

# Analysis of Teachers' Recruitment in Saudi Arabia Between 1437 and 1440

Data from Saudi Open Data Office

Musab Isah

6 March 2021

Group Number:

Group Members' Names (IDs):

1. Abc Xyz (00000000) – Team Leader
2.                      -- Member 1
3.                      -- Member 2
4.                      -- Member 3

The purpose of this project explore the data on teachers recruitment at different levels of schooling (with the exception of higher education, such as universities) between 1437 and 1440 Hijri in Saudi Arabia. The data is sourced from Saudi Open Data Office and you are expected to use the provided data to answer the questions below.

Remember that this is a collaborative work and each member must participate in completing the project.

**Note: We have been using RStudio Cloud (RStudio on web browser) for this course but for the purpose of the project, we will be using RStudio Desktop. Because we are using a dataset that is large in size, we will experience errors due to memory limitation if we are to continue using RStudio Cloud. For that we need to install Git program as well as RStudio Desktop.**

## Installing Git

For Windows users:

1. Download Git exe file from this link <https://github.com/git-for-windows/git/releases/download/v2.31.0.windows.1/Git-2.31.0-64-bit.exe>.
2. Install the downloaded file to you computer. Do not change any of the settings during the installation process.

For Macbook users, you need to use the terminal:

1. Open the Spotlight search window on your Mac. You could do this by clicking the search icon on the top-right corner of your Mac or by pressing the “Command” button and the spacebar on your keyboard.

2. Type the word “Terminal” (without the quote) and open the Terminal program that shows after the search (a blank screen with a few text).

3. Copy and paste the following command `git --version`. If git is not already installed, it will prompt you to install it. Follow the prompt and install the program.

4. Enter your Mac password if asked. If Git is already installed, it will show you the version of the installation.

For Linux Users, also through the terminal:

Open the Terminal and issue the command `sudo apt install git-all`. Enter your root password when asked.

## Installing RStudio Desktop

For Windows users:

1. Download the application by clicking this link <https://download1.rstudio.org/desktop/windows/RStudio-1.4.1106.exe>
2. Install the downloaded exe file on your system.

For Mac users:

1. Download the application by clicking this link <https://download1.rstudio.org/desktop/macos/RStudio-1.4.1106.dmg>
2. Install the downloaded dmg file on your system.

For Linux users:

1. Download the appropriate installation file for your distribution from here <https://rstudio.com/products/rstudio/download/#download>

## Linking Github to RStudio Desktop

1. Open your installed RStudio Desktop and go to the File menu. Select ‘New Project’.
2. A New Project Wizard will appear as shown in the figure. Select ‘Version Control’.

