VRIJE
UNIVERSITEIT
BRUSSEL

LANDSLIDE VISUALISATION

Information Visualisation project

Chapelle Alexendre
Dardenne Nicolas
Swirydowicz Szymon

March → May 2023

Sciences and Bio-Engineering Sciences

Contents

1 Introduction

Landslide visualisation is a project realised by a group of three students in the context of the `Information Visualisation - 008866` course at `The Vrije Universiteit Brussel` (`VUB`). The course is organised by Prof. Dr. Beat Signer and assisted by Yoshi Malaise.

The goal of the project is "to process and visualise a large dataset using the theoretical knowledge obtained during the course". As detailed further in the report, our group chose the theme of landslides, and young students as the target audience.

1.1 Source code

All resources used for the project including the visualisation source code are available on GitHub via `https://github.com/sswirydo/Visualisation-Landslides`.[1]

1.2 Visualisation setup

In order to run the project code from our GitHub repository, Python3 and its corresponding Pip manager are required.

[1] Please contact us if the link is not available anymore.

Packages and requirements The visualisation requires a series of python packages that are listed in the `requirements.txt` file, including packages such as dash and pandas. You can install them manually or simply run the following bash command:

```
pip3 install -r requirements.txt
```

that will install automatically all the necessary requirements.

Execution Once all installed, the visualisation can be launched using by simply running the main python file with:

```
python3 main.py
```

which will open the visualisation server/website, accessible by default through the 8050 port. It can thus be simply accessed via `http://127.0.0.1:8050`.

2 Targets

2.1 Target users

The target users are school students. We aim to provide a didactic, fun and youth-oriented visualisation that will engage and educate students.

2.2 Target task

We provide an interactive 2D Earth map that allows students to engage in an enjoyable exploration of landslides. The map serves as a sandbox that students can use to navigate and discover information about landslides and their impact in different regions. By interacting with the map and its various features, students can develop a deeper understanding of landslides, their causes, and their consequences.

In addition, the interactive map might serve as a valuable resource for teachers. They can assign specific tasks or projects to students using the dashboard as reference. For instance, teachers can ask to identify regions with a high incidence of landslides and analyse the possible contributing factors. They can also use it to discuss various topics as environmental impacts, preventive measures or case studies of notable landslides.

3 Dataset

Global Landslide Catalog Export [1]. Over 11k rows and 31 columns. Includes event descriptions, counts of fatalities and injuries, links to sources, photos, geolocalisation, etc.

sourc…	sourc…	event…	event…	event…	event…	event…	locati…	locati…	lands…	lands…	lands…	land
3 News	http://ww…	6,034	2014 May…		Codfish I…	A landslid…	Codfish I…	5km	landslide	downpour	small	natu
Otago Dai…	http://ww…	1,910	2010 May…		Berwick …	Elsewher…	Berwick …	10km	landslide	downpour	medium	unkn
Stuff.co	http://ww…	6,997	2015 Jun …		Highcliff …	Police est…	Highcliff …	5km	landslide	downpour	medium	belo
www.radi…	http://ww…	4,941	2013 Jun …		Portobell…	Two days…	Portobell…	5km	landslide	downpour	medium	unkn
3 News	http://ww…	6,008	2014 May…		Queen Str…	Five stud…	Queen Str…	1km	landslide	rain	small	unkn
momento…	http://mo…	1,510	2010 Feb …		Comodor…	A mudsli…	Comodor…	10km	mudslide	downpour	medium	unkn
odt	http://ww…	1,911	2010 May…		Portobell…	Flood con…	Portobell…	5km	landslide	downpour	medium	unkn
www.radi…	http://ww…	4,942	2013 Jun …		Harringto…	Two days…	Harringto…	5km	landslide	downpour	medium	unkn
tvnz.co.nz	http://tvn…	4,962	2013 Jun …		Upper Ju…	Two land…	Upper Ju…	exact	landslide	downpour	medium	unkn
www.3ne…	http://ww…	4,969	2013 Jun …		Blanket B…	Rain is sti…	Blanket B…	1km	mudslide	downpour	medium	unkn
livenews.…	http://live…	5,045	2013 Jul …		Sh8 Betw…	An active …	Sh8 Betw…	25km	landslide	downpour	small	unkn
Stuff	http://ww…	1,881	2010 May…		the narro…	A landslid…	the narro…	5km	rock_fall	continuo…	medium	abov
en.wikipe…	http://en.…	1,587	2010 Mar…		Fergus La…	More tha…	Fergus La…	5km	landslide	downpour	medium	unkn

