Comics and Diversity

Project for the Data Visualization course

Students:
Jelena Banjac,
Sofia Kypraiou,
Jelena Malić

Professor:
Kirell Benzi

Contents

1. Introduction

1.1 Overview

Comic books are intended to help many young people experience the first connection with the magical world of books and cinema. As such they have a profound impact on the way the mind, judgment and future preferences of a child develop during their coming of age. At the same time, comic books represent a way to better understand the social, economic and political reality they were created in. While doing this project, we wanted to see how comic books were used over time to show opinions and imagination of their creators and we hope our visualization would help in the discussion of the reasons some characters were represented in a certain manner. As a legend of the comic books world, Stan Lee once said:

"Some people will say, 'Why read a comic book? It stifles the imagination. If you read a novel you imagine what people are like. If you read a comic, it's showing you.' The only answer I can give is, 'You can read a Shakespeare play, but does that mean you wouldn't want to see it on the stage?'"

This quote was only one of the things that inspired us to dedicate this project to not only commenting on the way comic book characters evolved, but to also try to visualize it so that all of the discrepancies and differences would become even more obvious. The main objective would be to create a sequence of the way the representation of good and evil has shifted across time, but also of the ways the culture of diversity became more present in the entertainment industry.

We are all aiming to live in the world full of diversity and acceptance, but even with all of the improvements in recent years have brought us, what have we achieved so far? We hope our data visualizations will give us all some much-needed answers.

1.2 Motivation

"Words and pictures are yin and yang. Married, they produce a progeny more interesting than either parent." — Dr. Seuss

This is the quote that was the leitmotif of our whole project. Starting with the topic of comics, which themselves represent a special merge of the world of books and pictures, we knew that we wanted to achieve something that is as creative as that. We also knew that using only words to explain data that we found would not give enough inspiration and insight to our users, which is why our visualizations had to be both picturesque and interactive.

We hope that this project will give comic book fans a new way to explore their favorite characters, and motivate some of the users who were previously not that interested into this fascinating phenomena to discover the magical world of comics. At the same time, we are hoping it will encourage them to think about the issues of diversity and representation of minorities, not only in the imaginative world of comics, but in real life as well.

We were determined not to get scared of the fact that we chose the topic of diversity in comics, as a sort of taboo topic where you have to watch carefully at what you are saying if you do not want to offend or shock anyone. We were persistent in bringing this topic into light, but we decided to do it in a playful and colorful manner that is supposed to inspire and not to judge. We will leave it to our users to get their own conclusions from exploring the data story, and we will just give them the visualizations to lead them.

1.3   Target Audience

We wanted to make this visualization understandable for all of our potential users. Both the comics experts and the ones that are only starting to discover the colorful world of comic books will hopefully be equally entertained with our website and findings. We are also hoping to make people who never had any closer look into the comics universe more interested in this topic.

We would also be very happy if some of our ideas and representations from this project inspired our colleagues interested in data visualization to create new data stories on topics of either diversity, or comics in general, or something completely new and unexplored.

One specifically interesting use case would be if the creators of the comics themselves would take a look at our project, in order to either critically observe their and their colleagues' work. But it would be even more thrilling if they get inspiration to create some new and unique characters in some of their future issues.

2. Discovering what to do...

2.1 Inspiration and related work

Diversity and under-representation of minorities in comics are issues that have inspired many other data explorers and researchers. Before we started to grapple with our own dataset we tried to investigate different ways in which this topic was previously analyzed and represented.

At first we approached the issue of gender:

- One of the databases that we used came from the article *"Comic Books Are Still Made By Men, For Men And About Men"* from **FiveThirtyEight** [1]. There the authors approached the gender representation in the big two comic publishers, DC Comics and Marvel. It consisted mainly of statistics and text like shown on figure 2.1. The information and the graphs gave an a great insight about the gender representation and how it evolved though the year. However we wanted to make it more interactive.
- Another article regarding the gender was *"Analyzing the Gender Representation of 34,476 Comic Book Characters"* from **The Pudding**[2]. We found this article extremely informative and it had some great visualizations. This inspired us to use the database that was used for their analysis (ComicVine). Also the figure 2.2 served as an inspiration to explore the relations and the members that make a team. However something that we felt was missing was time-series data.

But diversity is not supposed to be only about the issues of gender. With this thought we decided to search for different ways race, sexual orientation, religion are represented in comic books:

- *"Race, Orientation, and Gender in Marvel Comics"* from **Marvel Wikia** [3] is a great article describing the above-mentioned concepts. Marvel Comics depicted visible minorities in ways that, for the times, were considered socially acceptable but by modern standards are considered to be offensive. The article identified these ways and put them into a historical context, for the purposes of understanding them as such. The article has great information, but no visualizations. To make this information more
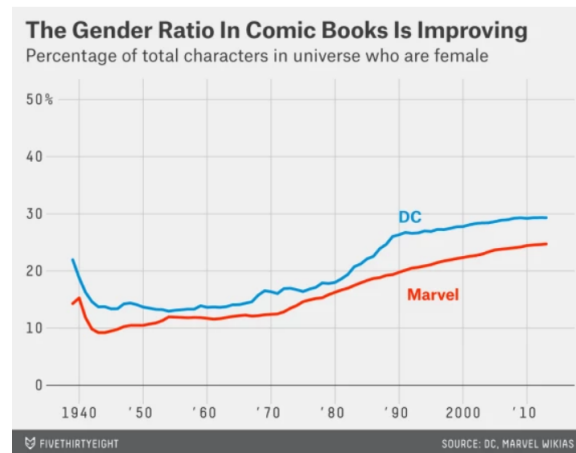
Figure 2.1:   Annual rate of the percentage of new characters introduced each year who are female.
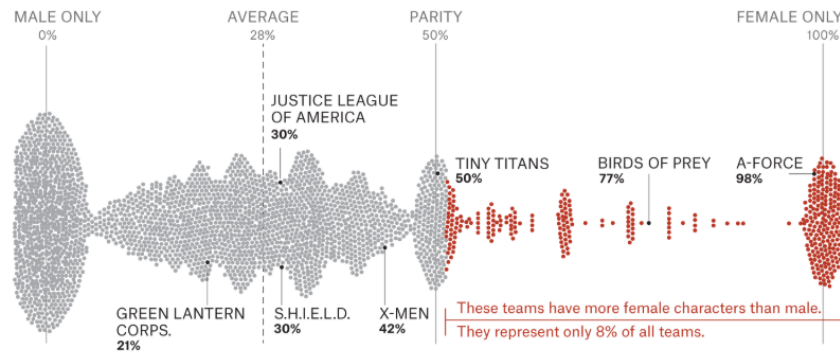

Figure 2.2: Female percentage of every team - Each dot represents one of 2,862 teams in DC and Marvel.

      accessible, this inspired us to create a timeline for the important minority-represented characters that were introduced though the years and contributed to enhancing the diversity in the comics universe.

- Also the site **Den of Geek!** released a series of articles taking a look at the portrayal of Minorities in the World of Comics. The series includes lists of 10 representative comic characters with *disabilities* [4], of *Color/Hispanic* [5], belonging to the *LGBTQ community* [6] etc. This inspired us to collect all this information and visualize it so that it becomes more approachable.

2.2   Objectives

The main objective is the exploration of the culture of diversity in comic characters in recent years and the way the representation of good and evil has shifted across time, but also of the ways the culture of diversity became more present in the entertainment industry. We want to make our data story interesting and informative.

