





	city	city_bge_code	date	estimated_population_2019	is_repeated	is_last	last_available_confirmed	last_available_confirmed_2
48	Aracati	2301109	2020-04-08	74547	False	False	2	
50	Aracati	2301109	2020-04-16	74547	False	True	2	
51	Aracati	2301109	2020-04-11	74547	False	False	2	
52	Aracati	2301109	2020-04-13	74547	False	False	3	
53	Aracati	2301109	2020-04-15	74547	False	False	3	
54	Aracati	2301109	2020-04-12	74547	False	False	2	
56	Aracati	2301109	2020-04-10	74547	False	False	2	
57	Aracati	2301109	2020-04-14	74547	False	False	3	
58	Aracati	2301109	2020-04-09	74547	False	False	2	
59	Aracati	2301109	2020-04-07	74547	False	False	2	
91	Capistrano	2302909	2020-04-16	17738	False	True	1	
92	Capistrano	2302909	2020-04-15	17738	False	False	1	
93	Capistrano	2302909	2020-04-14	17738	False	False	1	
97	Canis	2303303	2020-04-16	18699	False	True	1	
98	Canis	2303303	2020-04-15	18699	False	False	2	
99	Canis	2303303	2020-04-14	18699	False	False	2	
101	Canis	2303303	2020-04-13	18699	False	False	2	
102	Canis	2303303	2020-04-10	18699	False	False	2	
103	Canis	2303303	2020-04-12	18699	False	False	2	
104	Canis	2303303	2020-04-07	18699	False	False	2	
105	Canis	2303303	2020-04-08	18699	False	False	2	
106	Canis	2303303	2020-04-11	18699	False	False	2	
107	Canis	2303303	2020-04-09	18699	False	False	2	
108	Cascavel	2303501	2020-04-16	71743	False	True	3	
131	Caucaia	2303706	2020-04-16	361400	False	True	56	
