# Exercise 1 - LinkedIn Connections

Sheida Majidi

2024-03-18

**Introduction:**

I am analyzing my LinkedIn connections to understand my network better, focusing on the distribution of connections across different employers and the network's structure.

Data loading:

Starting with loading my LinkedIn connections data and perform initial cleaning and examination of the data.

```
file_path = "/Users/sheidamajidi/Desktop/Winter2024/Winter2024-
2/ORGB672/Exercises/1/Connections.csv"
connections = read_csv(file_path)
```

```
## Rows: 1413 Columns: 7
## ── Column specification
────────────────────────────────────────────────────────
## Delimiter: ","
## chr (7): First Name, Last Name, URL, Email Address, Company, Position,
Conne...
##
## 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.
```

```
head(connections)
```

```
## # A tibble: 6 × 7
##   `First Name` `Last Name` URL   `Email Address` Company Position
`Connected On`
##   <chr>        <chr>       <chr> <chr>           <chr>   <chr>    <chr>
## 1 Jeremy       Huang       http… <NA>            Alpine… Researc… 15-Mar-
24
## 2 Ali          Gorji       http… <NA>            iptiQ … Data En… 14-Mar-
24
## 3 Nima         Akbarzadeh  http… <NA>            Mila -… Postdoc… 14-Mar-
24
## 4 Wenjie       Zhan        http… <NA>            Desaut… Equity … 14-Mar-
24
## 5 Cynthia      Dugal       http… <NA>            IVADO … Executi… 13-Mar-
24
## 6 Alara        Buyukkorog… http… <NA>            KPMG C… Consult… 13-Mar-
24
```

**Data Analysis**

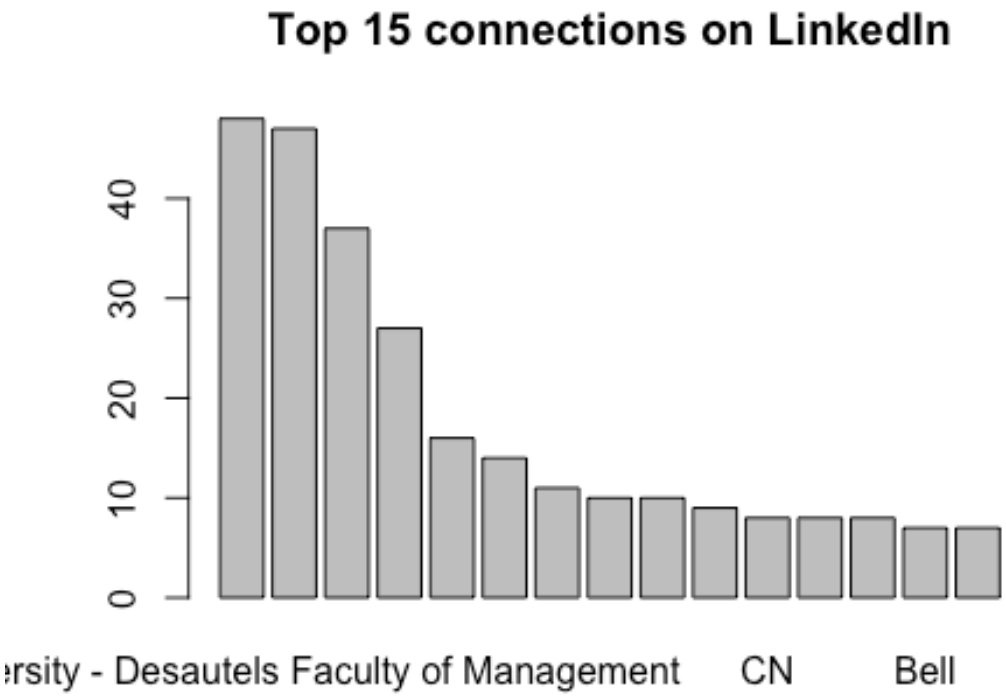
This is the analysis to count the number of contacts by their current employer and calculate the total number of contacts.

```
## # A tibble: 892 × 2
##    Company                                    Count
##    <chr>                                      <int>
##  1 360insights.com                                1
##  2 3CS Customer Centralized Customs Services      1
##  3 3SR Grenoble                                   1
##  4 52 Capital                                     1
##  5 5Y Capital                                     1
##  6 A & A Industries Inc.                          1
##  7 AAVAA                                          1
##  8 ABB                                            1
##  9 ABRPPVM                                        1
## 10 ACENSI                                         1
## # ℹ 882 more rows
```

Frequency table and bar chart

For a better visualizations, I want to create a frequency table of the companies and visualize the top 15 companies with the most connections

| Company | Connections |
|---|---|
| McGill University - Desautels Faculty of Management | 48 |
| McGill University | 47 |
| Desautels Capital Management | 37 |
| IVADO Labs | 27 |
| Sharif University of Technology | 16 |
| TD | 14 |
| Pratt & Whitney Canada | 11 |
| BOMBARDIER | 10 |
| RBC | 10 |
| BNP Paribas | 9 |
| CN | 8 |
| Deloitte | 8 |
| Thinkr Consulting | 8 |
| Bell | 7 |
| KPI Digital Solutions | 7 |




**Network creation**

Now I want to create nodes and edges for the network analysis. In this network, individuals are nodes and connections between individuals who work at the same company are edges.

