ANALYTICAL MODELING

MBD12 Group Project, Group No: 7

Farah Anis Pankhania Blanche De Saint Victor

Milestone 1.a

DATA SET
Natural Disasters: [link]

SOURCES

- Our World in Data based on EM-DAT, CRED / UCLouvain, Brussels, Belgium – www.emdat.be (D. Guha-Sapir)
- Link: www.emdat.be

COLLECTION METHODOLOGY

- Data was collected scraping of Our World in Data
- The data collected is observational.
- EM-DAT publishes comprehensive, global data on each individual disaster event – estimating the number of deaths; people affected; and economic damages, from UN reports; government records; expert opinion; and additional sources.
- Our World in Data has calculated annual aggregates, and decadal averages, for each country based on this raw event-by-event dataset.
- Per capita rates have been calculated using population figures from Gapminder (gapminder.org) and the UN World Population Prospects (https://population.un.org/wpp/).
- Economic damages data is provided by EM-DAT in current US$. It is calculated as a share of gross domestic product (GDP) using the World Bank's GDP figures (also in current US$) (https://data.worldbank.org/indicator).

Definitions of specific metrics are as follows:

- 'All disasters' includes all geophysical, meteorological and climate events including earthquakes, volcanic activity, landslides, drought, wildfires, storms, and flooding.
- The total number of people affected is the sum of injured, requiring assistance and homeless.

TERMS OF USE

- Credit the original authors if used in research work.
- Credit UCLouvain university and D. Guha-Sapir if used in any research work.

SUMMARY

This dataset provides information about various types of natural disasters that occurred between 2012 and 2022 in the world, per country. It contains 1,401 rows and 50 columns and more precisely, provides information on the number of deaths and the total number of people affected by each natural disaster (the latter ranges from economic impact to death, but we will not analyze its detailed types of impact). The data set also includes death rates and the number of people affected per 100,000 population for each type of disaster. There are 48 variables in the data set, each covering a specific natural disaster, which are drought, earthquakes, disasters, volcanic activity, floods, mass movements, storms, landslides, fog, wildfires, extreme temperatures, and glacial lake outbursts.

While data on the number of individuals affected can offer insights into the scope and magnitude of a disaster, data on the number of deaths can be used to gauge the severity of a disaster. A technique to assess the intensity of various disasters and their effects on the population is through death rates.
The data set also includes information on the percentage of the population that was affected by each disaster's total number of victims. In order to manage disasters and aid in recovery efforts, this information can be utilized to determine how much a particular disaster impacts a nation's population.

Milestone 1.b

COLUMN STATISTICS

Since the dataset is really huge and inorder to process it efficiently, we have selected the rows for years>=2012 and dropped columns which do not have ‘death’ or ‘total’ and ‘people’ - the latter ensuring that we keep all total numbers of people affected. This way we aim to analyze the trend of natural disasters for the last decade across different countries. In the below statistics, the index lists down all the columns that we have filtered from the original dataset and the statistical five number summary for every column.

index	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Country.name	1401	211	Afghanistan	11	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Year	1401.0	NaN	NaN	NaN	2017.0935046395432	3.146284758275549	2012.0	2014.0	2017.0	2020.0	2022.0
Number.of.deaths.from.drought	277.0	NaN	NaN	NaN	37.978339350180505	294.530551736026	0.0	0.0	0.0	0.0	2465.0
Death.rates.from.drought	277.0	NaN	NaN	NaN	0.022349931185523468	0.3145563773568665	0.0	0.0	0.0	0.0	5.2169766
Number.of.deaths.from.earthquakes	263.0	NaN	NaN	NaN	366.8212927756654	1322.3227403980825	0.0	1.0	8.0	116.0	9550.0
Death.rates.from.earthquakes	263.0	NaN	NaN	NaN	0.298651269675205	2.4530896666696225	0.0	0.000300465475	0.0055552465	0.027152290500000002	32.484222
Number.of.deaths.from.disasters	1401.0	NaN	NaN	NaN	416.39400428265526	1766.315060981404	0.0	1.0	12.0	81.0	22874.0
Death.rates.from.disasters	1401.0	NaN	NaN	NaN	0.584124367186296	4.2082173021454174	0.0	0.0067144684	0.05969048	0.18917824	90.88069
Number.of.deaths.from.volcanic.activity	103.0	NaN	NaN	NaN	43.85436893203884	138.00315482092873	0.0	0.0	0.0	5.0	914.0
Death.rates.from.volcanic.activity	103.0	NaN	NaN	NaN	0.07196049763566019	0.45571779033723064	0.0	0.0	0.0	0.0010913199000000002	3.7429702
Number.of.deaths.from.floods	987.0	NaN	NaN	NaN	218.87639311043566	835.5920681656197	0.0	2.0	13.0	60.0	9819.0
Death.rates.from.floods	987.0	NaN	NaN	NaN	0.13334190466994933	0.5072268849227624	0.0	0.0051820839	0.031001806	0.104117793	11.166011
Number.of.deaths.from.mass.movements	20.0	NaN	NaN	NaN	20.0	13.718984157574841	8.0	13.0	16.0	17.0	46.0
Death.rates.from.mass.movements	20.0	NaN	NaN	NaN	0.007312009652749999	0.021753355870627378	0.000105265615	0.000350748805	0.0005832641249999999	0.001416103925	0.09465565
Number.of.deaths.from.storms	646.0	NaN	NaN	NaN	174.0123839009288	706.6225156979945	0.0	1.0	8.0	59.0	8598.0
Death.rates.from.storms	646.0	NaN	NaN	NaN	0.7555488149399845	5.893649510744648	0.0	0.002466181425	0.018261702	0.057423084	90.88069
Number.of.deaths.from.landslides	223.0	NaN	NaN	NaN	132.06278026290583	258.93514719963827	0.0	16.0	40.0	128.0	2312.0
Death.rates.from.landslides	223.0	NaN	NaN	NaN	0.14926110600372197	0.978778279158157 1	0.0	0.004666659849999999	0.016809931	0.062572295	14.353499
Number.of.deaths.from.fog	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Death.rates.from.fog	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Number.of.deaths.from.wildfires	191.0	NaN	NaN	NaN	21.79581151832461	37.532374561956836	0.0	1.0	7.0	21.5	221.0
Death.rates.from.wildfires	191.0	NaN	NaN	NaN	0.023112877747031 41	0.1110052352648810 3	0.0	8.6213155e-05	0.0012142739	0.00711087695	1.0574795
Number.of.deaths.from.extreme.temperatures	240.0	NaN	NaN	NaN	451.42083333333335	1114.4257851137952	0.0	3.0	31.0	300.25	7425.0
Death.rates.from.extreme.temperatures	240.0	NaN	NaN	NaN	0.2898988750167917	1.0779449798908252	0.0	0.003137204200000003	0.021299659	0.13754005	12.627883
Number.of.deaths.from.glacial.lake.outbursts	9.0	NaN	NaN	NaN	117.77777777777777	125.4370138533457	12.0	12.0	12.0	250.0	250.0

