Topic: how aesthetics influence decision making

Title: A Cover Worth 1000 Words

Abstract:

Aesthetics tend to be something that tends to be taken for granted, even though the creation and appreciation of aesthetics is a fundamental part of what makes us human. Aesthetics play a huge role in our everyday lives, from affecting our mood, to even influencing the decisions that we make. Through this project, we'd like to create a page that is able to convey just how much aesthetics can influence decision making.

One way of showing this might be through book covers. There are datasets on the web with stats taken from goodreads.com (ex.

https://www.kaggle.com/jealousleopard/goodreadsbooks), that show what the most popular books are, and basic information about them. If we choose to go down this route, then our project's goal would be to visualize how much of a correlation there is between a book's cover (aesthetics) and its average rating (decision.

12/6

Team Check-in

- Whitespace
- Isabelle
 - Histogram and treemap transitions too abrupt
 - Violet text label too small
 - Add some data analysis narrative
 - Get some peer feedback
 - White color
 - Clean up code
 - Website hosting
- Lekshmi
 - Hook integration
 - Dynamic narrative
 - Percentage on tooltip
 - Clean up code

- Anastasia
 - Make links open in new tab
 - Add paragraph break intro
 - Spacing
 - Help Lekshmi with sizing
 - Outro sources
 - bold info in intro (make more prominent)
 - Script
 - webstroke to white text in innovative vis
 - description of ridgeline

12/5

Team Check-in

- Anastasia:
 - Still working with scrollytelling
 - Added to narrative
 - Changed interactive background
 - Also updated design
 - Working on getting percentages for color/genre

12/4 & 12/5

Team Meeting

Updated deadlines

- Update visualizations and fix any bugs EOD 12/5 (wed) ideally
- Storytelling: Methodology, Objects and genres, Aesthetics and Ratings,
 Psychology EOD 12/4 (wed)
- Website design/scrollytelling EOD 12/4 (wed)
- Fix sizing of vis/ tag selection EOD 12/4 (wed)
- Finish stack bar tooltip / dynamic titles/ book integration/ styling EOD 12/5
 (thu)
- Send Lekshmi material on psychology + genre EOD 12/5 (thu)
- Polish website/ spacing issues/ size issues EOD 12/5 (thu)

- Select auto complete/ finalize visualiza tions EOD 12/5 (thu)
- Toggling ridgeline between rating count and average rating EOD 12/5 (thu)
- Host website EOD 12/6 (fri)
- Draft of script EOD 12/6 (fri)
- Screencasting 1pm on 12/7 (sat)
- Update process book 1pm on 12/7 (sat)
- CHECK IN THURSDAY @ 10:30AM
- CHECK IN FRIDAY @ 3PM

Meeting with Laura

- Storytelling
- Design
- Interactivity in visualizations
- Color View
 - Show Color % instead of book counts
- Barchart
 - Consolidate item categories
 - Stacked bar graph with color frequency of book covers that have that object on the cover
 - Connect with treemap if select one color, highlight in other chart
 - In the middle show covers that have that object and color depending on what was selected in barchart or treemap
- Treemap
 - Could have a tooltip that shows some representative colors from category on hover
 - Mouse over a color 8 books stacked on side with that color as an example
- Histogram
 - Update with actual range on x-axis
 - Could do with dots instead of bars size of dots represents number of books
 - Compress vertically, margins come in
 - Transitions between genres

- Have in a translucent gray color compare to total distribution
- Make it more obvious that it's showing particular genre (title)
- See some of the titles of books corresponding to ratings take top 10 books and display on the side
- More clarifying tooltip number of books with average ___ rating
- Lab 9 responsiveness

_

12/1

Team Meeting

Deadlines

- Storytelling: Introduction, Background, Conclusion EOD 12/2 (mon)
- Book selection EOD 12/3 (tue)
- Innovative view/ tag selection all done by EOD 12/3 (tue)
- Treemap/bar graph/histogram EOD 12/3 (tue)
- Storytelling: Methodology, Objects and genres, Aesthetics and Ratings EOD 12/4 (wed)
- Storytelling: Color EOD 12/4 (wed)
- Website design/psycho storytelling/scrollytelling EOD 12/4 (wed)
 - at least begun by 12/3 (tue)
- Send Lekshmi material on psychology + genre EOD 12/5 (thu)
- CHECK IN WEDNESDAY @ NOON (design, check in about state of visualization, next steps, screencast)

