

AIT 580 – Assignment 15 – Cloud Computing

Due Sunday December 8, 2019

Submission to BOTH your Github and Blackboard

- **Overview**
 - Take a screenshot for each step of the followings and paste it in a Word Document in a chronological order. Part of the assignment is to submit the series of screenshots and the explanations of them as a PDF document.
 - If you have any difficulties in signing up to Google Cloud, let the instructor know as soon as possible (i.e., don't wait until the assignment due).
- **Getting Ready**
 - Pull or clone the course Git repo: <https://github.com/myeong/AIT580>
 - Copy the “assignment_15” folder to your own assignment Git repo.
 - Rename the “assignment_15.R” to “yourLastName_GNumber.R”.
- **Task 1 (You don't need screenshots for these tasks)**
 - a. Complete the tasks inside the given R file. There are three tasks that you need to do.
 - i. Printing out “titles” and “dates” from the scrapped data from NYTimes.
 - ii. Saving the titles and dates of articles as a CSV file (i.e., NYT_titles.csv).
 - iii. Commit and push the R and CSV files to your Github repository online.
- **Task 2 (You need a screenshot for each step in Task 2)**
 - b. Create a Google Cloud account. Then, \$300 credit is given (don't pay out of your pocket).
 - c. Create a Compute Engine instance. You can create a basic instance.
 - i. General-purpose, g1-small (1 vCPU, 1.7GB memory)
 - ii. 10 GB standard persistent disk
 - iii. Ubuntu 16.14 for the OS (i.e., Linux)
 - iv. In the options, check “Allow HTTP traffic”, “Allow HTTPS traffic”, and “Allow full access to all Cloud APIs”
 - v. Click “Create” button at the bottom.
 - d. After the Cloud instance is created, SSH into the server by clicking “SSH” button in the Console page (i.e., VM instances page under the “Compute Engine” menu). Your screenshot needs to show the IP address of your VM.
 - e. Clone your assignment Github repository to your VM machine by typing (hint: using the Git URL with “https” would work better if you don't have SSH key linked):

```
git clone [your Git repo URL]
```
 - f. Update the Linux APT repositories by typing (this updates your VM's software list):

```
sudo apt-get update
```
 - g. Install R inside the server by typing (this will take a while. Type “Y” if it asks whether to proceed with installation):

```
sudo apt install r-base
```
 - h. Install the following libraries by typing (these are essential for installing R packages):

```
sudo apt-get install libxml2-dev
```

```
sudo apt-get install libssl-dev
```

```
sudo apt-get install libcurl4-openssl-dev
```

- i. Get into your own folder for Assignment 15 using the “cd” command.
(e.g., `cd AIT580/assignment_15/`)
- j. Type “R” and hit the Enter key, so you can enter the R command line.
- k. Install required R packages by typing the followings (these may take some time. You can select “1” for the mirror):

```
install.packages("selectr")
```

```
install.packages("xml2")
```

```
install.packages("rvest")
```

- l. After installing the two packages, type `q()` and hit the Enter key to exit from the R command line (you don’t need to save the workplace image).
- m. Run the R script to see if it runs well. If you see your desired output from the command line, you’re successful. Your screenshot for this step should show all the printed outputs.
e.g., `Rscript assignment_15.R`
- n. Once you’re done, you just completed running R scripts on a cloud VM. Put all the screenshots in a MS Word document, and save it as a PDF file. Then, upload it to Blackboard.

Submission Guideline: Submission for this assignment is two ways.

1. Task 1: After writing scripts in a R file and its output in a CSV file, commit and push it to your own Git repository that you made for Assignment 3 (you shouldn’t use other repos. If the instructor or TA cannot see your repository, it is regarded as no-submission).
2. After pushing your assignment to your Github repository, submit the screenshot of your Github repository page that shows your uploaded files through Blackboard.
3. Task 2: upload the PDF file that contains all the screenshots of Task 2 to Blackboard.

After You’re Done with Assignment Submission

Make sure you turn off your VM and delete the instance. Otherwise, it could keep consuming your credits.