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.

Cloning the Project File

Actions by Team Leader

1. The team leader should clone the project files from 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. Clone the project files from the github page of your team leader https://github.com/team_leader_username/DSC200_202_course_project.git
2. Make the repo private during the creation process.

Action by Team Leader and members

Import the files to your rstudio cloud. The R Markdown for solving the tasks is called saudi_teachers.Rmd.

Packages

Once you import the project into rstudio.cloud, open the file saudi_teachers.Rmd. 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 `install.packages("tidyverse")` and `install.packages("readxl")`

Loading Libraries

Here we load the libraries that we will need to read the MS Excel file and to manipulate the dataset. Both the team leader and members should run the code below.

```
## 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.

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.*

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.*

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.*

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.*

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.*

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