Ekologia i ekonomia Samochodów

Michał Loska & Adrian Wysocki
15.12.2020

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
AGH University of Science and Technology

Zbiór danych: https://www.kaggle.com/mrmorj/car-fuel-emissions

Zbiór danych

- Dane opublikowane przez departament transportu w Wielkiej Brytanii
- Samochody z lat 2000 – 2013
- Dostępne atrybuty:
  - Marka oraz model
  - Rozmiar silnika
  - Paliwo (benzyna, diesel)
  - Skrzynia biegów (manualna, automatyczna)
  - Emisje spalin (emisje CO2 oraz norma EURO)
  - Spalanie (średnie w cyklu mieszanym)

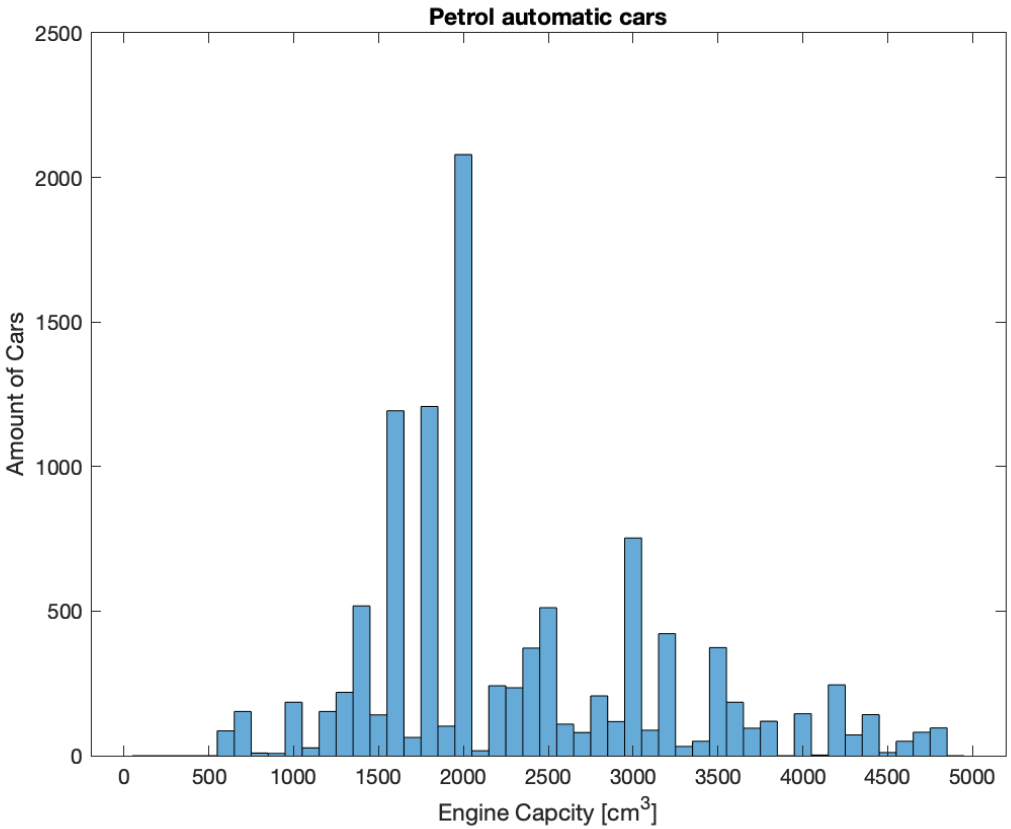
kod projektu: https://github.com/michalloska/DataExploration_Project

