Class Activity 7

Your name here

April 08 2024

Problem 1: Boolean Operators

Use Boolean operators to alter the code below to return only the rows that contain:

a. Girls named Rhea

b. Names that were used by exactly 5 or 6 children in 1990

c. Names that are one of Apple, Yoroi, Ada

d. Store the data tibble in part c into a new tibble and change all the character columns to upper case. Also, rename the `n` variable to `count`.

e. Change all the column names to upper case in the previous problem.

f. What do these commands do?

```
polluted_cities %>% select_if(is.numeric) #1
polluted_cities %>% rename_all(toupper) #2
polluted_cities %>% rename_if(is.character, toupper) #3
polluted_cities %>% rename_at(vars(contains("it")), toupper) #4
```

answer:

Let's look at an interesting example on how to join related information on various artists, bands, songs, and their labels.

```
artists <- tibble(first = c("Jimmy", "George", "Mick", "Tom", "Davy", "John",
                            "Paul", "Jimmy", "Joe", "Elvis", "Keith", "Paul",
                            "Ringo", "Joe", "Brian", "Nancy"),
                  last = c("Buffett", "Harrison", "Jagger", "Jones", "Jones",
                           "Lennon", "McCartney", "Page", "Perry", "Presley",
                           "Richards", "Simon", "Starr", "Walsh", "Wilson", "Wilson"),
                  instrument = c("Guitar", "Guitar", "Vocals", "Vocals", "Vocals",
                                 "Guitar", "Bass", "Guitar", "Guitar", "Vocals", "Guitar",
                                 "Guitar", "Drums", "Guitar", "Vocals", "Vocals"))


bands <- tibble(first = c("John", "John Paul", "Jimmy", "Robert", "George", "John",
                          "Paul", "Ringo", "Jimmy", "Mick", "Keith", "Charlie", "Ronnie"),
                last = c("Bonham", "Jones", "Page", "Plant", "Harrison", "Lennon",
                         "McCartney", "Starr", "Buffett", "Jagger", "Richards", "Watts", "Wood"),
                band = c("Led Zeppelin", "Led Zeppelin", "Led Zeppelin", "Led Zeppelin",
                         "The Beatles", "The Beatles", "The Beatles", "The Beatles",
```

```
                          "The Coral Reefers", "The Rolling Stones", "The Rolling Stones",
                          "The Rolling Stones", "The Rolling Stones"))

albums <- tibble(album = c("A Hard Day's Night", "Magical Mystery Tour", "Beggar's Banquet",
                           "Abbey Road", "Led Zeppelin IV", "The Dark Side of the Moon", "Aerosmith",
                           "Rumours", "Hotel California"),
                 band = c("The Beatles", "The Beatles", "The Rolling Stones", "The Beatles",
                          "Led Zeppelin", "Pink Floyd", "Aerosmith", "Fleetwood Mac", "Eagles"),
                 year = c(1964,1967,1968,1969,1971,1973,1973,1977,1982))


songs <- tibble(song = c("Come Together", "Dream On", "Hello, Goodbye", "It's Not Unusual"),
                album  = c("Abbey Road", "Aerosmith", "Magical Mystery Tour", "Along Came Jones"),
                first = c("John", "Steven", "Paul", "Tom"),
                last = c("Lennon", "Tyler", "McCartney", "Jones"))


labels <- tibble(album = c("Abbey Road", "A Hard Days Night", "Magical Mystery Tour",
                           "Led Zeppelin IV", "The Dark Side of the Moon", "Hotel California",
                           "Rumours", "Aerosmith", "Beggar's Banquet"),
                 label = c("Apple", "Parlophone", "Parlophone", "Atlantic", "Harvest",
                           "Asylum", "Warner Brothers", "Columbia", "Decca"))
```

Let's take a glimpse of the tibbles `artists` and `bands`. Notice that there are different number of rows in the dataset.

```
glimpse(artists)
Rows: 16
Columns: 3
$ first      <chr> "Jimmy", "George", "Mick", "Tom", "Davy", "John", "Paul", "~
$ last       <chr> "Buffett", "Harrison", "Jagger", "Jones", "Jones", "Lennon"~
$ instrument <chr> "Guitar", "Guitar", "Vocals", "Vocals", "Vocals", "Guitar",~
glimpse(bands)
Rows: 13
Columns: 3
$ first <chr> "John", "John Paul", "Jimmy", "Robert", "George", "John", "Paul"~
$ last  <chr> "Bonham", "Jones", "Page", "Plant", "Harrison", "Lennon", "McCar~
$ band  <chr> "Led Zeppelin", "Led Zeppelin", "Led Zeppelin", "Led Zeppelin", ~
glimpse(albums)
Rows: 9
Columns: 3
$ album <chr> "A Hard Day's Night", "Magical Mystery Tour", "Beggar's Banquet"~
$ band  <chr> "The Beatles", "The Beatles", "The Rolling Stones", "The Beatles~
$ year  <dbl> 1964, 1967, 1968, 1969, 1971, 1973, 1973, 1977, 1982
glimpse(songs)
Rows: 4
Columns: 4
$ song  <chr> "Come Together", "Dream On", "Hello, Goodbye", "It's Not Unusual"
$ album <chr> "Abbey Road", "Aerosmith", "Magical Mystery Tour", "Along Came J~
$ first <chr> "John", "Steven", "Paul", "Tom"
$ last  <chr> "Lennon", "Tyler", "McCartney", "Jones"
glimpse(labels)
Rows: 9
Columns: 2
```

```
$ album <chr> "Abbey Road", "A Hard Days Night", "Magical Mystery Tour", "Led ~
$ label <chr> "Apple", "Parlophone", "Parlophone", "Atlantic", "Harvest", "Asy~
```

Problem 2: Joining artists and bands data

a. Join the artists and bands tibbles using `left_join()`, `right_join()`, and `full_join()`. Verify that the datasets obtained from `left_join()` and `right_join()` are the same using `setequal()`.

b. Use the pipe operator, `%>%`, to create one table that combines all information from `artists`, `bands`, `albums`, `songs`, and `labels`.

Problem 3: Filtering and counting rows in the data

a. Collect artists that have songs provided, and return rows of artists that don't have bands info.

b. Collect the albums made by a band, count the number of rows, find the rows of songs that match a row in labels, and count the number of rows.