

Process Book

Our World in Tweets

Ashley Lan, Hibban Butt, Kelsey Nicholson

<https://github.com/nicholke/dataviscourse-pr-Our-World-in-Tweets>



Table of Contents

Our World in Tweets	1
Overview and Motivation	3
Related Work.....	3
Questions.....	4
Data.....	4
Exploratory Data Analysis	5
Design Evolution	5
<i>Prototype 1</i>	6
<i>Prototype 2</i>	7
<i>Prototype 3</i>	8
<i>Final Design</i>	9
Implementation	11
<i>WorldMap</i>	13
<i>Bubble Chart</i>	16
<i>Line Chart</i>	17
Evaluation	19

Overview and Motivation

The prevalence of social media in the past decade has changed the way people receive and exchange information. Social media has become one of the largest and most popular tools for trading information between friends, students, organizations, and businesses. A large network of public information is useful and important because it spreads new ideas, events, and broadens awareness of existing ideas.

Twitter is an especially powerful platform because of its ability to display trends in real-time. Trending data provides insight into not only what topics are important or popular, but specific sentiments towards news, issues, and interests. Twitter's algorithms use trending hashtags to prioritize displays of the most popular trending topics. Trending topics can be used to spread breaking news globally and because of Twitter's public API, trending data can be used for analysis. An analysis of trending topics could provide an extensive snapshot of the most important current events at a specific time on a global scale. This information is specifically interesting to our group because of our extensive use of social media for communication and news. There are many current events in the past few years that have been impactful and life altering, we are curious about tracking these trends and events across the globe.

Related Work

The Conversation: Twitter Trends 2022

https://marketing.twitter.com/en_gb/insights/the-conversation-twitter-trends-uk-2022

The report for trending twitter topics in 2022 uses animated bubble charts to display various categories of trending topics and a subsequent breakdown of each category. We plan to adopt a similar approach for our bubble chart; however, it is hard to tell if the scale of the bubbles in their chart is based on popularity of trend. In our bubble chart we plan to utilize the area of the bubble chart to display counts as well as topic.

Twitter Trend Worldwide

<https://getdaytrends.com/>

This website displays the order of trending topics with a numbered chart, the trending topics, and the order it is trending. #1 is the most trending topic and so on, this display also includes a small line chart with the trend over a 24-hour period. We plan to utilize a numbered chart with order of trending topic and instead of individual line charts to show trends overtime we plan to use one large line chart with the trends displayed together.

Data Viz Hw 4

<https://github.com/dataviscourse/2022-homework/tree/main/hw4>

The world map from hw4 inspired our use of the world map and may be used as the first implementation of the world map in the project.

Mapping the World one Tweet at a Time

<https://mislove.org/tweetmap/>

The world map visualization uses the geotag of tweets to map out the different locations that each tweet took place globally. The map uses a drag and zoom feature that we may implement on our world map as well. Instead of only using the geotag to display where the tweets were created, we will use the geotag and display the accompanying trending topic at that location.

Questions

Twitter is one of the most popular social media platforms available today. Twitter provides insights into popular trends and important cultural and political moments. We are interested in discovering the twitter trends, especially, trending topics across the globe. The project initially centered around finding the most popular issues globally in the year 2021. However, data constraints with the twitter API only allow us to grab trending topics from the past 24 hours. To provide enough relevant data from the project, we will be pull and store the twitter API data every day until the project submission, the visualization will display trending data for the past month instead of the past year. This slightly changes that project questions to:

- What were the most popular issues on Twitter in the last month?
- What were the different geographic breakdowns of these popular issues?
- How long are trending topics maintaining popularity?

Data

For the project we considered using twitter API to get 2021 trends, tweet volume, and location of tweet. While waiting to gain access to the data, the group found an existing dataset from Kaggle, called “Twitter: Trending Tweets per week”. This dataset would be used if we were not able to access enough relevant data from the API. However, the Kaggle dataset also has some limitations, it does not contain location information for all countries and has many missing values for tweet volume. This dataset would leave several countries blank in the world map.

link to the dataset is below:

<https://www.kaggle.com/datasets/rsrishav/twitter-trending-tweets>

Data Update

The group gained access to the twitter API and discovered that trending data can only be pulled for the last 24 hours. The group met to thoroughly review the differences between the data from the twitter API and the Kaggle dataset to decide how to proceed. The data from the twitter API contains trending topics for every country, unlike the Kaggle dataset which only contains info for several countries. Both datasets contain several missing values for the tweet volume, so there wasn't much benefit with either dataset. After meeting with the project TA, we decided to pivot the project topic slightly and use the 24-trending data from the twitter API. Instead of visualizing 2021 trending topics, our group will pull and store trending topics every 24 hours until the project submission. Data will be stored as a json file with date as the key in mm-dd-yy format and value with a nested dictionary that contains country name and relevant trending information. At the end of the project this will provide approximately a months' worth of trending topics and tweet volume for every country. Tweet volumes that are equal to None will be filtered out and additional URL links to the trending topic twitter page were also stored for potential use in future animation.

Exploratory Data Analysis

Exploratory data analysis was performed on the Kaggle set and twitter API data. Initial analysis used histograms to compare volume and country value distributions. The histograms were informative of the total missing values between datasets. This analysis led us away from the Kaggle dataset and helped us decide on the twitter API data. To explore this data, the group looked at the different types of data that we had access to with the API. The `api.trends_available()` had the most useful and complete information for the project. After exploring which data needed to be pulled, the "None" fields were filtered out, this brought the first day of data to a total of 1109 entries for all 50 countries.

Design Evolution

For the initial design phase of the project the group sketched three prototypes designs. Each design focused on a different component, the prototypes were separated into three different sections. The final design took the best versions of these designs to create the final design layout. The main components of the final project will be a world map, bubble chart, and line chart. The final design will be used as an outline for implementation. As the project continues, these designs will be updated in the implementation section as new ideas and better visualizations for the data are realized.

