Intro

1.6 million accidents and 16 years of traffic flow. It's a huge picture of a country undergoing change.

What are the best ways to map and analyse how a nation changes over 16 years?

The UK government amassed traffic data from 2000 - 2016, including the data on over 1.6 million accidents. This is one of the most comprehensive data sets on traffic. It's been compiled from a few UK government sources. More can be added to include casualty information, or vehicle damage information but I excldued them for now.

Two types of file

1. The Average Annual Daily Flow (ukTrafficAADF.csv) which tracks how much traffic there was on all major roads (split by bike, motorbike, car, HGV, lorries). This runs from 2000 - 2016.
2. The accidents data (This had to split into 3 files so it could be handled well in memory) - There are three sheets covering 1.6 million accidents this. The columns are exactly the same. Each file covers a different time period (in the file title). The total time period is 2005 - 2014 (2008 is missing).

n.b. accident data comes from police reports, so this data does not include minor incidents.

Goals & questions

* How has changing traffic flow impacted accidents?
* Can we predict accident rates over time?
* What might improve accident rates?
* Plot interactive maps of changing trends, e.g. How has London has changed for cyclists? Busiest roads in the nation?
* Which areas never change and why?
* North Vs. South, East Vs. West
* Identify infrastructure needs, failings and successes
* How have Rural and Urban areas differed (See RoadCategory column)
* Differences between England, Scotland, and Wales

The basics

AADF figures are presented as: Units = vehicles per day

AADF is the Annual Average Daily Flow, i.e. how many cars will drive on this road per day (averaged using yearly data)

AADF figures give the number of vehicles that will drive on that stretch of road on an average day of the year. For information on how AADFs are calculated, see the guidance on the Traffic Statistics pages on GOV.UK.

Columns for traffic data file

* AADFYear = The year data was recorded
* CP = Count point. A unique reference for the road link that links the AADFs to the road network
* Estimation\_method = Counted or estimated
* Estimation\_method\_detailed = How it was counted, or how it was estimated
* Region = 11 regions, essentially Wales, Scotland and then 9 English regions. No data from N.Ireland.
* LocalAuthority = Breaking down the regions further
* Road = Roads are either M for Motorway (the biggest roads), or A which are our major roads. B roads aren't included
* RoadCategory = See details below\*
* Easting = Cartesian X co-ordinate (lat and lon are in the last two columns)
* Northing = Cartesian Y co-ordinate (lat and lon are in the last two columns) \* Easting and Northing need to use epsg=27700. Explained in the Kernel 'Using Basemap for geographic data' <https://www.kaggle.com/daveianhickey/using-basemap-for-geographical-data/>
* StartJunction = Where the stretch of road begins
* EndJunction = Where the stretch ends
* LinkLength\_km = How far between the start and end
* LinkLength\_miles = How far between the start and end
* PedalCycles = Avergae volume per day
* Motorcycles = Avergae volume per day
* CarsTaxis = Avergae volume per day
* BusesCoaches = Avergae volume per day
* LightGoodsVehicles = Avergae volume per day
* V2AxleRigidHGV = HGV is a Lorry/Truck. It stands for Heavy Goods Vehicle. The number of Axles lets you know how big it is
* V3AxleRigidHGV
* V4or5AxleRigidHGV
* V3or4AxleArticHGV
* V5AxleArticHGV
* V6orMoreAxleArticHGV
* AllHGVs
* AllMotorVehicles
* Lat = Latitude
* Lon = Longitude

\*Road Category Description

* PM = M or Class A Principal Motorway
* PR = Class A Principal road in Rural area
* PU = Class A Principal road in Urban area
* TM = M or Class A Trunk Motorway
* TR = Class A Trunk road in Rural area
* TU = Class A Trunk road in Urban area

Missing data

South East is missing Estimation method

What about miles driven?

The UK government also like to look at miles driven. You can do this by multiplying the AADF by the corresponding length of road (link length) and by the number of days in the years.

Accident Columns

* Accident\_Index = Unique ID
* Location\_Easting\_OSGR = Local British coordinates
* Location\_Northing\_OSGR = Local British coordinates
* Longitude
* Latitude
* Police\_Force
* Accident\_Severity = The higher the number, the worse it is
* Number\_of\_Vehicles
* Number\_of\_Casualties
* Date = dd/mm/yyyy
* Day\_of\_Week = 1 is Monday, 2 is Tuesday, etc
* Time = GMT/UTC (Grenwich Mean Time and Coordinated Universal Time are the same)
* Local\_Authority\_(District)
* Local\_Authority\_(Highway)
* 1st\_Road\_Class = This field is only used for junctions
* 1st\_Road\_Number = This field is only used for junctions
* Road\_Type = Roundabout, One way, Dual Carriageway, Single carriageway, slip road, unknown
* Speed\_limit
* Junction\_Detail = Crossroads, roundabouts, private roads, not a junction, etc
* Junction\_Control = A person, a type of sign, automated, etc.
* 2nd\_Road\_Class = This field is only used for junctions
* 2nd\_Road\_Number = This field is only used for junctions
* Pedestrian\_Crossing-Human\_Control = Was there a human controller and what type?
* Pedestrian\_Crossing-Physical\_Facilities = was it a zebra crossing, or bridge, or another type
* Light\_Conditions = Day, night, street lights or not
* Weather\_Conditions = Wind, rain, snow, fog
* Road\_Surface\_Conditions = Wet, snow, ice, flood
* Special\_Conditions\_at\_Site = Was anything broken or defective, e.g. an obscured sign?
* Carriageway\_Hazards = Was something in the way, e.g. a pedestrian, another accident, something in the road?
* Urban\_or\_Rural\_Area
* Did\_Police\_Officer\_Attend\_Scene\_of\_Accident
* LSOA\_of\_Accident\_Location

2008 is missing.

Get more information

* Licence = Open Givernment Licence used by all data on data.gov.uk, <http://www.nationalarchives.gov.uk/doc/open-government-licence/version/3/>
* The data sets, <https://www.dft.gov.uk/traffic-counts/download.php>
* Related metadata, <http://data.dft.gov.uk/gb-traffic-matrix/all-traffic-data-metadata.pdf>