

courseWorkCS8631

Project Title: Cyber Security: Safety At Home, Online, and in Life

Project Overview:

This project investigates and provides insights on the dataset collected for the Cyber Security: (Safety At Home, Online, and in Life) course developed by Newcastle University and offered by the Future Learn Platform which is a massive open online course (MOOC). The course is 3 weeks long and was free to access. So we have to come up with two research question, and carry out our analysis by implement CRISP-DM and follow the steps to find a solution for the research question. Two cycles of CRISP-DM is carried out for this project.

Repository Overview:

Directory structure:

The repository is named as 'courseWorkCS8631', it has the following files

data -> this data folder has all the raw datafiles of the futurelearn cache -> cache has the pre-processed and intermediate datas munge -> munge folder has the files related to the pre-processing steps

1. 01-dataCleaning.R
2. 02-dataIntegration.R
3. 03-dataTransformation.R

graphs -> this folder has all the images of the visualisations that are plotted for the dataset for analysis purpose report -> this folder has the analysis report which needed to be knitted to visualise the output pdf of the report, and here is the main analysis is stored Analysis.Rmd Analysis.pdf

The README.md file has all the basic information about the project installation and also about the file structure. The gitlogdata.txt has the git commit history with which we can understand at which stage what all files and actions were carried out

The overall structure looks like > courseWorkCS8631 |__ cache > Intermediate dataset |__ data > Raw data files |__ gitlogdata.txt |__ graphs > images of the visualisations |__ munge > files related to pre processing of raw data |__ README.md |__ reports > the main analysis

Applicable Instruction:

Setup Instruction:

You will need Rstudio to execute this , so there is an attached link given below which you can follow to install the R studios based on your operating system

<https://rstudio-education.github.io/hopr/starting.html> <https://teacherscollege.screenstepslive.com/a/1135059-install-r-and-rstudio-for-mac>

After installing the Rstudio, Execution Instruction:

1. download the zip file and extract all the files into a new folder in your desired destination
2. open r studio in your system and file -> open project and select the directory in which you have extracted the project folder
3. set the working directory session

To load your new project, you'll first need to `setwd()` into the directory where this README file is located. For this you can choose session and select 'set working directory' and choose whichever directory is your current project.

Alt text

Figure 1: Alt text

4. open the report folder -> choose the analysisReport.rmd file and knit it
5. The pop up with saying 'Rendering Rmarkdown documents required to updated version of the folowing packages , do you want to instal it ?' give an yes to it.
6. If it says anyof the package is not installed I request to execute the command `install.packages("package name")` `library(packagename)` `renv::snapshot()` `renv::status()`
7. After this navigate to the reports folder -> AnalysisReport.pdf and scroll to line number 112 and you will find a green run icon , execute it and same process for line number 116 and you will find few packages getting installed wait for the process to complete.
8. Now, in the top section below the file name , there will be an option to knit the file (8th option) click it and choose knit to pdf and wait for it to generate a pdf file with the visualizations.

Additional resources: For more details about ProjectTemplate, see <http://projecttemplate.net>