## New Project Wizard

### Create Project



#### New Directory

Start a project in a brand new working directory



#### Existing Directory

Associate a project with an existing working directory



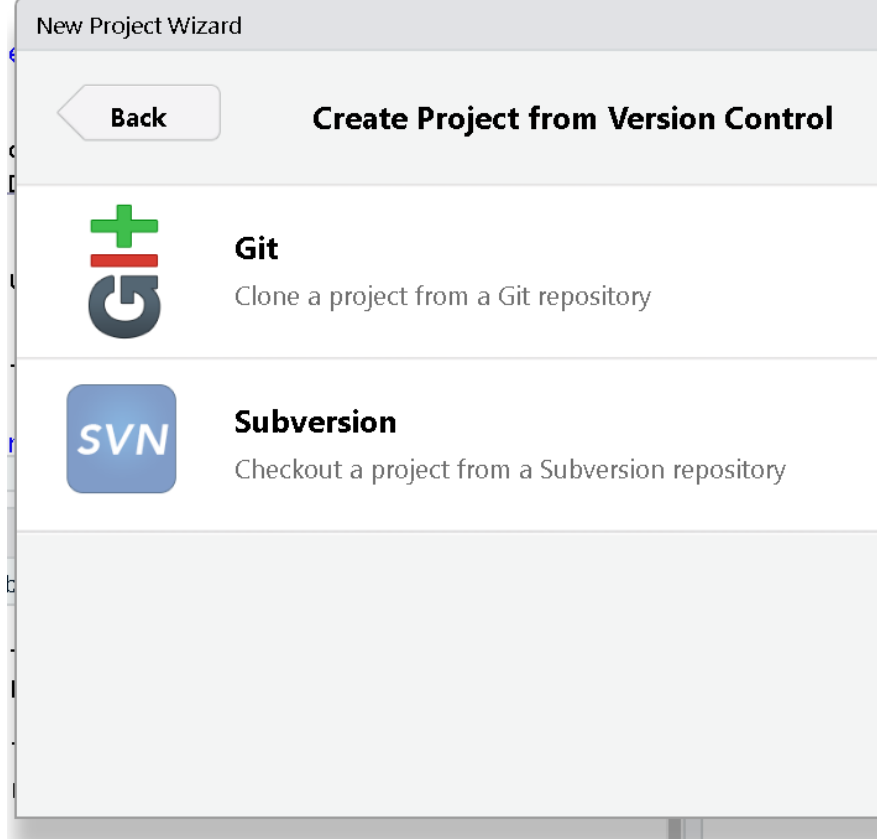
#### Version Control

Checkout a project from a version control repository



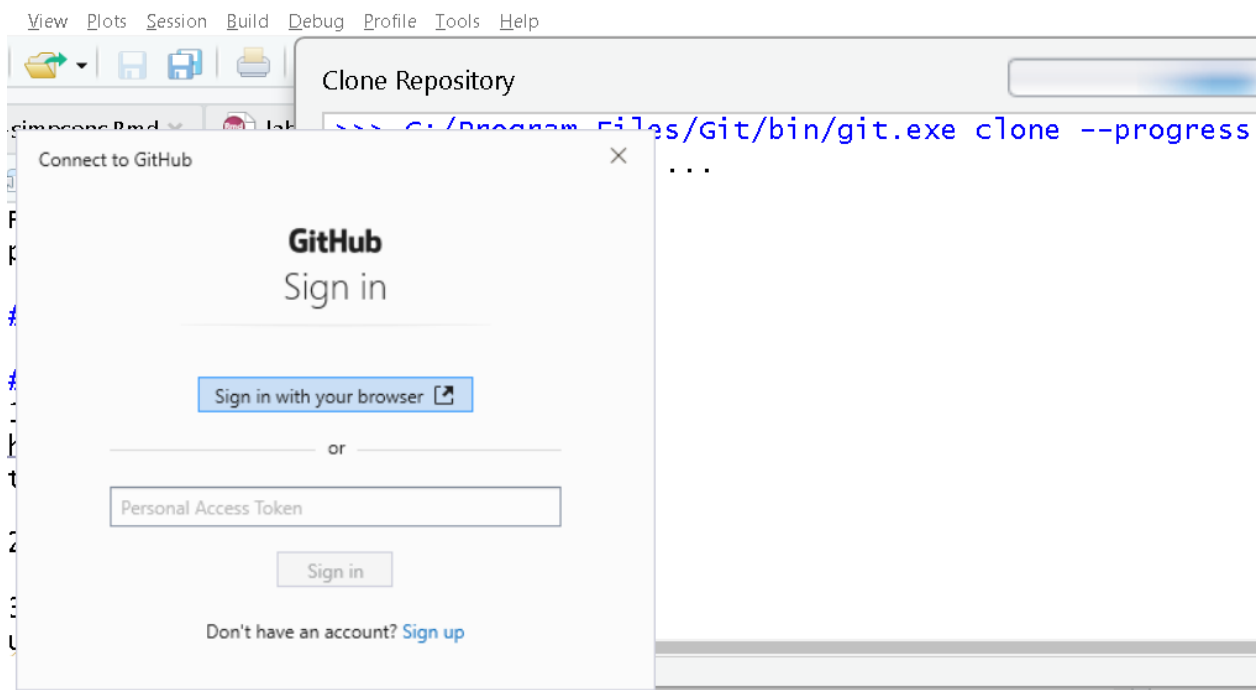
Cancel

the project.