132	Caucaia	2303706	2020-04-12	361400	False	False	26	
134	Caucaia	2303706	2020-04-10	361400	False	False	23	
136	Caucaia	2303706	2020-04-13	361400	False	False	31	
142	Caucaia	2303706	2020-04-15	361400	False	False	42	
143	Caucaia	2303706	2020-04-08	361400	False	False	16	
...	...	...	...	...	...	...	...	...
567	Santa Quiteria	2312205	2020-04-04	43703	False	False	1	
568	Santa Quiteria	2312205	2020-03-31	43703	False	False	1	
569	Santa Quiteria	2312205	2020-04-06	43703	False	False	1	
570	Santa Quiteria	2312205	2020-04-08	43703	False	False	2	
571	Santa Quiteria	2312205	2020-04-12	43703	False	False	2	
572	Santa Quiteria	2312205	2020-04-10	43703	False	False	2	
573	Santa Quiteria	2312205	2020-04-15	43703	False	False	3	
574	Santa Quiteria	2312205	2020-04-07	43703	False	False	1	
575	Santa Quiteria	2312205	2020-04-14	43703	False	False	1	
576	Santa Quiteria	2312205	2020-04-03	43703	False	False	1	
577	Santa Quiteria	2312205	2020-04-11	43703	False	False	2	
578	Santa Quiteria	2312205	2020-04-05	43703	False	False	1	
579	Santa Quiteria	2312205	2020-04-16	43703	False	True	3	
580	Santa Quiteria	2312205	2020-04-13	43703	False	False	1	
581	Santa Quiteria	2312205	2020-04-01	43703	False	False	1	
582	Santa Quiteria	2312205	2020-04-09	43703	False	False	2	
632	Tianguá	2313401	2020-04-03	75946	False	False	1	
633	Tianguá	2313401	2020-04-09	75946	False	False	1	
634	Tianguá	2313401	2020-04-13	75946	False	False	3	
635	Tianguá	2313401	2020-04-05	75946	False	False	1	
