

Introducción a ggplot2 y dplyr en R

Aprendamos un poco sobre `data.frames`, una de las estructuras más comunes en R y dos librerías, `dplyr` (manipulación de `data.frames`) y `ggplot2` (visualizaciones).]

Cuando importamos `tidyverse` importamos varias librerías incluyendo `ggplot2` y `dplyr`. Además importamos la librería `nycflights13` la cual usaremos para hacer ejercicios. Recuerda que si no tienes alguno de esto paquetes instalados puedes correr `install.packages("<LIBRARY_NAME>")`.

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.3.0      v purrr   0.3.3
## v tibble  2.1.3      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

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

library(datasets)
library(nycflights13)
data(iris)
```

Ggplot2

`ggplot` usa un estructura básica para generar sus visualizaciones donde primero definimos los datos a través de un `data.frame` y luego agregamos distintas partes, transformaciones, variables o separaciones a nuestro gráfico. La estructura es,

```
ggplot(data = <DATA>) +
  <GEOM_FUNCTION>(
    mapping = aes(<MAPPINGS>),
    stat = <STAT>,
    position = <POSITION>
  ) +
  <COORDINATE_FUNCTION> +
  <FACET_FUNCTION>
```

Estas partes de manera muy resumida significan:

1. `ggplot()`: construye la figura y define datos generales
2. `geom`: es el tipo de visualización (`scatter`, `línea`, `barras`, etc) Dentro de este además tenemos `mapping`: lo que conecta las variables a los elementos del gráfico, `stat`: generar nuevas variables como contar. `position`: arreglos a los gráficos `dodge`, `fill`, `jitter`, `stack` (revisar cheat sheet)
3. `coordinate`: nos permite cambiar el tipo de ejes: ej, transposición, coordenadas polares, etc.
4. `facet`: separar en varias subfiguras.

Realmente, para crear un gráfico sencillo, solo necesitamos lo siguiente

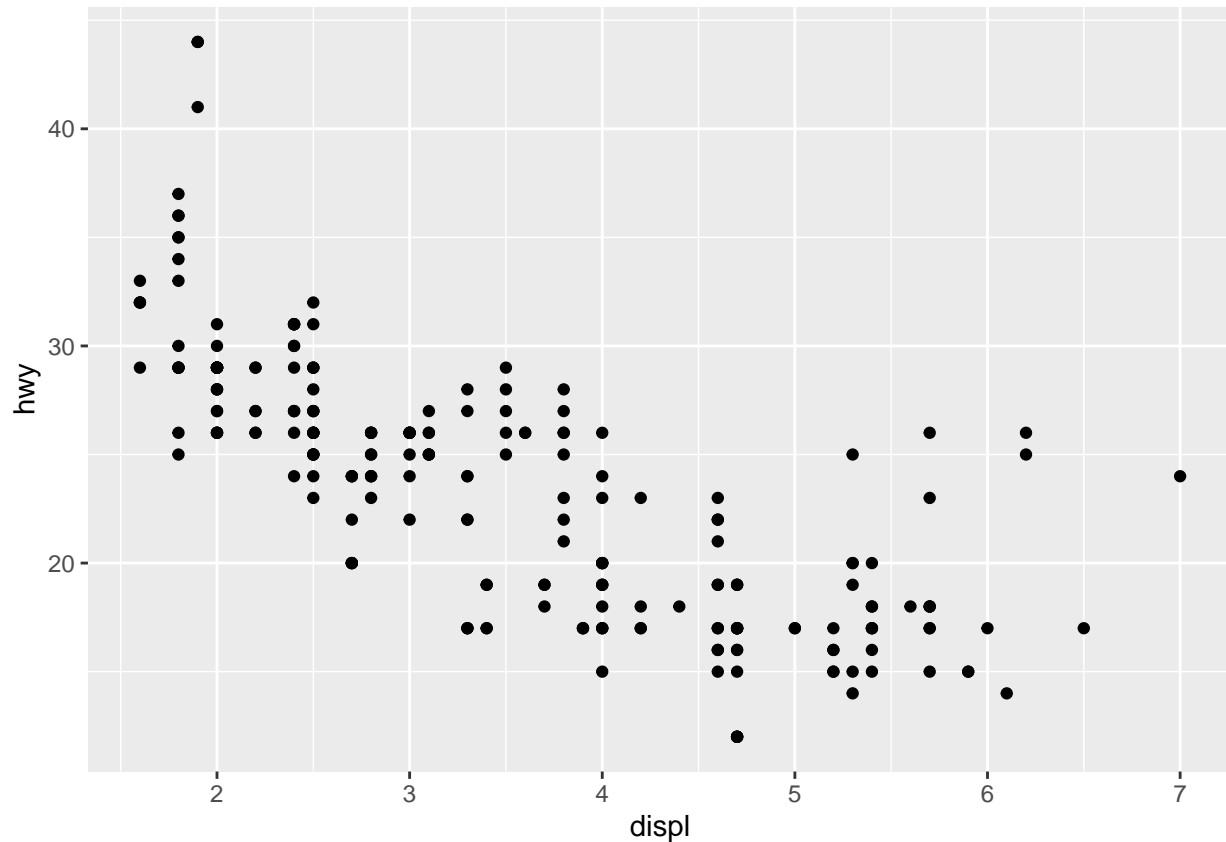
```
ggplot(data = <DATA>) +
```

```
<GEOM_FUNC>(mapping = aes(<MAPPING>))
```

Las otras partes las podemos ignorar por ahora pero las veremos en un momento.

Antes de empezar hagamos una visualización rápida y analicemos el resultado.

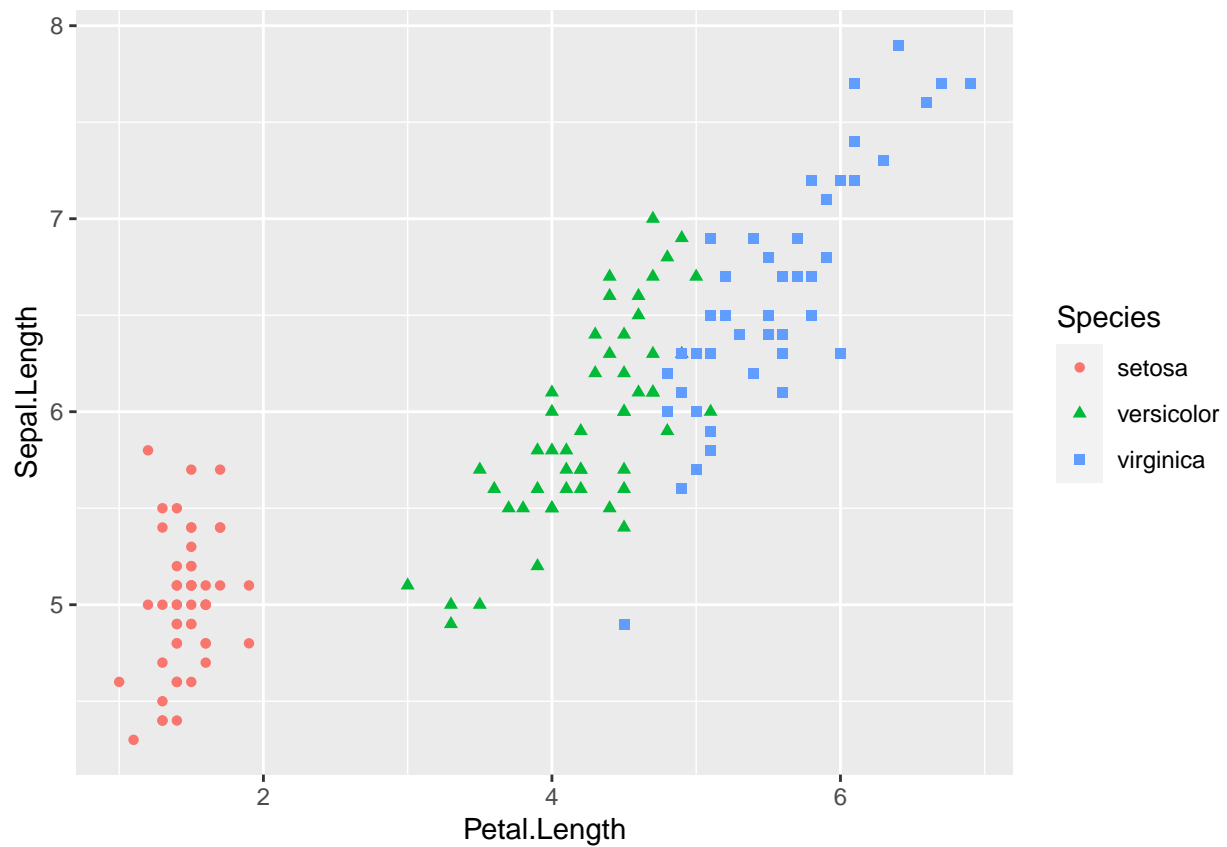
```
ggplot(data = mpg) +  
  geom_point(mapping = aes(x = displ, y = hwy))
```



Lo primero que hicimos fue definir los datos a utilizar en la función `ggplot()`. Siempre necesitamos iniciar una visualización con esta función, en ella podemos definir cosas básicas del gráfico, en este caso, los datos que se van a utilizar para graficar. Hasta ahora no hemos creado ninguna visualización.

Las visualizaciones como tal aparecen a la hora de agregar una función tipo `GEOM`. En este caso usamos `geom_point()` que nos sirve para hacer un scatter plot. Dentro de cada `geom` necesitamos definir un mapeo hacia los elementos del geom, que logramos con la función `aes()`. La cual amarra o “mappea” variables del `data.frame` a los elementos del gráfico. En el caso anterior lo hicimos solo con `x` y `y` pero podemos hacerlo también con una tercera variable para el color por ejemplo.

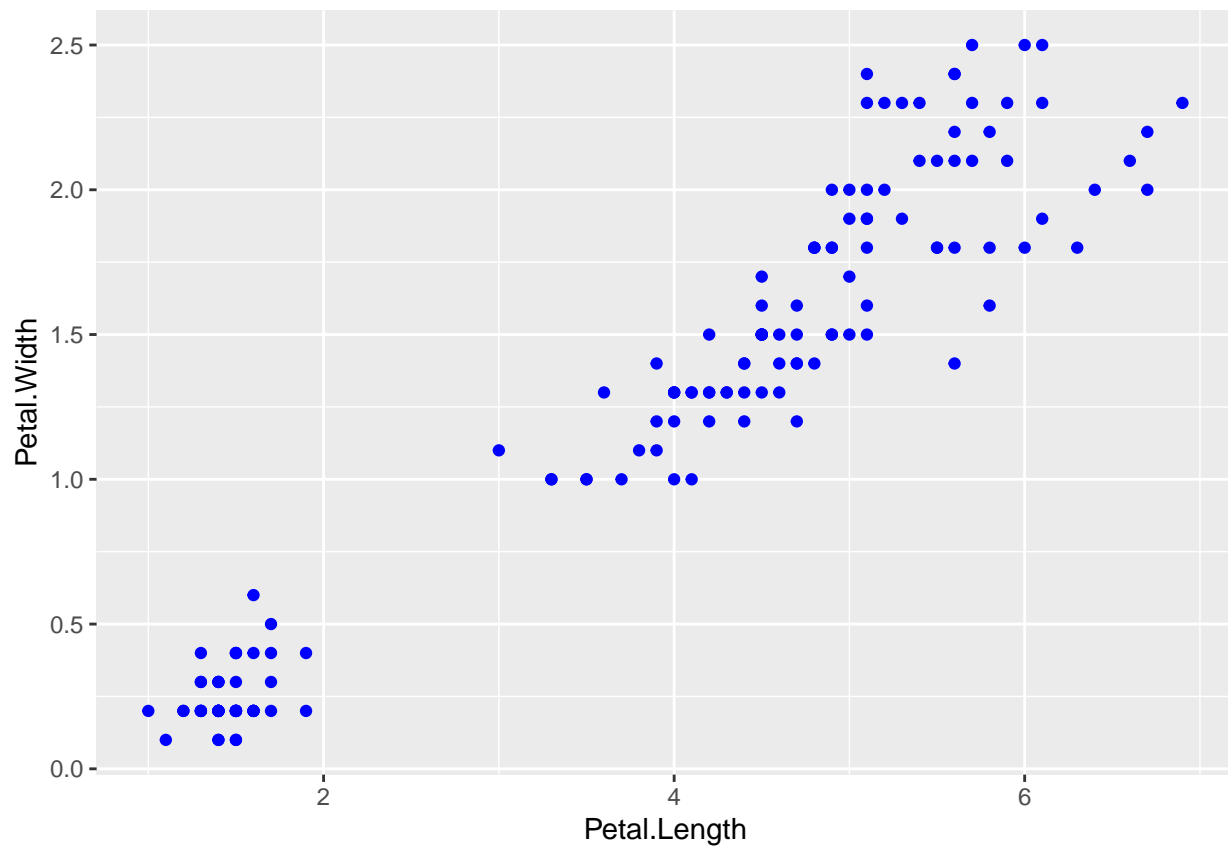
```
ggplot(data = iris) +  
  geom_point(mapping = aes(x = Petal.Length, y = Sepal.Length, color = Species, shape = Species))
```



Otros atributos que se pueden agregar al mapping son: **shape**, **alpha**, **size**

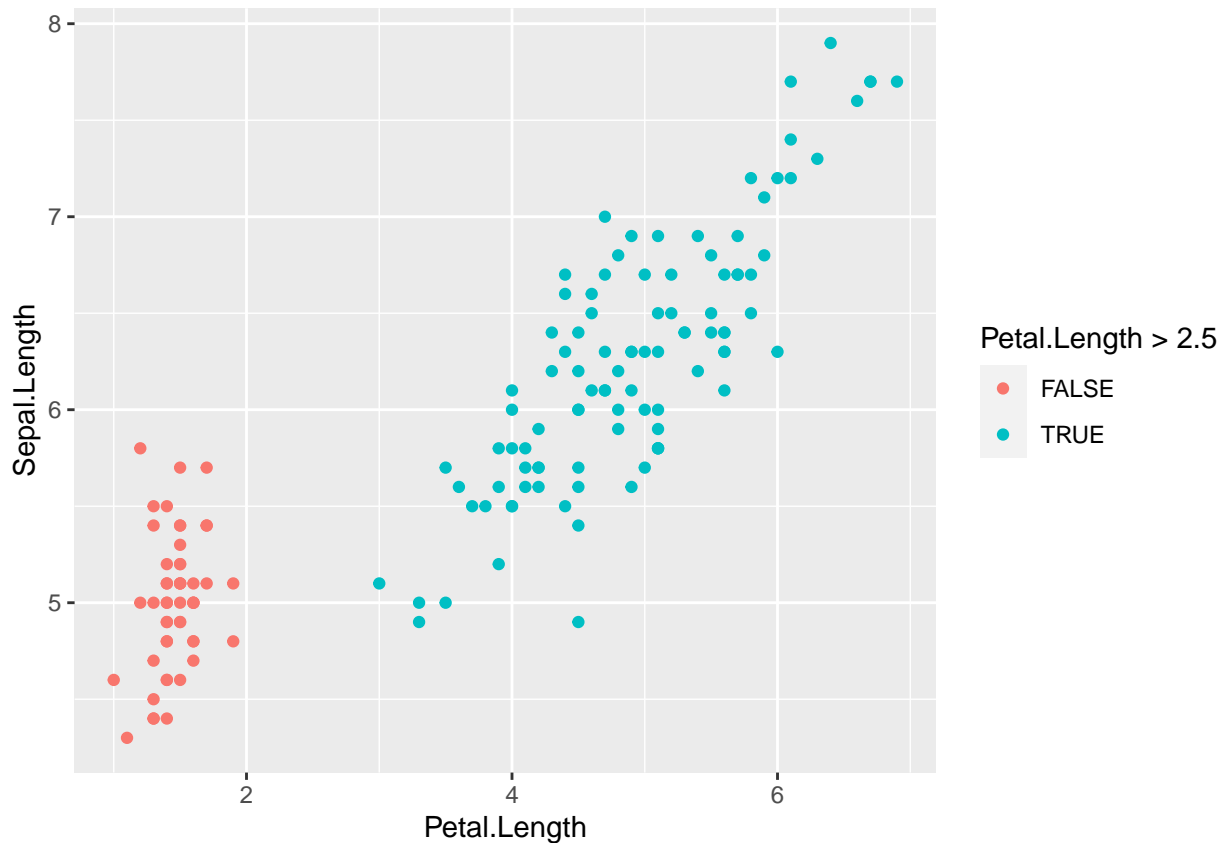
Estos atributos se pueden agregar fuera del mapping y entonces se aplicarán a cada punto por igual.

