

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

Data from Saudi Open Data Office

Musab Isah

11 April 2022

## Project's Total Points (64)

### Weightage (20%)

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. Please come to my office with your laptop if you struggle with installing RStudio Desktop and Git.**

### Installing Git

For Windows users:

1. Download Git exe file from this link <https://git-scm.com/download/win>
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 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.

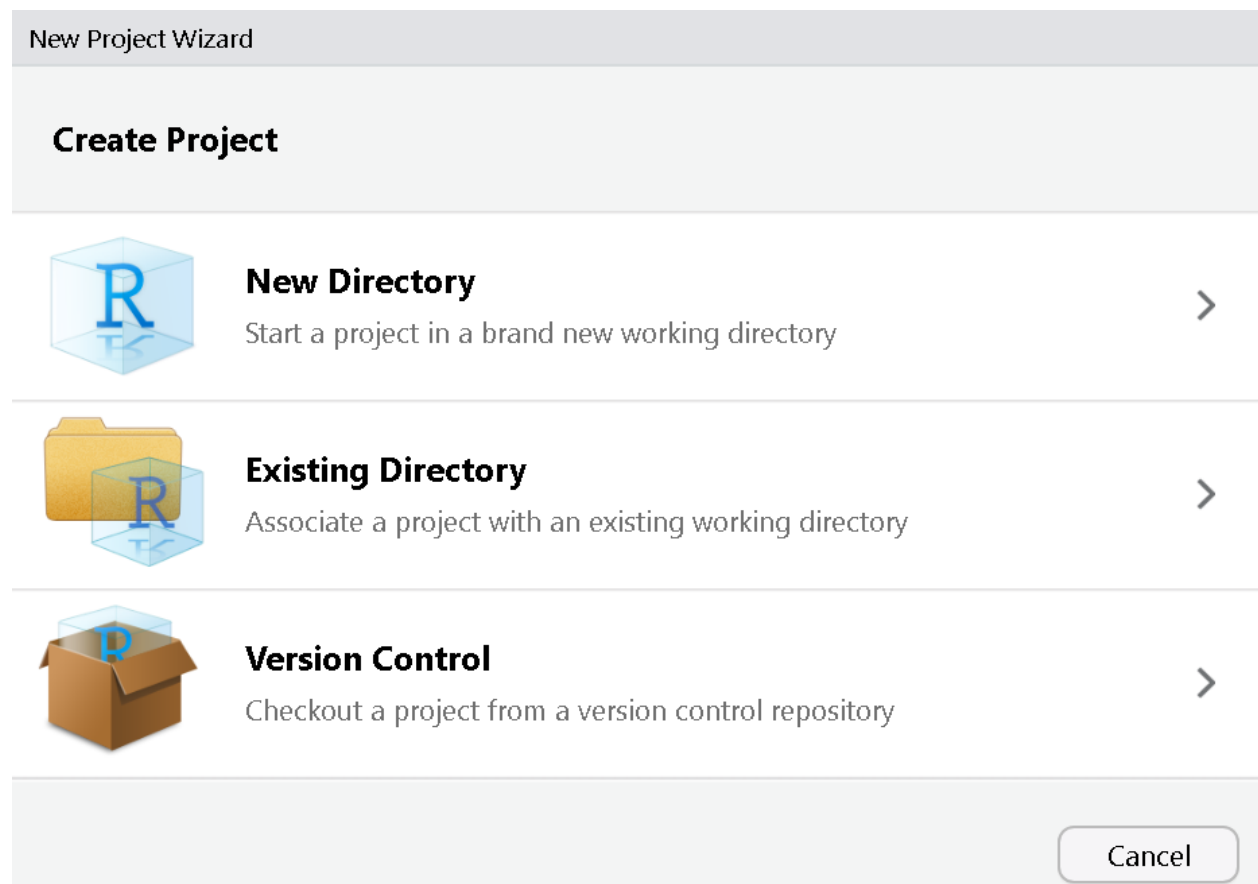
## Installing RStudio Desktop

For Windows users:

Download RStudio for your OS type from here <https://www.rstudio.com/products/rstudio/download/#download>

