Website: [https://heyhaiden.github.io/] Github Repo: [https://github.com/casa-ucl/casa0017-group-assignment-2022-casa-web-masters]

#Introduction

The aim of this web site is to investigate whether there is any visual or statistical correlation between changes in fine particulate matter (PM25) measurements between the two megacities London and Los Angeles leading up to and during the beginning of the Coronavirus pandemic in March of 2020. Due to the lockdown measures, would a car centric city have seen a more drastic improvement in its air quality than a non car centric city?

In order to test this question, the above two cities were chosen because they have similarities in total population, population density, and total square footage, but major differences in the commute culture and transportation modalities. The car-centric city of Los Angeles ranks highly in car ownership with 2,276,974 registered vehicles in 2019 for a population of just under 3.9 million people during the same time frame. This represents a 58.3% rate of vehicle ownership (European Commission, 2021). A non car centric city, London, with 9 million, boasted a much smaller 49.1% rate of vehicle ownership inhabitants during the same timeframe (European Commission, 2021). This hypothesis was posed due to the perceived negative impact of private vehicles on PM 2.5 measurements, which are composed of dissolved droplets of pollutants and other inhalants dangerous to human health (USC Environmental Health Centers , 2021).

The lockdown policies differed slightly between the two cities Los Angeles and London and that was taken into account with the data visualisation and analysis. Los Angeles issued a statewide stay at home order on March 19th, which included the shutting of schools, religious services, bars, gyms, and public or government property (Kandel, 2020). For London, the first lockdown was announced on the 16th of March and enforced beginning March 23rd, 2020 (Scott, 2021).

##Context of Air pollution component

Air pollution consists of many forms of natural and artificial pollutants. The five main sources are ozone, carbon monoxide (CO), nitrogen dioxide (NO2) and sulphur dioxide (SO2) and all fine particulate matter PM 2.5 is one of the main components of air pollution that is responsible for breathability problems. PM 2.5 consists of dirt, metals, mould, and smoke, which are released from industrial activities, cars and even the source of fireplaces. These are partly responsible for the health diseases that relate to the lungs, brain, and pancreas. The diseases have long-term effects on health.

The reason that the PM 2.5 is detrimental is that its size is 30 times smaller than the human hair. That PM 2.5 It is very easy to breathe it and then from the lungs it enters the blood system (World Health Organization, n.d.).

##The method of PM 2.5 measurement

The method of measuring the amount of PM 2.5 in the air is by the concentration of particular matters that are in the air. The amount of concentration is measured by satellites or lasers that count all the components that the air contains (UN Environment Program, 2022). The meaning of 1 µg/m3 is one cubic metre of air that includes between 6 to 10 micrograms of particulate matter of pollution. Moreover, the way that reports and websites show the amount of the PM 2.5 in the air is by the average concentration in the last 24 hours minimum. However, there is a differentiation between the duration of minimum time measurements among countries and companies (Bliss Air, n.d.).

The index of measures PM 2.5 is divided into 6 degrees of PM 2.5 levels which are: Good, Moderate, Unhealthy for sensitive groups, Unhealthy, Very Unhealthy and Hazardous. The scale is from 0 amount of PM 2.5 in the air which is the cleanest to 500 which is the highest amount of concentration of PM 2.5 (Air Quality Index, n.d.).

#Design

##Inspiration

The inspiration for the design of this website came primarily from map based narrative data visualisation projects and dashboards. The first of these was the live bike- sharing visualisation shown in class. Additional resources were ESRI's Story Map Swipe and Spyglass Gallery and the New York Times "'Hell on Earth': The First 12 Hours of California's Deadliest Wildfire" (Esri, n.d.) (The New York Times, 2018). These both used aerial and satellite imagery to provide visual representation of a large dataset evolving over time, which made larger trends easier to comprehend. ESRI was of particular interest because it allowed for the direct comparison of two maps within the same visual interface. Although it was a paid service, this was the visual template our data dashboard was based on.