Prototype 1

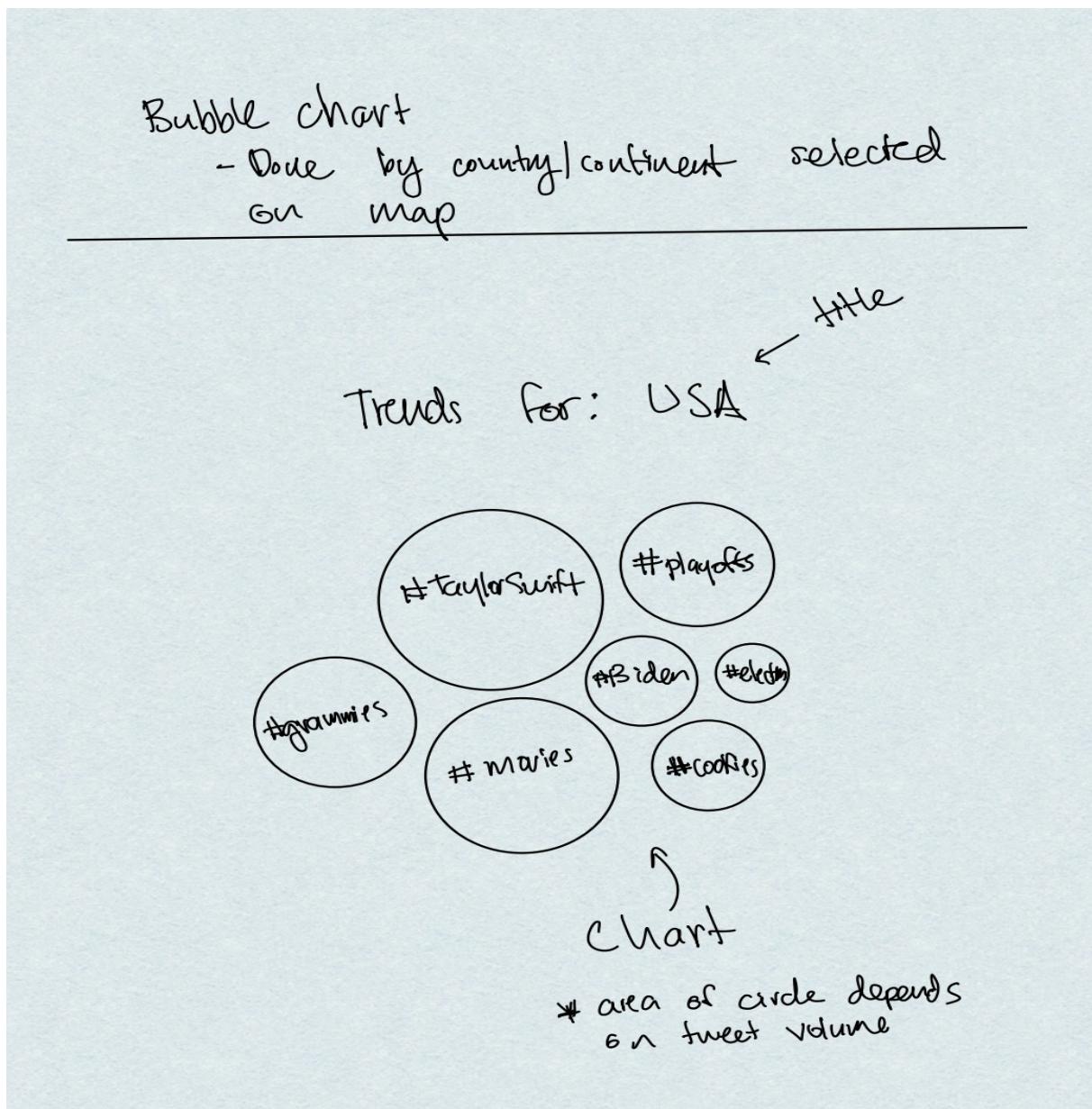


Prototype I will focus on the design of section I for the project. This prototype will use an interactive map and instead of an updating area chart, this design will utilize a bar chart to display varying trends between countries. The bar chart will use one country at a time and will update with the top trends and number of total counts for each trend.

This design will focus on looking at the trends in 2021 for each individual country. The bar chart will have a drop-down feature that allows the user to select how many trending topics they want to examine. As this parameter is updated, the bar chart will display animation that updates that