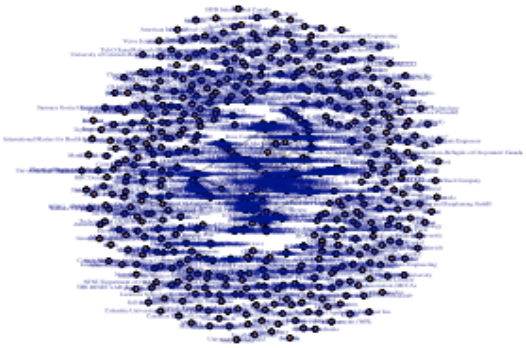


*Figure 1. My network visualized by igraph library tool*

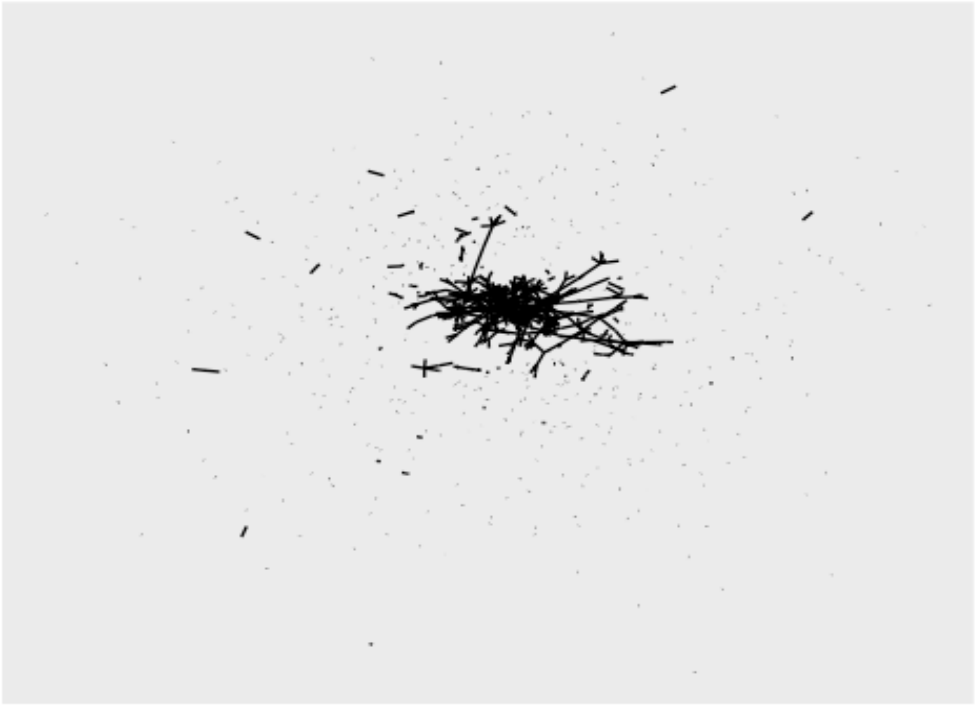


*Figure 2 My network visualized by ggraph tool*

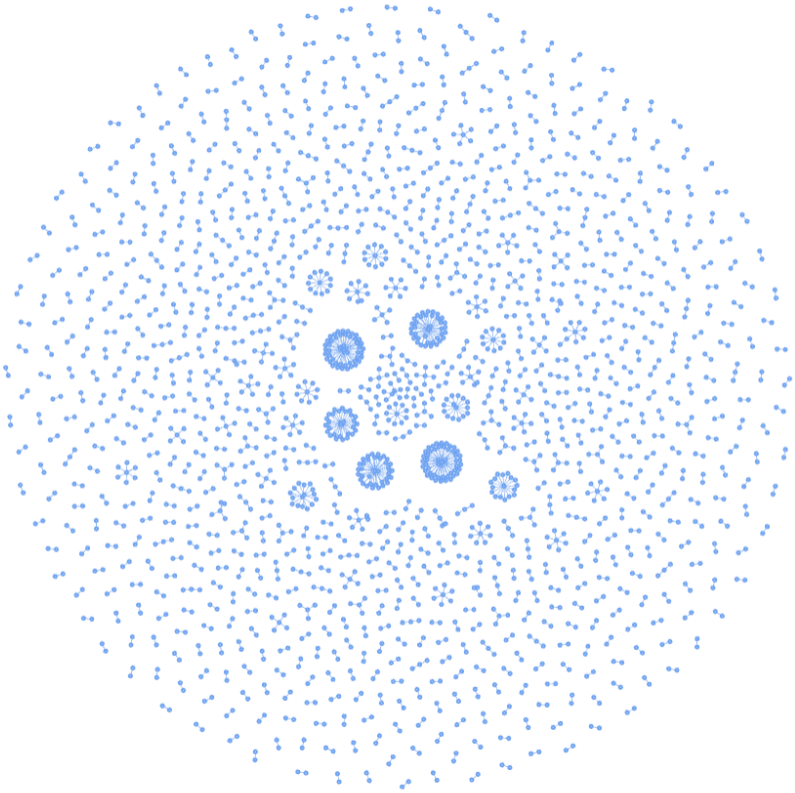


*Figure 3 My network visualized by vizNetwork library tool*

**My Classmates and I**

Below shows my classmates and I are