DATA VISUALIZATION

AIRCRASHES

ANDRIJA JELENKOVIC
ALEKSA MILISAVLJEVIC
LAZAR RADOJEVIC

INTRODUCTION

Airplane crashes contain loads of seemingly tedious data that needs to be carefully analyzed and understood. Visualizations about airplane crashes can be intriguing due to their ability to provide a **comprehensive** and **visually compelling** representation of complex data. They could also offer **a unique perspective** on the crashes and accidents, enabling viewers to delve deeper into the factors and circumstances involved. By combining various data features such as location, date, and airplane model, visualizations can create a cohesive narrative that helps researchers, investigators, and the general public gain a better understanding of the incident. Simply put, they allow us to **observe data patterns and trends in a more intuitive way**.

OUR GOALS 🎯

Determine temporal patterns of crashes: we want to provide a user with an interesting figure that would engage them. A user would be able to derive logical conclusions about yearly decreases and seasonal oscillations in the number of total crashes.

Interact with a detailed rendered map of the United States to observe spatial trends: one of the key factors in an airplane crash is where it happened. Therefore, by restricting only to data that have crash location coordinates we would like to **encourage users to explore** the spatial trends.

Show the spatial and temporal differences across multiple airlines: point out the differences between amateur and commercial airlines with interesting and engaging visualizations.

NOT INTERESTED?

Even though you might not be interested in airplane crashes, we tried to create a **thought-provoking story**. Most of the stories that our visualizations present are very intuitive and accessible to the broader audience. Therefore, we encourage a passive user to **think and interact** with our website.

DATASET

We will be visualizing the **NTSB aviation accident database** from the Kaggle. NTSB is an acronym for the National Transportation Safety Board, a federal agency in the United States that is responsible for investigating civil transportation accidents. They aim to uncover the cause of incidents and accidents and issue safety recommendations to prevent them from happening again.

As of 2014, they have issued as many as 14,000 safety recommendations. This database, therefore, provides one of **the most relevant sources of information** for anyone aspiring to delve into the topic of aviation accidents.

BRAINSTORMING

Geospatially, one could examine the distribution of crashes across different regions, identifying hotspots or areas with higher crash frequencies. **Temporally**, analyzing crash data over time can reveal patterns, such as seasonal variations or changes in safety measures. Additionally, by exploring **airline-specific trends**, such as accident rates based on airlines or aircraft models, we can provide insights into their differences.

HOW TO VISUALIZE?

We had several visualization ideas, being mindful of **selecting the most impactful ones** to create an engaging user experience. Considering the different features we aimed to present, we narrowed down our options and settled on the following visualizations:

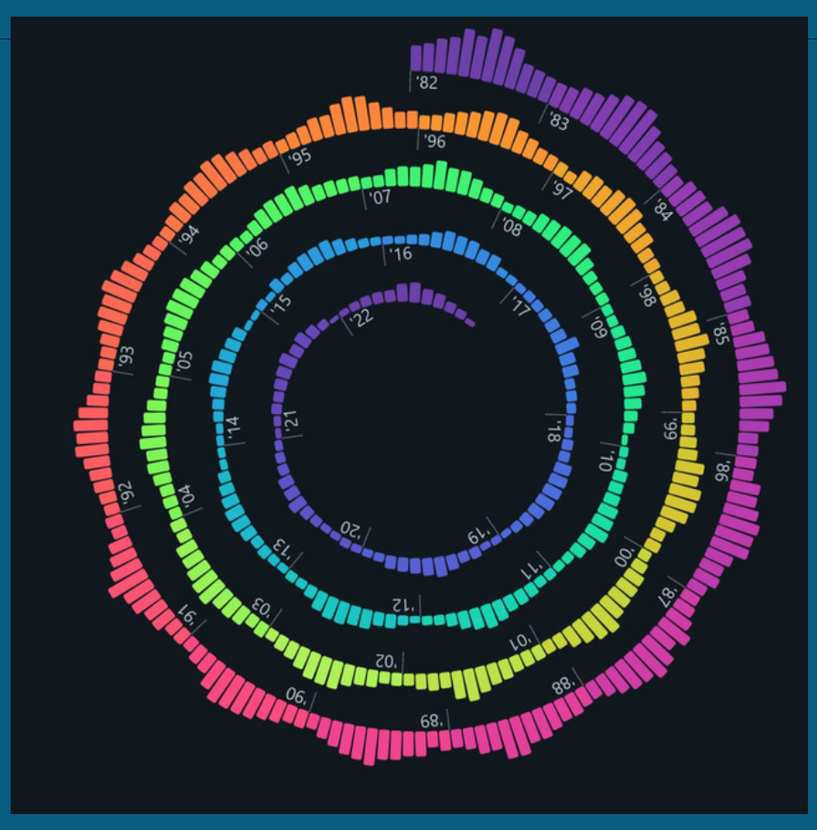
- **Condegram**: a spiral that is great for presenting the periodical and temporal trends in our data.
- **Map of the US**: detailed map of the US is used in the two visualizations. With additional interactivity, it becomes really interesting to play with.
- **Bar race**: it is always compelling to watch how changes happen over time. With bar race videos as inspiration, we created our own.