```
ggplot(data = iris) +  
  geom_point(mapping = aes(x = Petal.Length, y = Petal.Width), color = 'blue')
```



Podemos también agregar condiciones para visualizar,

```
ggplot(data = iris) +  
  geom_point(mapping = aes(x = Petal.Length, y = Sepal.Length, color = Petal.Length > 2.5))
```

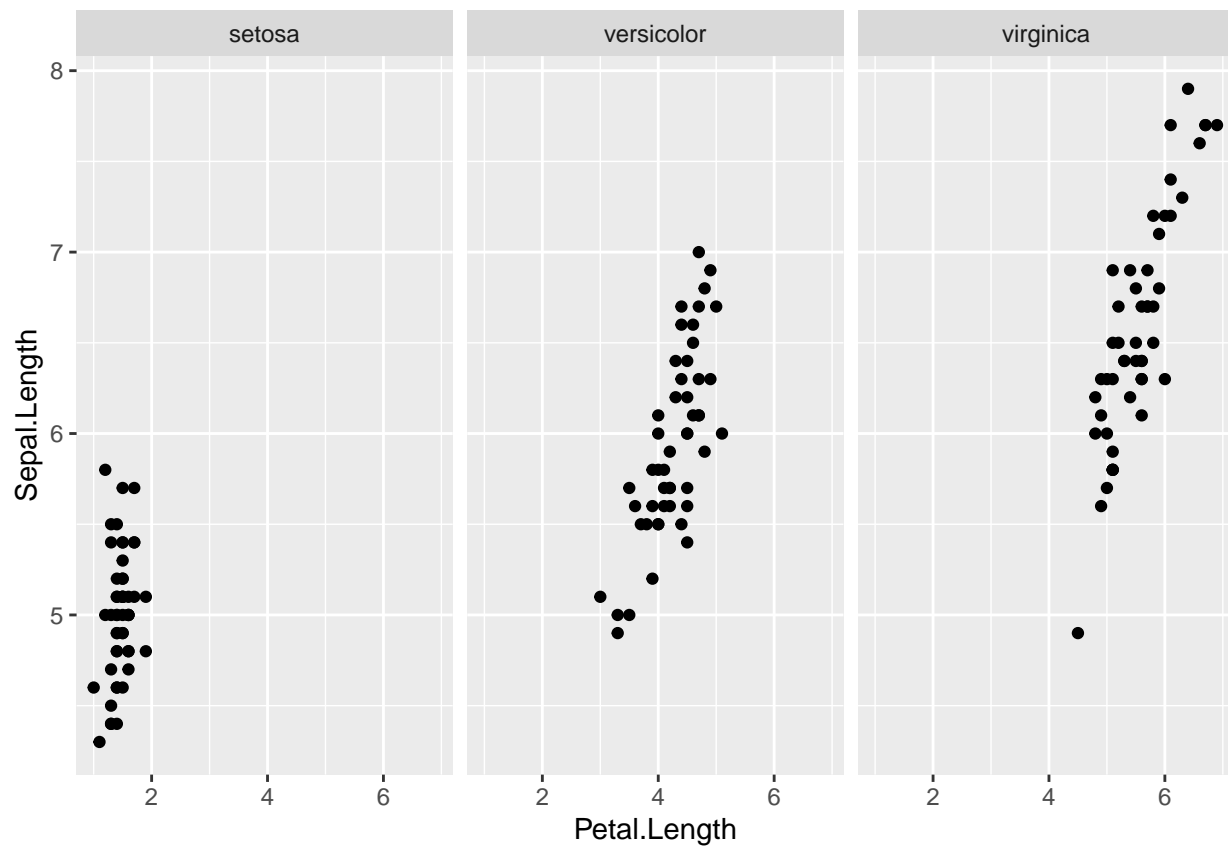


Facets

Los Facets nos permiten separar de manera sencilla el gráfico hay muchos tipos de facets pero concentrémonos en dos, `facet_wrap()` y `facet_grid()`

`facet_wrap()` solo necesita que definamos una variable con la cual separar los gráficos y hace el trabajo para nosotros, notarás que usamos el símbolo `~` para el primer argumento de la función. Por ahora solo necesitamos pensar en él como que le estamos indicando a la función que ella tendrá que cambiar en base a Species.

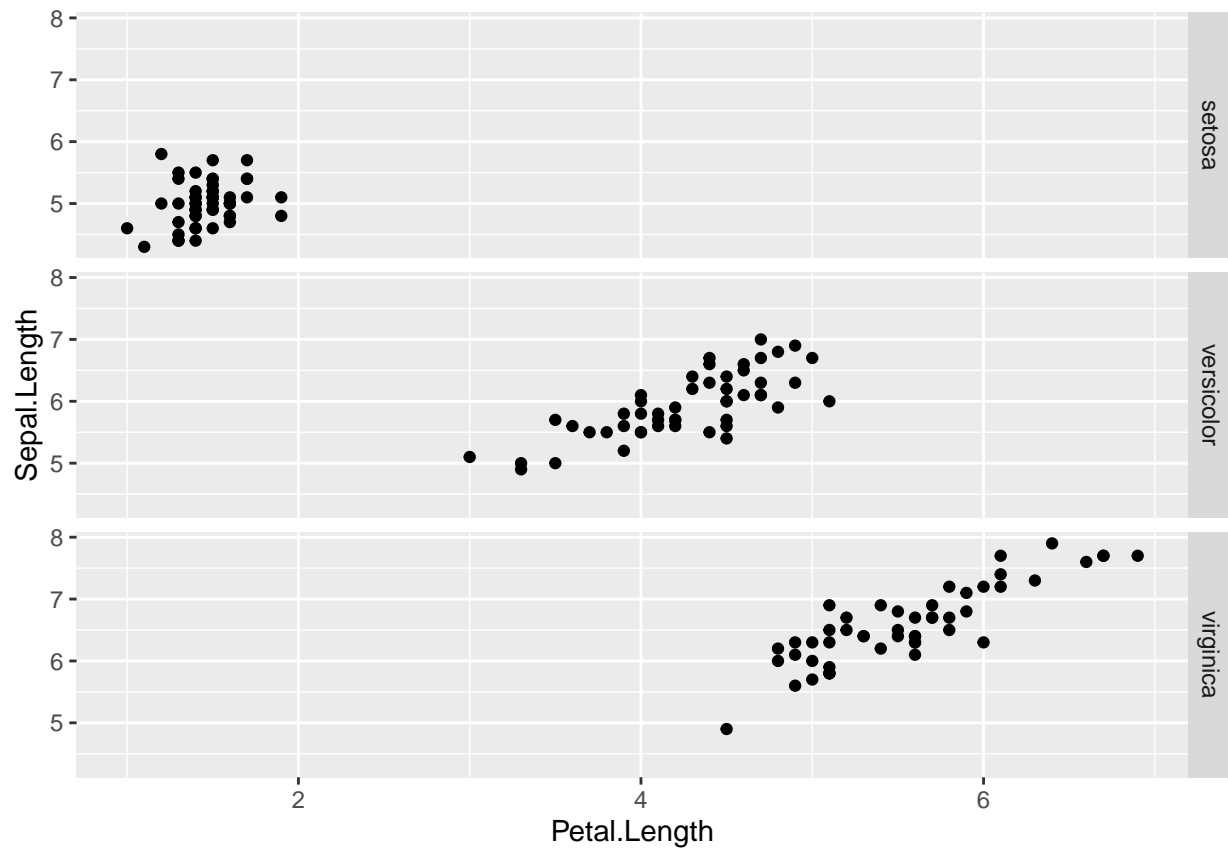
```
ggplot(data = iris) +  
  geom_point(mapping = aes(x = Petal.Length, y = Sepal.Length)) +  
  facet_wrap(~ Species, nrow = 1)
```



`facet_grid()` nos permite elegir el eje de la separación.

Veamos como podemos hacer la misma separación pero dividiendo por columnas,

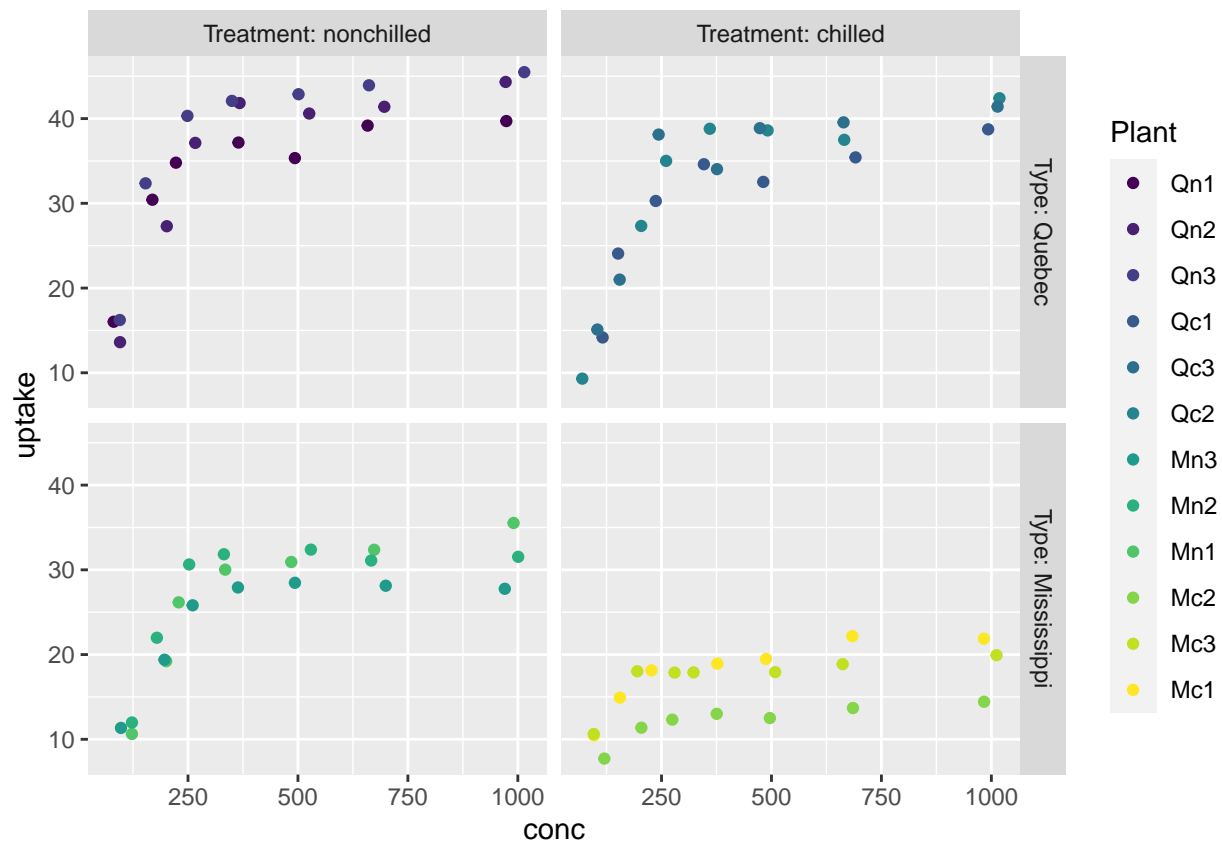
```
ggplot(data = iris) +  
  geom_point(mapping = aes(x = Petal.Length, y = Sepal.Length)) +  
  facet_grid(Species ~ .)
```



En general es buena idea separar un gráfico con base a variables categóricas, usemos otro dataset con más variables categóricas para visualizar esto,