The research topic was chosen because it combined readily available datasets with the ability to visually portray them over a large surface area. Adding a scrollable story element would allow a user to easily step through the data and understand at their own pace. They were chosen as visual templates due to the ease of use and clarity of data presented. The research topic was chosen first, emerging from previous experience on the topic of air quality, and subsequently assisted in the decision to display a heatmap as it appeared to be an effective way to contrast air quality between these two cities.

##Wider Context

There are many ongoing academic and government-funded studies that examine live air quality data, including the EPA's Interactive Map of Air Quality Monitors and California Air Resources Board (EPA, 2022). However, these data visualisations do not offer a narrative component to further contextualise historical data readings, rendering them helpful in some scientific aspects but deficient in effectively communicating hypotheses or a call to action to the general viewer.

##Methodology

The methodology for testing this hypothesis consisted of analysing historical PM2.5 readings reported by ground-level AQI stations from London and Los Angeles between the months of January and June 2020 - allowing for the long-term observation of air quality before and after lockdown measures were put into place during March 2020. This data was downloaded as a .csv from AQICN's Worldwide COVID-19 dataset and converted to JSON. (Aqicn, n.d.) (Flatfile csvjson, n.d.)

Additional files of Los Angeles County and London's physical perimeters were downloaded from an open source geoJSON Github project and overlaid as tilesets within the Mapbox Studio editor (Specious, 2016).

##User Experience Journey

In order to effectively get a point across with this website, consideration was given to how the information was presented and to walk a user through the website and its research in such a way that would allow them to experience the website for themselves - at their own pace. This involves a slow walk-through upon first visiting the website that details research about air quality, the purpose of the website, and how to use it. By

providing this background to the user immediately upon their entrance to the website, it ensures that any user - regardless of their existing knowledge of air quality - can fully experience the website, and take away something valuable. After this, the user is presented with a Dashboard page where they can experience the information more fully by interacting with sliders that visually represent the difference in air quality between these two cities on the same days. The user is also presented with a Teams page where they can learn more about our team and experience, and a resources page where they can take a deeper dive into the research that fueled our website and possibly develop a further understanding in the field of air quality.

##Storyboards, Wireframes & Prototypes

Prior to beginning development on this website, prototypes were sketched out by hand and then built using Figma. This allowed for a rough detailing of the website hierarchy and its functionality, including the general organisation and layout of the data dashboard. The page breakdown was determined so that the index.html would open to a scrolling story, with additional tabs to include a dashboard for comparing air quality over time, a team page, and a resources page to provide more information about the research that is covered by our website.

HTML, CSS, and Bootstrap elements were used to create a working prototype of these designs. They relied on the implementation of several rows and columns to keep all elements organised and easy to read. A header and footer were designed with these elements and subsequently implemented across all pages using Javascript and jQuery to prevent the need for duplication of their code and in order to make any changes to them easier to accomplish.

#Development

##Collection & Handling of Data

Currently, there is no shortage of weather and air quality APIs, but it is difficult to source granular, historical data without having to pay for location-specific calls. OpenWeatherMap integrates nicely with most mapping platforms, but the aforementioned reason made it difficult to justify for this project.

Therefore, The World Air Quality Index was chosen as the primary dataset. The entire project is open source, and COVID specific data was available as a city-specific .csv download, eliminating the need to make calls to an API (Waqi, n.d.).

Due to the historical, restricted scope of this project, the datasets integrated into the map visualisation were stored locally as .csv and .json files to remove the dependency on a live API, ensuring continuous uptime and minimal upkeep.

##Mapping & Data Visualization

Several mapping platforms and APIs were compared before settling on MapBox (Mapbox, n.d.). This tool allowed for the most flexibility in terms of building a scrollable, interactive story with customizable coordinate locations, as well as unlimited data overlays, including integrations with several well known open data sources. The storytelling template from Mapbox was referenced for our home page implementation, which included an index.html file and config.js file to act as a database for map positioning, styling, tile overlays, and scroll animations, which relied on a JavaScript library called Scrollama to simplify these chapters down to step events (Mapbox, n.d.) (Samora, 2017).

The dashboard page showcases a side by side comparison map of the cities Los Angeles and London, including a stepped timeline slider and a graphical comparison of PM2.5 readings during the same timeframe. Highcharts was used in conjunction with CodePen to stylize and animate the .csv data, then integrated into the webpage using Bootstrap.

