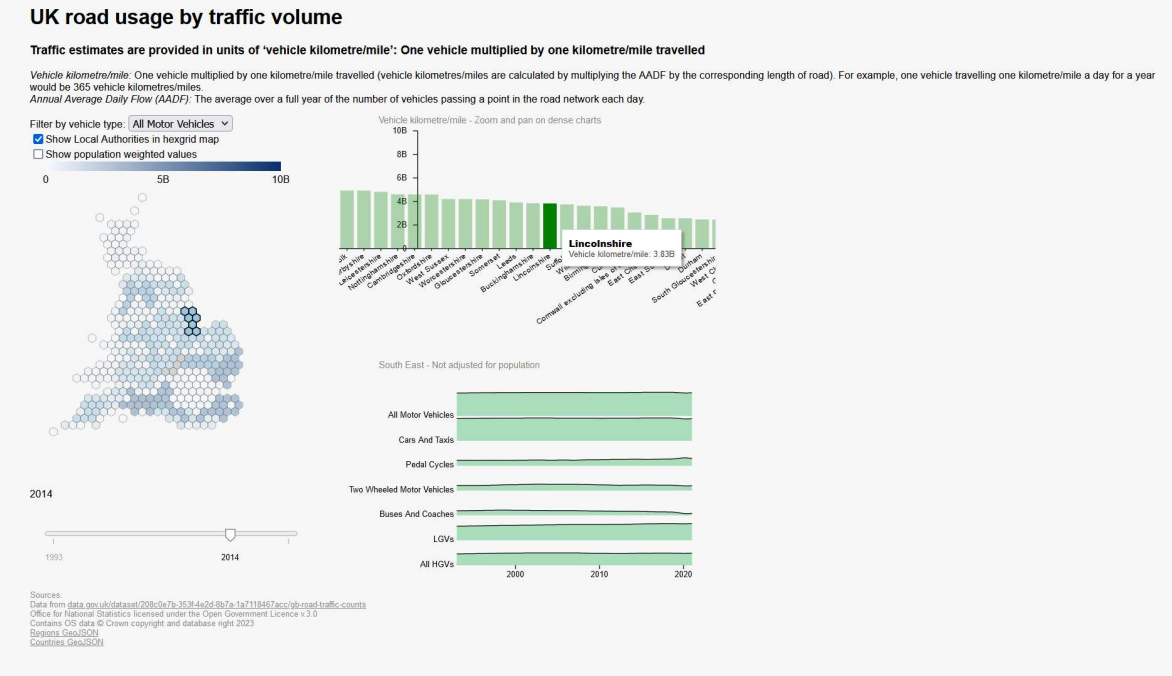
Data Visualisation Mini-Project A18331S1

UK road usage by Traffic Volume

Candidate Number: 1062380

I declare that, except where otherwise indicated, this mini-project is entirely my own work, and that it has not been previously submitted and/or assessed and is not due to be submitted in its entirety or in part for any other course, module or assignment.



Overview

This visualisation uses UK government data from across the country recorded over the last nearly 30 years at numerous count points, with the resulting estimates of total traffic volume given in units of vehicle kilometre/mile, as a representation of the numbers of vehicles in an area and an approximation of the distances travel – allocated into groups by regions (North West, West Midland, London etc.), and local authorities, as well as grouping by vehicle type.

From the dataset definitions:

"Vehicle kilometre/mile: One vehicle multiplied by one kilometre/mile travelled (vehicle kilometres/miles are calculated by multiplying the AADF by the corresponding length of road). For example, one vehicle travelling one kilometre/mile a day for a year would be 365 vehicle kilometres/miles.

Annual Average Daily Flow (AADF): The average over a full year of the number of vehicles passing a point in the road network each day."

The aim of this visualisation is to display all this information in a way that is easier for a person to perceive – using highlighting, tooltips and clear colourings; and giving a number of interactive tools to filter data to their needs or interests.

This means an understanding of road use and traffic distribution across the country can be gathered, so particularly targets this towards people making decisions about funding or development of national transport services, including those already interested in this dataset.