Figure 1: Landslide dataset preview [1]

3.1 Data processing

In order to have the most relevant data, we cleaned the basic data first. We removed non-essential columns, replaced some missing data (e.g., na is now 0 for numeric values) and renamed/formatted several columns to ensure that names do not contain underscores or other unwanted characters while they are presented to the users.

3.2 Temperature dataset

During the project we have also attempted to merge the landslide dataset with a temperature dataset in order to provide some additional insights as temperatures are often an important factor of landslides. However, we did not manage to successfully perform that task.

Indeed, our main difficulty concerned the size and related time complexity of the merge. On one side, we have a huge dataset of landslides, each described by a latitude and longitude value. On the other, a temperature data for each city (or major city) and is also labeled by latitude and longitude. The most accurate approach we found was to loop through each landslide and find the nearest city to obtain the temperature. The closest city is determined by latitude and longitude, as well as a date, since taking temperature from a different season or year would be irrelevant. As evoked before, the required time is too long even for a “do one-time and save” preprocessing. Some of our attempts included:

- Use a shape file of countries to determine to which temperature data the landslide belongs. The drawback is that it would be less accurate. Two landslides could occur close to one another but they could be labeled as two different countries, and similarly, two far away landslides might be labeled as the same country (US, outer-sea France, etc.) and thus be tagged with the same temperature. Plus, it would still require to find the closest date.

- Use the country columns of each dataset. However, the string values for the same countries are not the same, so it would require important manual renaming or some sort of edit distance comparison. Moreover, it presents the same drawbacks as the previous point.

Eventually, we chose to ignore the idea of merging the temperature dataset and focus on the landslide dataset instead, especially as our visualisation space and performance were already highly limited.

4 Visualisation

4.1 Tools and frameworks

There are lots of different open frameworks that allow to visualise and present our data. For our visualisation tool we chose to work with the free and open source python version of Dash by Plotly which allows to build interactive dashboards. Additionally, we integrate Leaflet into our dashboard which provides simple and interactive maps.

Additionally, we have also considered the following frameworks:

- Panel with Bokeh: allows to create interactive dashboard really simply with Python. However, from our experience, a drawback of this framework regarding our needs, is that the integration of interactive maps is complicated and requires multiple workarounds. As maps are a central part of our visualisation, we decided it was not ideal for our project.

- D3.js: allows to create complex web visualisations. However, it provides more capabilities than we required in our project and does not allow to create dashbord as simply as other frameworks. Moreover, as data processing and manipulation is way simpler with Python than with Javascript, we preferred to realise our whole project using Python.

4.2 UI

Our main objective for designing the desktop user interface was to present all pertinent information inside a single view without the need of scrolling the whole page. This is why we have chosen to organise the dashboard into three columns.