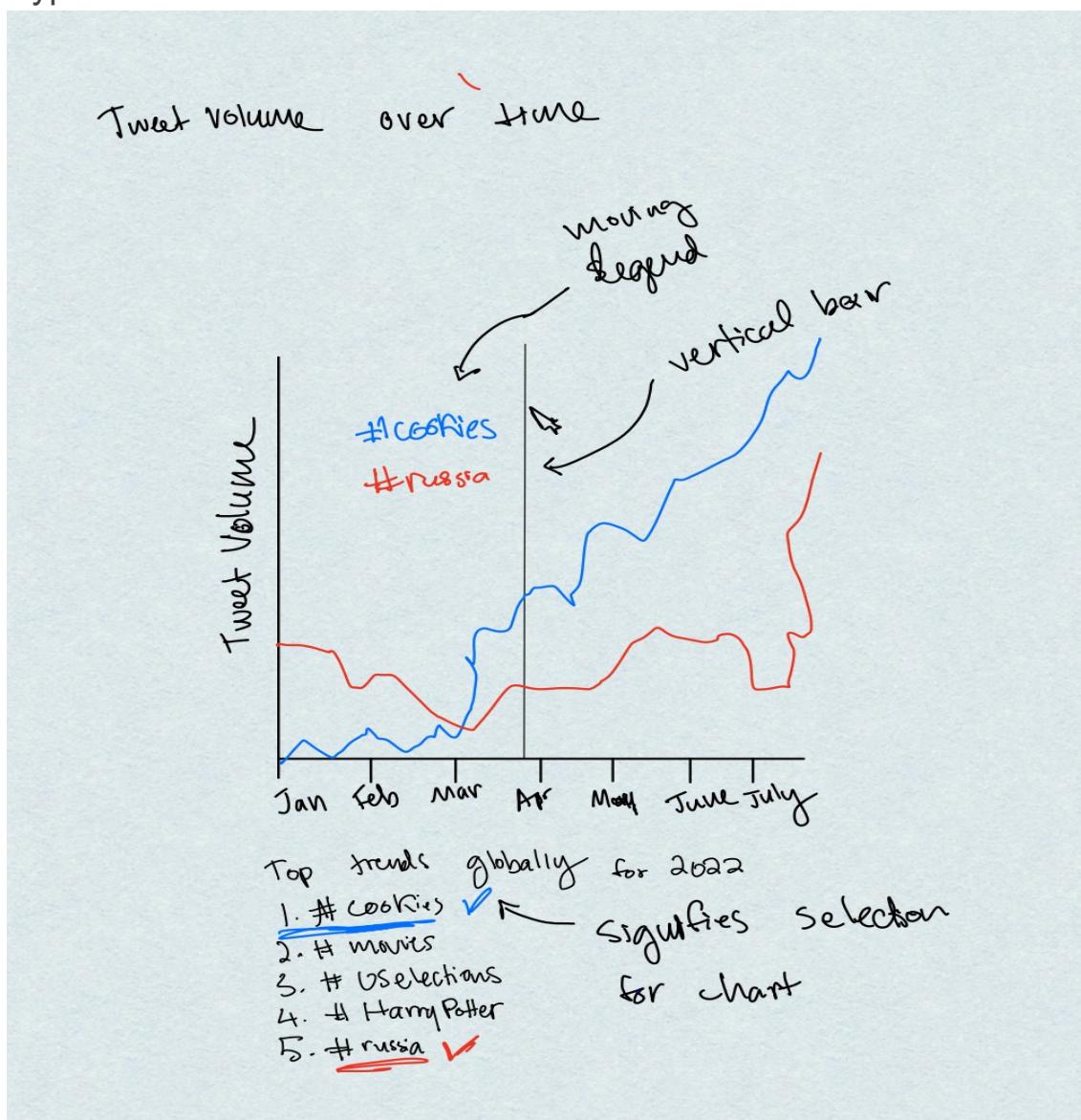
width of the bar charts depending on how many trending elements are displayed. If there are only a few trends displayed, the bars will get larger and the more trends selected, the smaller the bars will appear. For this prototype the second section of the visualization will be the same as described in the final section above. A line chart with top 50 trends of 2021 will update based on the specific trend selected by the user, this provides specific analysis of trends over time.

Prototype 2



Prototype 2 will focus on designs for section 2 of the final visualization. The first section of this design will incorporate the interactive world map and updating area/bubble chart thoroughly described in the final design section. This first section focuses on the top trends for each country in 2021. The alternative design for the second section will look at the top trending topics globally and provide an in-depth view of the trends at specified time. This design will utilize a scatterplot that will allow the user to filter the scatter based on a specific time provided in the drop down. This will populate the visualization with the 50-100 top trends at the given time. The dots will have a hover feature that displays text with the trending topic and number of instances in the dataset. The user will be able to examine data on a day, month, or yearly basis. Duplicate dots will be displayed with the same colors, so it is easier to see trending topics that span several days, weeks, or months

Prototype 3



The third prototype utilizes a world map with interactive hovers and pop outs. The first section of the design will only display an interactive world map. Each country will have a specific circle that can be clicked on to examine overall 2021 trending topics per country. When a country is clicked, the small colored circle will expand and cover up part of the world map. The expanded circle will display information about the top 5 most trending topics for each country globally. This provides an interesting interactive experience for the user. This allows them to click on any country of interest and see each country's 2021 trending topics in a quick snapshot.

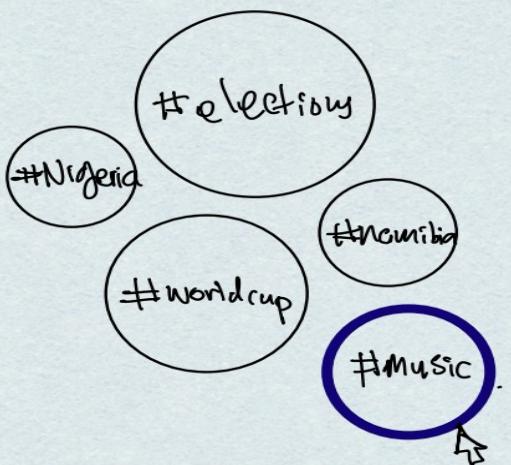
Final Design



World map + area chart combo with
popouts

Trends for Africa

from January to December 



MUSIC

The use of instruments and a voice to construct melodic sounds.

popout with
more info on the
trend from Google

For our final design we plan to have two distinct sections that will showcase different parts of the 2021 trending Twitter topics. For the first page of the website, we plan to have two different visualizations. Our project aims to look at trending topics from a global perspective, to showcase this we have decided to use an interactive world map that will update an accompanying area/bubble chart. The default state of the area/bubble chart will showcase the top 20-50 trending topics globally.

Differences in size and potentially color will be used to group trending topics by popularity. The larger the area, the more popular the topic. This method will be used to show trends from a broad perspective

and grab the attention of the user. This may prompt the user to further examine trends on the second page. The name of the trending topic will be displayed in the center of the circle. The area chart will also update based on the countries that are clicked from the world map. Once a country is clicked the default state will clear and the user can select one or several countries and then examine the most popular trends for the selected countries. Depending on the data available, we may also include a dropdown on this page that can filter data for a specific period in 2021 and then the user can click on countries to see trends at those specific times.