Data

Data from: roadtraffic.dft.gov.uk/downloads, specifically using Regional traffic: by vehicle type and, Local authority traffic data; across the whole range of years using only total volume estimates from both. Includes categorical data for area names and ONS codes (used to link values between datasets), and quantitative data.

Give data over 29 years, with 212 local authorities each with two data points, across values from around 8.51 billion to 52.28 million; and 11 regions each with 7 data points across values around 47 billion down to 111 million

And populations for both regions and local authorities; edited down to general values rather than the full set including age and gender groupings: www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/populationestimatesforukenglandandwalesscotlandandnorthernireland

Area plots for regions: geoportal.statistics.gov.uk/datasets/ons::regions-december-2022-en-bgc/about, to which plots for Scotland and Wales are added from the same site.

Hex map data for local authorities, used in full: github.com/odileeds/hexmaps/tree/gh-pages/maps

And data linking between lower and upper tier authorities, to join the datasets and map data where these are used interchangeably: geoportal.statistics.gov.uk/datasets/ons::lower-tier-local-authority-to-upper-tier-local-authority-december-2016-lookup-in-england-and-wales-1/explore

Data pre-processing is done within the visualisation to group and link data, so that mapped areas have the relevant statistics included and comparisons across encodings can be made.

Goals and Tasks

The main goal of the visualisation is to help analyse the datasets to aid users in discovering information – relating to traffic volumes across the UK over 28 years, grouping the information by area, year and vehicle type.

In a similar theme I aimed to help users search the datasets so they could locate information about specific target areas or browse for more general details – by looking across different areas and comparing adjacent values.

It was also an aim that they would be able query the data to summarise trends across different groups, and to identify outliers or extreme values within the set – searching for regions with the highest or lowest traffic volume, or identifying those with the greatest changes.

Visualisation

The visualisation uses four major distinct views, these are:

Choropleth Map

Encoding the traffic volume data in regions across a geographic map of the UK, this uses *geojson* data and the *d3.geoMercator* projection to give a accurate map – producing a very easy to read map with coloured areas that are simple to identify and compare; taking the values as a simple linear coloured scale and the geographic data for each region

- Only a small number of discriminable steps in the colour channel are need as there are only a few regions
- A standard colour scale, using the traffic volume, and a tooltip, providing the specific quantitative values, are used for multiple views - described in detail below

Hex Map

A specific variation of grid cartogram, encoding data for local authorities similar to choropleth map but represented each area as a hexagon-shaped point giving a rough representation of their geographic positioning – with bordering areas still approximately grouped together.

- This is done to improve the accuracy of the user's perception of data encoded in the colour channel by avoiding interference with information encoded by the area channel, particularly crucial here with a much larger number of areas and a wider range of values as well as a greater variety in size for the actual geographic area.
- The standard colour scale and tooltip is used as described below

As a result of nonpersistent definitions of local authorities and changes over time some data points are grouped together across multiple lower tier authorities where there is only data for the upper tier authority, or data is in some years not available.

Bar Chart

Linked to the map encoding showing areas along x-axis and the relevant vehicle type traffic volumes, making values simple to compare and identify.

- Ordered by traffic data values to help with task of querying
- Distinct colour scheme from map views avoiding any confusion and not adding unnecessary complexity with multiple channels encoding the same data

Ridgeline Chart

Encodes the selected region values across time, split into vehicle types so that corresponding changes can be seen, to find trends or pattern not immediately obvious from other views.

