

Assignment II Report on “Occurrences of Landslides”

Table of Contents

Introduction 3

Methods 3

Results.....3-8

Conclusion 8

References9-10

Appendices.....10-11

Introduction

A **landslide** is the movement of rock, earth, or debris down a sloped section of land. Several things can trigger landslides like Rain, Downpour, Earthquake, Dam Embankment Collapse, floods, mining, snowfall melt, volcano etc. The effects of landslides include property damage, injuries and fatalities and also there is an impact on a variety of resources like forests and roadways. The landslides can be distinguished by the types of material (rocks/debris etc.) involved and the mode of movement (rotational/translational etc).

The purpose of the report is to analyse the dataset concerning Global Landslides in order to draw inferences regarding common locations/ countries/ continents where the landslides occur and to compare the number of injured and fatalities under different landslide types and their trigger sources.

Methods

The dataset under study is associated with the landslide events occurring around the world, irrespective of size, type, impacts or location. The data has been collected over the years since 1988 up till 2016. The data has been taken from data.nasa.gov

For data analysis, the variables majorly used are:

Day, Month, Year of occurrence of landslide; continent code, country code, country name, location accuracy (in km), latitude and longitude, nearest places and location description of the place where the landslide has occurred

Then there are columns like unique id, cat id, cat src, size, type, trigger, storm name, injuries and fatalities caused for describing the landslide.

The method that has been used for analysing the data is SAS. Formulas such as PROC FREQ, PROC SORT, and PROC CONTENTS have been used.

Results

❖ Overview of the data set

The CONTENTS Procedure

Data Set Name	WORK.GLOBAL_LANDSLIDE	Observations	9471
Member Type	DATA	Variables	28
Engine	V9	Indexes	0
Created	09/17/2017 11:25:41	Observation Length	224
Last Modified	09/17/2017 11:25:41	Deleted Observations	0
Protection		Compressed	NO
Data Set Type		Sorted	NO
Label			
Data Representation	SOLARIS_X86_64, LINUX_X86_64, ALPHA_TRU64, LINUX_IA64		
Encoding	utf-8 Unicode (UTF-8)		

Engine/Host Dependent Information

Data Set Page Size	131072
Number of Data Set Pages	17
First Data Page	1
Max Obs per Page	584
Obs in First Data Page	561
Number of Data Set Repairs	0
Filename	/saswork/SAS_workACF700008D24_odaws01-prod-sg/SAS_work114900008D24_odaws01-prod-sg/global_landslide.sas7bdat
Release Created	9.0401M4
Host Created	Linux
Inode Number	14680078
Access Permission	rw-r--r--
Owner Name	s36314420
File Size	2MB
File Size (bytes)	2359296

Alphabetic List of Variables and Attributes

#	Variable	Type	Len
2	Day	Num	8
3	Month	Num	8
4	Year	Num	8
21	adminname1	Char	8
22	adminname2	Char	8
17	cat_id	Num	8
16	cat_src	Char	8
25	continent_code	Char	8
5	country	Char	8
24	country_code	Char	8
18	country_name	Char	8
20	distance	Num	8
11	fatalities	Num	8
7	htype	Char	8
1	id	Num	8
12	injuries	Num	8
27	latitude	Num	8
14	loc_acc_km	Num	8
13	loc_desc	Char	8
28	longitude	Num	8
15	ls_size	Char	8
8	ltype	Char	8
19	near	Char	8
6	nplaces	Char	8
23	pop	Num	8
10	storm_name	Char	8
9	trigger	Char	8
26	version	Num	8

Table 1: Shows variables in the dataset along with their attributes (Type and Length)

The dataset has 9471 observations and 28 variables.