Storytelling

- Maybe contrast ratio if we don't have enough material?
- Introduction
- Methodology
- Background: Psychology of color -
 - color and emotions
 - color in advertising/design
- Book cover analysis:
 - Color and genres

- Objects and genres
- Color, objects, and aesthetics affect sales/ratings
- Conclusion So what?
 - How we can be unconsciously affected by aesthetics/color REALIZING
 THIS
 - Being better consumers/more aware of these biases
 - How to design to grab attention

Action steps

Style/Storytelling

- Flesh out MUCH MORE storytelling
- I would also consider making the other elements of the webpage- the text boxes etc- a gray/black/white color scheme to give it a clean look since you are already using color a lot in the visualizations
- Scrolly-telling
- Conclusion
 - so what

Book selection

- add publication year to text description of selected book
- indicate visually which book was selected (e.g. with border, make image larger)
- Selecting a book at the beginning needs to connect with the rest of the narrative
 - Highlight year of selected book in innovative vis
 - create function of user's selection and use that to give the selection to innovative vis
 - Select genre of selected book in innvative vis
- Need to be able to select another book or refresh book selection
- Make books seem selectable
 - more subtle version of the effect on hover to make the selection process more fun.

- indicate visually which book was selected with a border, make the picture larger, something like that.
- More information about book after selecting it
 - add the publication year to the text description of the book that pops up once you select.

Visualizations

General

- Better directions for how to use/interpret visualizations
- Adding a border to the white color
 - the color for white seems a little dark/too similar to yellow
- Generally clearer and larger titles, names, labels and more explanatory text and axes
- Make elements of visualizations look clickable
 - change cursor if text is clickable

Innovative

- Innovative view should be entire width of page
- Innovative how size of circles connect to prevalence of color
- Color view directions making it clear that it's a stacked bar chart
- Area chart can be highlighted more
- The bar chart in the middle is cool- I think it could use some contextualization such as an x axis scale, some kind of indication of what it is measuring.
- Interaction between the middle and outer chart on hover/click would be really nice here.
- Have year text below the bars so there won't be any obstruction
- One way to address this is by adding some text captions pointing out anomalies, such as the one that has red as its #3 instead of blue.
- Alternatively, you can add a caption about how color doesn't really vary so much across genre. Either way, I think the viewer could use some help interpreting the results of this graph.

Bar graph/treemap

- Axes to stacked bar charts and description of units

- What are the different object types
- make this so it all fits on one page without the need to scroll- use a responsive svg size- look at lab 9 solutions for how to do this
- cut out the tail end of the graph- the categories with just one or two counts
- Include numbers/percentages in treemap
- Make bars bigger on stacked bar chart
- Bars shouldn't be orange
- treemap colors were used meaningfully in the bar chart such as coloring the bars stacked by the colors used in book covers featuring those subjects.
- view could also really use some interaction/tooltips. More information on hover/click, showing book covers or titles, something along those lines.

Tag selection

- Organize genre names better, alphabetical and no duplicates
 - Dropdown menu- alphabetize these or make it into an autocomplete search box.

11/25

Feedback

- Directions can stand out more and bigger
- Emphasizing that you need to scroll down
- Make transition of table and books appearing a bit slower (showed up before reading directions)
- Updating color view directions
- Adding axes to the stacked bar charts
- Emphasizing that color represents color
- Add option to reset
- Emphasis on being able to click middle circle to update stacked bar chart
- On hover would be nice if cursor changed on books
- Would be cool if you could choose different books
- Narrowing and organizing genre names, making them easy to read

- Define what object categories are what are packaged goods?
- Bars shouldn't be orange
- Include numbers in treemap
- Make bars bigger on stacked bar chart
- Adding a border to the white color
- Color view directions making it clear that it's a stacked bar chart

Presentations

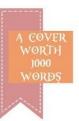
- Questions about the dataset/API we used
- Why are there more genres in the later graphs?
- Seems strange that yellow and packaged goods are so common
- When selecting a
- [Isabelle will finish]

11/24

Updates:

- Major changes to website layout, from feedback
- We have implemented all of our visualizations and user interactions, as well as styled our websites (e.g. defined color palette)





They say to never judge a book by its cover. But sometimes, we don't even consciously decide on which books get our attention in the first place. For example, ,

Chaose the book cover that most appeals to you - that makes you want to take the book home and read it right away - by clicking on it. Choose wively, since you won't be able to choose a different book without refreshing the page.



