

Exercise 1

Code

```
library("tidyverse")
library("igraph")
csv = read_csv('Connections.csv')
csv
View(csv)
attach(csv)
csv %>% drop_na(Company)

Counting

count = csv %>% count(Company, sort=TRUE)
count

# A tibble: 220 × 2
  Company                                n
  <chr>                                <int>
1 BrainBox AI                          37
2 McGill University - Desautels Facult... 13
3 EY                                    6
4 Nuance Communications                 6
5 Manulife                             5
6 Air Transat                           4
7 Autodesk                             4
8 Novartis                             4
9 Rogers Communications                 4
10 Accenture                            3
# ... with 210 more rows

csv$last_initial <- substr(csv$`Last Name`, 1, 1)
csv$node_label <- paste(csv$`First Name`, csv$last_initial)
nodes <- csv %>% distinct(node_label)
View(nodes)
nodes <- nodes %>% rowid_to_column('id')
nodes
copy <- csv
colnames(copy) <- paste(colnames(copy), "2", sep="_")
cross <- tidyr::crossing(csv, copy, .name_repair="minimal")
edges <- filter(cross, cross$Company == cross$Company_2 & cross$node_label != cross$node_label_2)
edges <- edges %>% select(node_label, Company, node_label_2, Company_2)
edges <- edges %>%
```

```

left_join(nodes, by = c("node_label" = "node_label")) %>%
  rename(node_1 = id)
edges <- edges %>%
  left_join(nodes, by = c("node_label_2" = "node_label")) %>%
  rename(node_2 = id)

```

```

edges <- select(edges, node_1, node_2)
edges

```

```

# A tibble: 1,716 × 2

```

	node_1	node_2
	<int>	<int>
1	115	107
2	115	94
3	43	23
4	178	167
5	178	163
6	178	203
7	178	165
8	178	179
9	178	193
10	178	198

```

# ... with 1,706 more rows

```

Graph Network

```

library("tidygraph")
library("ggraph")
network <- tbl_graph(nodes=nodes, edges=edges, directed=FALSE)
network

```

```

# A tbl_graph: 329 nodes and 1716 edges
#
# An undirected multigraph with 217 components
#
# Node Data: 329 x 2 (active)
  id node_label
  <int> <chr>
1     1 Zhi Cheng T
2     2 Nadine H
3     3 Jack L
4     4 Ramona M
5     5 Franck B
6     6 Omer M
# ... with 323 more rows
#
# Edge Data: 1,716 x 2
  from to
  <int> <int>
1   107  115
2    94  115
3    23   43
# ... with 1,713 more rows

```

```

ggraph(network) + geom_edge_link() + geom_node_point() + theme_graph()

```