```
data(CO2)

ggplot(data = CO2) +
  geom_point(aes(x = conc, y = uptake, color = Plant),
             position = 'jitter') +
  facet_grid(Type ~ Treatment, labeller = label_both)
```

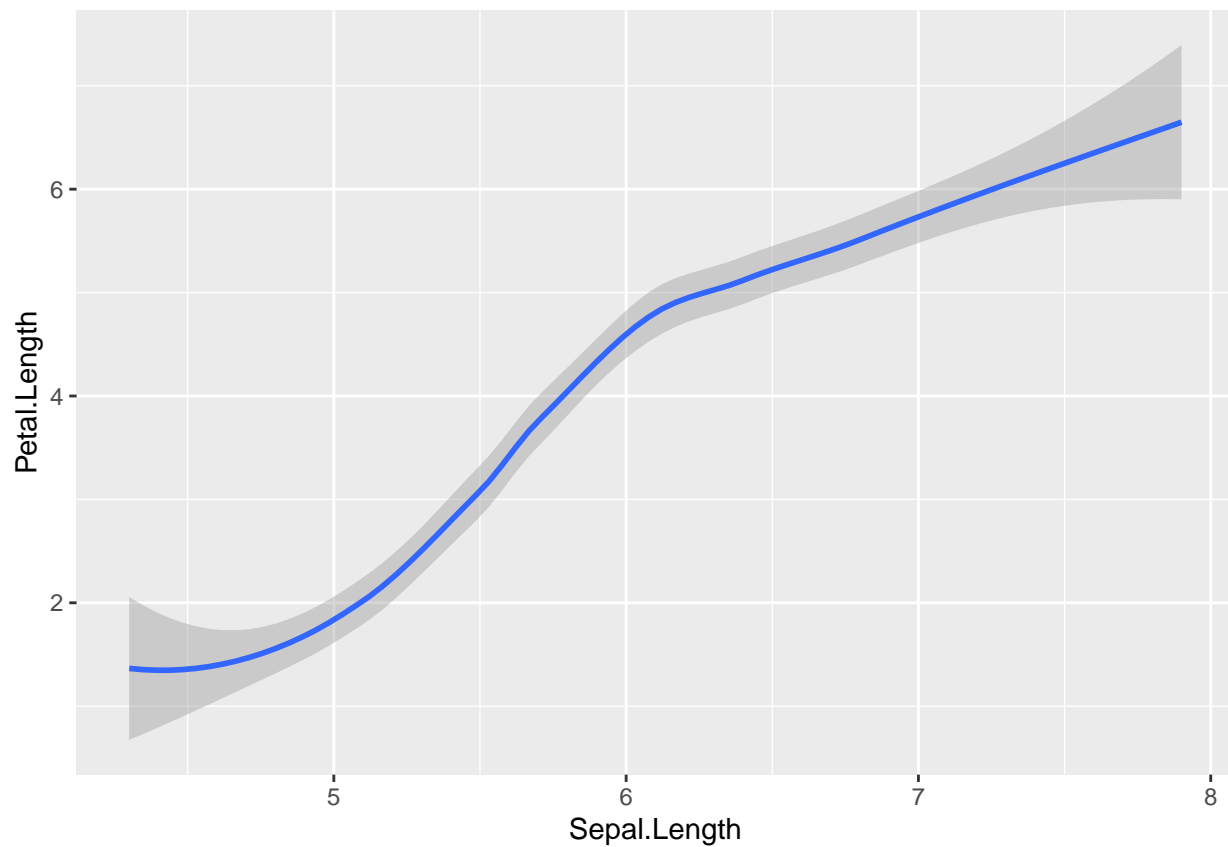


Otros tipos de geoms

Hasta ahora solo hemos utilizado `geom_point()` pero existen muchos otros, un ejemplo es `geom_smooth()` que dibuja una línea del promedio de un conjunto de puntos. En el siguiente ejemplo verás que los geoms funcionan de manera muy similar, en algunos casos puedes solo cambiar el nombre del geom y listo.

```
ggplot(data = iris) +
  geom_smooth(mapping = aes(x = Sepal.Length, y = Petal.Length))
```

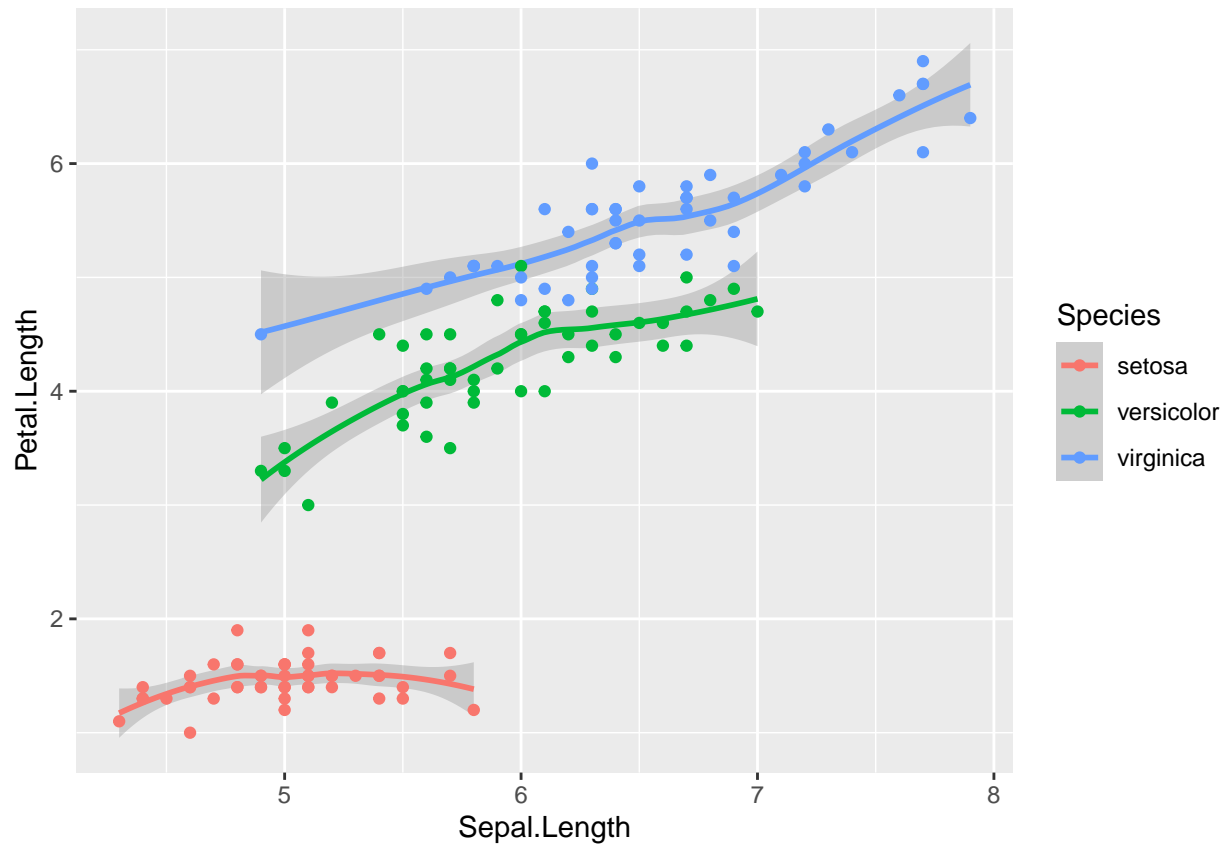
```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```

También hay ciertos mappings que funcionan con geoms y en otros no tienen sentido, un ejemplo con smooth sería `group` o `linetype`.

Otra cosa interesante es que podemos agregar más de un geom en un plot, solo hay que sumarlo

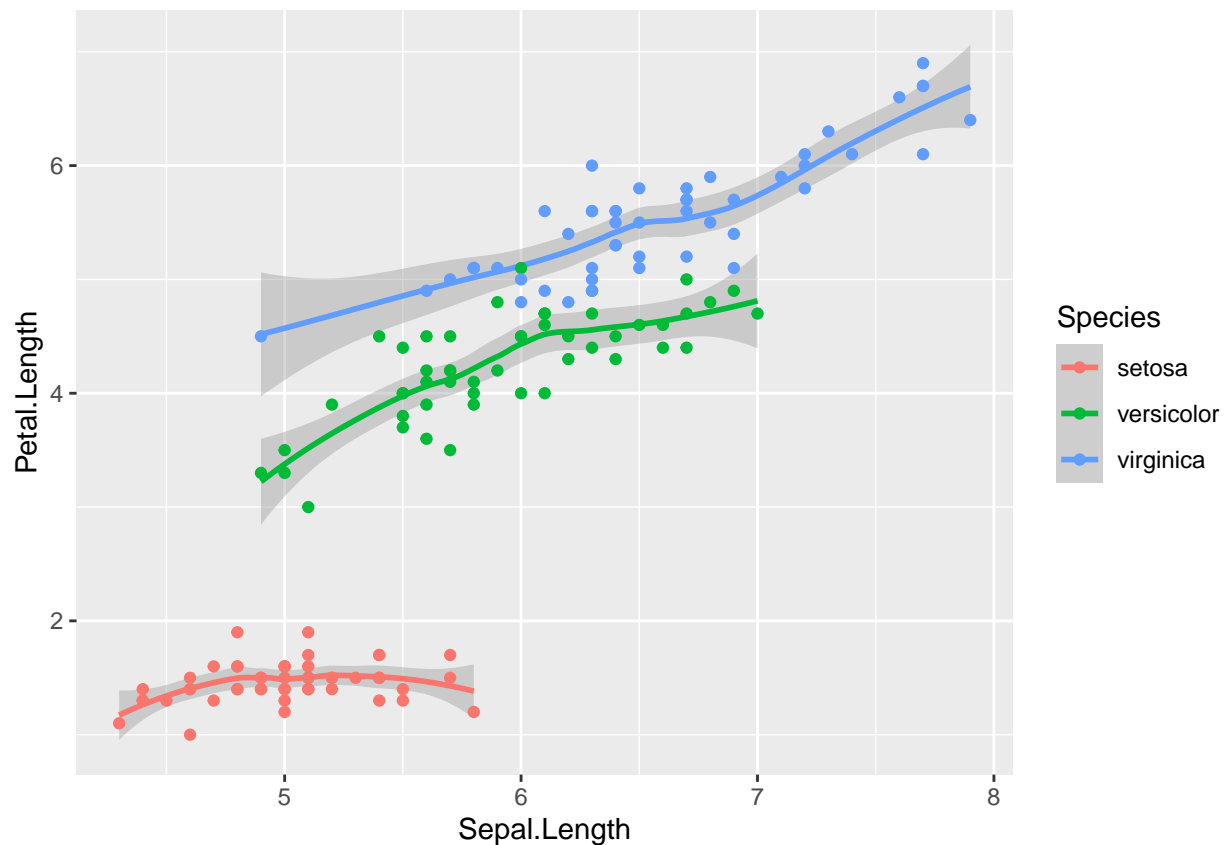
```
ggplot(data = iris) +  
  geom_smooth(mapping = aes(x = Sepal.Length, y = Petal.Length, group = Species, color = Species)) +  
  geom_point(mapping = aes(x = Sepal.Length, y = Petal.Length, color = Species))  
  
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



Tal vez te percataste que es un poco repetitivo tener que definir `x` y `y` en el mapping de cada geom, eso es cierto y por eso podemos también definir mapeos que compartirán los geoms desde que llamamos a `ggplot()`,

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