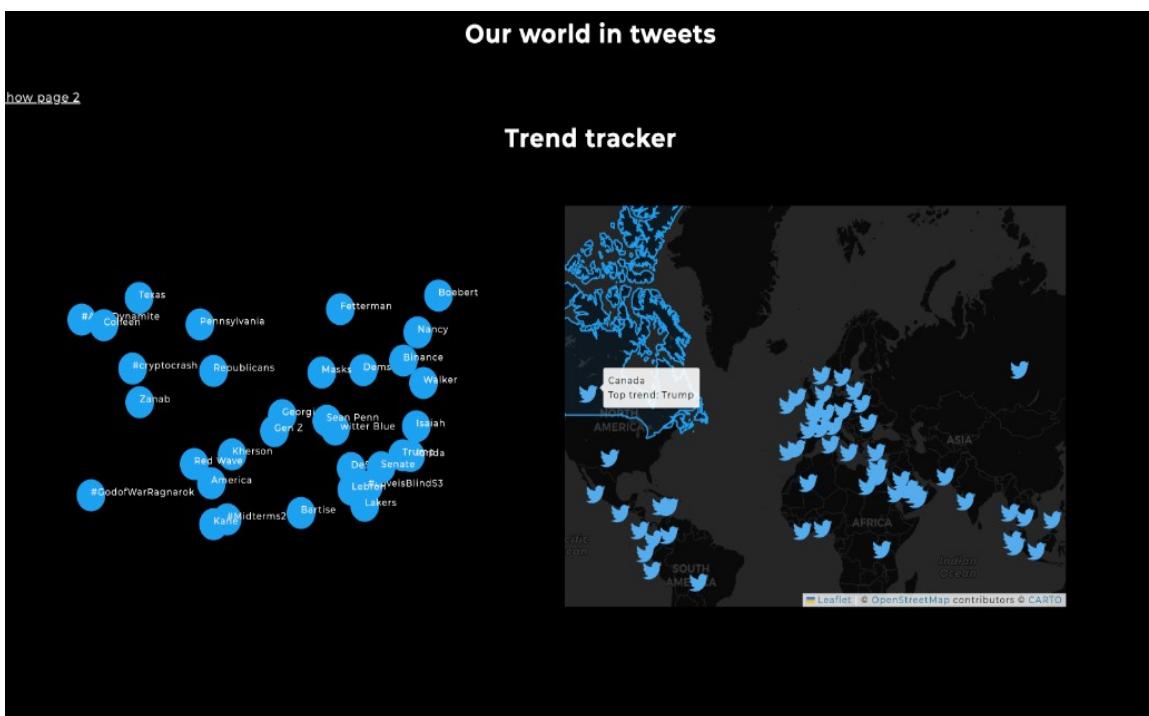
The second part of the project website will visualize the most popular trends in the past month.

The second page will display approximately the top 50 trends of 2021 in text on the right side of the page. The different trends will be represented by a color scale and will update a line graph on the left side of the page, making it easy for the user to see all the trends quickly. The user can then click on any number of the popular trends and visualize them in a time-series chart. The lines will be represented by different colors to allow the user to visually track the rise and fall of the trend's topics over 2021. The y-axis will contain popularity which will be measured by the total number of occurrences of each trending topic. This method can be used to track trend popularity over time and see when current events became popular and for how long.

Implementation

Overall Layout

First version for Project Milestone



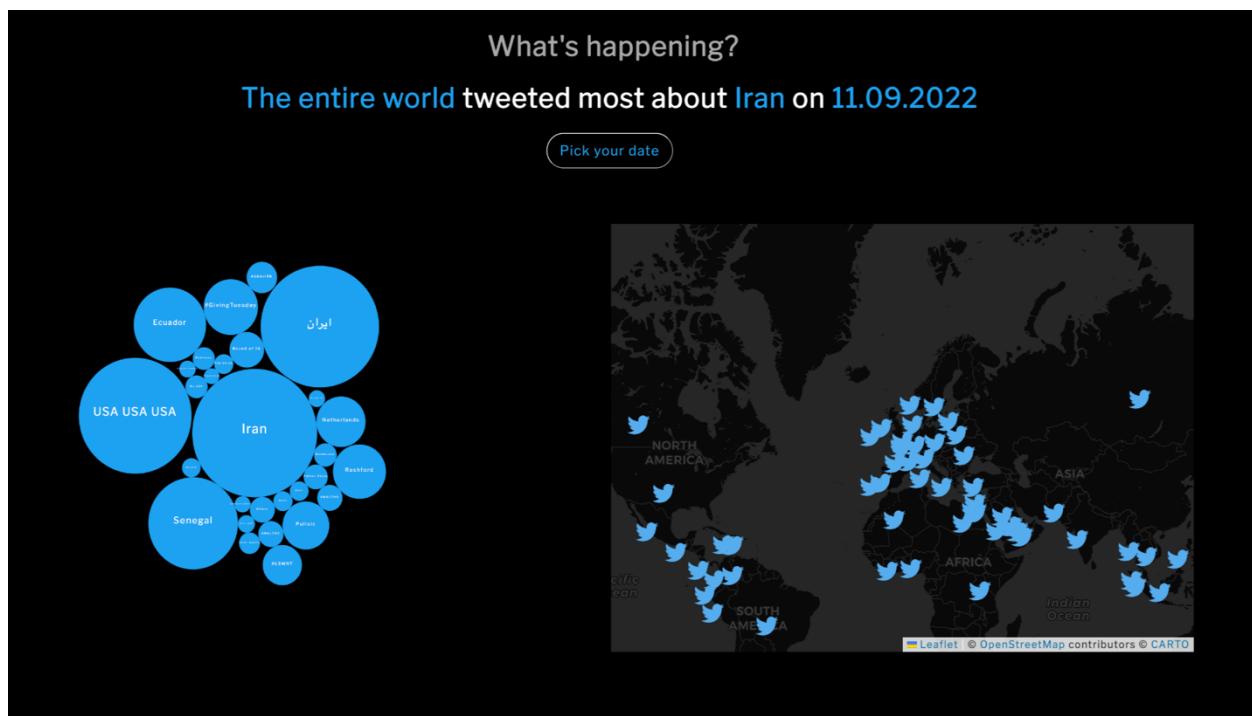
This is the current overall layout for the final project that was submitted as part of the project milestone, each section will be broken down and continually updated with design evolution and interaction implementation above. The next elements of the project will involve getting interactions and additional features functional.

