

An Eastern Fantasy World

-Data Visualization of *The Twelve Kingdoms*

By: Yutong Han
Advisor: Alberto Cairo



My Work

Preview the project

Concept Statement

Why I want to do data visualization of *The Twelve Kingdoms*

Precedents

Inspiration from Nadieh, Deb's works

Develop Process

Challenges I met and how to overcome them

What's Next

Things that wait for me to refine

1

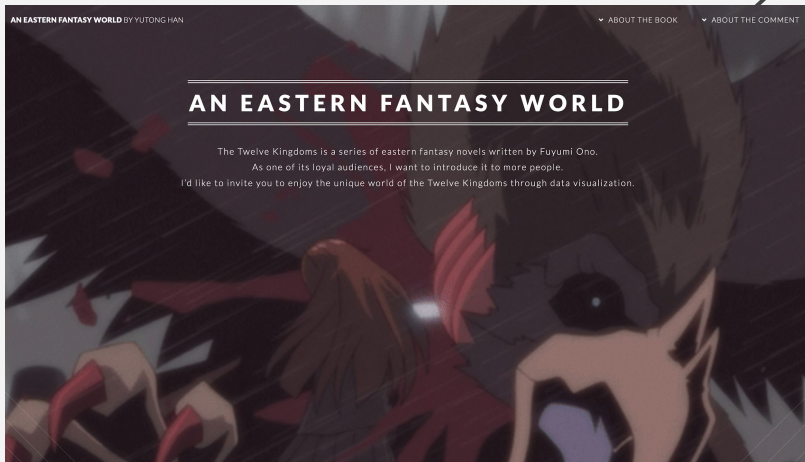
2

3

4

5

1



My Work

Preview the project

My Role

Designer, Developer

Timeframe

2020.02.01 - 2020.05.07

Link

[An Eastern Fantasy World](#)

Tool

Python, R, D3.js



2

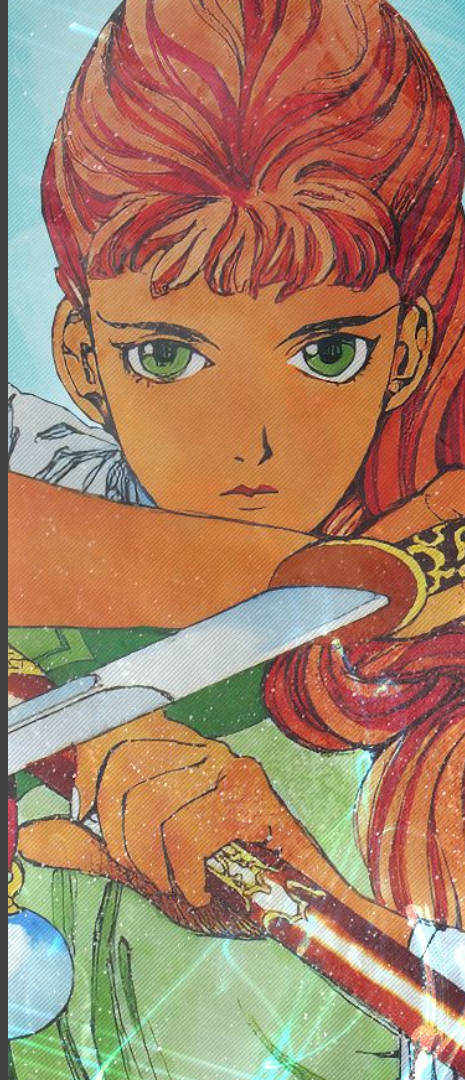
Concept Statement

Why I want to do data visualization of *The Twelve Kingdoms*

Concept Statement

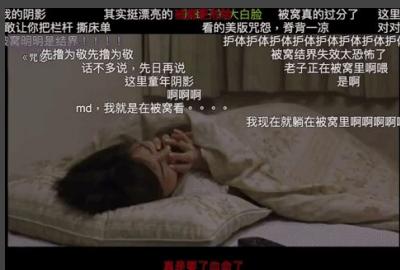
We are familiar with the western fantasy worlds, like *Harry Potter*, *Game of Thrones* and others. But do you know about the eastern fantasy world? Maybe you are an anime fan and you are familiar with *One Piece*, *Naruto* or something else. However, I bet you've never heard of ***The Twelve Kingdoms***.