To accomplish this objective, we analyzed our main goal in the following subcategories:

- Represent and explore the different characteristics that comic book characters have, like their gender, alignment, identity etc. The users can explore can gain insights about the characteristics that them.
- Find out how the average comic book character looks like. Explore and group the characters by their physical attributes, like hair and eye color.
- Explore how the origin of the character (human, alien, robot, etc.) influences its alignment in the world of comics (good, bad or neutral).
- See how the origin and the gender is represented by observing the most important characters in the most influential teams.
- Represent a timeline of the most important characters who contributed to widen the diversity in the comic world while also identifying the historical context in which they were created.

2.3 Exploratory data analysis

We begun our exploration from the FiveThirtyEight dataset, an overview of which can be found on figure 2.3. This database consisted of information regarding the physical attributes of the characters (like hair and eye color) and date of their first appearance in comic book issues.

Regarding the preprocessing step, for the prototype version we decided to merge the information about specific team members obtained from the Comic Vine API [7] of those teams with the data that is available in the 'characters' dataset also collected using the Comic Vine API. This way we obtained significantly more information about each character, which can be used to give more information to our readers. For the 'origin' column which is the only one containing null values we have decided that the characters for which the origin is not stated, will be observed as belonging to a new category named 'Undefined'. Additionally, in order to understand which characters belong to the same teams for a chord diagram we ended up visualizing, we created a 'relationships' dataset that is used only for representing the relations between characters.
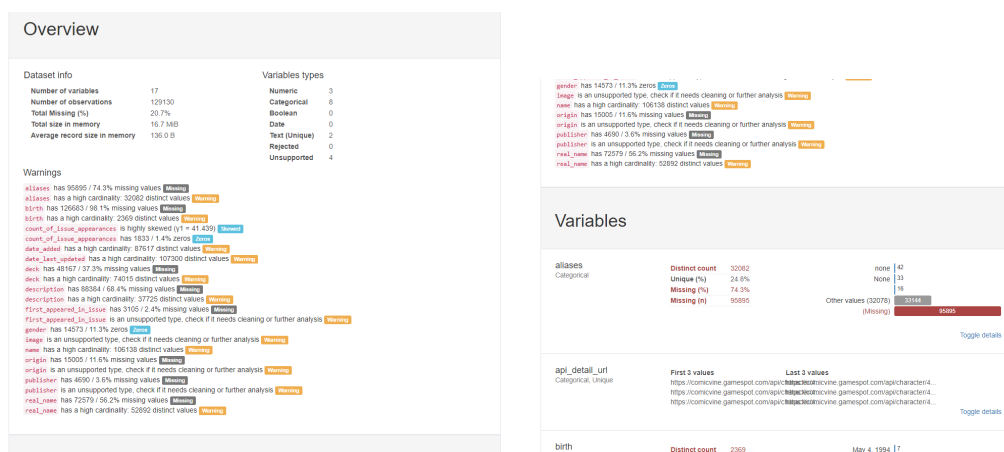

Figure 2.3: Screen-shots of the part of the report regarding FiveThirtyEight Data

Additionally, we also decided to merge the two datasets: Comic Vine and Five Thirty Eight [8], using their character name column. This way we managed to collect even more information that would be of interest to our readers. The Five Thirty Eight dataset contains mainly relationships as well as some appearance information. On the other hand, Comic Vine dataset contains the information such as what is the color of eyes of our character, what is their hair color, etc.

We also provide a directory *data/notebooks/reports/* on our GitHub page with the detailed reports about our datasets. The general look of these reports is shown on figure 2.3.

3. Design

3.1 Our journey

From the moment that we decided on our dataset, we knew that we wanted to represent as many different aspects of diversity as possible, while at the same time keeping our data story interesting and informative [9]. The process of creation made us overcome a few obstacles in the beginning stages, as we had to improve and rethink many of our approximate sketches and basic ideas.

Our first sketches were mostly based on some more common graphical representations, like different types of histograms enriched with the additional information (as can be seen on figure 3.1).
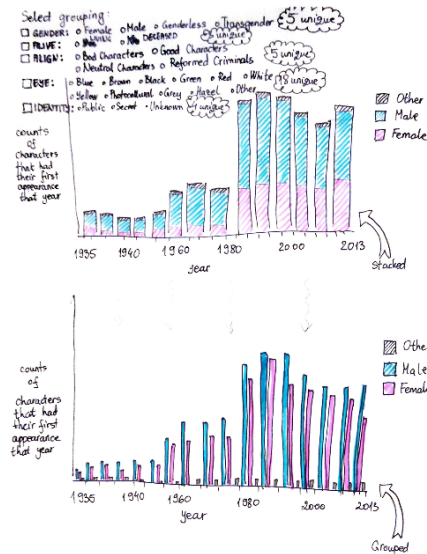

Figure 3.1: Histogram sketch

We then decided that a graphically innovative topic like comics needed some more visually challenging elements. This is why we started exploring different ways of using shapes, and contrasts between straight lines and smooth surfaces on different graphs in order to make our story more vivid and eye-catching. We were also certain that in addition to statistical data, our users will be motivated to observe specific attributes of some of the most famous characters that were ever drawn in comics. We came up with a simplified version of a chord diagram that is shown on figure 3.2.
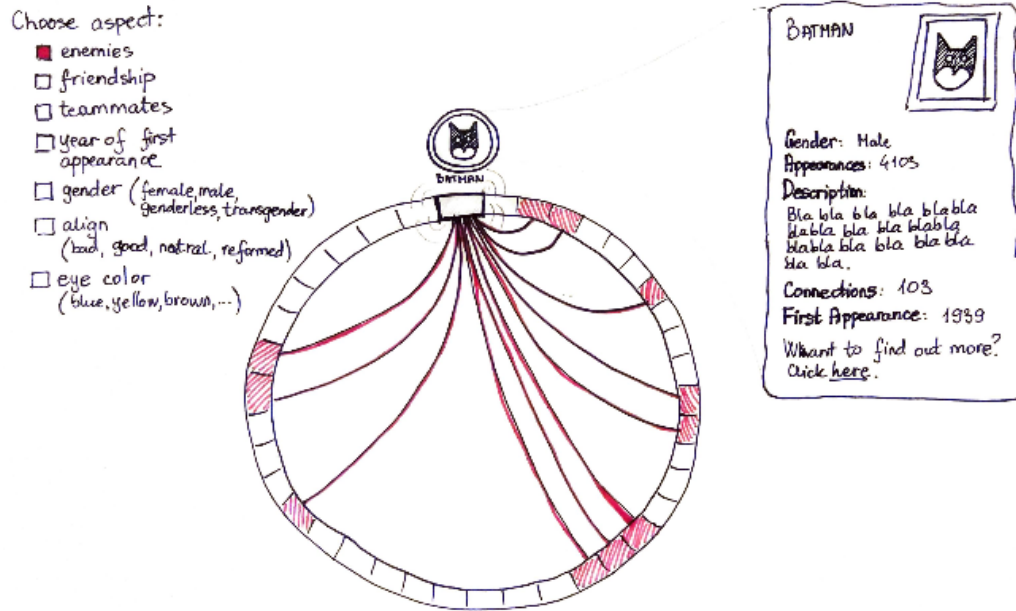

Figure 3.2: Chord diagram sketch

The other diagrams were added after determining that it was easier to perceive different patterns in our data if we add some variety and create several different visualizations, than if they were all shown using only one complex diagram, even if it was very interactive. These diagrams will incorporate the pictures of different comics characters in all of the appropriate places, since this will enable our users to relate our findings to the concepts and characters they are already familiar with.

