7CUSMTSD

Telling Stories with Data

GROUP COURSEWORK 2024/25

GROUP 7

HARA GEORGIOU

HAMZA LATIF

RAHAT AMIT HAIDER

YIFAN WANG

Contents

Summary

The report addressed eight questions, each exploring distinct aspects of urban data variables such as traffic, air quality, cycling, safety, and emissions in London. We distributed the questions among team members at random to promote individual effort while preserving group cooperation and fair workload distribution.

Our workflow followed a structured process to ensure data accuracy and insightful analysis:

1. **Dataset Identification**: Each member sourced datasets relevant to their assigned question from credible platforms like TfL Open Data, London Datastore, and Global Clean Air.
2. **Data Cleaning and Preprocessing**: Using Excel, we examined datasets to identify inconsistencies, missing values, and formatting issues. Placeholder values were removed, columns were standardized, and fields were normalised where required to align datasets for meaningful comparison.
3. **Visualization Creation**: We used Tableau to design interactive and intuitive visualisations. Members created initial drafts of their visualisations based on the cleaned datasets, ensuring clarity and relevance to the assigned questions.
4. **Feedback and Refinement**: Completed visualisations were shared with the group for peer review. Feedback focused on enhancing interactivity, readability, and interpretability. Based on this input, we refined the visualisations to better answer the questions.
5. **Integration and Presentation**: The final visualisations were compiled into dashboards, providing a cohesive view of each question's findings and ensuring consistency across the report.

We used the following tools to complete our task:

1. **Excel**:
    a. **Data Cleaning**: Addressed placeholders, resolved inconsistencies, and standardized formatting.
    b. **Preprocessing**: Aggregated data, calculated averages, and reorganised datasets to focus on relevant variables.
2. **Tableau**:
    a. **Visualisation Creation**: Designed diverse charts and maps, including bar-line charts, scatter plots, and choropleth maps, to explore trends, correlations, and anomalies.
    b. **Interactivity**: Enabled filters, hover features, and tooltips, enhancing the usability of dashboards for detailed exploration.
    c. **Dashboards**: Combined multiple visualizations into cohesive narratives to answer the assigned questions comprehensively.

The success of this report was largely due to teamwork. Members were able to exchange thoughts, talk about difficulties, and present their progress during weekly team meetings. The accuracy, impact, and alignment of the visualisations with the report's overall goals were guaranteed by this iterative feedback loop.
We were able to fully address all eight issues by using a systematic approach and efficient methods, offering practical insights into London's urban problems and showcasing the value of data visualisation in urban planning and policymaking.

Contribution Matrix:

	Visualizations	Video	Report
Hamza	Responsible for creating dashboards for Questions 1 and 4 using Tableau. I sourced relevant datasets from multiple platforms and conducted extensive preprocessing. For large datasets, I used Python to remove placeholders and handle missing values efficiently. Smaller datasets were cleaned and formatted using Excel, where I standardized names and ensured consistency across fields.	Recorded dashboards for **Questions 1 and 4**, adding a voiceover to explain the visualizations, interactivity, and key insights. Additionally, edited the videos for both peer review and final submission.	Wrote report sections for **Questions 1 and 4**, detailing the datasets, visualisations, and interpretations. Additionally, contributed to assembling the final report by writing the summary and ensuring consistent formatting and structural coherence throughout.
Amit	Created visualizations for **Questions 2 and 3**. Used the tools Excel, for preprocessing the datasets, and Tableau, to make visualizations. An additional tool was used, Tableau Prep Builder, for **Question 2**, to do additional preprocess and merging of the datasets.	Recorded the visualization and dashboard with voiceover for the **Questions 2 and 3**. Explained type of visualization, and analysis of the results.	Wrote the report answering the **Questions 2 and 3**, describing the datasets used, the visualizations and its findings. Additionally, briefly wrote the full analysis of the dashboard, with suggestions for **Question 3**.
Hara	Responsible for creating dashboards for **Questions 5 and 6** using tableau. Collected data from multiple sources and performed extensive preprocessing, such as cleaning, normalization, and integration, to create an Excel file suitable for analysis.	Recorded dashboards for **Questions 5 and 6** and added a voiceover to explain the visualizations, interactivity, and key insights.	Wrote report sections for **Questions 5 and 6**, including descriptions of datasets, visualizations, and interpretations. Additionally, edited and connected team contributions to ensure structural and stylistic coherence across the report.
Yifan	Responsible for creating the dashboard for **Question 7**, including the visualizations and interactivity in Tableau. Additionally, proposed the research question for **Question 8**, ensuring that the group met the requirement of providing a relevant question for the project.	Recorded dashboards for **Questions 7 and 8** and added a voiceover to explain the visualizations, interactivity, and key insights.	Wrote the report section for **Question 7 and 8**, detailing the datasets used, visualizations created, and the interpretation of results.

Question 1: Can we use traffic congestion as a means of predicting air quality?