The Twelve Kingdoms is a series of fantasy novels written by Fuyumi Ono, then made into an anime in 2002. The world of the Twelve Kingdoms is based on the coexistence of youma (monsters), and sages. This world is similar to ancient China, particularly during the Zhou dynasty, over 2000 years ago. It's a great chance to know more about **eastern philosophy, tradition and imagination**.



Concept Statement

In Bilibili (like Chinese YouTube), the anime *The Twelve Kingdoms* has over 106,000 bullet screen comments. They will accompany with you when you watch the video. Over 8000 bottom comments are left under the video.



As one of Fuyumi's big fans, I will analyze and visualize seven books, bullet screen comments and bottom comments. I'd like to invite you to enjoy this unique world through data visualization.

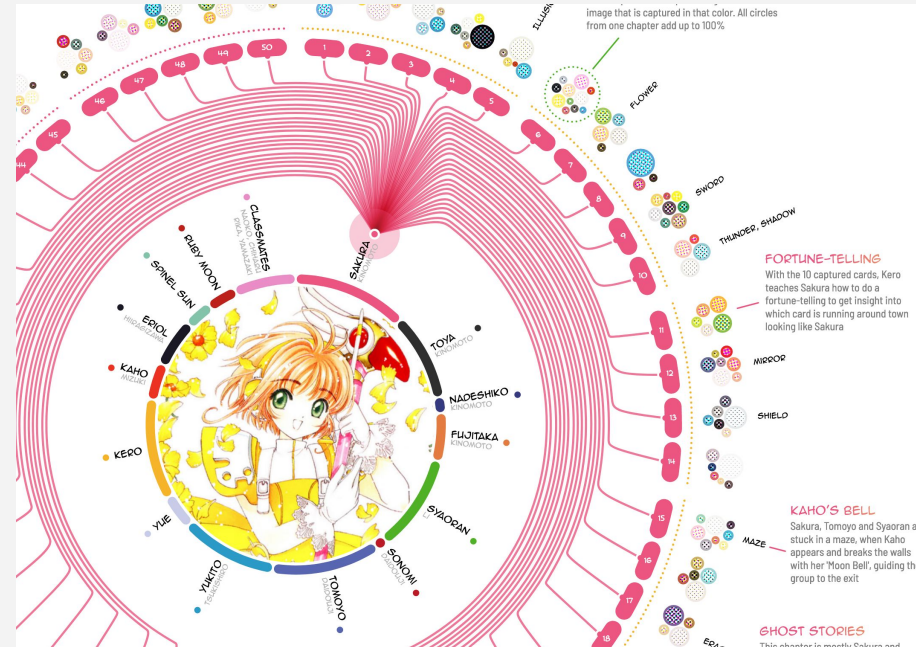


Abstract geometric lines in a dark gray color. One line starts from the top left, goes down to a point, then up to the top right. Another line starts from the bottom left, goes up to a point, then down to the bottom right. These lines create a large, irregular shape that frames the text on the right side of the slide.

3

Precedents

Inspiration from Nadieh, Deb's works



Anne Sexton: Poetry Between Pain

Deb Pang Davis

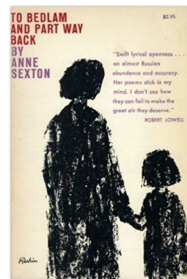
Text Analysis of Books

Sentiment Analysis of Words

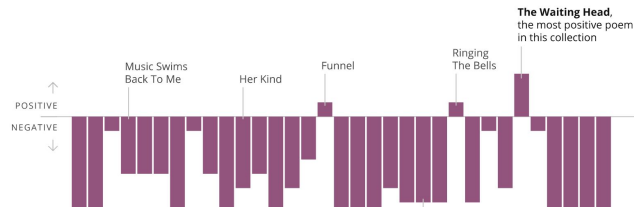
1960

A POET RISES FROM THE ASHES

To Bedlam and Part Way Back



To Bedlam and Part Way Back, published by Houghton Mifflin was Anne Sexton's first published collection. It is a collection of works that encapsulates her struggles with mental illness, postpartum depression, motherhood, being a housewife during the '50s and the grief of losing her beloved grandmother "Nana" in 1954. The poems in *Bedlam* exhibit the influence of W.D. Snodgrass who encouraged Anne to continue in the autobiographical style and find her voice^[1]. *Music Swims Back to Me* (1957) and *Double Image* are two of her signature poems in *Bedlam*. *The Division of Parts* is an elegy she wrote to her mother.

