Git & SQL Practical Assignment

Student Name

Assignment Overview

This assignment guides you through using Git and SQL in RStudio. You will clone a repository, work with a SQLite database, and use RStudio's Git interface to manage your work.

Step 1: Clone the Repository Using RStudio

1. Open RStudio.
2. Go to **File** → **New Project** → **Version Control** → **Git**.
3. Paste the repository URL provided by your instructor.
4. Choose a local folder and click **Create Project**.

Step 2: Copy the R Markdown Template

1. In the **Files** pane, locate `assignment_template.Rmd`.
2. Right-click and select **Copy**.
3. Rename the copy to `assignment_<your_student_name>.Rmd` (replace `<your_student_name>` with your actual name).

Step 3: Open Your Rmd File

Double-click your new `.Rmd` file to open it in the editor.

Step 4: Connect to the SQLite Database

```
library(DBI)
library(RSQLite)
```

```
## Warning: package 'RSQLite' was built under R version 4.4.2
```

```r
con <- dbConnect(RSQLite::SQLite(), "F:\\dsa_practical_db.sqlite")
```

Step 5: Perform SQL Queries

a. List all tables in the database

```r
dbListTables(con)
```

```
##  [1] "chemicals_clean"          "comments_clean"
##  [3] "demographics_clean"       "df_dictionary"
##  [5] "df_medications_drug_info" "dietary_clean"
##  [7] "medications_clean"        "mortality_clean"
##  [9] "occupation_clean"         "questionnaire_clean"
## [11] "response_clean"           "weights_clean"
```

b. Show the first 5 rows o a table

```r
dbGetQuery(con, "SELECT * FROM mortality_clean LIMIT 5;")
```

```
##   SEQN SEQN_new ELIGSTAT MORTSTAT UCOD_LEADING DIABETES HYPERTEN PERMTH_INT
## 1    1      C-1        2       NA           NA       NA       NA         NA
## 2    2      C-2        1        1            6        0        0        177
## 3    3      C-3        2       NA           NA       NA       NA         NA
## 4    4      C-4        2       NA           NA       NA       NA         NA
## 5    5      C-5        1        0           NA       NA       NA        244
##   PERMTH_EXM SDDSRVYR   VNELIGSTAT       VNMORTSTAT            VNUCOD_LEADING
## 1         NA        1 Under age 18             <NA>                      <NA>
## 2        177        1     Eligible Assumed deceased Alzheimer's disease (052)
## 3         NA        1 Under age 18             <NA>                      <NA>
## 4         NA        1 Under age 18             <NA>                      <NA>
## 5        244        1     Eligible    Assumed alive                      <NA>
##   VNDIABETES VNHYPERTEN
## 1       <NA>       <NA>
## 2         No         No
## 3       <NA>       <NA>
## 4       <NA>       <NA>
## 5       <NA>       <NA>
```

c. Count the number of records in another table

```r
a <- dbGetQuery(con, "SELECT COUNT(*) FROM demographics_clean;")
```

d. Join two tables and display the first 5 results

```r
b <- dbGetQuery(con, "SELECT * FROM mortality_clean JOIN demographics_clean ON mortality_clean.SEQN = d
```

Step 6: Disconnect from the Database

```
dbDisconnect(con)
```

Step 7: Generate a Personal Access Token(PAT)

First, you need to create a PAT on the GitHub website: - Log in to your GitHub account. If you don't have one create one. - Go to Settings by clicking your profile picture in the top-right corner. - In the left sidebar, scroll down and click Developer settings. - Click Personal access tokens, then select Tokens (classic). - Click Generate new token. - Give the token a descriptive Note . - Set an Expiration date. 30 days is a good starting point. - Under Select scopes, check the box next to repo. This grants permissions for full control of private repositories. - Scroll to the bottom and click Generate token.

Step 8: Commit and Push Your work Using RStudio

1. In the Git pane, check the box next to your .Rmd file.
2. Click Commit.
3. Enter a commit message (e.g., "Completed assignment by ").
4. Click Commit.
5. Click Push to upload your changes to the remote repository.