

Projeto Final

Cristiana Aparecida Nogueira Couto

06/2020

Ref:

<http://sape.inf.usi.ch/quick-reference/ggplot2/colour>

<http://users.dimi.uniud.it/~massimo.franceschet/ns/syllabus/make/ggraph/ggraph.html>

```
library(ggplot2)
library(outbreaks)
library(tidyverse)
```

```
## -- Attaching packages ----- tidyverse 1.3.0 --
```

```
## v tibble  3.0.1      v dplyr    0.8.5
## v tidyr   1.0.2      v stringr 1.4.0
## v readr   1.3.1      v forcats 0.5.0
## v purrr   0.3.4
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggraph)
library(tidygraph)
```

```
##
## Attaching package: 'tidygraph'

## The following object is masked from 'package:stats':
##
##     filter
```

```
data("measles_hagelloch_1861")
```

```
head(measles_hagelloch_1861)
```

```
##   case_ID infector date_of_prodrôme date_of_rash date_of_death age gender
## 1      1      45   1861-11-21   1861-11-25      <NA>  7      f
## 2      2      45   1861-11-23   1861-11-27      <NA>  6      f
## 3      3     172   1861-11-28   1861-12-02      <NA>  4      f
```

```
## 4      4      180      1861-11-27 1861-11-28      <NA> 13      m
## 5      5       45      1861-11-22 1861-11-27      <NA>  8      f
## 6      6      180      1861-11-26 1861-11-29      <NA> 12      m
##   family_ID class complications x_loc y_loc
## 1      41     1             yes 142.5 100.0
## 2      41     1             yes 142.5 100.0
## 3      41     0             yes 142.5 100.0
## 4      61     2             yes 165.0 102.5
## 5      42     1             yes 145.0 120.0
## 6      42     2             yes 145.0 120.0
```