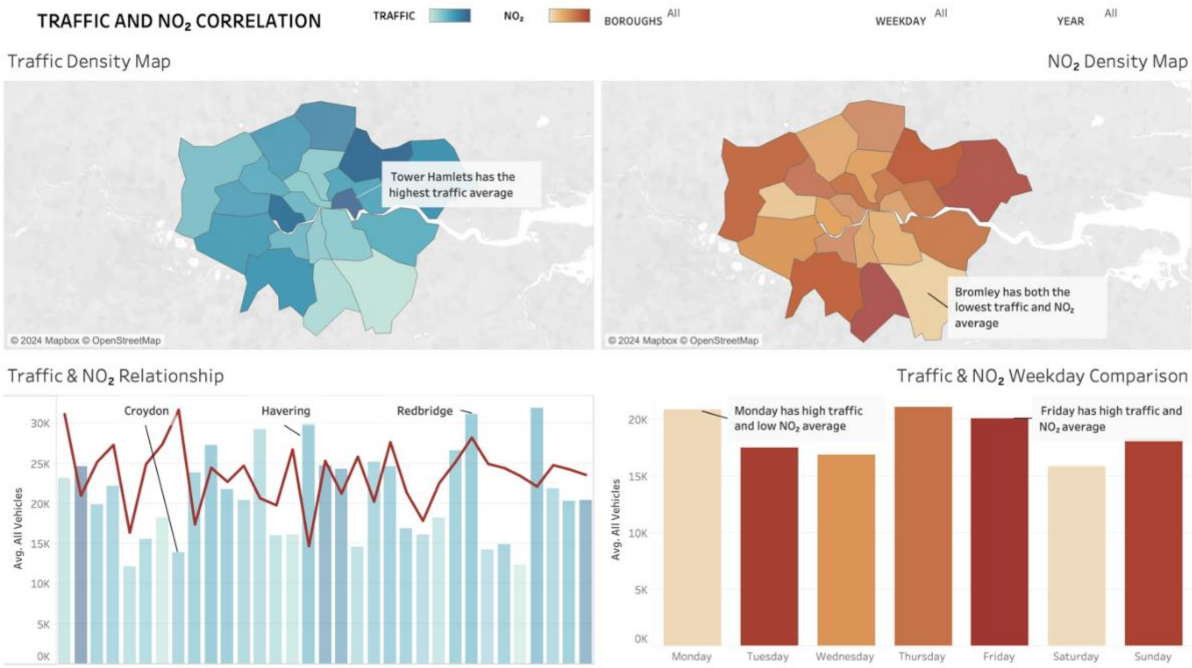

Figure 1: Dashboard answering the question "Can we use traffic congestion as a means of predicting air quality?"

1.1 Datasets Used and Changes Made

Two datasets were used for this analysis. NO_2 data from 2018 to 2020 was sourced from Global Clean Air [1], while traffic volume data per borough came from the Department for Transport [2]. Both datasets required preprocessing due to inconsistencies. In the NO_2 dataset, placeholder values (-999) were removed as they likely represented missing data. Additionally, borough names were not consistent between datasets, leading to potential data mismatches during plotting. These discrepancies were resolved by standardising borough names across both datasets. The NO_2 map only included the years 2018–2020, whereas the traffic dataset contained data from 2000–2023. Therefore, in order to establish a direct correlation between the two variables, the 2018–2020 timeframe was used for both datasets.

1.2 Visualizations Included in the Dashboard and Analysis

The dashboard contains four visualizations: two maps, a bar-line chart, and a bar chart, each offering unique insights:

1. **Traffic Density Map**
   The average traffic density in each London borough is shown on this interactive map; higher traffic volumes are indicated by a darker colour. To filter the dashboard, users can use the filters in the upper-right corner or click on a borough. The name and traffic values of a borough are displayed when you hover over it. "Tower Hamlets" has the highest traffic density, according to the data.
2. **NO_2 Density Map**
   The average NO_2 density each borough is displayed on this map in microgrammes per cubic meter ($\mu g/m^3$), with greater values denoted by darker colours. Although it concentrates on NO_2 levels, it has the same interactive capabilities as the traffic map. The hover and filtering features are still the same. According to the data, "Bromley" has the lowest averages for both traffic density and NO_2.

3. **Traffic and NO_2 Relationship Chart**
   The link between traffic volume and NO_2 density each borough can be seen in a bar-line graphic. The line chart displays NO_2 levels, and bars reflect traffic figures. Users can use the top-right filters or click on a bar or point to filter the dashboard. The chart reveals a positive correlation between traffic and NO_2 values, with some exceptions like ‘Havering’ and ‘Croydon.’
4. **Traffic and NO_2 Weekday Comparison**
   The average traffic on weekdays is compared in this bar chart, where NO_2 levels are represented by colour intensity. By clicking on a bar or using the filter in the upper-right corner, users can filter by weekday. The data indicates that traffic is higher on weekdays and confirms the positive relationship between traffic and NO_2. The highest NO_2 levels occur on Tuesday, Friday, and Sunday.

1.3 Interpretation of the Entire Dashboard

From the dashboard we can suggest that there is a positive correlation between traffic volumes and NO_2 levels. Boroughs like ‘Redbridge’ and ‘Bromley’ exemplify this relationship. However, anomalies such as ‘Croydon’ and ‘Havering’ stand out. Despite low traffic volumes, ‘Croydon’ exhibits high NO_2 levels, while ‘Havering’ shows the opposite, with high traffic but low NO_2 levels. Similarly, weekday data reveals patterns where Thursday, Friday, and Sunday typically see high traffic and NO_2 levels. Yet, Monday is an exception, with high traffic not correlating to high NO_2 levels.
In conclusion, traffic data can serve as a predictor for air quality, although specific anomalies highlight the need for further investigation into additional influencing factors.

Question 2: What is the relationship between air quality and cycling activities in London?