CONDEGRAM

INSPIRATION

We wanted to show the **temporal trends** of our data. However, we wanted to be unique with our approach. Kirel Benzi's Circadian Rhythm inspired us. The spiral is used to present the total crashes per month in order. A user can effectively notice the global decrease in the total number of crashes per month and seasonal trends within the same year.

DESIGN EVOLUTION



As we already knew what we want to implement we did not have any particular deviations from the plan. We tried to align years so that every full circle contains 10 years, but the trends were less visible. As the design progressed, we added interactivity. The condegram loads from the latest to the earliest year, in different colors. When **hovering over a bar**, the total number of crashes in that month is displayed in the middle of the spiral.

The design suggests that as we are coming closer to the center of the spiral, the number of total crashes decreases. If you were to predict the future and continue the spiral, the bars would be even smaller. That implicitly suggests that as years are going by we are **converging to the point of no airplane crashes**.

CHALLENGES

The main challenge was to implement the condegram in code. It is **not a standard visualization** so it took extra time to do it properly. The initial solution was concentric circles, but the transition from outer to inner did not make sense. Therefore, the only option was to implement a spiral.

After having the spiral, we had to put bars at the precise locations on it. After some mathematics, we derived the equation for the x and y coordinates for every month. The rest of the work, including interactivity, was straightforward.