```
#Quantidade de pessoas com erupção cutânea reportadas por dia
dates_prondrome <- measles_hagelloch_1861 %>%
  select(case_ID, date_of_prondrome) %>%
  group_by(date_of_prondrome) %>%
  summarise(frequency = n())

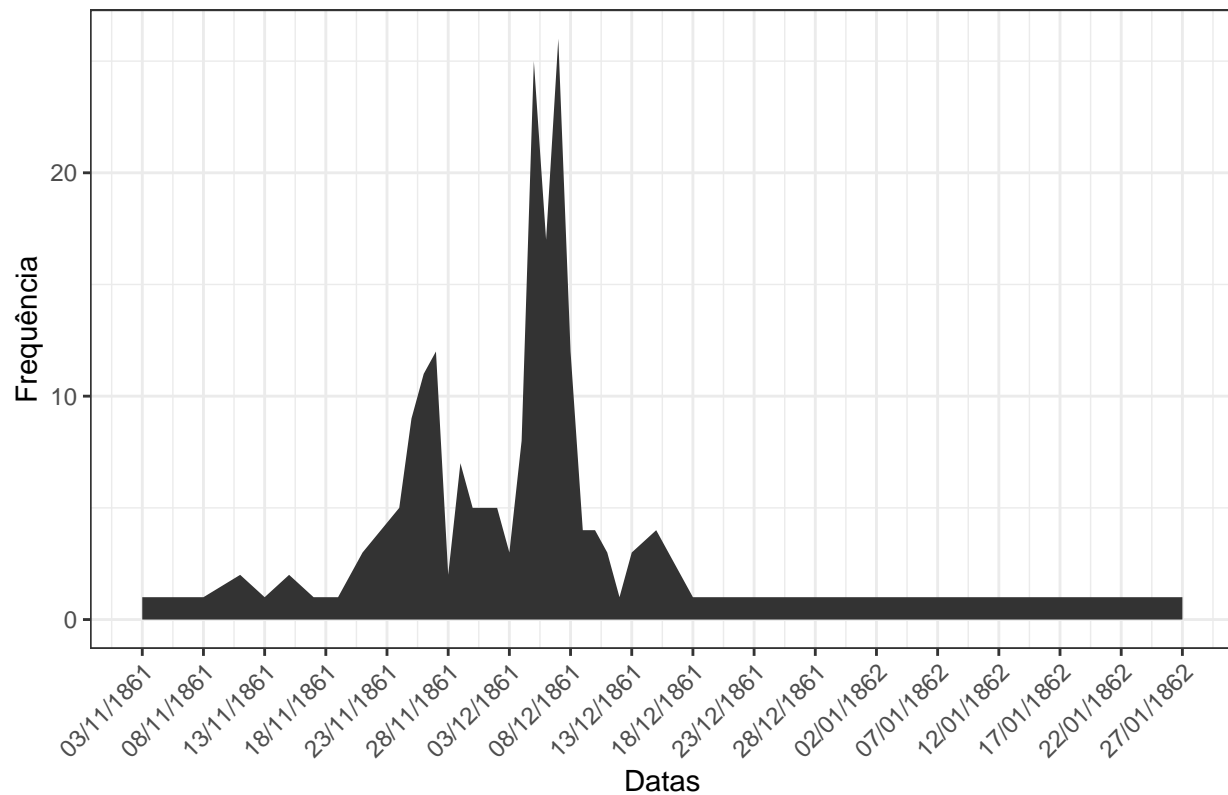
dates_prondrome
```

```
## # A tibble: 36 x 2
##   date_of_prondrome frequency
##   <date>             <int>
## 1 1861-10-30           1
## 2 1861-11-01           1
## 3 1861-11-07           2
## 4 1861-11-08           1
## 5 1861-11-11           2
## 6 1861-11-12           1
## 7 1861-11-13           1
## 8 1861-11-15           2
## 9 1861-11-17           1
## 10 1861-11-18          1
## # ... with 26 more rows
```

```
dates_rash <- measles_hagelloch_1861 %>%
  select(case_ID, date_of_rash) %>%
  group_by(date_of_rash) %>%
  summarise(frequency = n())

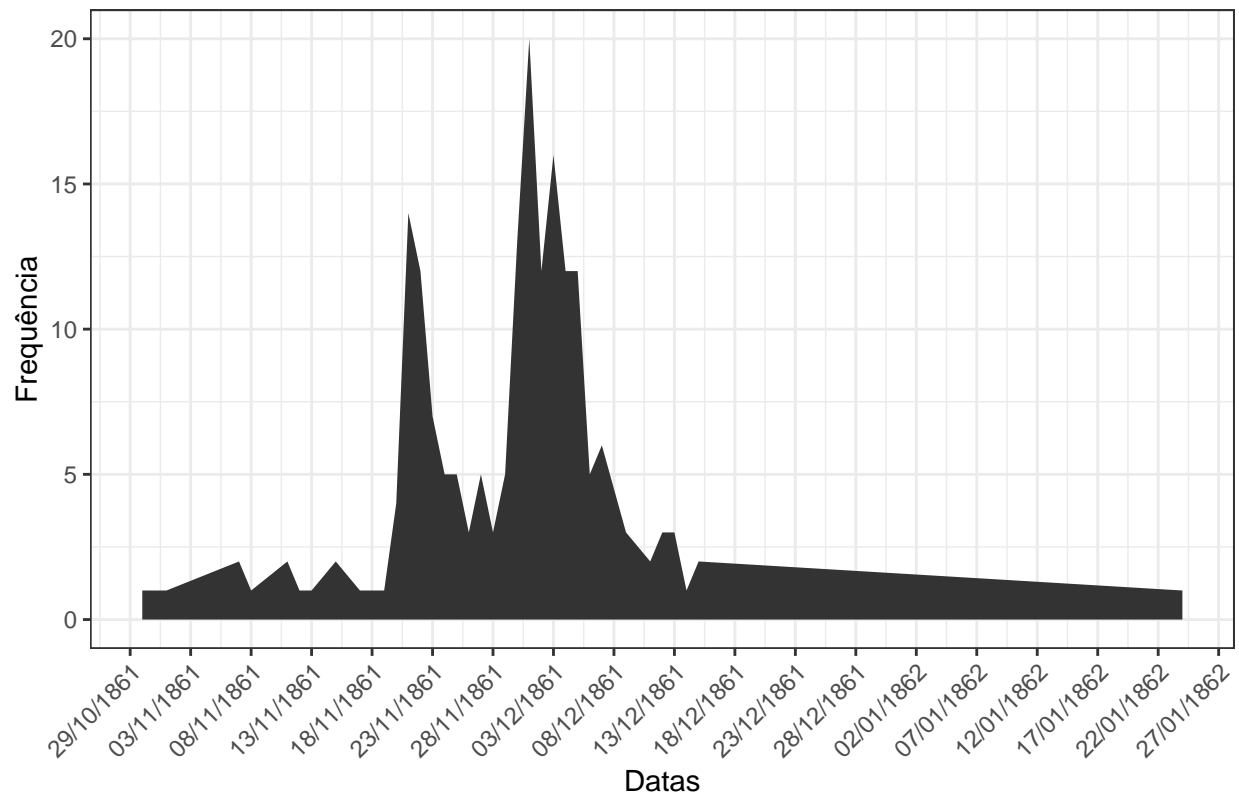
ggplot(dates_rash) +
  geom_area(aes(x = date_of_rash, y = frequency)) +
  theme_bw() +
  theme(axis.text.x = element_text(angle = 45, hjust = 1),
        plot.title = element_text(hjust = 0.5)) +
  scale_x_date(date_labels = "%d/%m/%Y", date_breaks = "5 day") +
  labs(x = "Datas", y = "Frequência",
       title = "Pessoas reportadas com erupção cutânea causada por sarampo")
```

