

# ggplot2

*David Criado Ramón*

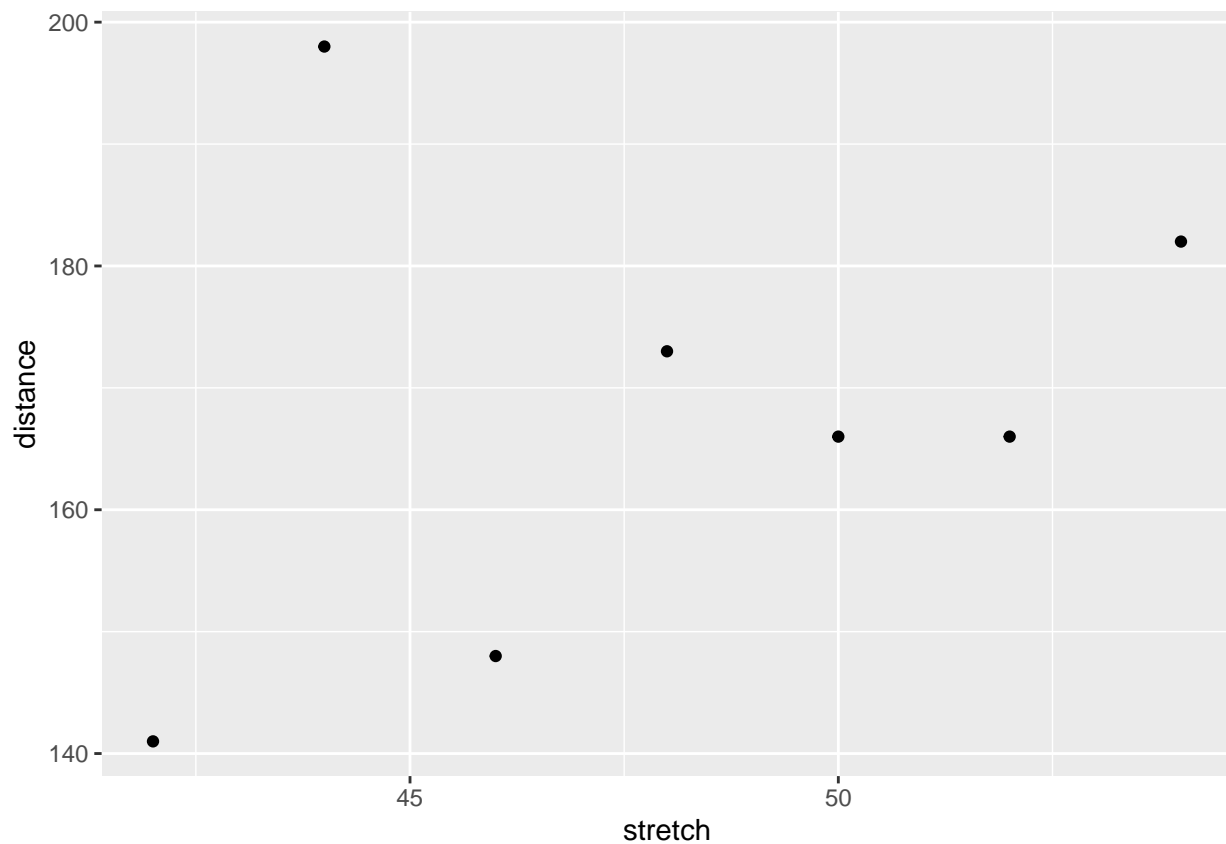
*11/11/2019*

## Ejercicio 1.

```
stretch <- c(46, 54, 48, 50, 44, 42, 52)
distance <- c(148, 182, 173, 166, 198, 141, 166)
df <- as.data.frame(cbind(stretch, distance))
```

Pinta distance contra stretch.

```
ggplot(df, aes(x=stretch, y=distance)) + geom_point()
```