❖ Considering oldest three records in the given dataset:

Three oldest Landslide Records																			
Obs	id	Day	Month	Year	country	nplaces	htype	ltype	trigger	storm_name	fatalities	injuries	loc_desc	loc_acc_km	ls_size	cat_src	cat_id	country_name	near
1	6306	0	1	1900		Coalbank	landslid	Landslid	Construc		0	0	Above_gi	5	Small	glc	6306	United S	Bunker H
2	6436	7	11	1988		Sta. Ros	landslid	Landslid	Continuo		0	0	Unknown	25	Medium	glc	6436	Philippi	Pasacao
3	7331	4	6	1993		Holbeck	landslid	Other	Rain		0	0	Deforest	1	Medium	glc	7331	United K	Scarboro

Table 2: Information for the three oldest landslide records depicting time, location, type, size and other details

As per the given dataset the oldest landslide in the given time period took place in Philippines in 1988. The following landslide was recorded for a location near Bunker Hill, United States in 1900. Both the above incidents have been the result of construction in or near those locations. The next landslide was registered for United Kingdom for the year 1993 and was caused by rain. But none of these three listed cases involved any case of injuries or fatalities.

❖ Comparing the number of landslides that have occurred in each continent for each landslide size category is as follows:

The FREQ Procedure						
Frequency Percent	Table of continent_code by ls_size					
	continent_code	ls_size				Total
		Large	Medium	Small	Very_Lar	unknown
	AF	32 0.37	145 1.68	19 0.22	5 0.06	0 0.00
	AS	382 4.42	3389 39.20	376 4.35	47 0.54	13 0.15
	EU	28 0.32	342 3.96	132 1.53	3 0.03	4 0.05
	NA	94 1.09	1574 18.21	1301 15.05	15 0.17	4 0.05
	OC	17 0.20	195 2.26	69 0.80	1 0.01	1 0.01
	SA	62 0.72	356 4.12	26 0.30	13 0.15	0 0.00
	Total	615 7.11	6001 69.42	1923 22.24	84 0.97	22 0.25
Frequency Missing = 826						

Table 3: Frequency Distribution for the number of landslides that have occurred in each continent (Africa (AF), Asia (AS), Europe (EU), North America (NA), Oceania (OC) and South America (SA)) for each landslide size category (viz. Very Large, Large, Medium, Small and unknown)

Largest number of landslides (48.66%), have been registered in Asia followed by North America (34.56%). The least have occurred in Africa (2.33%).

Most of the landslide occurrences have been medium sized (69.42%), followed by small sized with recorded 1923 cases (22.24%). The Very Large sized land slide occurrences have been very few contributing 0.97%

Note: The blank entries under landslide size (ls_size) have been omitted from the frequency table.

- ❖ **Relationship between the number of injuries and number of fatalities caused by each landslide.**

The CORR Procedure

2 Variables:	injuries fatalities	
---------------------	---------------------	--

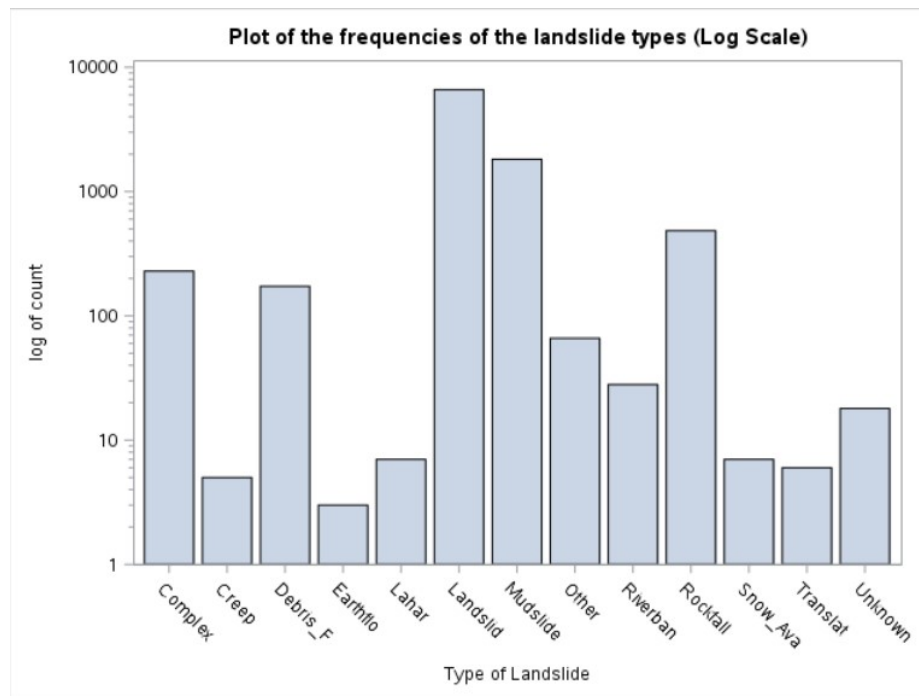
Pearson Correlation Coefficients		
Number of Observations		
	injuries	fatalities
injuries	1.00000 3925	0.14242 3918
fatalities	0.14242 3918	1.00000 8151