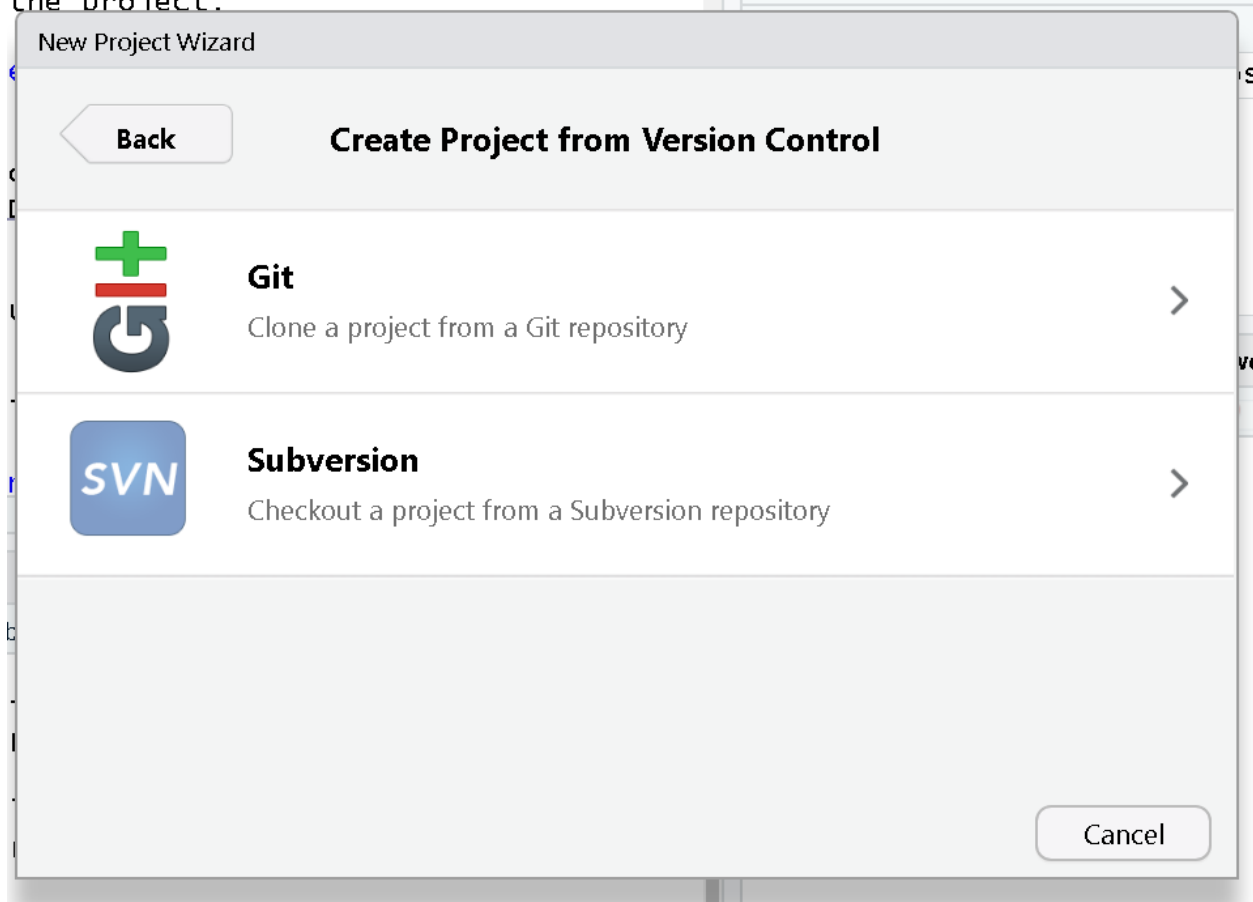
## Linking Github to RStudio Desktop

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'.