Final Implementation

Our world in tweets

Social media has revolutionized the way we communicate, allowing us to share thoughts and media about the trends that we feel matter most. This has generated loads of data on what people all over the world are passionate about. We collected just a sliver of what's possible in Twitter data to see what people cared about most in November 2022, and where. Scroll down to learn more.

↓



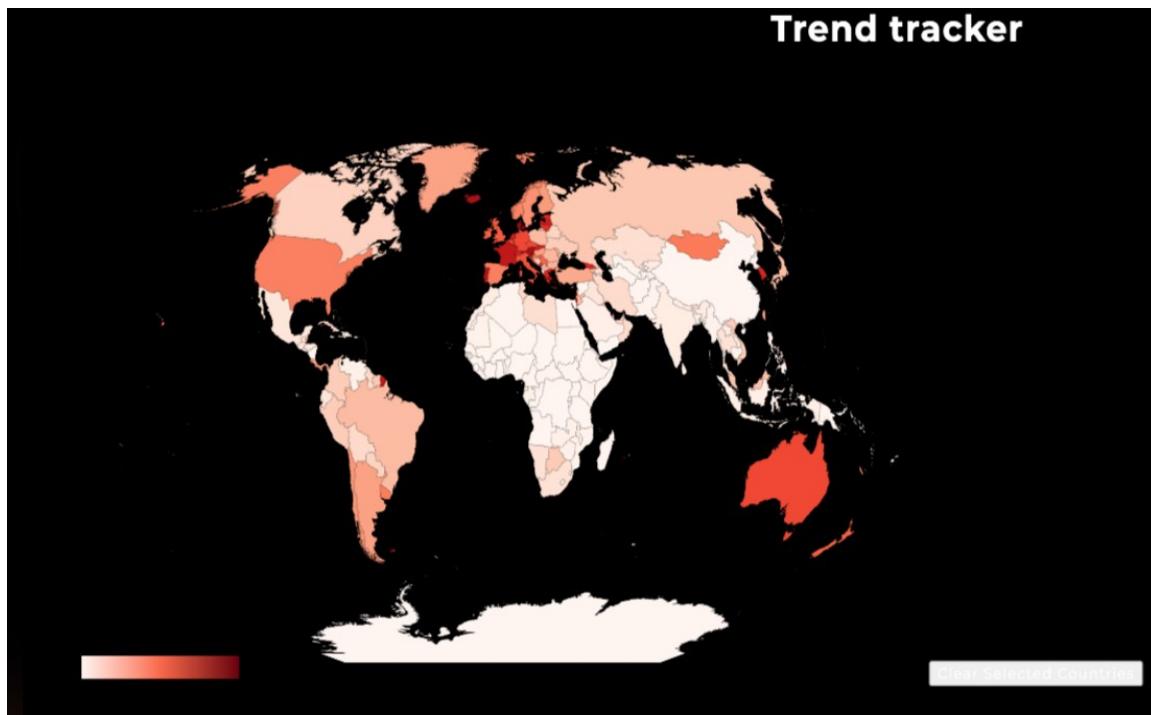
From the previous overall layout, the webpage has been updated with contextual information and a scrolling functionality. When the user navigates to the webpage the first image above will be the first display they will explore. The text on the first part of the display updates the user with the context of the visualizations and why exploring Twitter data is useful and interesting. When the user scrolls down the user will see the first two visualizations, the second display above. This display shows the bubble chart and accompanying world map. The text above the visualization updates when a country is selected, the text in blue changes depending on the information clicked. The country, trending topic, and date will be the only elements that are updated in blue. The default state of the visualizations will be the entire world. The graphs will display the trending topics for the entire world at first display and will update when the world map is clicked. Not only will the text update with the country clicked, but so will the



accompanying bubble chart. The text gives a broad overview of the contents of the bubble chart to help the user decipher the results. Another important functionality of the display is the pick your date from the drop down that will allow the user to select date from a range of dates. When the data is selected from the drop down, the world map and bubble chart will update with the most trending topics from the specific point in time.