The visualizations will take into account the aesthetic, functional, and sociopolitical dimensions as much as possible, while staying entertaining and practical. The website on which our visualizations will be shown is supposed to be quite simple and straightforward, in order to keep the attention of our users on the diagrams and data story. The website design will try to mimic the captivating and vibrant aspects of comics by using modern fonts, several pictures and suitable color schemes.

The first wireframe sketches were only representing a very simple one-visualization plans (like shown on figure 3.3), but the main idea of the text sections necessary for each diagram remained the same. The data story of course starts with an abstract explaining our main ideas and motivations. Then for each of the visualizations, we give a short explanation of the data and the findings that we have found interesting, with regard to the specific features

represented on the following diagram. In the end, right before the visualization there is a short section with instructions for the user, that are supposed to simplify the interaction to the user. The visualizations all have either clearly stated categories right on the diagrams themselves, or a legend where the color schemes are be explained in detail.
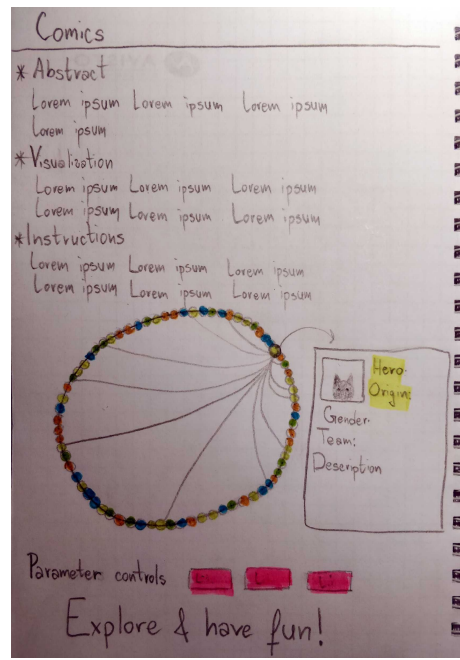

Figure 3.3: Wireframe sketch

3.2 Design decisions

At the end of the road with many twists and turns, we decided to implement five different visualizations for our project:

- Parallel-sets (parsets) diversity statistics visualization,
- Bubble appearance visualization,
- Bipartite comparison graph,
- Chord diagram with most important characters,
- Timeline of diversity.

We are hoping they will allow the users to have an overview of the general phenomena that can be observed from the dataset. At the same time, those who are interested will have a possibility to take some time to interact with the visualizations and get a deeper look into the behavior of specific features. Each of the diagrams are be focused on different aspects of our topic, but they will lead the user from the beginning which shows only the different statistical features, to the diagrams that represent some of the specific characters that embody the world of comics, and the diversity that it now contains.

3.2.1 Parsets visualization

For the first visualization of our data story we chose the parsets diagram (as shown on figure 3.4). This visualization represents the statistical frequency of different categorical features

that the comic book characters can have. It shows the relations between four different factors:

- Gender,
- Alignment,
- Alive,
- Identity.

The users are able to interact with this graph in order to get precise insights about specific subgroups that interest them. The axes of the parsets can be reordered, in order to enable discovering different overlaps in categorical data groups. In addition to this, the categories themselves can be reordered according to alphabetical order and size of the subcategory.
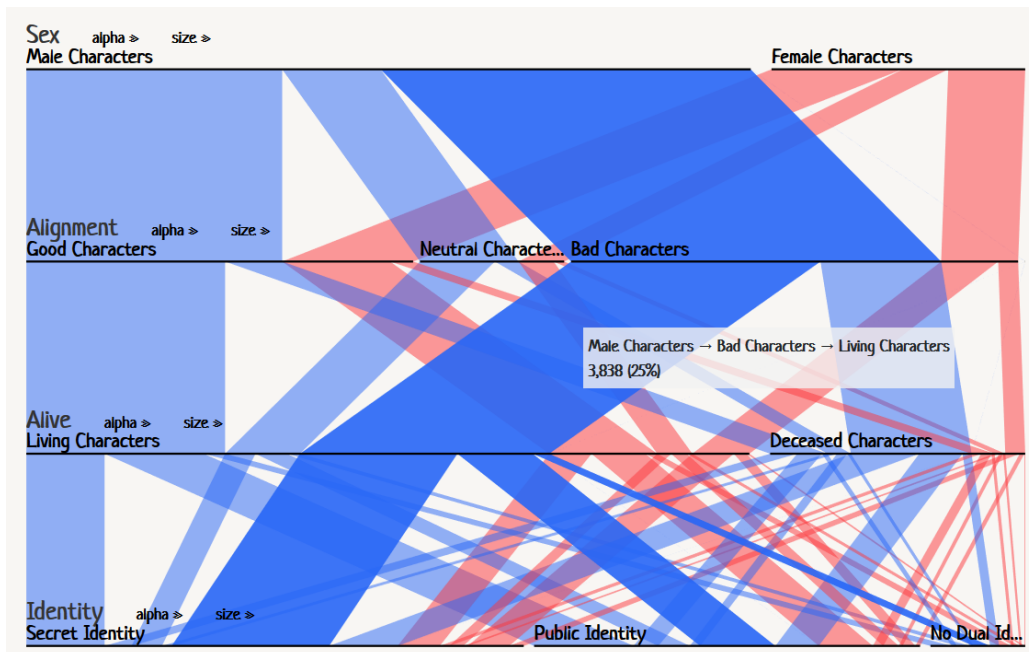

Figure 3.4: Parsets diversity statistics visualization

3.2.2   Bubble visualization

We had several ideas about how to represent different physical attributes of characters in comics. We wanted to merge our two categorical features (hair and eye colors) into a full appearance of a character, but at the same time enrich this with a notion about how common this feature combination is.

The traditional tabular data summaries took us in the direction of a very straightforward, but also not very informative mosaic diagram that we tried to implement. After some implementation issues, and reevaluating the usefulness of representing two-colored squares with numbers that give user concrete information only when hovered over we decided to drop this diagram.

However, we were still determined to represent the physical trait combinations, but we were facing a problem of finding which diagram gives us enough room to show the looks and size of subgroups at the same time. We decided to go with the relatively simple

representation using bubbles of different magnitude, as the users will in this way recognize the commonness of groups even from just glancing at the visualization. This also enabled us to put pictures of different characters as the background for the bubbles, but we also kept the option of hovering over circles in order to obtain more information about a specific subgroup (as shown on figure 3.5).

Figure 3.5: Bubbles appearance visualization

3.2.3 Bipartite graph

Another interesting path of discovery lead us toward comparing the representation of different characters depending on their origin. The visualization we implemented is representing the overlap of two categorical features: character origin and their alignment using a bipartite graph (represented on figure 3.6). On the left axis there are all of the subcategories of origins (Alien, Animal, Cyborg, God/Eternal, Human, Infection, Mutant, Other, Radiation, Robot) with the percentage of characters that have that origin. Similarly, on the right axis there are the subcategories of alignment (Good, Bad, Neutral Characters, and Reformed Criminals), along with their percentages.