Table 4: Correlation between injuries and fatalities for each landslide

There is a weak relationship between injuries and fatalities implying that in most of the cases the injuries didn't end up in fatalities. Most of the lives were saved despite getting injured by the landslide.

The result is based on 3918 records where there were non-missing values for both injuries and fatalities. The entries with either missing value for injuries or for fatalities or both have been ignored. By default, PROC CORR uses *pair-wise deletion* when observations contain missing values. PROC CORR includes all non-missing pairs of values for each pair of variables in the statistical computations.

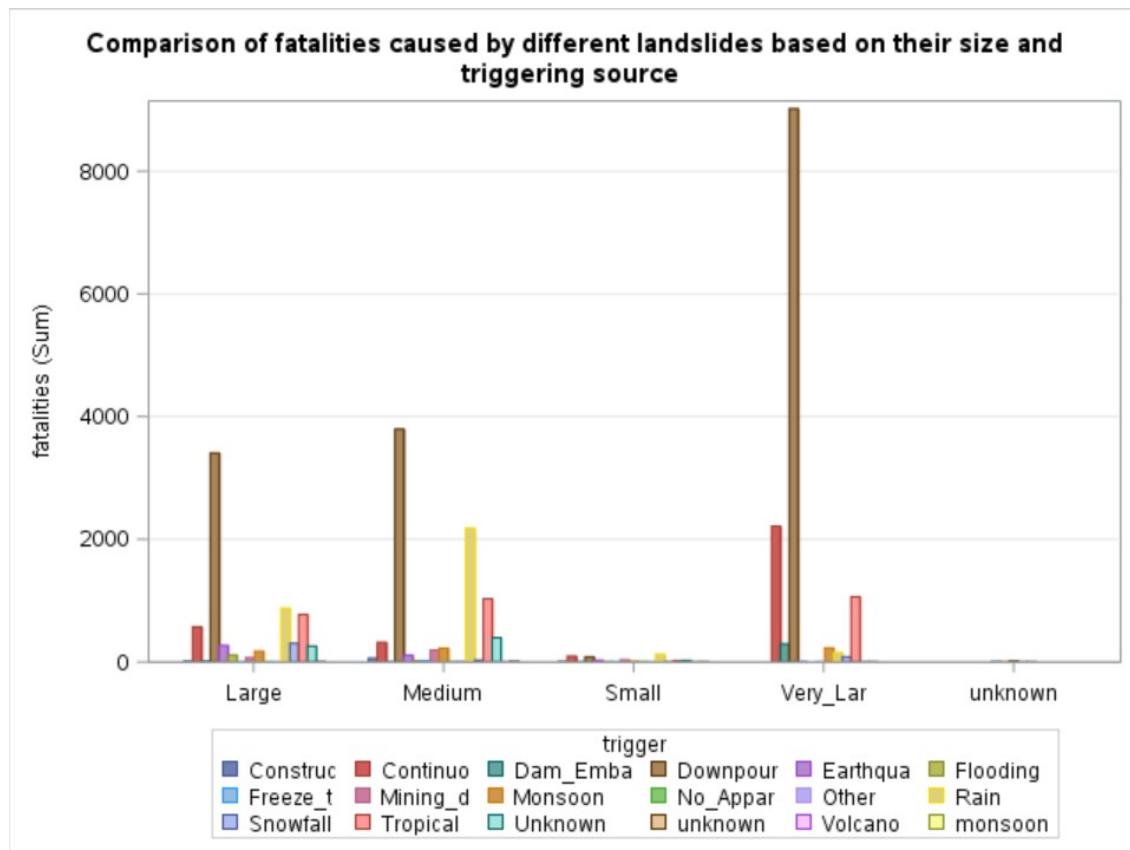
- ❖ **Getting counts of occurrences of various landslide types to know which ones are more common**



Plot 1: Bar plot showing the counts of various landslide types on log scale

As seen from the plot, the largest number of landslides occurrences is for land and mudslide type and the least one are of Earthflow type. Even rockfall, complex and Debris flow type landslides are quite common.

