CS416 Final Project Report

Name: Gabriel De Mesa

Project Url: https://jamesjellow.github.io/cs416_final_project/index.html

Repository: https://github.com/jamesjellow/cs416_final_project

Overview

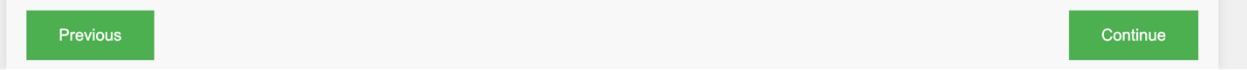
This project was built using only html, css, js, d3, and the d3 annotations library. The dataset chosen for this project is from Kaggle and encompasses flight data from the years 2003 to 2023. The purpose of this narrative visualization is to provide an interactive slideshow for readers to explore at their own pace.

Messaging

The message I am trying to communicate to the reader is that right from the beginning I want them to understand the trends of airline travel over time. This starts off very high level with an overview of total airline travel sorted by type and drills down progressively as each slide progresses. The goal is to help the reader understand airline travel within that recent 20-year period and to make them aware of trends, outliers, and patterns in the data.

Narrative Structure

The structure this narrative visualization was designed to follow was the interactive slide show narrative structure. My visualization supports this structure by firstly not overwhelming the reader upfront and providing them with next and previous buttons so that they can progressively work through each slide at their own pace.



Whilst they progress through the slides, they can hover their mouse over each datapoint and chart to get more details “on-demand” such that they can “drill-down” into the data. Additionally, the presentation follows a lineage of complexity such that the first being the least complex and the last slide being the most in-depth and detailed explanation and data analysis of the trends within the data. Reading through my narrative visualization is like that

of building a Lego set where you start with the base understanding and then work your up to the climax where you put all the pieces together.

Visual Structure

Scene zero is the overview and introductory page briefing the user on the visual journey they are about to embark on.

Scene one uses a line chart. This structure ensures the viewer can understand the data because not only is it coordinated but also has annotates that help highlight for example the massive dip in flights during the 2020 pandemic. It helps the subsequent scenes because it provides a great foundation and overview of knowledge to go in more detail in the later scenes.

Scene two uses a heatmap. This method is effective because the viewer can visualize the different months and years that were the most and least travelled. It helps the viewer transition to the next scene because we are able to convey to the viewer that there are patterns that are starting to emerge from the dataset and it starts to get the reader more engaged and start thinking analytically.

Scene three uses a triple stacked bar chart. This is the most data-rich as well as the most conclusions we can draw from hence the the in-depth reasoning and detail on that page. The very detailed descriptions, hover text, axis labels, and text help set the scene for this page as it can be very daunting to try to tackle the visualization without proper context from the previous scenes and the description. This is why I saved the most complex visualization for last.

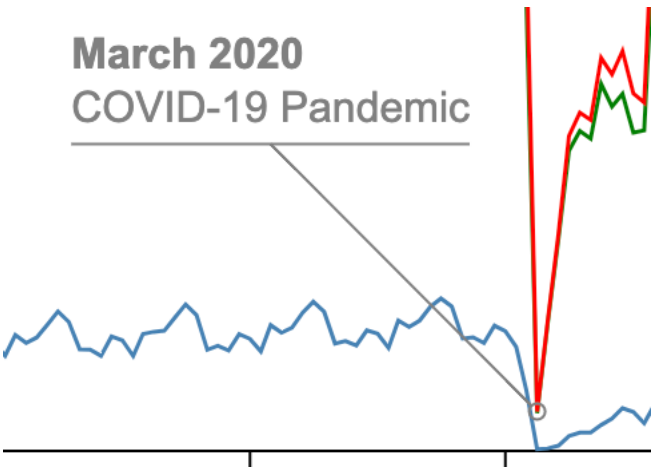
Scenes

1. Scene 0: Introduction
    a. The purpose of this scene is to welcome the viewer as well as give a brief overview of what is to come down the road.
2. Scene 1: Air Traffic Trends
    a. The second scene is a line chart depicting the various types of travel and a brief overview of the trends for that travel from 2003-2023. The purpose is to ensure the reader has a good background and foundation for what is to follow.
3. Scene 2: Heatmap of Total Travel per Year-Month

    a. The purpose of this scene is to dive 1 level deeper and drill into the per month and year travel to analyze which months and years were the most and least travelled. This helps the reader start thinking analytically and prepare themselves for the last and final scene.
4. Scene 3: Detailed Triple-Stack Bar Chart
    a. The most in-depth and in-detail of the scenes, the purpose is to dig to the lowest level of the data and really try to extract all we can in terms of knowledge and various categories to see what trends and patterns we can realize from this data. This scene also allows the user to change the type of data whether that be domestic, international, or total passenger flight data. This is the last scene because it is the most in-depth and the most detail-oriented of the set of scenes.

Annotations

The annotations in this narrative visualization follow the d3-annotation CalloutCircle template. This template is designed for highlighting specific data points and events on a chart, using a combination of text and graphical elements (circles and arrows). The CalloutCircle template is excellent for drawing attention to specific data points without cluttering the visualization. The use of circles around key points, combined with explanatory text, helps to clearly highlight and explain important events or anomalies in the data which in this case was the steep decline and rise of flights during the Covid-19 pandemic. These annotations do not change however they are hard to miss and provide greater clarity to the viewer because it helps point out patterns and key events in the data.



Parameters

The states of the narrative visualization are the different configurations or conditions that the visualization can be in, based on the parameters. The primary states include:

1. Initial State
    a. This is the initial state of the visualization without any prior filters, parameters loaded, or things highlighted
2. Filtered by Travel Category
    a. This usually happens on the last step when a viewer filters the data by the travel category (Domestic, International, Total)
3. Time-Filtered State
    a. During the heat-map phase, we are exploring the different time states in a grid like format seeing which months and years had the most travel
4. Metric focused state
    a. In scenes 1, we are focusing on the metrics and trends of the passengers over time and in scene 3, we are setting up the viewer to focus on the linear correlations between the revenue, load, and seating.

Triggers

1. Change of Scene
    a. This happens when viewer clicks either the “Next” or “Previous” button to either progress or go back from the scene they are currently on to the scene they are desiring to view.
2. Filtering
    a. On the last slide, the user can click buttons corresponding to which filter they want to apply and see the data reflect that filter as well as the label on the graph reminding them which filter is currently being applied.

3. Pop-Ups

a. Upon the viewer hovering over a data point, they can get more information on-demand resulting in a pop-up being shown with additional details about the specifics of that datapoint.