When user hovers over one of the subcategories on either axis, the visualization changes and shows the way that the specific subcategory is divided over categories of the opposite axis. That gives user the opportunity to explore the data, and come up with different discoveries about the representation of good and bad in comics.
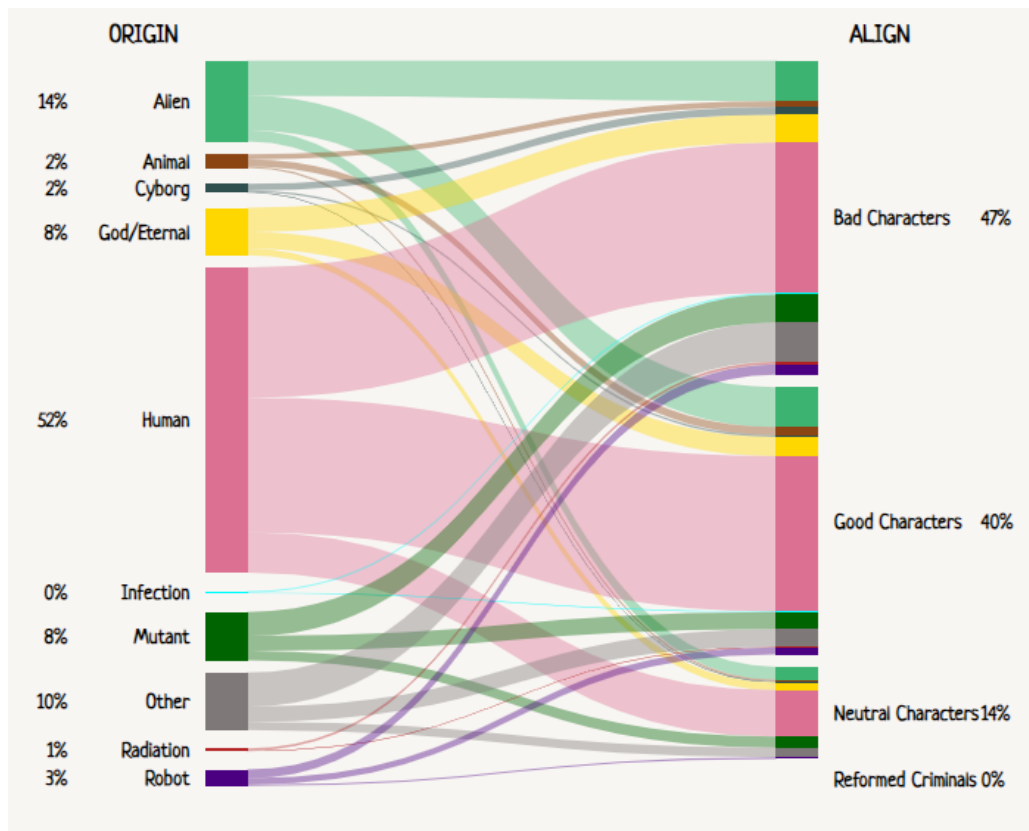
Figure 3.6: Bipartite comparison graph

3.2.4 Chord diagram

As we mentioned during the ideation process, it was important to us that we also represent many famous characters in person, not just through statistics. That is why we created a chord diagram shown on figure 3.7 that has 100 most relevant characters from comic books, and that belong to ten famous teams: "Avengers", "Defenders", "Guardians of the Galaxy", "HYDRA", "Injustice League", "Justice League", "Masters of Evil", "S.H.I.E.L.D.", "Suicide Squad", "X-Men". They can be sorted with regard to three different criteria: Team, Origin, Gender, which can be selected using buttons in upper left corner of the visualization. The fact that a specific character belongs to one of the subcategories of the criteria is represented with color coding.

We decided to use the Gestalt principles [10] and especially connectedness as one of the strongest natural grouping options to represent the teams that the characters belong to, even when they are sorted according to different criteria. This will enable our users to also observe the diversity within the teams, and explore the relations that exist between characters which are similar. The user can also hover over the character cell or click on its name in order to see more details about a specific character, or go to the corresponding Wiki page.
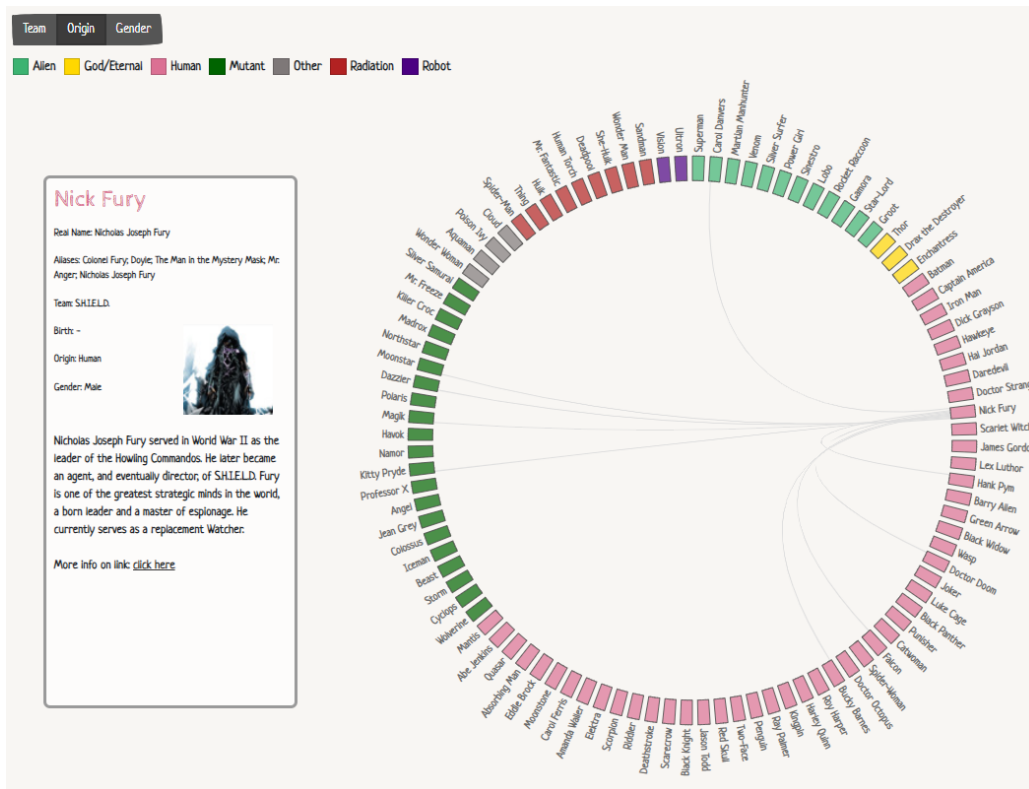
Figure 3.7: Chord diagram

3.2.5 Timeline of diversity

From the early steps of our research about the diversity and the minorities in the comics, we knew we wanted to create a timeline. We wanted to tell a story, we wanted to tell how new characters who were introduced widen the diversity in the comic books world.

The design and the implementation went through several stages until we reached the final result. Our initial design was to create a big timeline with all the characters that were introduced each year and a special mention to the ones who were representing a minority. However we felt that the information was overwhelming with this design and that it would be difficult for the user to navigate.

Our second approach was to create a bigger timeline with a navigation bar. A prototype implementation can be seen on figure 3.8. It was closer to the direction we wanted to take, but still not quite there. It was way more complicated for the user and it was not obvious how to interact with it. In short, it was over-designed.

