

Vulnerability Data Analysis

22 March, 2021

In this exercise, we will explore SDG's 13.1.1 indicator, [XXX], associated with target 13.1, [XXX]. The indicator's associated data is from the Emergency Events Database (EM-DAT), managed by the Centre for Research on the Epidemiology of Disasters (CRED).¹

Load libraries

First, we need to load the following libraries:

```
library(tidyverse)
library(readxl)
library(janitor)
library(maps)
library(ggmap)
library(ggplot2)
library(extrafont)
library(ggrepel)
library(directlabels)
library(Hmisc)
library(knitr)
library(kableExtra)
library(fuzzyjoin)
library(scales)
library(sf)
```

Download the dataset

You can access the full database here by submitting a query for the data you need. You will need to register with the database first before you can submit a query. For the purpose of this exercise, we will provide you with the full dataset that starts from 1970. This dataset is as of March 22, 2021.

Import dataset

Using the package `readxl` we can import the dataset that's in Excel format. If you preview the raw file in Excel, you will find that the first six rows are used as description of the dataset, therefore we need to make sure that R does not read rows that are not part of the dataset. The function `read_excel` provides a convenient option called `skip`, which tells R how many rows to skip before reading the dataset. In this case, we need to skip six rows.

```
em_dat <- readxl::read_excel("emdat_public_2021_03_22_query_uid-2i1xa8_raw.xlsx",
  skip = 6)
```

Clean dataset

Using the `janitor` package, we can use the function `clean_names` to create consistent-looking variable names.

¹“About EM-DAT” (Centre for Research on the Epidemiology of Disasters, March 2021), <https://www.emdat.be/about>.

```
em_dat <- janitor::clean_names(em_dat)
```

Instead of recreating the object `em_dat` again, we can combine the previous two functions using the ‘pipe’ operator, `%>%`, which is loaded with the `tidyverse` package. This operator uses the previous output as the new input of the subsequent function.

```
em_dat <- readxl::read_excel("emdat_public_2021_03_22_query_uid-2i1xa8_raw.xlsx",  
  skip = 6) %>% janitor::clean_names()
```

In the above code chunk, we did the following:

1. Imported the dataset using `read_excel` function.
2. Took that dataset and cleaned all the variables’ names using the `clean_names` function.
3. Called the output of the previous two processes `em_dat`.

The `pipe` operator is a powerful function that can reduce the amount of code you need to write, but you need to make sure

References

CRED. “About EM-DAT.” Centre for Research on the Epidemiology of Disasters, March 2021. <https://www.emdat.be/about>.