

SFO Air Traffic Analysis



This visualization project aims to analyze the patterns and trends of air travel departing from San Francisco International Airport, with a focus on the changes in passenger volume over time for various airlines and geographic regions.



STEP 1 : Milestone I

After choosing the dataset and the basic idea of the project, we did a bit of basic preprocessing to understand the data we will be working with. The next step was to do a brainstorming to conceive a view of the final product we will be delivering.

The first idea which popped off was to implement a heatmap which displays the different regions of the world, color-coded based on the number of flights departing from San Francisco to each region for the whole time period.

STEP 2 : Milestone II

The next step was to create a basic skeleton of our website with delimited areas for our future visualizations and so some sketches for the visualisations we want to implement.

At that time, we had a more precise idea of what our final result would look like. We wanted to implement 4 visualizations: **Air Traffic Heatmap, Total Number of Passengers Sankey Diagram, Ticket Fare Pie Chart** and a **Number of Passengers Chart**.

We also began searching for tools which will help putting these visualisations in place and localizing the best practices and teachings from the course content.

STEP 3 : Milestone III

We first decided to split the work so everyone could work individually and at their own pace. Also, after getting the feedback from Milestone 2, we decided to abandon the idea of an interactive pie chart for the ticket fares because it's difficult for the human eye to estimate quantities from angles only.

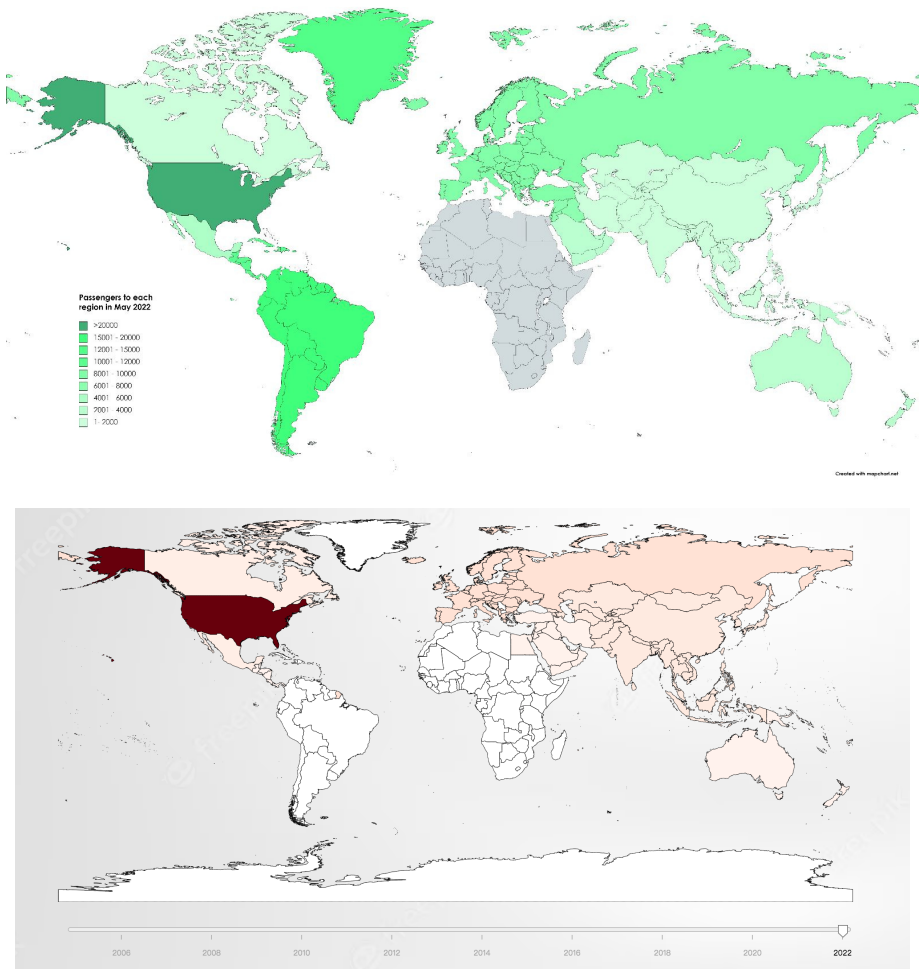
Instead of that, we decided to create an interactive bar chart showing the destinations across the months. We also decided to change the website skeleton that we had for Milestone 2 to have a cleaner website. Now that we had a clear and precise idea of what we wanted we could all start coding.

The discussions and brainstorming allowed to get to the desired final product, we could then concentrate on rendering this process book and the screencast.