## Ejercicio 2

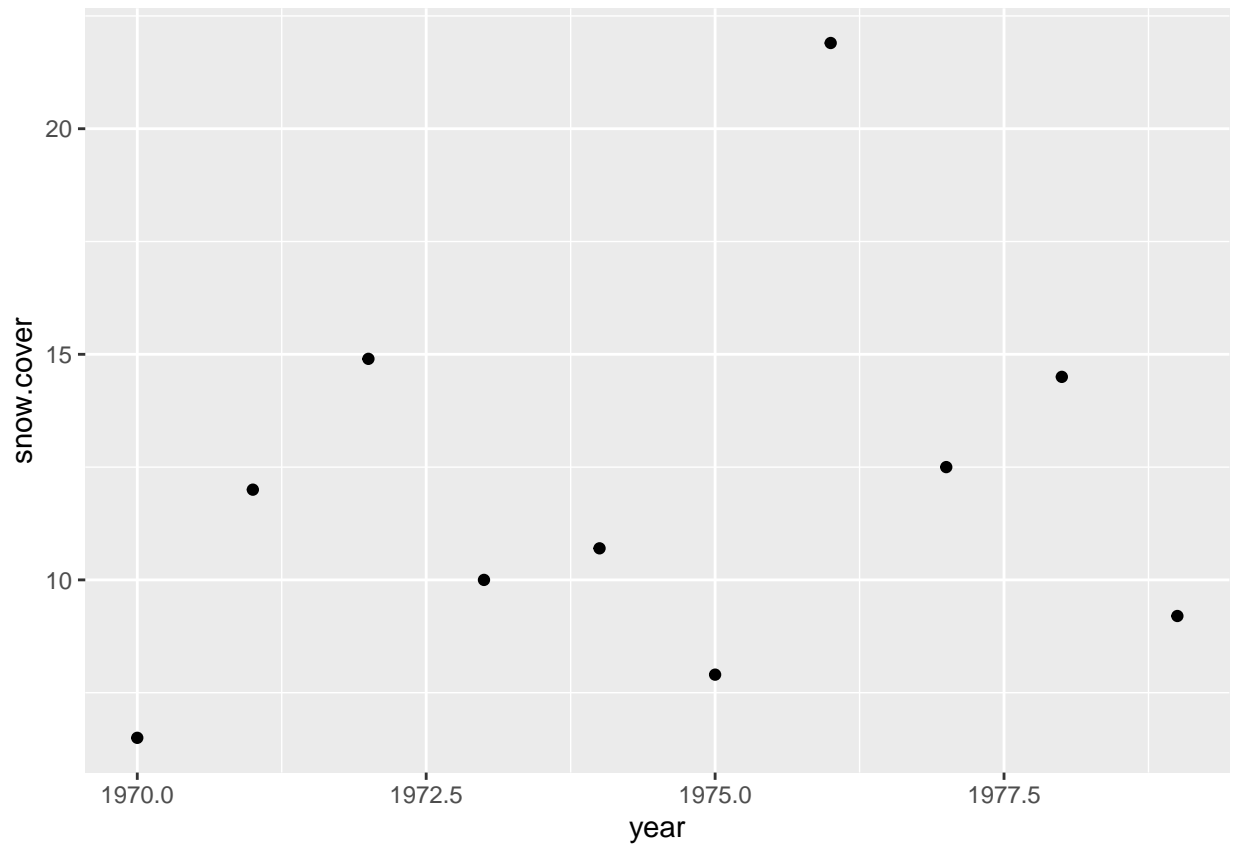
```
year <- 1970:1979
snow.cover <- c(6.5, 12, 14.9, 10, 10.7, 7.9, 21.9, 12.5, 14.5, 9.2)
df <- as.data.frame(cbind(year, snow.cover))
df
```

```
##   year snow.cover
## 1  1970         6.5
```

```
## 2 1971      12.0
## 3 1972      14.9
## 4 1973      10.0
## 5 1974      10.7
## 6 1975       7.9
## 7 1976      21.9
## 8 1977      12.5
## 9 1978      14.5
## 10 1979       9.2
```

a) Pinta snow.cover contra año

```
ggplot(df, aes(x=year, y=snow.cover)) + geom_point()
```



Puesto que es una progresión histórica a lo largo de los años parece tener más sentido usar un gráfico de líneas.

