INFO 4310 HW4 Report
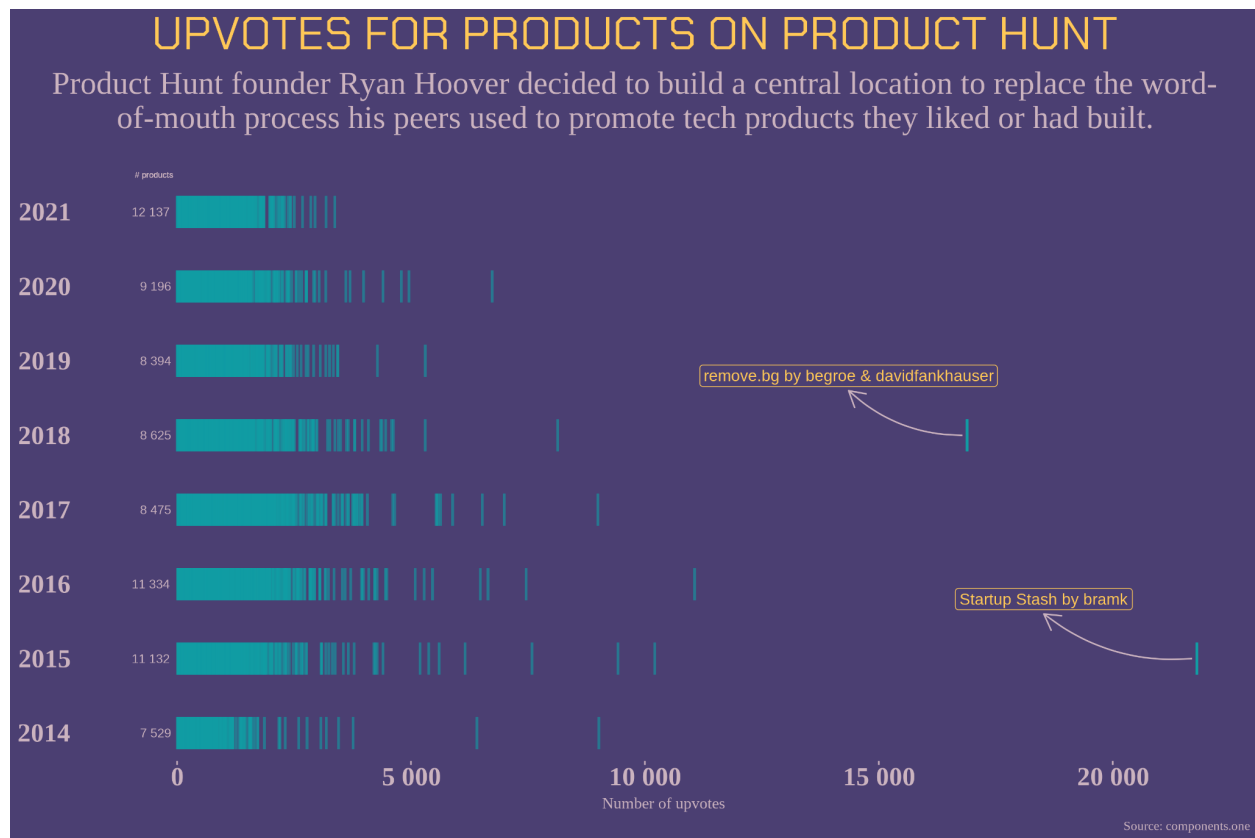
Critique

Visualization:
https://github.com/mvbloois/tidytuesday#2022-40product-hunt-products-by-componentsone
Dataset: https://components.one/datasets/product-hunt-products

This visualization was made for a Tidy Tuesday challenge. It uses a dataset of products on the website Product Hunt (https://www.producthunt.com/) and displays the products on a chart, with number of upvotes as the x-axis and year as the y-axis. The user information need is not specified.

No training or study is required to understand the visualization. The axes are labeled, as are two of the data points, which leads users to infer that each line represents an individual product.

Since lines are used as marks to represent products, it can be somewhat confusing at first glance since it doesn't fit the conventional and intuitive model of using circles for data points. There is no specific width for a line that humans would associate as representing a single unit,

meaning that users are required to take more mental steps to realize what they stand for. This is also exemplified by the intense clustering on the left of the chart, where all the lines blend together to form a continuous block, making it unclear where exactly one ends and another starts. Channels used are x position (number of upvotes) and y position (time, specifically year) on the respective scales.

The static visualization emphasizes the two products with the highest number of upvotes, which are obvious outliers, but omits most other info information. Aside from providing a rough overview of upvote distribution, it is unclear what the use cases might be.

No interactions are employed in this visualization. We think that a useful interaction would be details on demand, which would show further product details when a product is selected in some way.

We feel that this visualization is less effective:
- A linear scale is used, which results in the vast majority of lines clustering on the left, making it impossible to get an idea of the density/how many there are. We will use a log scale in our visualization in an effort to provide more clarity.
- There is too much data all at once to make any meaningful observations. We will add the functionality to take subsets with filter options.

