

Desafio 06

ME315-2S2025

brotto

2025-09-24

Funções

clearst

```
clearst = function(texto) {  
  texto = gsub("\n", " ", texto)  
  texto = gsub("\\s+", " ", texto)  
  texto = trimws(texto)  
  return(texto)  
}
```

Atividade

Instruções

```
file_path = "./disco.db"
```

```
conn = dbConnect(SQLite(), file_path)  
tabelas = dbListTables(conn)
```

```
kable(data.frame(tabelas = tabelas))
```

tabelas
albums
artists
customers
employees
genres
invoice_items
invoices
media_types
playlist_track
playlists
sqlite_sequence
sqlite_stat1
tracks

```
customers <- dbGetQuery(conn, paste0("PRAGMA table_info(", "customers", ")"))$name  
albums <- dbGetQuery(conn, paste0("PRAGMA table_info(", "albums", ")"))$name  
invoices <- dbGetQuery(conn, paste0("PRAGMA table_info(", "invoices", ")"))$name  
invoice_items <- dbGetQuery(conn, paste0("PRAGMA table_info(", "invoice_items", ")"))$name
```

```
kable(data.frame(customers = customers))
```

customers
CustomerId
FirstName
LastName
Company
Address
City
State
Country
PostalCode
Phone
Fax
Email
SupportRepId

```
kable(data.frame(albums = albums))
```

albums
AlbumId
Title
ArtistId

```
kable(data.frame(invoices = invoices))
```

invoices
InvoiceId
CustomerId
InvoiceDate
BillingAddress
BillingCity
BillingState
BillingCountry
BillingPostalCode
Total

```
kable(data.frame(invoice_items = invoice_items))
```

invoice_items
InvoiceLineId
InvoiceId
TrackId
UnitPrice
Quantity

```
s1 = "SELECT COUNT(*) AS n FROM customers"
kable(dbGetQuery(conn, s1))
```

n
59

```
s2 = "SELECT COUNT(DISTINCT Country) AS n_country FROM customers"
kable(dbGetQuery(conn, s2))
```

n_country
24

```
s3 = "SELECT Country, COUNT(*) AS n_clientes FROM customers GROUP BY Country ORDER BY n_clientes DESC"
kable(dbGetQuery(conn, s3))
```

Country	n_clientes
USA	13
Canada	8
France	5
Brazil	5
Germany	4
United Kingdom	3
Portugal	2
India	2
Czech Republic	2
Sweden	1
Spain	1
Poland	1
Norway	1
Netherlands	1
Italy	1
Ireland	1
Hungary	1
Finland	1
Denmark	1
Chile	1
Belgium	1
Austria	1
Australia	1
Argentina	1

```
s4 = paste(s3, "LIMIT 5")
kable(dbGetQuery(conn, s4))
```

Country	n_clientes
USA	13
Canada	8
France	5

Country	n_clientes
Brazil	5
Germany	4

```
s5 = "SELECT DISTINCT Country FROM customers WHERE Country GLOB '?????'"
kable(dbGetQuery(conn, s5))
```

Country
Brazil
Canada
Norway
France
Poland
Sweden

```
s6 = paste0("
SELECT
  c.CustomerId as id_cliente,
  c.Country as pais_cliente,
  t.Name as musica,
  ii.Quantity as quantidade
FROM customers c
JOIN invoices i ON c.CustomerId = i.CustomerId
JOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId
JOIN tracks t ON ii.TrackId = t.TrackId
WHERE c.Country = 'Brazil'")

kable(head(dbGetQuery(conn, s6)))
```

id_cliente	pais_cliente	musica	quantidade
1	Brazil	Experiment In Terra	1
1	Brazil	Take the Celestra	1
1	Brazil	Shout It Out Loud	1
1	Brazil	Calling Dr. Love	1
1	Brazil	Strutter	1
1	Brazil	Cold Gin	1