```
## `geom_smooth()` using method = 'loess' and formula 'y ~ x'
```



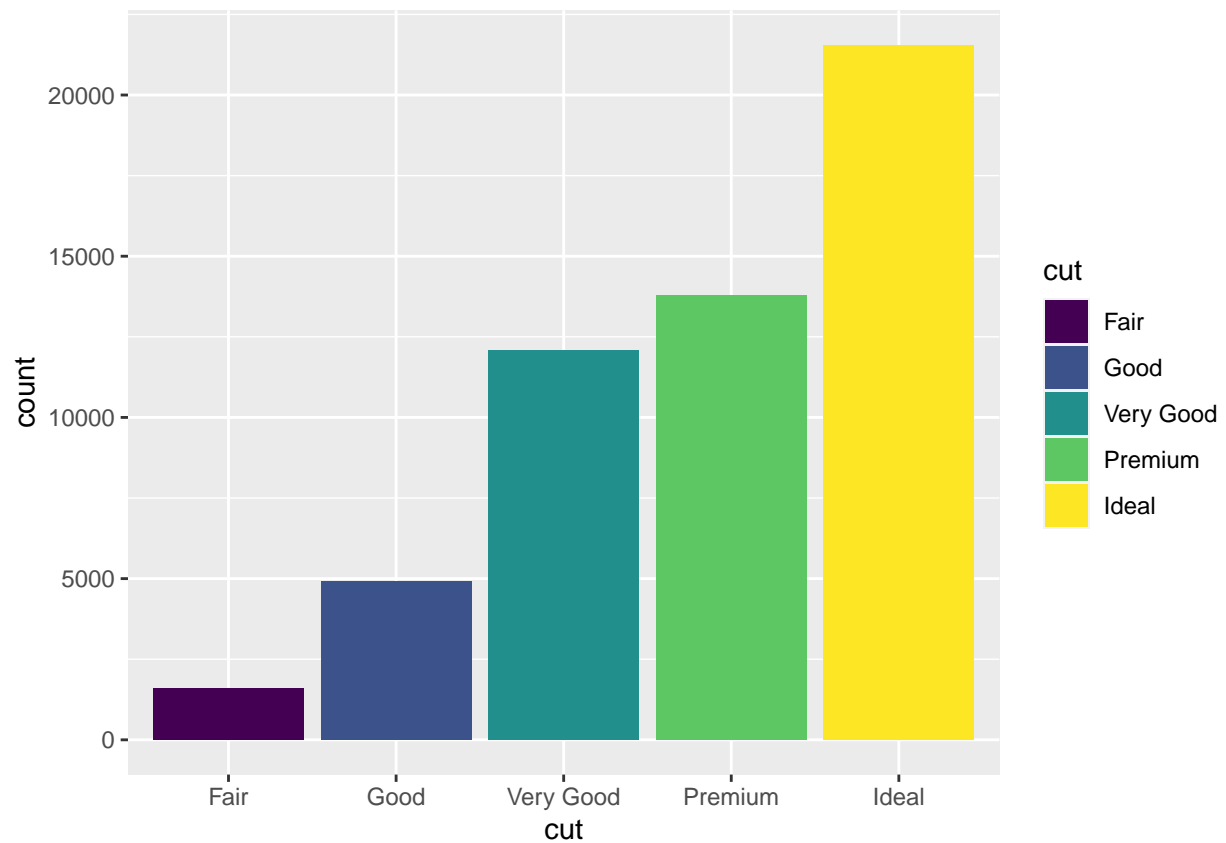
Stats

Algunos visualizaciones usan las variables presentes para generar nuevos datos como promedio, desviación estándar, mediana, conteo, entre otras de alguna variable. Los stats hacen precisamente eso

Un ejemplo común de gráfico que usa stats son los gráficos de barras e histogramas. Un gráfico de barras se suele usar con variables categóricas.

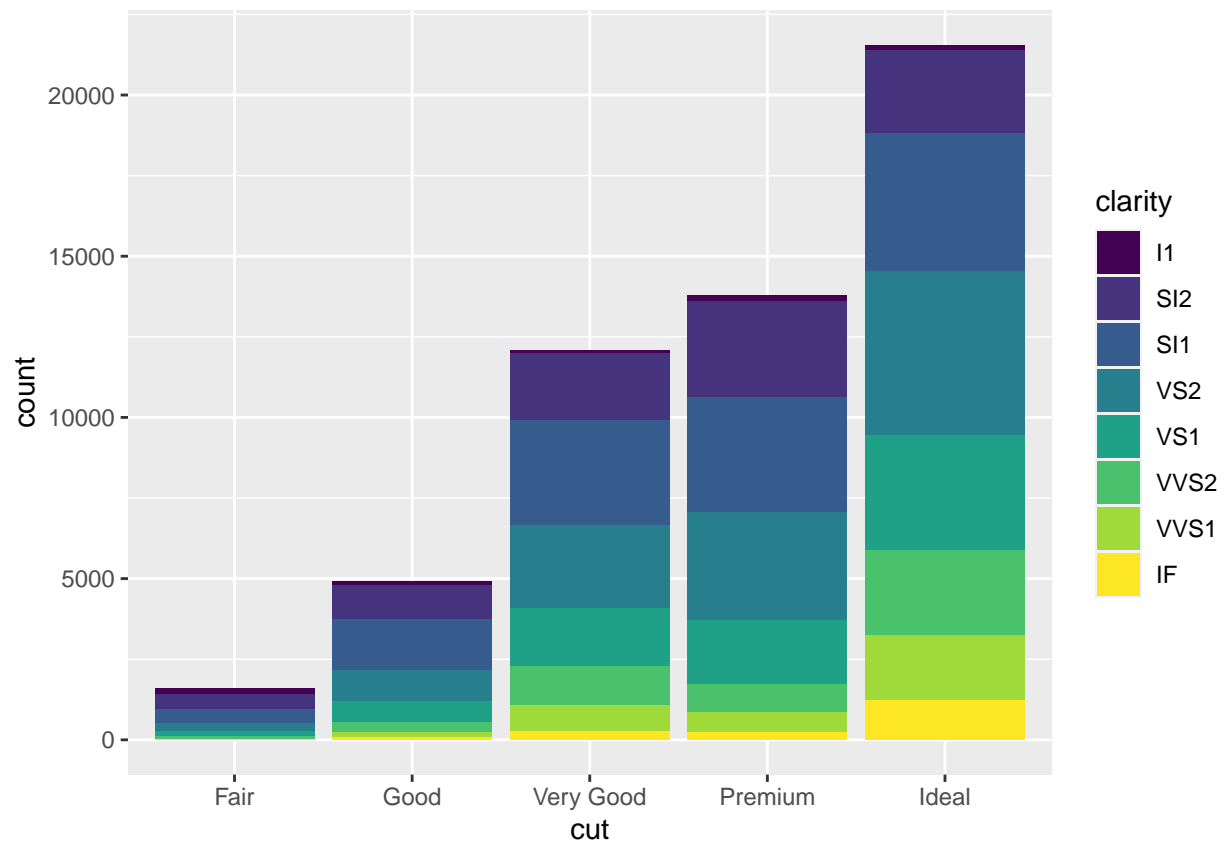
Nota que hay un línea de **geom** y otra de **stat** en general éstas son intercambiables y cambiarlas aquí produce el mismo resultado,

```
ggplot(data = diamonds) +  
  geom_bar(mapping = aes(x = cut, fill = cut))
```



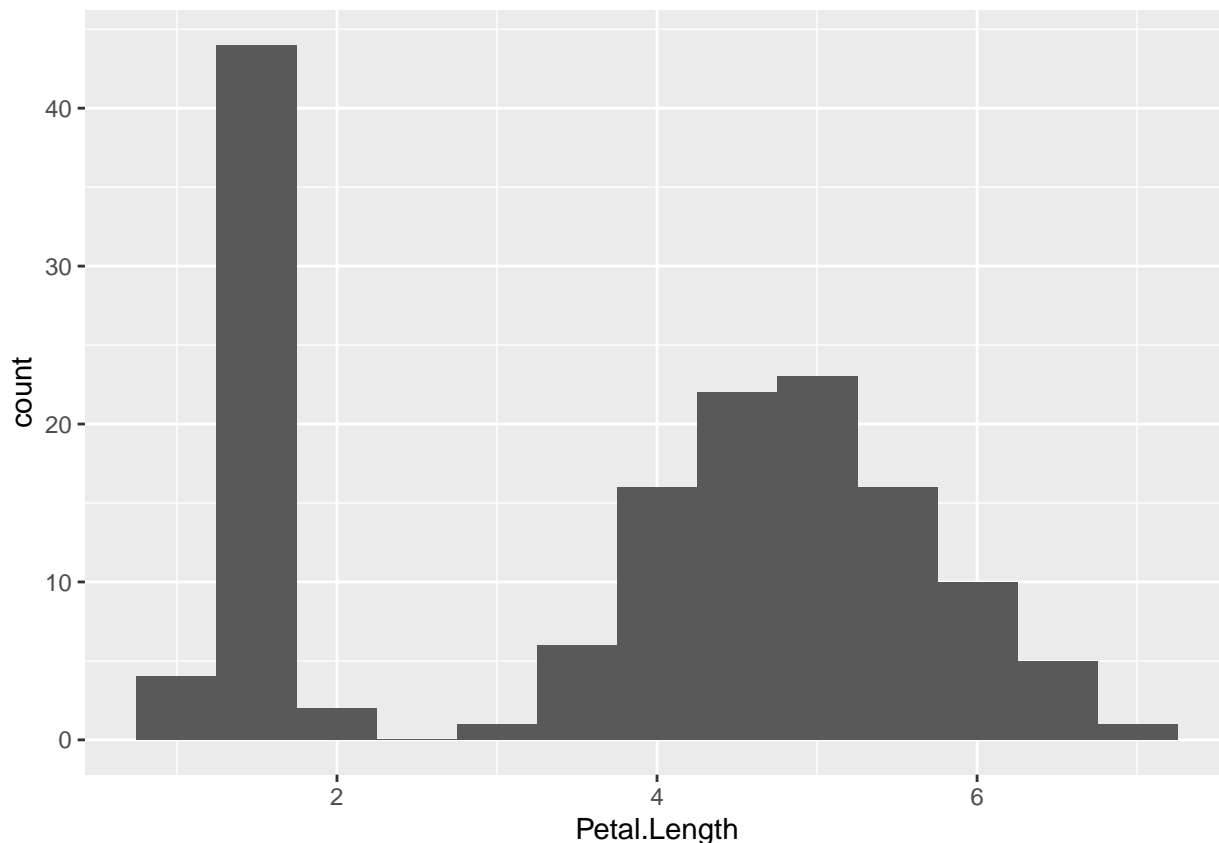
```
#stat_count(mapping = aes(x = cut, fill = cut))
```

```
ggplot(data = diamonds) +  
  geom_bar(mapping = aes(x = cut, fill = clarity))
```



Para variables continuas se puede usar el stat bin que separa la variable en segmentos. El geom histogram incluye el stat de bin en automático

```
ggplot(data = iris) +
  geom_histogram(mapping = aes(x = Petal.Length),
    binwidth = 0.5)
```



dplyr

La librería `dplyr` sirve para manejar `data.frames`, tenemos 3 métodos principales que revisar en este paquete,

1. `filter()`: quitar instancias (filas) con base a valores
2. `arrange()`: organizar según el valor
3. `select()`: seleccionar columnas
4. `mutate()/transmute()`: agregar nuevas columnas
5. `summarise()/groupby()`: agrupar y encontrar valores de tendencia

`filter()`

Como ya mencionamos con `filter()` podemos escoger solo algunos valores de un `data.frame` basados en el valor de alguna variable, usemos el dataset `flights`

Antes de filtrar veamos los datos originales, Iris es un dataset muy famoso para ciencia de datos, tenemos 4 variables numéricas (longitud y ancho de Pétalos y Sépalos) y una categórica (la especie de la planta: `setosa`, `virginica` y `versicolor`)

`iris`

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	5.1	3.5	1.4	0.2	setosa
## 2	4.9	3.0	1.4	0.2	setosa
## 3	4.7	3.2	1.3	0.2	setosa
## 4	4.6	3.1	1.5	0.2	setosa
## 5	5.0	3.6	1.4	0.2	setosa
## 6	5.4	3.9	1.7	0.4	setosa

## 7	4.6	3.4	1.4	0.3	setosa
## 8	5.0	3.4	1.5	0.2	setosa
## 9	4.4	2.9	1.4	0.2	setosa
## 10	4.9	3.1	1.5	0.1	setosa
## 11	5.4	3.7	1.5	0.2	setosa
## 12	4.8	3.4	1.6	0.2	setosa
## 13	4.8	3.0	1.4	0.1	setosa
## 14	4.3	3.0	1.1	0.1	setosa
## 15	5.8	4.0	1.2	0.2	setosa
## 16	5.7	4.4	1.5	0.4	setosa
## 17	5.4	3.9	1.3	0.4	setosa
## 18	5.1	3.5	1.4	0.3	setosa
## 19	5.7	3.8	1.7	0.3	setosa
## 20	5.1	3.8	1.5	0.3	setosa
## 21	5.4	3.4	1.7	0.2	setosa
## 22	5.1	3.7	1.5	0.4	setosa
## 23	4.6	3.6	1.0	0.2	setosa
## 24	5.1	3.3	1.7	0.5	setosa
## 25	4.8	3.4	1.9	0.2	setosa
## 26	5.0	3.0	1.6	0.2	setosa
## 27	5.0	3.4	1.6	0.4	setosa
## 28	5.2	3.5	1.5	0.2	setosa
## 29	5.2	3.4	1.4	0.2	setosa
## 30	4.7	3.2	1.6	0.2	setosa
## 31	4.8	3.1	1.6	0.2	setosa
## 32	5.4	3.4	1.5	0.4	setosa
## 33	5.2	4.1	1.5	0.1	setosa
## 34	5.5	4.2	1.4	0.2	setosa
## 35	4.9	3.1	1.5	0.2	setosa
## 36	5.0	3.2	1.2	0.2	setosa
## 37	5.5	3.5	1.3	0.2	setosa
## 38	4.9	3.6	1.4	0.1	setosa
## 39	4.4	3.0	1.3	0.2	setosa
## 40	5.1	3.4	1.5	0.2	setosa
## 41	5.0	3.5	1.3	0.3	setosa
## 42	4.5	2.3	1.3	0.3	setosa
## 43	4.4	3.2	1.3	0.2	setosa
## 44	5.0	3.5	1.6	0.6	setosa
## 45	5.1	3.8	1.9	0.4	setosa
## 46	4.8	3.0	1.4	0.3	setosa
## 47	5.1	3.8	1.6	0.2	setosa
## 48	4.6	3.2	1.4	0.2	setosa
## 49	5.3	3.7	1.5	0.2	setosa
## 50	5.0	3.3	1.4	0.2	setosa
## 51	7.0	3.2	4.7	1.4	versicolor
## 52	6.4	3.2	4.5	1.5	versicolor
## 53	6.9	3.1	4.9	1.5	versicolor
## 54	5.5	2.3	4.0	1.3	versicolor
## 55	6.5	2.8	4.6	1.5	versicolor
## 56	5.7	2.8	4.5	1.3	versicolor
## 57	6.3	3.3	4.7	1.6	versicolor
## 58	4.9	2.4	3.3	1.0	versicolor
## 59	6.6	2.9	4.6	1.3	versicolor
## 60	5.2	2.7	3.9	1.4	versicolor

## 61	5.0	2.0	3.5	1.0 versicolor
## 62	5.9	3.0	4.2	1.5 versicolor
## 63	6.0	2.2	4.0	1.0 versicolor
## 64	6.1	2.9	4.7	1.4 versicolor
## 65	5.6	2.9	3.6	1.3 versicolor
## 66	6.7	3.1	4.4	1.4 versicolor
## 67	5.6	3.0	4.5	1.5 versicolor
## 68	5.8	2.7	4.1	1.0 versicolor
## 69	6.2	2.2	4.5	1.5 versicolor
## 70	5.6	2.5	3.9	1.1 versicolor
## 71	5.9	3.2	4.8	1.8 versicolor
## 72	6.1	2.8	4.0	1.3 versicolor
## 73	6.3	2.5	4.9	1.5 versicolor
## 74	6.1	2.8	4.7	1.2 versicolor
## 75	6.4	2.9	4.3	1.3 versicolor
## 76	6.6	3.0	4.4	1.4 versicolor
## 77	6.8	2.8	4.8	1.4 versicolor
## 78	6.7	3.0	5.0	1.7 versicolor
## 79	6.0	2.9	4.5	1.5 versicolor
## 80	5.7	2.6	3.5	1.0 versicolor
## 81	5.5	2.4	3.8	1.1 versicolor
## 82	5.5	2.4	3.7	1.0 versicolor
## 83	5.8	2.7	3.9	1.2 versicolor
## 84	6.0	2.7	5.1	1.6 versicolor
## 85	5.4	3.0	4.5	1.5 versicolor
## 86	6.0	3.4	4.5	1.6 versicolor
## 87	6.7	3.1	4.7	1.5 versicolor
## 88	6.3	2.3	4.4	1.3 versicolor
## 89	5.6	3.0	4.1	1.3 versicolor
## 90	5.5	2.5	4.0	1.3 versicolor
## 91	5.5	2.6	4.4	1.2 versicolor
## 92	6.1	3.0	4.6	1.4 versicolor
## 93	5.8	2.6	4.0	1.2 versicolor
## 94	5.0	2.3	3.3	1.0 versicolor
## 95	5.6	2.7	4.2	1.3 versicolor
## 96	5.7	3.0	4.2	1.2 versicolor
## 97	5.7	2.9	4.2	1.3 versicolor
## 98	6.2	2.9	4.3	1.3 versicolor
## 99	5.1	2.5	3.0	1.1 versicolor
## 100	5.7	2.8	4.1	1.3 versicolor
## 101	6.3	3.3	6.0	2.5 virginica
## 102	5.8	2.7	5.1	1.9 virginica
## 103	7.1	3.0	5.9	2.1 virginica
## 104	6.3	2.9	5.6	1.8 virginica
## 105	6.5	3.0	5.8	2.2 virginica
## 106	7.6	3.0	6.6	2.1 virginica
## 107	4.9	2.5	4.5	1.7 virginica
## 108	7.3	2.9	6.3	1.8 virginica
## 109	6.7	2.5	5.8	1.8 virginica
## 110	7.2	3.6	6.1	2.5 virginica
## 111	6.5	3.2	5.1	2.0 virginica
## 112	6.4	2.7	5.3	1.9 virginica
## 113	6.8	3.0	5.5	2.1 virginica
## 114	5.7	2.5	5.0	2.0 virginica

## 115	5.8	2.8	5.1	2.4	virginica
## 116	6.4	3.2	5.3	2.3	virginica
## 117	6.5	3.0	5.5	1.8	virginica
## 118	7.7	3.8	6.7	2.2	virginica
## 119	7.7	2.6	6.9	2.3	virginica
## 120	6.0	2.2	5.0	1.5	virginica
## 121	6.9	3.2	5.7	2.3	virginica
## 122	5.6	2.8	4.9	2.0	virginica
## 123	7.7	2.8	6.7	2.0	virginica
## 124	6.3	2.7	4.9	1.8	virginica
## 125	6.7	3.3	5.7	2.1	virginica
## 126	7.2	3.2	6.0	1.8	virginica
## 127	6.2	2.8	4.8	1.8	virginica
## 128	6.1	3.0	4.9	1.8	virginica
## 129	6.4	2.8	5.6	2.1	virginica
## 130	7.2	3.0	5.8	1.6	virginica
## 131	7.4	2.8	6.1	1.9	virginica
## 132	7.9	3.8	6.4	2.0	virginica
## 133	6.4	2.8	5.6	2.2	virginica
## 134	6.3	2.8	5.1	1.5	virginica
## 135	6.1	2.6	5.6	1.4	virginica
## 136	7.7	3.0	6.1	2.3	virginica
## 137	6.3	3.4	5.6	2.4	virginica
## 138	6.4	3.1	5.5	1.8	virginica
## 139	6.0	3.0	4.8	1.8	virginica
## 140	6.9	3.1	5.4	2.1	virginica
## 141	6.7	3.1	5.6	2.4	virginica
## 142	6.9	3.1	5.1	2.3	virginica
## 143	5.8	2.7	5.1	1.9	virginica
## 144	6.8	3.2	5.9	2.3	virginica
## 145	6.7	3.3	5.7	2.5	virginica
## 146	6.7	3.0	5.2	2.3	virginica
## 147	6.3	2.5	5.0	1.9	virginica
## 148	6.5	3.0	5.2	2.0	virginica
## 149	6.2	3.4	5.4	2.3	virginica
## 150	5.9	3.0	5.1	1.8	virginica

Nota que tenemos 150 instancias.

Ahora usemos la función, su forma básica es,