Figure 2: Implemented dashboard

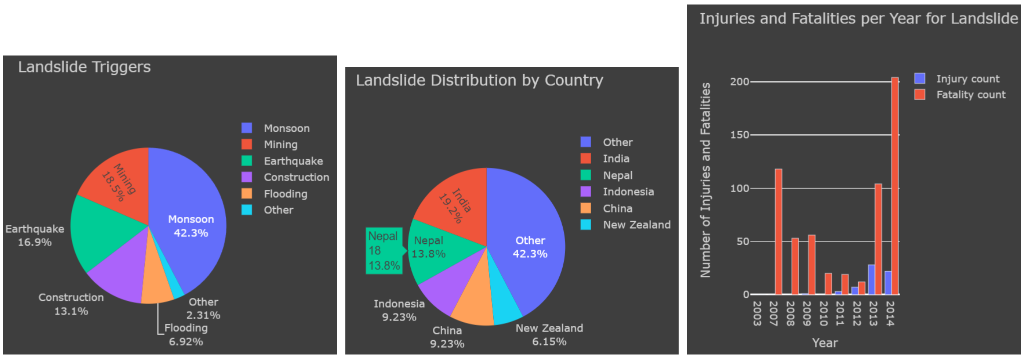

Figure 3: Graph box plots

4.2.1 Left column

Options The upper corner contains a set of options allowing to customise different elements of the visualisation. Those options might affect not only the markers placed on the map, but also the different graphs.

Those include:

- Date range slider. The slider allows to easily pick the start and end years of events through a range slider. The slider shows only the minimal and maximal years, and the selected ones as well. Other years are hidden to avoid the slider to be cluttered with too much text.
- Landslide triggers. Landslides can happen due to different circumstances or triggers. Users can choose multiple triggers (up to 18) through a drop down menu listing all triggers such as rain, downpour or earthquake. If no trigger is selected by the user, all triggers are shown by default. Proposed triggers are sorted alphabetically.
- Landslide sizes Same visualisation implementation as for landslide triggers but for the landslide sizes (up to 6). We preferred a dropdown instead of a slider in order to select multiple sizes at once and because in some cases the size is simply unknown. Proposed sizes are sorted by their semantic size from smallest to biggest.

Details The details box on the left is updated as the user clicks on a landslide marker on the map.

- Title: The title's event.
- Landslide info: Contains information about the landslide trigger and size.
- Source link: A link to the source of the event. Some links might be outdated.
- Event date and injuries: The date and the number of injuries and deaths.
- Description text: Text explaining what happened.
- Photo: Photo of the event. The photo is directly fetched from the Internet. The URL is provided in the dataset. Note that if the event has no photo, we chose to put a small sad-Pokémon GIF instead. It might inform the user that other events might still contain photos, and additionally gives a more fun ambiance to the visualisation which suits well our target audience.

Social networks A fun interaction we have added is social network sharing. A small text space was added that is filled automatically with details about the selected landslide and that can be edited. Its content can then be shared using the "share on Twitter" button.

Dataset reference A small button was added that opens the dataset source.

4.2.2 Middle column

Title A simple title called "Landslides Explorer" that allows to instantly understand the purpose of the dashboard. Theme-related emojis were added around the title to add an easygoing ambiance.

Interactive Map This is the central element of our dashboard, not only in terms of positioning but also of importance. We chose to put the map in the center as it is the element that the user will the most interact with. It is populated by a series of markers, each representing an individual landslide on which the user can click to obtain further details. Additionally, note that when the map is zoomed out, the markers are clustered together to provide better readability and performance. Similarly, the user can click on a marker cluster to zoom into smaller clusters until it reaches an individual marker.

Category tabs Landslides belong to different categories which can be switched by clicking on 1 of the 12 tabs. The map pins are then updated accordingly to the category. We chose this approach over dropdowns as tabs are more visual and allow to choose faster a desired category. As comparison between different categories is not particularly pertinent, not being able to choose multiple categories at once is not troublesome.

Category explanation As explained above, landslides belong to different categories. However, it is not always obvious how each category different from each other and what are their distinct properties. For this purpose, we have included a description zone that explains the currently selected landslide category. That way, it makes the dashboard self-contained and allows users to easily learn about the category, especially as in most cases user do not actively research those by themselves as they are overwhelmed by all the different categories. Note that some explanations are hard-coded and others fetched live from Wikipedia deepening on the quality.

4.2.3 Right column

Graphs and figures As we present multiple graphs, those were placed inside a scrollable box which avoids to scroll the whole page. We have decided to include 3 graphs that we found pertinent, which are

- Landslide trigger pie chart. Among the selected landslide triggers, it gives the percentage of landslide triggers among available map pins. Pie chart slices are sorted from biggest to lowest percentage and the importance of percentages is represented visually by the angle of the slice. As the user can choose up to 18 triggers, only the top 5 is retained, and the rest is grouped by the "Other" label to ensure the figure remains readable.
- Injuries and Fatalities per year (by category) bar chart. Bar chart compares the injuries (blue) and fatalities (red) per year. The time range of the graph follows the range selected by the year range picker. We chose to limit the the time dimension per year (and for

instance not per month) to ensure the bar chart remains readable when user selects the full year range.