index	count	unique	top	freq	mean	std	min	25%	50%	75%	max
Death.rates.from.glacial.lake.outbursts	9.0	NaN	NaN	NaN	0.006592059293333333	0.00741930687706597	0.00015046824	0.001611154	0.003160838	0.0073291487	0.020326074
Number.of.total.people.affected.by.drought	277.0	NaN	NaN	NaN	10243739.610108303	42420397.97926591	0.0	5800.0	823032.0	5864549.0	382985115.0
Total.number.of.people.affected.by.drought.per.100.000	277.0	NaN	NaN	NaN	7169.99953306101	14609.258527042964	0.0	1.1934079	719.0034	7546.37	93599.93
Number.of.total.people.affected.by.earthquakes	263.0	NaN	NaN	NaN	479158.40684410644	1190723.2259974252	0.0	2321.5	25610.0	280269.5	7228686.0
Total.number.of.people.affected.by.earthquakes.per.100.000	263.0	NaN	NaN	NaN	330.15473513911024	1601.6826971798791	0.0	2.2070123	16.34609	61.818833500000004	20434.926
Number.of.total.people.affected.by.disasters	1401.0	NaN	NaN	NaN	4588002.60385439	25200600.963650394	0.0	2282.0	30000.0	352498.0	430959537.0
Total.number.of.people.affected.by.disasters.per.100.000	1401.0	NaN	NaN	NaN	3404.8080962426825	10585.3551071558	0.0	13.17171	182.7388	1467.0697	101378.83
Number.of.total.people.affected.by.volcanic.activity	103.0	NaN	NaN	NaN	183071.84466019418	393683.57821388665	0.0	6888.0	18685.0	133327.5	1915398.0
Total.number.of.people.affected.by.volcanic.activity.per.100.000	103.0	NaN	NaN	NaN	1190.0257462303775	7918.996625288444	0.0	0.747842405	6.239709	44.064519000000004	78767.06
Number.of.total.people.affected.by.floods	987.0	NaN	NaN	NaN	1978463.3353596758	7997164.513687832	0.0	3465.5	26411.0	200367.0	79219326.0
Total.number.of.people.affected.by.floods.per.100.000	987.0	NaN	NaN	NaN	901.4217230652853	3974.9141156611176	0.0	10.724655	114.78683	490.43939	98453.734
Number.of.total.people.affected.by.mass.movements	20.0	NaN	NaN	NaN	40.4	81.87692946702	0.0	0.0	0.0	2.0	200.0
Total.number.of.people.affected.by.mass.movements.per.100.000	20.0	NaN	NaN	NaN	0.12063864109514999	0.5286395895616843	0.0	0.0	0.0	9.82322175e-05	2.3663912
Number.of.total.people.affected.by.storms	646.0	NaN	NaN	NaN	2185756.7445820435	8812146.058199825	0.0	251.25	10012.5	184118.5	94036842.0
Total.number.of.people.affected.by.storms.per.100.000	646.0	NaN	NaN	NaN	2462.133580683743	9996.030666534854	0.0	0.5697520225	31.850935	439.78716	101378.83
Number.of.total.people.affected.by.landslides	223.0	NaN	NaN	NaN	18909.905829596413	48147.520291223664	0.0	13.0	348.0	7173.5	269473.0
Total.number.of.people.affected.by.landslides.per.100.000	223.0	NaN	NaN	NaN	8.393820453916367	34.669908694187626	0.0	0.007649702999999999	0.1427439	1.86040345	304.90295
Number.of.total.people.affected.by.fog	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Total.number.of.people.affected.by.fog.per.100.000	0.0	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Number.of.total.people.affected.by.wildfires	191.0	NaN	NaN	NaN	260738.69109947645	1437040.158371979	0.0	540.0	4004.0	50503.0	10056764.0
Total.number.of.people.affected.by.wildfires.per.100.000	191.0	NaN	NaN	NaN	96.19775992849738	482.2102681996453	0.0	0.27061913000000004	2.4006104	13.1457605	4722.8535
Number.of.total.people.affected.by.extreme.temperatures	240.0	NaN	NaN	NaN	110998.5125	313994.7773399978	0.0	0.0	20.0	49000.0	1827759.0
Total.number.of.people.affected.by.extreme.temperatures.per.100.000	240.0	NaN	NaN	NaN	308.97240264560713	2272.684770043297	0.0	0.0	0.00315020685	16.789657	32549.053
Number.of.total.people.affected.by.glacial.lake.outbursts	9.0	NaN	NaN	NaN	10.666666666666666	12.649110640673518	0.0	0.0	0.0	24.0	24.0
Total.number.of.people.affected.by.glacial.lake.outbursts.per.100.000	9.0	NaN	NaN	NaN	0.0003581640044444444	0.0005702370294476076	0.0	0.0	0.0	0.0005113638	0.0017050735

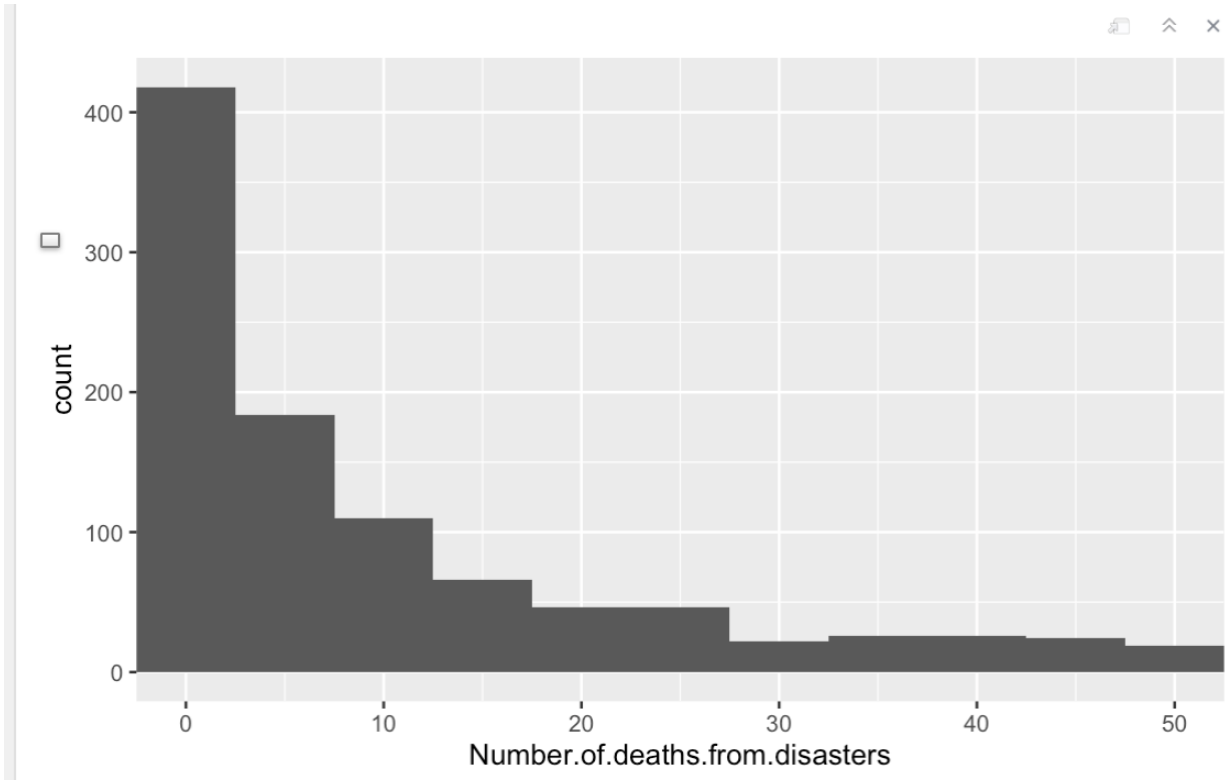
Statistical Insights:

- Droughts and Storms are the two main natural disasters which have affected the maximum number of people.
- The death rate due to storms and droughts stands at 90%
- The data shows resulting rows starting from the year 2012 until the year 2020.
- The total number of deaths due to natural disasters is approximately 22000 in the last decade.
- There are 211 unique countries where natural disasters have occurred in the past decade.