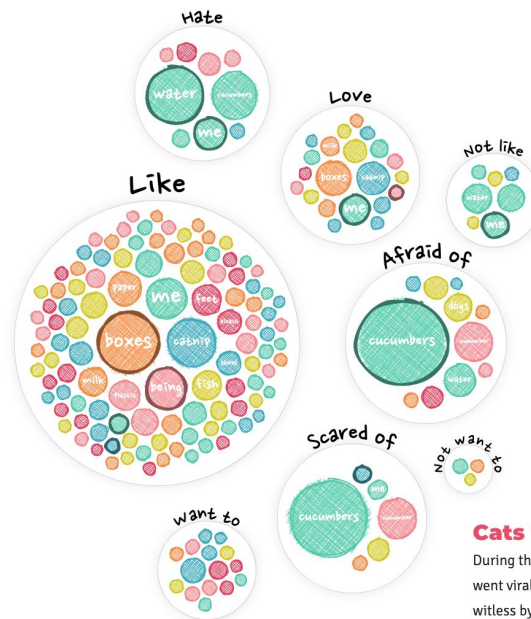


Why Do Cats & Dogs ...?

Nadieh Bremer

Forced D3 Bubble Charts

Nested Hierarchical JSON Structure





4

Develop Process

Challenges I met and how to overcome them

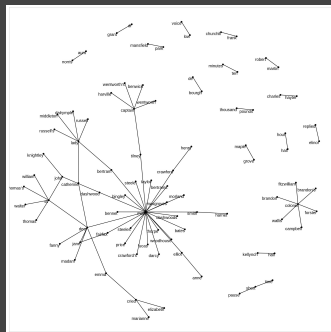
2.1 - 2.25

Scrape data of comments from Bilibili by Python



2.26 - 3.15

Clean and roughly visualize the data by R



Scrape and Clean the Data

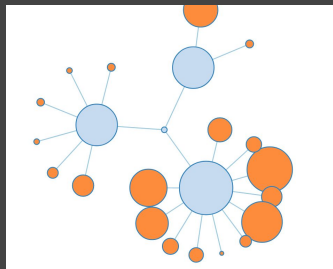
I met some troubles when I want to scrape mandarin from the website. Because it's different from English. And this website needs VPN to load comments.

Luckily, I found library pandas can help me scrape mandarin. And I asked my friend in China to run the code to solve the VPN problem.

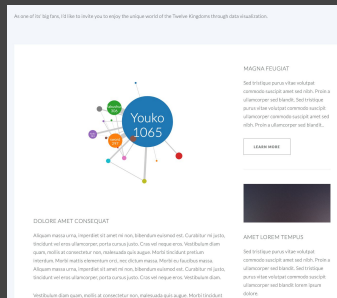
Another problem is I have no idea how to do n-gram in Python. Even though I watched some tutorials, it's still so hard for me to understand.

Fortunately, Deb recommended the book *Text Mining with R* to me. Do n-gram by R is much easier. In this way, I successfully found out the relationships between characters, and words.

3.10 - 3.25
Visualize the cleaned data by D3



3.16 - 4.10
Build the website, combine the charts and contents



Visualize the Data, Build the Website

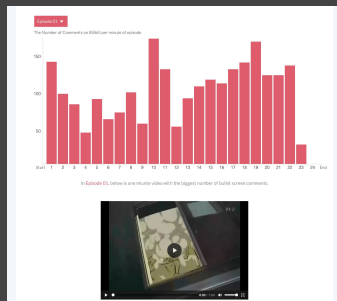
I tried to create bubbles by D3 after solving the n-gram problem. The challenge at this step is to create force among bubbles and change the stroke between them into different thickness. Because I created treemap charts by D3 last semester, I wasn't stuck with the json file for a long time.

