CM 1110 Fundamentals of Mathematics and Statistics

Assignment 1

11/05/2020

Introduction

We are still facing the global crisis of the COVID-19 pandemic. As individuals we can do our part in keeping others safe by following the rules suggested by the government and WHO.

As part of the data science community we can also provide our contribution, by analyzing the data and helping to communicate it. Such type of support is high demand among government and related institutions.

The purpose of this assignment is to fulfill that role. At the end of this assignment your blog post and the corresponding dataset will appear on a separate webpage devoted to this project.

Assignment guidelines:

1. Data preparation

~~Select (or create) a dataset related to Sri Lanka on the topic of COVID-19 from any sector~~ **Select a dataset on the topic of COVID-19 from any sector** (eg: education, agriculture, health, transportation, business, export, environment, telecommunication, retail, etc.) **You don't have to limit your research to Sri Lanka. You can work with any dataset from any region or country on the topic of COVID-19.**

Save the dataset as a csv file in the following format

`Data_(Sector)_(YourStudentIDNumber).csv`

Example:
If you select a dataset from education sector and your student ID number is XX1111111

Then your csv data file can be saved as `Data_Education_XX1111111.csv`

3. Data documentation

The documentation (i.e., the "Readme.txt" file) that accompanies each project data set is as important as the data itself. This information permits collaborators and other analysts to understand any limitations or special characteristics of the data that may impact its use. The following outline and content is recommended and should be adhered to as closely as possible to make the documentation consistent across all data sets.

Data Set Documentation/Readme Outline:

a) Save the text file in the following format

`Readme_(Sector)_(YourStudentIDNumber).txt`

b) Your readme file should contains the following

- **Sector**:

- **Region or Country**:

- **Title**: A descriptive title that match the content of the data set

- **Author**: Your name

- **Author_ID**: Student Index Number

- **Email**: Email address

- **Description**: Introduction to your dataset (Max 200 words)

- **Time**: Time period covered by the data: `Start date` and `end data`. (If not application put NA)

- **Source**: Data source. Provide any web address references (i.e., links for any publications, additional documentation such as Project web site,if available). If you create your own dataset then provide a description of data collection process)

- **Format**: `A data frame with xx rows and yy variables.`

Please replace `xx` and `yy` according to your dataset.

- Provide a **SEPERATE** description for **EACH column** of your dataset using the following format.

```
\item{column_1_name}{Varible description }
\item{column_2_name}{Varible description }
\item{column_3_name}{Varible description }
...
```

- **Other**: Other Remarks. Any other information related to the dataset, if any

3. Do a small analysis with the dataset using R, Excel, Python or any other preferred programming language.

4. Write a small blog post based on your findings.

5. References – This is the last section of your blog post. Here you should provide all the references that you used for your research and writing.

Ethical writers ALWAYS acknowledge the contributions of others and the source of their ideas. Any verbatim text taken from another author must be enclosed in quotation marks. Please acknowledge every source you use in your writing, whether you paraphrase it, summarize it or enclose it in quotations.

6. At the end of this activity you will have the following three documents

   1. `Data_(Sector)_(YourStudentIDNumber).csv`
   2. `Readme_(Sector)_(YourStudentIDNumber).txt`
   3. `Blogpost_(Sector)_(YourStudentIDNumber).doc`

I'll update you about the submission process later.

Deadline: 15 June 2020

Please be safe during this period and help one another. Your safety is my main priority. If you have any issues with submitting your assignment due to any unexpected circumstances, you can apply for an extension or special consideration. For special consideration or deadline extensions, please send me an email: priyangad@uom.lk informing your request with the email title: CM1110_Assignment1_deadline_extension_(yourIndexNumber)

Final Marks

This assignment carries 10 points towards your final grade

Discussion Forum

I initiated a discussion forum for this assignment (`Assignment 1- Forum`). Use this forum to discuss any questions, concerns related to this project. I am happy to participate in the forum discussion, but I want you to take part in this discussion and help each other.