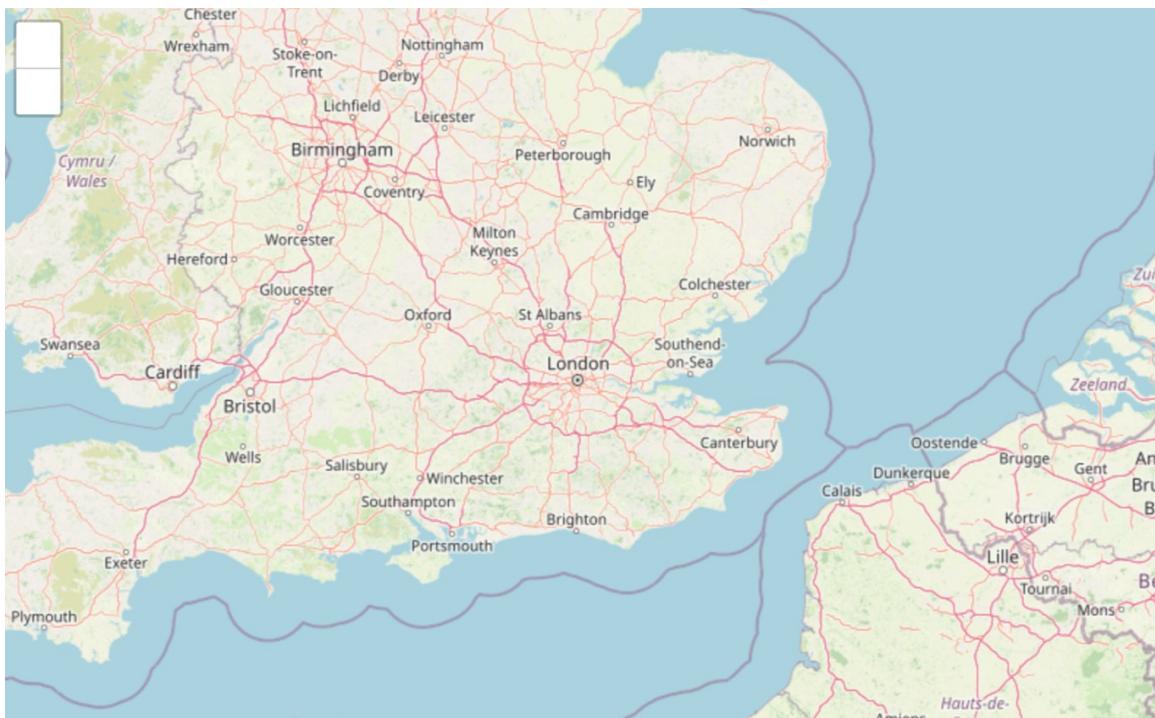
WorldMap

First version of World Map for Project milestone



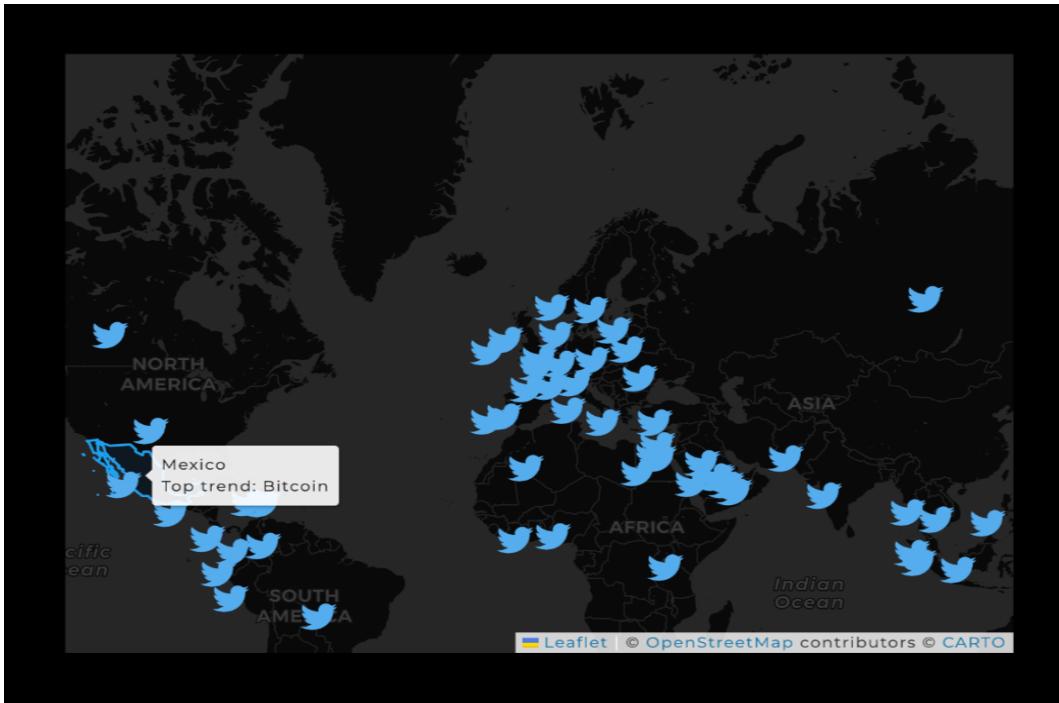
The first version of the interactive world map was inspired by the homework 4 world map. The trending topics and tweet volume will be linked to each country on the map. When the country is clicked it will eventually update the bubble chart with the trending data. The background display of the visualizations is black to mimic the Twitter display. The current color scale uses the interpolated red scale used in the hw, but this will be updated in the next version.

Potential candidate for final World Map



The next version of the world map that was tested as a potential fit for the final world map, is a topo map. This version would allow the user to zoom into specific sections of the country and could be interesting for the user to get more information about cities and land features associated with each country. This version of the map is commonly used in other map applications, like apple or google maps and is familiar to many users. This may be a good option for the final visualization, but the group will continue to meet and evaluate the overall layout of the project to decide the best display.