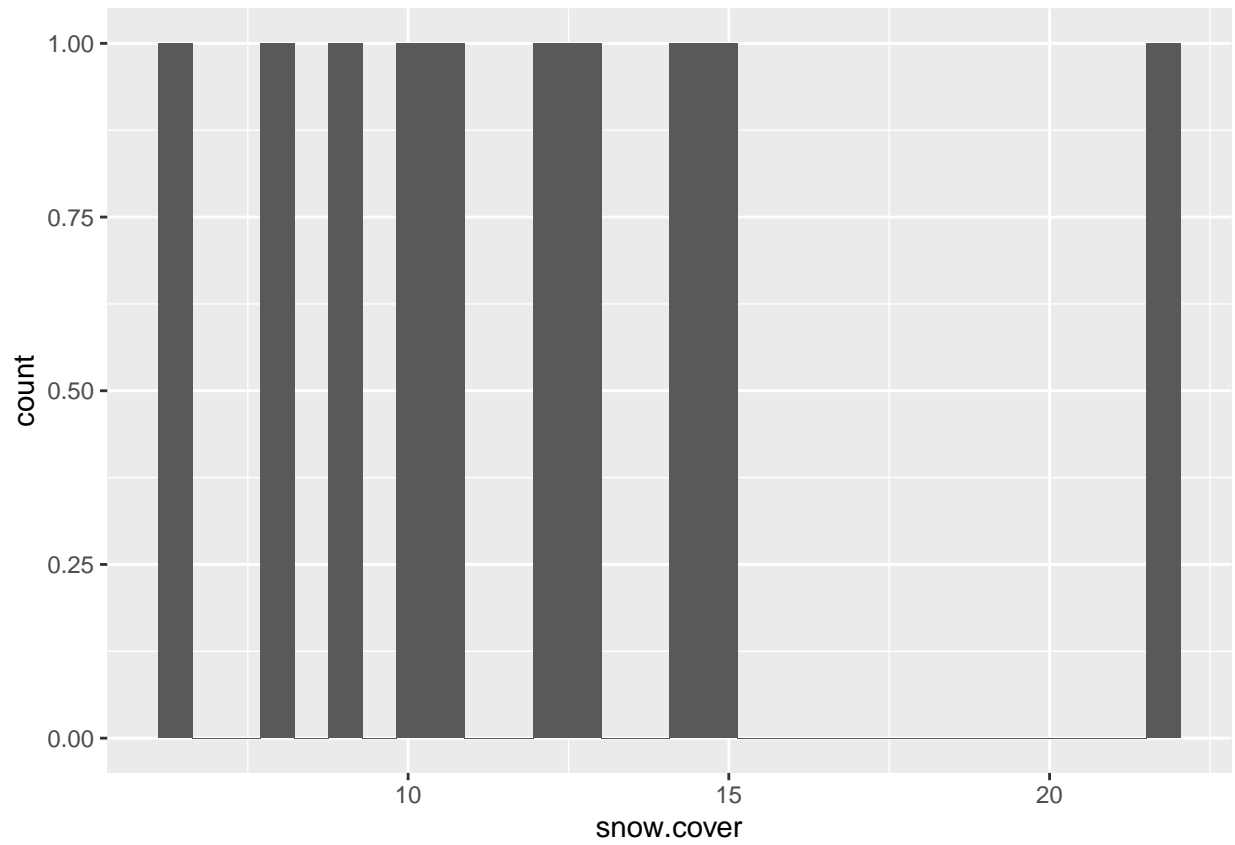
```
ggplot(df, aes(x=year, y=snow.cover)) + geom_line()
```



