

Introducción a gráficos en R

Armando Ocampo

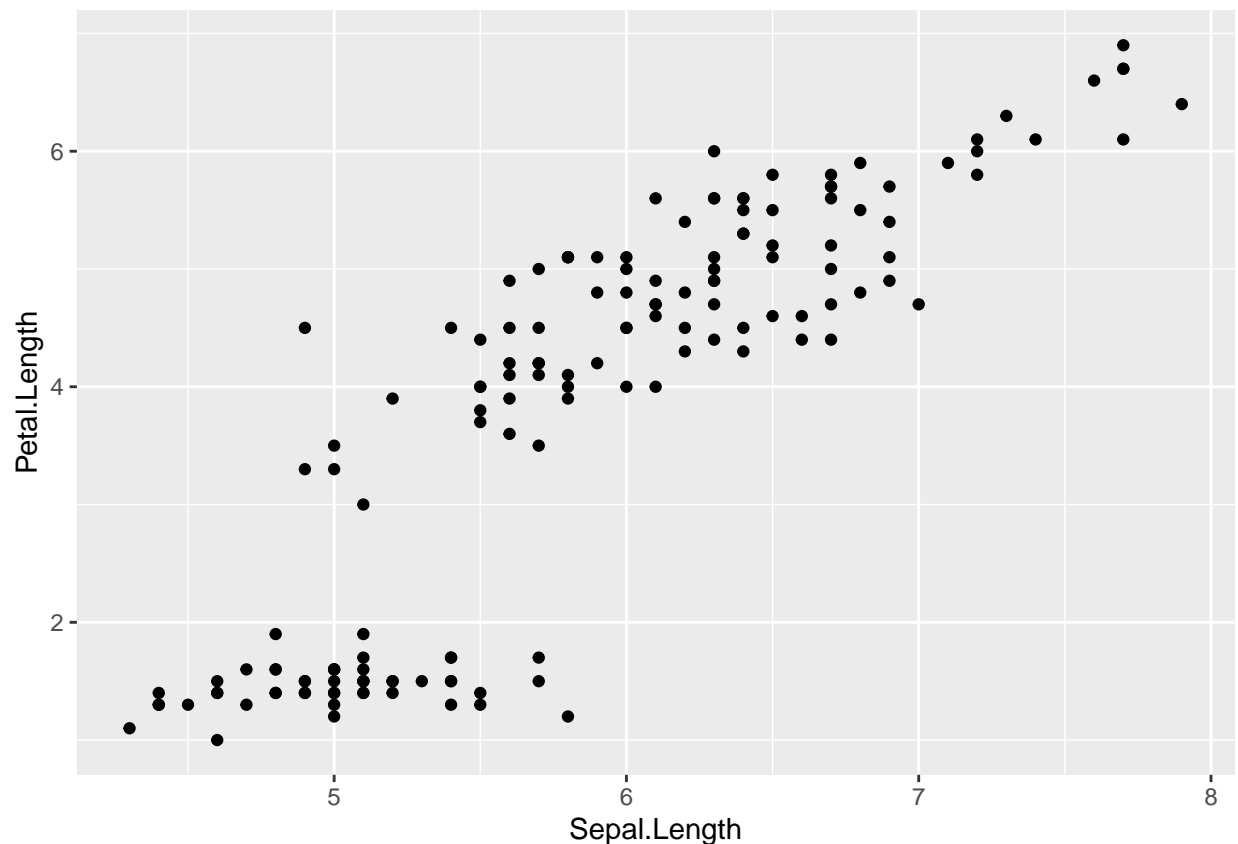
Características principales

En R, existen diferentes paqueterías para realizar gráficos. En esta clase utilizaremos la paquetería *ggplot2*, debido a la estética y manejo del lenguaje al momento de generar gráficos. Otra de las razones es que al pertenecer al conjunto de paqueterías de *tidyverse*, permite generar flujos de trabajo. Cargaremos la librería antes de comenzar a trabajar.