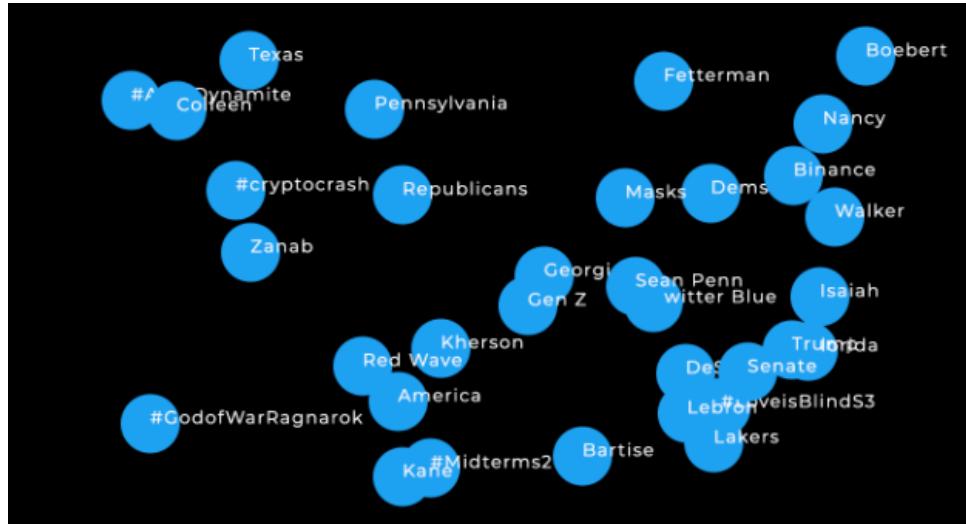
Final Implementation of World Map



The final version of the world map uses a similar design to the first visualization but contains points represented as twitter icons for each location that a trending topic was tweeted. This allows the user to hover over the specific location and examine what the trending topic was. This utilizes the black display and popular twitter icon to give the user a similar experience to the Twitter application. Also, the world map has a clicking feature that outlines the clicked country in blue. The selected country will then update the accompanying bubble chart with the trending topics for that country. Also, the map has a zoom function similar to the common application, google maps. The user can zoom in to specific countries and grab the map and drag to a specific location.

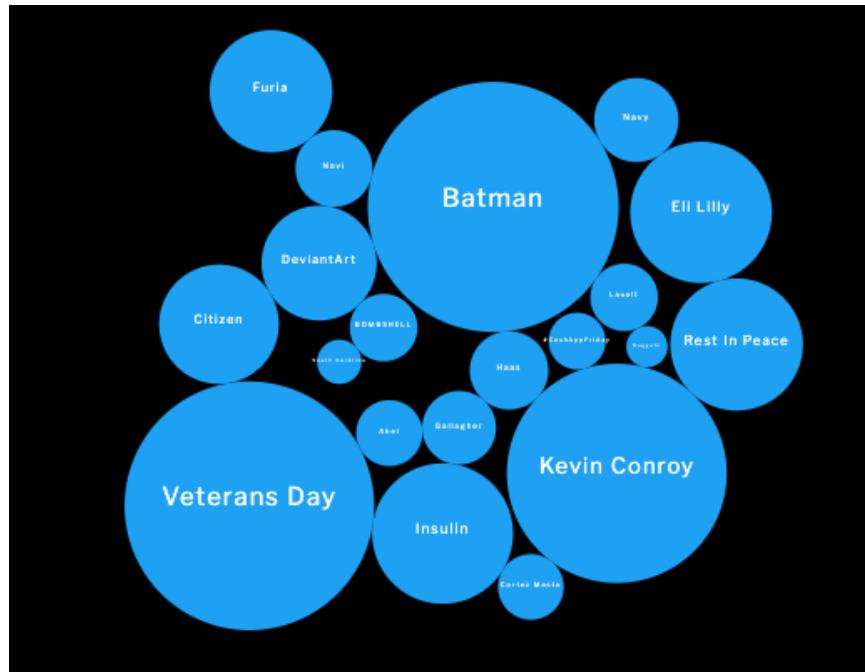
Bubble Chart

First implementation for Milestone



For the first version of the bubble chart each bubble is populated with random text. The bubble will update with the most popular trends for each country when the country is clicked. The final version of the bubble chart will use d3.force for animation and a proper area scale. The final area scale will use tweet volume, the more tweets with a specify topic, the greater the area and vice versa.

Final Bubble Chart

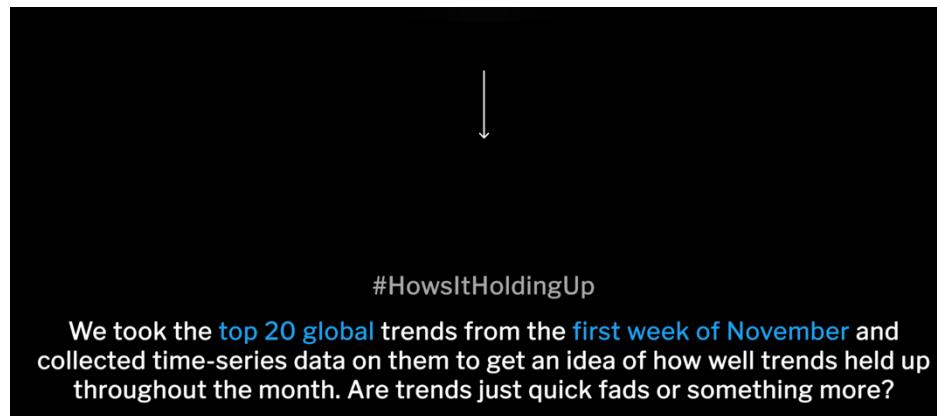


The final version of the bubble chart was updated from random text to corresponding trending topics for the selected country. The bubble sizes update depending on the number of occurrences of that trending topic and use animation to swarm the bubbles when a new country is clicked. More popular trends have a larger area, and the least popular trends have the smallest area. This bubble chart populates with the trending topics corresponding with the clicked country. With the example above, Twitter Blue is the most popular trend in the United States on a user selected date and based on the area of the bubble in comparison to other trends, it is significantly more popular than a smaller bubble displayed, like Louisville. When there is no data present the graph populates with a short text description displayed above. We were able to get data for most countries, but the few we couldn't we added the

text display shown.