b) Pinta un histograma de los valores de snow.cover

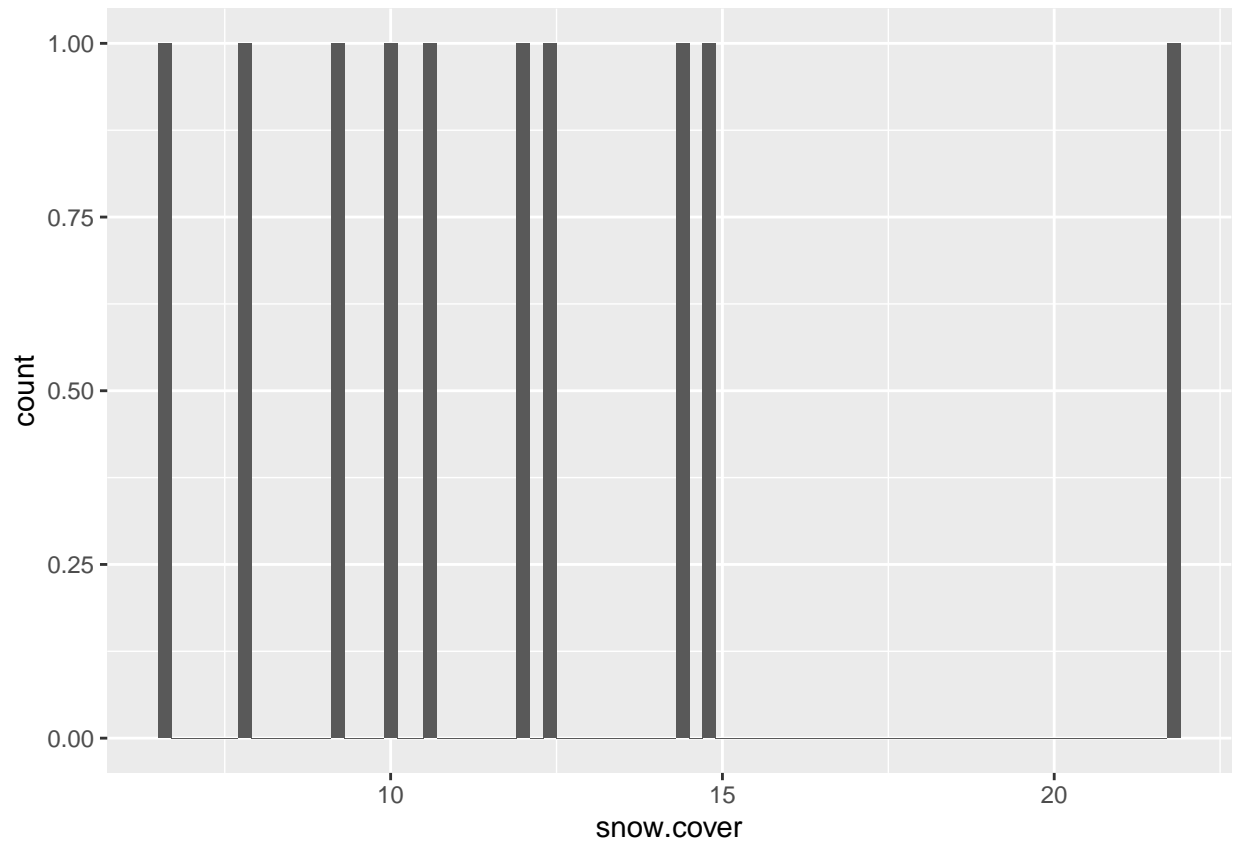
```
ggplot(df, aes(snow.cover)) + geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```



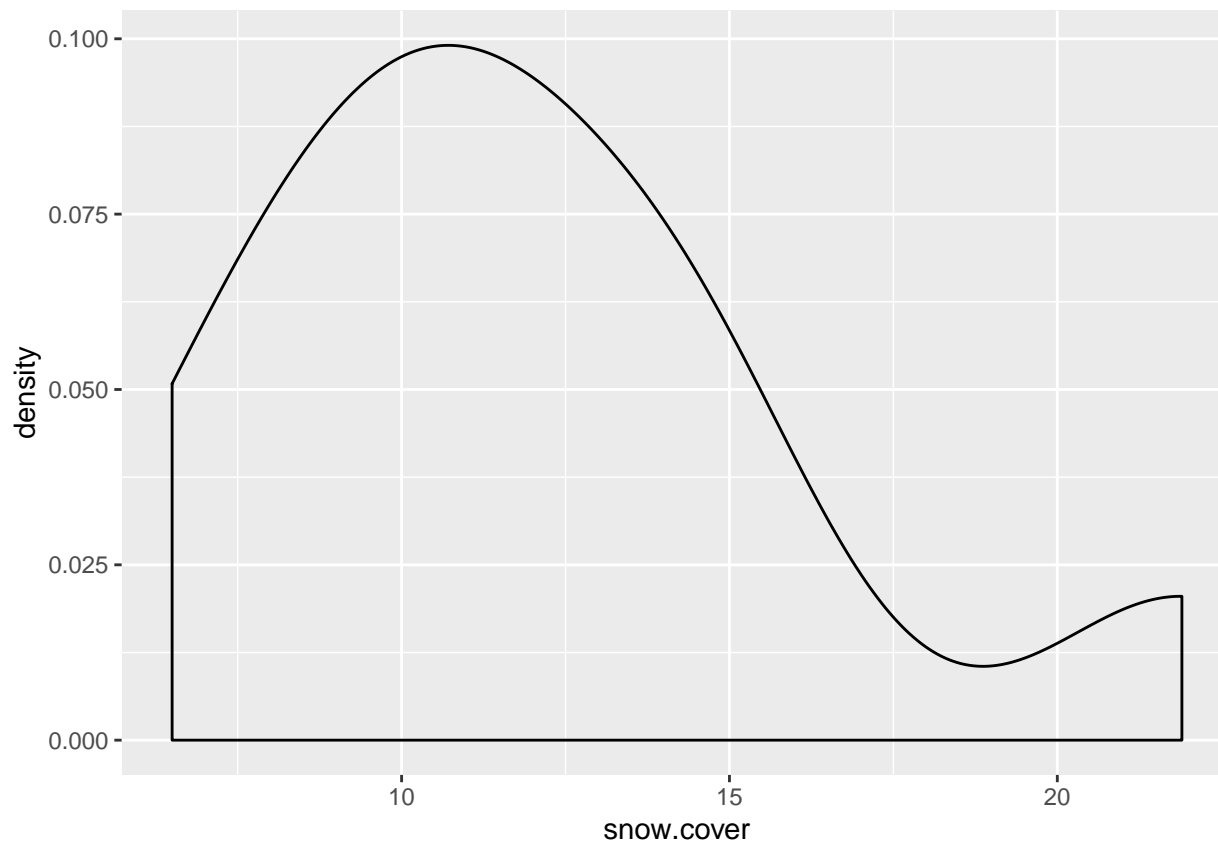
Probemos con un binwidth distinto de 30