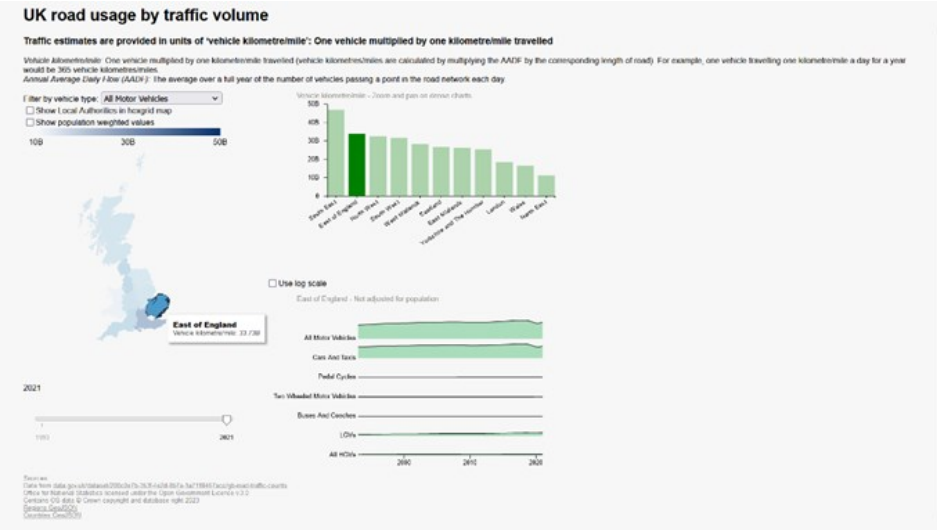
- Uses areas as marks, the channels encoding the data values are the lines' positions both vertically and horizontally – for the data traffic values and years.

- The years discretely split the data in categories but can also be used along a continuous scale as here showing the changes in traffic volume for a individual region.
- This view also includes a checkbox to use logarithmic scale along the y-axis values switching to show variation across orders of magnitude.

Other aspects of the visual encoding, including interactions:

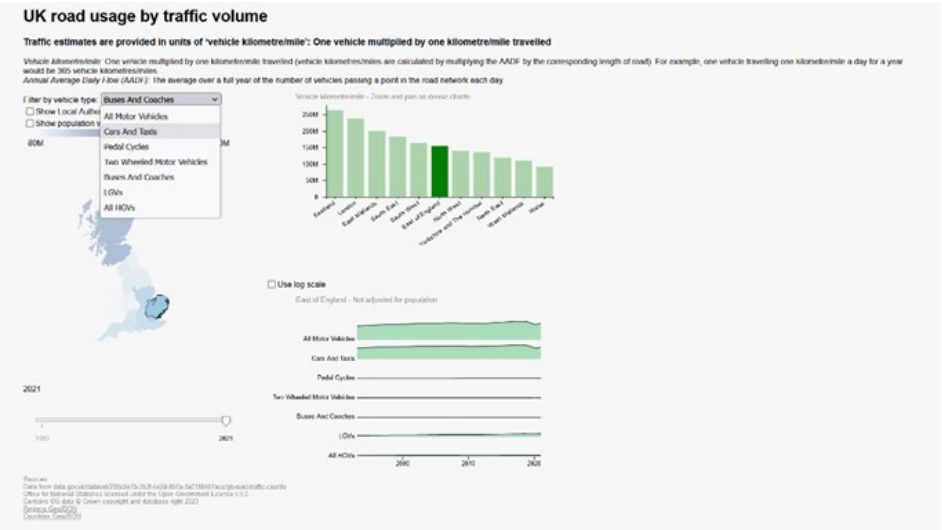
- **Tooltip**, when the cursor is over the maps or bar chart a simple popup block shows the area name and formatted value, as this is important information but it would be overly dense to present within the views
- **Time scale**, a basic slider beneath the views allows users to select the year for which data is shown, linked to the map and bar chart
- **Linked views**, selected values are linked across all the views with highlighting using the colour saturation channel and outlines of areas, to make them popout. And selections made by clicking on map areas or bars, also indicating the region selected for the ridgeline chart
- **Colour scale**, both maps use the simple *d3.interpolateBlues* scale which clearly shows the range of values with differences easy to perceive, as well as maintaining accessibility for any red-green colourblind users
- **Zooming**, to show more detail on choropleth map and, also with scrolling, on dense the bar chart for local authorities so that users can see specific bars and area names – *this is the source of a bug on the regions bar chart where zooming is disabled*
- **Dropdown / check boxes**, allow the users to filter by vehicle type, switch map view and include population weighting on area values *(although it may misrepresent the information as population data only covers a single year)*