Excellent choice! The Chemist by Stephenie Meyer is a mystery book with a dominantly whiteshade cover! How does the color of this cover compare with other covers of the same genre? Well, we can explore that by using the visualization below.

COLOR VIEW

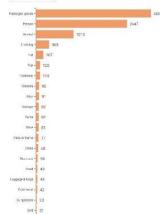
his visualization shows how colors are used in the covers of blooks of different genres. Each incit represents a different genre, and each bar on that circle represents... The graph in the enter shows how color has changed over time for. Hover over a bar for more detailed information, and disk on a bar to updated the timeline view.

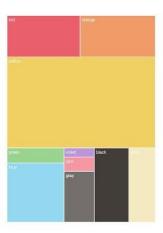


SYMBOL VIEW

Just as important as color, are the symbols that appear on book covers. In the visualization, the bar chart to the left shows the most popular symbols that are used on book covers, grouped by gerre. The tree resp to the right shows the main color of the books within that gerre. Feel free to use the dron-down born to filter by a specific covers.







"A Cover Worth 1000 Words" was created by Anastasia Lamothe '21, Lekshni Santhosh '21, and Isabelle Zheng '22 as the fina project for Harvard CS171 2019.

Sources go inste with truk tags

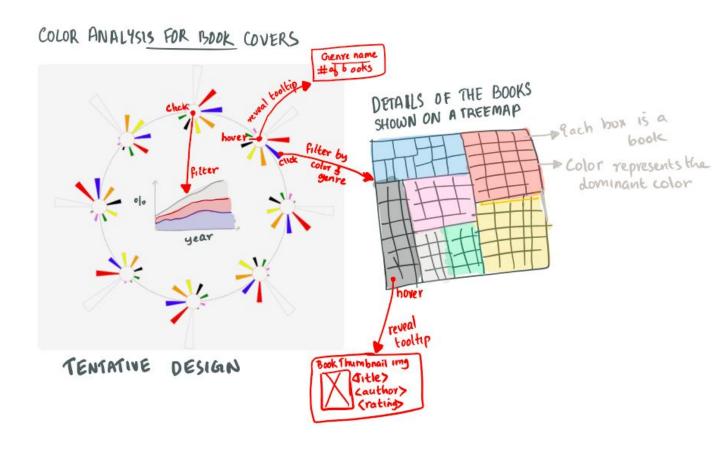
Future steps:

- treemap composed as a collage of book covers
- tooltip, subcategories, and zoom of treemap, as well as narrowing down genres in select box
- book covers in treemap
- innovated vis filter by using the colored circle of each genre; displaying extra data on hover of stacked bar chart
- changing opacity on hover in all visualizations to make object being hovered over stand out
- do data testing to make sure that all the data makes sense and that the code is implementing the data correctly, and that dominant color values are accurate
- add some more storytelling

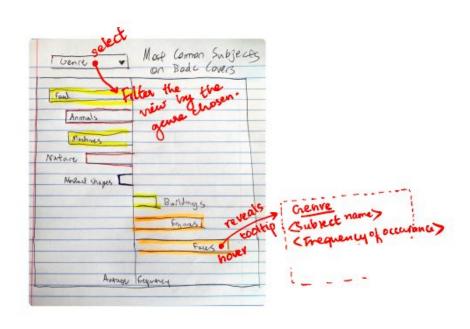
11/17

Updates to the interaction storyboard for the Color view

- We are using popular tags in Goodreads as a proxy for genre
- We assume that the dominant color identification by the Google Vision API is at least 90% accurate. Instead of visualizing all possible shades of color extracted, we are categorizing the color to the colors in the color spectrum alongside white and black. We convert the RGB value to HSV and use the HSV values to determine the broader color. This mapping has to be tested.
- We have two design ideas for the view
 - Encode circles as color and the spokes as genre
 - Replace radial plot with miniature barplots



Interaction storyboard for the Objects Bar plot



Meeting 3 (11/13):

Client/audience:

- o Audience is the general public; assume that viewers are not colorblind
- Could also be used by marketing/cover designers to figure out which features they should implement

• Specific goals:

- To provide a visualization tool that allows consumers to better understand the subconscious design details the influence how they purchases books
- To show trends and patterns of book covers, and to hypothesize what might cause those trends/ patterns