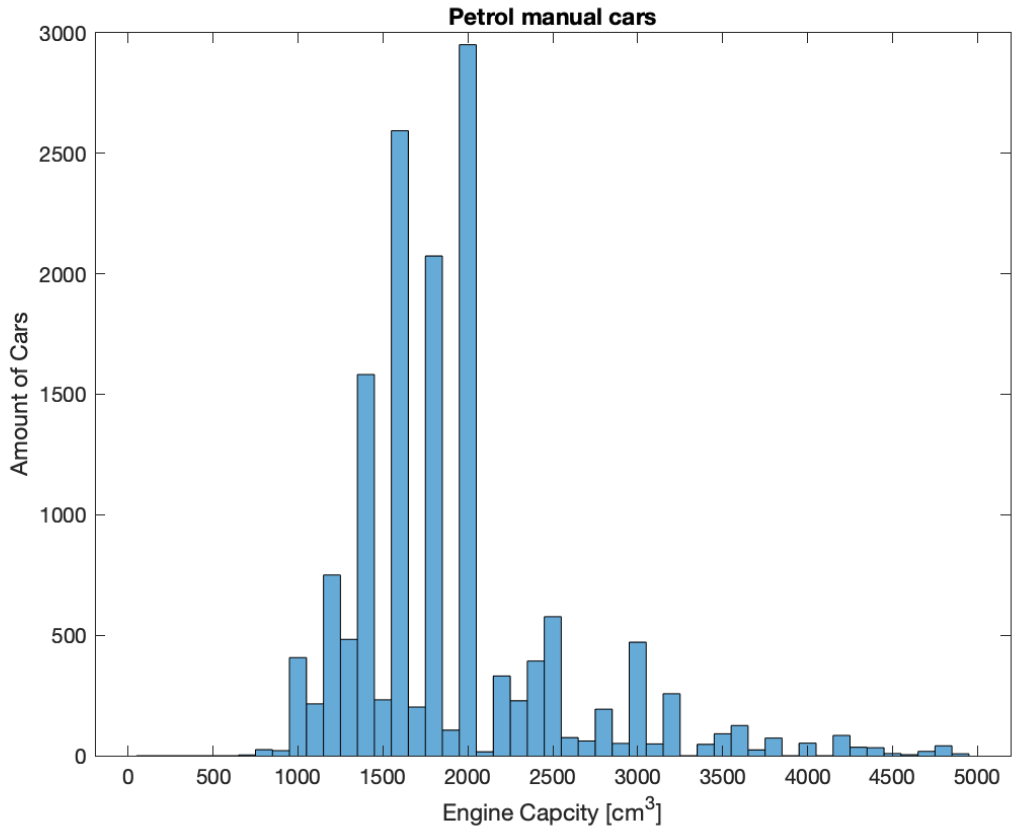
Przeprowadzone operacje

na zbiorze danych

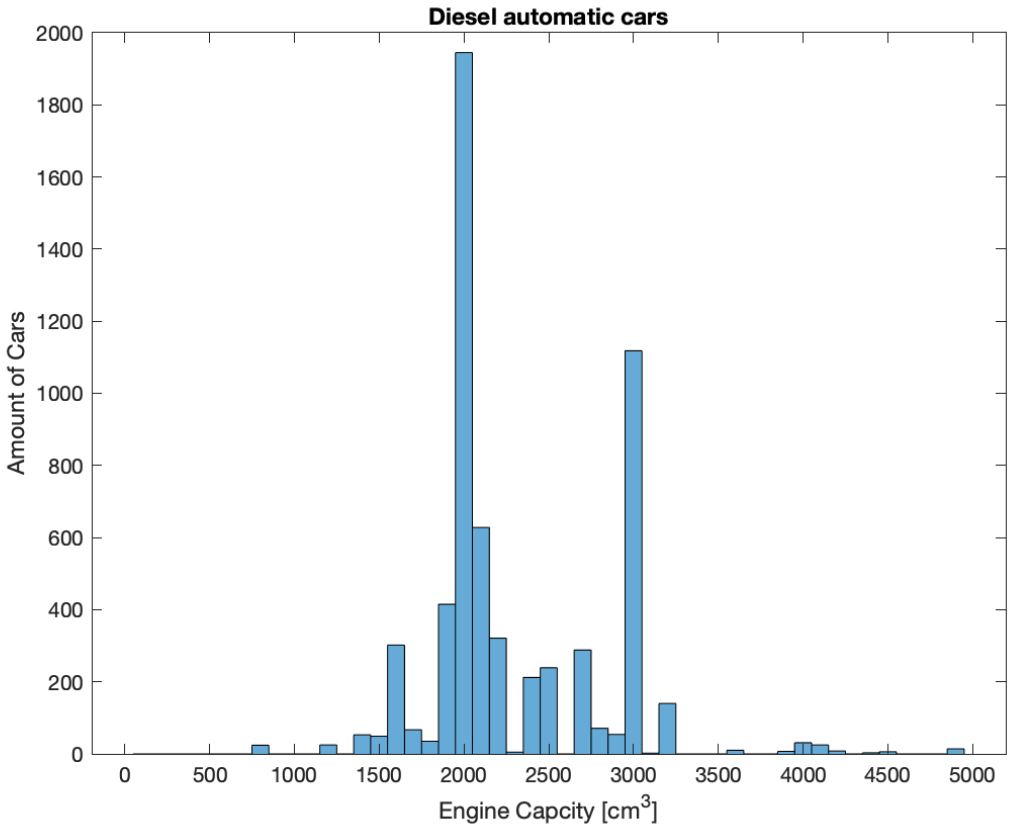
- Porównanie samochodów względem rodzaju paliwa oraz skrzyni biegów
- Prezentacja emisji spalin CO^2 dla wszystkich samochodów
- Porównanie spalania w zależności od wielkości silnika i paliwa
- Prezentacja elementów nietypowych i błędnych danych
- Redukcja wymiaru do 2D (dla wstępnego wyznaczenia il. Klastrów)
- Klasteryzacja i prezentacja ich zawartości
- Klasyfikacja
- Macierz pomyłek (Confusion Matrix)