```
ggplot(df, aes(snow.cover)) + geom_histogram(binwidth = 0.2)
```



Cuando tenemos tan pocos datos es más interesante generar una curva de densidad de probabilidad.

```
ggplot(df, aes(snow.cover)) + geom_density()
```



### Ejercicio 3.

Dados los de datos del tiempo en NY.

```
df <- read.csv("weather_data_nyc_centralpark_2016.csv", na.strings="T")
```

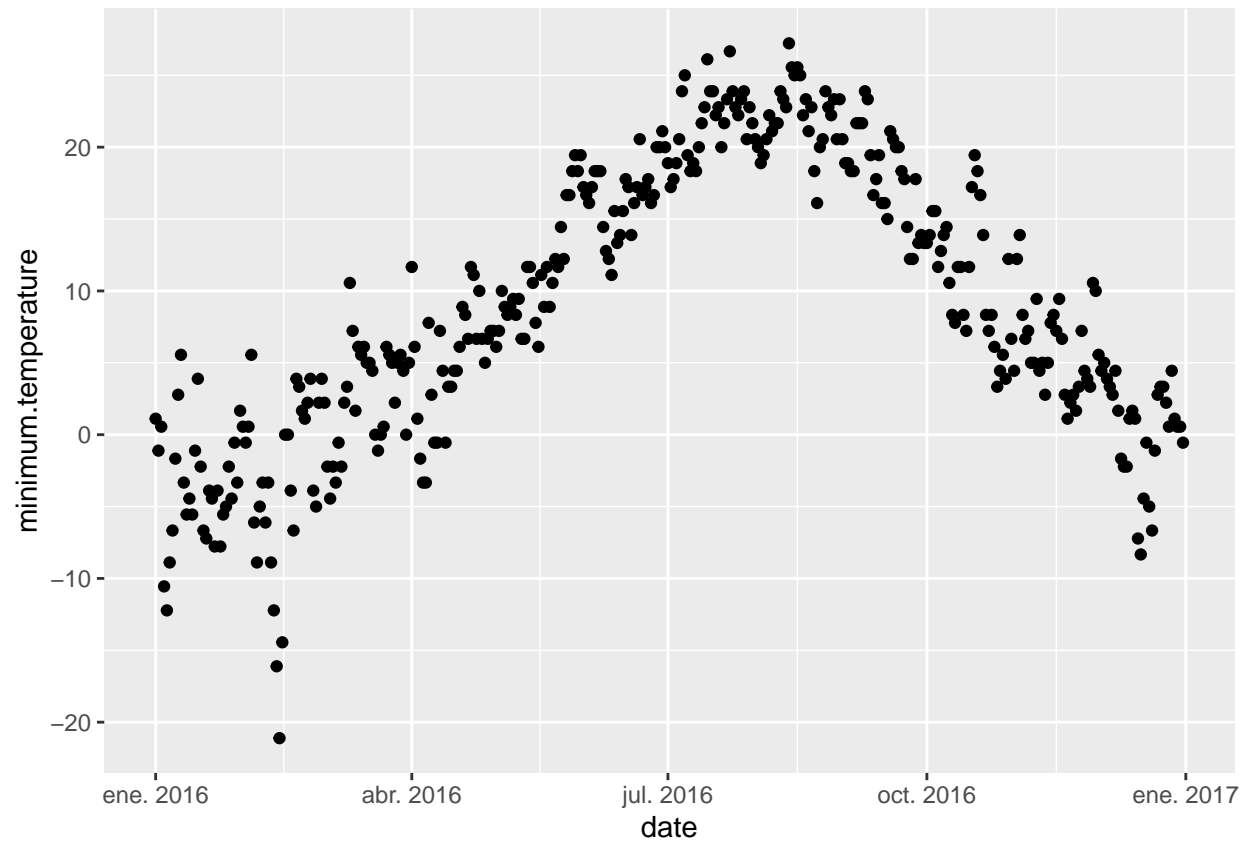
a) Convierte °F a °C e in a mm.

La T indica que ha habido alguna precipitación pero no suficiente para poder realizar una medida.