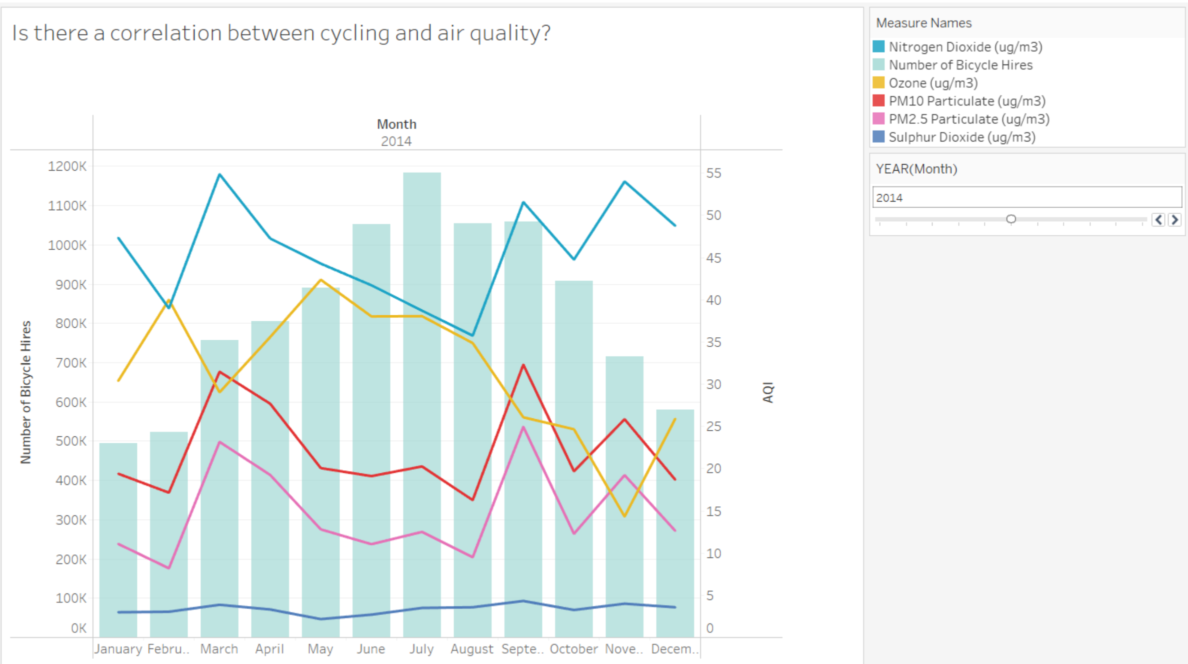

Figure 2 : Visualization answering the question "What is the relationship between air quality and cycling activities in London?"

2.1 Datasets Used and Changes Made

Two datasets have been used, one is for the air quality, while the other one is used for cycling data. Both datasets were retrieved from London Datastore [3]. For the air quality dataset, the columns representing each air quality for two different parts of London have been merged into one, by calculating the averages resulting in 5 columns showing the values of air quality for the whole city of London.

For the cycling usage data, a dataset containing the Santander cycle hire dataset has been used [4]. The file contained different tables, one for daily values, one for monthly values, and for yearly values. For the visualization, I wanted to represent only the monthly values for the cycle hires, across different years, therefore the table containing those values was extracted into a new spreadsheet, so it would be easier to import into Tableau.

Using Tableau Prep Builder, the two datasets have been combined into one, resulting in a table showing monthly values for air quality and cycling data in London for different years.

2.2 Visualization and Analysis

Using the merged dataset, the visualization illustrated in figure 2 was created; **a bar chart combined with a line chart**. The bar chart is indicating the number of cycles hired during the year, while the line chart is indicating the level of air quality during that period. Each line is representing an indicator of air quality (Nitrogen dioxide, Ozone, PM2.5, PM10, and Sulphur dioxide).

From the visualization, we can assume that there is some correlation between the usage of cycles and air quality. For example, in the year 2014, we can see that from March till July, there is an increase in cycling, and, at the same time, the air quality improves, but the moment the cycling usage decreases, pollution is rising again. Thus, showing that there is a correlation between these two variables. Furthermore, we can also determine the cause of the cycling number increase and decrease. As mentioned above, from March to July, there is an increase in the number of people cycling, this is most likely due to the weather in London getting warmer, since it's Spring and Summer time, and as we get closer to September, which is the period where the city starts to get cold, the public will opt for other types of transport that offers comfort and heat.

Question 3: How can we manage London’s conflicting goals of Vision Zero and cycling growth?

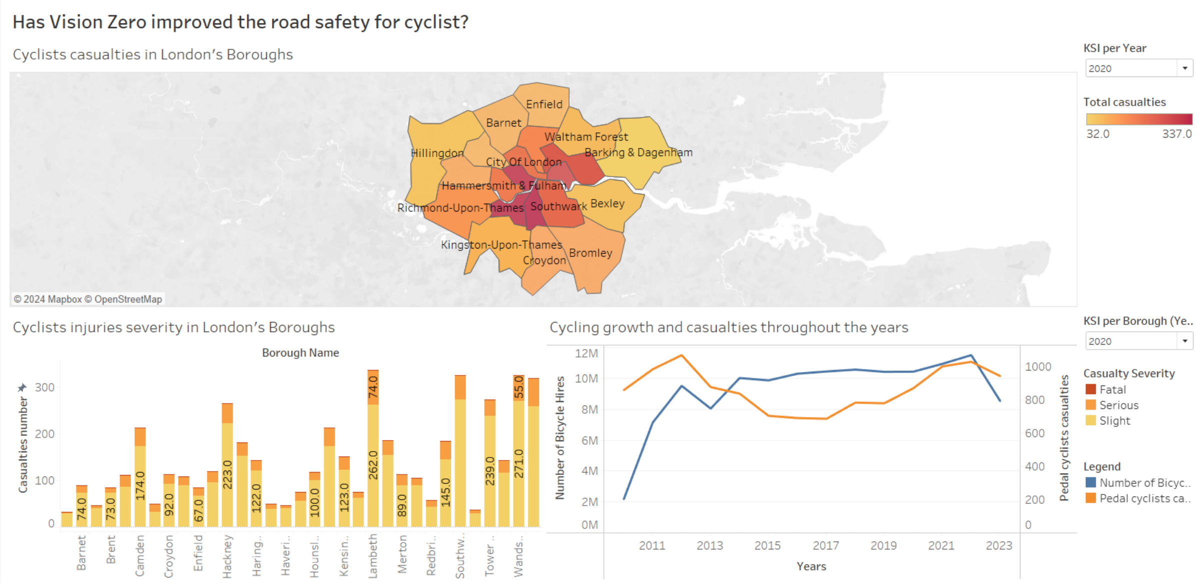

Figure 3: Dashboard answering the question "How can we manage London’s conflicting goals of Vision Zero and cycling growth?"

3.1 Datasets Used and Changes Made

For the following visualizations, the datasets used were sourced from London Datastore [3] and Transport for London (TfL) [5]. The first dataset contains the values of cyclist casualties in London, while the second dataset contains the cycling usage in London (same dataset from the previous question).

The file Excel with the casualty values had multiple tables, where each table showed different scenarios for the casualty numbers. To make a visualization that answers the question, only two tables have been extracted and used from these files and made into a new one. This new file has two tables, one showing the cyclist casualties for each year, from 2010 to 2023, and one table with casualty severity in each borough, from the year 2010-14 to 2023.

3.2 Visualizations Included in the Dashboard and Analysis

Using the two datasets, the dashboard included 3 visualizations: a **map** illustrating the intensity of accidents in each London Borough, a **segmented bar chart**, that shows the number and the severity of casualties in each borough and a **line chart** displaying the correlation between the growth of cycling and the number of accidents.

