Deaths in Game of Thrones

| Pedro Elias  89519  Lisbon, Portugal | Raquel Santos  89533  Lisbon, Portugal  [raquel.bento.santos@tecnico.ulisboa.pt](mailto:raquel.bento.santos@tecnico.ulisboa.pt) | Tiago Afonso  89436  Lisbon, Portugal |
| --- | --- | --- |

# ABSTRACT

This report describes the project develop for the Information Visualization course at Instituto Superior Técnico. The project consists of a visualization about the deaths of the *Game of Thrones* TV show and their relation with the book collection *A Song of Ice and Fire*. It is developed using HTML, CSS, JS and d3.js.

The visualization allows to analyze all the deaths along the different episodes and analyze the different conditions where they occur such as killing method, allegiance of the involved or location of the death.

Abstracts should be about 150 words and are required.

## Author Keywords

Information Visualization; InfoVis; Game of Thrones; A Song of Ice and Fire; Deaths; Kills.

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; See<http://acm.org/about/class/1998> for the full list of ACM classifiers. This section is required.

Human-centered computing: Visualization: Visualization application domains: Information visualization

# INTRODUCTION

The domain tackled are the deaths in the TV show *Game of Thrones* that occurred on-screen. *Game of Thrones* is one of the most famous shows of all time, and it is well known for frequently killing characters, either being part of the main cast, or background characters.

The popularity of the show makes it an interesting subject for data analysis and the creation of visualizations which allow the user to extract more information from each death.

Some questions that could be answered with visualizations regarding this problem domain are:

* How was the evolution of the number of kills for each character through the seasons?
* How was the evolution of the killing method for each character through the seasons?
* Which houses suffered the most casualties / have the most conflicts between each other?
* Which episode and season has the most deaths?
* Which houses have the most conflicts between each other?
* Where did most kills take place?
* Which species had the most deaths?
* Which gender has the most kills and deaths?
* There is a preferred killing method according to the killer’s house?
* Is the location of the deaths a factor when choosing the killing method?
* What percentage of TV show deaths were based on the books?
* Did the deaths in the TV show and the books occur by the same chronological order?

Currently there are no known tools to automatically collect and process data for deaths of characters in a TV Show. All the data needs to be imputed and verified manually.

Recording and analyzing the deaths in a TV Show whilst comparing them to the events of the original work (the books) not only allows for a better understanding of the evolution of characters, their locations throughout time, the interactions they had with other characters, but also the impact of battles or other important events on characters and houses.

Comparing the deaths in the TV show and in the books also helps seeing how much of the show was adapted from the books and where it diverges. Some characters only exist in the show, others only exist in the book, so it is interesting to analyze the difference in their fates.

# Related Work

There are many analyses on *Game of* *Thrones* available online, due to the popularity of the books and the show. While all focus on the *Game of Thrones* domain, some use completely different data (for example, the conversations between characters in the book), or try to answer different questions:

* Interactive visualization of *Game of Throne* deaths [1]. It uses the same initial dataset we used to analyze the show deaths. However, the visualization is very static and is too focused on a single character at a time, which does not allow the user to answer all the previous questions. Some idioms have scalability issues and the filters are insufficient for our needs.
* Narrative chart [2]. The domain of this visualization is the interactions between characters in the TV show and their locations throughout time. It also provides information on screen time. The interactions focused by the visualization were not used in our dataset since we mainly focus on deaths.
* The Chaos Ladder [3]. It focuses on characters and their number of appearances in the TV show. While it also shows their life status, it does not have any additional information on what happened in each death, so it does not answer our questions.
* Narrative structure of *A Song of Ice and Fire* [4] creates a fictional world with realistic measures of social complexity. This research paper thoroughly analyses the narrative and interactions between characters to measure the complexity of the fictional world with our own. Since we are mostly interested in the deaths and comparison between show and book, it does not answer our questions.
* Illustrated *Game of Thrones Deaths* [5] uses a similar dataset to analyze each death in the TV show. Even though it uses idioms such as a map and dot plots, it lacks the filters needed to answer our questions and its purpose is to guide the user to a predetermined analysis, lacking interactivity.

# The data

The data was obtained from two different datasets being the first one about the deaths in the TV show [6] and the second one referring to the deaths that occurred in the book collection [7].

Initially, we kept all the data from the first dataset and joined the columns of the second with an outer merge. Then, we removed some lines to keep only the characters that have died in the TV show and appeared in the first dataset.