Finally, after the lecture on "Storytelling", what really caught our attention was the connected scaterplot from the **NY times** article *"Driving Safety, in Fits and Starts"*. This served as an inspiration to scrap our previous complex design and aim to something more simple but more clear for the user. So we selected fewer but more important characters who played a significant role in history of diversity, like the first minority superheros introduced by Marvel and DC Comics. The figure 3.9 shows our final design. The timeline is strictly defined and there are labels that show the year and the name of the most important, in

Figure 3.8: A prof of concept for an initial timeline design

terms of minority history, characters introduced. As comics are a depiction of the word we live in, we wanted to explore the background and the sociopolitical events that were taking place when each character was introduced. This lead us to add a history description of the decade on the left of the timeline which appears when the user clicks on a label for the specific character.

Figure 3.9: Timeline when selecting the Silhouette label. At the left of the time there is information about the decade the character was introduced and the political events, and on the right there is information about the character

4. Implementation

In this section we will elaborate the implementation details, based on the final implementation model described in the previous section. We will also 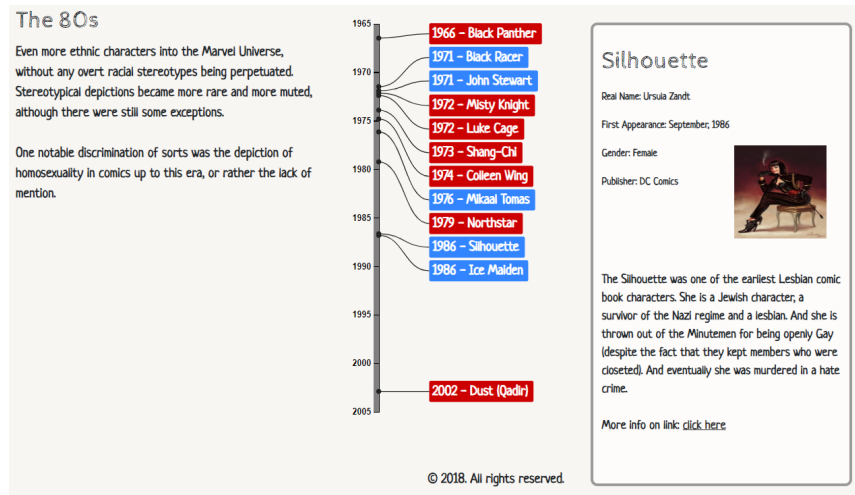cover technical details: libraries used, the functionality and reasoning behind it, as well as the overview and uses of the visualization.

4.1 Project architecture

On figure 4.1 the organization of our project structure is displayed and we can see the elementary parts contained. We tried to make it as modular as possible.

From the mentioned figure, we can see that there is a *brainstorming/* folder that contains the files with all the information about our initial ideas and brainstorming sessions regarding design requirements before the coding started.

Directory *data/* contains the files from our data repository. In our data repository, we have collected data from the ComicVine website and FiveThirtyEight website. The directory *data/notebooks/* contains all the data analysis that was performed after the collection of the data for our data visualizations. Therefore, the data analysis notebooks include the data collection, data cleaning, and data analysis.

Directory *images/* contains the images for our README files only. Directory *prototype_data/* contains data that was initially used to make visualization running as well as test our initial implementations of them.

Directory *src/app/* is the root folder of our web application. As usual, it contains *src/app/css/* containing our style sheets, *src/app/data/* containing our data for the final data visualizations, *src/app/img/* contain images for our website, *src/app/js/* JavaScript files that are used for our visualization implementations. These visualizations include bipartite graph, bubble graph, chord, parsets, and timeline visualization. We also use the webpack package manager.

```
ExploringComics.github.io
├── brainstorming/
│       ├── BrainstormDesignRequirements.pdf
│       ├── DesignAndDevelopingProcess.pdf
│       └── WireFrame.pdf
├── data/
│       ├── ComicVine/
│       ├── FiveThirtyEight/
│       ├── notebooks/
│               ├── data/
│               ├── reports/
│               ├── DataAnalysis-ComicVine.ipynb
│               ├── DataAnalysis-FiveThirtyEight.ipynb
│               ├── DataCleaning-Bipartite.ipynb
│               ├── DataCleaning-ChordVisualization.ipynb
│               ├── collector.py
│               └── helpers.py
├── images/
├── prototype_data/
├── src/
│       ├── app/
│               ├── css/
│                       ├── bipartite.css
│                       ├── bubble.css
│                       ├── chord.css
│                       ├── parsets.css
│                       └── timelines.css
│               ├── data/
│               ├── img/
│               ├── js/
│                       ├── bipartite.js
│                       ├── bubble.js
│                       ├── core.js
│                       ├── parsets.js
│                       ├── data-handlers.js
│                       ├── main.js
│                       └── timeline.js
│               └── index.html
│       ├── package.json
│       ├── package-lock.json
│       └── webpack.config.js
├── index.html
├── LICENCE
└── README.md
```
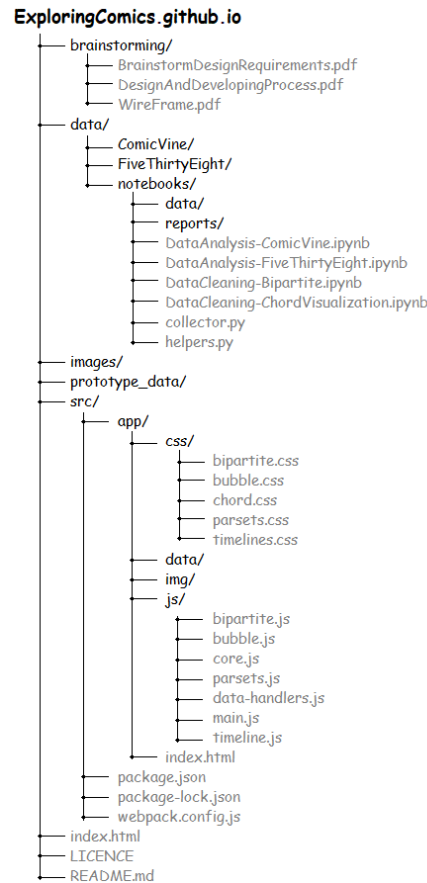
Figure 4.1: Project Architecture

4.2 Design and interaction elements

In this section we will describe what libraries we used as well as other technical details regarding the code implementation.

It will describe in detail the functionality and implementation of all five visualization, as well as the interaction with each of them.

4.2.1 Parsets visualization

For the parset visualization, see figure 3.4, since we wanted to visualize the proportions between female and male characters and how many of them have which kind of alignment (good, bad, or neutral), we thought that using **Parallel Sets** type of visualization will be the best way to visualize it.

Regarding the user interaction, the visualization provides option of hovering the mouse over the colored area. This way, the exact percentage of hovered group is shown. In addition, hovering the mouse above the text representing groups (e.g. Female, Male, Bad Characters, etc) also shows the percentage of that group.

Besides the hovering option, the parsets visualization offers the possibility for the user to manually determine the appearance of the visualization by changing the order of groups.

That is done by clicking on a specific group and moving the group in the left or right direction. This provides a bit of flexibility to the users, in case they are interested in finding out specific information.

4.2.2 Bubbles visualization

For the bubbles visualization, we were inspired by bubbles visualization implemented using D3.js in general. The visualization could be seen on this figure 3.5.

Regarding the interaction with bubbles, hovering the mouse above the bubbles makes the information box pop up. This information box contains more information about how many characters are grouped in that bubble, based on their gender, hair color, etc.

4.2.3 Bipartite graph

For the bipartite visualization, besides **d3.js** we used **viz.js**.