```
filter(<DATA>, <CRITERIA>)
```

```
(iris_g7 <- filter(iris, Sepal.Length > 7))
```

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	7.1	3.0	5.9	2.1	virginica
## 2	7.6	3.0	6.6	2.1	virginica
## 3	7.3	2.9	6.3	1.8	virginica
## 4	7.2	3.6	6.1	2.5	virginica
## 5	7.7	3.8	6.7	2.2	virginica
## 6	7.7	2.6	6.9	2.3	virginica
## 7	7.7	2.8	6.7	2.0	virginica
## 8	7.2	3.2	6.0	1.8	virginica
## 9	7.2	3.0	5.8	1.6	virginica
## 10	7.4	2.8	6.1	1.9	virginica

```
## 11      7.9      3.8      6.4      2.0 virginica
## 12      7.7      3.0      6.1      2.3 virginica
```

Ahora que filtramos para obtener sólo las instancias con Sepal.Length mayor a 7 y vemos que tenemos solo 12 instancias ahora. Podemos con comas añadir nuevas condiciones o juntarlas con operadores lógicos como and y or.

Otra opción para filtrar es el comando %in% el cual usamos para encontrar todos los valores en un vector, en este caso lo utilizamos para encontrar los ejemplos de iris setosa o virginica.

```
filter(iris, Species %in% c('setosa','virginica'))
```

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	5.1	3.5	1.4	0.2	setosa
## 2	4.9	3.0	1.4	0.2	setosa
## 3	4.7	3.2	1.3	0.2	setosa
## 4	4.6	3.1	1.5	0.2	setosa
## 5	5.0	3.6	1.4	0.2	setosa
## 6	5.4	3.9	1.7	0.4	setosa
## 7	4.6	3.4	1.4	0.3	setosa
## 8	5.0	3.4	1.5	0.2	setosa
## 9	4.4	2.9	1.4	0.2	setosa
## 10	4.9	3.1	1.5	0.1	setosa
## 11	5.4	3.7	1.5	0.2	setosa
## 12	4.8	3.4	1.6	0.2	setosa
## 13	4.8	3.0	1.4	0.1	setosa
## 14	4.3	3.0	1.1	0.1	setosa
## 15	5.8	4.0	1.2	0.2	setosa
## 16	5.7	4.4	1.5	0.4	setosa
## 17	5.4	3.9	1.3	0.4	setosa
## 18	5.1	3.5	1.4	0.3	setosa
## 19	5.7	3.8	1.7	0.3	setosa
## 20	5.1	3.8	1.5	0.3	setosa
## 21	5.4	3.4	1.7	0.2	setosa
## 22	5.1	3.7	1.5	0.4	setosa
## 23	4.6	3.6	1.0	0.2	setosa
## 24	5.1	3.3	1.7	0.5	setosa
## 25	4.8	3.4	1.9	0.2	setosa
## 26	5.0	3.0	1.6	0.2	setosa
## 27	5.0	3.4	1.6	0.4	setosa
## 28	5.2	3.5	1.5	0.2	setosa
## 29	5.2	3.4	1.4	0.2	setosa
## 30	4.7	3.2	1.6	0.2	setosa
## 31	4.8	3.1	1.6	0.2	setosa
## 32	5.4	3.4	1.5	0.4	setosa
## 33	5.2	4.1	1.5	0.1	setosa
## 34	5.5	4.2	1.4	0.2	setosa
## 35	4.9	3.1	1.5	0.2	setosa
## 36	5.0	3.2	1.2	0.2	setosa
## 37	5.5	3.5	1.3	0.2	setosa
## 38	4.9	3.6	1.4	0.1	setosa
## 39	4.4	3.0	1.3	0.2	setosa
## 40	5.1	3.4	1.5	0.2	setosa
## 41	5.0	3.5	1.3	0.3	setosa
## 42	4.5	2.3	1.3	0.3	setosa
## 43	4.4	3.2	1.3	0.2	setosa

## 44	5.0	3.5	1.6	0.6	setosa
## 45	5.1	3.8	1.9	0.4	setosa
## 46	4.8	3.0	1.4	0.3	setosa
## 47	5.1	3.8	1.6	0.2	setosa
## 48	4.6	3.2	1.4	0.2	setosa
## 49	5.3	3.7	1.5	0.2	setosa
## 50	5.0	3.3	1.4	0.2	setosa
## 51	6.3	3.3	6.0	2.5	virginica
## 52	5.8	2.7	5.1	1.9	virginica
## 53	7.1	3.0	5.9	2.1	virginica
## 54	6.3	2.9	5.6	1.8	virginica
## 55	6.5	3.0	5.8	2.2	virginica
## 56	7.6	3.0	6.6	2.1	virginica
## 57	4.9	2.5	4.5	1.7	virginica
## 58	7.3	2.9	6.3	1.8	virginica
## 59	6.7	2.5	5.8	1.8	virginica
## 60	7.2	3.6	6.1	2.5	virginica
## 61	6.5	3.2	5.1	2.0	virginica
## 62	6.4	2.7	5.3	1.9	virginica
## 63	6.8	3.0	5.5	2.1	virginica
## 64	5.7	2.5	5.0	2.0	virginica
## 65	5.8	2.8	5.1	2.4	virginica
## 66	6.4	3.2	5.3	2.3	virginica
## 67	6.5	3.0	5.5	1.8	virginica
## 68	7.7	3.8	6.7	2.2	virginica
## 69	7.7	2.6	6.9	2.3	virginica
## 70	6.0	2.2	5.0	1.5	virginica
## 71	6.9	3.2	5.7	2.3	virginica
## 72	5.6	2.8	4.9	2.0	virginica
## 73	7.7	2.8	6.7	2.0	virginica
## 74	6.3	2.7	4.9	1.8	virginica
## 75	6.7	3.3	5.7	2.1	virginica
## 76	7.2	3.2	6.0	1.8	virginica
## 77	6.2	2.8	4.8	1.8	virginica
## 78	6.1	3.0	4.9	1.8	virginica
## 79	6.4	2.8	5.6	2.1	virginica
## 80	7.2	3.0	5.8	1.6	virginica
## 81	7.4	2.8	6.1	1.9	virginica
## 82	7.9	3.8	6.4	2.0	virginica
## 83	6.4	2.8	5.6	2.2	virginica
## 84	6.3	2.8	5.1	1.5	virginica
## 85	6.1	2.6	5.6	1.4	virginica
## 86	7.7	3.0	6.1	2.3	virginica
## 87	6.3	3.4	5.6	2.4	virginica
## 88	6.4	3.1	5.5	1.8	virginica
## 89	6.0	3.0	4.8	1.8	virginica
## 90	6.9	3.1	5.4	2.1	virginica
## 91	6.7	3.1	5.6	2.4	virginica
## 92	6.9	3.1	5.1	2.3	virginica
## 93	5.8	2.7	5.1	1.9	virginica
## 94	6.8	3.2	5.9	2.3	virginica
## 95	6.7	3.3	5.7	2.5	virginica
## 96	6.7	3.0	5.2	2.3	virginica
## 97	6.3	2.5	5.0	1.9	virginica

```
## 98      6.5      3.0      5.2      2.0 virginica
## 99      6.2      3.4      5.4      2.3 virginica
## 100     5.9      3.0      5.1      1.8 virginica
```

Ejercicios Filter

```
flights
```

```
## # A tibble: 336,776 x 19
##   year month   day dep_time sched_dep_time dep_delay arr_time sched_arr_time
##   <int> <int> <int>   <int>         <int>       <dbl>   <int>         <int>
## 1  2013     1     1     517           515         2       830           819
## 2  2013     1     1     533           529         4       850           830
## 3  2013     1     1     542           540         2       923           850
## 4  2013     1     1     544           545        -1      1004          1022
## 5  2013     1     1     554           600        -6       812           837
## 6  2013     1     1     554           558        -4       740           728
## 7  2013     1     1     555           600        -5       913           854
## 8  2013     1     1     557           600        -3       709           723
## 9  2013     1     1     557           600        -3       838           846
## 10 2013     1     1     558           600        -2       753           745
## # ... with 336,766 more rows, and 11 more variables: arr_delay <dbl>,
## #   carrier <chr>, flight <int>, tailnum <chr>, origin <chr>, dest <chr>,
## #   air_time <dbl>, distance <dbl>, hour <dbl>, minute <dbl>, time_hour <dtm>
```

`arrange()`

Podemos usar `arrange` para organizar por una variable.