We removed the repeated columns and introduced two derived measures to analyze if the character appeared in the books and to differentiate the deaths of the animals.

As part of the cleaning process, we removed repeated characters that appeared with different names, corrected wrong values and introduced the missing values based on research on some wikis dedicated to Game of Thrones [8] and A Song of Ice and Fire [9]. Also, it was added sentinel values for the missing values that did not apply.

During the implementation, some changes were made to adjust the data to the problems that start appearing and that were not predicted at first.

Considering that we are treating the killer and the killed characters in the same way, we introduced columns with the characteristics of the killer. All this information was manually introduced based on our knowledge from the TV show and the previously mentioned wikis.

Additionally, we transformed the columns related to the death in the book into the estimated death, converting the chapters into episodes according to the narrative [10].

To simplify the representation of the deaths of the animals, we introduced the allegiance “Animal” and changed this value for all.

To keep the legibility of the chord diagram, considering that we had 278 characters, we had to introduce two new columns with the names that would be represented. The characters with higher number of kills are represented by their name and the remaining are aggregated by their allegiance.

To avoid having redundance in the killing methods, we aggregated some similar methods and removed the additional explanations.

To represent the points in the map, it was introduced two columns with the coordinates of each location. These coordinates were estimated considering the annotated maps that we found online.

Finally, we removed the columns related to the introduction of the character in the book and the death year considering that it did not follow the theme of the visualization. Also, we ended up deleting the two derived measures initially introduced, considering that we could represent the same data without them after the modifications.

Besides the reunited data, we also wanted to find the episode of introduction for each character but, while it was possible to manually find for the main characters, it would be impossible for a large amount of characters. This would be used to compare the presence of the characters in the TV show and in the books considering that we also had the book chapter of introduction. To compare both narratives, we tried to find additionally the killer of each character in the book, but we couldn’t find it.

# Visualization

In this section it will be described all the idioms implemented in the visualization and it will be explained all the choices that led to the final result.

## Overall Description

The visualization is showed in Figure 1. It is composed by the “Deaths/Kills” button and the filter menu on top. The button is used to choose whether kills or deaths are shown, in cases where there is a difference between them (a specific house’s kills is different from their deaths, but a season has the same amount of kills and deaths). The filters allow to aggregate the deaths or kills by season, books, character, house, killing method, gender, nobility or animals. There is also the indication of the current filter and a button to remove it on the right side to simplify the perception of the current state and to ease the change to the default overall view. Besides this filtering, by clicking in each idiom, it will apply the correspondent filter to the