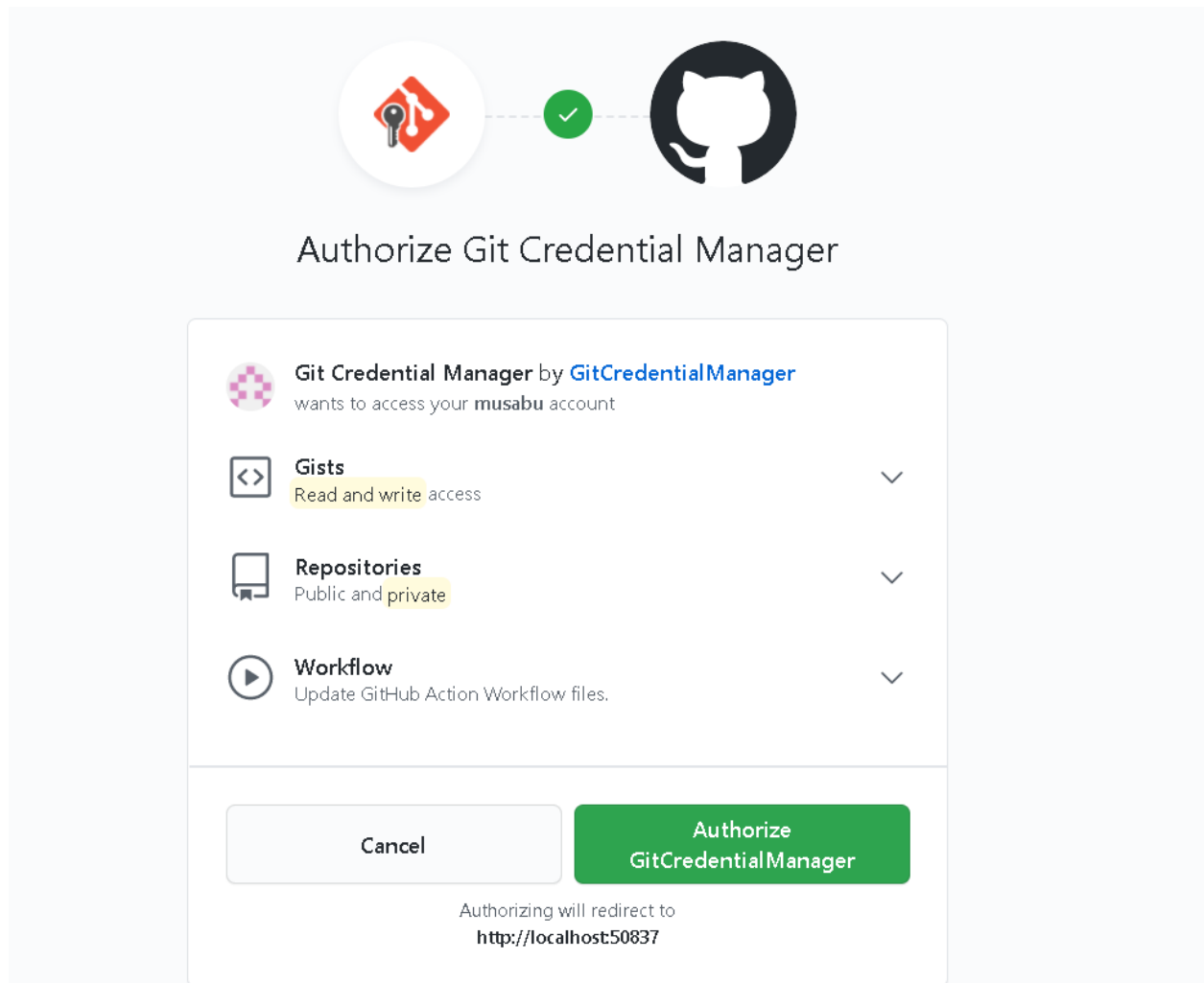


3. Select Git in the next Window (See image).

4. The next window (see image) will ask you to sign into Github with your browser. Click on 'Sign in with your browser'.



5. The Github page window will be opened on your default browser. Enter you login details if asked and then click on ‘Authorize GitCredentialsManager’ (see image).