##Interactivity

The website was designed for readability and understanding across both desktop and mobile. Scrolling is the primary user interaction, which allows for a measured delivery of information and illusion of narrative flow. The primary map animates based on the speed of user input, positioning the viewer at relevant coordinates while context is given through the delivery of title cards and overlaid images.

On the dashboard, a timeline slider was implemented to allow the user to change and compare dates - allowing users to spend more time on certain dates which may be important to answering their own questions about the information provided. This timeline slider is connected to the maps using javascript and depending on the date (and slider index) will present minimum, median, and maximum PPM values in statistic bars located below each map. Similarly, the maps' location will change colour using a transparent overlay whose colour is dependent on the median PPM value for that day in each respective location. All of this can be interacted with via the slider. This colour coordination aligns with the charts located below the timeline slider. These charts are a graphical representation of the same data and provide basic levels of interactivity. The user is able to single out specific city data by selecting it below the x-axis. This provides a different way for the users to compare data between the two cities.

#Conclusion

##Difficulties & Solutions

There were several difficulties encountered during the development of the website. Primarily, the heatmap visualisation proved difficult to complete because it was dependent on using city wide data arrays to show daily change over a sizable area. Unfortunately, our primary data source restricted access to the specific coordinates of the weather stations in the API, instead giving daily ranges of values including the minimum, maximum, and mean value. An attempt was made to supplement this dataset with another resource, but most ended up being behind a paywall. An arbitrary array of values could have been used to visualise this on the scrolling map on the home page, but we ran into time and resourcing issues. Therefore, a basic colour overlay was used over the maps in the dashboard as a stand in for the heat map visualisation.

Several weather APIs including Breezometer and Plume Labs boasted stunning visual representation, including streamlined integration into Mapbox, but both are paid services that were outside of the scope of this project (Breezometer, n.d.) (PlumeLabs, n.d.).

It became clear early on in the project that while there is an ever growing list of mapping platforms and visualisation tools, not all solutions use the same file formats and input data, especially within the world of satellite imagery. Dataset cleaning became a timely initiative that inhibited the integration of additional sources for the completion of the project. Mapbox is an extremely robust tool that allows for the amalgamation of several different types of data arrays, but many errors occurred while converting JSON files to geoJSON. In the future, more effort will be put into sourcing compatible datasets at the origination of the project.

##Improvements

In further iterations of this website, there are many aspects that could be improved upon. First and foremost the visual design of the user interface could be streamlined and improved. Another relevant addition would be a dropdown menu on the dashboard page listing different cities throughout the world. This would allow users to compare more than just London and Los Angeles - providing them with more value and a greater understanding of lockdown's effect on air quality globally. Swapping the maps' coloured overlay with a heatmap would be a great inclusion that should be possible by finding similar air quality data that also includes geographic coordinates. Additional overlays could be added to the scroll map as well, including a static heatmap overlay, population density, points indicating where weather station sites are located, and more.

The implementation of real time air quality index readings along with the ability to view historical ones would provide increased value to the user and provide a way for the website to continue growing - accumulating captured data which can subsequently be stored and added to the historic section of the website. An animation over the dashboard when the user first arrives to that page could be used to provide instruction to the user and ensure proper use of the resources that the website provides - it could also highlight certain points in the data or time periods that relate back to our research question. Finally, a comparison tool that is visual and primarily colour based is not very accessible to those that are visually-impaired. Although it's possible to read the air quality readings in certain parts of the dashboard, adding other methods of accessibility to our website, such as a video discussing our research and conclusions, would go a long way to helping those people interact with our website and research.

Lastly, if given more time, the datasets within the codebase could be separated out and added to a database such as mySQL. This would allow for a cleaner codebase and an easier way to call and manipulate data within our dashboard elements. The classroom server that we intended to host on was down in the final weeks of the project and unfortunately another solution was not implemented before the deadline.