Usage Scenario

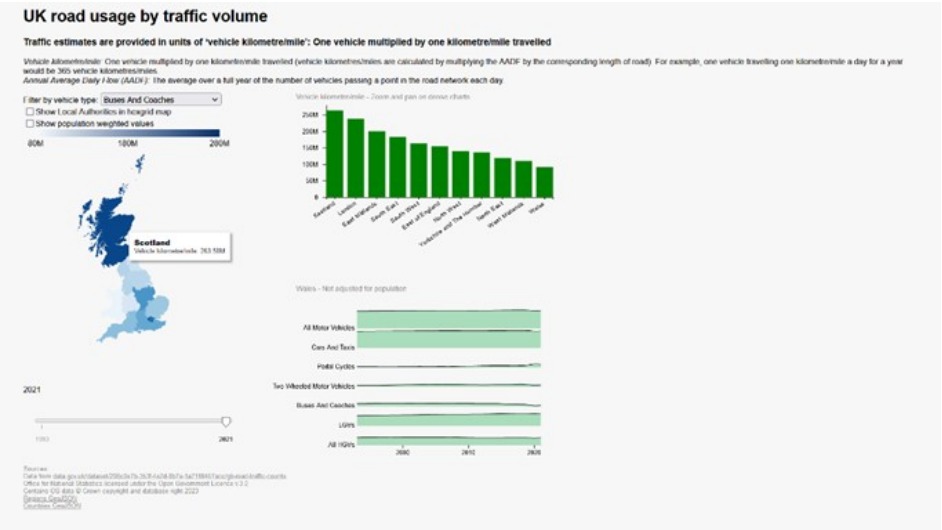


Explore region traffic volumes with map and tooltip.

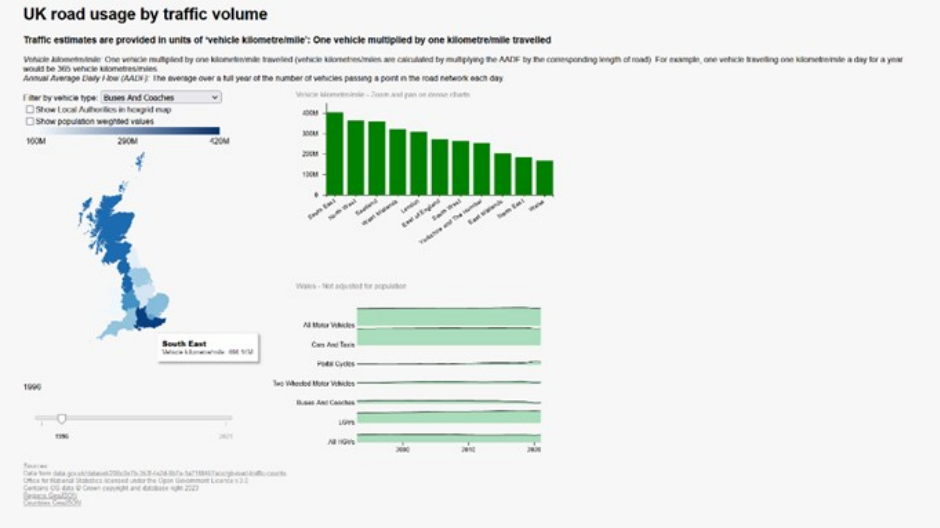
Linked selection makes area and bar popout for comparison, and brings up corresponding ridgeline chart



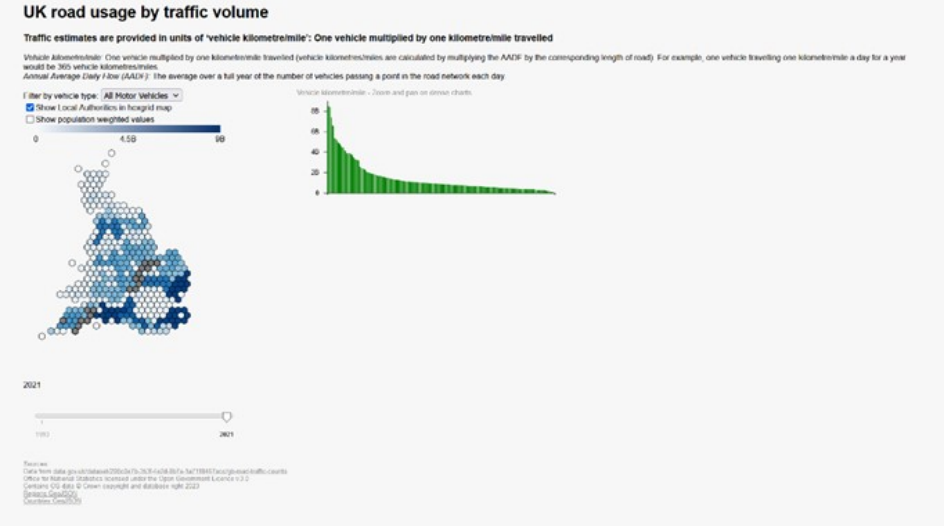
Dropdown menu can filter for the relevant vehicle type



