

Ejercicio dplyr

Con el dataset diamonds:

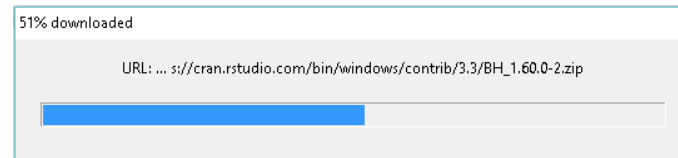
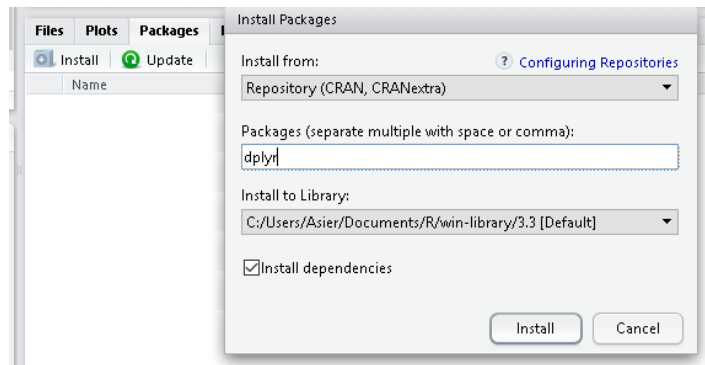
I Filtrar los diamantes con corte \Ideal".

En el R básico seria:

```
diamonds2 <- diamonds  copia de seguridad de diamonds
```

```
corte <- data.frame ( diamonds2 [diamonds2$cut == "Ideal", ]) poner en corte un dataframe nuevo con datos de diamonds 2 , pero donde en la columna cut de diamonds sea igual a Ideal
```

Pero como esto va de dplyr pues primero cargamos libreria



```
library(dplyr)
```

```
corte2 <- filter(diamonds2, cut == "Ideal")
```

	carat	cut	color	clarity	depth	table	price	x	y	z
1	0.23	Ideal	E	SI2	61.5	55.0	326	3.95	3.98	2.43
2	0.23	Ideal	J	VS1	62.8	56.0	340	3.93	3.90	2.46
3	0.31	Ideal	J	SI2	62.2	54.0	344	4.35	4.37	2.71
4	0.30	Ideal	I	SI2	62.0	54.0	348	4.31	4.34	2.68
5	0.33	Ideal	I	SI2	61.8	55.0	403	4.49	4.51	2.78
6	0.33	Ideal	I	SI2	61.2	56.0	403	4.49	4.50	2.75
7	0.33	Ideal	J	SI1	61.1	56.0	403	4.49	4.55	2.76

Seleccíon las columnas carat, cut, color, price y clarity.

```
seleccion <- select(diamonds2,carat,cut,color, price,clarity)
```

seleccion Nombre del data frame

select() que haces, seleccionar

diamonds2 de que data frame

carat,cut,color, price,clarity nombre columnas

	carat	cut	color	price	clarity
1	0.23	Ideal	E	326	SI2
2	0.21	Premium	E	326	SI1
3	0.23	Good	E	327	VS1
4	0.29	Premium	I	334	VS2

1 Crear una nueva columna precio/quilate.

```
seleccion <- mutate(seleccion,precio_quilate=price/carat)
```

`seleccion <-` Modificamos el data frame seleccion

`mutate()` Mutate para añadir columna

`seleccion` Donde añadimos columna

`,precio_quilate` como se llama la columna

`=price/carat`) contenido columna que es el resultado de dividir la columna price con la de carat

	carat	cut	color	price	clarity	precio_quilate
1	0.23	Ideal	E	326	SI2	1417.391
2	0.21	Premium	E	326	SI1	1552.381

1 Agrupar los diamantes por color.

```
selecccion <- group_by(seleccion,color)
```

`selecccion <-` Modificamos el data frame seleccion

`group_by()` agrupar por

`seleccion`, Donde agrupamos

`color`) en base a que

1 Calcular la media del precio/quilate para cada uno de los grupos anteriores.

```
media <- summarise(seleccion2, mean(precio_quilate))
```

`media <-` creamos el data frame media

`summarise()` utilizamos la función summarise para hacer una operación aritmetica

`seleccion2`, donde utilizamos la función

`mean(precio_quilate)` donY aque ddecimos que la operación aritmetica es la media en la columna precio_quilate (ya la teniamos agrupada y declarada que es en seleccion 2)

```

> media
# A tibble: 7 x 2
  color mean(precio_quilate)
  <ord>          <dbl>
1     D          3952.564
2     E          3804.611
3     F          4134.731
4     G          4163.412
5     H          4008.027
6     I          3996.402
7     J          3825.649
> |

```

1 Ordenar por precio/quilate de forma descendente.

```
ordenado <- arrange(seleccion2, desc(precio_quilate))
```

ordenado <- creamos el data frame ordenado
 arrange(funcion para ordenar
 seleccion2, en que data frame la aplicamos
 desc(en que orden se ordena
 precio_quilate)) cual es la columna

	carat	cut	color	price	clarity	precio_quilate
1	1.04	Very Good	D	18542	IF	17828.85
2	1.07	Premium	D	18279	IF	17083.18
3	1.03	Ideal	D	17590	IF	17077.67
4	1.07	Very Good	D	18114	IF	16928.97
5	1.02	Very Good	D	17100	IF	16764.71