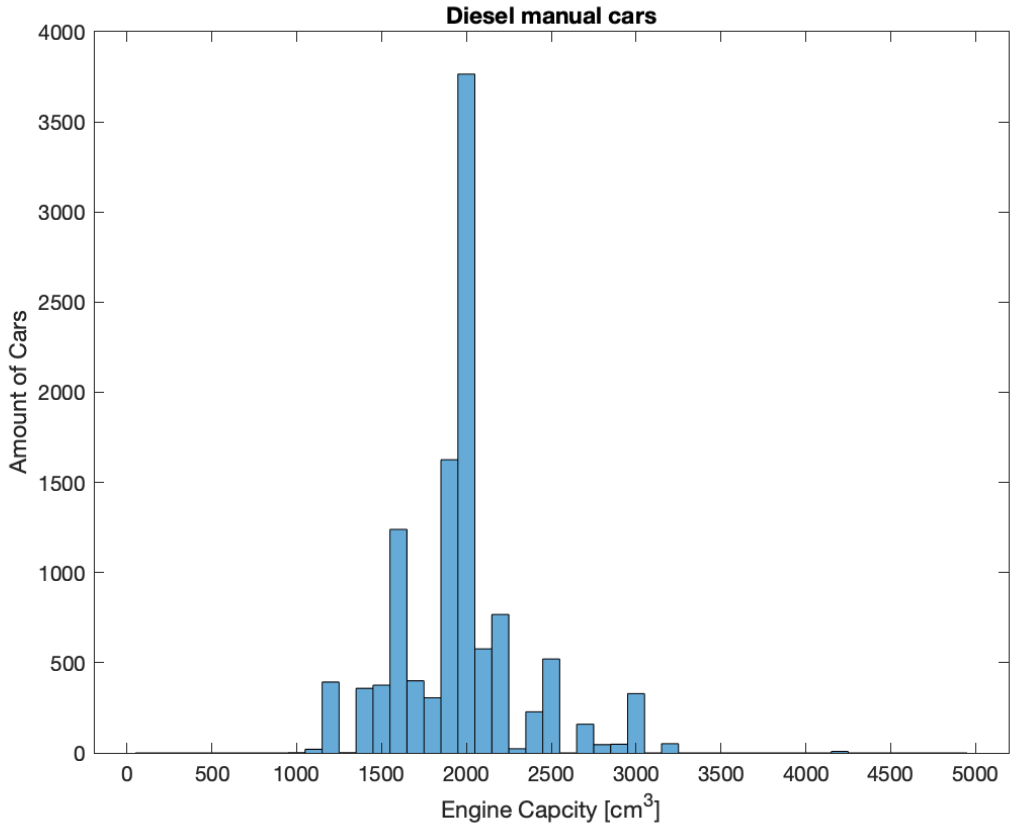
Samochody benzynowe



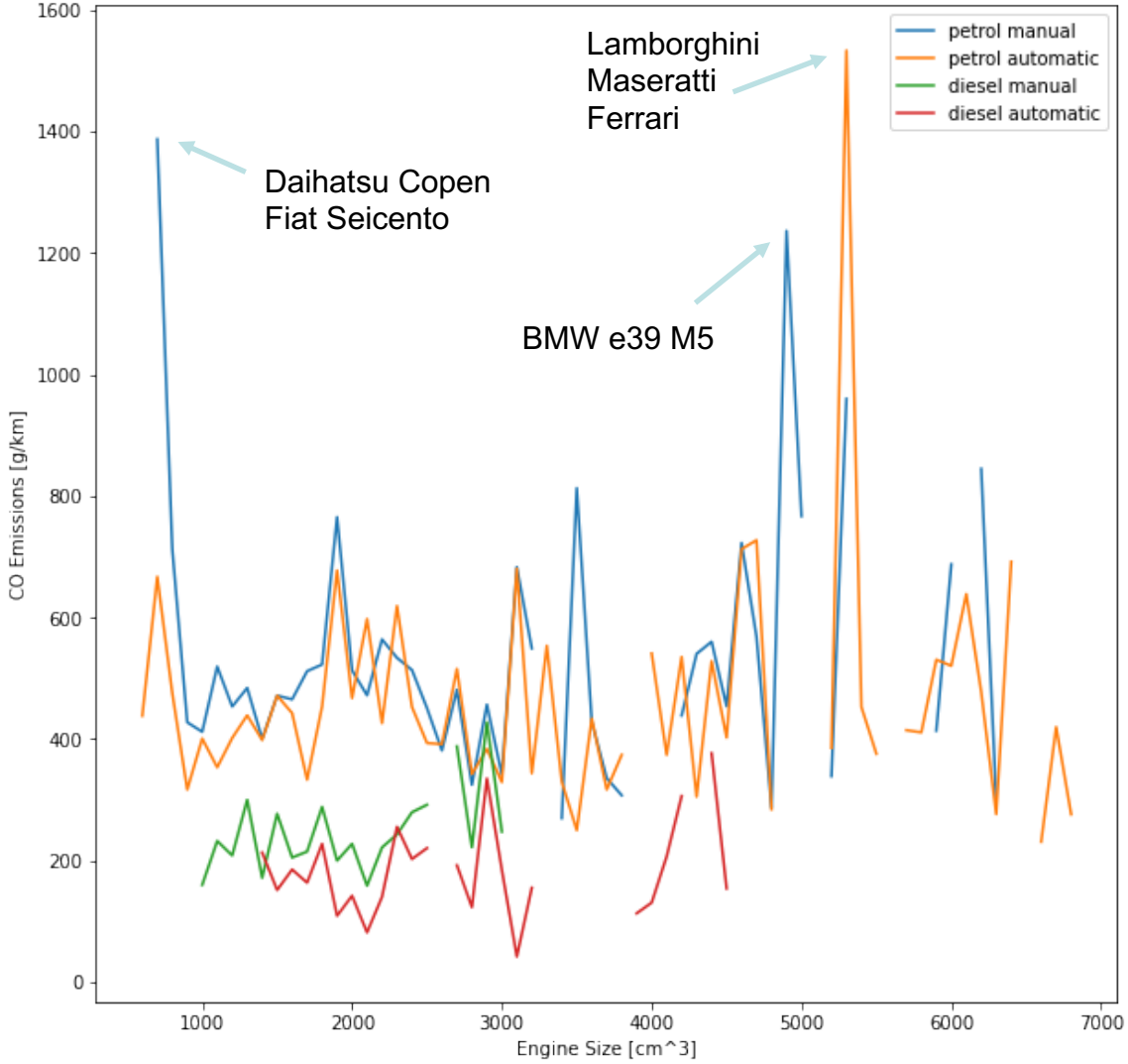


Samochody z silnikiem diesla





Emisje spalin dla całego zbioru danych



Daihatsu Copen

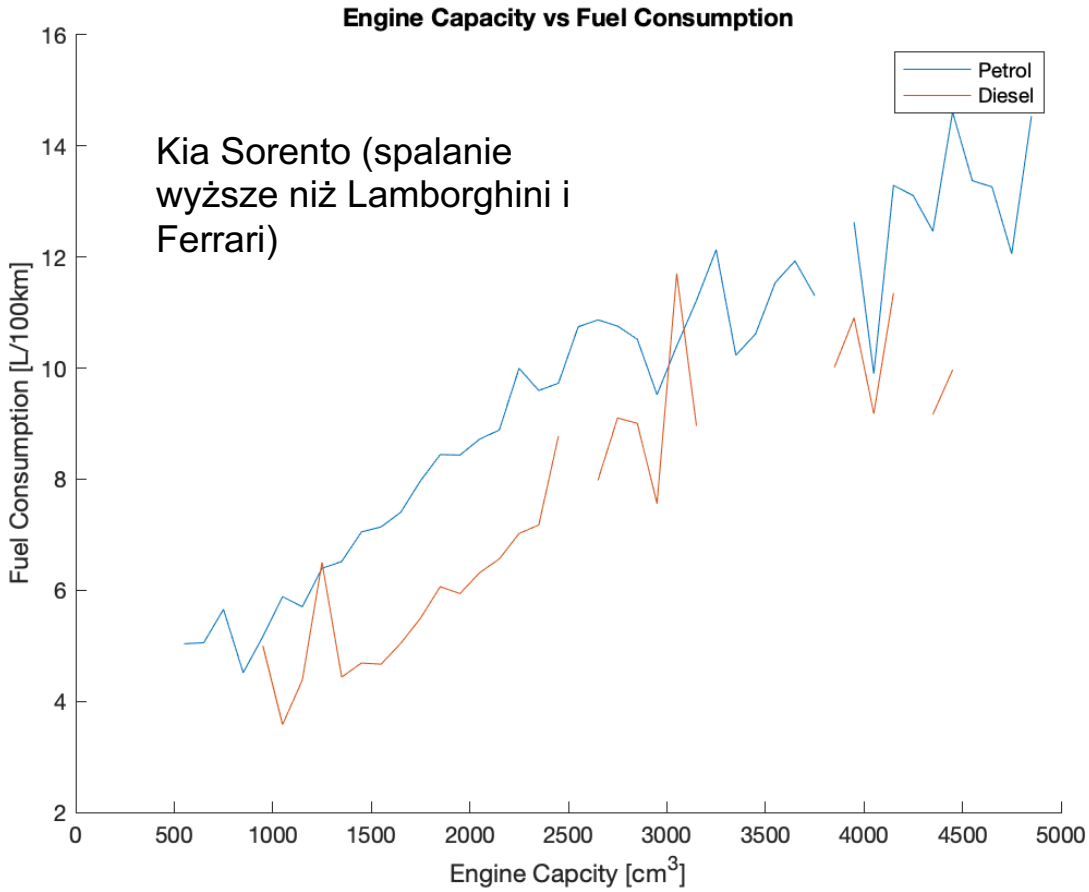
Regulacje Kei Car

First generation (L880; 2002) [edit]

The Copen was originally designed with a 660 cc turbocharged engine in order to meet Japanese kei car regulations. However, since this engine did not meet environmental emissions standards in several other countries, the Copen was fitted with a more powerful 1.3 L non-turbo engine in these markets in 2007. Styling was inspired by the older Nissan Figaro, which has key elements found in the Copen.