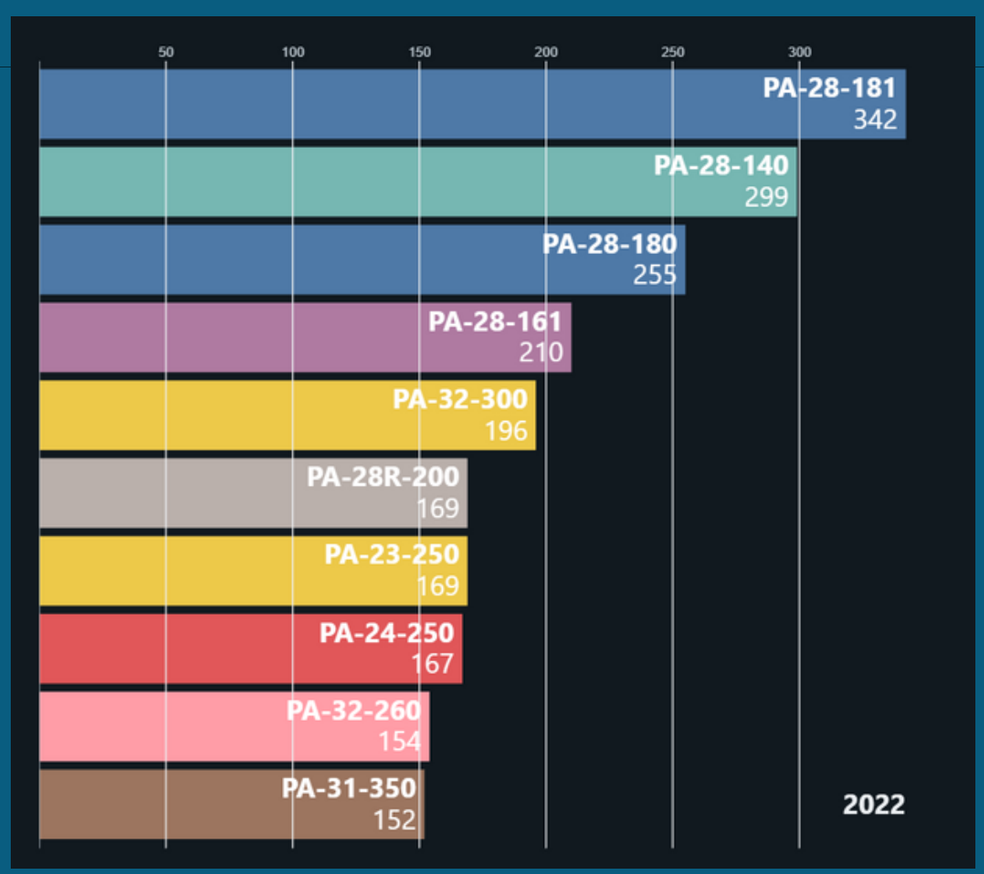
BAR RACE

INSPIRATION

It is always very **engaging** to look at a bar race video of any kind. A viewer can effectively see changes that happened over time in multiple entities. That is exactly what we needed when we wanted to show the differences across airplane models. It perfectly presents the **temporal trends of multiple airline companies**.

DESIGN EVOLUTION

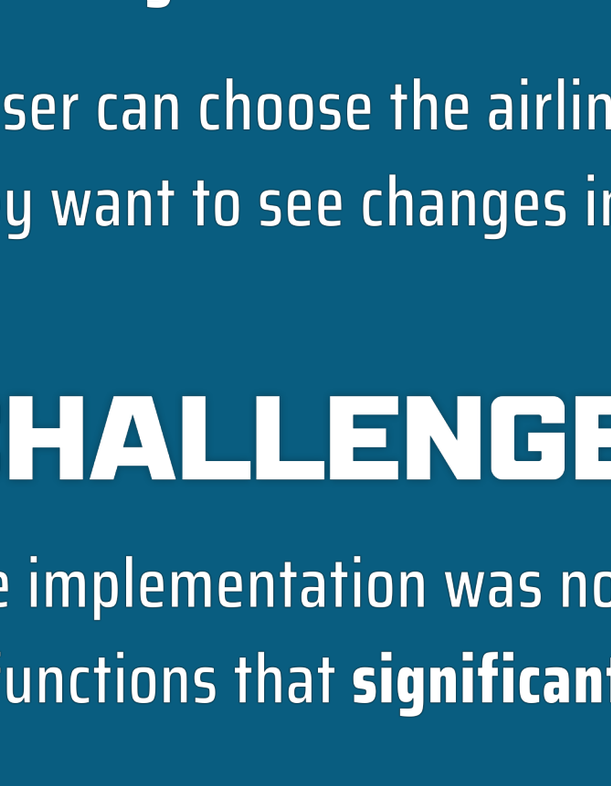
We did not have this visualization in our initial plan. However, we agreed that it can be interesting and useful to show what happened to different airplane models over time. The design itself is not hard, so we implemented it quickly. We decided to use a specific palette to make transitions during the race even more **appealing**.



A user can choose the airline from the combo box. They can also select whether they want to see changes in **the total number of crashes** or **death toll**.

CHALLENGES

The implementation was not a problem with this visualization, as there are built-in functions that **significantly alleviate** the task of implementing such a bar race.



However, the thought process of designing a visualization that depicts the temporal trends across dozen of models was challenging. We tried our best to make it **interesting** for the user. In addition, we had to be careful with the speed of transition when an overtake happens. If it happens too fast, a user cannot follow, if it happens too slow, multiple overtakes can happen at a time.