Final Product

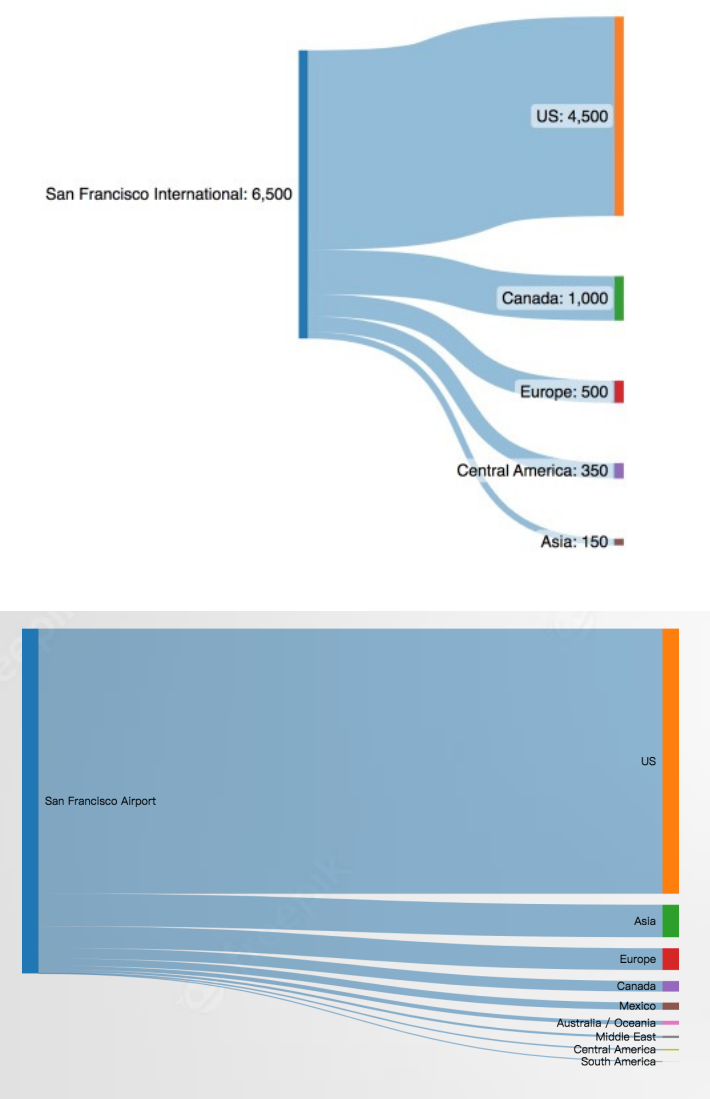
Let's us now take a look at the final visualisations we put in place and compare them to the initial sketches. First we see the sketch and then the final product just beneath it.

- Air Traffic Heatmap**



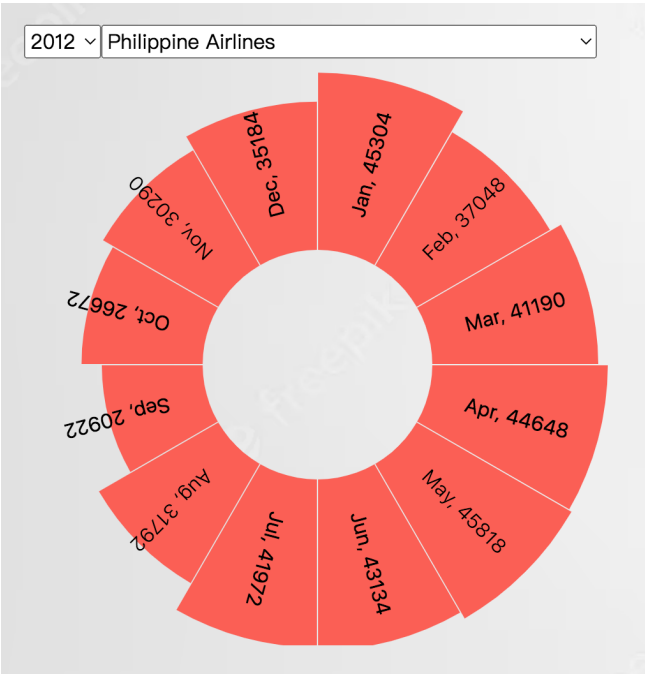
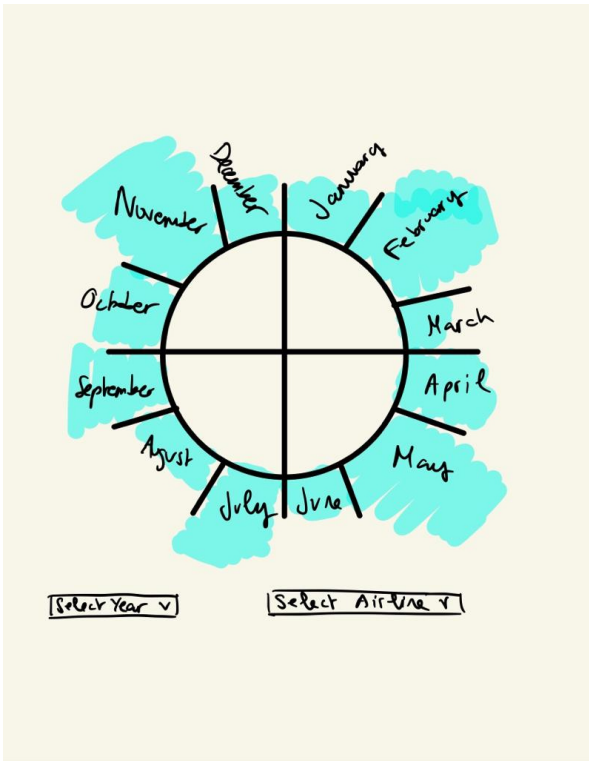
I has been decided that the heatmap would be the first thing that user sees when entering the website. It's a clear visualizations and gives a direct idea of where you can directly fly to from the San Francisco airport. It includes a time slider to navigate through the years. The passenger volume density is represented in a red color gradient.