```
library(ggplot2)
```

Para realizar un gráfico son necesarios tres elementos. El conjunto de datos, los ejes (en R se define cómo *aesthetic*) y el tipo de gráfico. A continuación definiremos un gráfico de dispersión con información de longitud de sépalo y pétalo presente en el dataset *iris*.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length)) +  
  geom_point()
```

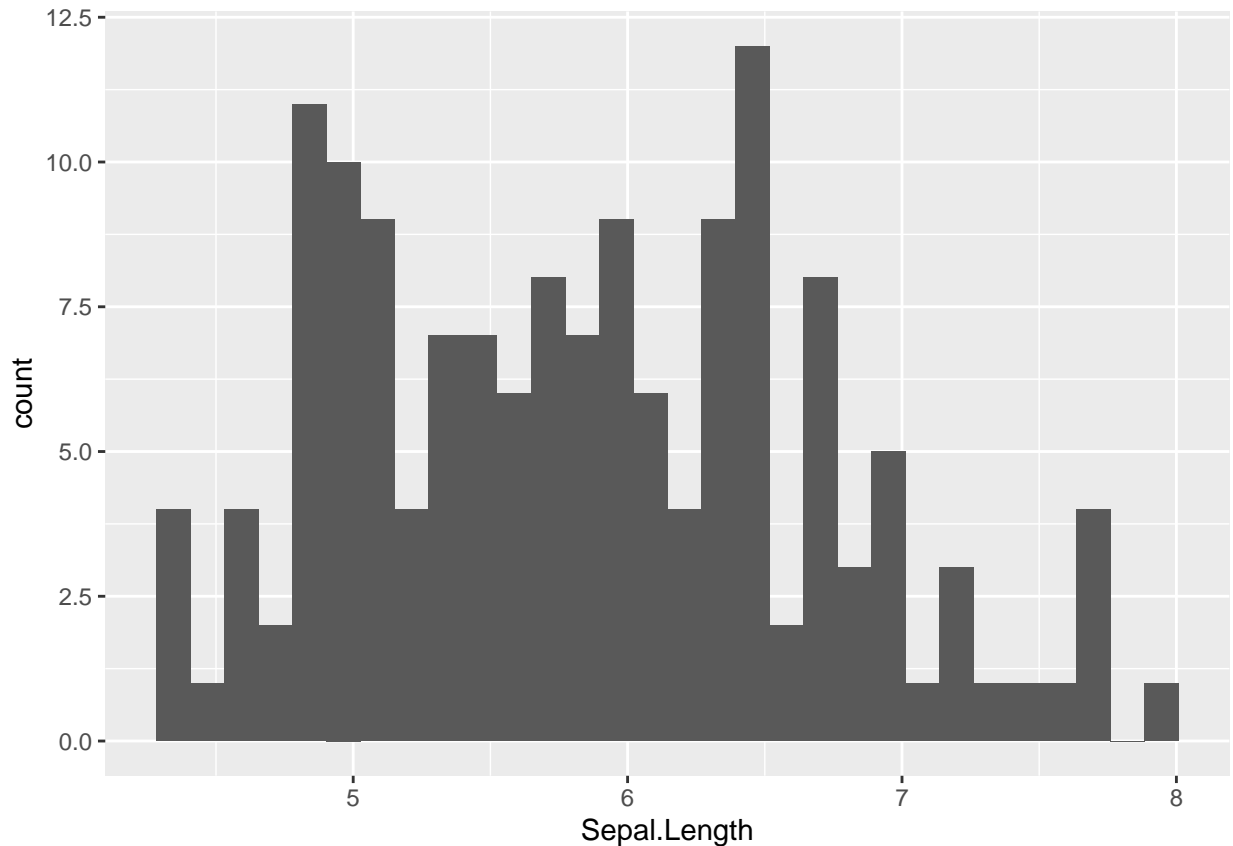


Ahora, describiremos cada uno de los tres elementos. La función principal es *ggplot()*, como primer argumento se escribe el nombre del dataset que vamos a utilizar. Posteriormente, la función *aes()* define la información

presente en cada eje. Para continuar con el gráfico se agrega el signo '+', como siguiente elemento se añade el tipo de gráfico con la función `geom_x()`. Para realizar un gráfico de puntos utilizamos la función `geom_point()`.

El tipo de gráfico a utilizar depende de las variables de trabajo. Por ejemplo, para los histogramas solo necesitamos definir una variable en el eje de las equis.