Pessoas reportadas com erupção cutânea causada por sarampo



```
ggplot(dates_prondrome) +  
  geom_area(aes(x = date_of_prondrome, y = frequency)) +  
  theme_bw() +  
  theme(axis.text.x = element_text(angle = 45, hjust = 1),  
        plot.title = element_text(hjust = 0.5)) +  
  scale_x_date(date_labels = "%d/%m/%Y", date_breaks = "5 day") +  
  labs(x = "Datas", y = "Frequência",  
       title = "Pessoas reportadas com sinais de sarampo")
```

Pessoas reportadas com sinais de sarampo



#Para quantas pessoas cada indivíduo infectado transmitiu a doença?

```
infectior <- table(measles_hagelloch_1861$infectior)
infectior <- data.frame(Indivíduo = names(infectior),
                        Infectados_por_ele = as.vector(infectior))
```

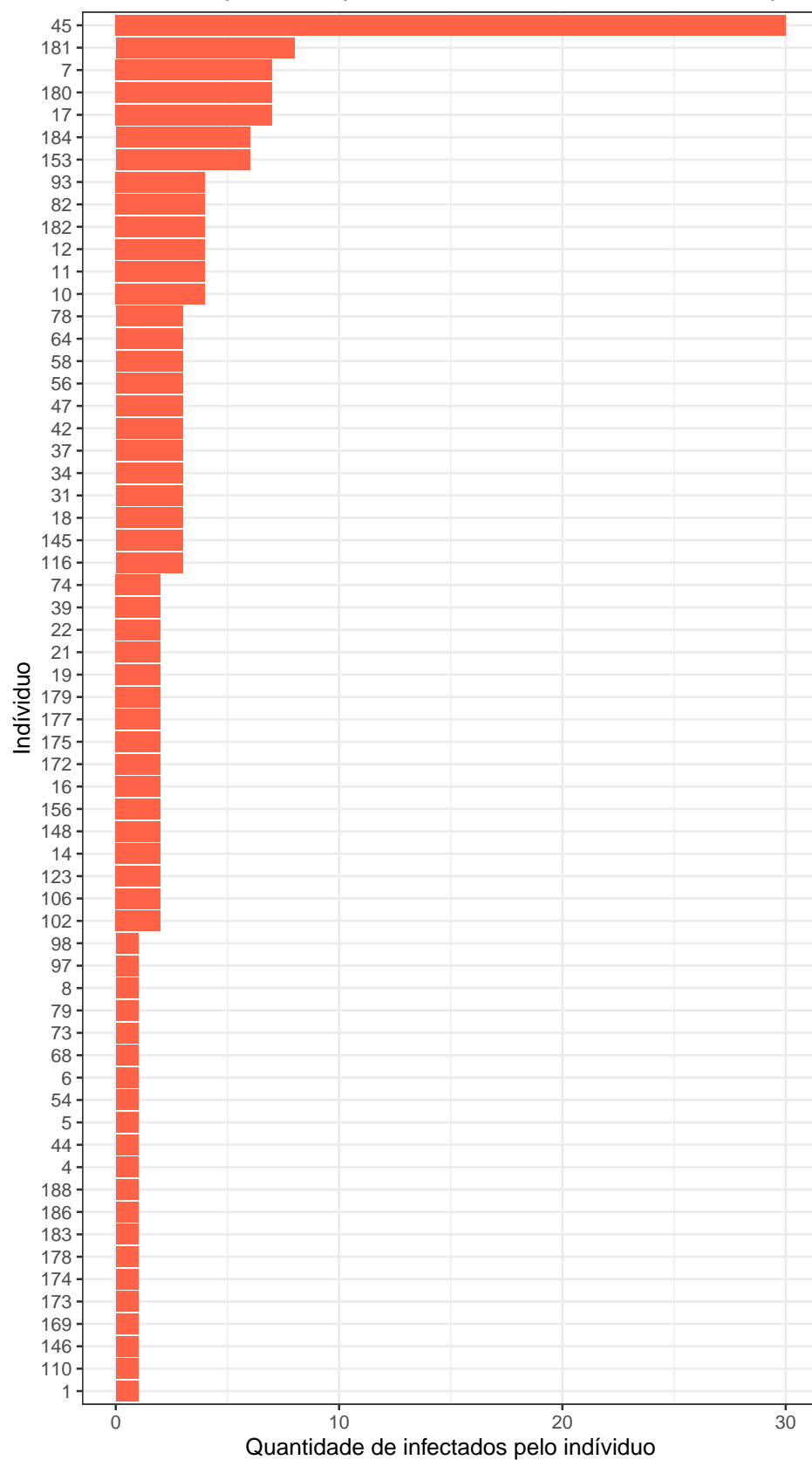
```
infectior
```

##	Indivíduo	Infectados_por_ele
## 1	1	1
## 2	4	1
## 3	5	1
## 4	6	1
## 5	7	7
## 6	8	1
## 7	10	4
## 8	11	4
## 9	12	4
## 10	14	2
## 11	16	2
## 12	17	7
## 13	18	3
## 14	19	2
## 15	21	2
## 16	22	2
## 17	31	3
## 18	34	3
## 19	37	3