##Bibliography

1. Air Quality Index, n.d. About the Air Quality and Pollution Measurement: [Online] Available at: https://aqicn.org/ [Accessed 13 11 2022].
2. Aqicn, n.d. Air Quality Open Data Platform. [Online] Available at: https://aqicn.org/data-platform/covid19 [Accessed 2 11 2022].
3. Bliss Air, n.d. What is PM2.5 and Why You Should Care. [Online] Available at: https://blissair.com/what-is-pm-2-5.htm [Accessed 13 11 2022].
4. Breezometer, n.d. Heatmap Tile Overlay. [Online] Available at: https://docs.breezometer.com/api-documentation/heatmap-tile-overlay/v1/#overview [Accessed 3 11 2022].
5. EPA, 2022. Interactive Map of Air Quality Monitors. [Online] Available at: https://www.epa.gov/outdoor-air-quality-data/interactive-map-air-quality-monitors [Accessed 2 11 2022].
6. Esri, n.d. Browse a curated selection of stories created using the Story Map Swipe and Spyglass app template [Online] Available at: https://storymaps-classic.arcgis.com/en/app-list/swipe-spyglass/gallery/#s=30&md=storymaps-apps:swipe-spyglass
7. European Commission, 2021. EU transport in figures - Statistical Pocketbook 2021. [Online] [Accessed 6 11 2022]. Flatfile csvjson, n.d. CSV or TSV JSON. [Online] Available at: https://csvjson.com/csv2json [Accessed 6 11 2022].
8. Kandel, J., 2020. Timeline: A Look at Key Coronavirus Pandemic Events and Milestones in California. [Online] Available at: https://www.nbclosangeles.com/news/coronavirus/2020-2021-california-coronavirus-pandemic-timeline-key-events/2334100/ [Accessed 9 11 2022].
9. Mapbox, n.d. [Online] Available at: https://www.mapbox.com/ [Accessed 1 11 2022].

10. Mapbox, n.d. Interactive Storytelling. [Online] Available at:
    https://www.mapbox.com/solutions/interactive-storytelling [Accessed 2 11 2022].

11. PlumeLabs, n.d. Air quality map. [Online] Available at: https://air.plumelabs.com/air-quality-map
    [Accessed 4 11 2022].

12. Samora, R., 2017. An Introduction to Scrollama.js. [Online] Available at:
    https://pudding.cool/process/introducing-scrollama/ [Accessed 2 11 2022].

13. Scott, E., 2021. Lockdown 1.0 and the pandemic one year on: What do we know about the impacts?.
    [Online] Available at: https://lordslibrary.parliament.uk/lockdown-1-0-and-the-pandemic-one-year-on-
    what-do-we-know-about-the-impacts/ [Accessed 8 11 2022].

14. Specious, 2016. click_that_hood. [Online] Available at:
    https://github.com/codeforgermany/click_that_hood [Accessed 1 11 2022]. The New York Times, 2018.
    'Hell on Earth': The First 12 Hours of California's Deadliest Wildfire. [Online] Available at:
    https://www.nytimes.com/interactive/2018/11/18/us/california-camp-fire-paradise.html [Accessed 1 11
    2022].

15. UN Environment Program, 2022. How is air quality measured?. [Online] Available
    at:https://www.unep.org/news-and-stories/story/how-air-quality-
    measured#:~:text=Some%20use%20lasers%20to%20scan,nitrogen%20dioxide%20and%20ulfur%20dio
    xide [Accessed 14 11 2022].

16. USC Environmental Health Centers , 2021. PM2.5 PurpleAir Data Visualization Project: Long Beach and
    South/Southeast Los Angeles. [Online] Available at: https://envhealthcenters.usc.edu/pm-2-5-shiny
    [Accessed 8 11 2022].

17. Waqi, n.d. World's Air Pollution: Real-time Air Quality Index. [Online] Available at: http://waqi.info/
    [Accessed 1 11 2022].

18. World Health Organization, n.d. Air quality and health. [Online] Available at:
    https://www.who.int/teams/environment-climate-change-and-health/air quality-and health/health-
    impacts/types-of-pollutants [Accessed 13 11 2022].

Declaration of Authorship

We, Gil Adda, Andre Bourgeois, and Haiden McGill confirm that the work presented in this assessment is my own. Where information has been derived from other sources, we confirm that this has been indicated in the work.

16 November 2022