```
arrange(iris, Sepal.Length)
```

```
##   Sepal.Length Sepal.Width Petal.Length Petal.Width Species
## 1           4.3          3.0          1.1          0.1   setosa
## 2           4.4          2.9          1.4          0.2   setosa
## 3           4.4          3.0          1.3          0.2   setosa
## 4           4.4          3.2          1.3          0.2   setosa
## 5           4.5          2.3          1.3          0.3   setosa
## 6           4.6          3.1          1.5          0.2   setosa
## 7           4.6          3.4          1.4          0.3   setosa
## 8           4.6          3.6          1.0          0.2   setosa
## 9           4.6          3.2          1.4          0.2   setosa
## 10          4.7          3.2          1.3          0.2   setosa
## 11          4.7          3.2          1.6          0.2   setosa
## 12          4.8          3.4          1.6          0.2   setosa
## 13          4.8          3.0          1.4          0.1   setosa
## 14          4.8          3.4          1.9          0.2   setosa
## 15          4.8          3.1          1.6          0.2   setosa
## 16          4.8          3.0          1.4          0.3   setosa
## 17          4.9          3.0          1.4          0.2   setosa
## 18          4.9          3.1          1.5          0.1   setosa
## 19          4.9          3.1          1.5          0.2   setosa
## 20          4.9          3.6          1.4          0.1   setosa
## 21          4.9          2.4          3.3          1.0 versicolor
## 22          4.9          2.5          4.5          1.7   virginica
## 23          5.0          3.6          1.4          0.2   setosa
## 24          5.0          3.4          1.5          0.2   setosa
```

## 25	5.0	3.0	1.6	0.2	setosa
## 26	5.0	3.4	1.6	0.4	setosa
## 27	5.0	3.2	1.2	0.2	setosa
## 28	5.0	3.5	1.3	0.3	setosa
## 29	5.0	3.5	1.6	0.6	setosa
## 30	5.0	3.3	1.4	0.2	setosa
## 31	5.0	2.0	3.5	1.0	versicolor
## 32	5.0	2.3	3.3	1.0	versicolor
## 33	5.1	3.5	1.4	0.2	setosa
## 34	5.1	3.5	1.4	0.3	setosa
## 35	5.1	3.8	1.5	0.3	setosa
## 36	5.1	3.7	1.5	0.4	setosa
## 37	5.1	3.3	1.7	0.5	setosa
## 38	5.1	3.4	1.5	0.2	setosa
## 39	5.1	3.8	1.9	0.4	setosa
## 40	5.1	3.8	1.6	0.2	setosa
## 41	5.1	2.5	3.0	1.1	versicolor
## 42	5.2	3.5	1.5	0.2	setosa
## 43	5.2	3.4	1.4	0.2	setosa
## 44	5.2	4.1	1.5	0.1	setosa
## 45	5.2	2.7	3.9	1.4	versicolor
## 46	5.3	3.7	1.5	0.2	setosa
## 47	5.4	3.9	1.7	0.4	setosa
## 48	5.4	3.7	1.5	0.2	setosa
## 49	5.4	3.9	1.3	0.4	setosa
## 50	5.4	3.4	1.7	0.2	setosa
## 51	5.4	3.4	1.5	0.4	setosa
## 52	5.4	3.0	4.5	1.5	versicolor
## 53	5.5	4.2	1.4	0.2	setosa
## 54	5.5	3.5	1.3	0.2	setosa
## 55	5.5	2.3	4.0	1.3	versicolor
## 56	5.5	2.4	3.8	1.1	versicolor
## 57	5.5	2.4	3.7	1.0	versicolor
## 58	5.5	2.5	4.0	1.3	versicolor
## 59	5.5	2.6	4.4	1.2	versicolor
## 60	5.6	2.9	3.6	1.3	versicolor
## 61	5.6	3.0	4.5	1.5	versicolor
## 62	5.6	2.5	3.9	1.1	versicolor
## 63	5.6	3.0	4.1	1.3	versicolor
## 64	5.6	2.7	4.2	1.3	versicolor
## 65	5.6	2.8	4.9	2.0	virginica
## 66	5.7	4.4	1.5	0.4	setosa
## 67	5.7	3.8	1.7	0.3	setosa
## 68	5.7	2.8	4.5	1.3	versicolor
## 69	5.7	2.6	3.5	1.0	versicolor
## 70	5.7	3.0	4.2	1.2	versicolor
## 71	5.7	2.9	4.2	1.3	versicolor
## 72	5.7	2.8	4.1	1.3	versicolor
## 73	5.7	2.5	5.0	2.0	virginica
## 74	5.8	4.0	1.2	0.2	setosa
## 75	5.8	2.7	4.1	1.0	versicolor
## 76	5.8	2.7	3.9	1.2	versicolor
## 77	5.8	2.6	4.0	1.2	versicolor
## 78	5.8	2.7	5.1	1.9	virginica

## 79	5.8	2.8	5.1	2.4 virginica
## 80	5.8	2.7	5.1	1.9 virginica
## 81	5.9	3.0	4.2	1.5 versicolor
## 82	5.9	3.2	4.8	1.8 versicolor
## 83	5.9	3.0	5.1	1.8 virginica
## 84	6.0	2.2	4.0	1.0 versicolor
## 85	6.0	2.9	4.5	1.5 versicolor
## 86	6.0	2.7	5.1	1.6 versicolor
## 87	6.0	3.4	4.5	1.6 versicolor
## 88	6.0	2.2	5.0	1.5 virginica
## 89	6.0	3.0	4.8	1.8 virginica
## 90	6.1	2.9	4.7	1.4 versicolor
## 91	6.1	2.8	4.0	1.3 versicolor
## 92	6.1	2.8	4.7	1.2 versicolor
## 93	6.1	3.0	4.6	1.4 versicolor
## 94	6.1	3.0	4.9	1.8 virginica
## 95	6.1	2.6	5.6	1.4 virginica
## 96	6.2	2.2	4.5	1.5 versicolor
## 97	6.2	2.9	4.3	1.3 versicolor
## 98	6.2	2.8	4.8	1.8 virginica
## 99	6.2	3.4	5.4	2.3 virginica
## 100	6.3	3.3	4.7	1.6 versicolor
## 101	6.3	2.5	4.9	1.5 versicolor
## 102	6.3	2.3	4.4	1.3 versicolor
## 103	6.3	3.3	6.0	2.5 virginica
## 104	6.3	2.9	5.6	1.8 virginica
## 105	6.3	2.7	4.9	1.8 virginica
## 106	6.3	2.8	5.1	1.5 virginica
## 107	6.3	3.4	5.6	2.4 virginica
## 108	6.3	2.5	5.0	1.9 virginica
## 109	6.4	3.2	4.5	1.5 versicolor
## 110	6.4	2.9	4.3	1.3 versicolor
## 111	6.4	2.7	5.3	1.9 virginica
## 112	6.4	3.2	5.3	2.3 virginica
## 113	6.4	2.8	5.6	2.1 virginica
## 114	6.4	2.8	5.6	2.2 virginica
## 115	6.4	3.1	5.5	1.8 virginica
## 116	6.5	2.8	4.6	1.5 versicolor
## 117	6.5	3.0	5.8	2.2 virginica
## 118	6.5	3.2	5.1	2.0 virginica
## 119	6.5	3.0	5.5	1.8 virginica
## 120	6.5	3.0	5.2	2.0 virginica
## 121	6.6	2.9	4.6	1.3 versicolor
## 122	6.6	3.0	4.4	1.4 versicolor
## 123	6.7	3.1	4.4	1.4 versicolor
## 124	6.7	3.0	5.0	1.7 versicolor
## 125	6.7	3.1	4.7	1.5 versicolor
## 126	6.7	2.5	5.8	1.8 virginica
## 127	6.7	3.3	5.7	2.1 virginica
## 128	6.7	3.1	5.6	2.4 virginica
## 129	6.7	3.3	5.7	2.5 virginica
## 130	6.7	3.0	5.2	2.3 virginica
## 131	6.8	2.8	4.8	1.4 versicolor
## 132	6.8	3.0	5.5	2.1 virginica

## 133	6.8	3.2	5.9	2.3	virginica
## 134	6.9	3.1	4.9	1.5	versicolor
## 135	6.9	3.2	5.7	2.3	virginica
## 136	6.9	3.1	5.4	2.1	virginica
## 137	6.9	3.1	5.1	2.3	virginica
## 138	7.0	3.2	4.7	1.4	versicolor
## 139	7.1	3.0	5.9	2.1	virginica
## 140	7.2	3.6	6.1	2.5	virginica
## 141	7.2	3.2	6.0	1.8	virginica
## 142	7.2	3.0	5.8	1.6	virginica
## 143	7.3	2.9	6.3	1.8	virginica
## 144	7.4	2.8	6.1	1.9	virginica
## 145	7.6	3.0	6.6	2.1	virginica
## 146	7.7	3.8	6.7	2.2	virginica
## 147	7.7	2.6	6.9	2.3	virginica
## 148	7.7	2.8	6.7	2.0	virginica
## 149	7.7	3.0	6.1	2.3	virginica
## 150	7.9	3.8	6.4	2.0	virginica

El default es orden ascendente pero podemos usar descendiente,

```
arrange(iris, desc(Sepal.Length))
```

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
## 1	7.9	3.8	6.4	2.0	virginica
## 2	7.7	3.8	6.7	2.2	virginica
## 3	7.7	2.6	6.9	2.3	virginica
## 4	7.7	2.8	6.7	2.0	virginica
## 5	7.7	3.0	6.1	2.3	virginica
## 6	7.6	3.0	6.6	2.1	virginica
## 7	7.4	2.8	6.1	1.9	virginica
## 8	7.3	2.9	6.3	1.8	virginica
## 9	7.2	3.6	6.1	2.5	virginica
## 10	7.2	3.2	6.0	1.8	virginica
## 11	7.2	3.0	5.8	1.6	virginica
## 12	7.1	3.0	5.9	2.1	virginica
## 13	7.0	3.2	4.7	1.4	versicolor
## 14	6.9	3.1	4.9	1.5	versicolor
## 15	6.9	3.2	5.7	2.3	virginica
## 16	6.9	3.1	5.4	2.1	virginica
## 17	6.9	3.1	5.1	2.3	virginica
## 18	6.8	2.8	4.8	1.4	versicolor
## 19	6.8	3.0	5.5	2.1	virginica
## 20	6.8	3.2	5.9	2.3	virginica
## 21	6.7	3.1	4.4	1.4	versicolor
## 22	6.7	3.0	5.0	1.7	versicolor
## 23	6.7	3.1	4.7	1.5	versicolor
## 24	6.7	2.5	5.8	1.8	virginica
## 25	6.7	3.3	5.7	2.1	virginica
## 26	6.7	3.1	5.6	2.4	virginica
## 27	6.7	3.3	5.7	2.5	virginica
## 28	6.7	3.0	5.2	2.3	virginica
## 29	6.6	2.9	4.6	1.3	versicolor
## 30	6.6	3.0	4.4	1.4	versicolor
## 31	6.5	2.8	4.6	1.5	versicolor

## 32	6.5	3.0	5.8	2.2	virginica
## 33	6.5	3.2	5.1	2.0	virginica
## 34	6.5	3.0	5.5	1.8	virginica
## 35	6.5	3.0	5.2	2.0	virginica
## 36	6.4	3.2	4.5	1.5	versicolor
## 37	6.4	2.9	4.3	1.3	versicolor
## 38	6.4	2.7	5.3	1.9	virginica
## 39	6.4	3.2	5.3	2.3	virginica
## 40	6.4	2.8	5.6	2.1	virginica
## 41	6.4	2.8	5.6	2.2	virginica
## 42	6.4	3.1	5.5	1.8	virginica
## 43	6.3	3.3	4.7	1.6	versicolor
## 44	6.3	2.5	4.9	1.5	versicolor
## 45	6.3	2.3	4.4	1.3	versicolor
## 46	6.3	3.3	6.0	2.5	virginica
## 47	6.3	2.9	5.6	1.8	virginica
## 48	6.3	2.7	4.9	1.8	virginica
## 49	6.3	2.8	5.1	1.5	virginica
## 50	6.3	3.4	5.6	2.4	virginica
## 51	6.3	2.5	5.0	1.9	virginica
## 52	6.2	2.2	4.5	1.5	versicolor
## 53	6.2	2.9	4.3	1.3	versicolor
## 54	6.2	2.8	4.8	1.8	virginica
## 55	6.2	3.4	5.4	2.3	virginica
## 56	6.1	2.9	4.7	1.4	versicolor
## 57	6.1	2.8	4.0	1.3	versicolor
## 58	6.1	2.8	4.7	1.2	versicolor
## 59	6.1	3.0	4.6	1.4	versicolor
## 60	6.1	3.0	4.9	1.8	virginica
## 61	6.1	2.6	5.6	1.4	virginica
## 62	6.0	2.2	4.0	1.0	versicolor
## 63	6.0	2.9	4.5	1.5	versicolor
## 64	6.0	2.7	5.1	1.6	versicolor
## 65	6.0	3.4	4.5	1.6	versicolor
## 66	6.0	2.2	5.0	1.5	virginica
## 67	6.0	3.0	4.8	1.8	virginica
## 68	5.9	3.0	4.2	1.5	versicolor
## 69	5.9	3.2	4.8	1.8	versicolor
## 70	5.9	3.0	5.1	1.8	virginica
## 71	5.8	4.0	1.2	0.2	setosa
## 72	5.8	2.7	4.1	1.0	versicolor
## 73	5.8	2.7	3.9	1.2	versicolor
## 74	5.8	2.6	4.0	1.2	versicolor
## 75	5.8	2.7	5.1	1.9	virginica
## 76	5.8	2.8	5.1	2.4	virginica
## 77	5.8	2.7	5.1	1.9	virginica
## 78	5.7	4.4	1.5	0.4	setosa
## 79	5.7	3.8	1.7	0.3	setosa
## 80	5.7	2.8	4.5	1.3	versicolor
## 81	5.7	2.6	3.5	1.0	versicolor
## 82	5.7	3.0	4.2	1.2	versicolor
## 83	5.7	2.9	4.2	1.3	versicolor
## 84	5.7	2.8	4.1	1.3	versicolor
## 85	5.7	2.5	5.0	2.0	virginica

## 86	5.6	2.9	3.6	1.3 versicolor
## 87	5.6	3.0	4.5	1.5 versicolor
## 88	5.6	2.5	3.9	1.1 versicolor
## 89	5.6	3.0	4.1	1.3 versicolor
## 90	5.6	2.7	4.2	1.3 versicolor
## 91	5.6	2.8	4.9	2.0 virginica
## 92	5.5	4.2	1.4	0.2 setosa
## 93	5.5	3.5	1.3	0.2 setosa
## 94	5.5	2.3	4.0	1.3 versicolor
## 95	5.5	2.4	3.8	1.1 versicolor
## 96	5.5	2.4	3.7	1.0 versicolor
## 97	5.5	2.5	4.0	1.3 versicolor
## 98	5.5	2.6	4.4	1.2 versicolor
## 99	5.4	3.9	1.7	0.4 setosa
## 100	5.4	3.7	1.5	0.2 setosa
## 101	5.4	3.9	1.3	0.4 setosa
## 102	5.4	3.4	1.7	0.2 setosa
## 103	5.4	3.4	1.5	0.4 setosa
## 104	5.4	3.0	4.5	1.5 versicolor
## 105	5.3	3.7	1.5	0.2 setosa
## 106	5.2	3.5	1.5	0.2 setosa
## 107	5.2	3.4	1.4	0.2 setosa
## 108	5.2	4.1	1.5	0.1 setosa
## 109	5.2	2.7	3.9	1.4 versicolor
## 110	5.1	3.5	1.4	0.2 setosa
## 111	5.1	3.5	1.4	0.3 setosa
## 112	5.1	3.8	1.5	0.3 setosa
## 113	5.1	3.7	1.5	0.4 setosa
## 114	5.1	3.3	1.7	0.5 setosa
## 115	5.1	3.4	1.5	0.2 setosa
## 116	5.1	3.8	1.9	0.4 setosa
## 117	5.1	3.8	1.6	0.2 setosa
## 118	5.1	2.5	3.0	1.1 versicolor
## 119	5.0	3.6	1.4	0.2 setosa
## 120	5.0	3.4	1.5	0.2 setosa
## 121	5.0	3.0	1.6	0.2 setosa
## 122	5.0	3.4	1.6	0.4 setosa
## 123	5.0	3.2	1.2	0.2 setosa
## 124	5.0	3.5	1.3	0.3 setosa
## 125	5.0	3.5	1.6	0.6 setosa
## 126	5.0	3.3	1.4	0.2 setosa
## 127	5.0	2.0	3.5	1.0 versicolor
## 128	5.0	2.3	3.3	1.0 versicolor
## 129	4.9	3.0	1.4	0.2 setosa
## 130	4.9	3.1	1.5	0.1 setosa
## 131	4.9	3.1	1.5	0.2 setosa
## 132	4.9	3.6	1.4	0.1 setosa
## 133	4.9	2.4	3.3	1.0 versicolor
## 134	4.9	2.5	4.5	1.7 virginica
## 135	4.8	3.4	1.6	0.2 setosa
## 136	4.8	3.0	1.4	0.1 setosa
## 137	4.8	3.4	1.9	0.2 setosa
## 138	4.8	3.1	1.6	0.2 setosa
## 139	4.8	3.0	1.4	0.3 setosa

## 140	4.7	3.2	1.3	0.2	setosa
## 141	4.7	3.2	1.6	0.2	setosa
## 142	4.6	3.1	1.5	0.2	setosa
## 143	4.6	3.4	1.4	0.3	setosa
## 144	4.6	3.6	1.0	0.2	setosa
## 145	4.6	3.2	1.4	0.2	setosa
## 146	4.5	2.3	1.3	0.3	setosa
## 147	4.4	2.9	1.4	0.2	setosa
## 148	4.4	3.0	1.3	0.2	setosa
## 149	4.4	3.2	1.3	0.2	setosa
## 150	4.3	3.0	1.1	0.1	setosa

`select()`

Select es para seleccionar columnas en lugar de filas/instancias. Experimenta qué sucede cuando usas las líneas comentadas.