## 20	39	2
## 21	42	3
## 22	44	1
## 23	45	30
## 24	47	3
## 25	54	1
## 26	56	3
## 27	58	3
## 28	64	3
## 29	68	1
## 30	73	1
## 31	74	2
## 32	78	3
## 33	79	1
## 34	82	4
## 35	93	4
## 36	97	1
## 37	98	1
## 38	102	2
## 39	106	2
## 40	110	1
## 41	116	3
## 42	123	2
## 43	145	3
## 44	146	1
## 45	148	2
## 46	153	6
## 47	156	2
## 48	169	1
## 49	172	2
## 50	173	1
## 51	174	1
## 52	175	2
## 53	177	2
## 54	178	1
## 55	179	2
## 56	180	7
## 57	181	8
## 58	182	4
## 59	183	1
## 60	184	6
## 61	186	1
## 62	188	1

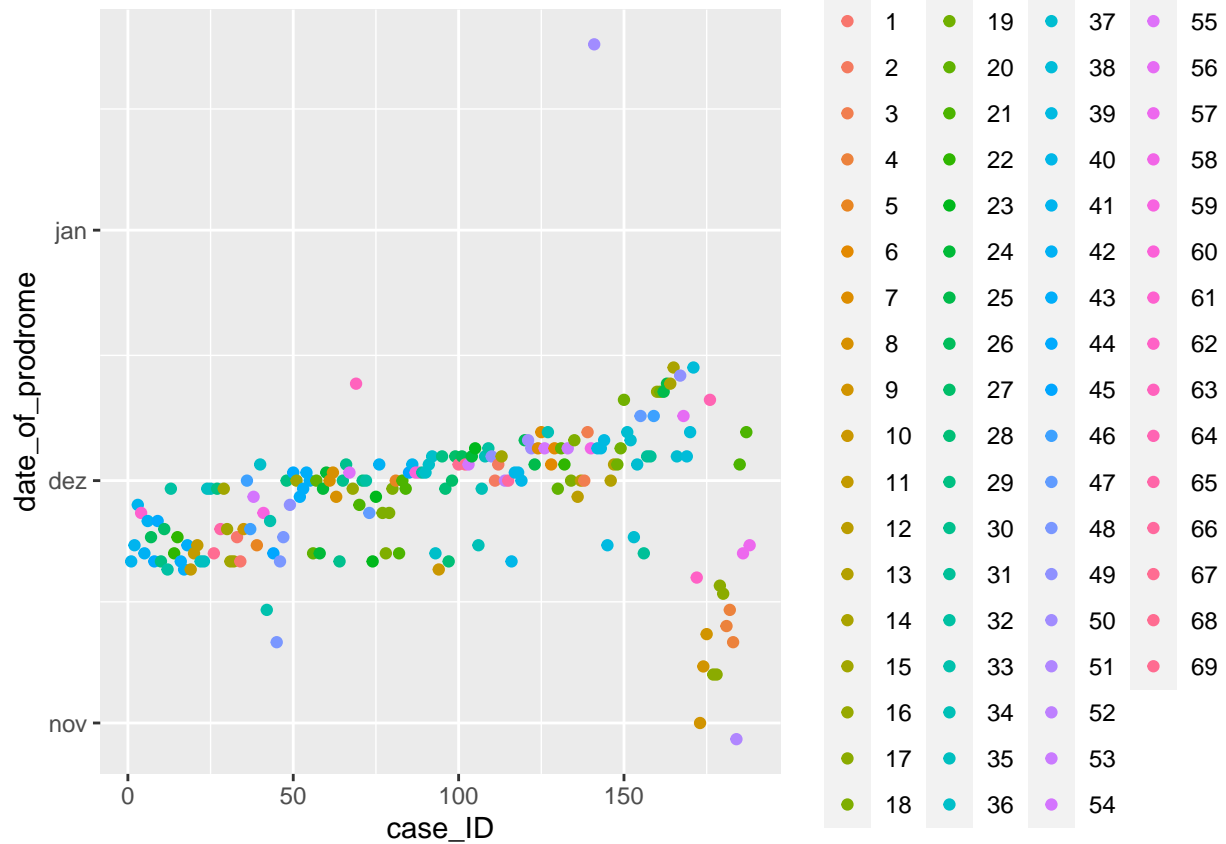
```
#Melhorar labels
ggplot(infeccion) +
  geom_col(aes(x = reorder(Indivíduo, Infectados_por_ele),
                    y = Infectados_por_ele), fill = "tomato1") +
  theme_bw() +
  coord_flip() +
  labs(y = "Quantidade de infectados pelo indivíduo", x = "Indivíduo",
        title = "Número de pessoas que cada indivíduo transmitiu sarampo")
```