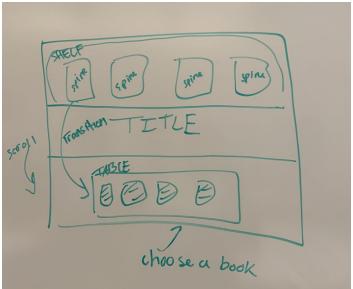
• Interesting questions:

- Are there certain colors that are more common for certain genres of books?
- Are there certain types of subjects that are more common for certain genres of books?
- Do books with covers that have certain characteristics (certain colors, contrast, etc.) get better reviews?
- Has the distribution of colors on book covers changed over time?
- Do different countries tend to have different distributions of book cover colors among the best selling books?
- o In general, what tend to be the most common features of book covers?
- Do certain genres tend to get better reviews? If so, can book cover feature distribution help to explain that?
- What are possible explanations for the distribution of book cover features?
- What are the most common objects shown on book covers across all genres?
- If there are certain types of objects that are more popular than others,
 then how do specific elements of those objects vary over time (i.e. how

do features apply to symbols themselves as opposed to covers as a whole)?

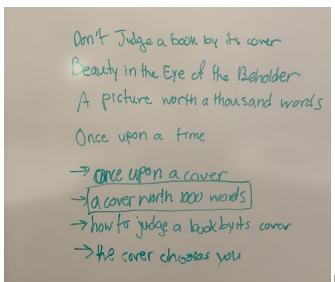
- https://www.kaggle.com/zygmunt/goodbooks-10k/data
- https://cloud.google.com/vision/docs/detecting-properties
- Need to find genre metadata as well
 - Possibly need larger images?
- Dominant vs. secondary colors, etc.
 - can use dominary color in some visualizations and then on click/
 mouse-over have a tooltip/ info box appear with book cover itself
 - o use this for any tooltips used
- Custom visualization probably should not use for quantitative data other than color, could also work for luminance
 - o Could also be non-radial plots might be more readable
- Circle sizes could encode something otherwise, get rid of the numbers (too many things encoded)
- Do countries differ— any dominant colors above average frequency?
- number of total ratings vs. average rating → aesthetics probably less related to ratings because ratings are typically provided by people who read the book
 - Popularity is more important than rating
 - What colors sell more? Which books are coming off the shelf?
 - le. are red books actually better rated or just more likely to sell?
 - disparity between what people buy, and books that people consider to be good
- middle visualization in innovative chart → have area chart show popularity over time?
- data wrangling and cleaning needs to be complete by Sunday && should have some kind of visualization started in d3 (does not need to be completed) → planning on implementing dual bar chart && clear sketches of rest of webpage && pay attention with interactions/ interactive views && storytelling
 (!!) → some lesson learned, or some new thing that you can point out

- shelf and table in one svg with transitions → image of spine and merge/ transition into image of book cover → does d3 support image?



updated intro interaction → same

svg



Project title thought process → **A cover**

worth a thousand words

Meeting 2 (11/12):

- Limit scope to only book covers

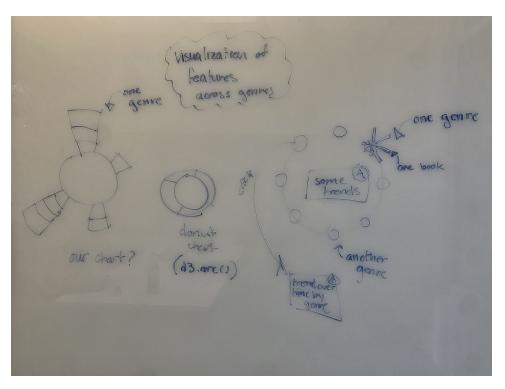
- Limit genres to most common 8 or so
- Lekshmi starts work on innovative visualization
 - using d3.arch() as a base
- Work on area chart or bar chart (simple) in parallel to get it implemented
- Isabelle data scraping (associate book covers with metadata, run images through analyzers
 - Most prominent color (RGB/HSL) Vision API
 - Object types Vision API
 - background/foreground contrast OR general contrast convert to greyscale and find difference in brightness
- group colors by broad category e.g. crayola
 10-pack colors
- group genres by broad categories



- At the top of the website: show people a couple of books and how we're analyzing them

