

# ESTUDIO DEL CONSUMO DE VEHÍCULOS ESTATALES DE REINO UNIDO

Knowledge Graphs

Laura García Perrín, Prisco García-Consuegra Martín, Mario Serrano  
Sánchez, Xiya Sun y Yimin Zhou

Universidad Politécnica de Madrid

December 14, 2021

# Introducción

## Assignment 1

### Objetivos:

- Análisis del consumo de vehículos
- Evaluación marcas de coche
- Proporcionar información sobre proyecciones de mercado
  - ▶ Sector automovilístico
  - ▶ Sectores regulados: banca



# Dataset

## Assignment 1

data.gov.uk | Find open data

- fuel-performance-april-19-march-20.csv
- Fichero .csv: identificador, matrícula, modelo, combustible, cuentakilómetros, distancia recorrida, unidad de medida, MPG

Number	Registration	Details	Product	Odometer	Distance	Unit	MPG
0001	U1	Jaguar XJL Portfolio V6 D Limousine	Diesel	67691	0	Miles	0
0012	FL65FBA	Vauxhall CORSA CDTi Car	Diesel	9329	2849	Miles	50.04
0013	MJ16XMY	Vauxhall Mokka 4x4 Utility Vehicle	Diesel	29064	5369	Miles	39.29
0016	BP16MKK	Nissan Note Car	Diesel	42233	1052	Miles	56.17
0017	YD16NAL	Peugeot car Car	Diesel	20475	3319	Miles	108.32
0018	YD16NAK	Peugeot 308 Car Car	Diesel	115020	36100	Miles	50.79
0019	BP16OPD	Nissan Note Car	Diesel	15559	948	Miles	36.04
0020	YH17VCF	Peugeot Expert Combi 9 seat MPV	Diesel	37352	14202	Miles	34.75
0021	MF68FNR	Renault Traffic Welfare	Diesel	112	0	Miles	0
0023	YE19FLS	Volkswagon Caddy Maxi Welfare	Diesel	16270	16244	Miles	37.09
0024	YG69UKP	Caddy Maxi Life Wav	Diesel	8103	8066	Miles	42.22
0025	YG69UKR	Caddy Maxi Life Van	Diesel	6187	6160	Miles	36.93
0026	YH69ZYT	Volkswagen Caddy Maxi Life Van	Diesel	6378	6361	Miles	35.21
0027	YL69KJN	VW Caddy	Diesel	1060	1032	Miles	26.6
0028	YL69KJV	VW Caddy	Diesel	360	307	Miles	15.77
0029	YL69KJO	VW Caddy	Diesel	967	884	Miles	24.48
0030	YL69KJU	VW Caddy	Diesel	925	895	Miles	28.76
0031	YG20GRF	Volkswagen Caddy 7 seater	Diesel	691	656	Miles	20.03
0032	YG20AHU	Volkswagen Caddy 7 seater	Diesel	366	299	Miles	16.18
0033	YG20GRU	Volkswagen Caddy 7 seater	Diesel	1075	1044	Miles	29.36

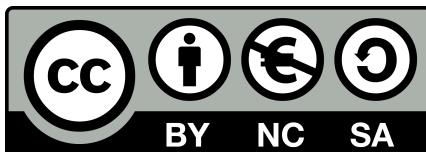
# Analysis of the Data

## Assignment 2

- **Información general**

- ▶ **Licencia del dataset:** Open Government v3.0
- ▶ **Titular de los derechos:** UK Government
- ▶ **Editor:** The National Archives of the UK Government

- **Nuestra licencia** vamos a utilizar creative commons license (CC BY-NC-SA), que establece una serie de condiciones.



# Data characteristics

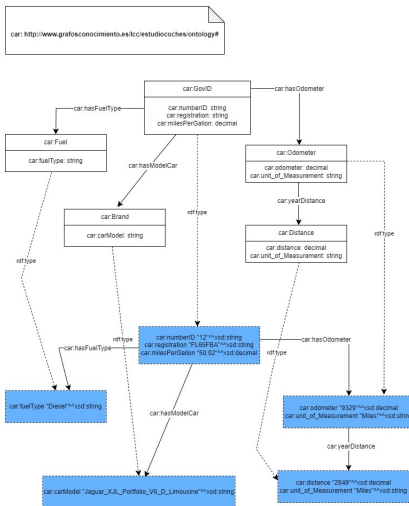
## Assignment 2

Column	Type	Range	Problems
Number	String	Unique? Empty values	Repeated values
Registration	String	Unique? Empty values (1 cell)	Some registrations are repeated
Details (vehicle's model)	String	Not unique (293 models), empty values	Eight cells with empty values
Product (type of fuel)	String	Not unique (2 types considering that gasoil and diesel are the same type), empty values	
Odometer (Km)	Decimal	0-424905	Some cells with negative values (discarded)
Distance (Km)	Decimal	0-49463	Some cells with negative values (discarded)
Unit	String	Not Unique (Hours, Kilometres, Miles, No Limit), empty values	Two cells with the value No Limit?
Miles per Gallon (MPG)	Decimal	0-824,85	Some cells with negative values (discarded)

**Note:** For the analysis of these data, we have discarded the last cells which doesn't provide any information to the data. We also discarded those cells with negative values in the parameters of Distance, Odometer and Miles per Gallon.