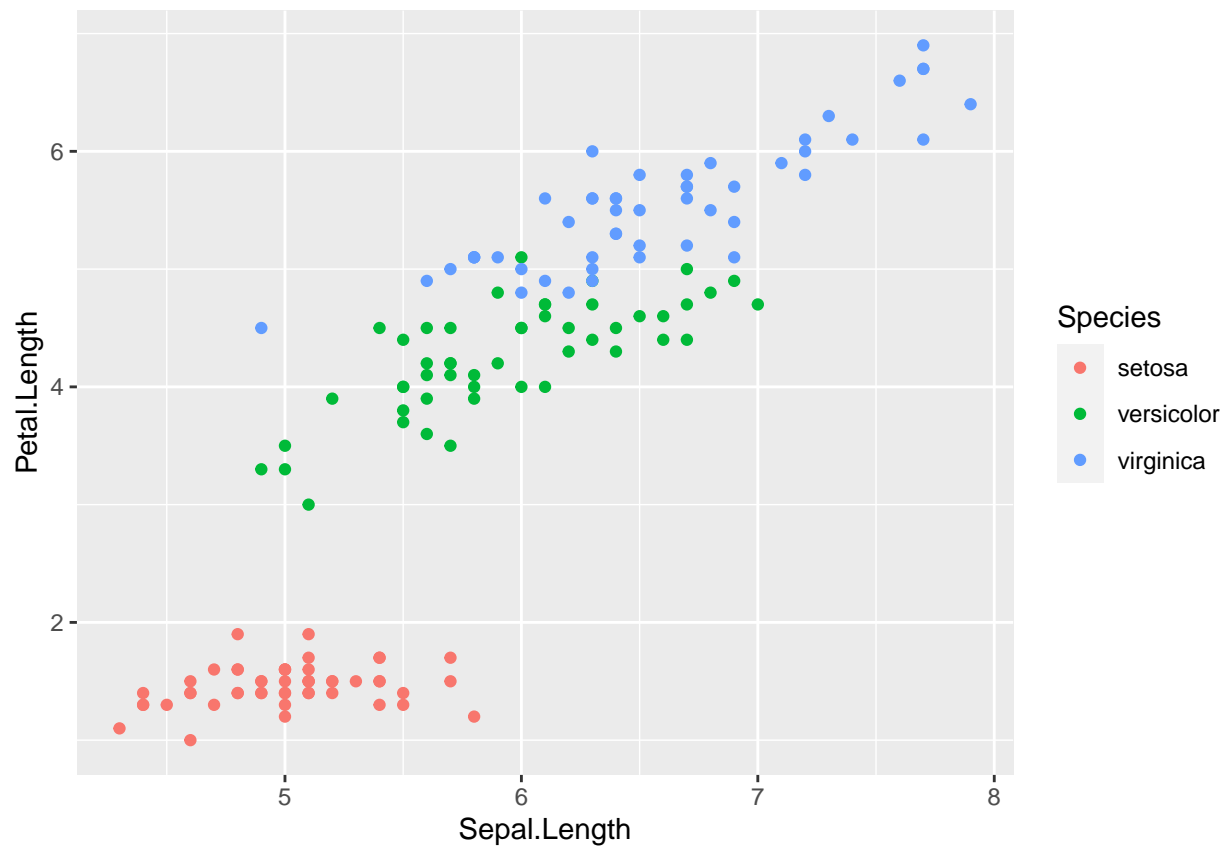
```
ggplot(iris, aes(x = Sepal.Length)) +  
  geom_histogram()
```



Otras características

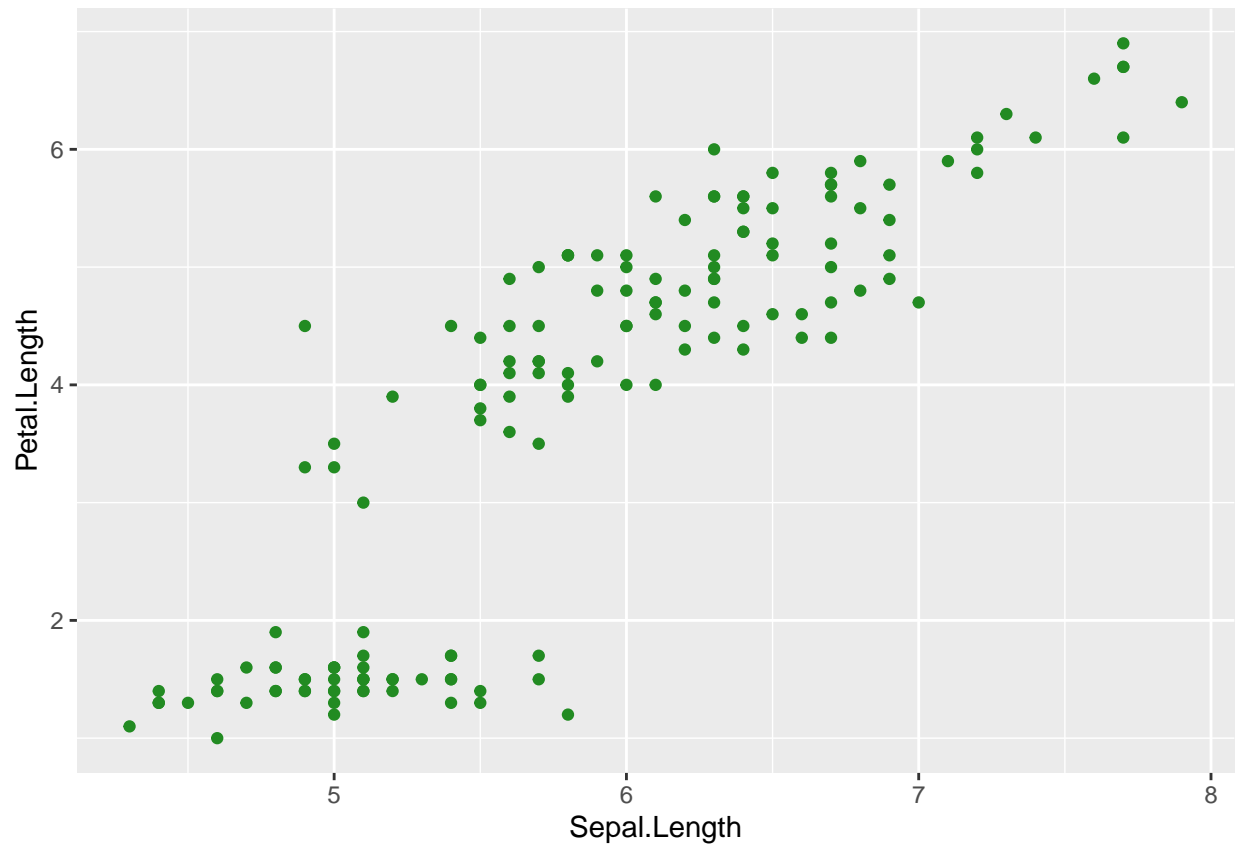
El color de los elementos se puede definir de dos maneras. De forma general o agregando una tercera variable. Para este último caso el color se define dentro de la estética añadiendo la variable de la cual se va a agregar el color. En el siguiente ejemplo, se retomará el gráfico de puntos y se coloreará cada punto acorde al tipo de especie.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length, color = Species))+  