Spalanie diesel a benzyna

W cyklu mieszanym



Usunięte dane:

	year	manufacturer	model	description	euro_standard	transmission_type	engine_capacity	fuel_type	combined_metric	co2	co_emissions
45356	2013	Chrysler Jeep	Jeep Compass, MY2014	2.4 4x4	6	NaN	209.0	Petrol	9.0	209	366.0
3574	2001	Toyota	Landcruiser	4.2 24v TD Auto	2	Automatic	4164.0	Diesel	12.6	340	170.0
1390	2000	Toyota	Landcruiser	4.2 TD Auto	2	Automatic	4164.0	Diesel	12.6	340	170.0
25497	2008	Volkswagen	Touareg	5.0 TDI V10 (313 PS) Tiptronic with DPF	4	Automatic	4921.0	Diesel	12.6	333	424.0
20390	2007	Kia	Sorento	2.5 CRDi	4	Manual	2497.0	Diesel	35.8	209	302.0
42908	2013	Mercedes-Benz	B-Class, Model Year 2013	B 180 CDI with 16" rear wheels	5	Manual	1796.0	Diesel	44.0	114	348.0
12370	2005	Audi	Audi TT (Coupé)	1.8 T quattro (225 PS)	3	Manual	1781.0	Petrol	9.4	226	73000.0
12376	2005	Audi	Audi TT (Roadster)	1.8 T quattro (225 PS)	3	Manual	1781.0	Petrol	9.5	228	73000.0
12369	2005	Audi	Audi TT (Coupé)	1.8 T quattro (180 PS)	3	Manual	1781.0	Petrol	9.5	228	74000.0
12375	2005	Audi	Audi TT (Roadster)	1.8 T quattro (180 PS)	3	Manual	1781.0	Petrol	9.6	230	74000.0