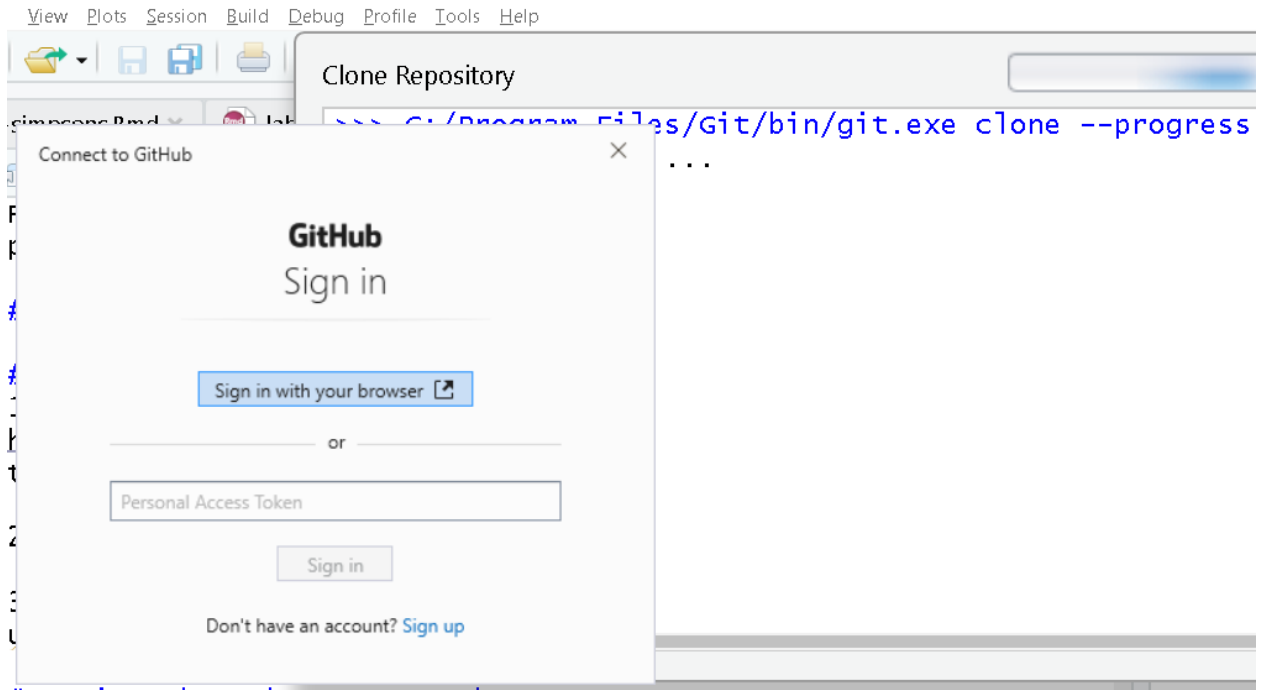


3. Select Git in the next Window (See image).