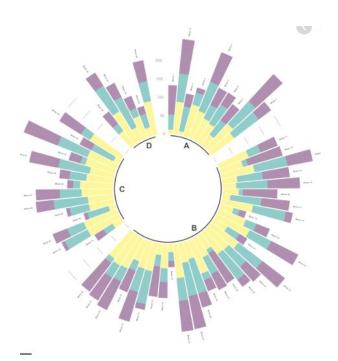
Data scraping and cleaning complete (using the real data sets)	Isabelle
Storytelling clear	Anastasia
First design of an innovative view (using mockup data)	Lekshmi
Interactions (e.g., filtering, brushing, etc.) have to be designed (either in an interaction storyboard, or in a textual description and some sketches)	Lekshmi (4th interaction storyboard)
At least one D3 visualization already partly implemented, and drafts for 2 more visualizations	Bar chart - Isabelle Area chart - Isabelle Treemap - Hold off on

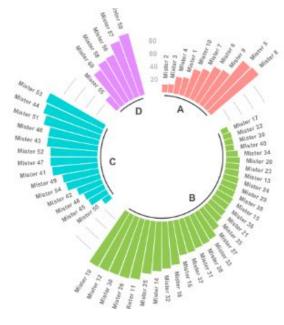
Rough webpage design and structure has to be done and implemented (placeholders for visualizations, text and images allowed) Anastasia



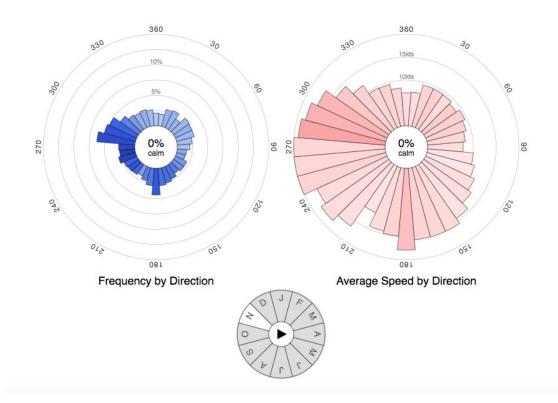
Innovative view sketch (circular bar chart)

Inspiration

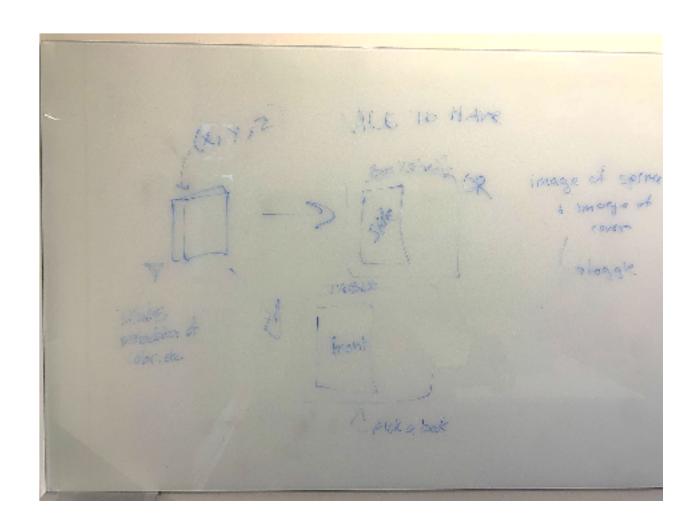


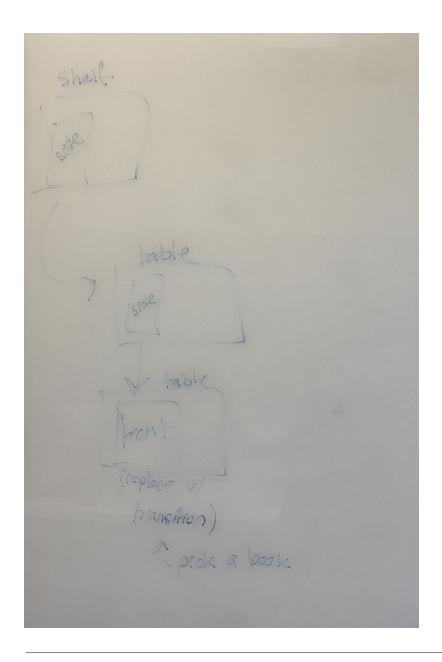


CWQP: Point Petre , Ont.



Sketches for hook (transition books for user selection)





11/10 - Detailed Project Plan

Names of all students who have actively worked on the project plan. Only those students will get points for the submission.