  geom_point()
```



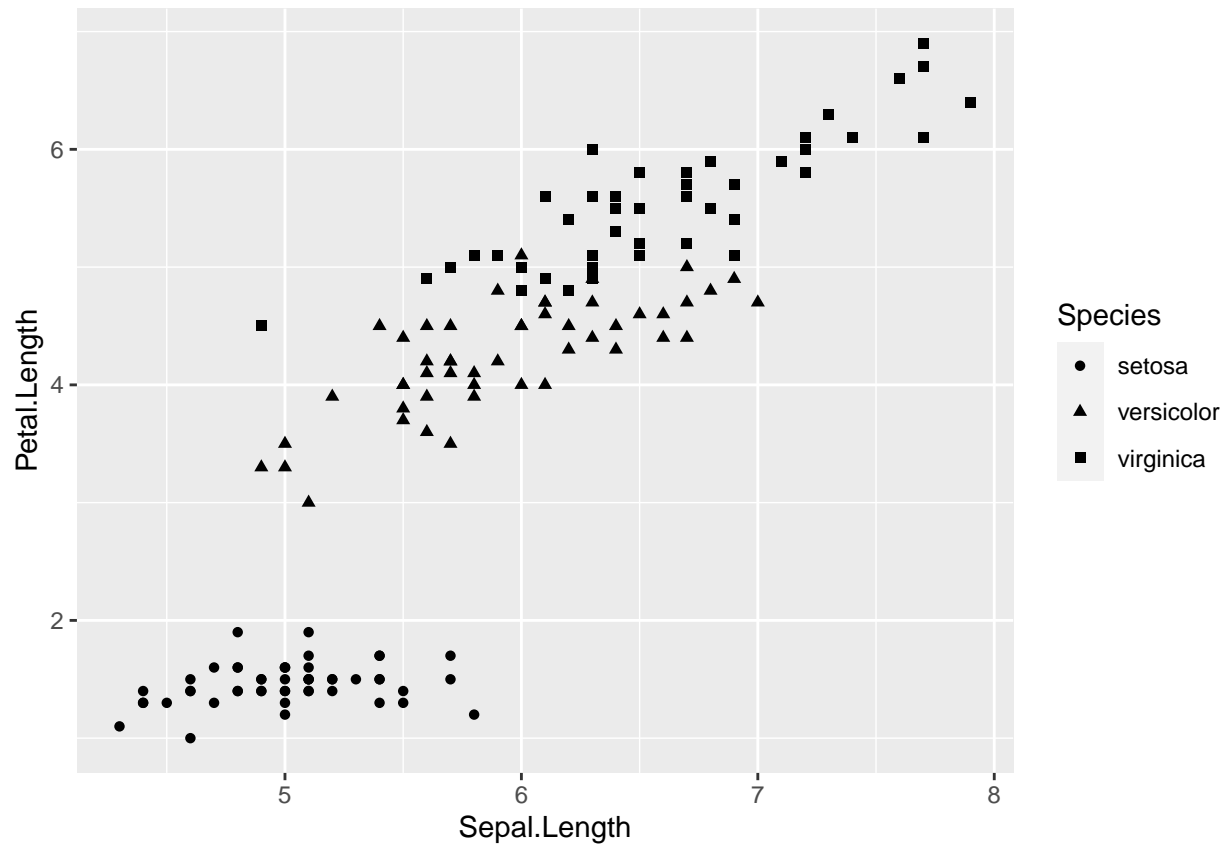
En caso de querer colorear de manera general cada elemento, es necesario colocar el argumento `color =` dentro de la función `geom_x()`. En el siguiente ejemplo, retomaremos el gráfico de puntos y se coloreará cada elemento de manera general.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length))+  
  geom_point(color = 'forestgreen')
```



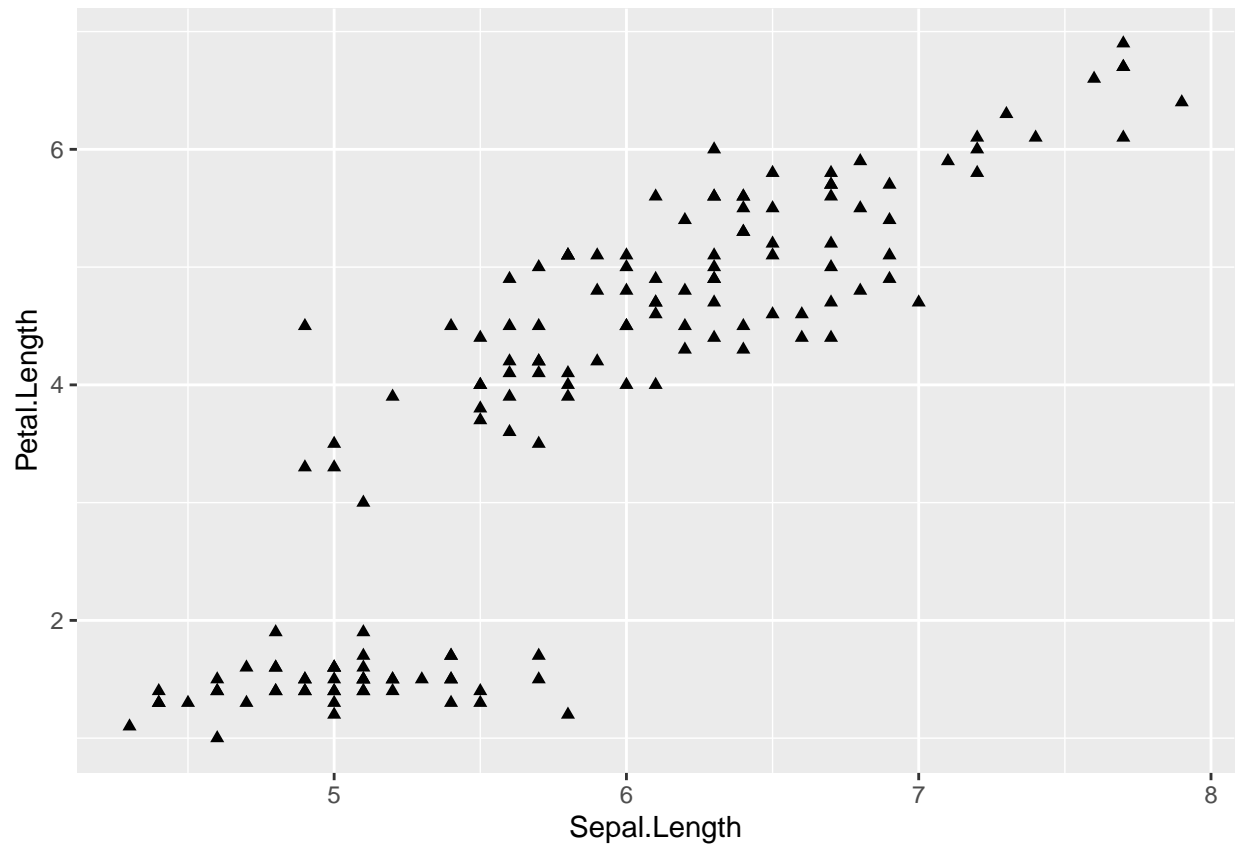