# Data characteristics

## Assignment 2



# De modelos a marcas

## Assignment 3

### Problemas:

- Modelos ambiguos y nada específicos
- La solución a nuestra cuestión era trivial
- La reconciliación sería imposible

Registration	Details	Fuel type	Occ
YA19TMX	Large Van	Diesel	6234
YA19TNF	Large Van	Diesel	6714
YA19TNJ	Large Van	Diesel	9000
YK19TDG	Large Van	Diesel	11238
YB19GUE	Large Van	Diesel	15559
YA19TNN	Large Van	Diesel	8435
YA19TNL	Large Van	Diesel	8892
YB19GTZ	Large Van	Diesel	4356

Registration	Details	Fuel type	Odometer
96YX	Welfare Bus	Diesel	9110
96YY	Welfare Bus	Diesel	7384
96YZ	Welfare Bus	Diesel	15509
96YW	Welfare Bus	Diesel	7284
96ZB	Welfare Bus	Diesel	4897
96YR	Welfare Bus	Diesel	1766
96YN	Welfare Bus	Diesel	3448
96YM	Welfare Bus	Diesel	4145
96ZE	Welfare Bus	Diesel	4847

Details	
145 choices Sort by: name count Cluster	
18t insulated	
Tipper/Gritter/Snowplough 5	
26t Tipper/Gritter/Snowplough 1	
3 Axle Bin 4	
3 Axle RCV 33	
50412 1	
7.5t Cage 2	
7.5t Tipper 2	
Coachbuilt Bus 1	

Details	
13 choices Sort by: name count Cluster	
Fiat 15	
Ford Motor Company 280	
Iveco 28	
John Deere 2	
Mercedes 105	
Mitsubishi Fuso Truck and Bus Corporation 19	
Mitsubishi Motors 2	
Nissan 2	
Peugeot 169	
Renault 63	
Vauxhall 87	
Volkswagen 138	
Volvo 2	

# Modificaciones de los datos

## Assignment 3

Hemos eliminado todos los datos que no nos eran útiles para nuestro estudio

- 1 Vehículos sin marca
- 2 Matrículas repetidas
- 3 Datos en blanco
- 4 etc.



# OpenRefine



# RDF Generation

## Assignment 4

- Mapping csv-rdf
  - ▶ R2RML y RML
- Helio
- RDF



```
<http://www.grafosconocimiento.es/resources/Distance/6273>
  <http://www.grafosconocimiento.es/lcc/ontology/estudiocoches#distance>
    "6273"^^<http://www.w3.org/2001/XMLSchema#string> ;
  <http://www.grafosconocimiento.es/lcc/ontology/estudiocoches#unit_of_measurement>
    "Miles"^^<http://www.w3.org/2001/XMLSchema#string> .
```

```
<#TriplesMapS>
  rml:logicalSource [
    rml:source "./data/fuel-performance-update.csv" ;
    rml:referenceFormulation ql:CSV
  ];
  rr:subjectMap [
    rr:template "http://www.grafosconocimiento.es/resources/Distance/{Distance}";
    rr:termType rr:IRI;
    rr:class car:Distance
  ];

  rr:predicateObjectMap [
    rr:predicate car:distance;
    rr:objectMap [
      rml:reference "Distance";
      rr:datatype xsd:decimal
    ]
  ];

  rr:predicateObjectMap [
    rr:predicate car:unit_of_measurement;
    rr:objectMap [
      rml:reference "Unit";
      rr:datatype xsd:string
    ]
  ]
];
```

# Data linking y consulta

## Assignment 5

En este apartado, ha sido necesario realizar las siguientes acciones:




- Identificar aquellas clases cuyas instancias podían enlazarse.
  - ▶ Marca de coche
  - ▶ Tipo de combustible
- Reconciliar los datos con los datasets identificados. En nuestro caso, con wikidata.

Number	Registration	Details	same_as_wikidata_marcas	Product	same_as_wikidata_combustibles	Odometer	Distance	Unit	MPG
12	FL65FBA	Vauxhall	<a href="https://wikidata.org/entity/Q59187">https://wikidata.org/entity/Q59187</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	9329	2849 Miles		50.04
13	MJ16XMY	Vauxhall	<a href="https://wikidata.org/entity/Q59187">https://wikidata.org/entity/Q59187</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	29064	5369 Miles		39.29
16	BP16MKK	Nissan	<a href="https://wikidata.org/entity/Q20165">https://wikidata.org/entity/Q20165</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	42233	1052 Miles		56.17
18	YD16NAK	Peugeot	<a href="https://wikidata.org/entity/Q6742">https://wikidata.org/entity/Q6742</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	115020	36100 Miles		50.79
19	BP16OPD	Nissan	<a href="https://wikidata.org/entity/Q20165">https://wikidata.org/entity/Q20165</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	15559	948 Miles		36.04
20	YH17VCF	Peugeot	<a href="https://wikidata.org/entity/Q6742">https://wikidata.org/entity/Q6742</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	37352	14202 Miles		34.75
23	YE19FLS	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	16270	16244 Miles		37.09
24	YG69UKP	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	8103	8066 Miles		42.22
25	YG69UKR	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	6187	6160 Miles		36.93
26	YH69ZYT	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	6378	6361 Miles		35.21
31	YG20GRF	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	691	656 Miles		20.03
32	YG20AHU	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	366	299 Miles		16.18
33	YG20GRU	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	1075	1044 Miles		29.36
34	YG20GRK	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	784	756 Miles		26.26
35	YG20VYZ	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	393	318 Miles		18.5
39	YG20WFR	Volkswagen	<a href="https://wikidata.org/entity/Q246">https://wikidata.org/entity/Q246</a>	Diesel	<a href="https://wikidata.org/entity/Q38423">https://wikidata.org/entity/Q38423</a>	350	300 Miles		17.98

# Data linking y consulta

## Assignment 5

- Publicación del dataset final con links
- Helio
- Consultas que verifiquen los links enlazados

 helio-storage	13/12/2021 16:19	Carpeta de archivos
 mappings	30/11/2021 16:22	Carpeta de archivos
 publisher-0.3.12	30/11/2021 16:26	Executable Jar File