1. **Cyclists’ casualties in London’s Boroughs (Choropleth Map)**
   This the map of London Boroughs, where each borough shows the intensity of cyclist casualties during the year. These boroughs were coloured with different shades depending on the number of casualties. Lighter shades mean the number of casualties in that borrow is low, while darker shades mean the number of casualties is high. In terms of interactivity, this visualization has a filter where you can choose to show the data from a specific year, for example 2020. Furthermore, while hovering over the map, you can see the number of accidents occurred in each borough.
   By analysing this visualization, we can see that the centre of London has the areas with the highest number of collisions involving cyclist. The cause of these high numbers is related to the fact that area of London is one of the most populate area, as people travel

there to work, study or to even just hang out with friends, thus potentially increasing the number of accidents

2. **Cyclists' injuries severity in London's Boroughs (Bar-line chart)**
   This bar chart is a more detailed version of the map visualization, as it not only shows the number of casualties in each borough but also includes the severity of the casualty. To represent this, each bar is segmented in three different categories based on the severity of the casualty: fatal, serious and slight.
   We can observe that, for the year 2017, most of the accidents in each borough resulted with slight injuries, with Westminster being the borough with the highest number of casualties.
3. **Cycling growth and casualties throughout the years (Dual-axis line chart)**
   This visualization aims to examine whether there is a correlation between the growth in cycling and the rise in cyclist casualties. The blue line represents the bicycle usage over the years, while the orange line depicts cyclist casualties throughout the years. From this, we observe that there is a correlation between these two variables, as the number of cycling increases, the number of casualties increases as well.

3.3 Interpretation of the Entire Dashboard

The Vision Zero plan, launched in 2018, aims to eliminate road casualties. The dashboard shows limited progress, with injuries increasing post-COVID lockdowns but beginning to decline slightly by 2022–2023. The data highlights the challenge of balancing cycling growth with safety, as more cyclists on the road correlate with higher casualty rates.
Expanding the Cycleway network, which has grown from 90km in 2016 to 350km in 2023 [14], along with stricter road safety rules, additional cycle lanes, and cyclist training programs, could address these issues. These measures would support safer cycling while advancing Vision Zero goals.

Question 4: How does traffic affect health outcomes?

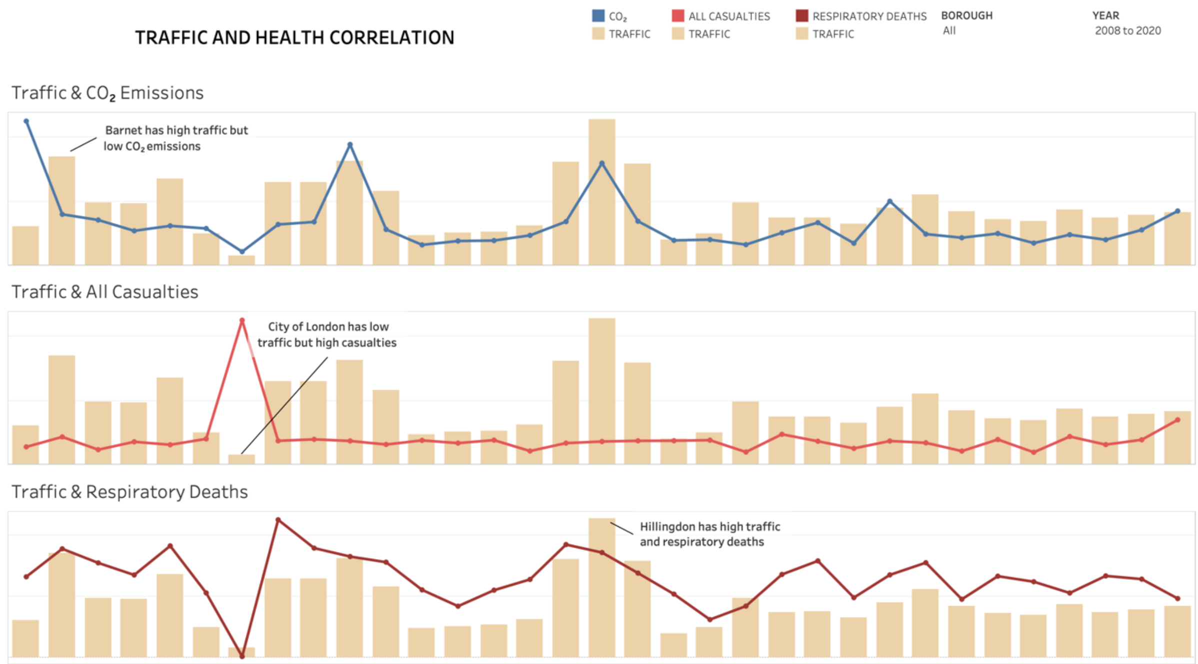

Figure 4: Dashboard answering the question: "How does traffic affect health outcomes?"

4.1 Datasets Used and Changes Made

The dataset used contains values for the parameters mentioned above for the years 2008, 2010, 2012, 2015, and 2020. A small formatting adjustment was made to enable the use of different years as filters in Tableau. Initially, the dataset combined year and parameter information, such as "CO_2 2008," with all values grouped by borough. To address this, the reformatted dataset includes separate columns for year, CO_2, traffic, casualties, and respiratory deaths. The sources used to create this dataset were from TfL Open Data [5], Air Quality Statistics [8], National Archives Open Government Licence [9], TfL Casualties in Greater London Report 2023 [10].

4.2 Visualizations Included in the Dashboard and Analysis

The dashboard includes three bar-line charts, each illustrating distinct relationships between traffic and various metrics across London boroughs:

1. **Traffic & CO_2 Emissions**
   This graph illustrates the connection between CO_2 emissions and traffic volume. CO_2 emissions are displayed as a line, and traffic values are displayed as bars. By using the filters in the upper-right corner or by clicking on a bar to filter the dashboard, users can interact with the chart. According to the visualisation, traffic and CO_2 emissions are positively correlated, rising simultaneously.
2. **Traffic & All Casualties**
   The number of casualties of all kinds (line) and traffic volume (bars) are compared in this graphic. Using the dashboard filters or the bar choices, users can filter by borough. The graph indicates that the level of traffic has no apparent effect on the number of fatalities in London.
3. **Traffic & Respiratory Deaths**
   The association between respiratory fatalities (line) and traffic volume (bars) is explored in this chart. Interactive features enable dashboard controls or borough-based filtering. According to this visualisation, there is a strong association between higher traffic levels and more mortality from respiratory diseases.