in gray, while our shared company "McGill University – Desautels Faculty of Management" is in red for a better visualization.

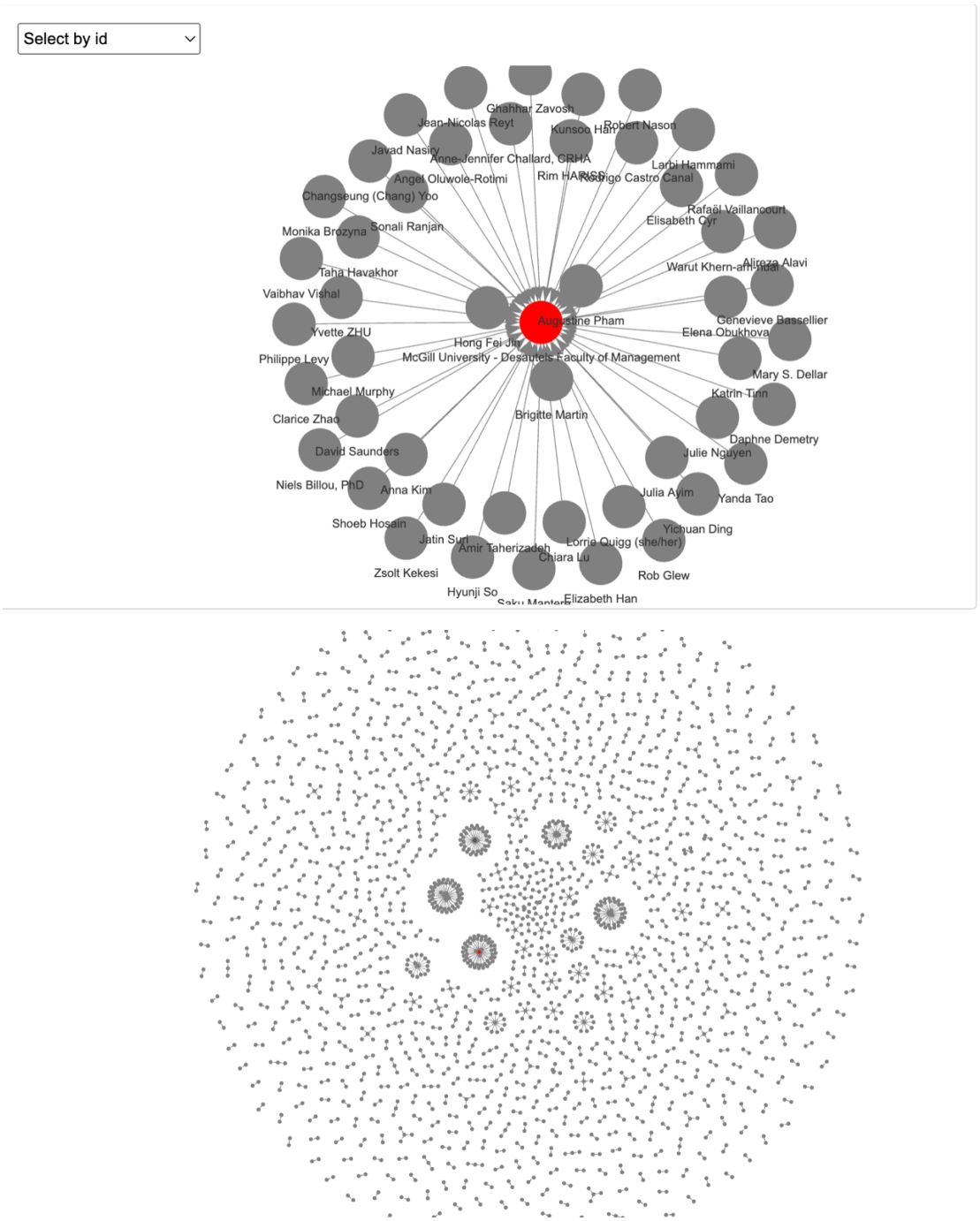



*Figure 4 My classmates and I as a network within my LinkedIn network, showing in Red for our shared company "McGill University- Desautels Faculty of Management" and nodes are Grey as us*