Número de pessoas que cada indivíduo transmitiu sarampo



#Quantas famílias foram afetadas?

```
ggplot(measles_hagelloch_1861) +  
  geom_point(aes(x = case_ID, y = date_of_prodrone, color = factor(family_ID)))
```



#Encontrando quem é a família do infector

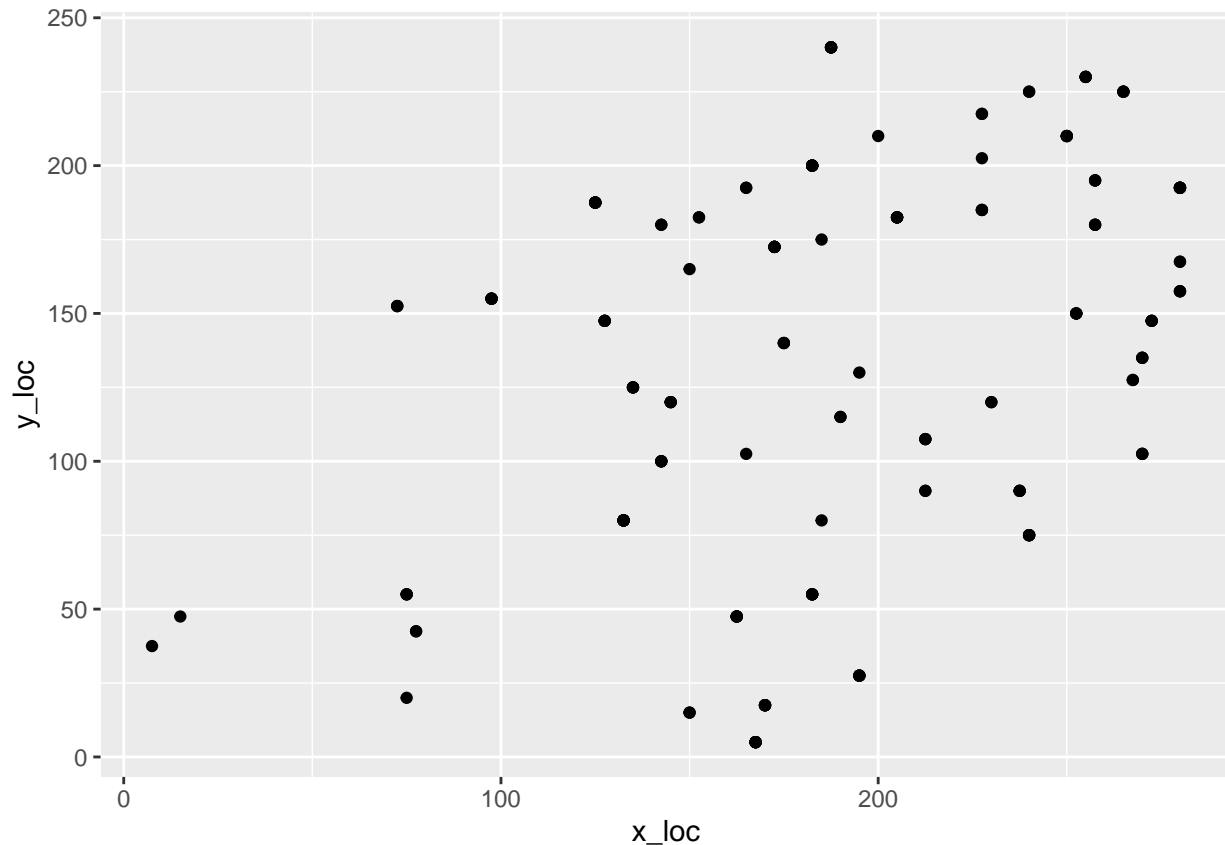
```
infecções_fora_da_familia <- cbind(Family_infeutor= measles_hagelloch_1861$family_ID[measles_hagelloch_1861$case_ID == case_ID],  
  Family_person = measles_hagelloch_1861$family_ID)  
  
head(infecções_fora_da_familia)
```

```
##      Family_infeutor Family_person  
## [1,]             48             41  
## [2,]             48             41  
## [3,]             62             41  
## [4,]             17             61  
## [5,]             48             42  
## [6,]             17             42
```