4.3 Interpretation of the Entire Dashboard

The first line-column chart shows a clear relationship between traffic volume and CO_2 levels, where an increase in traffic corresponds to higher CO_2 emissions. The second chart indicates no apparent association between traffic and total casualties. The third chart suggests a positive correlation between traffic volume and respiratory deaths, though the relationship is not as strong as the one with CO_2 levels.

Three boroughs demonstrate particularly interesting patterns. 'Hillingdon' has the highest average traffic volume, which correlates with high CO_2 levels and respiratory death rates, yet it records one of the lowest casualty rates. The 'City of London', with the lowest average traffic volume, has the lowest CO_2 levels and respiratory death rates. However, it exhibits a casualty rate 10 times higher than the average, likely due to its central location and status as a major tourist destination. 'Barnet', despite having high traffic volumes, shows average CO_2 levels, suggesting that factors other than traffic volume may influence CO_2 emissions in this borough.

In conclusion, traffic volume appears to affect CO_2 levels and respiratory deaths, although there are some exceptions.

Question 5: Are there correlations between transport volumes and other urban data metrics?

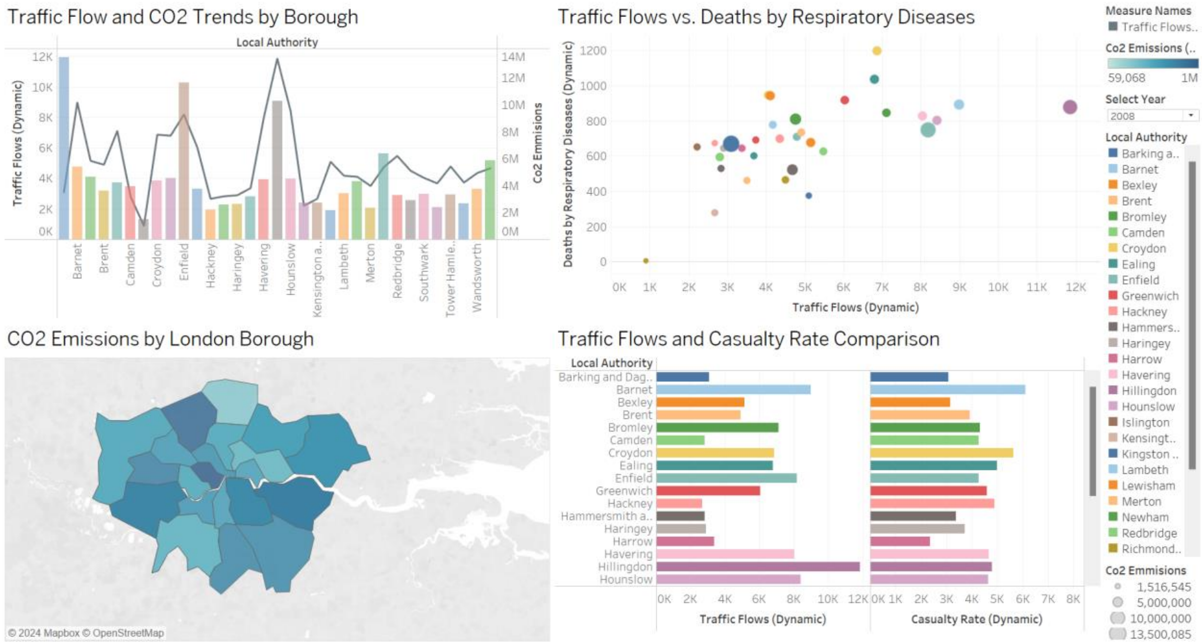

Figure 5: Dashboard answering the question "Are there correlations between transport volumes and the other urban data metrics? "

5.1 Datasets Used and Changes Made

The dataset was sourced from publicly available data, including the TfL Open Data [5], Air Quality Statistics [8], National Archives Open Government Licence [9], TfL Casualties in Greater London Report 2023 [10]. The raw data from these sources underwent substantial preprocessing. **Normalization** was applied to ensure consistency across metrics, such as borough population sizes and traffic volumes, allowing meaningful comparisons. The datasets were **cleaned** by removing irrelevant fields, resolving missing values, and standardizing formats. The processed datasets were then combined into a single Excel file, which served as the foundation for the visualizations. These steps ensured that the data was accurate, comparable, and ready for visualization.

5.2 Visualizations Included in the Dashboard and Analysis

The dashboard includes four visualizations, each exploring relationships between traffic flows, CO_2 emissions, casualty rates, and respiratory deaths across London boroughs:

1. **Traffic Flows vs. Respiratory Deaths (Scatter Plot)**
   This scatter plot highlights the correlation between traffic flow and deaths caused by respiratory diseases. Boroughs are represented as points, with size reflecting CO_2 emissions and colour coding distinguishing regions. The chart includes a year parameter, enabling users to view trends for specific years like 2008, 2012, or 2020. Boroughs with higher traffic volumes, such as Westminster, often exhibit increased respiratory-related deaths, suggesting links between transport activity and health impacts.
2. **Traffic Flows & CO_2 Emissions (Bar and Line Chart)**
   This dual-axis chart compares traffic volumes (bars) with CO_2 emissions (line) for each borough. Interactive features allow users to hover for specific values and filter by year. Results reveal boroughs like Hillingdon as hotspots of transport activity and emissions, highlighting the environmental burden of traffic-heavy areas.

3. **CO_2 Emissions by Borough (Geographic Heatmap)**
   A geographic heatmap displays the spatial distribution of CO_2 emissions, with darker shades indicating higher levels. Users can explore emissions trends over time using year filters. Boroughs in Inner London, such as Camden and Westminster, exhibit the highest emissions, aligning with their dense urban activities.
4. **Traffic Flow and Casualty Rate Comparison (Side-by-Side Bar Chart)**
   This bar chart compares traffic flows and casualty rates by borough, offering insights into safety concerns. Filters enable users to explore borough-specific trends, and tooltips provide precise metrics. Boroughs like Camden show disproportionately high casualty rates relative to traffic, raising questions about road safety infrastructure.