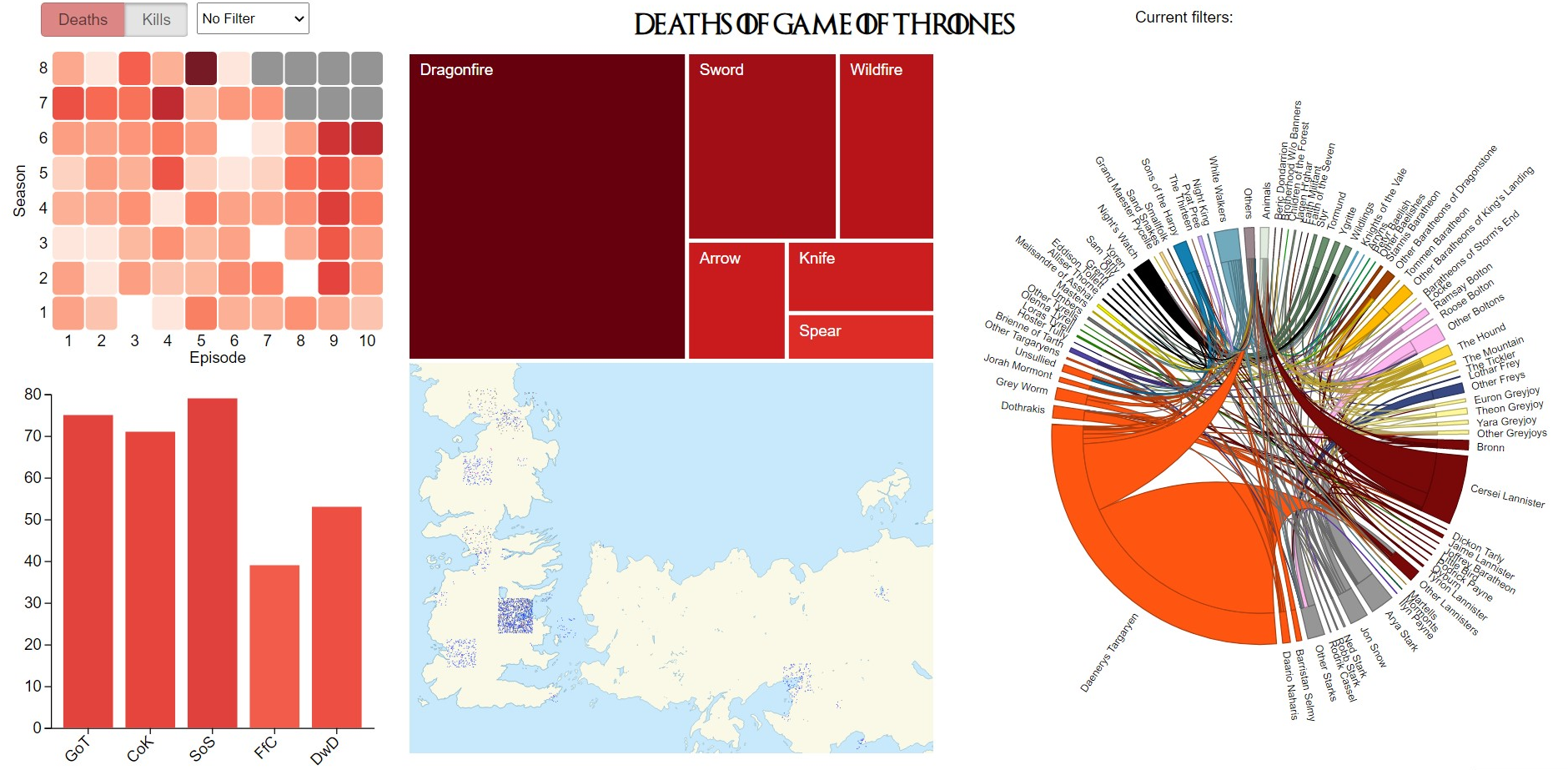


Figure 1. Overview of the visualization.

remaining. Clicking on a selected option will remove the applied filter for all idioms.

The visualization is divided in 5 idioms, those being:

* A **heatmap** representing the number of deaths or kills by each episode of the different seasons, being the episode on the x-axis and the seasons on the y-axis. For a character that died in the books, it will appear a red boarder on the square correspondent to the estimated episode of death in the book. By clicking on one of the squares it will apply the filter to the selected episode and by hovering it will show the exact number of deaths or kills.
* A **treemap** representing the six most frequent killing methods where the size of each square corresponds o the quantity of occurrences. The click and hover are similar to the described for the heatmap, applied for the selected killing method.
* A **chord diagram** representing the relations between the killer and the deceased character and their allegiances. Each node represents a character or an aggregation of characters per house, for legibility purposes as mentioned in the previous section. By clicking on a border or a name it will only show the arcs for that character or group, fading the remaining and apply the filter for its allegiance. By hovering a border or a name, it will show the total percentage of kills of the character or group and on the border will fade the arcs correspondent to the remaining characters. By hovering an arc it will show the percentage of kills between the two characters or groups represented by it.
* A **bar chart** representing the amount of characters that appear in each one of the five books where each bar corresponds to a book. The click and hover are similar to the described for the heatmap and treemap, applied for the selected book.
* A **map** where each pin represents the location of each death. The shape of the pin represents the nobility of the killed character being a square for noble characters and a circle for the remaining. The gender is represented by color, being pink for female characters, blue for male characters and grey for the unspecified. By clicking on each point, it will apply the filter to its location. By hovering it will show the names of the killer and the deceased, the location of the death, the method used and the season and episode where it occurred.

## Rationale

A

Why did you think your techniques would work? What visual encodings did you use and why (and why not others)? What alternatives did you consider, even if they turned out not to work? Especially, discuss how you managed the complexity of real data, and matters of scalability. Also, include in your discussion the evolution of the prototype, from the initial sketches to the last version highlighting what you learned from version to version and how that influenced your design.

## Demonstrate the Potential

A

Why did you think your techniques would work? What visual encodings did you use and why (and why not others)? What alternatives did you consider, even if they turned out not to work? Especially, discuss how you managed the complexity of real data, and matters of scalability. Also, include in your discussion the evolution of the prototype, from the initial sketches to the last version highlighting what you learned from version to version and how that influenced your design.