Regarding the interaction with the bipartite graph, the possibility of hovering is provided. For example, hovering the mouse above the origin type, user can see the amount of characters (in percentage) that have certain align type. If we hover above human, we can see how many humans are good, bad, or neutral. For this example, see figure 4.2. The same applies to the other side of the bipartite graph, see figure 4.3.
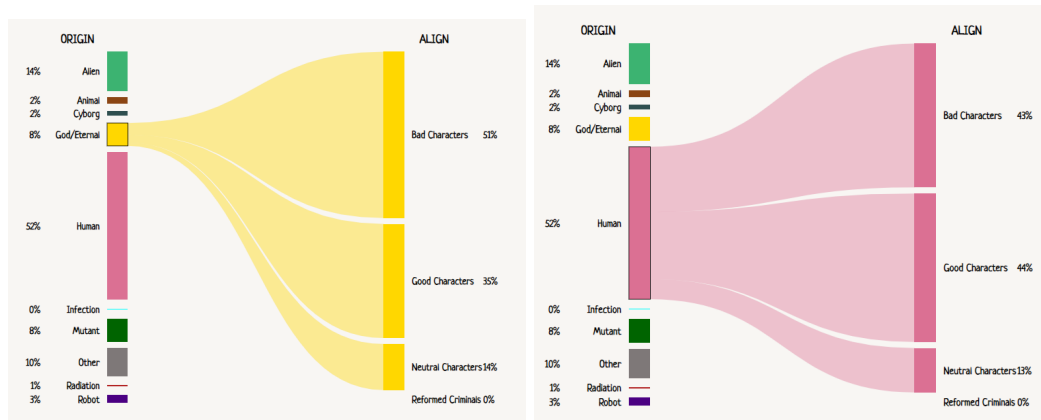

Figure 4.2: Clicking on the left side of the bipartite plot - *ORIGIN* (on the left screenshot the God is selected, on the right one the Human is selected).

4.2.4 Chord

The chord visualization was inspired by the **d3-chord** example visualization, even if in the end we ended up extending the code to fully suit our needs.

Regarding the interaction with this chord visualization, the user has several possibilities to play with this visualization. The first one is making the choice of what will be visualized: team, origin or gender. Clicking the button determines the coloring of the chord, more precisely, the user will be able to distinguish different genders based on the color. Meaning that the characters of the same gender will be colored with the same color. In addition, clicking the buttons: team, origin and gender, will also rearrange the chord graph in a way that the characters that belong to the same group are next to each other. See figure 4.4
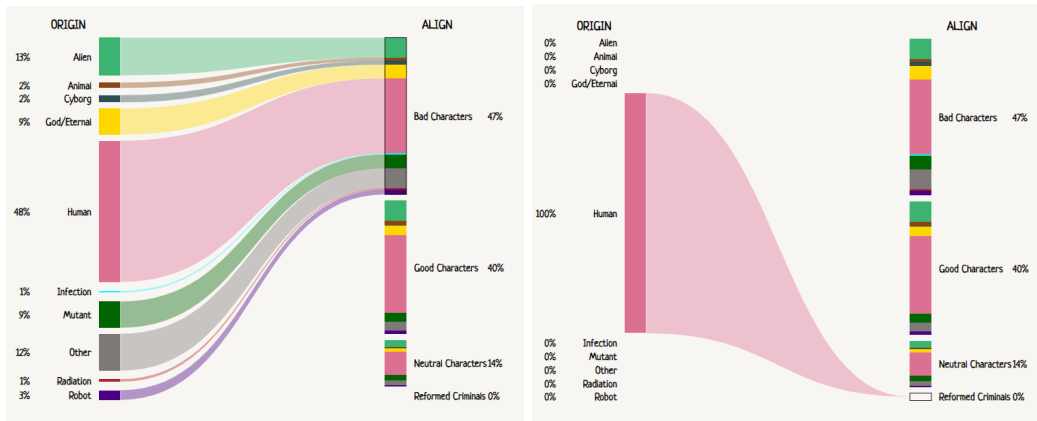
Figure 4.3: Clicking on the right side of the bipartite plot - *ALIGN* (on the left screen shot the Bad Characters type is selected, on the right one the Reformed Criminals type is selected).

Another option of interaction will be to hover the mouse above the chord node. This will temporarily show information box on the left side of the screen. The information provided in the information box of the clicked character includes the character name, picture, gender, etc.

The last option gives the user the possibility to click on one of the character names and this way the information box will stay there until user clicks on the same character name again, or clicks on another one. See figure 4.5
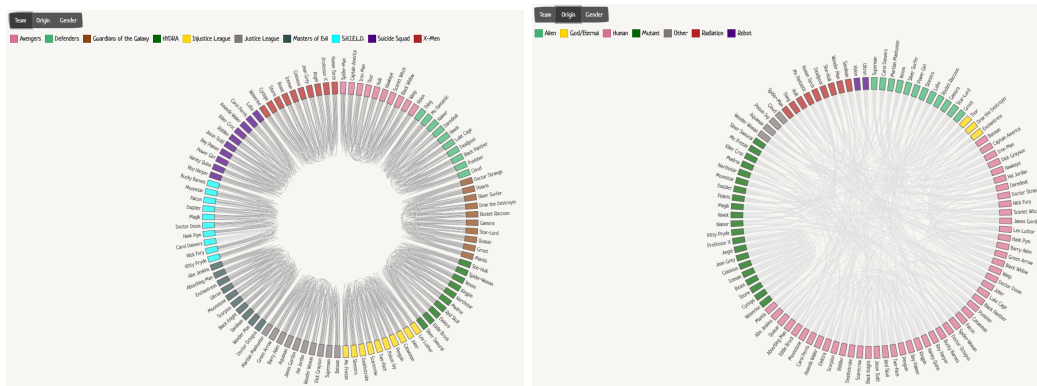

Figure 4.4: Chord with selected *TEAM* and *ORIGIN* button respectively

4.2.5 Timeline of diversity

For the final version of the timeline we used the library **d3Kit-timeline**. The library implements a simple timeline that labels do not overlap and it's built on top of *D3*, *d3Kit* and *Labella.js*.

When viewing the page, only the timeline is shown as it can be seen from left figure 4.6. The labels on the timeline show the year a character was introduced and their name.
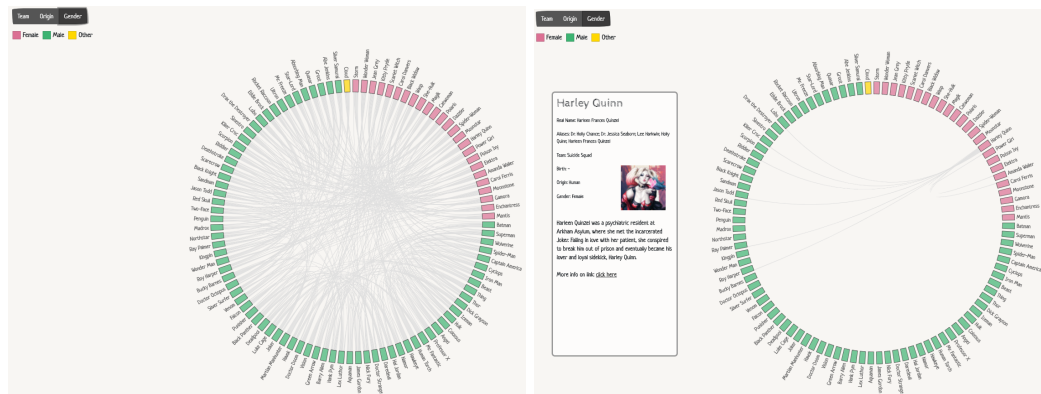
Figure 4.5: Left: Chord with selected *GENDER* button; Right: Clicking on a node of the chord

