

desafio6

```
library(RSQLite)
```

Warning: pacote 'RSQLite' foi compilado no R versão 4.4.3

```
path <- '../ME315'  
fname <- file.path(path, 'disco.db')  
  
conn <- dbConnect(SQLite(), fname)
```

```
# Exercicios do slide
```

```
# Contagem do total de clientes cadastrados  
sql <- 'SELECT COUNT(customerid) FROM customers'  
head(dbGetQuery(conn, sql))
```

```
      COUNT(customerid)  
1                59
```

```
# Contagem do total de paises distintos  
sql2 <- 'SELECT COUNT(DISTINCT country) FROM customers'  
head(dbGetQuery(conn, sql2))
```

```
      COUNT(DISTINCT country)  
1                24
```

```
# Contagem de clientes por país  
sql3 <- 'SELECT country, COUNT(customerid) FROM customers GROUP BY country'  
head(dbGetQuery(conn, sql3))
```

	Country	COUNT(customerid)
1	Argentina	1
2	Australia	1
3	Austria	1
4	Belgium	1
5	Brazil	5
6	Canada	8

```
# 5 paises com mais clientes
sql4 <- paste('SELECT country, COUNT(customerid) FROM customers GROUP BY country',
              'ORDER BY COUNT(customerid) DESC LIMIT 5')
head(dbGetQuery(conn, sql4))
```

	Country	COUNT(customerid)
1	USA	13
2	Canada	8
3	France	5
4	Brazil	5
5	Germany	4

```
# Paises com 6 letras
sql5 <- 'SELECT DISTINCT country FROM customers WHERE LENGTH(country) = 6'
head(dbGetQuery(conn, sql5))
```

	Country
1	Brazil
2	Canada
3	Norway
4	France
5	Poland
6	Sweden

```
# Musicas compradas por brasileiros
sql6 <- paste('SELECT tracks.name FROM customers',
              'JOIN invoices ON customers.customerid = invoices.customerid',
              'JOIN invoice_items ON invoices.invoiceid = invoice_items.invoiceid',
              'JOIN tracks ON invoice_items.trackid = tracks.trackid',
              'WHERE customers.country = 'Brazil' GROUP BY tracks.name',
              'ORDER BY tracks.name')
head(dbGetQuery(conn, sql6))
```

	Name
1	1/2 Full
2	2 Minutes To Midnight
3	A Cor Do Sol
4	A Cura
5	A Menina Dança
6	Abraham, Martin And John

```
# Desafios do classroom
```

```
# Qual o álbum mais tocado por país?
```

```
# Tabela com a contagem de compras de cada album por cada país
```

```
albums_count_query <- paste('SELECT albums.title, country, COUNT(*) as count FROM customers'
                             # titulo do album, país dos usuarios, contagem futura
                             'JOIN invoices ON customers.customerid = invoices.customerid',
                             # adicionando a tabela de compras pelo id dos usuarios que realizou a compra
                             'JOIN invoice_items ON invoices.invoiceid = invoice_items.invoiceid',
                             # adicionando a tabela das outras informacoes das compras através do id da compra
                             'JOIN tracks ON invoice_items.trackid = tracks.trackid',
                             # adicionando a tabela de musicas através do id da musica contida no album
                             'JOIN albums ON tracks.albumid = albums.albumid',
                             # adicionando a tabela de albums através do id do album contido no track
                             'GROUP BY albums.title, customers.country',
                             # agrupando por titulo do album e país
                             'ORDER BY country, count DESC')
                             # ordenando por país e a contagem de albums
```

```
# selecionando apenas a maior contagem em cada país
```

```
album_count_bycountry_query <- paste('SELECT title as Album, country as Pais,
                                     count as Vendas FROM (', albums_count_query, ')',
                                     # mantendo as colunas necessarias
                                     'GROUP BY country',
                                     # Agrupando por país
                                     'HAVING count = MAX(count)')
                                     # em cada grupo (país) seleciona a linha de maior valor
```

```
most_stremead_albums <- dbGetQuery(conn, album_count_bycountry_query)
head(most_stremead_albums)
```

	Album	Pais	Vendas
1	Acústico	Argentina	4

2	The X Factor	Australia	4
3	My Generation - The Very Best Of The Who	Austria	4
4	Big Ones	Belgium	4
5	Use Your Illusion I	Brazil	7
6	Arquivo II	Canada	9

```
# Qual o artista mais tocado por pais?
```

```
total_artists_count <- paste('SELECT artists.name, country, COUNT(*) as count FROM customers
# nome do artista, país dos usuarios, contagem futura
'JOIN invoices ON customers.customerid = invoices.customerid',
# adicionando a tabela de compras pelo id dos usuarios que real
'JOIN invoice_items ON invoices.invoiceid = invoice_items.invoic
# adicionando a tabela das outras informacoes das compras atrav
'JOIN tracks ON invoice_items.trackid = tracks.trackid',
# adicionando a tabela de musicas atraves do id da musica contido
'JOIN albums ON tracks.albumid = albums.albumid',
# adicionando a tabela de albums atraves do id do album contido
'JOIN artists ON albums.artistid = artists.artistid',
# adicionando a tabela de artistas pelo id do artista contido j
'GROUP BY artists.name, customers.country',
# agrupando por artista e pais dos usuarios que o ouvem
'ORDER BY country, count DESC')
# ordenando por país e a contagem de usuarios que ouvem cada ar

# Selecionando o artista com mais contagem de compras por pais
most_stremead_bycountry <- paste('SELECT name as Artista, country as Pais,
count as Vendas FROM (', total_artists_count, ')',
# Varivaeis de interesse
'GROUP BY country',
# AGrupando por pais
'HAVING count = MAX(count)')
# Selecionando o artista de maior numero de vendas em cada g

most_stremead_artists <- dbGetQuery(conn, most_stremead_bycountry)
head(most_stremead_artists)
```

	Artista	Pais	Vendas
1	Metallica	Argentina	6
2	Iron Maiden	Australia	18
3	U2	Austria	9
4	Faith No More	Belgium	9

5 Os Paralamas Do Sucesso	Brazil	11
6 Os Paralamas Do Sucesso	Canada	16