Para el gráfico de puntos, también es posible modificar la forma. De la misma manera, el cambio puede ser tomando una variable o general. En el siguiente ejemplo se modificará la forma acorde al tipo de especie.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length, shape= Species))+  
  geom_point()
```



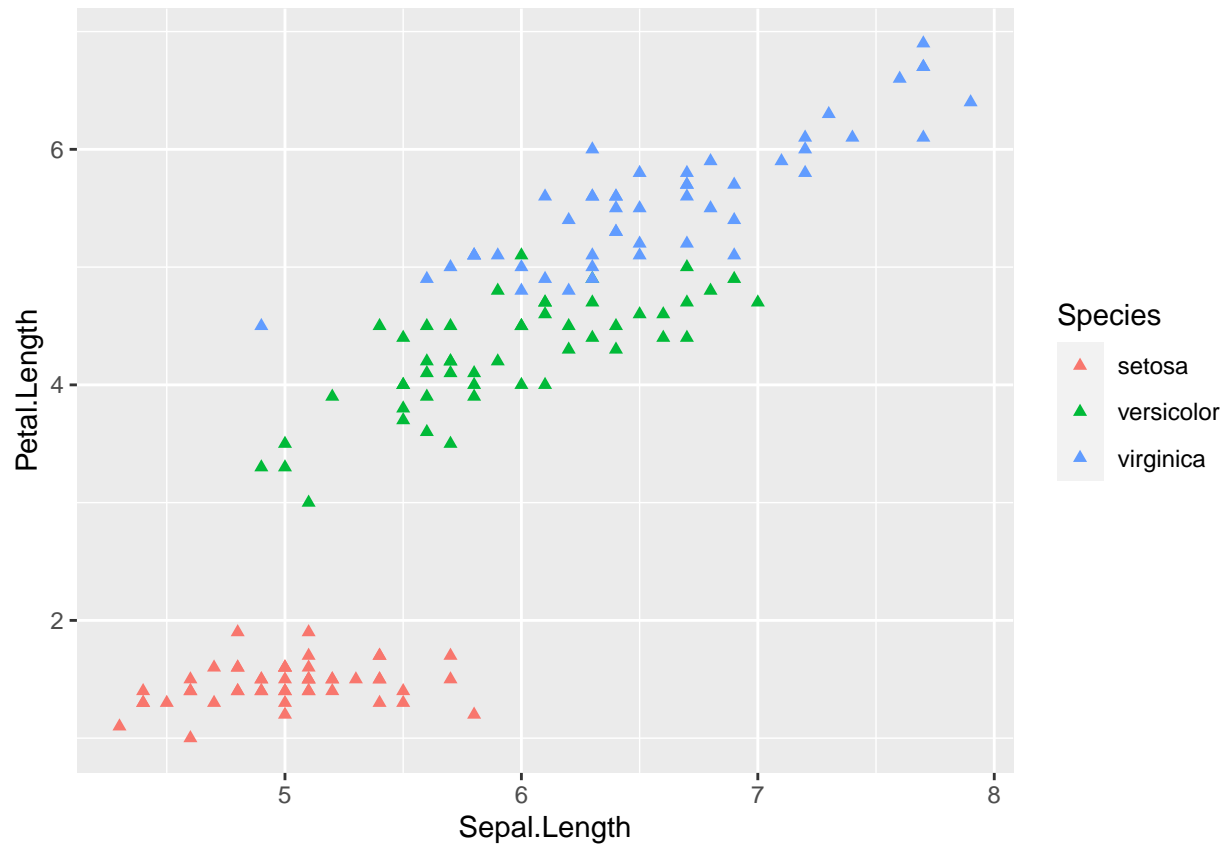
A continuación, el cambio de forma se realizará de manera global.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length))+  
  geom_point(shape = 17)
```



Los gráficos de *ggplot* permiten combinar estas características. Teniendo puntos con una sola forma y colores diferentes acorde a la especie.