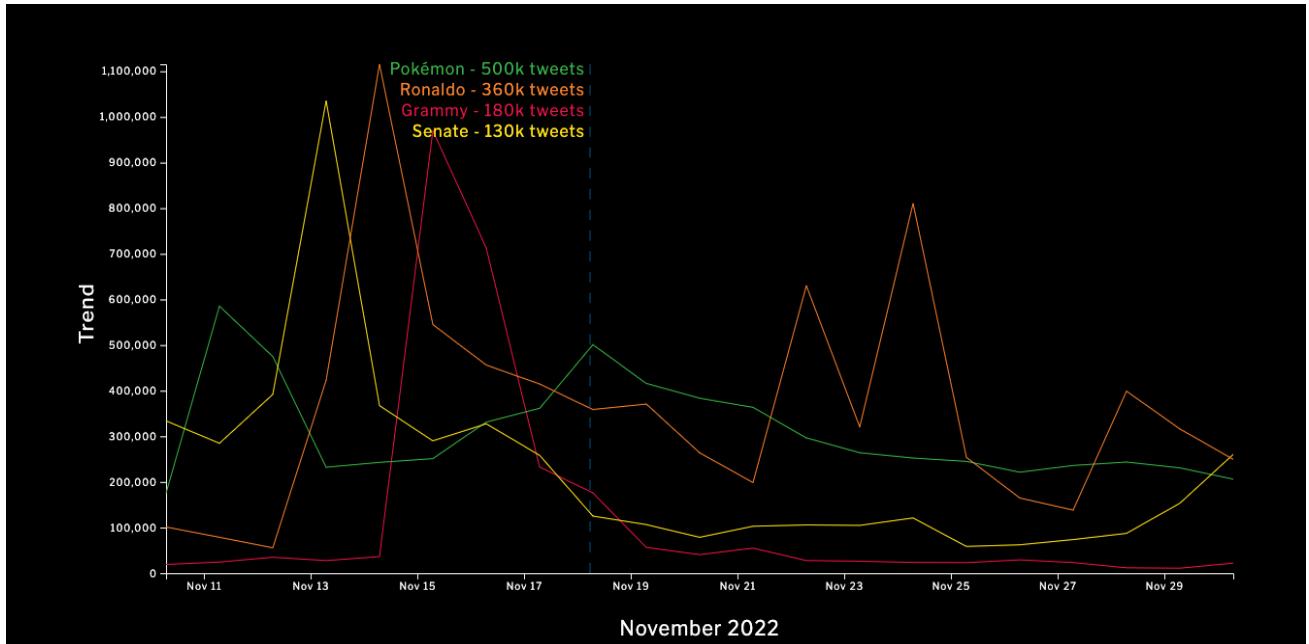
The Twitter API gave us no data on Greenland :(

Line Chart



This is the base layout for the time series section of the website. As the user continues to scroll down the page this text will display. This allows the user to continue to the time series visualization and informs them of the types of data they will examine. The next visualization will include top 20 trending topics globally from the beginning of November to project submission.

Final Implementation of Line Chart



The implementation of the Line Chart appears after the user scrolls past the text indicating that the top 20 global trends will be displayed. The default state of the line chart showcases the top three trending topics globally. The trends populate the y-axis, and the time is populated on the x-axis. Since we were only able to collect data for November 2022 the axis starts at the beginning of collection, November 11 through date of submission Nov 30. When the mouse is detected on the line chart a scroll line appears. The scroll bar updates with the trend and total number of trends for that topic at each point along the line. When the mouse reaches the end of the graph, the text on the scroll line flips.

1. Grammy
2. Pokémon
3. Senate
4. Russia
5. Ronaldo
6. YET TO COME
7. Batman
8. Veterans Day
9. Qatar
10. DeSantis
11. BTS WORLD CUP
12. Ticketmaster
13. Album of the Year
14. Kevin Conroy
15. Poland
16. ポッキー

At the bottom of the line chart, the top 20 trends are listed. When the page is opened the first three trending topics are highlighted with a different color. When the user hovers over the text it will highlight in blue. When the text is clicked it will highlight with a unique color. The color corresponds to a line present on the line chart. As the trending topics are clicked, the line chart will continue to populate with time series data for each trend. When the colored texted is clicked again the line will be removed from the line chart.

Evaluation

From the data we can learn a lot about the important events and topics happening around the globe. This project only contains data from November, but it is an interesting time to follow the trending topics as Twitter acquired new ownership in November. From the visualization we were able to see similar trending topics and moments are

important across the globe, especially the massive changes that occurred to Twitter during this time. These visualizations could be useful for tracking current events and topics across the globe over longer periods of time. The visualizations make it easy and simple to visualize all the trends in certain locations and places instead of combing through Twitter API data to find specific instances of information.

For the first version of the project our group sought out to find the most trending topics of 2021. Many large current events occurred in 2021 and we thought it would be interesting to see how similar and different these events were across the globe. However, one of the main difficulties of our project was acquiring the right data. The group had to gain access to the Twitter API before we could determine the information, we would have access to. In the meantime, we found datasets on Kaggle, but were limited by the trend counts and number of countries that contained trending data. Upon gaining access to the Twitter API, we discovered that trending data could only be pulled for the past 24 hours, so we were not able to access the trending data for 2021. The group pivoted the project and we have been collecting and storing the data every day for the month of November. The project will only include data for the month of November, but it would be interesting to have and explore data for a longer period.

Throughout our implementation of the visualizations, we have had many evolutions of the different displays and visualizations of our projects. Our original design prototypes were implemented and either changed because the project pivoted, or we found alternative designs that fit better with the overall look of the project. The evolution of the world map started with a similar design to the world map used in hw 4. Originally, the group did not decide on a specific color palette or theme of the webpage. After implementation of several world maps seen above, it was suggested we follow a similar layout to Twitter. Black background displays with hints of blue and twitter icons were used as design choices. We thought these design choices would grab the user's attention and provide quick context about the source of our data if the user was familiar with Twitter. Also, originally the display outline was going to have separate pages for each section of the project, however, it was decided that a scrolling feature would better suit our needs and give the user a better flow from one aspect of the data to the next. With the scrolling it allowed us to display the text elements before the user sees the visualization. This provides some general context and understanding of the designs before the user gets to the page. This way the user won't be overwhelmed with the number of elements on the page, and they may be able to orient themselves with the visualization quicker than a separate page display.