When clicking on a label, two types on information appear as it can be seen form the figure on the right side 4.6. On the left there is the historical background and the sociopolitical events that were taking place the decade when each character was introduced.
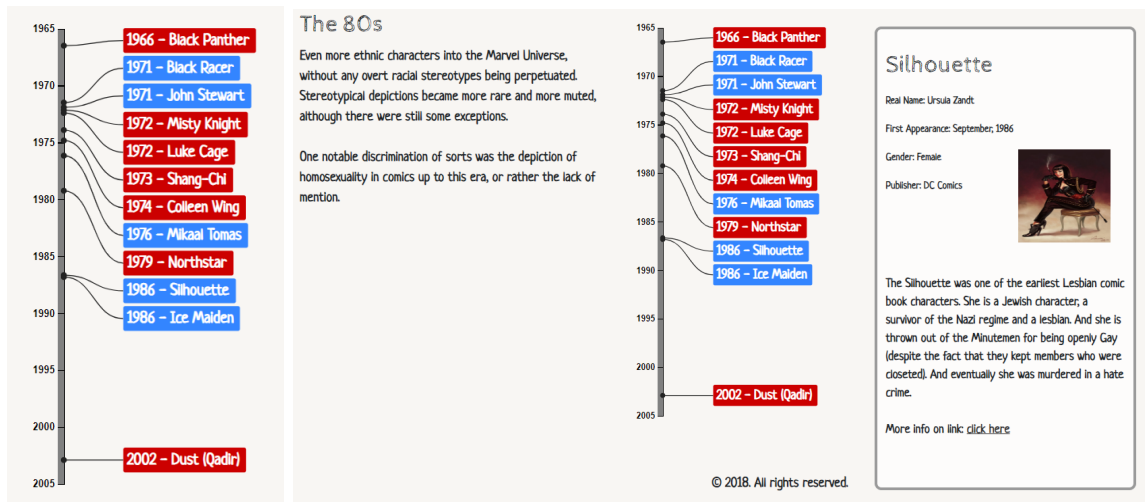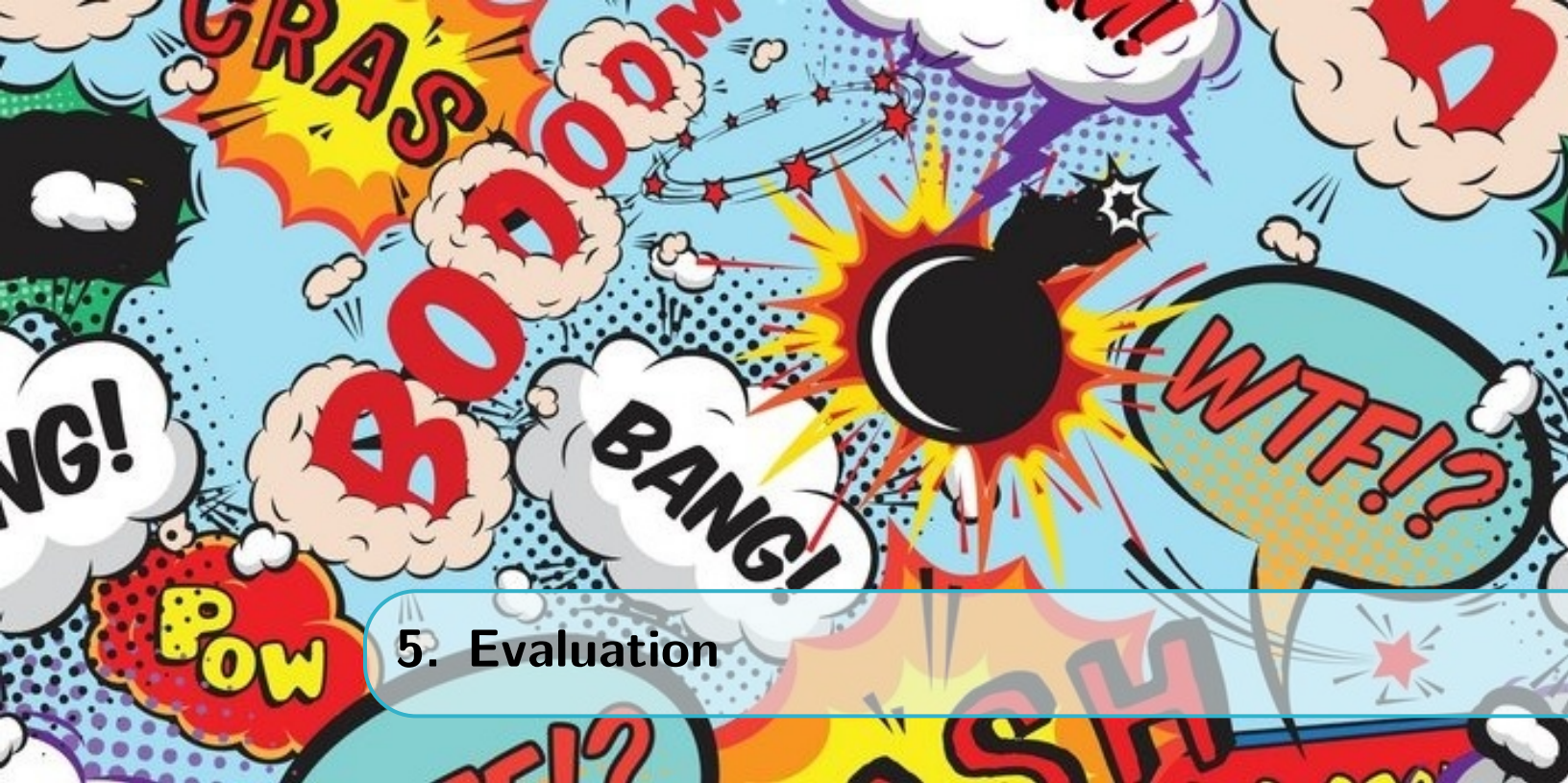
Figure 4.6: Left: The timeline view when navigating to the page; Right: The timeline when the user hovers on "Silhouette"'s label. On the left of the timeline there is the description of the 80s in the US and on the right there is information about "Silhouette"

5. Evaluation

5.1 Lessons learned

Before this project, other courses at EPFL only gave us very dull and mathematical tools to use when trying to explore our data. Without reducing the importance of all of those methods, this is the first course that gave us another, creative new look into the world of data and the way it can be used to represent concepts to non-technical audiences. Even though we spent countless hours coding, searching for inspiration and intensely discussing our ideas, we enjoyed every minute of working on this extraordinary project.

Of course, we learned, or got reminded of many other important aspects of human impressions by debating about our visualization with different possible users. They also reminded us to always think of the principles of the user-friendly design, and the fact that we should make our programs as easy to understand and interactive as possible. The whole process of creating this data story was a very unique experience that inspired us, and we are hoping will influence the perception of others as much as it did ours.

Figure 5.1: Our new motto

5.2 Further improvements

In the future, we would firstly like to enrich our visualization with even more data. Finding reliable datasets on the topic of comics was not an easy task, as they are mostly made by the volunteers from fan communities. That is why some of the data in the datasets was not consistent, some of the entries were missing, and even though we made some generalizations in order to keep as many characters as possible, some of them were dropped during the preprocessing phase. It would be nice if we could find a way to scrape the missing data from the official websites of the publishers, or even from online issues directly. We believe this would make the quality of our analysis even better, and this dataset could then be used by other programmers interested to explore the comic book characters.