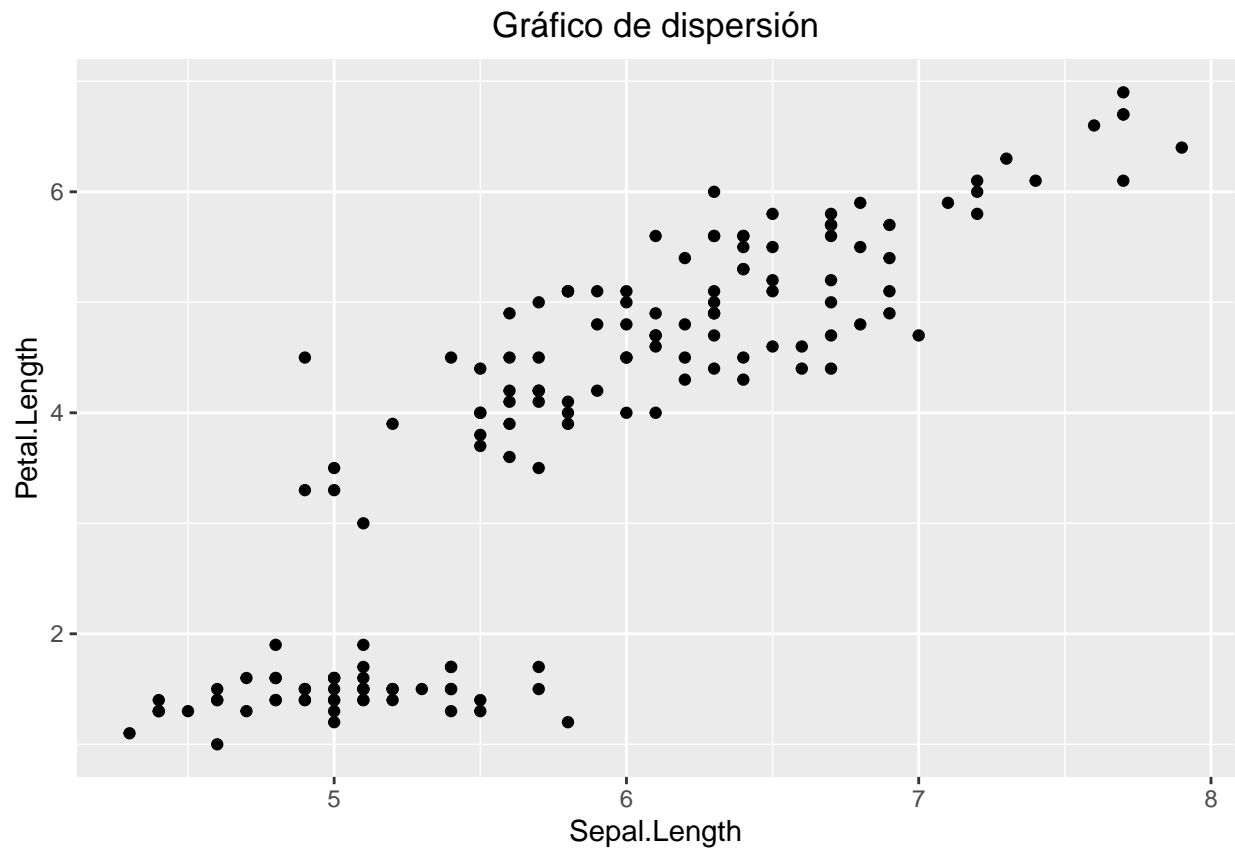
```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length, color= Species))+  
  geom_point(shape = 17)
```



Títulos

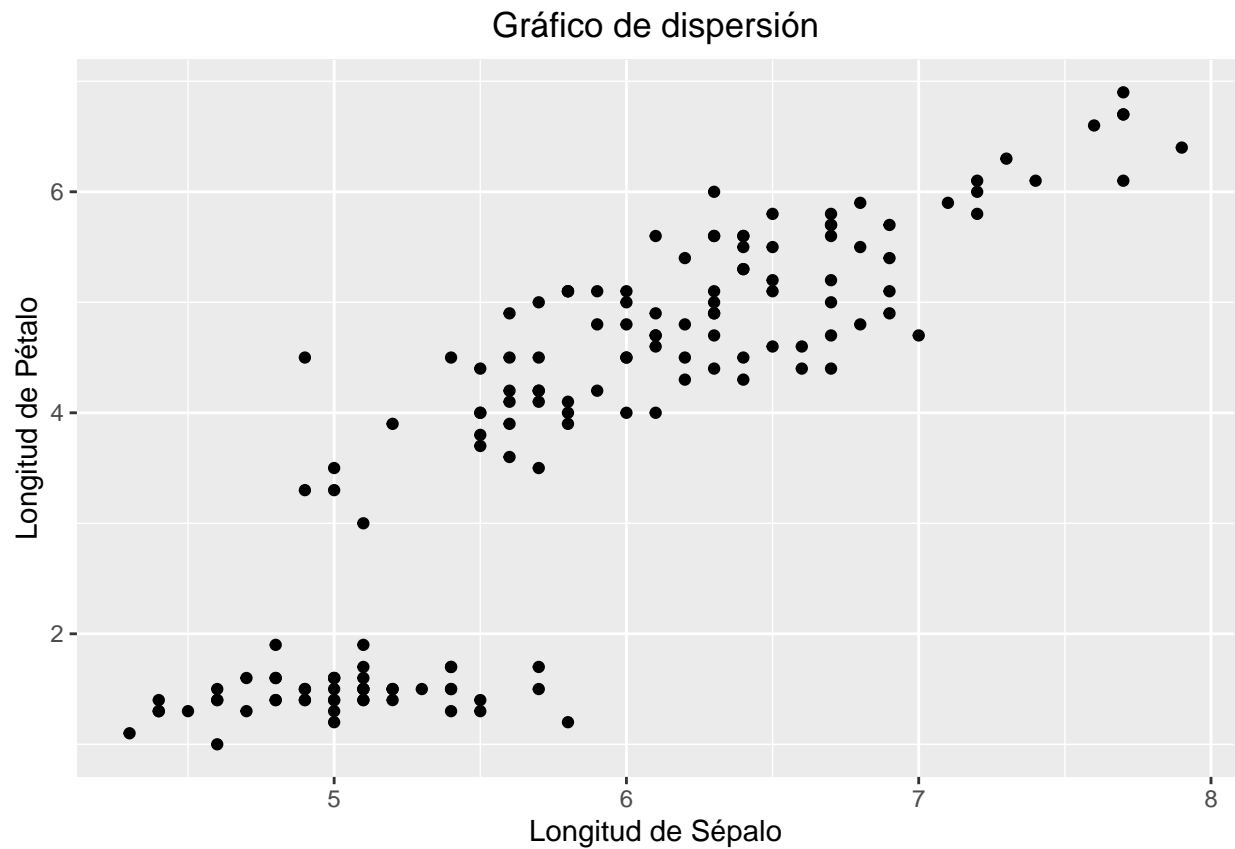
El título del gráfico se agrega utilizando la función `ggtitle()`. Dentro de la función se coloca el conjunto de caracteres entre comillas.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length))+  
  geom_point()+  
  ggtitle('Gráfico de dispersión')+  
  theme(plot.title = element_text(hjust = 0.5)) #centrar título
