

iDAV COURSEWORK REPORT

BTC Presentation

By Khadija Baffa

INTRODUCTION

In November 2021, Portsmouth's Clean Air Zone (CAZ) came to effect to make the air we breathe cleaner for everyone. After two and a half years in operation, the British Transport Consortium (BTC) now want to understand if the CAZ in its current configuration has made any difference to pollution levels in Portsmouth.

In response to the British Transport Consortium's (BTC) request, this report analyzes air quality data from Portsmouth's Clean Air Zone (CAZ). The BTC has requested four visualizations to assess pollution levels on specific dates, as well as monthly averages for the year 2020.

Furthermore, the BTC has entrusted the data analysis specialist with generating one visualization that shows a pattern discovered during exploratory investigation.

This report outlines the methods employed and the rationale behind each step.

Visualisation 1:

1. Data preparation

- - Merged multiple datasets into one for easier accessibility.
- - Used Lubridate to convert dates to the day-month-year (dmy) format, which complies with accepted date formats.
- - Renamed a column with a long name for better clarity and readability.
- - Filtered the dataset to produce a smaller, more focused subset for analysis.

2. Data Cleaning:

- - Removed null values from the PM10 column using tidyr's cheatsheet as a guide, which accounted for just 4 items across the dataset.

3. Visualisation Technique:

- - Used a scatter plot to show PM10 concentration trends over time.

4. Rationale:

- - The visualization aimed to examine whether PM10 levels reduced post-implementation of the Clean Air Zone (CAZ). The scatter plot revealed that the dates with the highest PM10 concentrations occurred before November 2021, indicating potential effectiveness of the CAZ in reducing PM10 pollution.

Visualisations 2 and 3:

1. Research and Implementation:

- - Conducted research online and discovered the capability of combining Shiny and Plotly for interactive visualizations.
- -Studied the cheat sheets that were provided to us and even found some more
- - went through plotly and shiny's websites to learn the different types of charts they have and how to use them

2. Technique:

- - Created an interactive visualisation using Shiny and Plotly, allowing for dynamic exploration by switching the x and y variables.
- -used the codes i found on the internet to make mine better

3. Rationale:

- Chose These techniques because it is versatile and easy to use, allowing users to interactively study data.

Visualisation 4

1. Data Cleaning:

- -Removed null values from the subset i used in by filtering the rows (not including them in the data set)

2. Techniques used:

- -using group by and summarise to calculate the mean of each month and removing the few null values i have left in the columns.
- -Created another table with the months listed as strings and the values being vectors to make plotting easier for me.

3. Rationale:

- Easy to understand as you can compare the averages of all three at the same time.
-

CONCLUSION

Through the curation of these visualizations, key insights into PM10 concentration trends and the effectiveness of the Clean Air Zone were revealed. The utilization of advanced visualization techniques enhances the interpretability and accessibility of the data, facilitating informed decision-making processes.

REFERENCES

1. Shiny <https://shiny.posit.co/r/gallery/>
2. <https://mgcodesandstats.medium.com/creating-an-interactive-bubble-chart-with-shiny-and-pyplot-59d1b9b7bbe6>
3. <https://plotly.com/r/basic-charts/>

4.