- Per country distribution pie chart. Same implementation as the landslide trigger pie chart above, but shows the top 5 of countries where landslide take place, which allows to understand which countries are the most impacted by landslides. We chose to do this per country and not per continent, as the map itself already provides an insight of the density of landslides per continent by pin grouping.

4.3 Mobile/Phone version

As our aim are target are young students, an important point in our dashboard implementation was to make the visualisation accessible using mobile devices. For this purpose, the above column organisation is replaced by a single and scrollable column when a small and narrow view-port is detected. A drawback of this approach is that all elements can not be viewed at all time, but we believe this remains more user friendly and readable than simply down-scaling the whole dashboard.

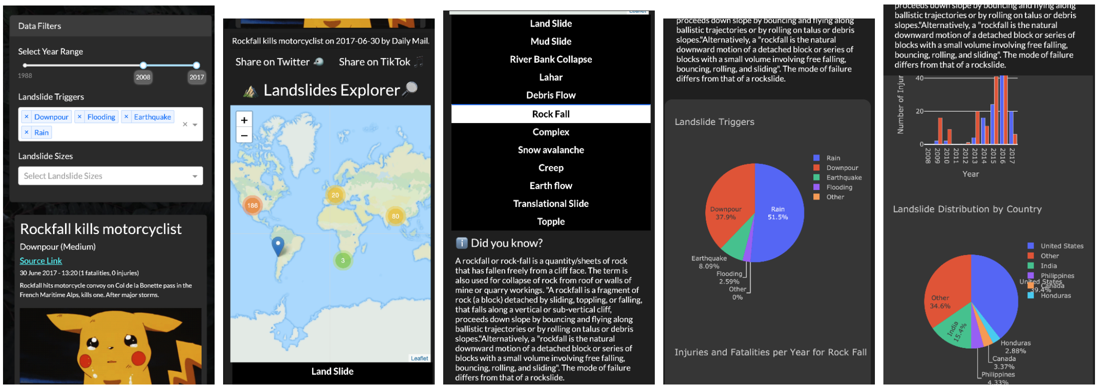

Figure 4: Mobile version - Phone

4.4 Details

4.4.1 Internet access required

Note that our visualisation requires internet access. Indeed, the photos illustrating the landslides are loaded through an image URL included in the data set. A solution would be downloading all the photos in advance, however this would require an important amount of storage. Additionally, it is still required for different data interactions such as social network sharing or fetching some landslide category details from Wikipedia.

4.5 Details versus Graphs positioning

Our main focus was to fit everything into one screen. As the `details` element contains all the different information about specific landslides, it is the main source of information of the visualisation. We could thus argue between exchanging `graphs` and `details` position to give more space for `details`. However, the reason we choose to put `graphs` on the right side is mostly because they take way more vertical space than `details`, and thus putting them inside a medium-size box on the left would require way more scrolling making it less ergonomic.

5 Evaluation

5.1 Similar visualisations

We continuously compared our implementation with existing one in order to improve our solution. A remarkable resource which highly influenced our work was the COVID dashboard [2] linked in project's instructions. The following characteristics particularly caught our attention:

- Strategic Placement of Selectors and Information (left)
- Central Placement of Interactive Map (center)
- Convenient Location of Plots (right)
- Utilisation of tabs (below)