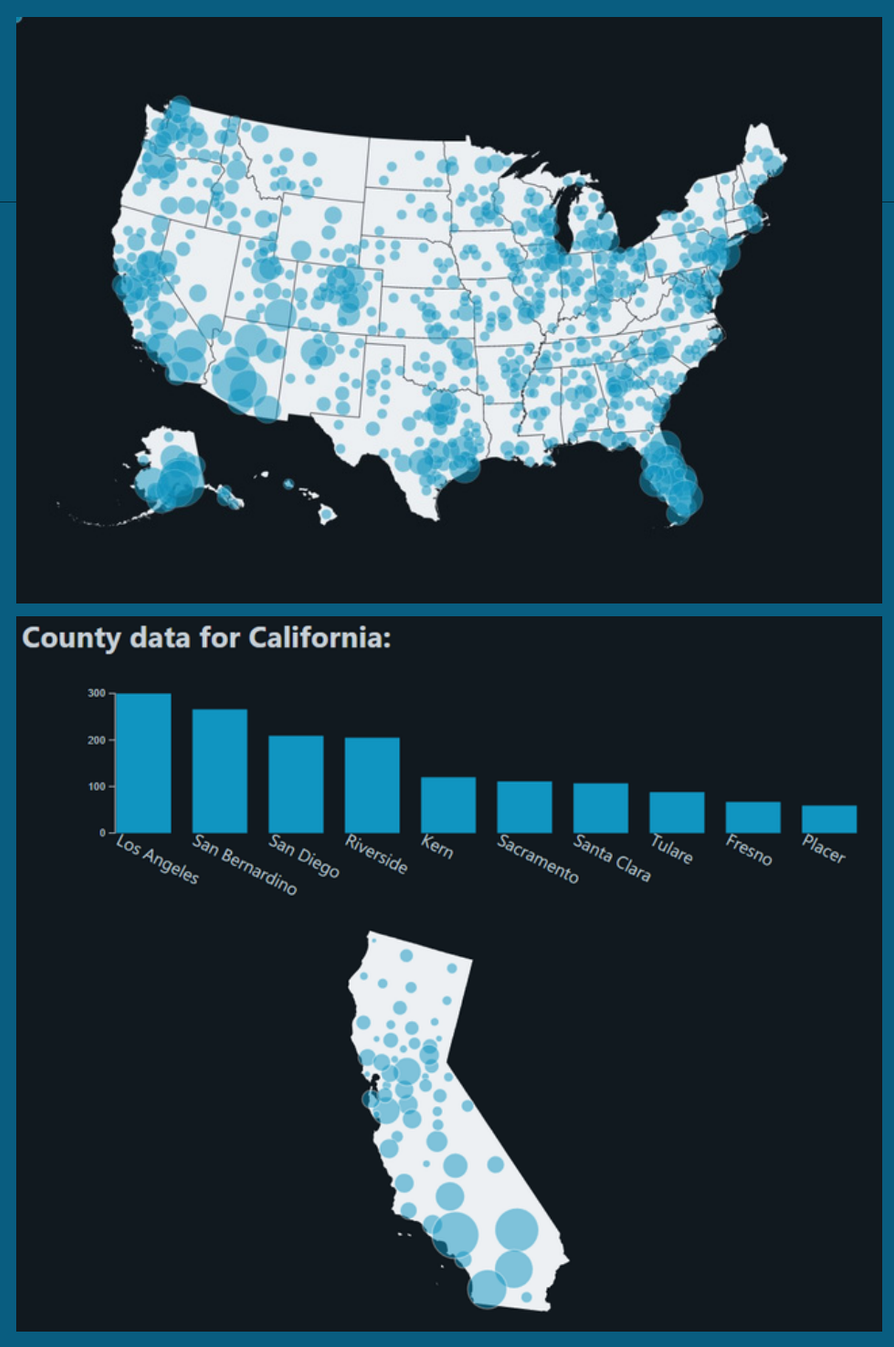
BUBBLE MAP

INSPIRATION

To present the **geospatial** aspects of our data we knew that we need to implement a US map. The main source of inspiration was the lectures throughout the course, especially **Lecture 8: Maps**. We knew exactly how we wanted this visualization to look, so we designed it right away. We noticed that it was almost impossible to see all the bubbles across the whole US. Therefore, we used detailed borders of the states making zooming in and out fluid. When a user clicks on a state, we even show the distribution of the top ten counties with the most crashes.

