data

Biggest US airports

```
airports <- c("ATL", "DFW", "DEN", "ORD", "LAX", "JFK", "LAS", "MCO", "MIA",
              "CLT", "SEA", "PHX", "EWR", "SFO", "IAH", "BOS", "FLL", "MSP",
              "LGA", "DTW")
```

Flights data set

```
flights <- read_csv("../../data/routes.csv")
```

```
## Rows: 67663 Columns: 9
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (8): airline, airline ID, source airport, source airport id, destination...
## dbl (1): stops
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
flights <- flights %>%
  filter(`source airport` %in% airports, `destination apirport` %in% airports) %>%
  group_by(`source airport`, `destination apirport`) %>% summarise(count=n())
```

```
## `summarise()` has grouped output by 'source airport'. You can override using
## the `.groups` argument.
```

Create a network using the airports and the number of flights from and to each airport as edges and vertices.

Make a plot of the network.