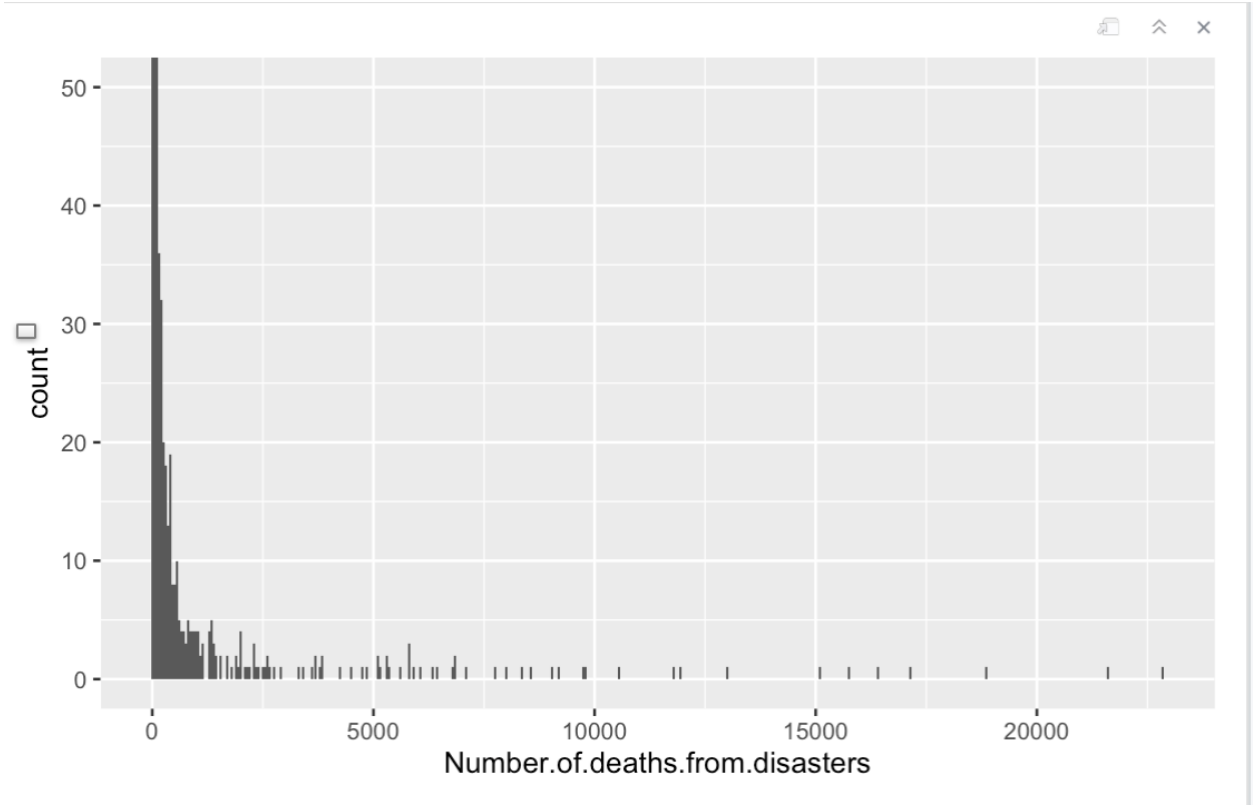
Milestone 2.a

The sequence of questions and findings. Only include the questions and findings, not how you answered the question; for example, do not describe the code that you wrote to answer the question.

To ensure a thorough understanding of our dataset, we started by asking ourselves general questions. We start by visualizing the distribution of Number of deaths from disasters. The count for the number of deaths due to disasters is the highest when number of deaths ranges between 0 to 10 for each country.



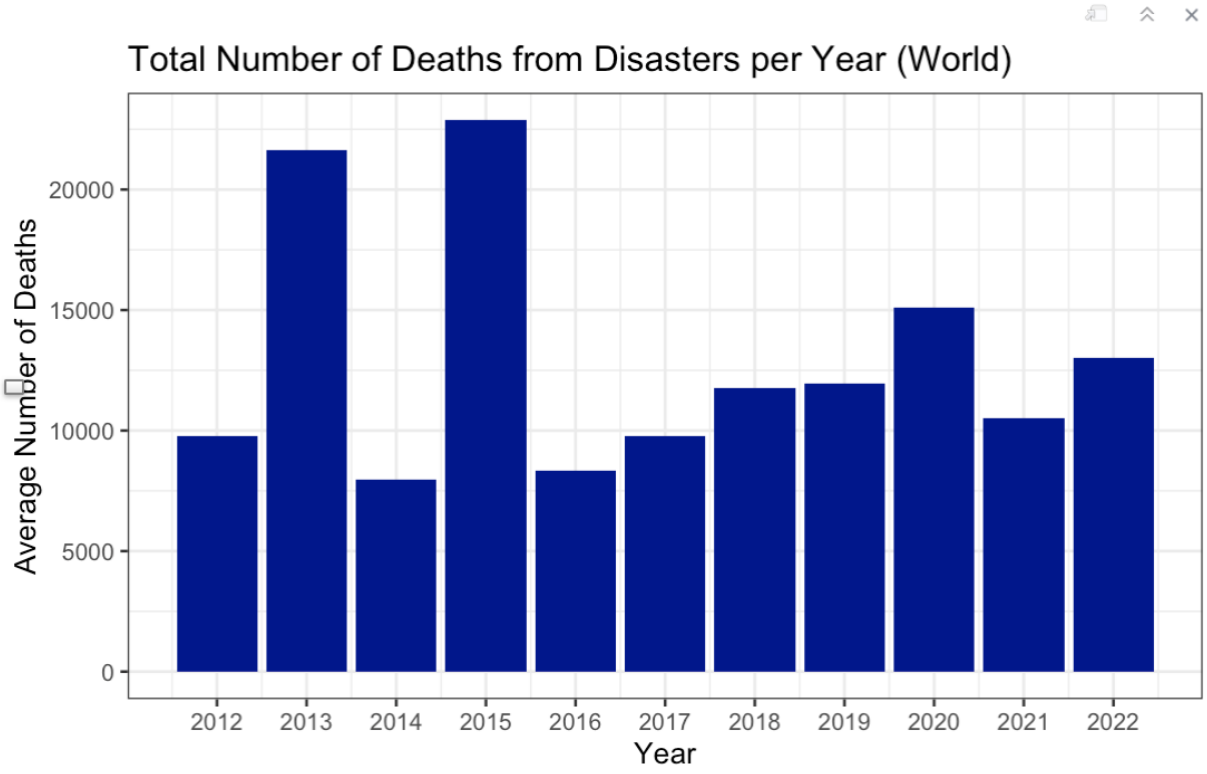
Observing the outliers where the number of deaths from disasters is very high:



After Analyzing the rows where number.of.deaths.from.disasters is greater than 5000. We could see Nepal in 2015 had the highest number of deaths. Thus, we decided to start our analysis by finding the total number of deaths from disasters per year worldwide.