DESIGN EVOLUTION

Initially, our goal was to create a standard bubble map. After implementing it, we realized that we can tell a more interesting and compelling story by adding **more interactivity**. As we wanted to explore spatial trends with more details, we introduced **zoom-in and zoom-out** to the states. We felt that the zoomed-in state was missing a piece, so we designed the bar plot of the air crash per county distribution. That completed our story of spatial visualization.



CHALLENGES

This was definitely our most challenging visualization. We had to deal with large **GeoJSON** files that contains precise state borders. Map rendering was an expensive operation so we had to be extra careful not to make a bad user experience.

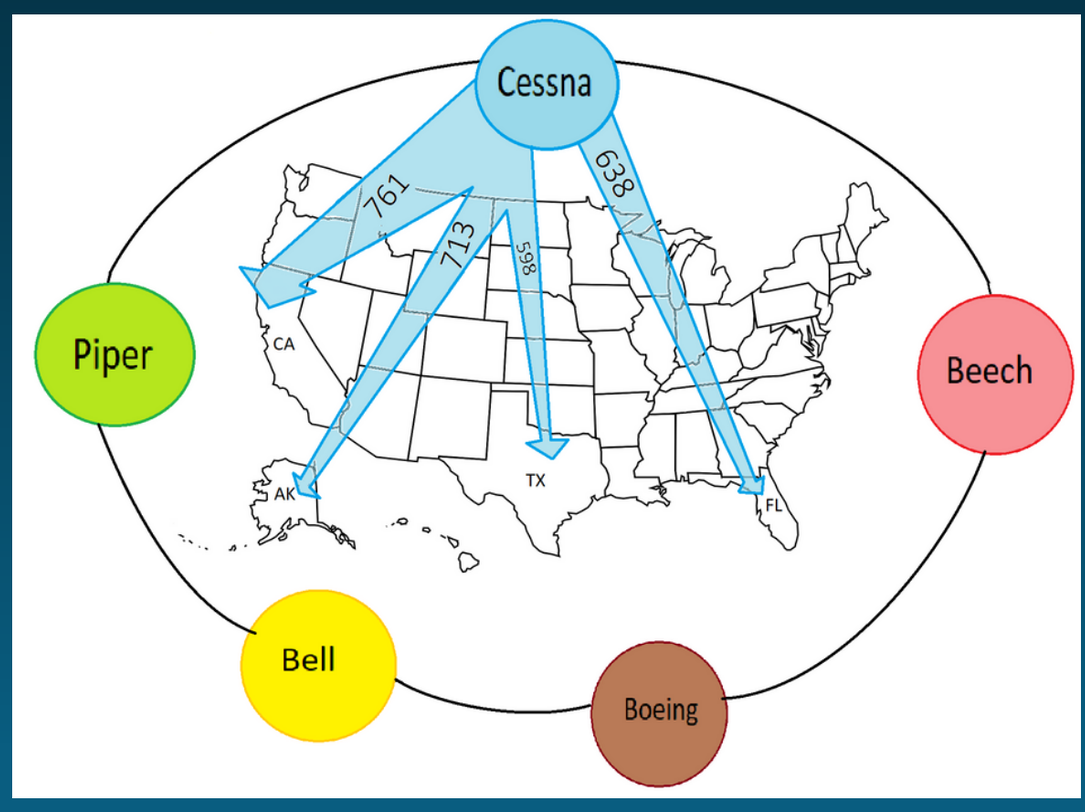
Zooming was the next major milestone. We had to rerender the circles when a state is clicked on since their size is relative to the others. We had to **refactor our initial code and adapt it**.

Finally, it was hard for us to create an idea of how to **complete** the zoomed state visualization. The county distribution was not a hard solution to implement but a hard one to come up with.