5.3 Interpretation of the Entire Dashboard

This dashboard effectively answers the question by demonstrating strong correlations between transport volumes and urban metrics. The scatter plot establishes potential links between traffic flow and respiratory health issues, while the dual-axis chart and heatmap provide insights into the environmental impact of traffic. The comparative bar chart highlights road safety disparities, emphasizing the need for targeted interventions. The interactive features allow users to explore these relationships over time, making the dashboard a valuable tool for policymakers and researchers seeking to understand and mitigate the consequences of urban transport activity.

Question 6: Can or do urban traffic restrictions work?

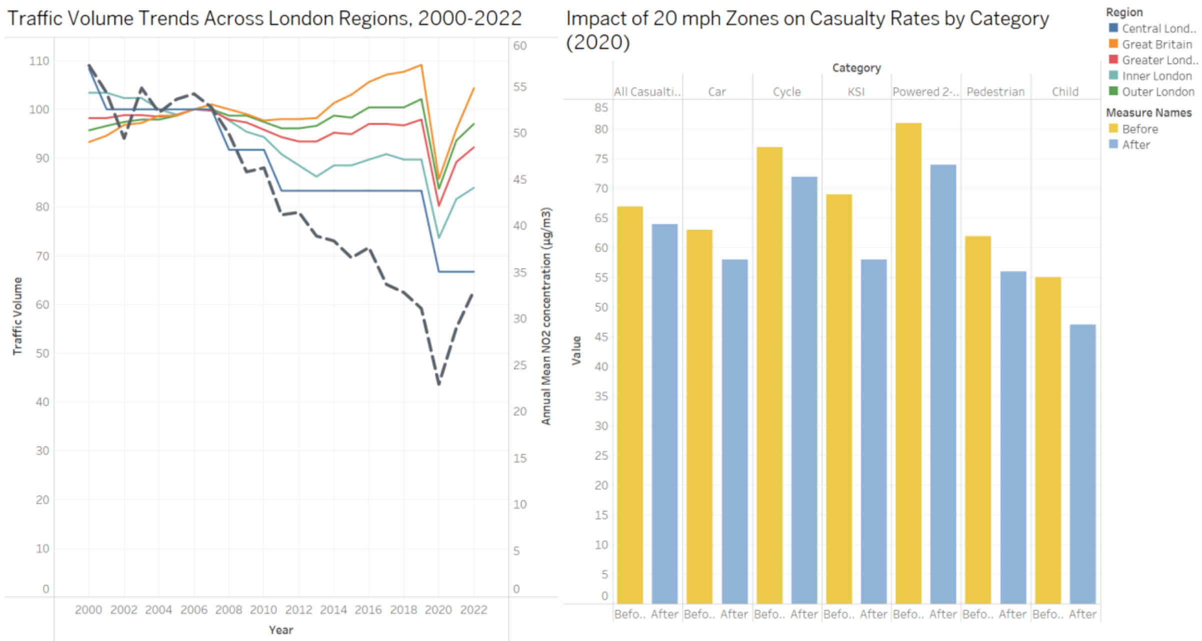

Figure 6: Dashboard answering the question: "Can or do urban traffic restrictions work?"

6.1 Datasets Used and Changes Made

The dataset for this dashboard was built using publicly available data from the following sources: TfL Open Data [5] and London Road Casualties Severity Data [11]. To ensure consistency and comparability, the raw data underwent a rigorous preparation process. **Normalization** was applied to casualty rates and traffic indices to account for variations in borough populations and geographic sizes. The datasets were also cleaned by removing irrelevant fields, addressing missing values, and standardizing formats to align across metrics. Additionally, traffic-related air pollution data (NO_2 levels) was processed and cross-referenced with traffic trends. The resulting smaller, individual datasets were used as the basis for the visualizations, ensuring flexibility and accuracy in analysing specific metric.

6.2 Visualizations Included in the Dashboard and Analysis

The dashboard includes two visualizations, each exploring the impact of urban traffic restrictions on safety and environmental outcomes across London:

1. **Traffic Volume Trends Across London Regions (2000–2022) (Dual-Axis Line Chart)**
   This dual-axis line chart tracks long-term trends in traffic volumes and NO_2 pollution levels across Central, Inner, and Outer London. Filters enable users to focus on specific regions, while tooltips provide detailed yearly data. The chart shows a steady decline in traffic volumes, especially in Central London, since the early 2000s. A concurrent decrease in NO_2 levels suggests that traffic reductions contributed to improved air quality in these areas.
2. **Impact of 20 mph Zones on Casualty Rates by Category (Grouped Bar Chart)**
   This grouped bar chart evaluates the effect of 20 mph zones on casualty rates among cyclists, pedestrians, and motorists. Users can interact with the chart to view specific road user groups and use hover tooltips to compare "before" and "after" rates. Results indicate significant reductions in casualties for all categories, with the most notable improvements seen for vulnerable users like cyclists and pedestrians. This demonstrates the positive safety impact of implementing 20 mph zones.

6.3 Interpretation of the Entire Dashboard

This dashboard effectively demonstrates the potential of urban traffic restrictions in improving road safety and reducing environmental pollution. The grouped bar chart reveals the success of 20 mph zones in decreasing casualty rates, particularly for vulnerable road users. The dual-axis line chart provides further evidence by correlating traffic reductions with air quality improvements over time. By integrating these insights, the dashboard highlights the multidimensional benefits of traffic restrictions and serves as a valuable resource for policymakers evaluating the impact of urban interventions.

Question 7: How do different urban areas compare in one or more of these variables?

