

# Process Book - Dataviz

**Important!** This project only works on Chromium-based browsers. As a result, it will not work on Safari, but this shouldn't be an issue as we are supposed to use Chrome for this course.

## Introduction

Our project tackles the utilization of profanity in Quentin Tarantino's films. Renowned for his bold and uninhibited storytelling, Tarantino's movies often showcase a great amount of violence, both in action and language. So much, so that people thought it would be interesting to mark down each time a character in his movies swears. So they did, and ended up finding a little over 1700 occurrences of usage of profanity across only 7 movies he directed.

While many tried showcasing this data before, we decided to take a creative turn and give an interesting user experience to view said content.

## Path to Final Result

### Collecting the data

We began by gathering relevant data from reliable sources on the internet. This involved researching Quentin Tarantino's movies and their dialogue, specifically focusing on instances of swear words. This was obtained from a publicly available dataset, containing the movie name, the swear word, and the death occurs, along with a timestamp in a CSV file. We ended up with a total of 7 movies to showcase.

Fortunately, we did not have to preprocess the data as it was already clean and ready to be used. We did some basic analysis to get a better understanding of the data and to get some ideas for the visualizations however as explained in the 1st Milestone report.

We also had to use a small dataset that contained the movies and other facts about them to show as extra information alongside the profanities. This wasn't too hard to make sure we created a simple JSON file containing information such as the movie synopsis, length, year of appearance, actors, images, and a couple of fun facts as well :)

### Visualization Strategy

We deliberated on the most effective way to visualize the data to provide meaningful insights to the audience. Taking into consideration the nature of the project and the goals we set, we decided on a visual representation that would simply showcase the frequency and distribution of swear words across Tarantino's movies.

What would set this apart from a simple table view? Interaction and context. We want the viewer to get a grasp of the quantity of profanity that is thrown around. 1700 words sounds like a lot but can one actually visualize the quantity? **The idea was to make the user feel the number of words as the time unfolds in a movie in an impactful way.**

We also wanted to give design context to the visualization and we made sure to convey that it was about cinema/movies.

## Code

To create our chosen visualization strategy, we developed functions that generated HTML code dynamically through Javascript. This code was responsible for rendering the visual elements, such as the reel (chart), graphs, or interactive components (movie tabs/information/popups), based on all of the processed data. The functions ensured the seamless integration of the visual representation into the final report.

## Enhancing General Visual Appeal

To make the data more captivating, easy to read, and engaging, we incorporated additional visual components. These enhancements included animated effects, color schemes, or interactive elements, providing an immersive experience for the viewers some were added later. These proved to be very important and gave the final project a lot of liveliness.

## Design Decisions

Throughout the project, we encountered some challenges, mostly related to design decisions. This was about how to be original and expressive in the way we showcase our data.

### Giving context/story to the data

We wanted to make sure that the website is interesting and immersive enough for the viewer to have the best experience: we discussed how we could use cinema and movie icons/visualization to present the data, given its context.

Immersion was reached mainly through styling and design. We used specific fonts:

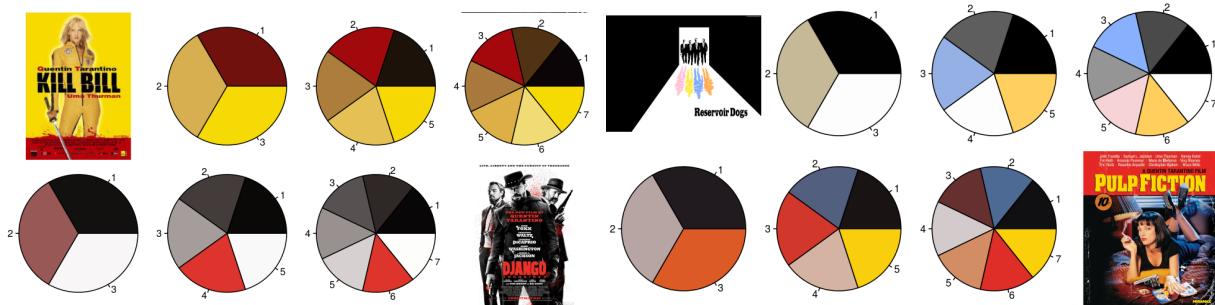
- Typewriter-like to give a storyboard effect,
- A font resembling the Hollywood signs for capitalized letters, titles, and other important elements
- And a slim font occasionally to give a polished look to the website overall.

There are two main color palettes on the website which also gives a very targeted tone to our website.