Desafio

Querys extras

Qual o álbum mais tocado por país?

```
s7 = paste0("
SELECT
  pais,
  album,
  artista,
  MAX(total_compras) as compras
FROM (
```

```

SELECT
  c.Country as pais,
  alb.Title as album,
  ar.Name as artista,
  COUNT(ii.InvoiceLineId) as total_compras
FROM customers c
JOIN invoices i ON c.CustomerId = i.CustomerId
JOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId
JOIN tracks t ON ii.TrackId = t.TrackId
JOIN albums alb ON t.AlbumId = alb.AlbumId
JOIN artists ar ON alb.ArtistId = ar.ArtistId
GROUP BY c.Country, alb.AlbumId
)
GROUP BY pais
ORDER BY compras DESC")

```

```
kable(dbGetQuery(conn, clearst(s7)))
```

pais	album	artista	com- pras
USA	The Office, Season 3	The Office	14
Canada	Arquivo II	Os Paralamas Do Sucesso	9
Ireland	Lost, Season 2	Lost	7
France	Minha Historia	Chico Buarque	7
Brazil	Use Your Illusion I	Guns N' Roses	7
India	Up An' Atom	Gene Krupa	6
United Kingdom	No Prayer For The Dying	Iron Maiden	5
Spain	My Generation - The Very Best Of The Who	The Who	5
Germany	Djavan Ao Vivo - Vol. 1	Djavan	5
Finland	Greatest Kiss	Kiss	5
Czech Republic	Prenda Minha	Caetano Veloso	5
Sweden	Volume Dois	Titãs	4
Portugal	Raul Seixas	Raul Seixas	4
Poland	Rattle And Hum	U2	4
Norway	For Those About To Rock We Salute You	AC/DC	4
Netherlands	In Step	Stevie Ray Vaughan & Double Trouble	4
Italy	Acústico	Titãs	4
Hungary	Pop	U2	4
Denmark	Chronicle, Vol. 1	Creedence Clearwater Revival	4
Belgium	Big Ones	Aerosmith	4
Austria	My Generation - The Very Best Of The Who	The Who	4
Australia	The X Factor	Iron Maiden	4
Argentina	Acústico	Titãs	4
Chile	My Way: The Best Of Frank Sinatra [Disc 1]	Frank Sinatra	3

Qual o artista mais tocado por pais?

```
s8 = paste0("
SELECT
  pais,
  artista,
  MAX(total_compras) as compras
FROM (
  SELECT
    c.Country as pais,
    ar.Name as artista,
    COUNT(ii.InvoiceLineId) as total_compras
  FROM customers c
  JOIN invoices i ON c.CustomerId = i.CustomerId
  JOIN invoice_items ii ON i.InvoiceId = ii.InvoiceId
  JOIN tracks t ON ii.TrackId = t.TrackId
  JOIN albums alb ON t.AlbumId = alb.AlbumId
  JOIN artists ar ON alb.ArtistId = ar.ArtistId
  GROUP BY c.Country, ar.ArtistId
)
GROUP BY pais
ORDER BY compras DESC, pais")

kable(dbGetQuery(conn, s8))
```

pais	artista	compras
USA	Iron Maiden	34
Australia	Iron Maiden	18
Canada	Os Paralamas Do Sucesso	16
Portugal	Iron Maiden	16
Germany	Iron Maiden	14
Brazil	Os Paralamas Do Sucesso	11
France	Creedence Clearwater Revival	11
Austria	U2	9
Belgium	Faith No More	9
Czech Republic	Iron Maiden	9
Norway	Led Zeppelin	9
United Kingdom	Metallica	9
India	Iron Maiden	8
Ireland	Lost	8
Denmark	Creedence Clearwater Revival	7
Argentina	Metallica	6
Chile	Led Zeppelin	6
Finland	Van Halen	6
Netherlands	Red Hot Chili Peppers	6
Poland	U2	6
Spain	Pearl Jam	6
Hungary	Jamiroquai	5
Italy	Faith No More	5
Sweden	Titãs	5