```
select(iris, Petal.Length, Species)
```

##	Petal.Length	Species
## 1	1.4	setosa
## 2	1.4	setosa
## 3	1.3	setosa
## 4	1.5	setosa
## 5	1.4	setosa
## 6	1.7	setosa
## 7	1.4	setosa
## 8	1.5	setosa
## 9	1.4	setosa
## 10	1.5	setosa
## 11	1.5	setosa
## 12	1.6	setosa
## 13	1.4	setosa
## 14	1.1	setosa
## 15	1.2	setosa
## 16	1.5	setosa
## 17	1.3	setosa
## 18	1.4	setosa
## 19	1.7	setosa
## 20	1.5	setosa
## 21	1.7	setosa
## 22	1.5	setosa
## 23	1.0	setosa
## 24	1.7	setosa
## 25	1.9	setosa
## 26	1.6	setosa
## 27	1.6	setosa
## 28	1.5	setosa
## 29	1.4	setosa
## 30	1.6	setosa
## 31	1.6	setosa
## 32	1.5	setosa
## 33	1.5	setosa
## 34	1.4	setosa
## 35	1.5	setosa

## 36	1.2	setosa
## 37	1.3	setosa
## 38	1.4	setosa
## 39	1.3	setosa
## 40	1.5	setosa
## 41	1.3	setosa
## 42	1.3	setosa
## 43	1.3	setosa
## 44	1.6	setosa
## 45	1.9	setosa
## 46	1.4	setosa
## 47	1.6	setosa
## 48	1.4	setosa
## 49	1.5	setosa
## 50	1.4	setosa
## 51	4.7	versicolor
## 52	4.5	versicolor
## 53	4.9	versicolor
## 54	4.0	versicolor
## 55	4.6	versicolor
## 56	4.5	versicolor
## 57	4.7	versicolor
## 58	3.3	versicolor
## 59	4.6	versicolor
## 60	3.9	versicolor
## 61	3.5	versicolor
## 62	4.2	versicolor
## 63	4.0	versicolor
## 64	4.7	versicolor
## 65	3.6	versicolor
## 66	4.4	versicolor
## 67	4.5	versicolor
## 68	4.1	versicolor
## 69	4.5	versicolor
## 70	3.9	versicolor
## 71	4.8	versicolor
## 72	4.0	versicolor
## 73	4.9	versicolor
## 74	4.7	versicolor
## 75	4.3	versicolor
## 76	4.4	versicolor
## 77	4.8	versicolor
## 78	5.0	versicolor
## 79	4.5	versicolor
## 80	3.5	versicolor
## 81	3.8	versicolor
## 82	3.7	versicolor
## 83	3.9	versicolor
## 84	5.1	versicolor
## 85	4.5	versicolor
## 86	4.5	versicolor
## 87	4.7	versicolor
## 88	4.4	versicolor
## 89	4.1	versicolor

## 90	4.0 versicolor
## 91	4.4 versicolor
## 92	4.6 versicolor
## 93	4.0 versicolor
## 94	3.3 versicolor
## 95	4.2 versicolor
## 96	4.2 versicolor
## 97	4.2 versicolor
## 98	4.3 versicolor
## 99	3.0 versicolor
## 100	4.1 versicolor
## 101	6.0 virginica
## 102	5.1 virginica
## 103	5.9 virginica
## 104	5.6 virginica
## 105	5.8 virginica
## 106	6.6 virginica
## 107	4.5 virginica
## 108	6.3 virginica
## 109	5.8 virginica
## 110	6.1 virginica
## 111	5.1 virginica
## 112	5.3 virginica
## 113	5.5 virginica
## 114	5.0 virginica
## 115	5.1 virginica
## 116	5.3 virginica
## 117	5.5 virginica
## 118	6.7 virginica
## 119	6.9 virginica
## 120	5.0 virginica
## 121	5.7 virginica
## 122	4.9 virginica
## 123	6.7 virginica
## 124	4.9 virginica
## 125	5.7 virginica
## 126	6.0 virginica
## 127	4.8 virginica
## 128	4.9 virginica
## 129	5.6 virginica
## 130	5.8 virginica
## 131	6.1 virginica
## 132	6.4 virginica
## 133	5.6 virginica
## 134	5.1 virginica
## 135	5.6 virginica
## 136	6.1 virginica
## 137	5.6 virginica
## 138	5.5 virginica
## 139	4.8 virginica
## 140	5.4 virginica
## 141	5.6 virginica
## 142	5.1 virginica
## 143	5.1 virginica

```
## 144      5.9 virginica
## 145      5.7 virginica
## 146      5.2 virginica
## 147      5.0 virginica
## 148      5.2 virginica
## 149      5.4 virginica
## 150      5.1 virginica
```

```
#select(iris, Petal.Length:Sepal.Length, Species)
#select(iris, ends_with('Length'))
#select(iris, starts_with('Petal'))
```

mutate()

Agregar nuevas columnas calculando valores desde columnas anteriores

```
mutate(iris,
  Sepal.Ratio = Sepal.Length/Sepal.Width,
  Petal.Ratio = Petal.Length/Petal.Width)
```

##	Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species	Sepal.Ratio
## 1	5.1	3.5	1.4	0.2	setosa	1.457143
## 2	4.9	3.0	1.4	0.2	setosa	1.633333
## 3	4.7	3.2	1.3	0.2	setosa	1.468750
## 4	4.6	3.1	1.5	0.2	setosa	1.483871
## 5	5.0	3.6	1.4	0.2	setosa	1.388889
## 6	5.4	3.9	1.7	0.4	setosa	1.384615
## 7	4.6	3.4	1.4	0.3	setosa	1.352941
## 8	5.0	3.4	1.5	0.2	setosa	1.470588
## 9	4.4	2.9	1.4	0.2	setosa	1.517241
## 10	4.9	3.1	1.5	0.1	setosa	1.580645
## 11	5.4	3.7	1.5	0.2	setosa	1.459459
## 12	4.8	3.4	1.6	0.2	setosa	1.411765
## 13	4.8	3.0	1.4	0.1	setosa	1.600000
## 14	4.3	3.0	1.1	0.1	setosa	1.433333
## 15	5.8	4.0	1.2	0.2	setosa	1.450000
## 16	5.7	4.4	1.5	0.4	setosa	1.295455
## 17	5.4	3.9	1.3	0.4	setosa	1.384615
## 18	5.1	3.5	1.4	0.3	setosa	1.457143
## 19	5.7	3.8	1.7	0.3	setosa	1.500000
## 20	5.1	3.8	1.5	0.3	setosa	1.342105
## 21	5.4	3.4	1.7	0.2	setosa	1.588235
## 22	5.1	3.7	1.5	0.4	setosa	1.378378
## 23	4.6	3.6	1.0	0.2	setosa	1.277778
## 24	5.1	3.3	1.7	0.5	setosa	1.545455
## 25	4.8	3.4	1.9	0.2	setosa	1.411765
## 26	5.0	3.0	1.6	0.2	setosa	1.666667
## 27	5.0	3.4	1.6	0.4	setosa	1.470588
## 28	5.2	3.5	1.5	0.2	setosa	1.485714
## 29	5.2	3.4	1.4	0.2	setosa	1.529412
## 30	4.7	3.2	1.6	0.2	setosa	1.468750
## 31	4.8	3.1	1.6	0.2	setosa	1.548387
## 32	5.4	3.4	1.5	0.4	setosa	1.588235
## 33	5.2	4.1	1.5	0.1	setosa	1.268293
## 34	5.5	4.2	1.4	0.2	setosa	1.309524

## 35	4.9	3.1	1.5	0.2	setosa	1.580645
## 36	5.0	3.2	1.2	0.2	setosa	1.562500
## 37	5.5	3.5	1.3	0.2	setosa	1.571429
## 38	4.9	3.6	1.4	0.1	setosa	1.361111
## 39	4.4	3.0	1.3	0.2	setosa	1.466667
## 40	5.1	3.4	1.5	0.2	setosa	1.500000
## 41	5.0	3.5	1.3	0.3	setosa	1.428571
## 42	4.5	2.3	1.3	0.3	setosa	1.956522
## 43	4.4	3.2	1.3	0.2	setosa	1.375000
## 44	5.0	3.5	1.6	0.6	setosa	1.428571
## 45	5.1	3.8	1.9	0.4	setosa	1.342105
## 46	4.8	3.0	1.4	0.3	setosa	1.600000
## 47	5.1	3.8	1.6	0.2	setosa	1.342105
## 48	4.6	3.2	1.4	0.2	setosa	1.437500
## 49	5.3	3.7	1.5	0.2	setosa	1.432432
## 50	5.0	3.3	1.4	0.2	setosa	1.515152
## 51	7.0	3.2	4.7	1.4	versicolor	2.187500
## 52	6.4	3.2	4.5	1.5	versicolor	2.000000
## 53	6.9	3.1	4.9	1.5	versicolor	2.225806
## 54	5.5	2.3	4.0	1.3	versicolor	2.391304
## 55	6.5	2.8	4.6	1.5	versicolor	2.321429
## 56	5.7	2.8	4.5	1.3	versicolor	2.035714
## 57	6.3	3.3	4.7	1.6	versicolor	1.909091
## 58	4.9	2.4	3.3	1.0	versicolor	2.041667
## 59	6.6	2.9	4.6	1.3	versicolor	2.275862
## 60	5.2	2.7	3.9	1.4	versicolor	1.925926
## 61	5.0	2.0	3.5	1.0	versicolor	2.500000
## 62	5.9	3.0	4.2	1.5	versicolor	1.966667
## 63	6.0	2.2	4.0	1.0	versicolor	2.727273
## 64	6.1	2.9	4.7	1.4	versicolor	2.103448
## 65	5.6	2.9	3.6	1.3	versicolor	1.931034
## 66	6.7	3.1	4.4	1.4	versicolor	2.161290
## 67	5.6	3.0	4.5	1.5	versicolor	1.866667
## 68	5.8	2.7	4.1	1.0	versicolor	2.148148
## 69	6.2	2.2	4.5	1.5	versicolor	2.818182
## 70	5.6	2.5	3.9	1.1	versicolor	2.240000
## 71	5.9	3.2	4.8	1.8	versicolor	1.843750
## 72	6.1	2.8	4.0	1.3	versicolor	2.178571
## 73	6.3	2.5	4.9	1.5	versicolor	2.520000
## 74	6.1	2.8	4.7	1.2	versicolor	2.178571
## 75	6.4	2.9	4.3	1.3	versicolor	2.206897
## 76	6.6	3.0	4.4	1.4	versicolor	2.200000
## 77	6.8	2.8	4.8	1.4	versicolor	2.428571
## 78	6.7	3.0	5.0	1.7	versicolor	2.233333
## 79	6.0	2.9	4.5	1.5	versicolor	2.068966
## 80	5.7	2.6	3.5	1.0	versicolor	2.192308
## 81	5.5	2.4	3.8	1.1	versicolor	2.291667
## 82	5.5	2.4	3.7	1.0	versicolor	2.291667
## 83	5.8	2.7	3.9	1.2	versicolor	2.148148
## 84	6.0	2.7	5.1	1.6	versicolor	2.222222
## 85	5.4	3.0	4.5	1.5	versicolor	1.800000
## 86	6.0	3.4	4.5	1.6	versicolor	1.764706
## 87	6.7	3.1	4.7	1.5	versicolor	2.161290
## 88	6.3	2.3	4.4	1.3	versicolor	2.739130

## 89	5.6	3.0	4.1	1.3 versicolor	1.866667
## 90	5.5	2.5	4.0	1.3 versicolor	2.200000
## 91	5.5	2.6	4.4	1.2 versicolor	2.115385
## 92	6.1	3.0	4.6	1.4 versicolor	2.033333
## 93	5.8	2.6	4.0	1.2 versicolor	2.230769
## 94	5.0	2.3	3.3	1.0 versicolor	2.173913
## 95	5.6	2.7	4.2	1.3 versicolor	2.074074
## 96	5.7	3.0	4.2	1.2 versicolor	1.900000
## 97	5.7	2.9	4.2	1.3 versicolor	1.965517
## 98	6.2	2.9	4.3	1.3 versicolor	2.137931
## 99	5.1	2.5	3.0	1.1 versicolor	2.040000
## 100	5.7	2.8	4.1	1.3 versicolor	2.035714
## 101	6.3	3.3	6.0	2.5 virginica	1.909091
## 102	5.8	2.7	5.1	1.9 virginica	2.148148
## 103	7.1	3.0	5.9	2.1 virginica	2.366667
## 104	6.3	2.9	5.6	1.8 virginica	2.172414
## 105	6.5	3.0	5.8	2.2 virginica	2.166667
## 106	7.6	3.0	6.6	2.1 virginica	2.533333
## 107	4.9	2.5	4.5	1.7 virginica	1.960000
## 108	7.3	2.9	6.3	1.8 virginica	2.517241
## 109	6.7	2.5	5.8	1.8 virginica	2.680000
## 110	7.2	3.6	6.1	2.5 virginica	2.000000
## 111	6.5	3.2	5.1	2.0 virginica	2.031250
## 112	6.4	2.7	5.3	1.9 virginica	2.370370
## 113	6.8	3.0	5.5	2.1 virginica	2.266667
## 114	5.7	2.5	5.0	2.0 virginica	2.280000
## 115	5.8	2.8	5.1	2.4 virginica	2.071429
## 116	6.4	3.2	5.3	2.3 virginica	2.000000
## 117	6.5	3.0	5.5	1.8 virginica	2.166667
## 118	7.7	3.8	6.7	2.2 virginica	2.026316
## 119	7.7	2.6	6.9	2.3 virginica	2.961538
## 120	6.0	2.2	5.0	1.5 virginica	2.727273
## 121	6.9	3.2	5.7	2.3 virginica	2.156250
## 122	5.6	2.8	4.9	2.0 virginica	2.000000
## 123	7.7	2.8	6.7	2.0 virginica	2.750000
## 124	6.3	2.7	4.9	1.8 virginica	2.333333
## 125	6.7	3.3	5.7	2.1 virginica	2.030303
## 126	7.2	3.2	6.0	1.8 virginica	2.250000
## 127	6.2	2.8	4.8	1.8 virginica	2.214286
## 128	6.1	3.0	4.9	1.8 virginica	2.033333
## 129	6.4	2.8	5.6	2.1 virginica	2.285714
## 130	7.2	3.0	5.8	1.6 virginica	2.400000
## 131	7.4	2.8	6.1	1.9 virginica	2.642857
## 132	7.9	3.8	6.4	2.0 virginica	2.078947
## 133	6.4	2.8	5.6	2.2 virginica	2.285714
## 134	6.3	2.8	5.1	1.5 virginica	2.250000
## 135	6.1	2.6	5.6	1.4 virginica	2.346154
## 136	7.7	3.0	6.1	2.3 virginica	2.566667
## 137	6.3	3.4	5.6	2.4 virginica	1.852941
## 138	6.4	3.1	5.5	1.8 virginica	2.064516
## 139	6.0	3.0	4.8	1.8 virginica	2.000000
## 140	6.9	3.1	5.4	2.1 virginica	2.225806
## 141	6.7	3.1	5.6	2.4 virginica	2.161290
## 142	6.9	3.1	5.1	2.3 virginica	2.225806

## 143	5.8	2.7	5.1	1.9	virginica	2.148148
## 144	6.8	3.2	5.9	2.3	virginica	2.125000
## 145	6.7	3.3	5.7	2.5	virginica	2.030303
## 146	6.7	3.0	5.2	2.3	virginica	2.233333
## 147	6.3	2.5	5.0	1.9	virginica	2.520000
## 148	6.5	3.0	5.2	2.0	virginica	2.166667
## 149	6.2	3.4	5.4	2.3	virginica	1.823529
## 150	5.9	3.0	5.1	1.8	virginica	1.966667
##	Petal.Ratio					
## 1	7.000000					
## 2	7.000000					
## 3	6.500000					
## 4	7.500000					
## 5	7.000000					
## 6	4.250000					
## 7	4.666667					
## 8	7.500000					
## 9	7.000000					
## 10	15.000000					
## 11	7.500000					
## 12	8.000000					
## 13	14.000000					
## 14	11.000000					
## 15	6.000000					
## 16	3.750000					
## 17	3.250000					
## 18	4.666667					
## 19	5.666667					
## 20	5.000000					
## 21	8.500000					
## 22	3.750000					
## 23	5.000000					
## 24	3.400000					
## 25	9.500000					
## 26	8.000000					
## 27	4.000000					
## 28	7.500000					
## 29	7.000000					
## 30	8.000000					
## 31	8.000000					
## 32	3.750000					
## 33	15.000000					
## 34	7.000000					
## 35	7.500000					
## 36	6.000000					
## 37	6.500000					
## 38	14.000000					
## 39	6.500000					
## 40	7.500000					
## 41	4.333333					
## 42	4.333333					
## 43	6.500000					
## 44	2.666667					
## 45	4.750000					

## 46	4.666667
## 47	8.000000
## 48	7.000000
## 49	7.500000
## 50	7.000000
## 51	3.357143
## 52	3.000000
## 53	3.266667
## 54	3.076923
## 55	3.066667
## 56	3.461538
## 57	2.937500
## 58	3.300000
## 59	3.538462
## 60	2.785714
## 61	3.500000
## 62	2.800000
## 63	4.000000
## 64	3.357143
## 65	2.769231
## 66	3.142857
## 67	3.000000
## 68	4.100000
## 69	3.000000
## 70	3.545455
## 71	2.666667
## 72	3.076923
## 73	3.266667
## 74	3.916667
## 75	3.307692
## 76	3.142857
## 77	3.428571
## 78	2.941176
## 79	3.000000
## 80	3.500000
## 81	3.454545
## 82	3.700000
## 83	3.250000
## 84	3.187500
## 85	3.000000
## 86	2.812500
## 87	3.133333
## 88	3.384615
## 89	3.153846
## 90	3.076923
## 91	3.666667
## 92	3.285714
## 93	3.333333
## 94	3.300000
## 95	3.230769
## 96	3.500000
## 97	3.230769
## 98	3.307692
## 99	2.727273