- Dark/Yellow/White that conveys “danger” or a sign of some warning. This is due to the nature of profanity, the strong language. We made elements on the site sharp/strong to convey the violence of the words.



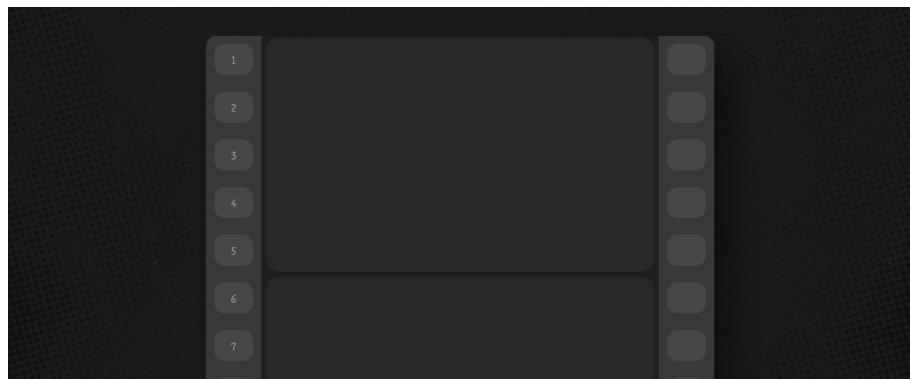
- Color palettes directly from Tarantino's movies. They are often very Western-like, vibrant, industrial (hence the background), and very unique.



## Repartition of the data

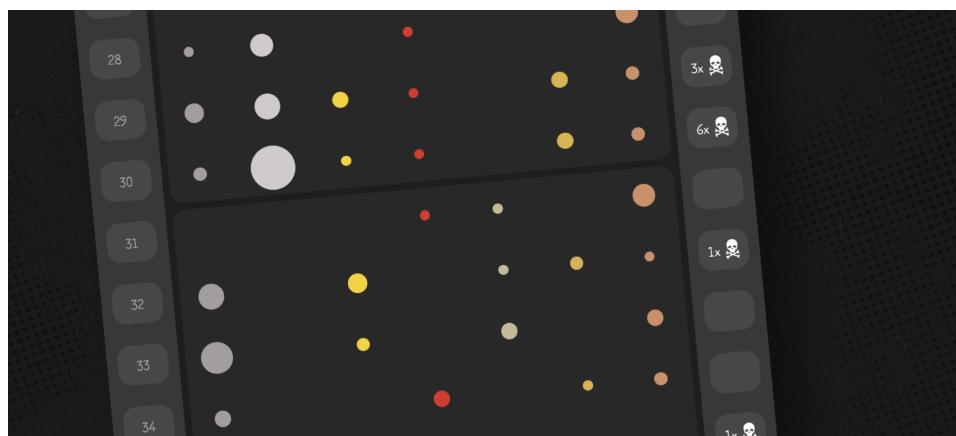
### Movie Reel

For the main component, we decided to use a long film reel to showcase the data. We thought that having the viewer scroll vertically would make them feel like they are actually watching the movie and would be the most impactful (and fun!) way, as mentioned previously.



We managed to show the important data on the reel, with the minutes being indicated on the left and the actual swears depicted as circles on the reel: the bigger the circle, the more swears a movie has at that specific minute. On the right, we indicate the number of deaths for every minute (if any).