- Anastasia Lamothe
- Isabelle Zheng
- Lekshmi Santhosh

Definition of goals and tasks of the final project (1-2 pages)

Background and Motivation

Aesthetics tend to be something that is taken for granted, even though the creation and appreciation of aesthetics is a fundamental part of what makes us human. Aesthetics play a huge role in our everyday lives, from affecting our mood, even influencing the decisions that we make (e.g. what t.v. show we watch on Netflix, what book we choose to buy, which product to purchase). Through this project, we'd like to create a page that is able to convey just how much aesthetics can influence decision making in our lives as information and product consumers. We are designing this interactive webpage for all internet users. Book authors, artists, cover designers and publishers can particularly benefit from this analysis.

Project Objectives and Goals

Our goal is to do an aesthetic analysis of different types of designed media to see what attributes are common between different types, genres, and levels of quality. In particular, we want to look at book covers, movie posters, and visual art (e.g. museum exhibits). We will start by building several visualizations for book covers and continue on from there. For our different types of media, the aesthetic attributes we would like to look at are: contrast, color, typography, composition, size, and subject. We will analyze images that we scrape from various databases using machine learning visual analysis libraries (such as Google's Vision API, CNN's) to find data for these attributes. In addition, we will gather metadata (such as title, description, genre, author, # sold/commercial success, ratings, country of origin, year of release, covers, etc.) from other datasets so we can use them to compare our aesthetic attributes. For example, we could see what the distribution of primary colors used on book covers and/ or movie posters is based off of their years of release.

Key Tasks

Listed below are the key tasks (organized by design sprint phase) we need to complete for this project

Data Collection

- 1. Identify and finalize data sources
- 2. Set up Github repo for collaboration, version control and website hosting
- 3. Extract aesthetic features from images This task would involve gathering book cover images and extracting features such as most prominent color, title size, foreground

to background contrast differences etc using Google Vision API and other Image Segmentation models.

Map

- 4. Identify the target audience for the website
- 5. Explore the dataset and identify main questions for the project

Sketch and Decide

- 6. Sketch out visuals for each question
- 7. Decide and finalize the visual for each question
- 8. Decide the website layout
- 9. Create content for the website This includes writing the storyline (intro, analysis and conclusion) for the website and collecting supporting photographs/graphics.

Prototype

- 10. Build the main visualizations and custom visualizations using D3
- 11. Test the interactivity and other functionality of the visuals
- 12. Integrate D3 visuals with the storyline contents in the website

Test

- 13. Present V1 version to a tester and solicit feedback
- 14. Incorporate the feedback to the viz design and build the final version of the website

Final Presentation

- 15. Capture the details of the design sprint in the process book and complete any additional mandatory project documentation
- 16. Record a 2 minute screencast of the website

Related Work

We saw a <u>Pudding.cool visual story</u> that analyzed colorism in fashion magazine covers. We were inspired by the way they analyzed the average primary skin tone of the subjects on the cover and graphed their distribution, especially over time. We also liked their storytelling, especially when they isolated individual models and analyzed their depicted skin tone over time.

FiveThirtyEight did an analysis on MoMA's collection data set (<u>link</u>), but more on the metadata, including the size of paintings, the data of creation and acquisition, the medium, and the artist. We could definitely improve on the data visualizations, but this gave us some ideas of the metadata attributes we could study.

A final source of inspiration came from a talk with the director of technology at the Harvard Art Museums that Isabelle attended for the Harvard Open Data Project. His presentation talked about the vast array of data that museums have about their collections, and did a fun comparison of different image analyses from different machine learning/image processing APIs. This brought our attention to the possibility of using these machine learning libraries to find visual attributes in various images, and using this data for our visualizations.

A description of your data and where you will get the data from (at least concrete ideas on where to acquire the data)

Data Sources

We will be gathering images from various data sources (such as the Open Library Book cover API, museum APIs, and movie poster APIs). Then, we will run these images through machine learning image analysis libraries to gather data for various visual attributes for the images, especially the ones that we wanted to look at above. Finally, we will associate these images to their metadata using their ISBNs or some other form of ID so that we can better analyze the data.