```

## 100    3.153846
## 101    2.400000
## 102    2.684211
## 103    2.809524
## 104    3.111111
## 105    2.636364
## 106    3.142857
## 107    2.647059
## 108    3.500000
## 109    3.222222
## 110    2.440000
## 111    2.550000
## 112    2.789474
## 113    2.619048
## 114    2.500000
## 115    2.125000
## 116    2.304348
## 117    3.055556
## 118    3.045455
## 119    3.000000
## 120    3.333333
## 121    2.478261
## 122    2.450000
## 123    3.350000
## 124    2.722222
## 125    2.714286
## 126    3.333333
## 127    2.666667
## 128    2.722222
## 129    2.666667
## 130    3.625000
## 131    3.210526
## 132    3.200000
## 133    2.545455
## 134    3.400000
## 135    4.000000
## 136    2.652174
## 137    2.333333
## 138    3.055556
## 139    2.666667
## 140    2.571429
## 141    2.333333
## 142    2.217391
## 143    2.684211
## 144    2.565217
## 145    2.280000
## 146    2.260870
## 147    2.631579
## 148    2.600000
## 149    2.347826
## 150    2.833333

```

Si solo quieres conservar las nuevas columnas puedes usar `transmute`

```
transmute(iris,
  Sepal.Ratio = Sepal.Length/Sepal.Width,
  Petal.Ratio = Petal.Length/Petal.Width)
```

##	Sepal.Ratio	Petal.Ratio
## 1	1.457143	7.000000
## 2	1.633333	7.000000
## 3	1.468750	6.500000
## 4	1.483871	7.500000
## 5	1.388889	7.000000
## 6	1.384615	4.250000
## 7	1.352941	4.666667
## 8	1.470588	7.500000
## 9	1.517241	7.000000
## 10	1.580645	15.000000
## 11	1.459459	7.500000
## 12	1.411765	8.000000
## 13	1.600000	14.000000
## 14	1.433333	11.000000
## 15	1.450000	6.000000
## 16	1.295455	3.750000
## 17	1.384615	3.250000
## 18	1.457143	4.666667
## 19	1.500000	5.666667
## 20	1.342105	5.000000
## 21	1.588235	8.500000
## 22	1.378378	3.750000
## 23	1.277778	5.000000
## 24	1.545455	3.400000
## 25	1.411765	9.500000
## 26	1.666667	8.000000
## 27	1.470588	4.000000
## 28	1.485714	7.500000
## 29	1.529412	7.000000
## 30	1.468750	8.000000
## 31	1.548387	8.000000
## 32	1.588235	3.750000
## 33	1.268293	15.000000
## 34	1.309524	7.000000
## 35	1.580645	7.500000
## 36	1.562500	6.000000
## 37	1.571429	6.500000
## 38	1.361111	14.000000
## 39	1.466667	6.500000
## 40	1.500000	7.500000
## 41	1.428571	4.333333
## 42	1.956522	4.333333
## 43	1.375000	6.500000
## 44	1.428571	2.666667
## 45	1.342105	4.750000
## 46	1.600000	4.666667
## 47	1.342105	8.000000
## 48	1.437500	7.000000
## 49	1.432432	7.500000

## 50	1.515152	7.000000
## 51	2.187500	3.357143
## 52	2.000000	3.000000
## 53	2.225806	3.266667
## 54	2.391304	3.076923
## 55	2.321429	3.066667
## 56	2.035714	3.461538
## 57	1.909091	2.937500
## 58	2.041667	3.300000
## 59	2.275862	3.538462
## 60	1.925926	2.785714
## 61	2.500000	3.500000
## 62	1.966667	2.800000
## 63	2.727273	4.000000
## 64	2.103448	3.357143
## 65	1.931034	2.769231
## 66	2.161290	3.142857
## 67	1.866667	3.000000
## 68	2.148148	4.100000
## 69	2.818182	3.000000
## 70	2.240000	3.545455
## 71	1.843750	2.666667
## 72	2.178571	3.076923
## 73	2.520000	3.266667
## 74	2.178571	3.916667
## 75	2.206897	3.307692
## 76	2.200000	3.142857
## 77	2.428571	3.428571
## 78	2.233333	2.941176
## 79	2.068966	3.000000
## 80	2.192308	3.500000
## 81	2.291667	3.454545
## 82	2.291667	3.700000
## 83	2.148148	3.250000
## 84	2.222222	3.187500
## 85	1.800000	3.000000
## 86	1.764706	2.812500
## 87	2.161290	3.133333
## 88	2.739130	3.384615
## 89	1.866667	3.153846
## 90	2.200000	3.076923
## 91	2.115385	3.666667
## 92	2.033333	3.285714
## 93	2.230769	3.333333
## 94	2.173913	3.300000
## 95	2.074074	3.230769
## 96	1.900000	3.500000
## 97	1.965517	3.230769
## 98	2.137931	3.307692
## 99	2.040000	2.727273
## 100	2.035714	3.153846
## 101	1.909091	2.400000
## 102	2.148148	2.684211
## 103	2.366667	2.809524

```
## 104      2.172414      3.111111
## 105      2.166667      2.636364
## 106      2.533333      3.142857
## 107      1.960000      2.647059
## 108      2.517241      3.500000
## 109      2.680000      3.222222
## 110      2.000000      2.440000
## 111      2.031250      2.550000
## 112      2.370370      2.789474
## 113      2.266667      2.619048
## 114      2.280000      2.500000
## 115      2.071429      2.125000
## 116      2.000000      2.304348
## 117      2.166667      3.055556
## 118      2.026316      3.045455
## 119      2.961538      3.000000
## 120      2.727273      3.333333
## 121      2.156250      2.478261
## 122      2.000000      2.450000
## 123      2.750000      3.350000
## 124      2.333333      2.722222
## 125      2.030303      2.714286
## 126      2.250000      3.333333
## 127      2.214286      2.666667
## 128      2.033333      2.722222
## 129      2.285714      2.666667
## 130      2.400000      3.625000
## 131      2.642857      3.210526
## 132      2.078947      3.200000
## 133      2.285714      2.545455
## 134      2.250000      3.400000
## 135      2.346154      4.000000
## 136      2.566667      2.652174
## 137      1.852941      2.333333
## 138      2.064516      3.055556
## 139      2.000000      2.666667
## 140      2.225806      2.571429
## 141      2.161290      2.333333
## 142      2.225806      2.217391
## 143      2.148148      2.684211
## 144      2.125000      2.565217
## 145      2.030303      2.280000
## 146      2.233333      2.260870
## 147      2.520000      2.631579
## 148      2.166667      2.600000
## 149      1.823529      2.347826
## 150      1.966667      2.833333
```

summarise()

Summarise es de las funciones más poderosas, en general la combinamos con un groupby para generar cosas más interesantes.

```
by_species <- group_by(iris, Species)
```

```
summarise(by_species,
  count = n(),
  Petal.Length.Mean = mean(Petal.Length),
  Petal.Length.Standard_Deviation = sd(Petal.Length))
```

```
## # A tibble: 3 x 4
##   Species    count Petal.Length.Mean Petal.Length.Standard_Deviation
##   <fct>      <int>          <dbl>                      <dbl>
## 1 setosa        50            1.46                      0.174
## 2 versicolor   50            4.26                      0.470
## 3 virginica    50            5.55                      0.552
```

Ejercicios

Has lo que se pide para el dataset flights.

1. Filtra los vuelos que ocurrieron en Halloween, y crea un gráfico de barras que muestre cuántos vuelos salieron por hora.
2. Filtra todos los vuelos que salieron el día primero de Julio y Agosto y has un scatter plot con el tiempo de salida (dep_time) en el eje x y el retraso de llegada (arr_delay) en el eje y. Haz que cambie el color de los puntos en base al tiempo que el avión estuvo en el aire (air_time).

En este dataset, cada que un vuelo no salió se usó el valor NA. Aunque ggplot se deshace solo de los datos, vale la pena agregar `!is.na(<variable>)` a la lista de valores en filter para asegurarte que no regrese valores de NA.

3. Filtra los vuelos que llegaron más de dos horas tarde (arr_delay) pero no salieron más de dos horas tarde (dep_delay) y luego enseña la tabla organizada por el retraso de llegada donde el primer valor es el que tiene mayor retraso. Que tu tabla solo contenga las variables de flight, carrier, arr_delay y dep_delay en ese orden.
4. Filtra los vuelos que salieron al menos una hora tarde pero recuperaron al menos 30 minutos durante el vuelo y crea una nueva columna con el tiempo que recuperó el vuelo. Ahora preséntalo en una tabla donde por cada aerolínea diga en promedio cuánto recuperaron por vuelo, la desviación estándar.

Si no quitas los NA entonces no podrás encontrar los valores que buscas