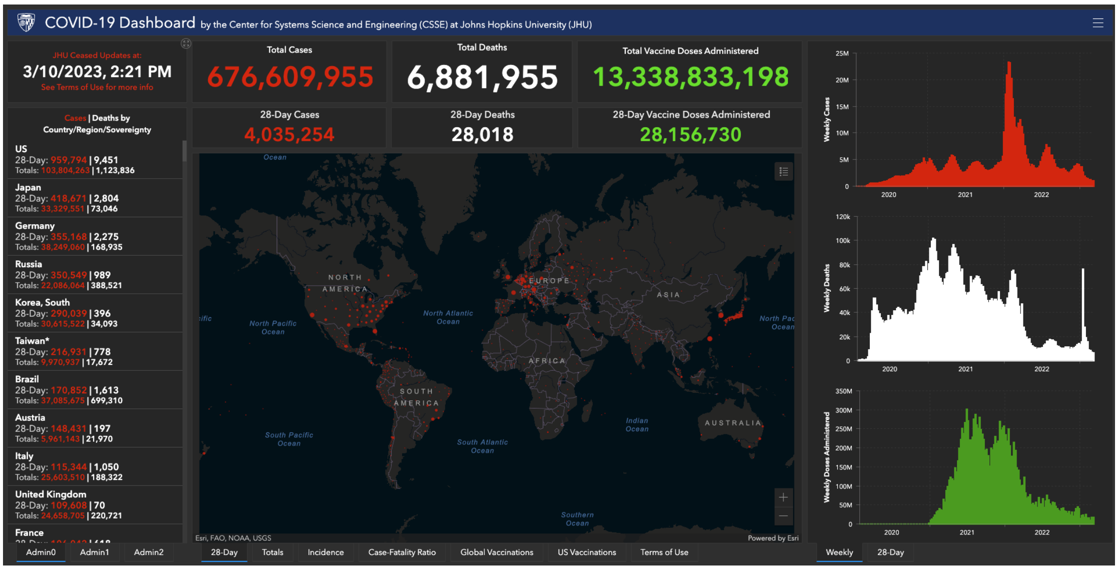

Figure 5: COVID-19 Dashboard [2]

5.2 User validation

To evaluate our dashboard, we opted for a continuous testing approach with the assistance of our peers and friends. Testing was done continuously at different stages of development. Each stage had taken into account previous feedback and contained relatively significant changes in the application. Below we provide some relevant improvements we made following the received feedback.

Scrollable page At first the whole page was scrollable because we had too much information that could fit in one page. Because of good information visualisation principles, we wanted to have a static dashboard that would fit in one web page for a more user-friendly interface. Therefore, the components that were making the page "too long" become scrollable, such as details of landslides and the charts container.

Date pickers Initially, the way we chose time intervals to display landslides was a casual date picker component with day, month and year. Although that was a fancy component, it turned out from the feedback of our peers that it was cumbersome and frustrating to use several clicks to select a day, a month and a year (twice for start and end). Moreover, we came to the conclusion that specific days were meaningless for our application. Therefore, selecting only a range of years would be enough. Ultimately, a slider was found to be the best option from a user point of view.

Description pop-up The details of selected landslides are sometimes very large, while others can be described in once sentence. To tackle this, we decided to display at first a "show more" button when the details description would exceed 500 characters. Once pressed, that button would pop-up a window with the full description. Later on, we came to the conclusion that it would be easier to just make the component scrollable, allowing the user to still see the rest of the dashboard while reading the details description.

Concerning the charts, we basically had the same idea: we wanted to make a pop-up with several charts. once again, making them scrollable ended up to be a better alternative.

Landslides types tabs We wanted to reduce as much as possible the number of different dropdown boxes for the user to feel comfortable and do less monotonous actions. To achieve this, we decided to include tabs below the map for the type of landslides. Switching a tab will update the markers on the map as well as the charts.

Dropdown choice ordering Concerning the first dropdown box about landslides triggers, we sorted it alphabetically so that finding a specific trigger is not too complicated. On the other hand, for landslide sizes, we sorted it from the smaller to the bigger. This makes the user experience more enjoyable and satisfying. Although it is a small detail, this sorting was suggested during a feedback session by one of our peers.

Moreover, we realised a small survey of the final product. The survey was composed of 17 different people, and were asked to be as neutral as possible, knowing it would only help us improve the application. Note that around a half of surveyed people were in the age interval of our target user base (12-18 years old in high school musical). The rest were a mix of IT fellows, friends and family.

Here are the criteria scored from 0 to 5 included:

- **Ease of use**, how easy is it to navigate through the app and understand the different components?
- **Design**, how good looking is the application?
- **Features**, how relevant are the features?
- **Performances**, loading times, correctness of the information being displayed.
- **Overall Satisfaction**, The average of the previous points.

This helped us improve several things in the whole application. Concerning the ratings, the averaged scores are the following:

Criteria	Average Rating
Ease of use	4.65
Design	4.24
Features	4.12
Performances	3.24
Overall Satisfaction	*4.06*

Table 1: Average ratings over a scoring from 0 to 5

We can assume the continuous feedback process was successful as the rating are relatively high, especially concerning the ease of use and general visualisation. The performance is lacking a bit behind as further explained in sections 5.3 and 6.1.

Summary Many changes were brought on the table, and the design of the project has changed significantly over the development process. This was due to the feedback we received, but also from a perpetual questioning about what would be the best user experience. Indeed, spending a lot of time on the same project can easily prevent from thinking outside the box and take a step back.

5.3 Loading times

The average performance of our dashboard is quite efficient, thanks to the implementation of certain restrictions and map optimisations. By allowing users to filter the data based on triggers, sizes, and time range, we are able to optimize the loading time. The map pins, descriptions, and graphs load quickly and almost instantaneously. This ensures a smooth user experience, enabling users to explore the information without significant delays.

However, in the worst-case scenario where all triggers and sizes are selected, along with the full time range, the number of pins exceeds 7500 and the loading time is drastically longer. In such cases, it may take a good couple of minutes for the dashboard to fully load and display all the necessary data. While the general page layout and graphs load quickly, the main bottleneck in this situation is the rendering of the map pins and especially the retrieval of specific pin information which is used to fill the description box.

6 Further improvements

6.1 Fix performance bottlenecks

Choosing too large time intervals or big landslides categories will cause the amount of data to be processed to rise drastically. In the current state of our web application, those amounts of information can take from seconds up to minutes to load completely. We did not find a solution of optimization to this problem in the given time we had to implement the project. The loading time impact the map, the charts, but mainly the details of clickable events. Note that these are extreme cases and that the application is perfectly usable overall. A solution would have been to put a limit on the maximum number of events loaded, but we would have lost the relevancy of the application by not displaying all of the information.

6.2 Include additional datasets

We could consider adding more datasets to the visualisation such as global temperatures changes or rainfalls in order to find correlations with landslides.

6.3 Implement additional features

Themed map pin icons One idea that we wanted to include was changing the markers icons based on the trigger of a landslide. For instance if a landslide was caused by a rainfall, an icon of rain would have been displayed instead. We could do this for multiple triggers and therefore have a nice visualization on the map that would allow to determine certain areas were specific triggers appear the most. Since we already had some optimization problems, we did not include this feature.

Landslide search bar In order to look for a specific landslide, we need to know its date, type and trigger. In order to speed up the process of looking for specific events, we could have implemented a search bar that would display a list of events based on a text prompt, such as a google research.

Accessibility features Once we have a fully stable application, we could tackle accessibility features such as allowing to change font-size, different languages or text-to-speech.

Fatalities Bar Chart Concerning the chart with fatalities and injuries count, we could superpose those data in order to get the sum of the two, which would me more visual instead of manually computing the two different bars. This could be done by putting 2 bars, the smaller inside the bigger one, which would be separated by two colors.

7 Conclusion

During this landslides project, we learnt how to represent specific data to targeted users following the information visualisation principles presented in the course Information Visualisation. Thanks to continuous feedback from our peers and the intermediate presentation of the project to the assistant, we could improve the solution bits by bits to achieve a user-friendly dashboard interface.

References

[1] NASA. *Global landslide catalog export.* URL: https://data.nasa.gov/Earth-Science/Global-Landslide-Catalog-Export/dd9e-wu2v.

[2] Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). *COVID-19 Dashboard.* URL: https://gisanddata.maps.arcgis.com/apps/dashboards/bda7594740fd40299423467b48e9ecf6.