- **Total Number of Passengers Sankey Diagram**



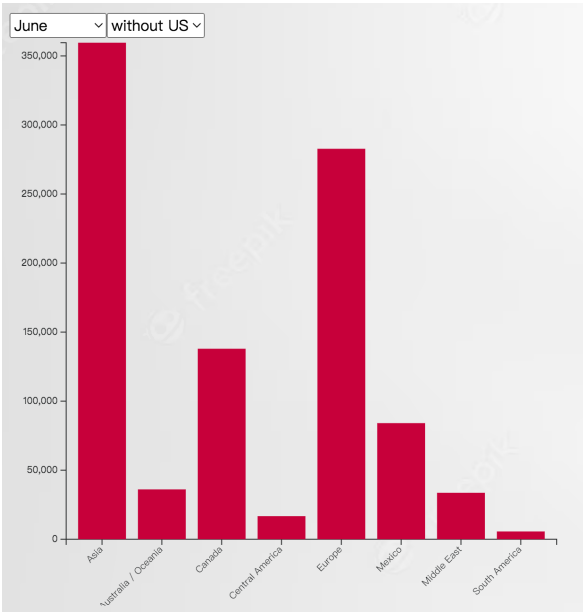
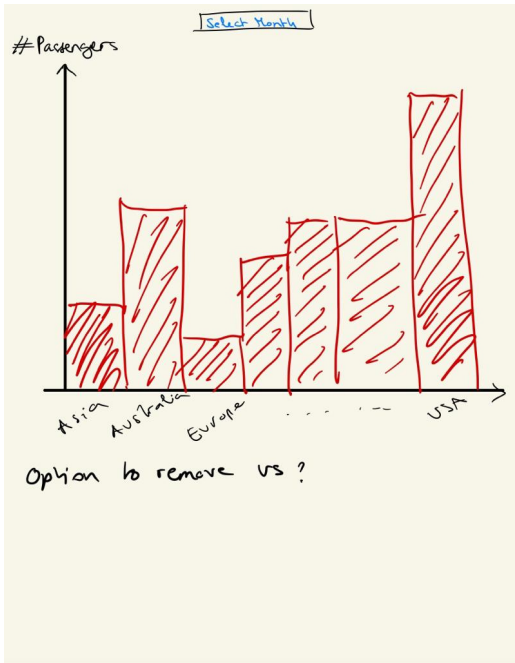
The Sankey diagram is clear way of showing the air traffic flow departing from the San Francisco airport as a global view. When mouse hovering over a link or a target node, we can see the actual total number of passengers to the destination region.

- Passengers Circular Bar Chart



The sketch from Milestone II has been updated to have a more concrete view of the aimed visualisation. The main idea is to see the number of passengers for a selected year and airline for every month. This allows, for example, the users to see the impact of COVID on air traffic. Some of the airlines don't have flights on certain months but we decided to keep all the data so that the users can see whether their target airline flew on that specific month or not.

• New visualisation : Passengers Per Destination Region Bar Chart



As a complement to the heatmap and the sankey, here we have the passenger flow in a more granular way. We can see the top destinations for each month, averaged across the year to highlight seasonal trends.

We noticed that the number that travelled within the US (domestic flights) was significantly higher than for other parts of the world so we decided to add an option to keep or remove the US to be able to better compare the number of passengers across regions.

Challenges

The dataset we chose was very straightforward and didn't need a lot of preprocessing. The whole data has been processed with Pandas and put into several ready to use csv files for the different visualisations. However, because the data that we have is only from one airport we needed to take that into account for our visualizations and our explanations. For example, for the bar chart and the heat map we don't have data of the exact country of destination but only the region.

The main challenge that we faced was to understand how the d3.js library worked because we used it for all our visualizations. Also, we decided to split the work so everyone did one (or two) visualization and when we had to put everything together we needed to be careful with all the dependencies between our code snippets. Coding in html and javascript is a first for us so it was quite an interesting challenge.

Peer Assessment

Sara	<ul style="list-style-type: none">• Circular bart chart, top Destinations Bar Chart• Website code merging• Process book
Mahdi	<ul style="list-style-type: none">• Heatmap and Sankey Diagram• Website design and rendering• Process book
David	<ul style="list-style-type: none">• Heatmap• Screencast

Enjoy the website 😊