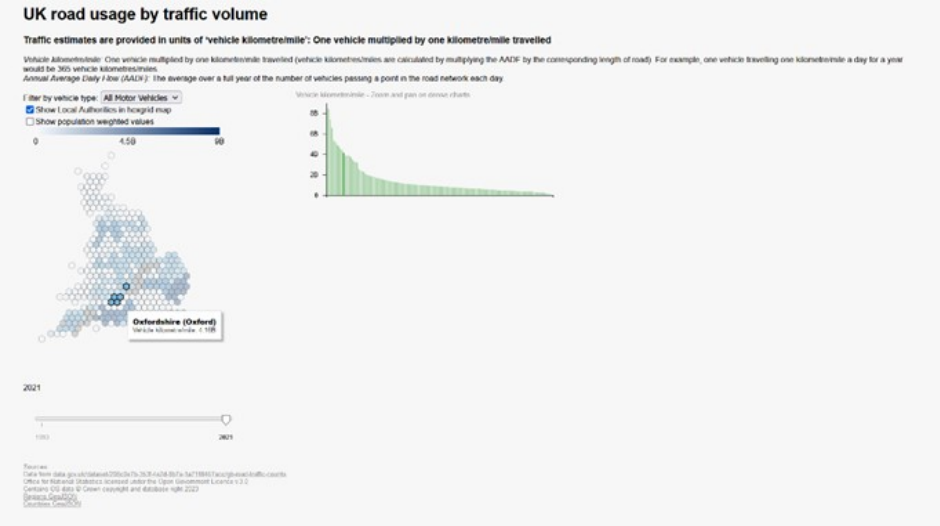
With a new filter, looking now at bus and coach value, the user can again look at region values and compare



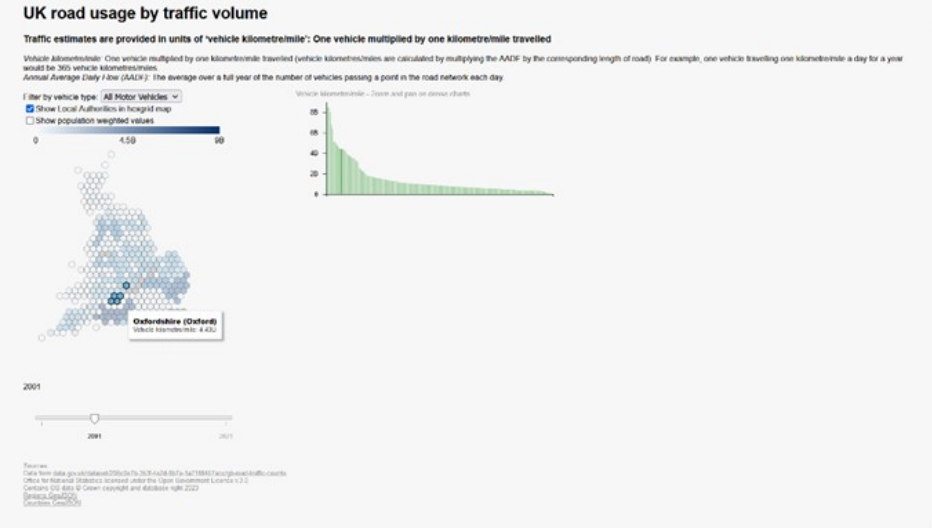
Time slider shows visual change in relative values by colours on map and as bars change size and order. Specific values visible in tooltip. Here the user can see changes on a national level.