#Os indivíduos que estão na mesma família tem a mesma localização

#Há algo que posso acrescentar na visualização das posições?

```
ggplot(measles_hagelloch_1861) +  
  geom_point(aes(x = x_loc, y = y_loc))
```



```
#ref:  
#http://www.hellomister.com.au/data-blog/2019/7/30/creating-an-arc-plot-part-1
```

```
nodes <- measles_hagelloch_1861 %>%  
  select(case_ID, class)  
  
infector_edge <- measles_hagelloch_1861 %>%  
  select(infector, case_ID, class, x_loc, y_loc) %>%  
  rename(from = infector) %>%  
  rename(to = case_ID) %>%  
  na.omit()  
  
head(infector_edge)
```

```
##   from to class x_loc y_loc  
## 1   45  1     1 142.5 100.0  
## 2   45  2     1 142.5 100.0  
## 3  172  3     0 142.5 100.0  
## 4  180  4     2 165.0 102.5  
## 5   45  5     1 145.0 120.0  
## 6  180  6     2 145.0 120.0
```

```
#Quem foi infectado por quem?  
infector_tidy <- tbl_graph(edges = infector_edge, directed = TRUE, nodes = nodes)
```

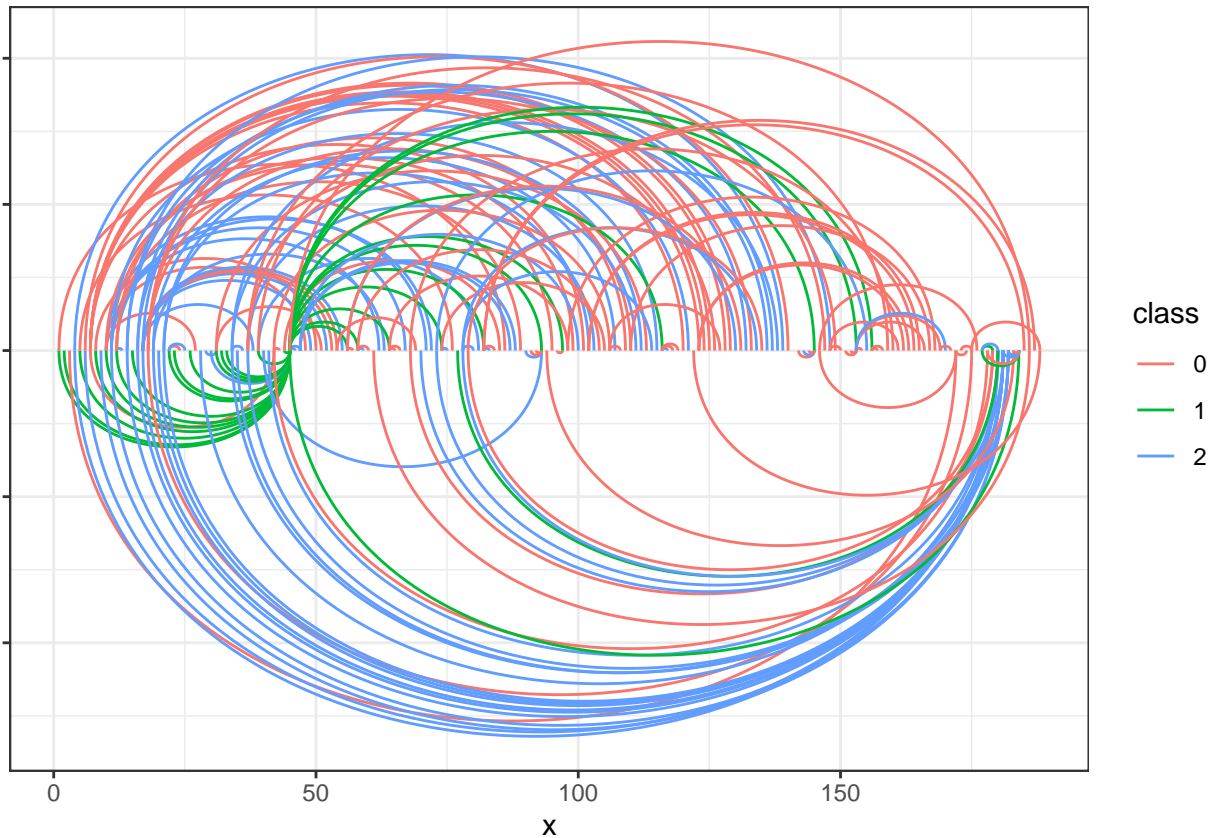


```

infector1 <- ggraph(infector_tidy, layout = 'linear') +
  geom_edge_arc(aes(color = class)) +
  theme_bw() +
  theme(axis.title.y = element_blank(),
        axis.text.y = element_blank())