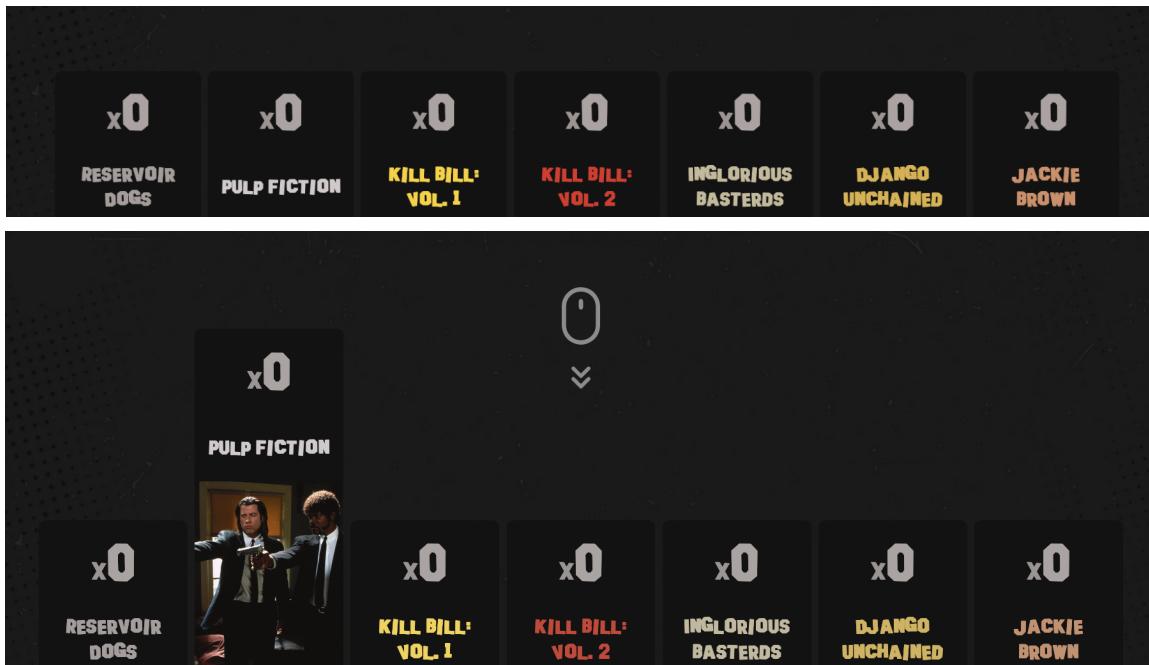
This design decision induced the need for programmatically generated HTML code, which was a nice challenge and directly related to the course content.



## Movies tab

We added a tab at the bottom of the reel that follows the scroll and sticks to the bottom of the screen. This shows the number of swear words that appeared up until a minute (how much the user scrolled), and the name of the movie. The color of the movie names is used for the colors of the dots to keep the data matched to its appropriate movie. The movies are all docked in this tab, but if the user hovers over a movie, it pans up and reveals its cover.

This tab ends up unsticking to the bottom of the page and will become a normal nav bar allowing the viewer to still select a movie for more info once the user scrolled down completely, along with the total number of swear words. It is arranged in a card stack effect.



## Movies “More info” module

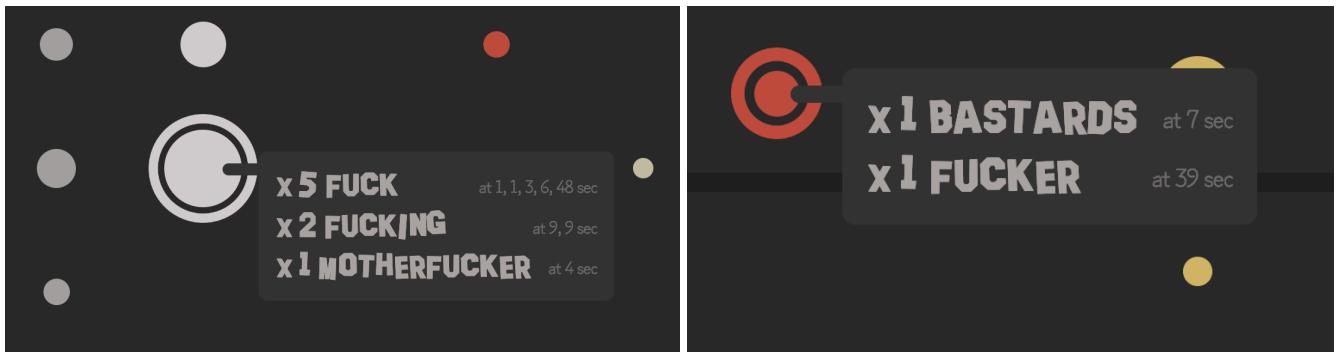
Once the user clicks on a movie from the movie tab, an overlay screen glides up that blurs the background and displays information about the movie. Once the viewer reads it, they can casually close it by clicking the back arrow, clicking on the blurred background, hitting Esc on their keyboard, or casually continuing to scroll. It allows a very smooth integration of this extra module.