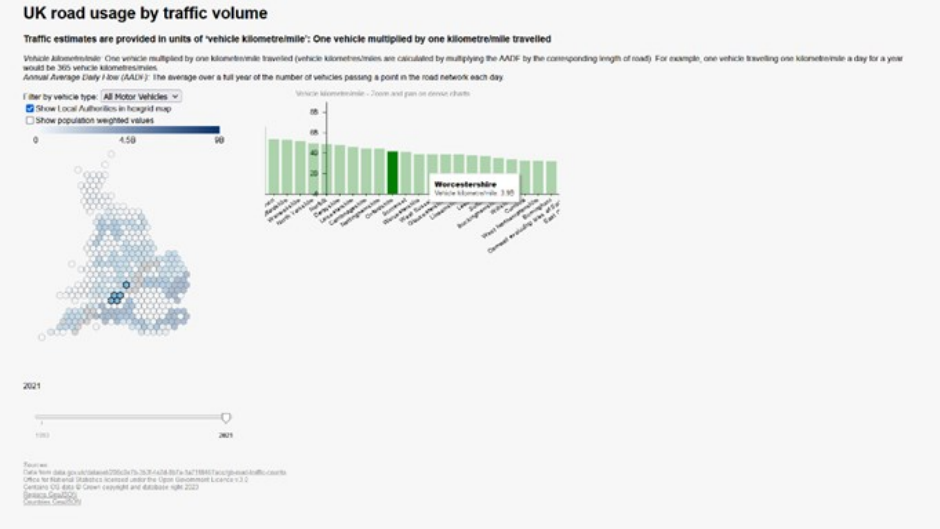
Local authority hex map, breaks down regions into smaller areas giving a more resolution to the data.



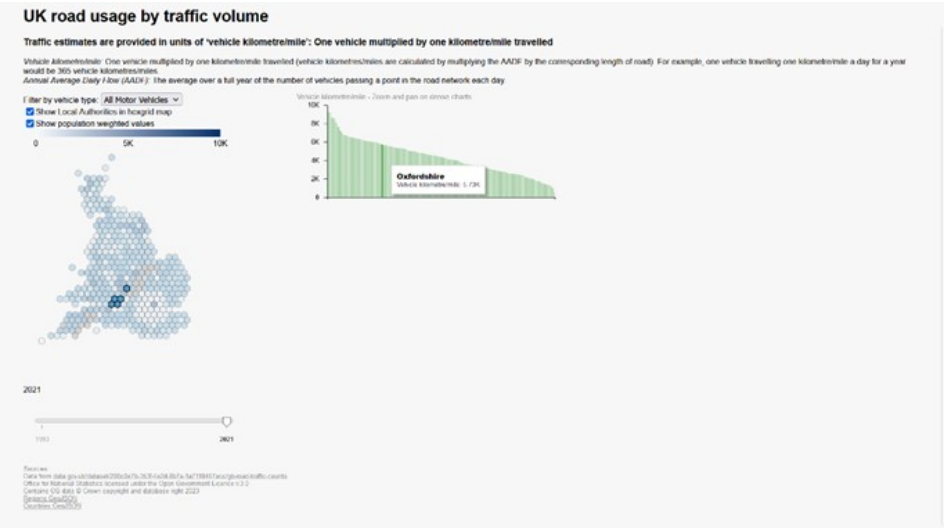
Again can select particular area of interest or browse views with tooltip, here Oxfordshire is selected and is equally highlighted on the bar chart where the values can be perceived more clearly on a national scale.



Again the time slider can be used to understand changes over time



Zoom on bar chart directly shows other local authorities with similar traffic volumes that year, or can find the minimum or maximum values



Finally, population weighting can be included across both area set, in order to avoid the traffic volume data being only representative of the population density

A18331S1

Credits

Hex map using file format and library from github.com/olihawkins/d3-hexjson

Time slider uses library from unpkg.com/d3-simple-slider and adapts the implementation from vizhub.com/gotyou007/0691b8d1dd884517af7697eb7f33f956

Ridgeline chart adapted from observablehq.com/@d3/ridgeline-plot

Bar chart zoom functionality integrated from observablehq.com/@d3/zoomable-bar-chart

And re-use of code developed for problem sheets and from tutorials, in particular for the colour legend bar.