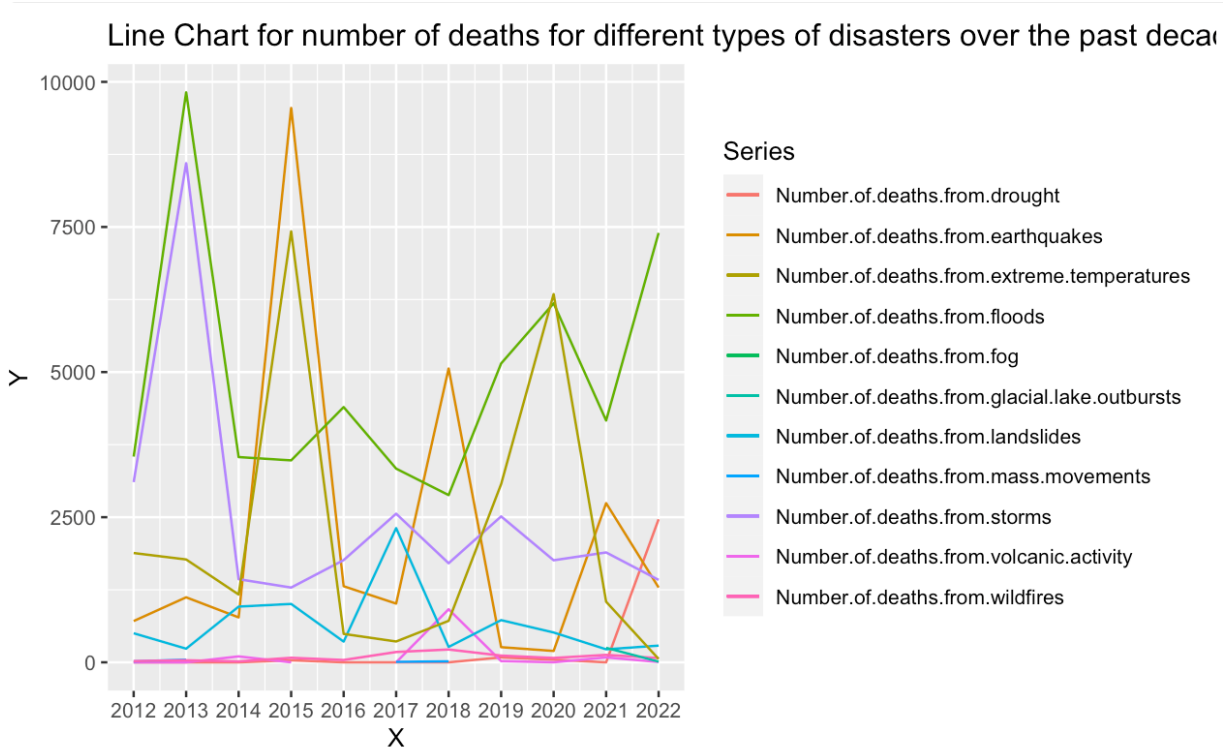
Question 1: What is the total number of deaths from disasters per year worldwide?

Highest number of deaths due to natural disasters occurred in the years 2013 and 2015 in the past decade.



Question 2: What types of natural disasters had the highest number of deaths in the past decade?

We can see from the below line chart that the global number of deaths in 2013 were the highest due to floods and in 2015 due to earthquakes.



Question 3: Which country was most affected by the floods in 2013?

In 2013, the country that was most affected by floods was India with number of deaths = 6453

Question 4: Which country was most affected by the storms in 2013?

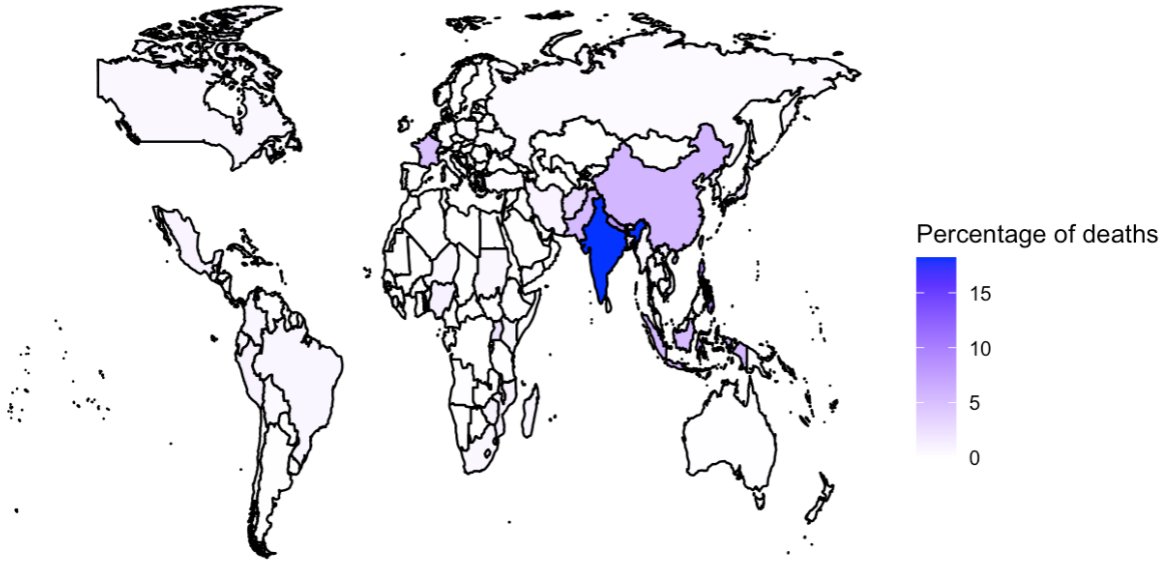
In 2013, the country that was most affected by storms was the Philippines with number of deaths = 3542

Question 5: Which country was most affected by the earthquakes in 2015?

In 2015, the country that was most affected by earthquakes was Nepal with number of deaths = 8969

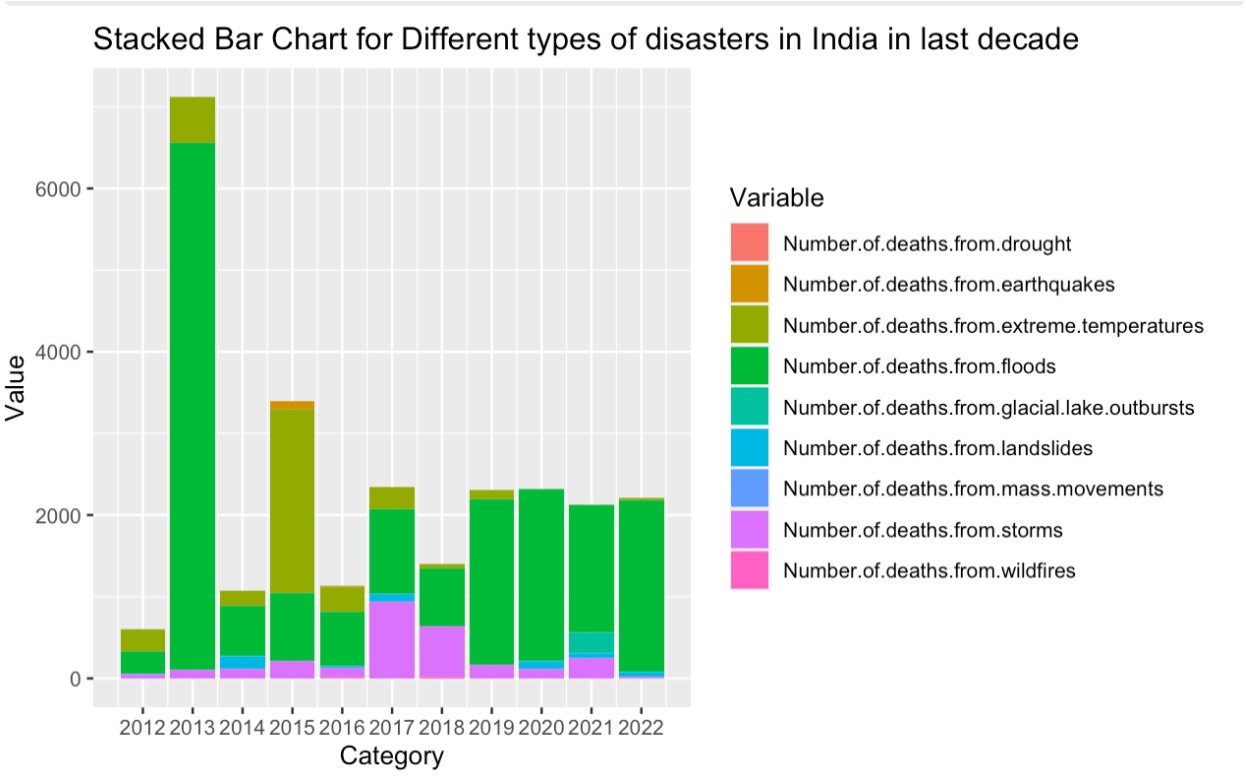
Question 6: Which country has the highest proportion of deaths due to natural disasters as compared to the total number of global deaths due to disasters?

From the below figure, India and China have the highest proportion of deaths due to natural disasters in the past decade.



Question 7: Which type of disaster is the most fatal in India?

The most fatal disaster in India has been floods in the last decade.

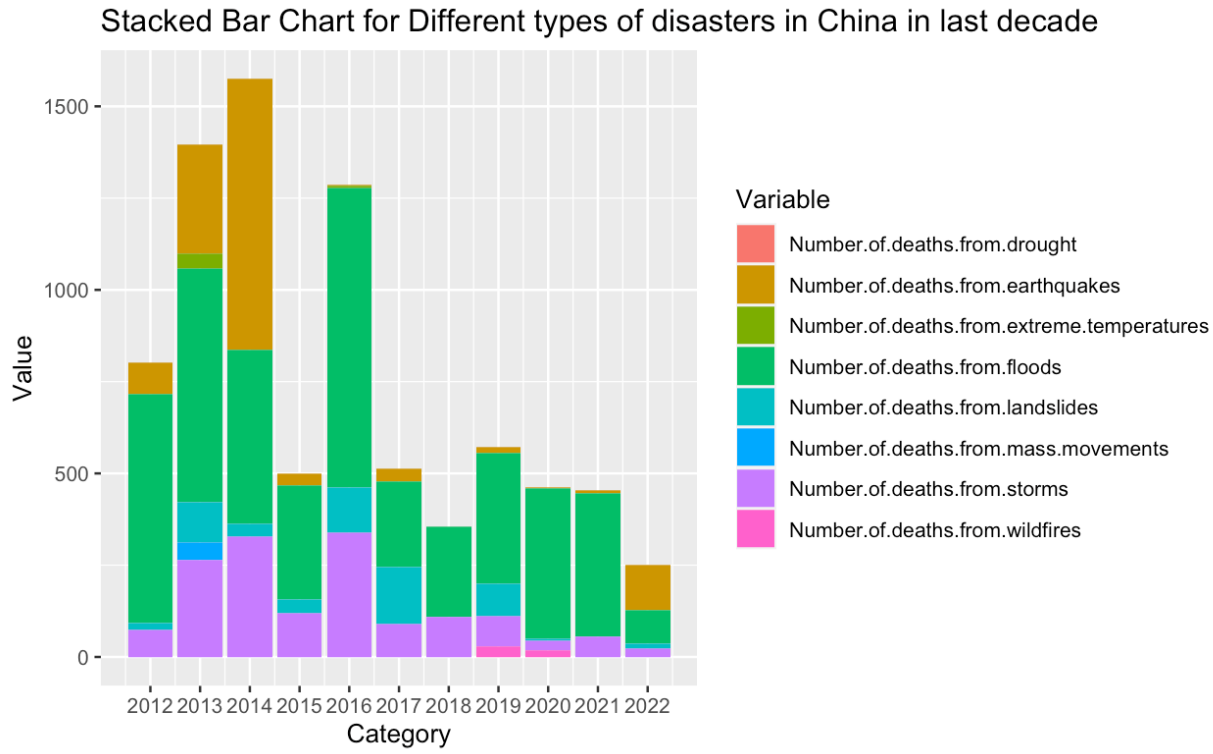


Question 8: What was the total number of deaths in India due to floods?

There were 18397 deaths due to earthquakes in India in the last decade.

Question 9: Which type of disaster is the most fatal in China?

The most fatal disaster in China has also been due to floods in the last decade.

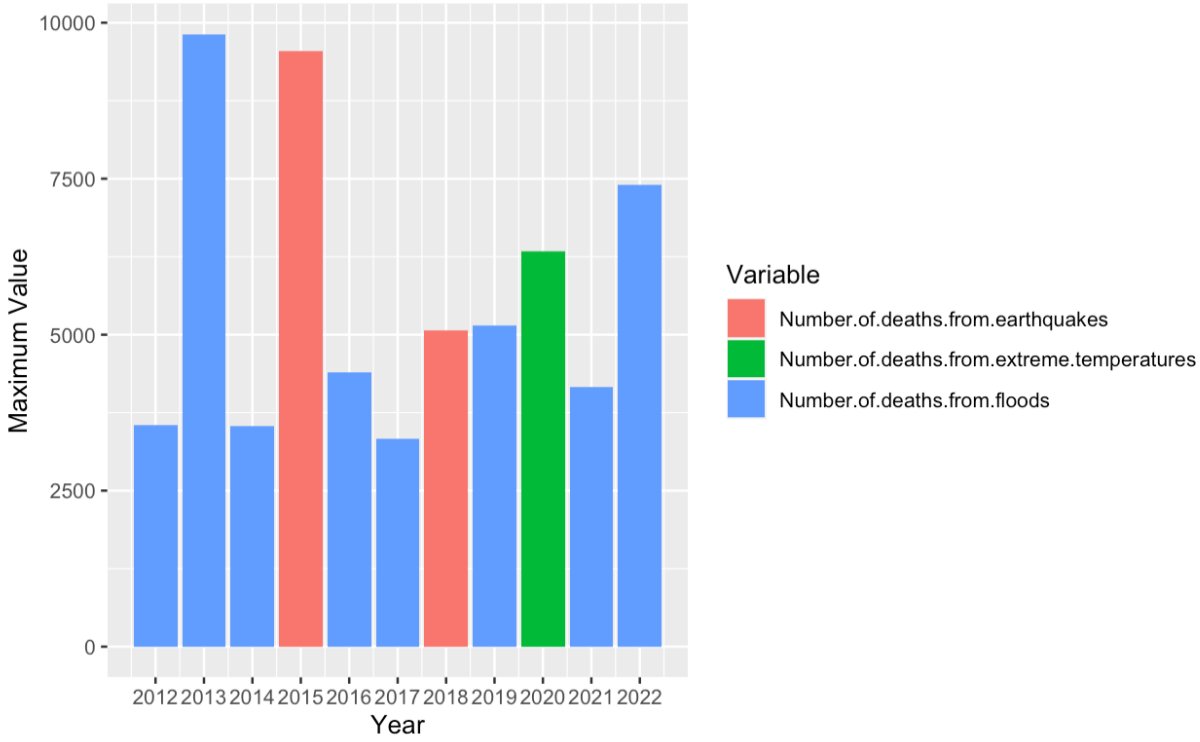


Question 10: What was the total number of deaths in China due to floods?

There were 4580 deaths due to floods in China in the last decade.

Question 11: Which were the most fatal disasters each year in the last decade?

Most fatal deaths in the last decade were due to earthquakes, extreme temperatures and floods.