AIRLINES PER STATE

INSPIRATION

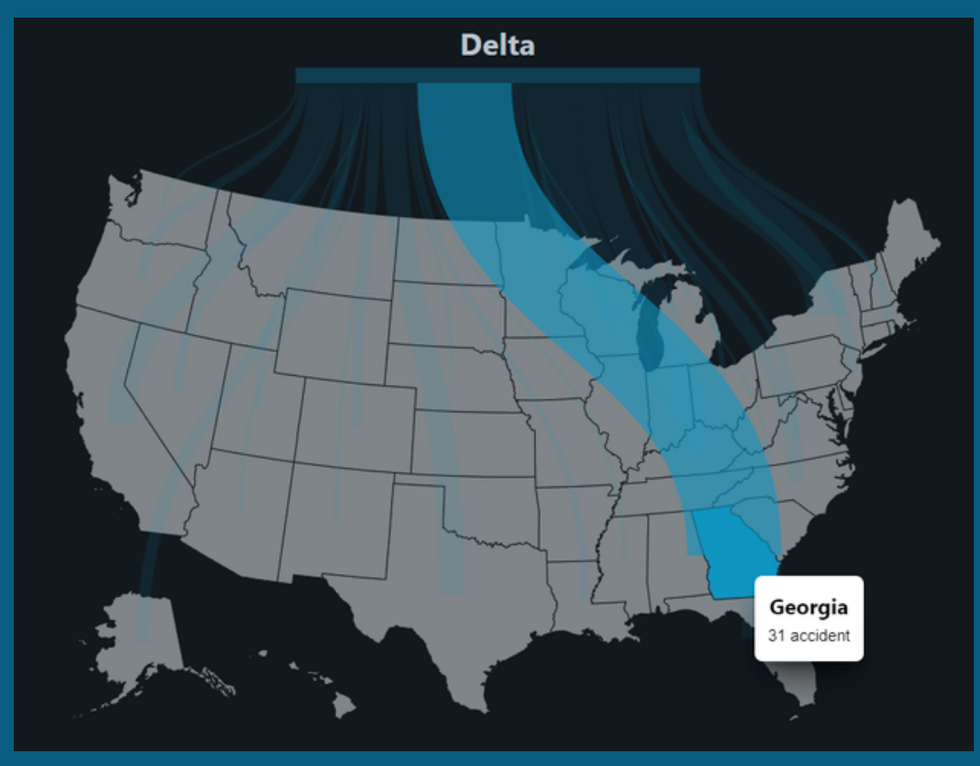
One of the most fascinating visualizations ever made is Minard's Visualization Of Napoleon's 1812 March. It was the first time when we realized that not just the color, points, or direction of a line can show the information, but the **width** as well. During the course, we saw this idea multiple times so we decided to implement it ourselves.



As we have previously shown temporal trends across different airlines, now we want to do the same for **geospatial**. We started with the simple scratch.

DESIGN EVOLUTION

It was hard to get started. Firstly, we decided to fix one airline since more of them make the visualization **messy**. You can choose it from the combo box. Next, we had to implement lines with varying widths. We did not like the straight lines so we redesigned them to have **curves**. With many lines over the map, we decided to point out only the one that leads to a state that is **hovered**. You can see the final result.



CHALLENGES

We had the idea from the start, but the **implementation was hard**. It took a lot of time to figure out how to make curvy lines. We also had to be careful with the starting point of the line for every state. Imagine if the beginning of the lines connecting states on the East were on the left, they would all overlap with the ones connecting the Western states.

WHAT HAVE WE LEARNED?

This project gave us a chance to practice presenting data in an **interactive and engaging way**. Previously, we mainly focused on the technical aspects of data, statistics, and machine learning. However, this course allowed us to shift our perspective and concentrate on the **artistic** side. It helped us discover new, effective, and creative approaches when working with data.

HOW TO IMPROVE?

It is possible to modify the visualizations to ensure they can be viewed on **mobile devices**. While they function well on larger screens, adapting them for mobile requires a complete overhaul of their layout. Creating a mobile-friendly layout involves more than just resizing components - it requires a **thorough redesign**.

We all contributed to every part of this project. Whether it was implementation, design, or improvement, we shared our ideas and work, as a **great team** does!

Andrija Jelenkovic

- implemented 'Condegram'
- designed the histogram in 'Bubble map'
- website structure and deployment
- text for Process book

Aleksa Milisavljevic

- developed the idea for the 'Condegram'
- created the 'Bar race' and 'Airlines per state'
- improved the interactivity of 'Airlines per state'
- text for Milestone 1, Milestone 2 and Process book

Lazar Radojevic

- preprocessed data
- built the 'Bubble map' with zooming feature
- redesigned the 'Bar race'
- designed the Process book