One other addition that we would like to create is a new visualization that allows the exploration of not only the diversity and relations between the most popular 100 characters, like in our chord diagram, but all of the characters that appear in the dataset. It would also be nice to add a search option to this visualization, in order to facilitate the finding of specific characters or teams and their properties to the user.

Additionally, one part of the dataset that we did not manage to explore for this 'release version', and would love to in the future are the powers different characters have. Diversity in terms of powers, and different strengths and weaknesses given to different types of characters could probably also provide us with invaluable insight into roles given to characters.

6. Peer assessment

In general, we had a very good distribution of tasks. Also, during the development process we were helping each other when help was needed. We were also able to share the experience that we gained from the other tasks, and apply it to issues that arose. We were meeting every week and setting up new goals for the following week. In case of disagreements or misunderstandings, we overcame them as a team, with discussing them over and considering the pros and cons of every decision. The communication was good and we learned a lot about the process necessary to create a data story both as individuals and as a team.

6.1   Jelena Banjac

- *Preparation – were they prepared during team meetings?* Yes.

- *Contribution – did they contribute productively to the team discussion and work?* Yes.

- *Respect for others' ideas – did they encourage others to contribute their ideas?* Yes.

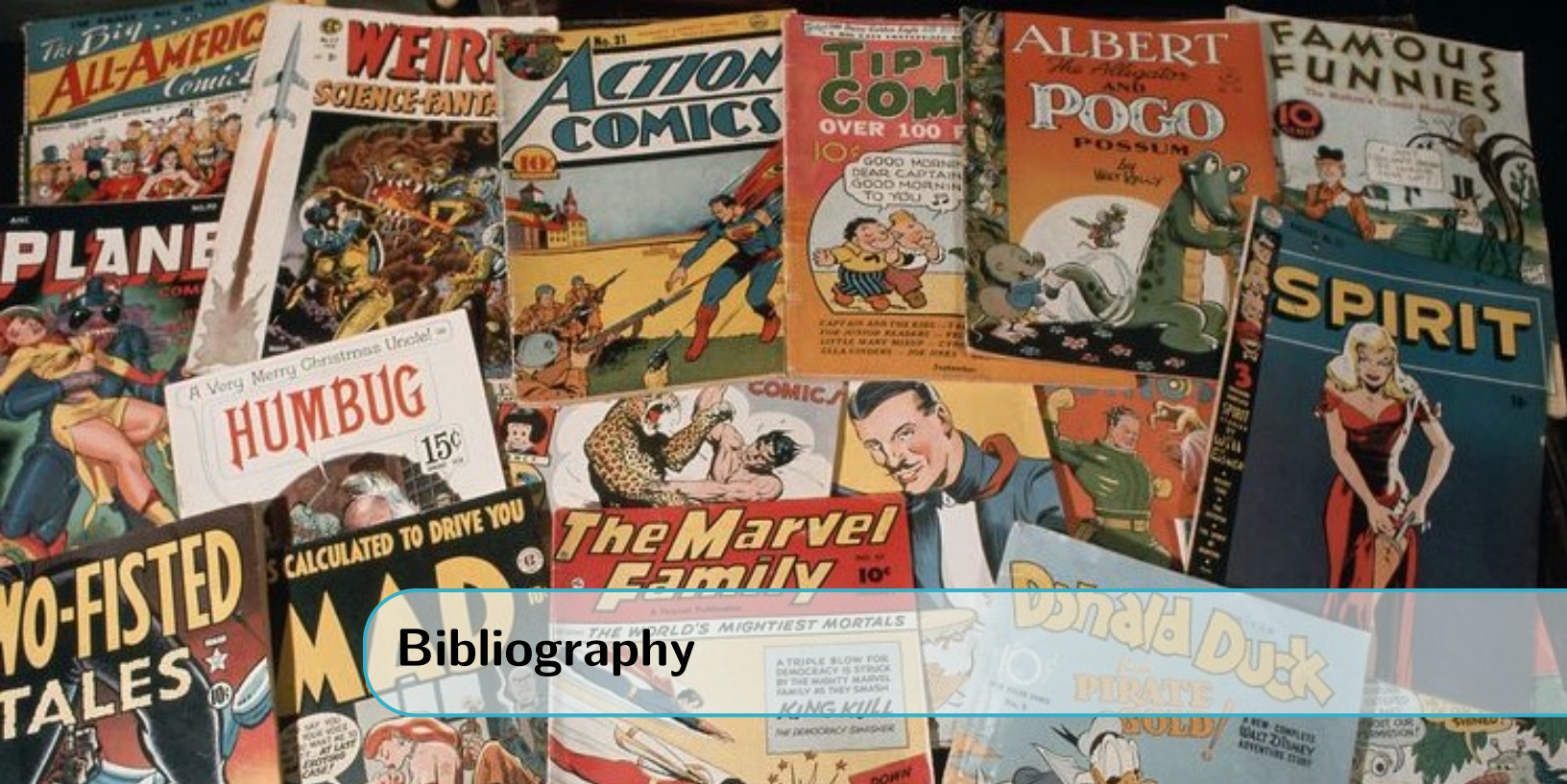- *Flexibility – were they flexible when disagreements occurred?* Yes.

6.2   Sofia Kypraiou

- *Preparation – were they prepared during team meetings?* Yes.

- *Contribution – did they contribute productively to the team discussion and work?* Yes.

- *Respect for others' ideas – did they encourage others to contribute their ideas?* Yes.

- *Flexibility – were they flexible when disagreements occurred?* Yes.

6.3 Jelena Malić

- *Preparation – were they prepared during team meetings?* Yes.

- *Contribution – did they contribute productively to the team discussion and work?* Yes.

- *Respect for others' ideas – did they encourage others to contribute their ideas?* Yes.

- *Flexibility – were they flexible when disagreements occurred?* Yes.

Bibliography

[1] W. Hickey, "Comic books are still made by men, for men and about men." `https://fivethirtyeight.com/features/women-in-comic-books/`.

[2] A. Shendruk, "Analyzing the gender representation of 34,476 comic book characters." `https://pudding.cool/2017/07/comics/`.

[3] M. Database, "Race, orientation, and gender in marvel comics." `http://marvel.wikia.com/wiki/Race,_Orientation,_and_Gender_in_Marvel_Comics`.

[4] E. Lewis, "Minorities in comics part 1 of 7: 10 comic characters with disabilities." `https://www.denofgeek.com/us/books-comics/minorities-in-comics-part-1-of-7-10-comic-characters-with-disabilities/45334/minorities-in-comics-part-1-of-7-10-comic-characters-with-disabilities`.

[5] E. Lewis, "Minorities in comics part 2 of 7: 10 people of color/hispanics in comics,"

[6] E. Lewis, "10 lgbtq comic characters." `https://www.denofgeek.com/us/books-comics/lgbtq-comic-characters/55874/10-lgbtq-comic-characters`.

[7] C. Vine, "Comic vine website." `https://comicvine.gamespot.com/api`.

[8] F. T. Eight, "Five thirty eight." `https://fivethirtyeight.com/`.

[9] R. Kosara, "An argument structure for data stories," 2017.

[10] K. Benzi, "Materials from the course of data visualization at epfl 2018,"