```

infector1



```

infector_edge45 <- measles_hagelloch_1861 %>%
  select(infector, case_ID, class, x_loc, y_loc) %>%
  rename(from = infector) %>%
  rename(to = case_ID) %>%
  filter(from == 45) %>%
  na.omit()

```

```

infector_tidy45 <- tbl_graph(edges = infector_edge45, directed = TRUE, nodes = nodes)

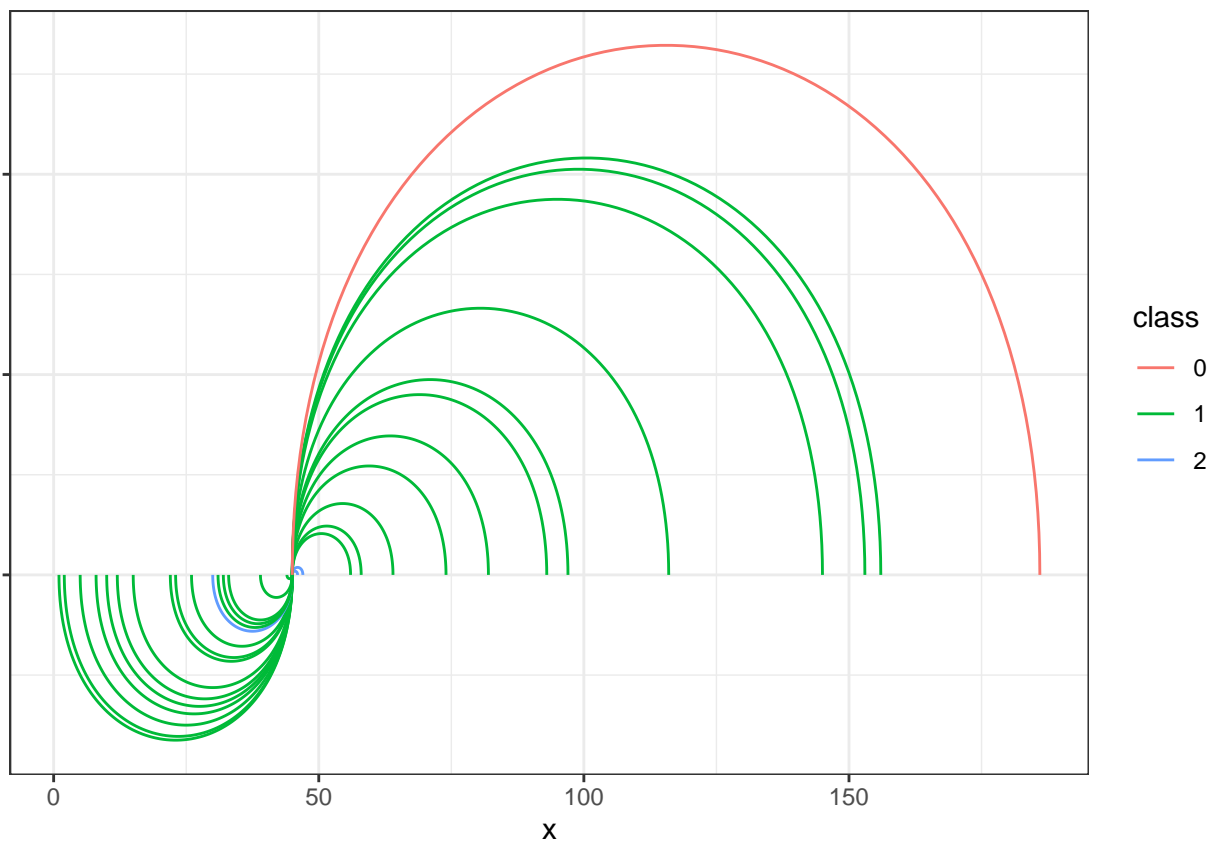
```

```

infector45 <- ggraph(infector_tidy45, layout = 'linear') +
  geom_edge_arc(aes(color = class)) +
  theme_bw() +
  theme(axis.title.y = element_blank(),
        axis.text.y = element_blank())