- ❖ **Comparing the number of fatalities caused by different landslide triggering sources and the size**



Plot 2: Stacked Bar chart showing the comparison between fatalities caused by different trigger sources and sizes of landslides

Looking on the plot we can say that:

Downpour, Rain, tropical cyclone and Continuous rain triggered landslides cause more deaths in comparison to other triggers.

For very large sized landslides more than 8000 deaths have been caused by downpour, around 2000 and 1000 by continuous rain and tropical cyclone respectively.

For Large sized land slide majority fatalities (>3000) are caused by downpour and rain (>500).

For medium sized landslide major triggers causing life losses are downpour (around 4000), rain (>2000) and tropical cyclone (~1000).

Conclusion

Amongst all the continents Asia has witnessed majority of landslides while Africa has stayed least affected. Majorly occurring landslides with respect to size are the medium and small sized ones while considering the type of it, the three most common ones are basic landslides, mudslides and rockfall slides. The injuries caused in such incidents rarely lead to fatalities. The two are almost independent of each other. Besides, this it is seen that majority of the landslide events have been rainfall triggered (be it due to downpour, continuous rain, rain).

References

<https://data.nasa.gov/dataset/Global-Landslide-Catalog-Export/dd9e-wu2v/data>

www.nationalgeographic.org/encyclopedia/landslide/

http://support.sas.com/documentation/cdl/en/procstat/66703/HTML/default/viewer.htm#procstat_corr_details20.htm

Appendices

/* Read the dataset into SAS using code and prepare an Overview. */

DATA Global_Landslide;

```
INFILE "/home/s36314420/Assignment II/Global_Landslide.csv" dlm=',' DSD firstobs=2;
input id Day Month Year country$ nplaces$ htype$ ltype$ trigger$ storm_name$
fatalities injuries loc_desc$ loc_acc_km ls_size$ cat_src$ cat_id country_name$ near$
distance adminname1$ adminname2$ pop country_code$ continent_code$ version latitude
longitude;
run;
```

PROC CONTENTS DATA=Global_Landslide;

RUN;

/* Produce a table that shows the three oldest records in the dataset and include a suitable title.*/

PROC SORT DATA= Global_Landslide;

BY Year Month Day;

RUN;

PROC PRINT DATA = Global_Landslide (obs=3);

TITLE Three oldest Landslide Records;

RUN;

/* Produce a table that compares the number of landslides that have occurred in each continent for each landslide size category. Show just the frequencies and the overall percentage in your output. */

```
PROC FREQ DATA=Global_Landslide;  
TABLES continent_code*ls_size/ nocol nocum norow;  
RUN;
```

/* Test the correlation between the number of injuries and the number of fatalities caused by each landslide recorded. */

```
PROC SORT DATA=Global_Landslide;  
BY ltype;  
RUN;
```

```
PROC CORR data=Global_Landslide noprob nosimple;  
VAR injuries fatalities;  
RUN;
```

/* Produce a plot of the frequencies of the landslide types. Use a log scale on axis and label appropriately.*/

```
ODS GRAPHICS ON;  
PROC FREQ DATA=Global_Landslide;  
TABLE ltype;  
RUN;  
  
PROC SGPLOT DATA=Global_Landslide;  
VBAR ltype/ BASELINE=1;  
XAXIS LABEL="Type of Landslide";  
YAXIS LABEL="log of count" type=log logbase=10 logstyle=logexpand minor;  
Title Plot of the frequencies of the landslide types (Log Scale)  
RUN;
```

/* Produce a plot of the number of fatalities based on size and triggering source*/

```
proc sgplot data=Global_Landslide ;  
    title 'Comparison of fatalities caused by different landslides based on their size and triggering  
source';  
    vbar ls_size/response=fatalities group=trigger groupdisplay=cluster barwidth=1;  
    xaxis display=(nolabel);  
    yaxis grid;  
run;
```