636	Tianguá	2313401	2020-04-10	75946	False	False	2	
638	Tianguá	2313401	2020-04-07	75946	False	False	1	
639	Tianguá	2313401	2020-04-04	75946	False	False	1	
640	Tianguá	2313401	2020-04-12	75946	False	False	3	
641	Tianguá	2313401	2020-04-08	75946	False	False	1	

```
In [158]: novos_casos <- df$new_confirmed # Variável resposta
idhm <- d$IDHM # Variável explicativa
dados <- data.frame(novos_casos,idhm)

modelo.regressao <- lm(novos_casos ~ idhm, data= dados)
summary(modelo.regressao) # Estimativa dos parâmetros, Erro, R2 do modelo

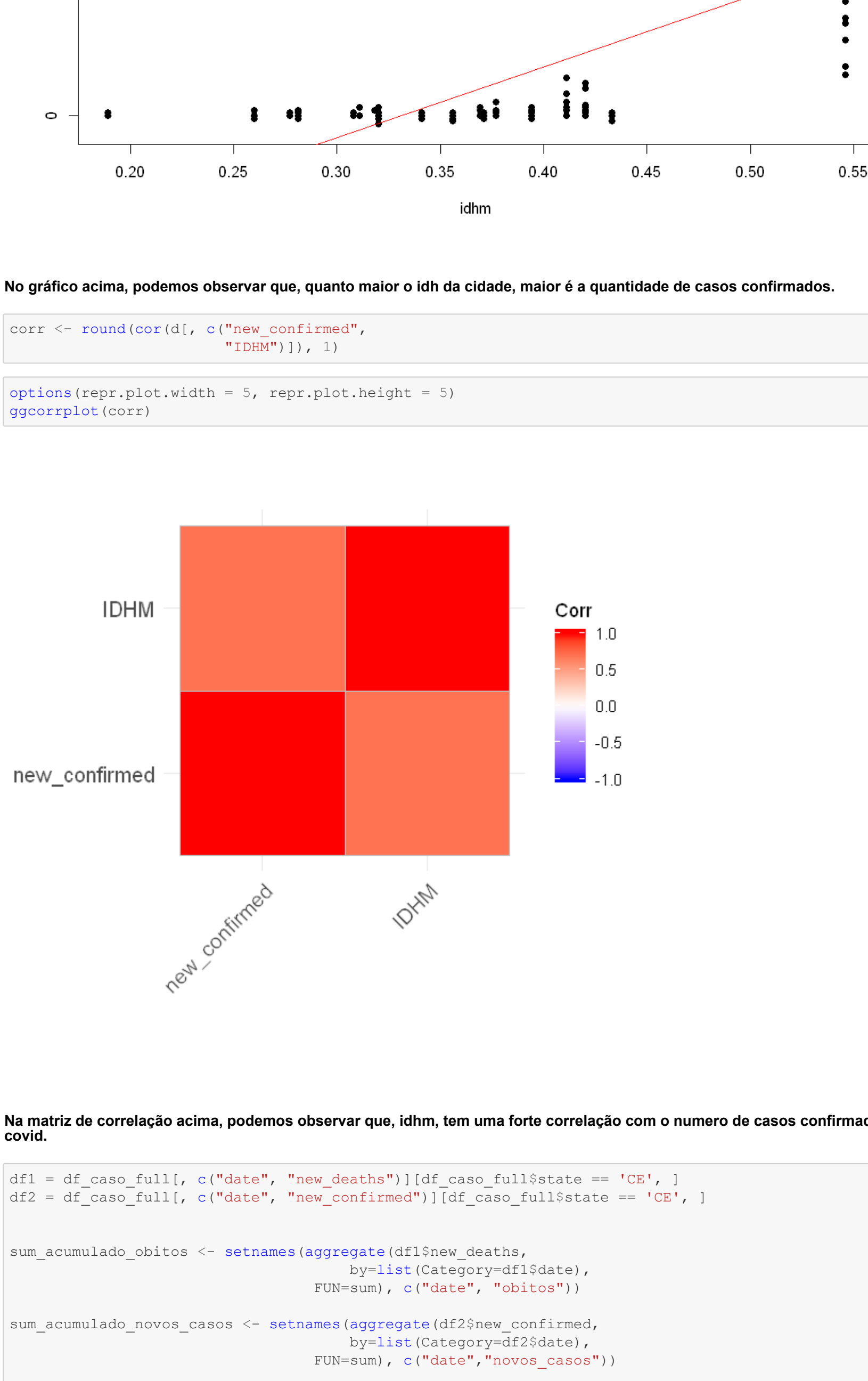
Call:
lm(formula = novos_casos ~ idhm, data = dados)

Residuals:
    Min       1Q   Median       3Q      Max
-40.996 -12.017  -6.552   5.939 132.004

Coefficients:
(Intercept)  Estimate Std. Error t value Pr(>|t|)
            -86.090      8.109   -10.62   <2e-16 ***
            idhm      260.230     12.148    21.43   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 23.37 on 184 degrees of freedom
Multiple R-squared:  0.4445, Adjusted R-squared:  0.4414
F-statistic: 147.2 on 1 and 184 DF, p-value: < 2.2e-16

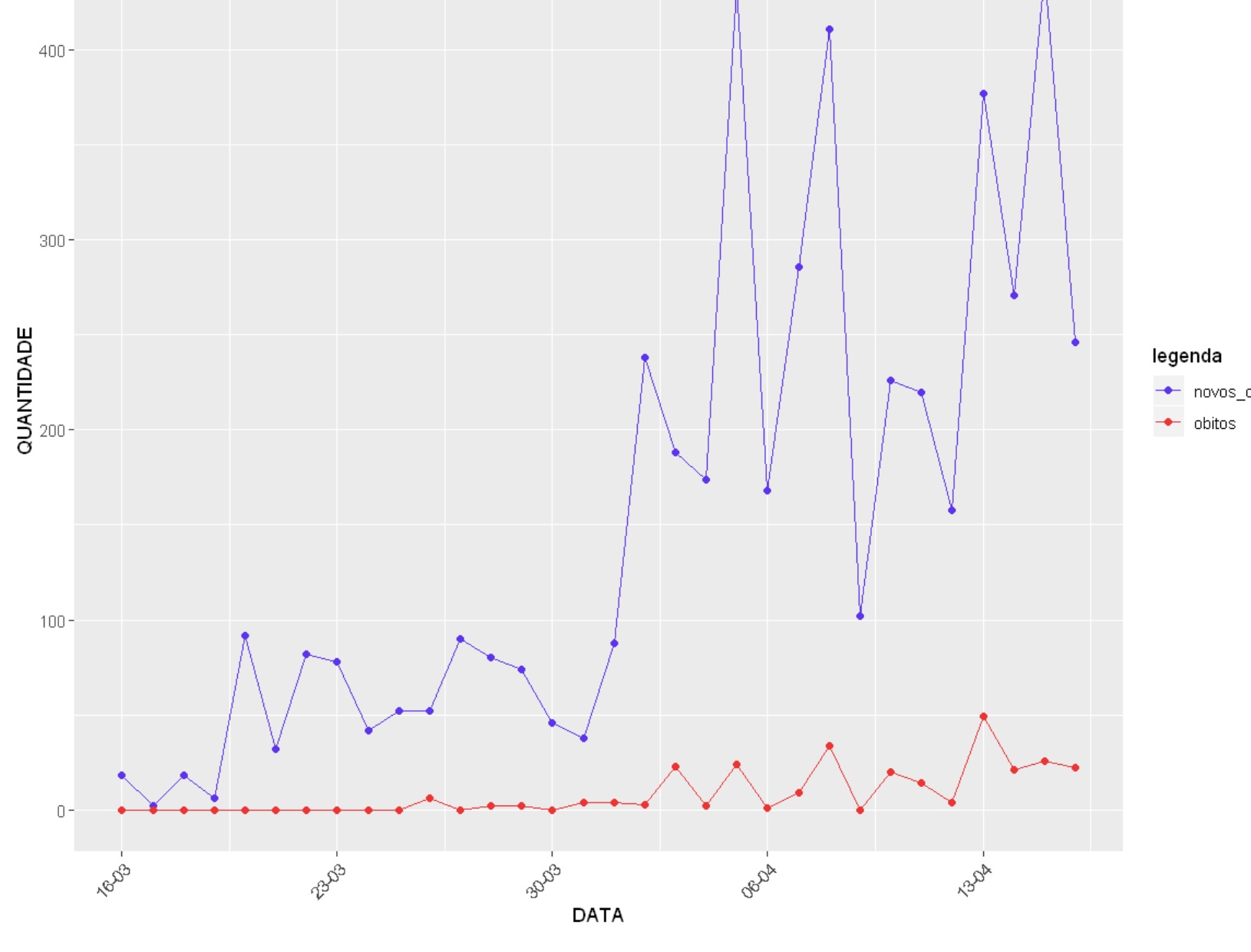
In [159]: plot(novos_casos ~ idhm, pch=16, data = dados)
abline(modelo.regressao,col="red") # Esta função ajusta a reta do modelo aos dados
```



No gráfico acima, podemos observar que, quanto maior o idh da cidade, maior é a quantidade de casos confirmados.