```

infector45



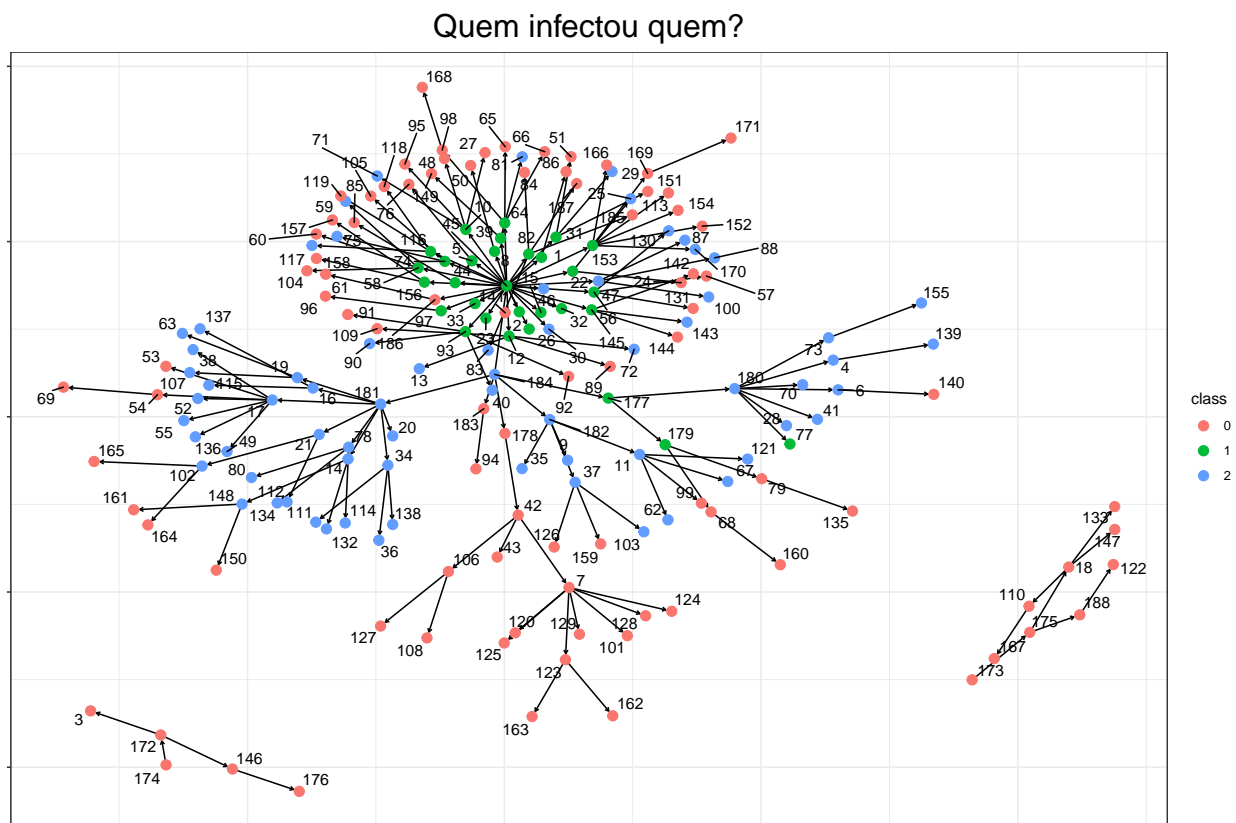
```
#Melhorar o titulo, mudar a paleta de cores, mudar a estetica do class
infectior2 <- ggraph(infectior_tidy, layout = 'kk') +
  geom_edge_link(arrow = arrow(length = unit(1, 'mm')),
    end_cap = circle(1.5, 'mm')) +
  geom_node_point(aes(colour = class), size = 3) +
  geom_node_text(aes(label = case_ID), repel = TRUE) +
  theme_bw() +
  theme(axis.title.y = element_blank(),
    axis.text.y = element_blank(),
    axis.text.x = element_blank(),
    axis.title.x = element_blank(),
    plot.title = element_text(hjust = 0.5, size = 22)) +
  labs(title = "Quem infectou quem?")

infectior_tidy
```

```
## # A tbl_graph: 188 nodes and 184 edges
## #
## # A rooted forest with 4 trees
## #
## # Node Data: 188 x 2 (active)
##   case_ID class
##   <int> <fct>
## 1      1  1
## 2      2  1
## 3      3  0
```

```
## 4      4 2
## 5      5 1
## 6      6 2
## # ... with 182 more rows
## #
## # Edge Data: 184 x 5
##   from   to class x_loc y_loc
##   <int> <int> <fct> <dbl> <dbl>
## 1    45    1 1     142.   100
## 2    45    2 1     142.   100
## 3   172    3 0     142.   100
## # ... with 181 more rows
```

```
infectior2
```



Perguntas

Quando foi o pico de casos?

Para quantas pessoas cada indivíduo infectado transmitiu a doença?

Quem foi infectado por quem?

Existe alguma infecção fora da esfera familiar?

Qual a localização dos infectados?

Quantos dias em média levam da aparição do primeiro sinal até o início da erupção cutânea?

Quanto tempo durou o surto?

Quantas famílias foram afetadas?