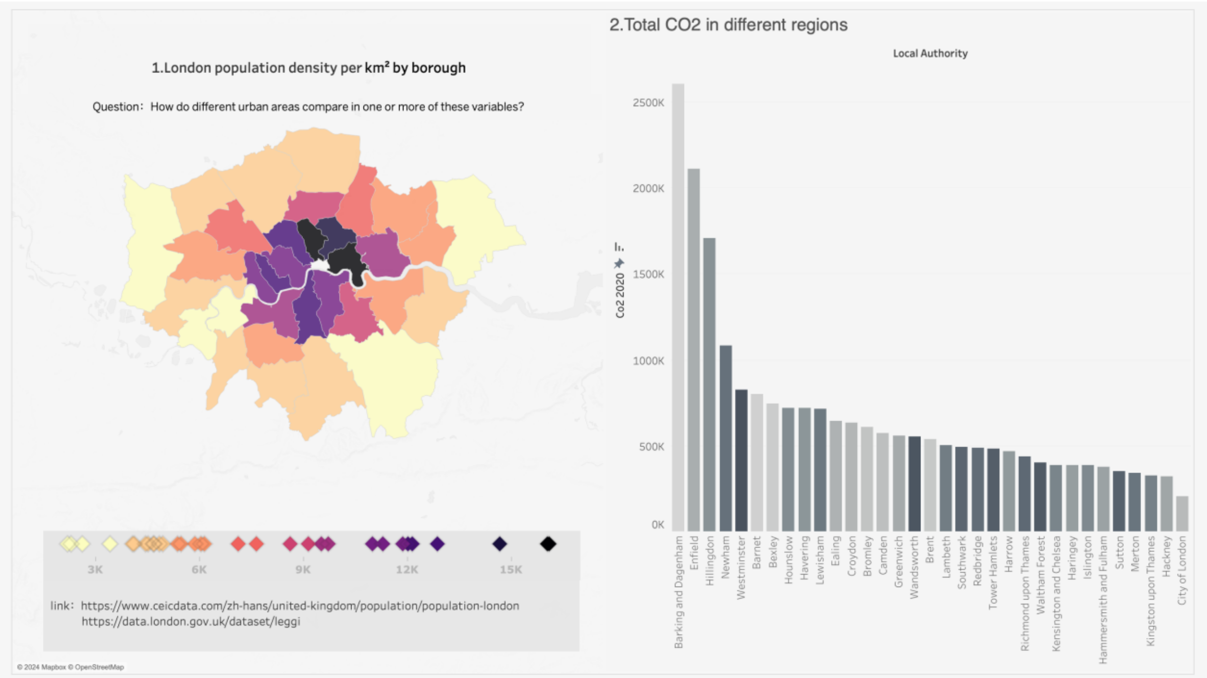

Figure 7: Dashboard answering question: ": How do different urban areas compare in one or more of these variables?"

7.1 Datasets Used and Changes Made

The dataset includes population data sourced from the [CEIC (Office for National Statistics)] [12] and CO_2 emissions data obtained from the [London Energy and Greenhouse Gas Inventory (LEGGI)] [15]. Preprocessing was essential to ensure accuracy and comparability. Population density was calculated by dividing the total population by the area of each borough. The CO_2 emissions dataset was aggregated at the borough level, with missing or incomplete data interpolated where necessary. The processed datasets were then integrated into a single file to facilitate dashboard creation and analysis.

7.2 Visualizations Included in the Dashboard and Analysis

The dashboard includes two visualizations that explore the relationship between population density and CO_2 emissions across London boroughs:

1. **Population Density Map (Choropleth Map)**
   This map visualizes population density (people per km^2) for each borough using a gradient colour scale. Darker shades indicate higher densities, with boroughs like Camden and Westminster standing out as urban hubs. Interactivity allows users to hover over boroughs to view specific population density values, enabling easy exploration of the data.
2. **CO_2 Emissions (Bar Chart)**
   A vertical bar chart shows CO_2 emissions by borough. Taller bars represent higher emissions, with industrial and transport-heavy boroughs like Barking and Dagenham, Hillingdon, and Newham leading. The chart helps users compare emission levels across boroughs, highlighting areas with significant environmental impact.

7.3 Interpretation of the Entire Dashboard

The dashboard effectively answers the question by showcasing clear patterns in population density and CO_2 emissions across London boroughs. High-density central boroughs, such as

Camden and Westminster, exhibit higher emissions, likely due to increased urban activity and transport use. Conversely, suburban boroughs like Bromley and Richmond upon Thames have lower population densities and CO_2 emissions, reflecting their residential and green-space character.
These insights highlight the interplay between urban density and environmental impact, emphasizing the importance of tailored policies to address emissions in densely populated areas while preserving the liveability of suburban regions.

Question 8: Distribution and number of tree species in different regions of London