6. Enter your password if asked and your RStudio Desktop is now linked to your Github account.

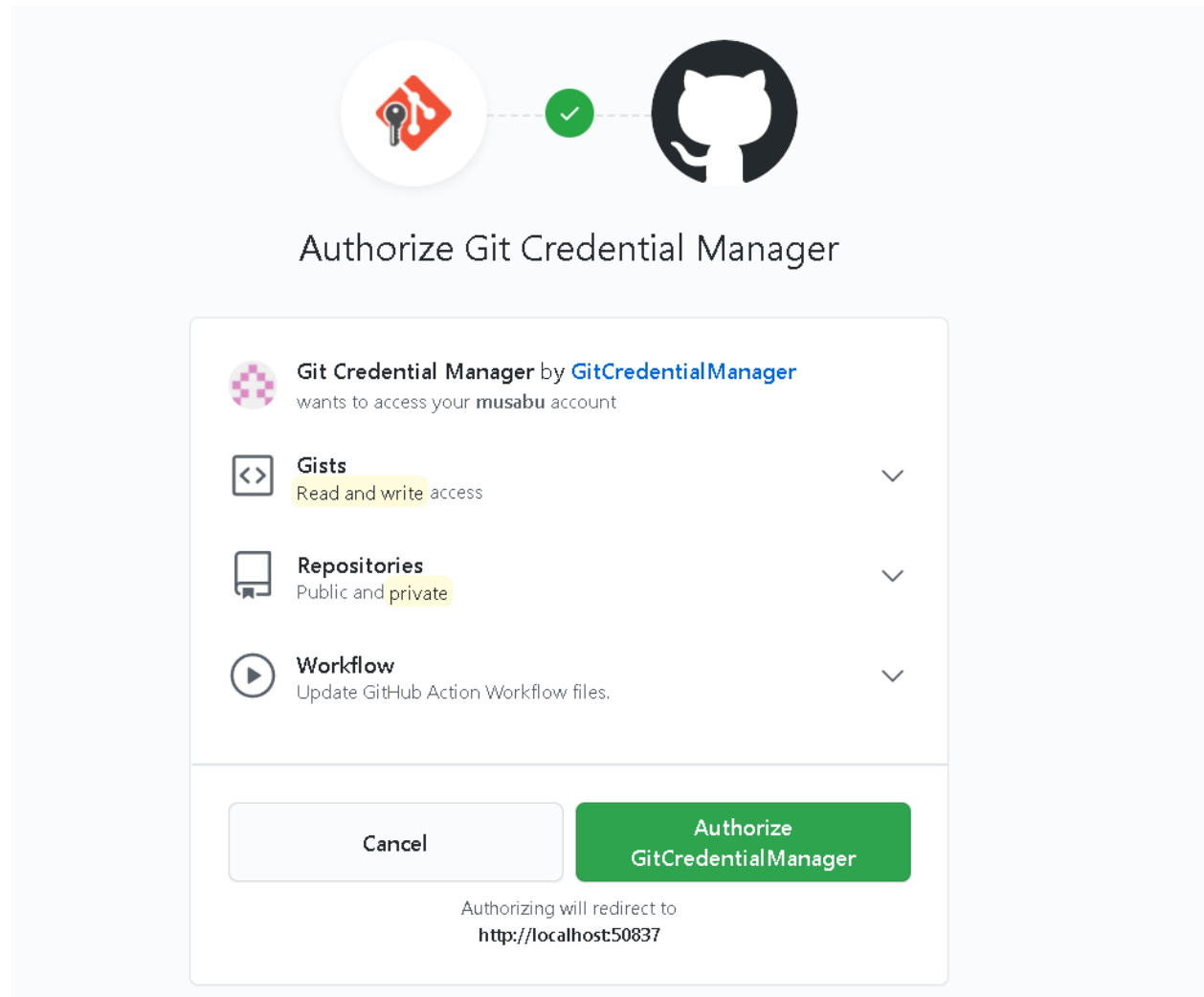
## Cloning the Project Files

### Actions by Team Leader

1. The team leader should clone the project files from [https://github.com/musabu/DSC200\\_202\\_course\\_project.git](https://github.com/musabu/DSC200_202_course_project.git) to their github account.
2. Make the repo private during the creation process.
3. Add the members of your team and your instructor (my username is musabu) as collaborators on the project.

## Actions by other Team Members

1. Accept the invitation to collaborate that was sent to your email after the Team Leader added you to the project.
2. Check your Github account to confirm that the project is now available on your Github account.
3. Click on the repository and copy the git link of the project (see image).



## Importing the Cloned Files to RStudio Desktop

1. Again, go to File -> New Project -> Version Control -> Git
2. In the Repository URL textfield, paste the link you have copied in the previous section. It will be something like this [https://github.com/team\\_leader\\_username/DSC200\\_202\\_course\\_project.git](https://github.com/team_leader_username/DSC200_202_course_project.git).  
*Note that the username must be the username of your team leader.*
3. Project Directory Name will be filled automatically. Browse and select the folder where you want the project to be saved.

4. Click on Create Project. You should have all the files downloaded to your PC. If you are asked to provide username and password at any point. Enter your Github credentials to continue.