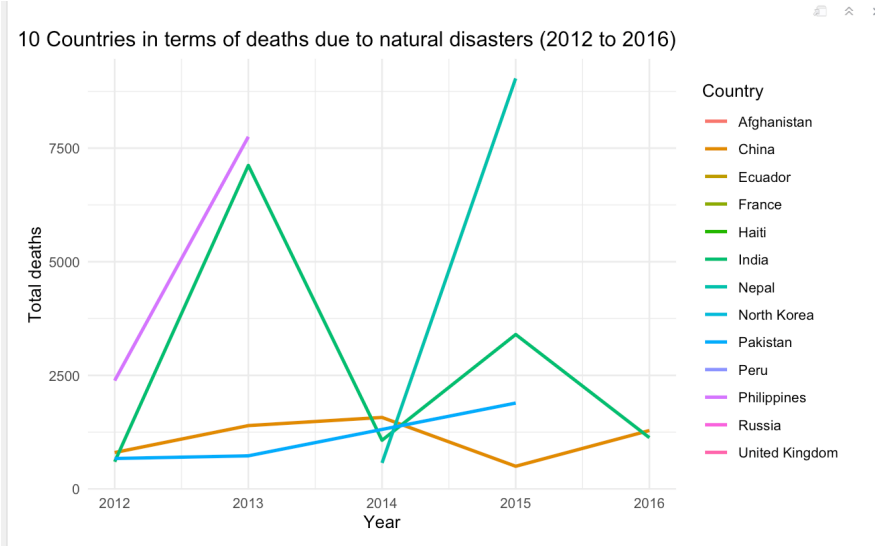


Question 12: Which was the most fatal disaster in 2022?

The most fatal disaster in the year 2022 was due to floods.

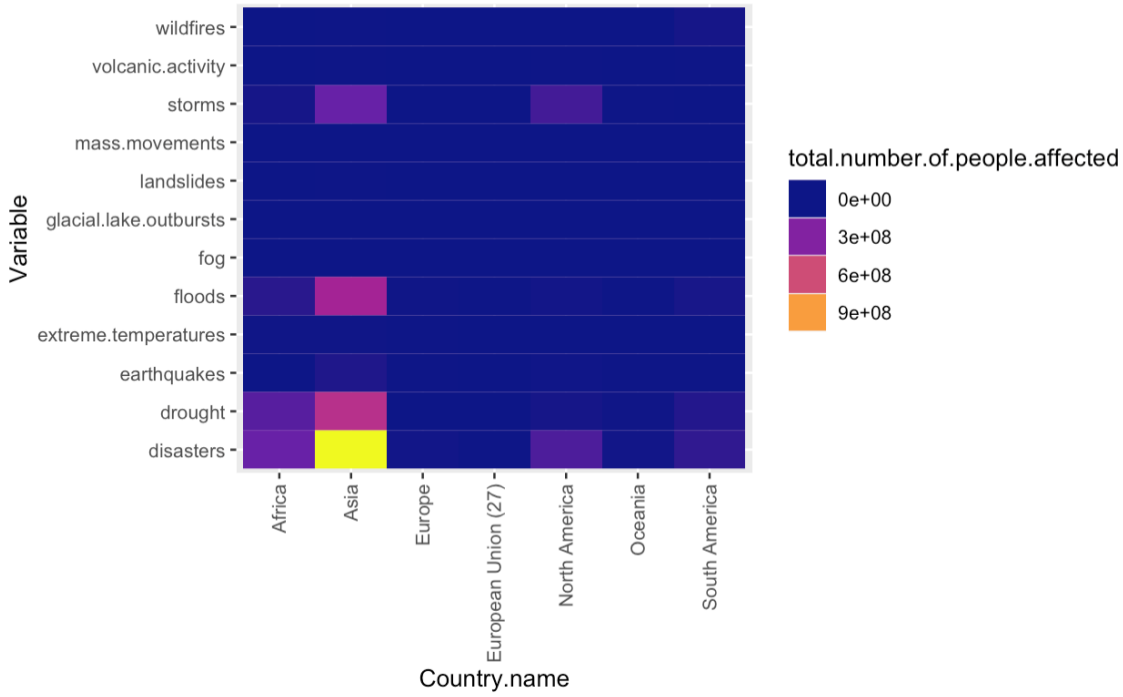
Question 13: What are the top 10 countries in terms of death due to natural disasters from 2012 to 2016?

We can see that the top county in terms of death due to natural disasters in Nepal and India is closely followed by the Philippines.



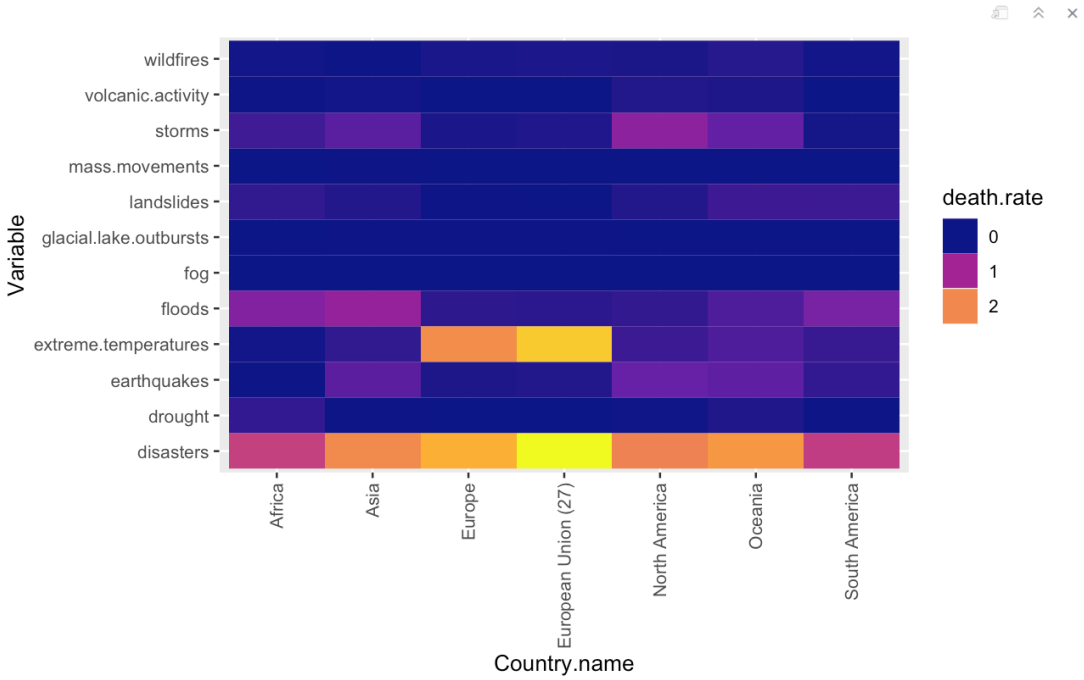
Question 14: What are the total number of people affected due to natural disasters region wise in the last decade?

We can see that Asia has been the most affected by natural disasters mainly due to droughts, floods and storms.