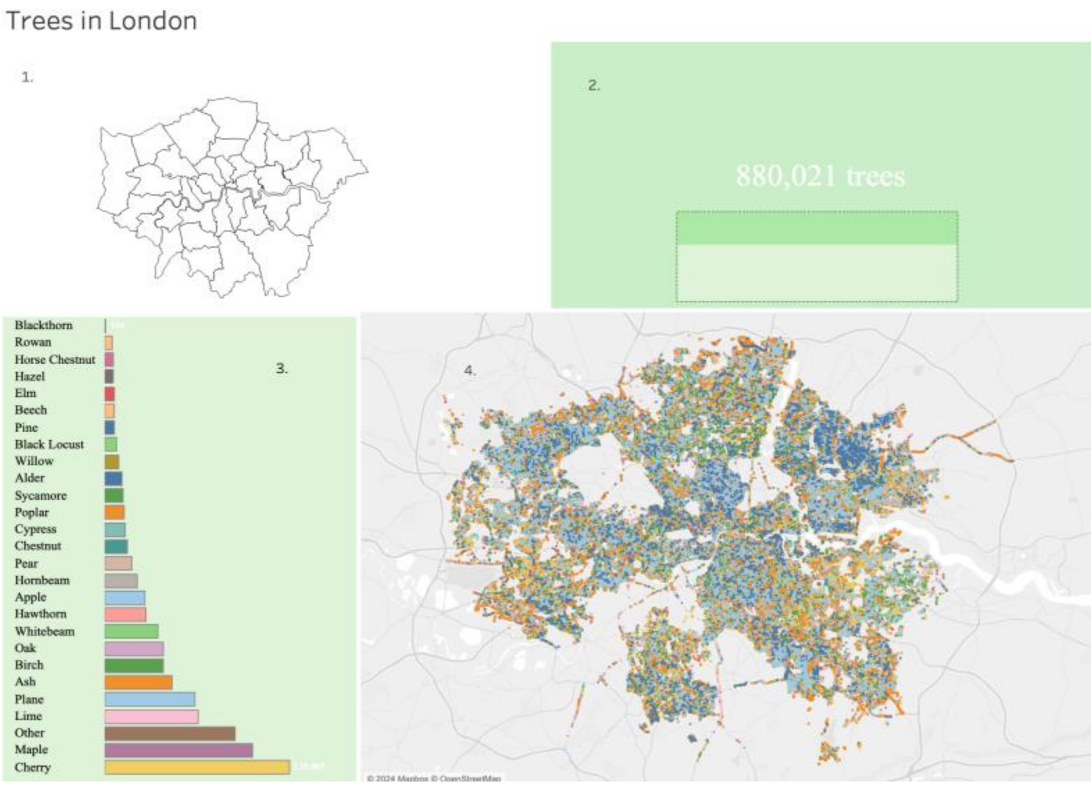

Figure 8: Dashboard answering question "Distribution and number of tree species in different regions of London"

8.1 Datasets Used and Changes Made

The dataset is sourced from the [London Datastore (Local Authority Maintained Trees) [4]], which provides information on over 880,000 trees maintained by local authorities across London. The dataset includes species names, borough locations, and tree counts. Preprocessing involved filtering the data to include only 26 of London's 33 boroughs, along with the City of London and Transport for London data, to improve query performance. Lesser-represented tree species were excluded to focus on meaningful insights. The dataset was cleaned to ensure consistency in borough names and formats, allowing for smooth integration into the dashboard.

8.2 Visualizations Included in the Dashboard and Analysis

The dashboard includes four visualizations that together provide insights into the distribution and variety of tree species in London:

1. **Borough Map (Top Left)**
   A map of London boroughs serves as a dashboard filter. Users can click on a borough to view its specific tree data across other visualizations. This feature helps users focus on specific regions of interest.
2. **Total Tree Count Display (Top Right)**
   A numerical display shows the total number of trees in the selected borough(s). For instance, the dashboard reveals that London has 880,021 recorded trees. This visualization provides a quick summary of tree abundance.
3. **Tree Species Bar Chart (Bottom Left)**
   A vertical bar chart ranks tree species by count within the selected borough(s). Colours represent different tree types, corresponding to the spatial map visualization. Cherry trees dominate as the most abundant species, while Blackthorn is the least represented.
4. **Tree Species Map (Bottom Right)**
   A spatial map plots the locations of tree species within the selected borough(s). Tree species are color-coded to match the bar chart, enabling easy correlation. Users can identify where specific species, such as Oak or Maple, are concentrated.

8.3 Interpretation of the Entire Dashboard

This dashboard provides a comprehensive view of London's urban forest, highlighting both the distribution and variety of tree species across boroughs. Central boroughs with dense urban areas, such as Westminster and Camden, exhibit fewer trees compared to outer boroughs like Richmond upon Thames and Bromley, known for their green spaces.

The bar chart and spatial map reveal Cherry as the most abundant tree species, while Blackthorn is the least. This reflects planting preferences in urban planning. Interactive features, such as borough selection and species-specific data, make the dashboard a valuable tool for city planners and environmental researchers aiming to manage and expand London's tree coverage strategically.

Conclusion

This report analysed urban data in London through interactive Tableau dashboards, addressing eight key questions on transport, air quality, cycling, health, and urban density. The dashboards revealed critical insights, such as the correlation between traffic and NO_2 emissions and disparities in tree coverage across boroughs, highlighting the importance of data-driven urban planning.

This report analyzed urban data in London through interactive Tableau dashboards, addressing eight key questions on transport, air quality, cycling, health, and urban density. The dashboards revealed critical insights, such as the correlation between traffic and NO_2 emissions and disparities in tree coverage across boroughs, highlighting the importance of data-driven urban planning.

Team collaboration was essential, with contributions ranging from data preprocessing and visualization design to report writing and video creation. This project demonstrates the value of visualization tools in uncovering actionable insights and supporting strategies for sustainable urban development.

This report analyzed urban data in London through interactive Tableau dashboards, addressing eight key questions on transport, air quality, cycling, health, and urban density. The dashboards revealed critical insights, such as the correlation between traffic and NO_2 emissions and disparities in tree coverage across boroughs, highlighting the importance of data-driven urban planning.

References

[1] Global Clean Air ~ https://globalcleanair.org/innovative-air-quality-monitoring/london-uk/breathe-london-data/
[2] Department for Transport ~ https://roadtraffic.dft.gov.uk/regions/6
[3] London Datastore ~ https://data.london.gov.uk/
[4] Santander's cycling hire dataset ~ https://data.london.gov.uk/dataset/number-bicycle-hires
[5] TfL Open Data ~ https://tfl.gov.uk/info-for/open-data-users/
[6] London's air quality dataset ~ https://data.london.gov.uk/dataset/london-average-air-quality-levels
[7] Cyclist casualties in London ~ https://tfl.gov.uk/corporate/publications-and-reports/road-safety#
[8] Air Quality Statistics ~ https://tfl.gov.uk/info-for/open-data-users/our-open-data
[9] National Archives Open Government Licence ~ https://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/
[10] TfL Casualties in Greater London Report 2023 ~ https://content.tfl.gov.uk/casualties-in-greater-london-2023.pdf
[11] London Road Casualties Severity Data ~ https://data.london.gov.uk/dataset/road-casualties-severity-borough
[12] CEIC (Office for National Statistics) ~ https://www.ceicdata.com/en/united-kingdom/population
[13] London Authority https://data.london.gov.uk/dataset/local-authority-maintained-trees
[14] Transport for London | Every Journey Matters, Encouraging Cycling & Walking, Transport for London. Available at: https://tfl.gov.uk/corporate/about-tfl/how-we-work/planning-for-the-future/encouraging-cycling-and-walking (Accessed: 26 November 2024)
[15] London Energy and Greenhouse Gas Inventory (LEGGI) ~ https://data.london.gov.uk/dataset/leggi#:~:text=In%202022%2C%20London's%20emissions%20were,to%2028.40%20MtCO%202%20e.