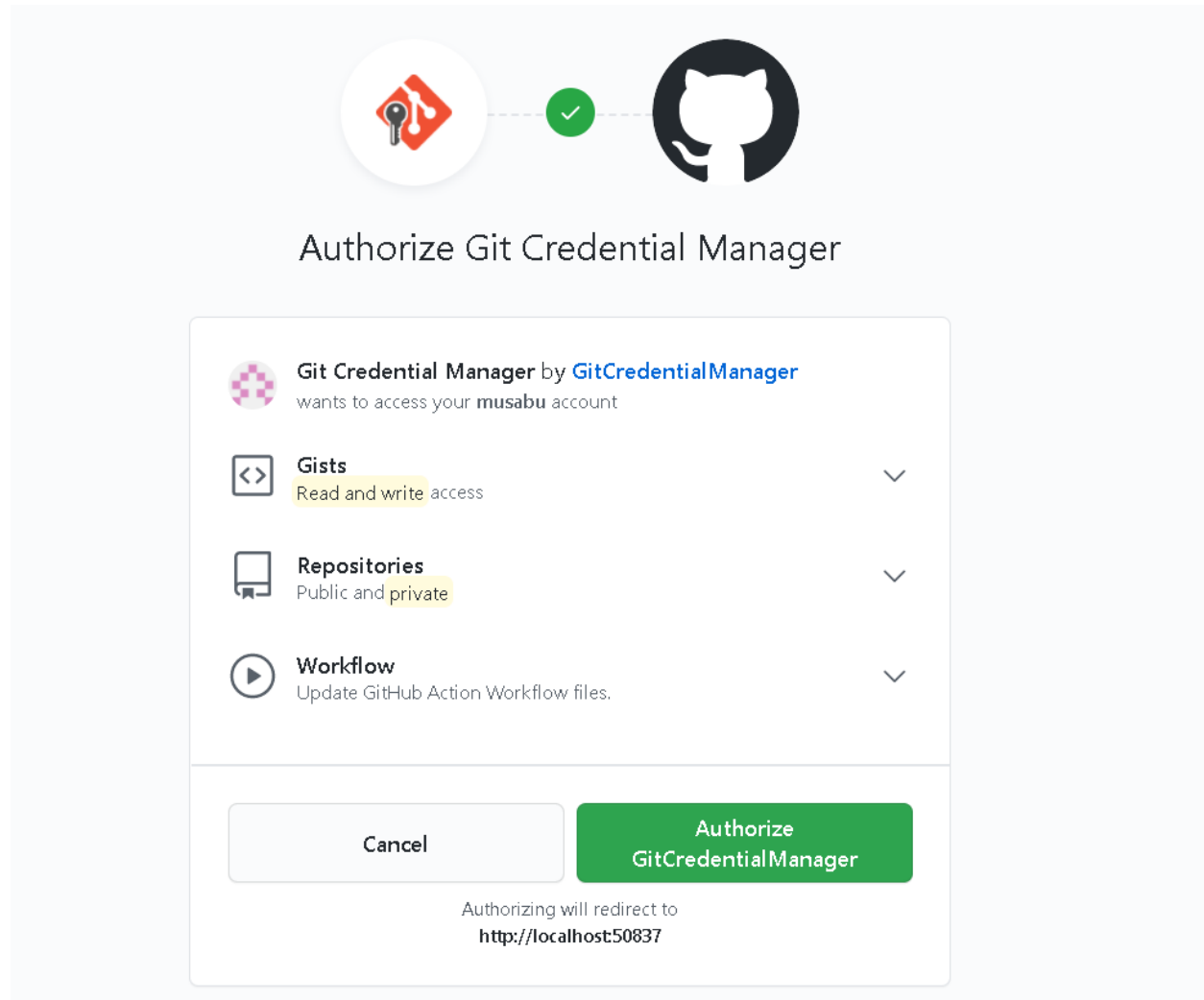
the project.



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.