Usunięte dane:

	year	manufacturer	model	description	euro_standard	transmission_type	engine_capacity	fuel_type	combined_metric	co2	co_emissions
12339	2005	Audi	A8	4.0 TDI V8 quattro (275 PS) LWB	3	Automatic	3936.0	Diesel	9.8	265	70000.0
12340	2005	Audi	A8	4.0 V8 TDI quattro (275 PS)	3	Automatic	3936.0	Diesel	9.8	265	70000.0
12352	2005	Audi	Audi Allroad (Up to June 2005)	2.5 TDI V6 quattro (180 PS)	3	Automatic	2496.0	Diesel	9.5	257	71000.0
12353	2005	Audi	Audi Allroad (Up to June 2005)	2.5 TDI V6 quattro (180 PS)(Low & High Ratio T...	3	Manual	2496.0	Diesel	8.8	238	72000.0

Usunięte dane:

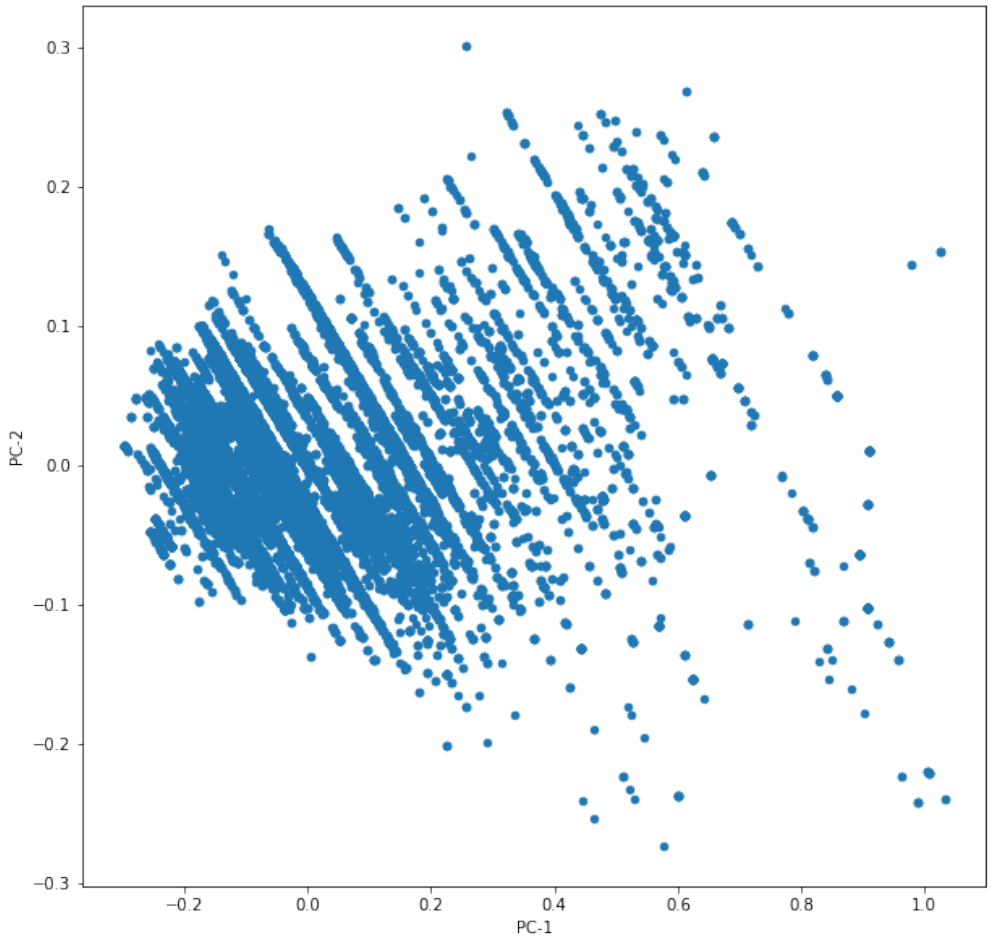
	year	manufacturer	model	description	euro_standard	transmission_type	engine_capacity	fuel_type	combined_metric	co2	co_emissions	fuel_cost
45356	2013	Chrysler Jeep	Jeep Compass, MY2014	2.4 4x4	6	NaN	209.0	Petrol	9.0	209	366.0	
2872	2001	Micro Compact Car	MCC Smart City Coupe Hatchback	Smart and Passion	3	Automatic	599.0	Petrol	4.9	118	343.0	
8265	2003	Smart	Smart Cabrio Hatchback (03MY)	Smart Cabrio and Pulse	3	Automatic	599.0	Petrol	4.7	113	500.0	
8266	2003	Smart	Smart Cabrio Hatchback (03MY)	Smart Cabrio and Pulse	3	Automatic	599.0	Petrol	4.9	118	300.0	
14574	2005	Smart	Smart Cabrio	Crossblade	3	Automatic	599.0	Petrol	5.7	136	523.0	