# Implementation details

A

# conclusion & future work

A

# ACKNOWLEDGMENTS

Sample text: We thank all the volunteers, and all publications support and staff, who wrote and provided helpful comments on previous versions of this document. Authors 1, 2, and 3 gratefully acknowledge the grant from NSF (#1234-2012-ABC). This is just an example.

# REFERENCES

1. Game of Thrones Interactive Viz by David Murphy. Last accessed in 16/12/2020 at <https://datasaurus-rex.com/gallery/gotviz-mkiii>
2. Game of Thrones Narrative Chart by Jeffrey Lancaster. Last accessed in 16/12/2020 at <https://jeffreylancaster.github.io/game-of-thrones/map/>
3. The Chaos Ladder by Patrick Gillespie. Last accessed in 16/12/2020 at <http://patorjk.com/vis/chaos-ladder/>
4. Narrative structure of *A Song of Ice and Fire* creates a fictional world with realistic measures of social complexity. Thomas Gessey-Jones, Colm Connaughton, Robin Dunbar, Ralph Kenna, Pádraig MacCarron, Cathal O’Conchobhair, Joseph Yose. Last accessed in 16/12/2020 at https://www.pnas.org/content/117/46/28582
5. Illustrated *Game of Thrones* Deaths by Shelly Tan. Last accessed in 16/12/2020 at https://www.washingtonpost.com/graphics/entertainment/game-of-thrones/
6. Deaths in the TV show *Game of Thrones*. Retrieved September 26, 2020 from <https://data.world/makeovermonday/2019w27/workspace/file?filename=game-of-thones-deaths.xlsx>
7. Deaths in the book collection *A Song of Ice and Fire*. Retrieved September 26, 2020 from <https://www.kaggle.com/mylesoneill/game-of-thrones?select=character-deaths.csv>
8. *Game of Thrones* wiki. Retrieved October 25, 2020 from <https://gameofthrones.fandom.com/>
9. *A Song of Ice and Fire* wiki. Retrieved October 25, 2020 from <https://awoiaf.westeros.org/>
10. *Game of Thrones* Episode Chapter Coverage. Retrieved October 25, 2020 from <https://joeltronics.github.io/got-book-show/bookshow.html>
11. ~~Ronald E. Anderson. 1992. Social impacts of computing: Codes of professional ethics.~~ *~~Soc Sci Comput Rev~~* ~~10, 2: 453-469.~~

# Accessibility

The Executive Council of SIGCHI has committed to making SIGCHI conferences more inclusive for researchers, practitioners, and educators with disabilities. As a part of this goal, the all authors are asked to work on improving the accessibility of their submissions. Specifically, we encourage authors to carry out the following five steps:

1. Add alternative text to all figures
2. Mark table headings
3. Generate a tagged PDF
4. Verify the default language
5. Set the tab order to “Use Document Structure”

For more information and links to instructions and resources, please see: <http://chi2016.acm.org/accessibility>.

# Page Numbering, Headers, and Footers

Your final submission should not contain footer or header information at the top or bottom of each page. Specifically, your final submission should not include page numbers. Initial submissions may include page numbers, but these *must* be removed for camera-ready. Page numbers will be added to the PDF when the proceedings are assembled.

On pages beyond the first, start at the top of the page and continue in double-column format. The two columns on the last page should be of approximately equal length.



Figure 1. Use high-resolution images, 300+ dpi, legible if printed in color or black-and-white. Number all figures and include captions below, using Insert, Caption.

Use a numbered list of references at the end of the article, ordered alphabetically by last name of first author, and referenced by numbers in brackets [1,3,4].

| Objects | **Caption – pre-2002** | **Caption – 2003 and afterwards** |
| --- | --- | --- |
| Tables | Above | Below |
| Figures | Below | Below |

Table 1. Table captions should be placed below the table. We recommend table lines be 1 point, 25% black. Minimize use of unnecessary table lines.

# Producing and testing PDF files

We recommend that you produce a PDF version of your submission well before the final deadline. Your PDF file must be ACM DL Compliant. The requirements for an ACM Compliant PDF are available at:

<http://sheridanprinting.com/typedept/ACM-distilling-settings.htm>

## Subsections

Headings of subsections should be in Arial 9-point bold with initial letters capitalized (Heading 2 style). For sub-sections and sub-subsections, a word like *the* or *of* is not capitalized unless it is the first word of the heading.

### Sub-subsections

Headings for sub-subsections should be in Arial 9-point italic with initial letters capitalized (Heading 3 style).