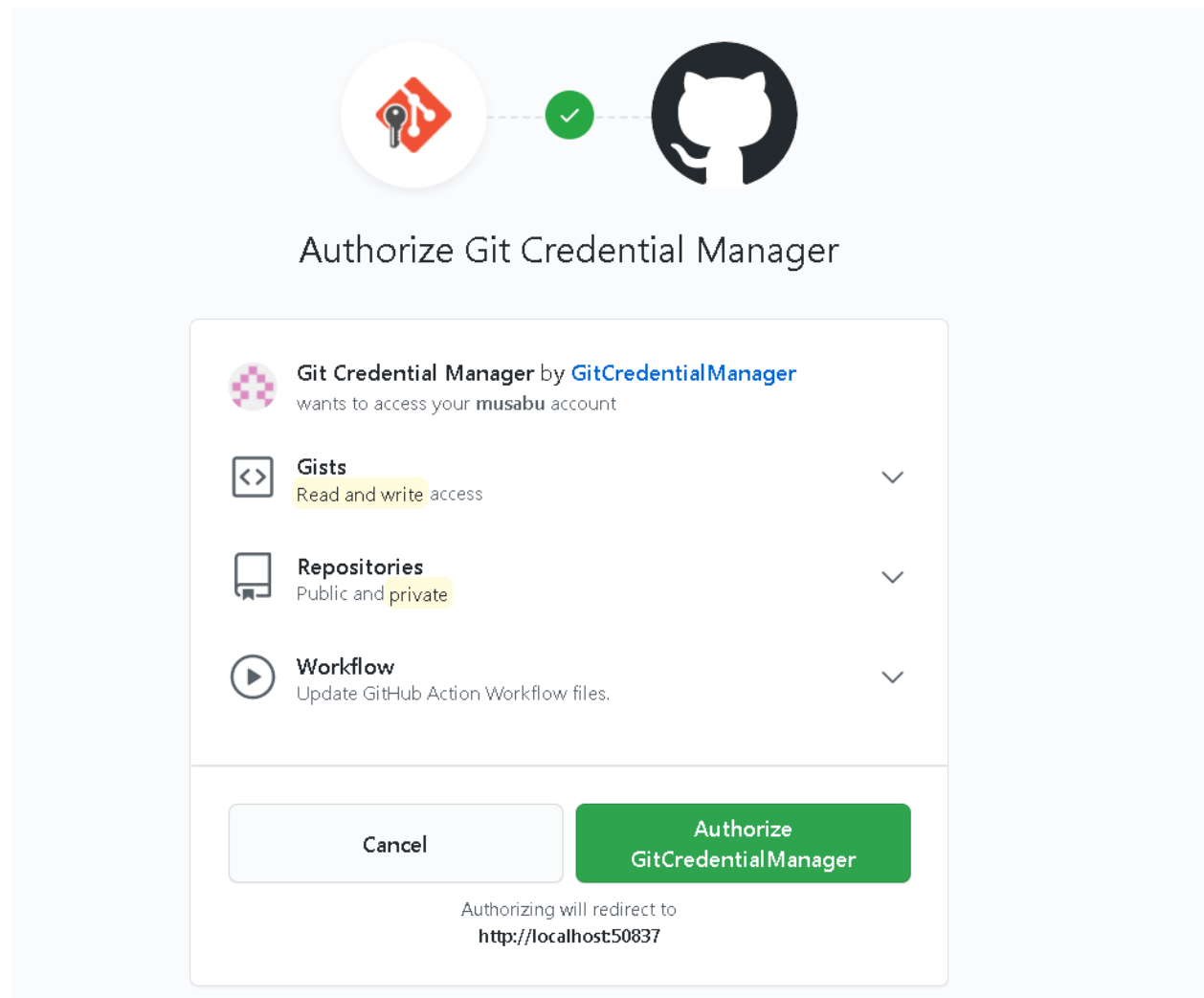
## Cloning the Project Files

### Actions by Team Leader

1. The team leader should clone the project files from [https://github.com/musabu/DSC200\\_212\\_course\\_\\_project.git](https://github.com/musabu/DSC200_212_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.
4. Create a personal access token (PAT) and share it with your team members.

## 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. Click on the repository cloned by your team leader and copy the git link of the project (see image).



3. Import the cloned files to RStudio Desktop by:

- going to File -> New Project -> Version Control -> Git
- 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\\_212\\_course\\_project.git](https://github.com/team_leader_username/DSC200_212_course_project.git).

*Note that the username must be the username of your team leader.* 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 username and the PAT provided by your team leader.

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

**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 - DEADLINE: 16 April 2022**

(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 English names for all the columns below. Ensure that there is no space for variables with two words. For instance administrative regions should be written as `administrative_regions`.

(“ID”, “Administrative\_Region”, “Levels”, “Authority”, “Education\_Office”, “Sex”, “School\_Type”, “Education\_Type”, “School\_System”, “Saudis”, “Expats”, “Total\_Numb\_Teachers”, “Year”)

### **Task2 - DEADLINE: 23 April 2022**

(8 points)

- (a) What are the different `administrative_regions` found in the data? (4 points)
- (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 to see if all of them are listed. (4 points)

### **Task3 - DEADLINE: 12 May 2022**

(20 points)

- (a) Saudi Arabian education is divided into different stages/levels, list the different levels showed in the data in descending order of frequency. (3 points)
- (b) Which level recruit the most teachers?? (2 points)
- (c) Plot a scatter graph of region versus number of **Saudi** teachers recruited over the period. (5 points)
- (d) Plot a scatter graph of region versus number of **Non-Saudi** teachers recruited over the period. (5 points)
- (e) Plot the graph of region versus number of **all teachers** recruited over the period. (5 points)

#### **Task 4 - DEADLINE: 15 May 2022**

(12 points)

- (a) How many teachers are recruited in boys' and girls' schools? (3 points)
- (b) Between boys and girls schools, which recruit more teachers? (1 point)
- (c) What is the correlation between recruitment at boys schools and recruitment of girls schools? (8 points)

#### **Task 5 - DEADLINE: 19 May 2022**

(16 points)

- (a) How many teachers are recruited each year, 1437, 1438, 1439, and 1440? (8 points)
- (b) Which region recruited the most teachers in each of the years under review? (8 points)

**End of DSC 200 course project.**