```



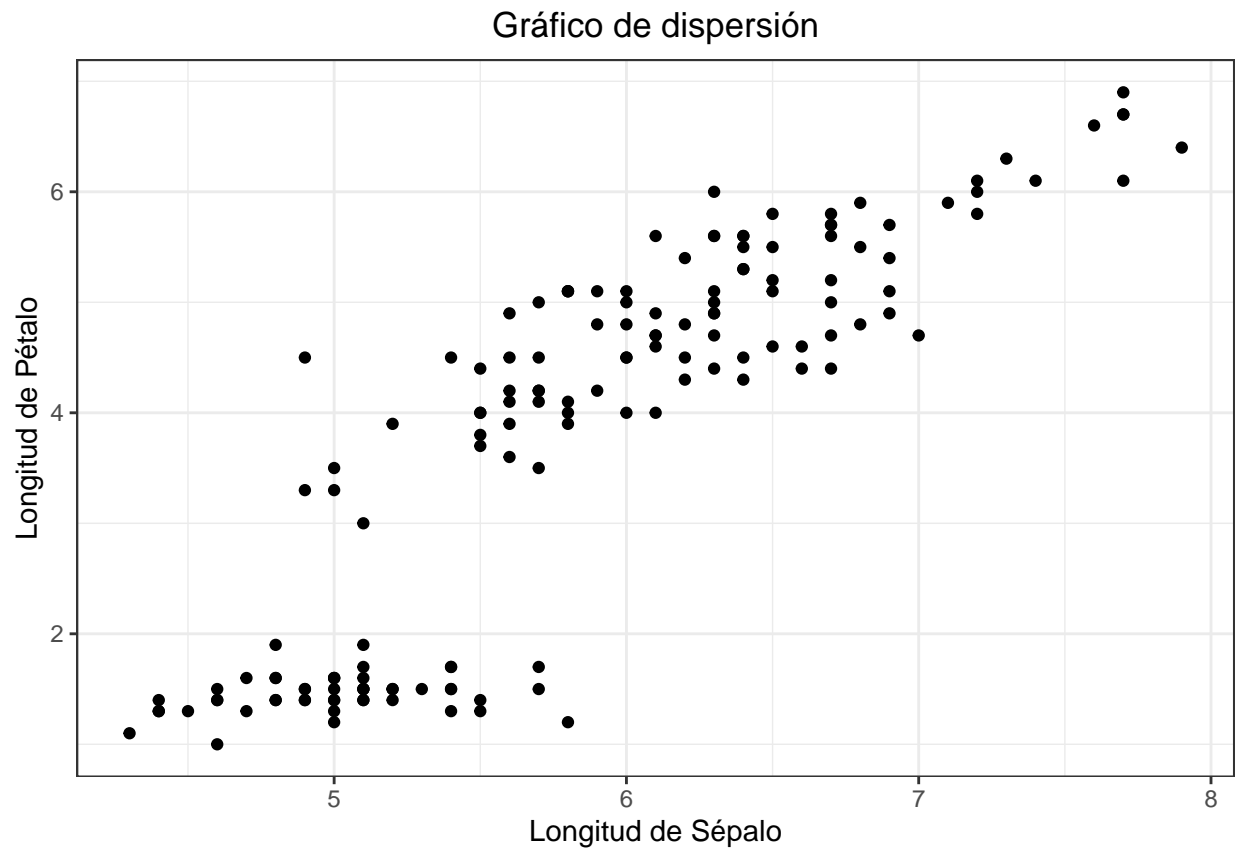
Por su parte, el nombre de los ejes se modifican con las funciones `xlab()` y `ylab()`.

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length))+  
  geom_point()+  
  ggtitle('Gráfico de dispersión')+  
  theme(plot.title = element_text(hjust = 0.5))+  
  xlab('Longitud de Sépalo')+  
  ylab('Longitud de Pétalo')
```

El fondo del gráfico se modifica utilizando la función `theme_x()`. En el siguiente ejemplo, se agregará un fondo blanco al gráfico, con la función `theme_bw()`

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length))+  
  geom_point()+  
  ggtitle('Gráfico de dispersión')+  
  theme_bw()+  
  theme(plot.title = element_text(hjust = 0.5))+  
  xlab('Longitud de Sépalo')+  
  ylab('Longitud de Pétalo')
```



Combinando todo

```
ggplot(iris, aes(x = Sepal.Length, y = Petal.Length, color=Species))+  
  geom_point(shape=17)+  
  ggtitle('Gráfico de dispersión')+  
  theme_bw() +  
  theme(plot.title = element_text(hjust = 0.5))+  
  xlab('Longitud de Sépalo')+  
  ylab('Longitud de Pétalo')
```