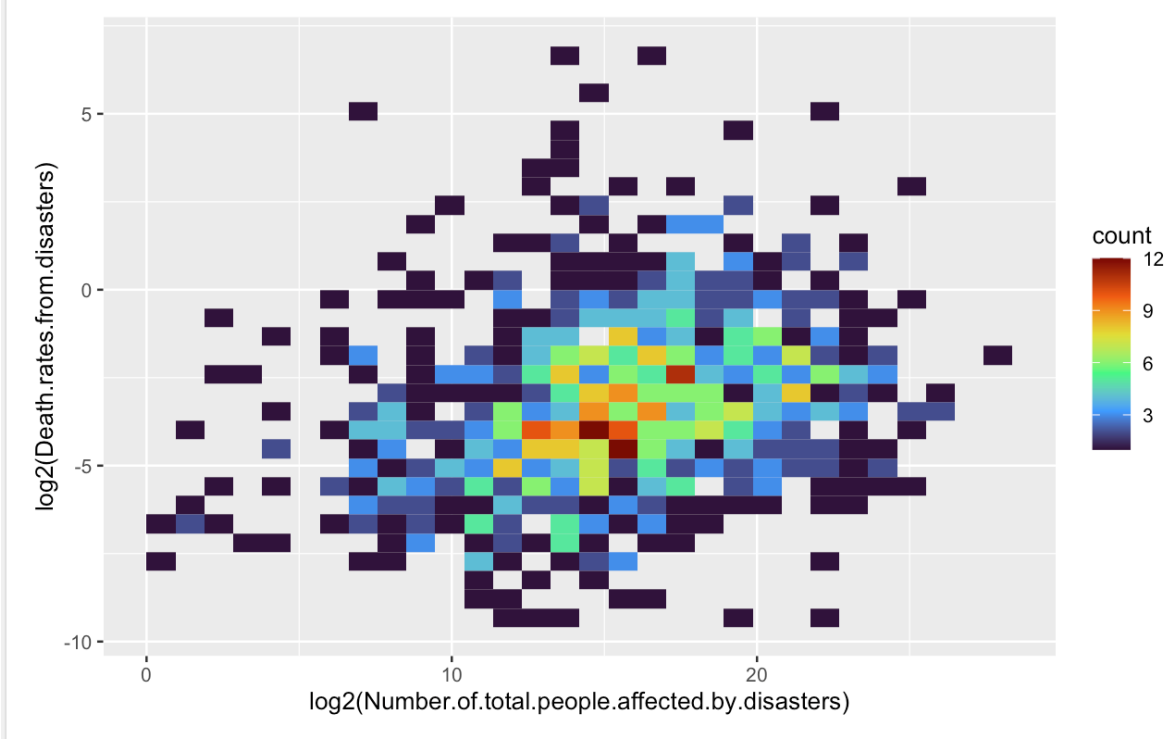
Question 15: What is the death rate from natural disasters region wise in the last decade?

We see that the death rate is higher in Europe and the European union due to extreme temperatures, followed by floods in other regions.



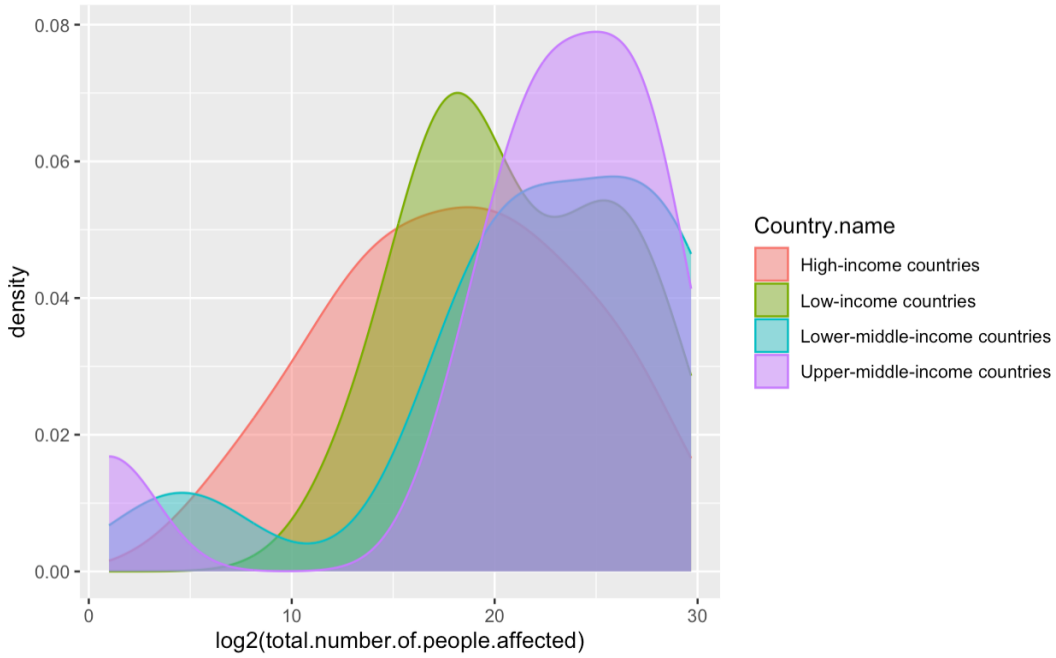
Question 16: What is the correlation between death rate and total number of people affected country wise?

We can see that as the total number of people affected by disasters increases, the death rate also increases.



Question 17: How does the economic condition of a country impact the total number of people affected due to natural disasters?

The low-income countries have a comparatively higher number of people affected due to natural disasters.



Milestone 2.b

Table 1: Top 10 countries affected by floods in 2013 which contributed to the highest number of deaths from disasters in the last decade.

Top 10 countries affected by floods in 2013 which contributed to the highest number of deaths from disasters in the last decade.

Country.name	Year	Number.of.deaths	Number.of.people.affected
India	2013	6453	3419473
China	2013	637	7684030
Pakistan	2013	268	1497782
Cambodia	2013	200	1500000
Nepal	2013	195	16823
Vietnam	2013	141	2161001
Kenya	2013	128	115800
Zimbabwe	2013	125	9700
Brazil	2013	119	247410
Mozambique	2013	119	240000

Table 2: Top 10 countries which have the highest proportion of deaths from natural disasters in the past decade.

Top 10 countries which have the highest proportion of deaths from natural disasters in the past decade.

Country.name	Number.of.deaths.from.disasters.in.last.decade	death.proportion
India	26020	18.22
Philippines	12736	8.92
Nepal	11391	7.98
China	8166	5.72
Pakistan	7621	5.34
Indonesia	7517	5.27
France	6770	4.74
Afghanistan	4258	2.98
United Kingdom	3521	2.47
Haiti	3454	2.42