Project Timeline

A project timeline (with milestones when you are planning to finish which feature)

Week	Tasks and Key Dates
Week 1 Nov 4 - Nov 10	Nov 6 - Finalize the data sources
	Nov 10 - Complete any necessary web scraping for book data and complete detailed project plan
Week 2 Nov 11 - Nov 17	Nov 11 - Run book covers through Vision API data and generate aesthetic attribute data. Align on the visual design for each question.
	Nov 12 - Finish up any additional feature generation needed from book images

	Nov 14 - Decide on the storyline and gather images/write paragraphs needed for the story Nov 16 - Complete the D3 implementation of visual design Nov 17 - Complete and submit the V1 prototype of the website with 2 key visualizations pertaining to books data
Week 3 Nov 18 - Nov 24	Nov 18 - Gather additional features relevant for movie posters and/ or artwork. Capture any additional feedback in the process book.
	Nov 19 - Incorporate any additional feedback from Viz mentor TF
	Nov 21 - Incorporate 1-2 visualizations for posters and/ or artwork onto the website
	Nov 24 - Complete and submit the V2 prototype of the website with 1-2 additional visualizations pertaining to artworks and/ or movie posters. Present the V2 to a tester.
Week 4 Nov 25 - Dec 1	Nov 25 Capture additional feedback obtained in the process book. Update the process book as needed.
	Dec 1 - Incorporate any additional feedback from Viz mentor TF and tester
Week 5 Dec 2 - Dec 7	Dec 5 - Finalize the CSS detailing and the storyline and record the 2 minute screencast.
	Dec 6 - Finalize the process book content and finish up any remaining edits
	Dec 7 - Submit the final version of the project on canvas

Feature List

A feature list (with must-have, good-to-have, and optional items)

Feature Priority	Feature ID	Feature
Must have	1	An Interactive Visualization which summarizes the aesthetic trends by book genre

		Sample questions answered - Is there any noticeable correlation between the cover color and a book's popularity? - Do books with profanity titles fare well in ratings? - How do cover color choices trend over time for sci-fi?
	2	An Interactive Visualization which summarizes the objects used on covers Sample questions answered - Does a certain genre tend to use more human faces on covers?
	3	An Interactive Visualization which summarizes aesthetic trends for movie posters or artworks
Good to have	4	Additional visualizations (at least 2) based on movie data
	5	Additional visualizations (at least 2) based on artwork data
Optional	6	A visualization which integrates all 3 media forms (books, movies and artwork) in one visualization

Roles

A description of team roles

Role	Responsibilities
Project Submitter (Anastasia)	 Submits all documents pertaining to this project on canvas
Project Manager (Anastasia)	 Coordinates meeting times across team members and hosts the weekly team meetings Keeps the process book updated as we progress through the design sprints Creates relevant project artifacts Manages communication with the team, TA and tester Keeps the spirits up when the going gets tough:)

Viz Developer (All of us)	 Generates features needed for the viz Develops interactive vizzes in D3 adhering to the spec requirements Performs unit testing on the visualizations developed
Website Hoster (Isabelle)	 Integrates the visualizations and the storyline content. Sets up the Git environment for hosting the website Conduct any website performance tuning if needed Maintains any data generation scripts in Git
Data Wrangler (Isabelle)	 Leads web scraping and data manipulation to get data into the right form for visualizations Writes scripts to scrape and change data Calls image processing APIs to gather that data for each image
Website/Viz Tester (Lekshmi)	 Conducts data and functionality testing on all vizzes Validates brushing and linking alongside other interactions on the visuals Validates the storyboard interactions and evaluate the overall website performance
Writer/Content Developer (Lekshmi)	 Authors the content (intro, analysis summary, conclusions etc) to support the storyline Collects any relevant images needed for the storyline

Meeting 1 11/8:

Possible datasets:

- Goodread books: https://github.com/zygmuntz/goodbooks-10k
- Book covers API: https://openlibrary.org/dev/docs/api/covers
- MoMA collection: https://github.com/MuseumofModernArt/collection
- Movie posters: https://www.cinematerial.com/, https://theposterdb.com/, http://www.impawards.com/
- Amazon reviews: https://nijianmo.github.io/amazon/index.html

Logos: https://data.vision.ee.ethz.ch/sagea/lld/

Insights from images: https://cloud.google.com/vision/

Aesthetic Analysis: https://github.com/aimerykong/deepImageAestheticsAnalysis

Color summarizer: http://mkweb.bcgsc.ca/color-summarizer/?api

Attributes:

- Contrast
- Color
 - Hue
 - Luminance
- Typography
- Composition
 - Foreground and background
 - Perspective
- Size
- Subject People? Vs. Objects? Landscapes?
- Title (Words

Analyze:

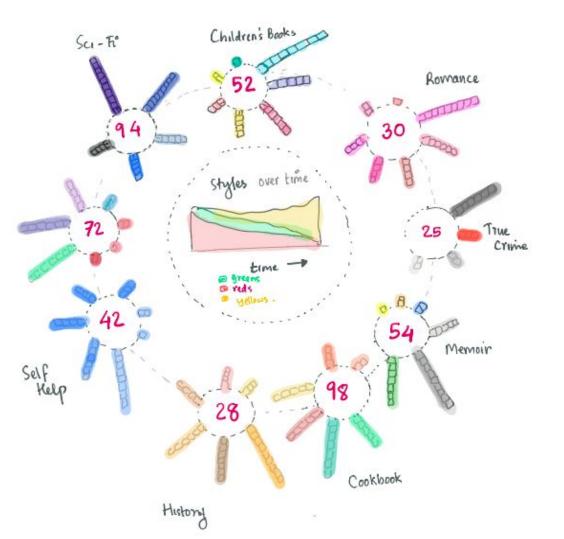
- Genre
- Rating
- Frequency
- Country of origin
- Year of release
- book rack —> select book you noticed first/ that you think is most appealing
- amazon page —> most popular page
- multiple categories, and seeing what's most popular per category
 - Start with looking at one category and building visualizations for that
- joining data: joining with Tableau (or Python?) merge data sets (for books) based off of ISBN number

Sketches



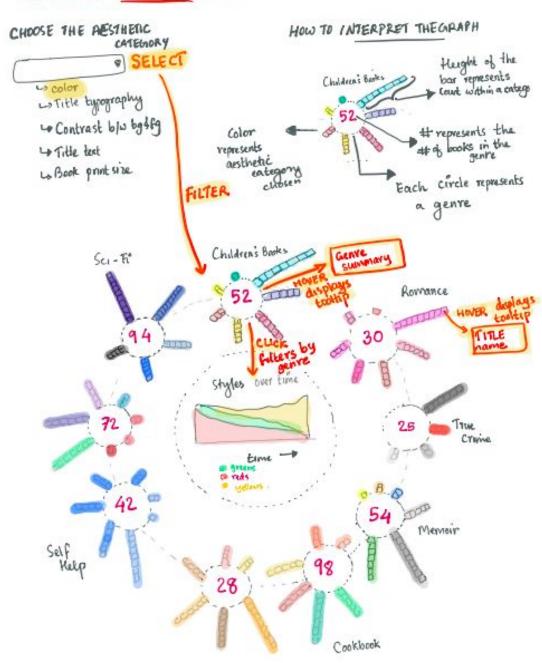
Tibense V	Most Common Subjects on Bode Covers
	on Bod Covers
15.	
Food	
Animals	
Machines	
Vature	
Aboland shapes	
	Bulldays
	frigues !
1	Faces
Awrage	Fearence

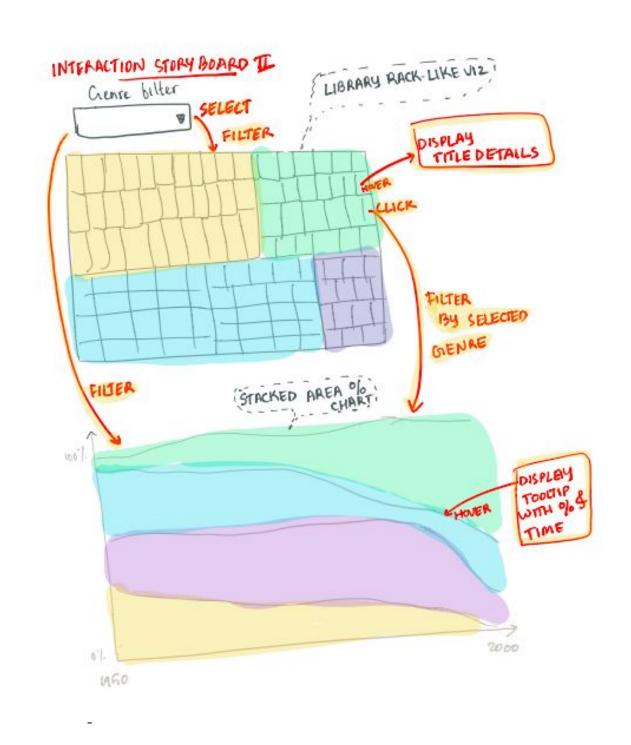
CHOOSE THE AESTHETIC CATELORY Color Title typography Contrast blw bg & fg Title text Book print size Color represents aesthetic eategory chosen Fach circle represents a genre



Interaction Storyboards:

INTERACTION STORYBOARD I





Website Layout:

