



Department
for Transport

RAP Validation

Where did it all go so wrong?





Disclaimer!

I'm quite new to R...



Agenda

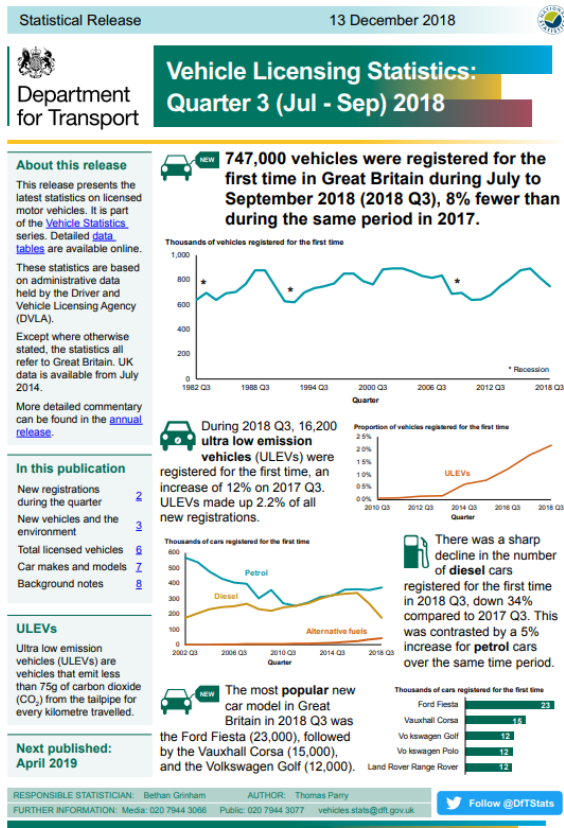
- ▶ Overview of the statistics and the current
- ▶ Demo SQL stored procedure scripts
- ▶ Demo R scripts
 - ▶ Import data
 - ▶ Validation steps
 - ▶ Table production steps
- ▶ Upcoming issues
- ▶ Thoughts and suggestions

Questions as we go...





Vehicle Licensing Statistics



Department for Transport statistics

Vehicle Licensing Statistics (<https://www.gov.uk/government/collections/vehicles-statistics>)

Table VEH0150

Vehicles registered for the first time by body type, including average CO₂ emissions for cars and breakdown by ULEVs and plug-in grant eligible

The Department for Transport are running a short survey on the use of table VEH0150 to understand our user needs and requirements. If you wish to comment, please email https://www.smartsurvey.co.uk/s/EOPQW/

Great Britain				Thousand	Percentage	Grams per kilometre			
			Light Goods Vehicles	Heavy Goods Vehicles	Buses & coaches	Other vehicles	Total	Year-on-year change in total vehicles ³	Average CO ₂ emissions for newly registered cars
Date	Cars	Motorcycles							
2001	2586.0	180.5	274.0	52.5	10.6	34.2	3137.7		
2002	2682.1	165.8	283.6	48.7	12.3	36.9	3229.4	2.9	
2003	2646.1	160.9	317.9	53.2	13.7	40.2	3231.9	0.1	172.6
2004	2599.1	137.1	341.0	53.1	13.2	41.9	3185.4	-1.4	171.3
2005	2443.5	135.6	329.7	56.9	13.1	42.6	3021.4	-5.2	169.7
2006	2340.0	135.2	328.3	54.5	12.3	43.3	2913.6	-3.6	167.7
2007	2390.1	146.4	340.1	47.3	12.1	60.9	2996.9	2.9	164.7
2008	2112.0	142.0	291.0	52.8	11.7	62.8	2672.2	-10.8	158.2
2009	1968.3	114.9	189.7	31.4	9.5	57.4	2371.2	-11.3	149.7
2010	1996.3	100.1	226.1	30.3	8.9	56.1	2417.8	2.0	144.3
2011	1907.4	99.3	263.0	40.7	8.5	62.5	2381.5	-1.5	138.2
2012	2010.8	100.1	242.4	41.7	9.8	65.0	2469.8	3.7	133.0
2013	2225.1	98.4	274.4	53.2	9.1	56.0	2716.1	10.0	128.4
2014	2438.3	108.3	324.4	38.5	8.3	55.8	2973.7	9.5	124.8
2015	2602.1	123.4	374.9	49.1	9.6	52.9	3212.0	8.0	121.4
2016	2665.3	138.2	378.8	51.3	10.0	52.4	3296.0	2.6	120.3
2017	2509.3	113.6	364.5	50.7	8.7	56.1	3102.9	-5.9	121.2
2001 Q1	707.9	39.9	69.5	11.9	2.6	7.3	839.1		
2001 Q2	618.6	55.1	66.9	13.2	2.8	9.5	766.0		