I know my article will be very long, since seven books, bullet screen comments and bottom comments will be visualized and analyzed. So I need to organize my website very convenient for users to go through.

I add the top sticky navigation, supplementary instructions in the sidebar, tooltips to make them freely read.

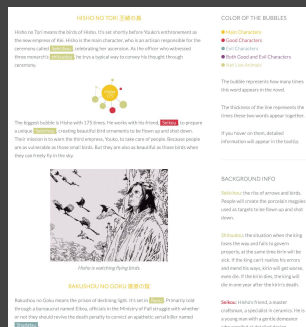
4.11 - 4.25

Add the button for users to jump among episodes.



4.16 - 5.6

Add texts and images in the context.



Add button for Episodes, Improve the Website

Because I want to know which moment is the climax part per episode, I combined the chart and the video of the minute with the highest number of comments. You can use the button to change episode to see the change.

For book 6 and 7, they are composed of several short stories. So I cut them into separate bubble charts.

And I keep refining the website according to Professor Cairo, Clay, Lenny's suggestions.

Final Work:

An Eastern Fantasy World

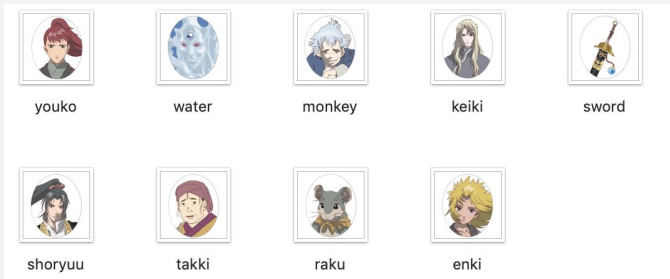
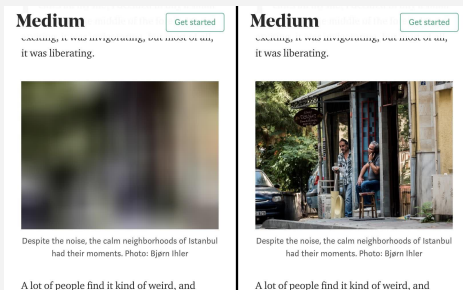
Abstract geometric lines in dark gray, forming a large, irregular shape that frames the text on the right side of the slide.

5

What's Next

Things that wait for me to refine

Next Step



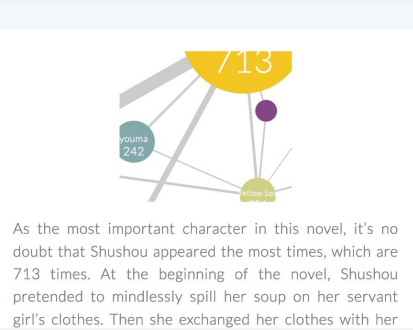
Page load time is too long.

I will try “Lazy Loading” to make the website more fluent.

The images in tooltip is not consistent.

I will redraw other characters in my style. Then I will change the images in the tooltip.

Next Step



The chart is not mobile friendly.

The content is mobile friendly, but the chart is not. I adjusted first two charts size on phone. But the result is not ideal. I will find a new way to solve it.

The auto translator is not good enough.

Because the comments are mandarin, I wanted to do mandarin n-gram in the beginning.

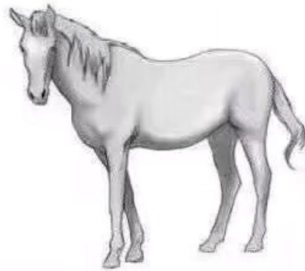
However, I ask the R experts, and they told me there is no way to do that. So I used Google auto translator. But the result is not great enough. I'm looking for a new way to analyze the data.



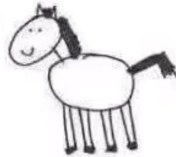
Beginning



Midterm



Final Work



Q & A

Thank You



Hometown: Guiyang
Email: yxh703@miami.edu