Milestone 2.c

Figure 1: Top 3 countries per Year in terms of death due to Natural Disaster

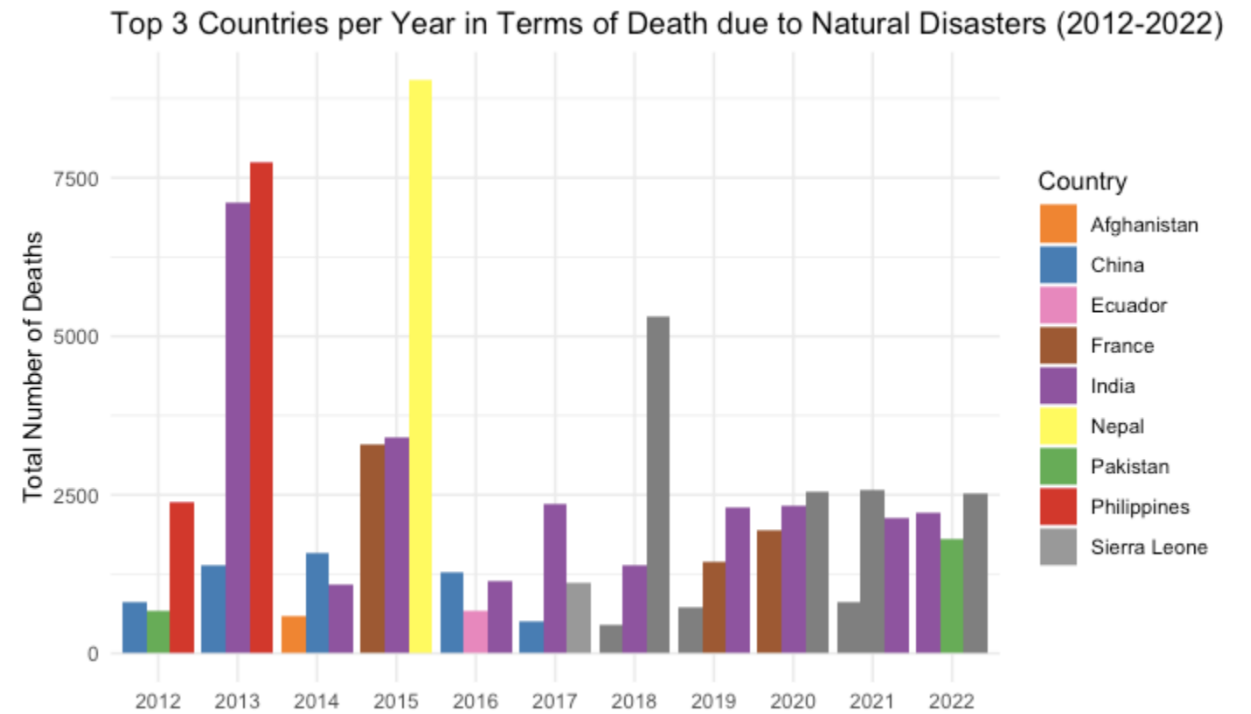


Figure one shows that Nepal most likely went through a major natural disaster in 2015. Additionally, we see that India and Sierra Leone consistently have a high number of deaths due to natural disasters.

Figure 2: Number of deaths per type of disaster between 2012 and 2022

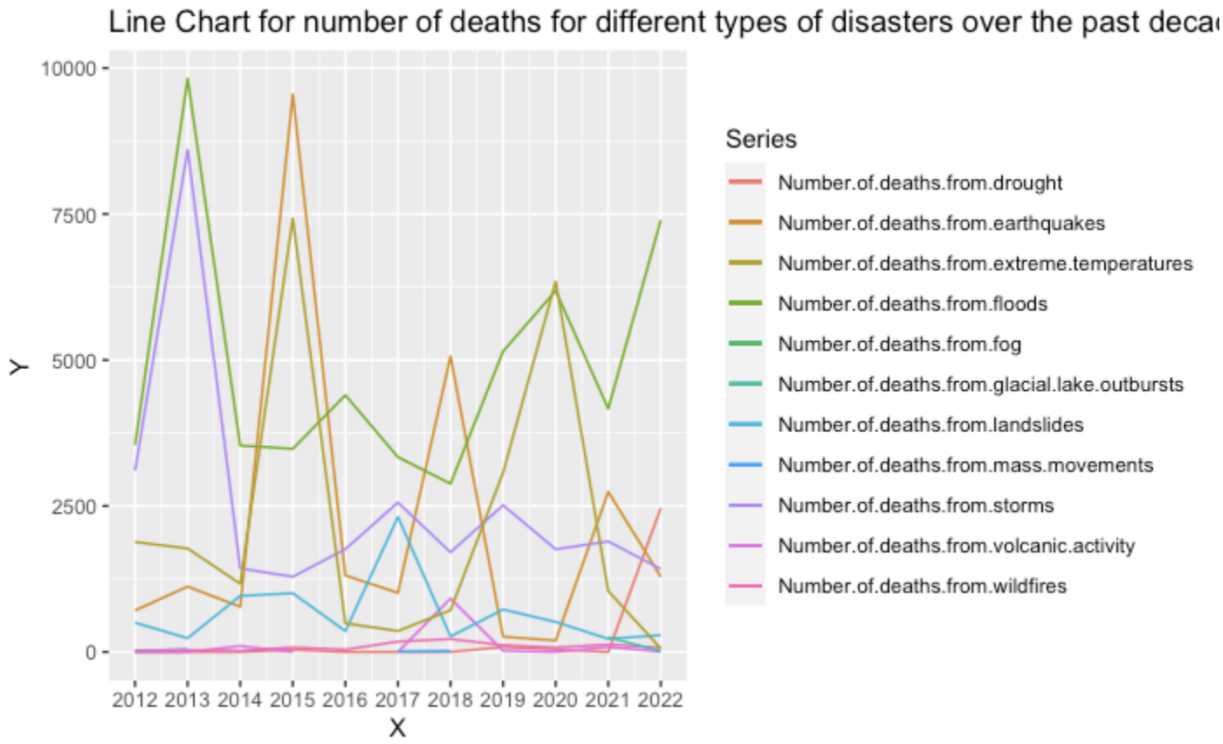


Figure 2 shows that the number of deaths from floods are consistently high, and getting higher over time, and should most likely increase due to climate change. Additionally, the variation in deaths from earthquakes from a year to another is very high, which makes sense since earthquakes do not happen on a recurring basis, but cause enormous death tolls. Finally, deaths from storms remain quite high over time.

Figure 3: Correlation between countries' income and number of people affected by natural disasters

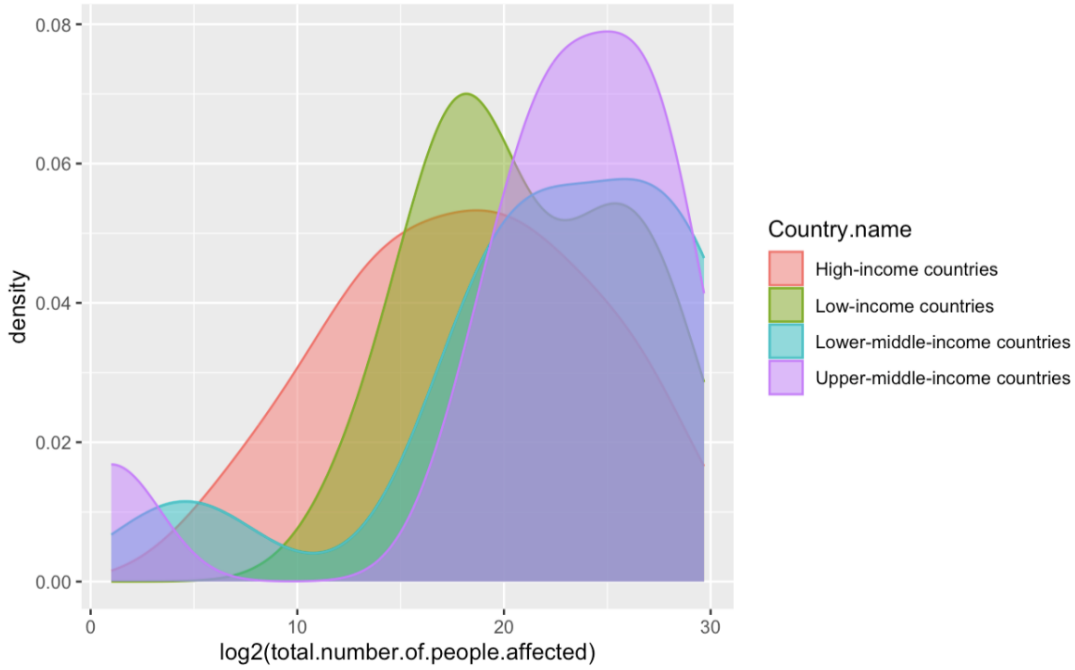


Figure 3 shows that lower income countries are more impacted by natural disasters.