```
to_degrees <- function(f) (f-32)/1.8
to_mm <- function(i) i*25.4
names(df)[1] <- "date"
df <- df %>% mutate(date=as.Date(date, format="%d-%m-%y"),
                    maximum.temperature=to_degrees(maximum.temperature),
                    minimum.temperature=to_degrees(minimum.temperature),
                    average.temperature=to_degrees(average.temperature),
                    precipitation=to_mm(precipitation),
                    snow.fall=to_mm(snow.fall),
                    snow.depth=to_mm(snow.depth))
```

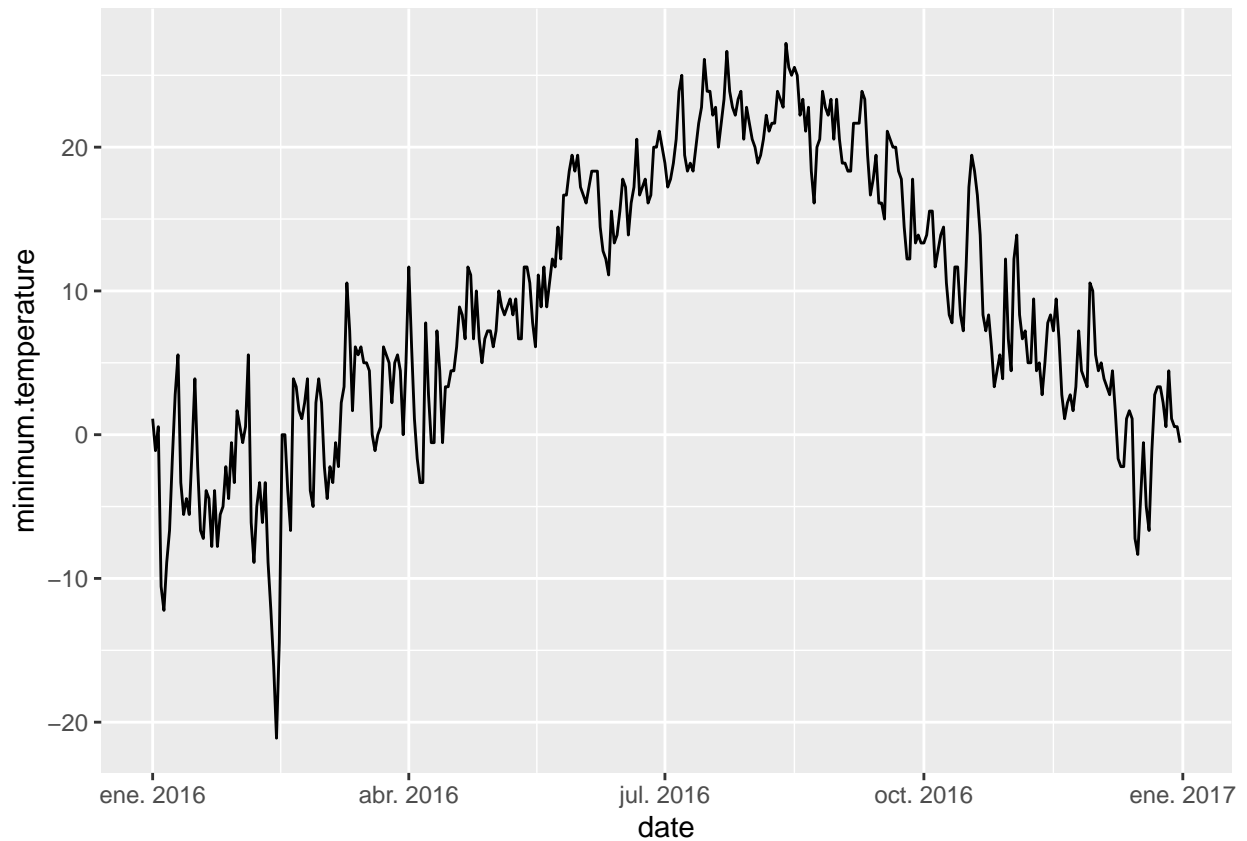
b) Pinta fecha vs temperatura mínima más caliente.

```
ggplot(df, aes(x=date, y=minimum.temperature)) + geom_point()
```



Siendo una progresión temporal, me parece más informativo hacerlo con un gráfico de líneas.

```
ggplot(df, aes(x=date, y=minimum.temperature)) + geom_line()
```



c) Pinta fecha vs temperatura mínima más caliente y temperatura mínima más fría.

Ahora utilizamos dos colores para diferenciar.

```
ggplot(df, aes(x=date)) + geom_line(aes(y=minimum.temperature), col="blue") +  
  geom_line(aes(y=maximum.temerature), col="red") + ggtitle("Temperatura máxima y mínima según f
```



Temperatura máxima y mínima según fecha

