

# Day Four: Data Fundamentals and Intro to RStudio Environment

SDS 192: Introduction to Data Science

Lindsay Poirier Statistical & Data Sciences, Smith College

Spring 2022

```
global_landslide <- read.csv("https://data.nasa.gov/api/views/dd9e-wu2v/rows.csv")
```

1. Run the code chunk above, and then start to review the dataset using strategies we discussed in class. You might check its dimensions (number of columns and number of row), column names, and print the first 6 rows. You can use the code chunk below to help you.

```
# Check number of columns  
  
# Check the column names  
  
# Print the first six rows  
  
# View the data frame
```

2. Find one nominal variable in the dataset, one ordinal, one discrete, and one continuous. Uncomment the lines below, and store the entire column of values in the variables.

```
#nominal <-  
#ordinal <-  
#discrete <-  
#continuous <-
```

3. Check the class of each of these vectors. Are they what you expected?

```
# Check the class of the nominal variable below.  
  
# Check the class of the ordinal variable below.  
  
# Check the class of the continuous variable below.  
  
# Check the class of the discrete variable below.
```

4. Determine the number of distinct possible values in the nominal and ordinal variables.

```
# Calculate the number of distinct values in the nominal variable.  
  
# Calculate the number of distinct values in the ordinal variable.
```