Once you import the project into RStudio, open the file `steachers_solution.Rmd` where you will answer the question asked in the project.

## Loading Libraries

Ensure you have the `tidyverse` and `readxl` packages installed by running the 2 lines of code below. Both the team leader and members should run the code below in the console of rstudio.

```
## Warning: package 'readxl' was built under R version 4.0.4
```

We then need to load the data from the MS Excel (.xlsx) file, named `teachers_data.xlsx`, to the R environment for processing.

*Note: Only Team Leader should run the code below*

```
tdata <- read_excel("teachers_data.xlsx")
```

Wherever you see the text ‘`#Add some code`’, it means you should delete the comment and add your code for the task.

**2 Marks will be deducted for every deadline missed.**

### Task 1

(8 points)

The first thing we need to do is to tidy the data by changing the Arabic column names to English names that you could easily use in your code. There should be no space in the column name. For example, change the arabic text *Al-Sanah* to *year*. I have provided you with some English names you could use in the tasks below. You could use Google Translate to find the right English words for the rest of the variables. Ensure that there is no space for variables with two words. For instance administrative regions should be written as `administrative_regions`.

```
#Add some code
```

Team Leader should *Knit, commit, and push changes to GitHub with an appropriate commit message. Make sure to commit and push all changed files so that your Git pane is cleared up afterwards.*

**DEADLINE: END OF WEEK 9**

### Task2

(8 points)

Next Team Member should *Pull the changes made by Team Leader before proceeding.*

- (a) What are the different `administrative_regions` (Al-Mandaqah Al-idariyyah) found in the data?

#Add some code

Comments about your answer:

- (b) Based on your knowledge of the administrative regions in Saudi Arabia, is there any region that has not recruited any teacher during the period under review (1437 - 1440 Hijri)? List the Saudi regions shown in the data.

#Add some code

Comments about your answer:

The Team Member should *Knit, commit, and push changes to GitHub with an appropriate commit message.*

**DEADLINE: END OF WEEK 10**

### Task3

(20 points)

The Next Team Member should *Pull the changes made by Member 1 before proceeding.*

- (a) Saudi Arabian education is divided into different stages/levels (Al-Marhala), list the different levels showed in the data in descENDING order of frequency. (4 points)

#Add some code

Comments about your answer:

- (b) Which level recruit the most teachers? (4 points)

#Add some code

Comments about your answer:

The Team Member should *Knit, commit, and push changes to GitHub with an appropriate commit message.*

The Next Team member should *Pull the changes made by Team Leader before proceeding.*

- (c) Plot the graph of region versus number of **Saudi** teachers recruited over the period. (4 points)

#Add some code

Comments about your answer:

- (d) Plot the graph of region versus number of **Non-Saudi** teachers recruited over the period. (4 points)

#Add some code

Comments about your answer:

- (e) Plot the graph of region versus number of **all teachers** recruited over the period. (4 points)



#Add some code

Comments about your answer:

The Team Member should *Knit, commit, and push changes to GitHub with an appropriate commit message.*

**DEADLINE: END OF WEEK 12**

## Task 4

(12 points)

The Next Team member should *Pull the changes made previously before proceeding.*

- (a) How many boys and how many girls schools are listed in the data? (4 points)

#Add some code

Comments about your answer:

- (b) Between boys and girls schools, which recruit more teachers? (4 points)

#Add some code

Comments about your answer:

- (c) What is the correlation between recruitment at boys schools and recruitment of girls schools? (4 points)

#Add some code

Comments about your answer:

The Team Member should *Knit, commit, and push changes to GitHub with an appropriate commit message.*

**DEADLINE: END OF WEEK 13**

## Task 5

(12 points)

The Next Team member should *Pull the changes made previously before proceeding.*

- (a) How many teachers are recruited each year, 1437, 1438, 1439, and 1440?

#Add some code

Comments about your answer:

- (b) Which region recruited the most teachers in each of the years under review? (4 points)

```
#Add some code
```

Comments about your answer:

- (c) Plot the graph of number of teachers recruited vs the number of schools in an administrative region? What type or relationship exist between the two variables? (4 points)

```
#Add some code
```

Comments about your answer:

The Team Member should *Knit, commit, and push changes to GitHub with an appropriate commit message.*

**DEADLINE: END OF WEEK 14**

**If you are here, then congratulations!! You have completed the DSC 200 course project.**