The y-axis is somewhat confusing because it should represent the year, but the number of products in each year are also labeled in the vicinity, so that at first glance it is less clear what exactly the smaller number means. After some data exploration, we found no huge variations in the number of products across years, and decided not to include this information in our own visualization.
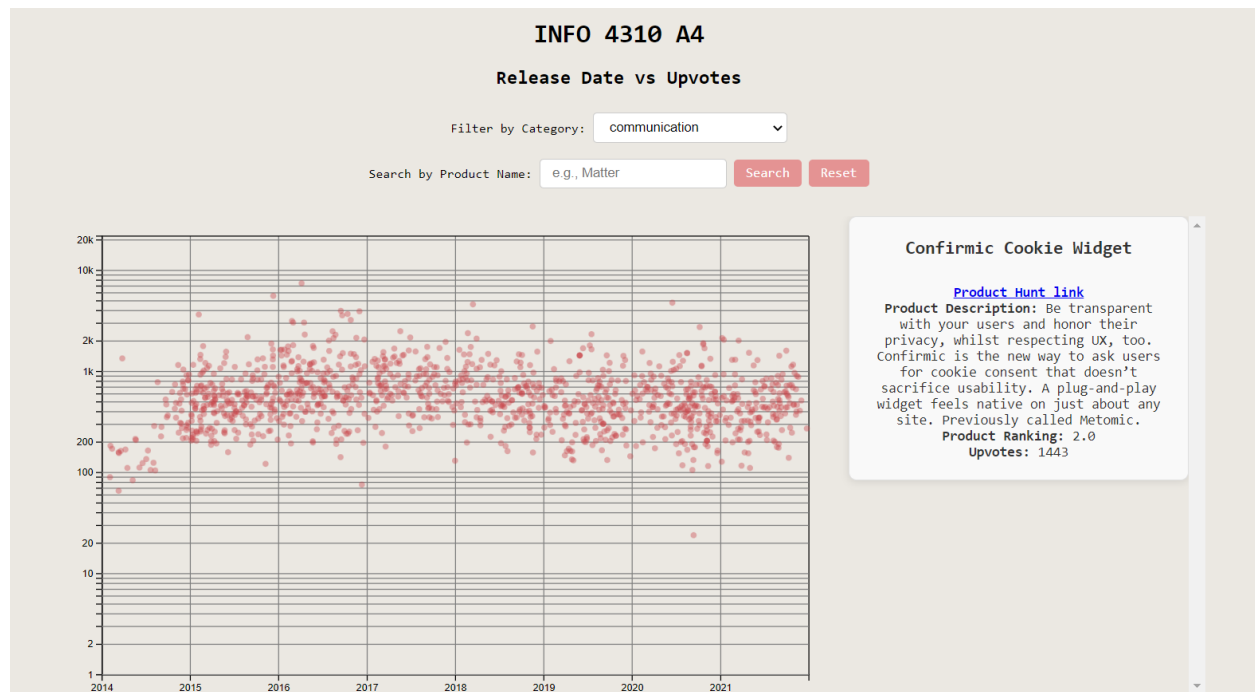
Visualization

Link: https://info4310-a4.onrender.com/

Drafts
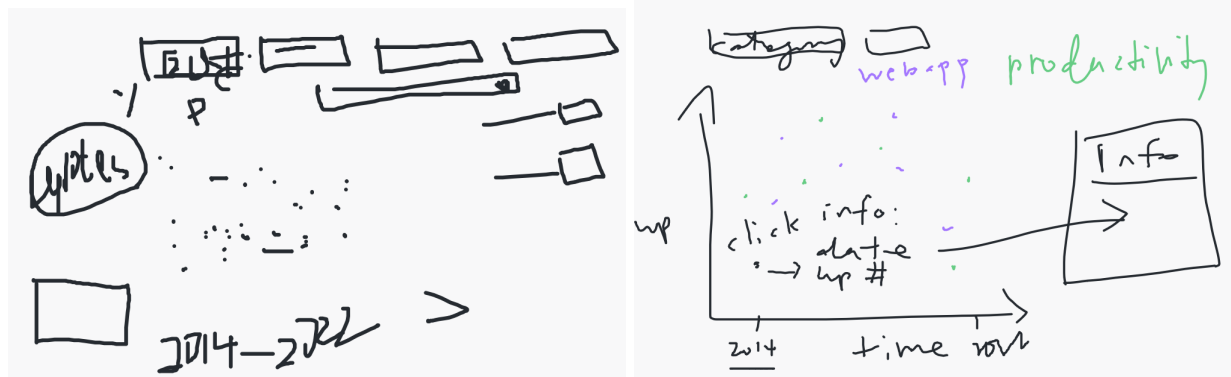
Final Visualization

Preprocessing was done on the data to narrow down the categories: there were originally 244 distinct categories, but we were able to simplify it to 15 broad ones for usability. Out of 14682 entries, only 1205 have two or more category tags after processing, which is a reasonably small percentage to ignore when coloring circles by category (the first tag in the list is used).

Our goal was to make the visualization more interactive, user-friendly, and informative, with the following design rationale:

1. **Interactive category selection**: we included a category dropdown menu that allows users to filter the data by specific categories. This feature enables users to focus on a particular subset of the data, which significantly trims the number of data points present on the screen, and may reveal insights that are obscured when looking at the entire dataset. This added interactivity helps users explore the data more thoroughly.

2. **Hover tooltip**: In our design, when the user hovers over a point, a tooltip appears which shows basic information on the product, such as release date and category. This is useful as a first impression for the user to decide whether they are interested in learning more about the product.

3. **Dynamic info-box**: in addition to the hover tooltip, an info-box is present to the right of the scatterplot, which provides a more convenient way to show product details. The info-box displays information about the selected data point, such as the product link, product image, and description. By having the info-box update dynamically as users click

on data points, they can quickly access additional details without disrupting their exploration of the visualization.

4. **Basic search functionality**: in the case that users know the exact product they want to find, we provide a basic search box that dynamically shows all products with the name input by the user. Moreover, we show the details of the retrieved product(s) in the scrollable info box.