## Popup revealing more info on profanity

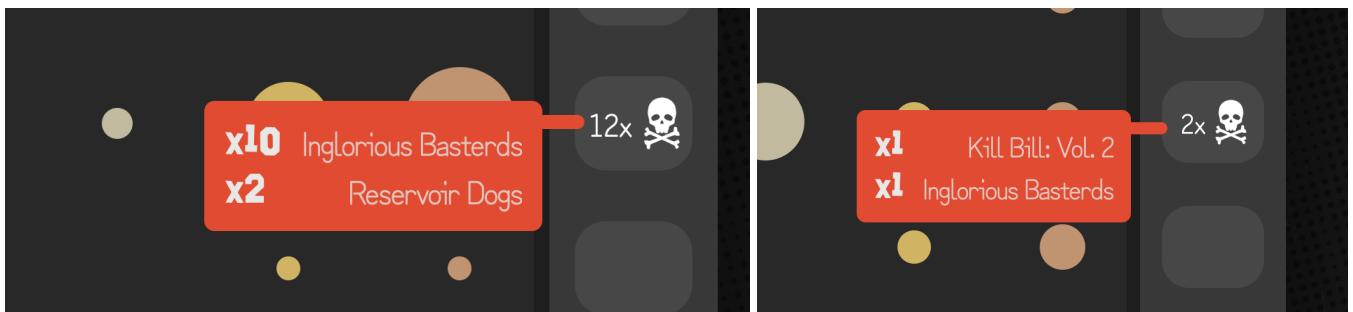
If the number of swear words for each movie wasn't enough, we've added the possibility to hover a circle to display what words exactly were used in the minute for that specific movie.

This displays an aggregate view of the words (ordered by most used) with a glide-in animation, and also tells at what second in the minute this word was said. This box will follow your mouse until you un-hover the circle.



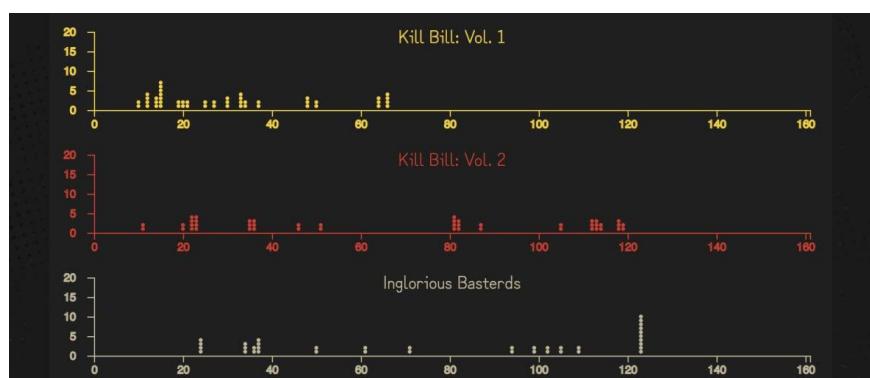
## Popup revealing more info on the death

Similarly to the info on profanity, we added a way to see what movie featured a death, because the deaths are summed up per row (minute).



## Overview data section

Finally, to recap the project and to give even more insight into the data, more in a “bird’s view” way, we decided to add interactive tables. These tables are per movie, where in each the number of profanity is measured by time. The more dots the more words a minute. Upon hovering one word, the words (dots) take a color across all movies according to certain profanity groups (f-words, racial slurs, ...). This allows the viewer to see correlations in movies, to get an idea of what types of words there are.



## A lot of interactivity/animation

To make the viewer feel the actual impact of the words, we decided to use advanced animation as they scroll the reel:

- Each time a new line (minute) is reached and appears at the bottom of the page, all dots from that line appear by popping up. This happens each time the user scrolls by: this gives a feeling of liveliness to the data.
- When a dot appears for a movie, the movie tab below associated with the movie shakes and increases the number of word count for a movie: this is again for liveliness and to give visual feedback to the user that "something happened".
- When more than 8 words are counted for a minute, the movie tab strongly shakes and flashes: this gives the effect of heavy impact indicating a lot of swearing happened.

## Challenges

One major issue that led to harder-to-maintain code was data colors. Instead of using classes to properly set colors for the different movies, we had to use the style attribute. This leads to some extra style attributes getting generated in the HTML for the elements. This is not too big of an issue.

Another challenge was properly scaling the data. It was hard to make sure that it was not too long to scroll, but at the same time not too tight so that the data was still visible.

## Changes compared to previous milestones

### The data “flows”

The idea of the reel is to show the data unrolling at the movie unroll/runs. There is a notion of a flow of time that when you reach the end of all the movies, the data (profanity count) will add up and show you the final result through the docked movie tabs. We did not account for this at first, but after playing around with static prototypes of the website, we decided that including something like this would be extremely beneficial and adds to the experience.

The user is essentially the one who makes the data add up while scrolling, this boosts interactivity and puts the viewer in the center as the “discoverer”, and we believe what makes this visualization powerful. This turned a simple static visualization experience into a dynamic and engaging one.

The extra animations and effects for the reel give an additional layer of satisfaction when scrolling through the page.

We now will let the reader try out the website to experience this feature for visual confirmation. The website can be [found here](#).

### Expanding on movie information

While we wanted to include information about each individual movie, the prototype did not contain what specifically we wanted to add. So we had to further expand on that. We added:

- The list of main actors, their names, and the character they portrayed
- Fun facts about the movie
- A frame of the movie is displayed on the right
- A call to action button to go back next to the title for intuitiveness

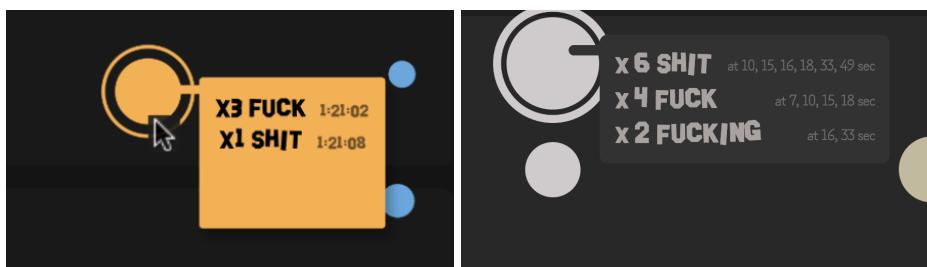
Here are the visual differences:

The image shows two side-by-side movie information pages. The left page is for 'KILL BILL (2004)' and the right page is for 'KILL BILL: VOL. I (2003)'. Both pages feature a large movie title at the top, followed by a runtime indicator ('RUNTIME: 1H38' and '1H51' respectively). Below the runtime are sections for 'SYNOPSIS' and 'ACTORS'. The 'SYNOPSIS' section contains placeholder text ('Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat.') and a small white rectangular box. The 'ACTORS' section also contains placeholder text. The right page includes a small thumbnail image of a woman holding a sword.

The popup reveals more info on cursing

At first, this module displayed only one timestamp for each word on the right, which could have been a better design because there are multiple words occurring in that minute. We decided to change the time indicator to the “at 2, 3, and 10 seconds in” format which indicates at what time the words were said in the given minutes. We only show the seconds.

We also removed the color to avoid overloading the site too much, an example of before/after is shown below.



## Peer Assessment

**Idea/brainstorming:** Both Daniel and Louca participated equally during this task

**Design drafts:** Both Daniel and Louca participated equally during this task. Daniel took care to design the overall reel and graph drafts, Louca drew the extras and header/title and came up with the data presentation.

**Data collection:** Daniel took care of finding and importing the cussing data or creating it in case of the movie information.

**Website implementation:** Daniel took care of the HTML and project skeleton, theme, and style.

**Data manipulation:** Louca took care of aggregating and making the data usable, which allowed him to tackle the reel and graph generation. He also worked on the bottom graph's intractability.

**Website polishing:** Daniel made the reel visualization interactive and engaging by adding animations, effects, pop-ups, and movie information on the site.

**Progress book:** Louca wrote the skeleton of the report and Daniel added the subsequent content.

**Screencast:** Daniel took care of the demonstration

**Overall assessment:** We believe we both spread the workload in an equal manner and tried to work together as much as possible so we would stay aligned and motivated to produce the best possible result. And so we did!

## Conclusion

In conclusion, our project on profanity in Quentin Tarantino's films has been an exciting and creative journey for us! We aimed to provide an immersive user experience, showcasing the frequency and distribution of swear words across Tarantino's movies, which overall is not the average dataset that you would encounter on a regular basis.

We collected data and conducted basic analysis to gain insights into the information at hand. Through dynamic HTML code generation, we seamlessly integrated our visual representation into the final report.

To captivate the audience, we incorporated animated effects, color schemes, and interactive elements. Our design decisions, influenced by cinema and movies, aimed to give context and tell a story through the data. Fonts, color palettes, and design elements were carefully selected to enhance the visual appeal and convey the main theme.

The film reel served as the main component, allowing viewers to scroll through the data and experience the unfolding quantity of profanity, resembling a movie-watching experience where the user "creates" the data by advancing. Additional features, such as movie tabs, "More info" popups, and word hover functionality, added more interactivity and insight into the subject.

Our project successfully transformed a static visualization into a dynamic and engaging experience. By immersing viewers in Tarantino's world, we hope to captivate their attention and provide a fun, out-of-the-ordinary glimpse into profanity utilization in his movies!