Vehicle Licensing Statistics

It's chunky

- ▶ 76 tables due out for annual in April
- ▶ 19 tables for quarterly releases
- ▶ Two heavily related but different data sources
 - ▶ New registrations
 - ▶ Vehicle stock

Quick stats

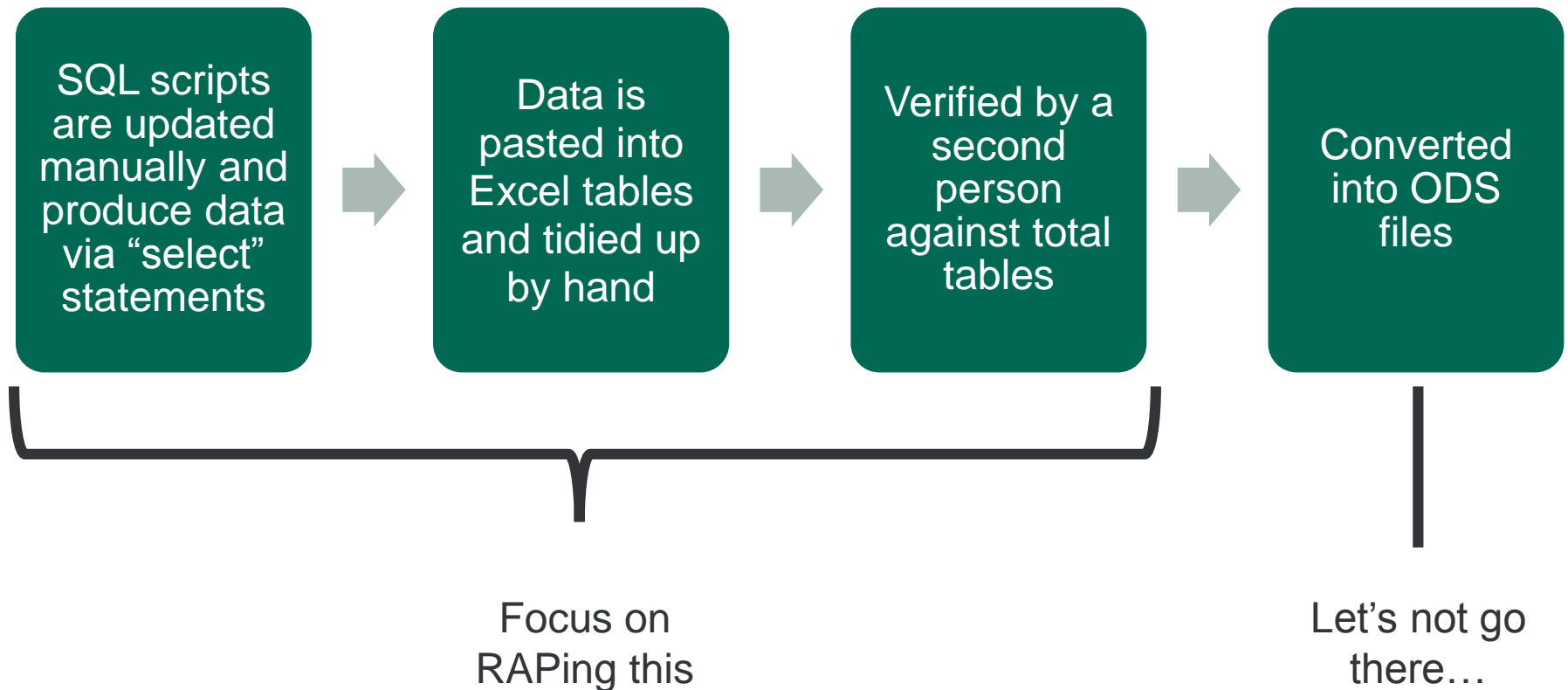
- ▶ 747,000 vehicles were registered for the first time in 2018 Q3 in GB
- ▶ At the end of September 2018, there were 38.4 million licensed vehicles in Great Britain

Quick stats about the stats

- ▶ The largest table we publish is VEH0220, licensed cars by make, model, fuel type and engine size:
 - ▶ 54Mb
 - ▶ 57 columns
 - ▶ circa 48k rows
 - ▶ 8 versions in the file
- ▶ GB licensed stock data tables refer to 1.9 billion vehicle records
- ▶ GB SORN stock data tables refer to 110 million vehicle records



Current process





Department
for Transport

Demo - SQL



Department
for Transport

Demo - R



Upcoming Issues

- ▶ Time!
- ▶ Trying to apply a consistent approach to tables that are structured differently – do we change their structure?
- ▶ How to present GB vs UK data
- ▶ RAP gives us tables to spare (e.g. HGV by fuel type)
- ▶ Current validation output is helpful for 1 table but not for 76!
- ▶ How do we convey all these changes to users?





Department
for Transport

Questions?
Suggestions?