```
In [160]: corr <- round(corr(d[, c("new_confirmed",
                                "IDHM")]), 1)

In [161]: options(repr.plot.width = 5, repr.plot.height = 5)
ggcorplot(corr)
```



Na matriz de correlação acima, podemos observar que, idhm, tem uma forte correlação com o numero de casos confirmados de covid.

```
In [162]: df1 = df_caso_full[, c("date", "new_deaths")] [df_caso_full$state == 'CE', ]
df2 = df_caso_full[, c("date", "new_confirmed")] [df_caso_full$state == 'CE', ]

sum_acumulado_obitos <- setnames(aggregate(df1$new_deaths,
                                           by=list(Category=df1$date),
                                           FUN=sum), c("date", "obitos"))

sum_acumulado_novos_casos <- setnames(aggregate(df2$new_confirmed,
                                                by=list(Category=df2$date),
                                                FUN=sum), c("date", "novos_casos"))

df_final = merge(sum_acumulado_novos_casos, sum_acumulado_obitos, by = "date")
long <- reshape2::melt(df_final, id.vars = "date")

#renomeia colunas
long <- setnames(long, c("date", "Legenda", "value"))

In [163]: options(repr.plot.width = 10, repr.plot.height = 8)
plot <- ggplot(long, aes(x=(date), y = value,
                        group = Legenda, colour = Legenda)) +
  scale_color_manual(values=c("#5b344b", "#eb3434")) +
  geom_line() +
  geom_point() +
  labs(title = "Acumulado de Casos X Obitos no Ceará",
        x = "DATA", y = "QUANTIDADE")

plot + scale_x_date(date_labels = "%d-%m") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
```



```
In [164]: df_hist <- read.csv("HIST_PAINEL_COVIDBR_28fev2021.csv", header = TRUE, sep = ";")

In [165]: df_hist <- df_hist[ (df_hist$estado == 'CE' & df_hist$casosNovos == 1), ]

In [166]: df_hist[, ]
```

regiao	estado	municipio	coduf	codmun	codRegiaoSaude	nomeRegiaoSaude	data	semanaEpi	populacaoTCU2019	casosA
314118	Nordeste	CE	23	230000	NA	2020-04-07	15	NA	NA	

```
In [167]: #Converte factor to date
newDates <- as.Date(df_hist$date)
df_hist <- mutate(df_hist, data= newDates)
df_hist["mes_dia"] <- format(newDates, format = "%d-%m")

newDates <- as.Date(df_caso_full$date, format = "%d-%m-%Y")
df_caso_full["mes_dia"] <- format(newDates, format = "%d-%m")

In [168]: df_hist[, ]
```

regiao	estado	municipio	coduf	codmun	codRegiaoSaude	nomeRegiaoSaude	data	semanaEpi	populacaoTCU2019	casosAcumulados
Nordeste	CE	23	230000	NA	2020-04-07	15	NA	NA	1	

```
In [169]: df1 = df_caso_full[, c("mes_dia", "new_confirmed")] [df_caso_full$state == 'CE', ]
df2 = df_hist[, c("mes_dia", "casosNovos")] [df_hist$mes_dia %in% c(df_caso_full$mes_dia) & df_hist$date >= "2021-01-01", ]

sum_novos_casos_2020 <- setnames(aggregate(df1$new_confirmed,
                                           by=list(Category=df1$mes_dia),
                                           FUN=sum), c("date", "2020"))

sum_novos_casos_2021 <- setnames(aggregate(df2$casosNovos,
                                             by=list(Category=df2$mes_dia),
                                             FUN=sum), c("date", "2021"))

df_final = merge(sum_novos_casos_2020, sum_novos_casos_2021, by = "date")
long <- reshape2::melt(df_final, id.vars = "date")

#renomeia colunas
long <- setnames(long, c("date", "Legenda", "value"))

In [170]: #options(repr.plot.width = 10, repr.plot.height = 8)
plot <- ggplot(long, aes(x=(date), y = value,
                        group = Legenda, colour = Legenda)) +
  scale_color_manual(values=c("#5b344b", "#eb3434")) +
  geom_line() +
  geom_point() +
  labs(title = "Comparativo de novos casos no mês de fevereiro de 2020 e 2021",
        x = "DATA", y = "QUANTIDADE")

#plot + scale_x_date(date_labels = "%d-%m") + theme(axis.text.x = element_text(angle = 45, hjust = 1))
plot
```