Usunięte dane:

	year	manufacturer	model	description	euro_standard	transmission_type	engine_capacity	fuel_type	combined_metric	co2	co_emissions	fu
11491	2004	Vauxhall	Astra, Model Year 2003	2.0 DTI 5 Door Estate, From VIN: W0L0TGF353800...	3	Manual	1686.0	Diesel	6.9	186	-200.0	
13313	2005	Land Rover	Discovery 3	2.7 V6	3	Manual	2720.0	Diesel	9.4	249	608.0	
4273	2002	Fiat	Doblo Range	1.9 JTD	3	Manual	1910.0	Diesel	6.4	168	611.0	
2540	2001	Kia	Sedona	2.9 TDI	2	Automatic	2902.0	Diesel	9.1	240	620.0	
689	2000	Kia	Sedona	2.9 TDI S/SX/GSX/Executive	2	Automatic	2902.0	Diesel	9.1	240	620.0	
...	...	...	...	...	...	...	...	...	...	...	...	
2385	2001	Ford	Transit Tourneo SWB Bus	2.5 Di Diesel (76PS) 4.11 FDR	2	Manual	2496.0	Diesel	8.3	218	1274.0	

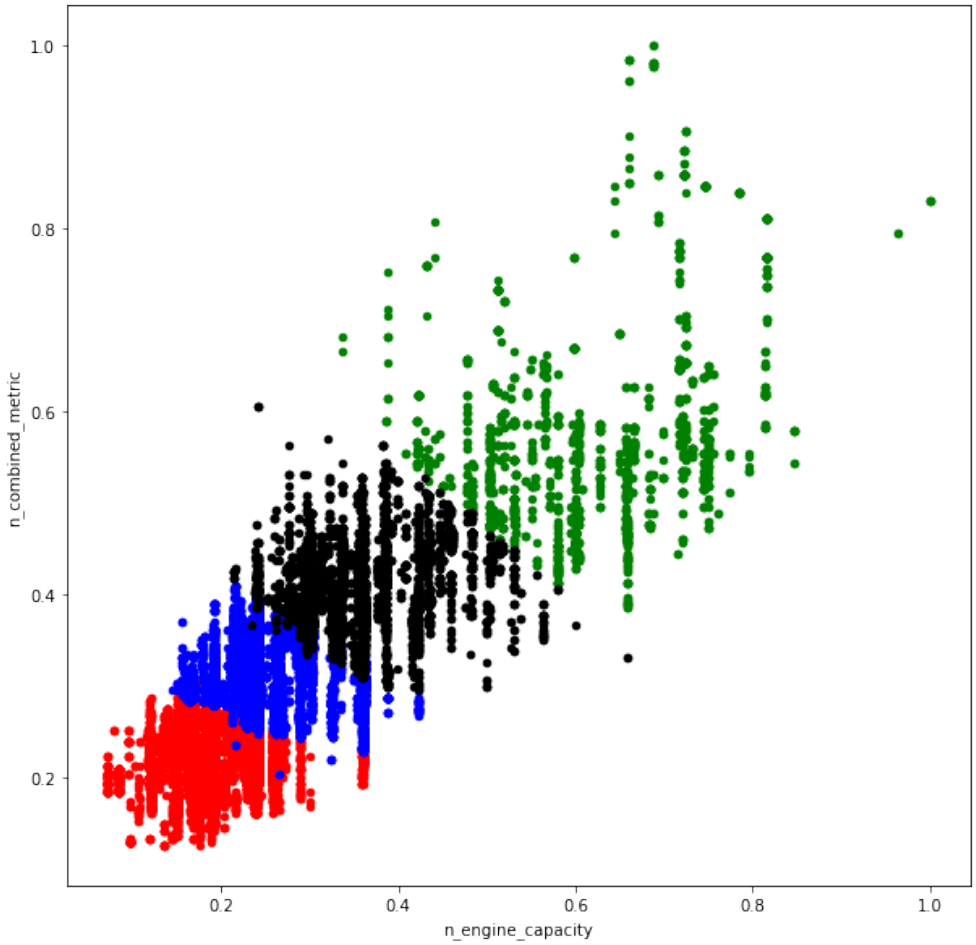
Redukcja wymiarów do 2D

Celem wstępnego oszacowania ilości potrzebnych klastrów



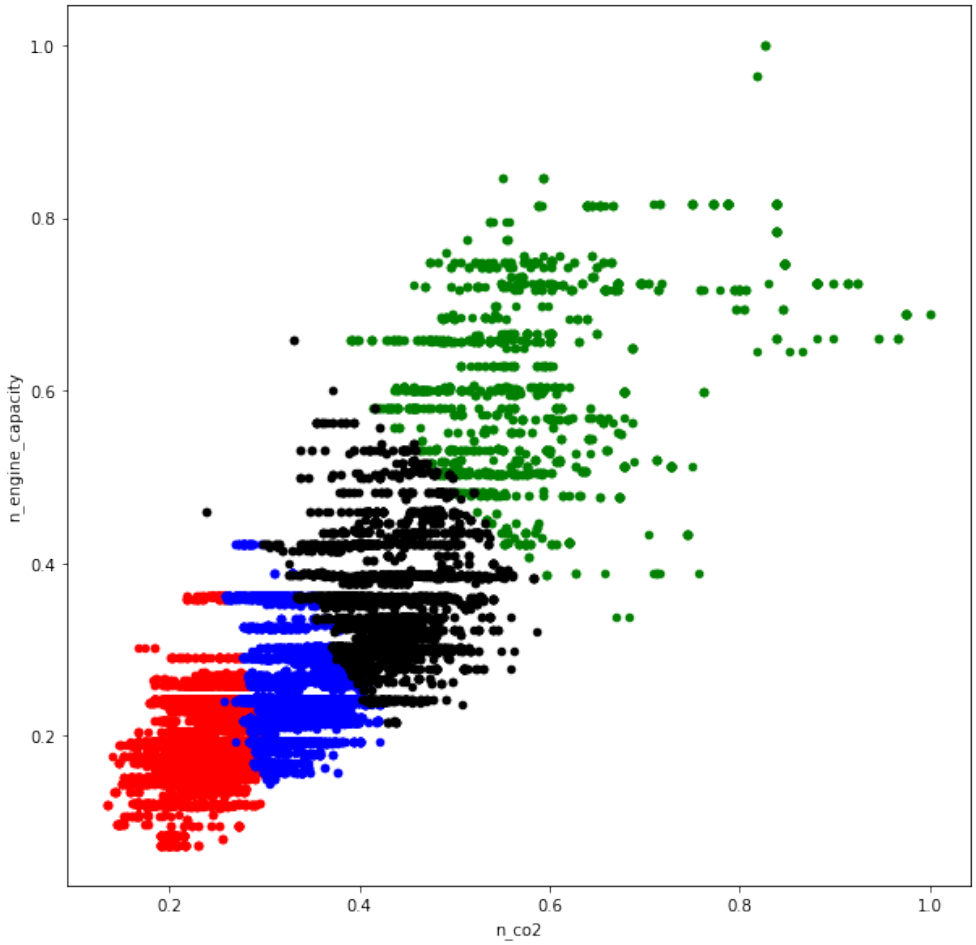
Klasteryzacja

rozmiar silnika a spalanie



Klasteryzacja

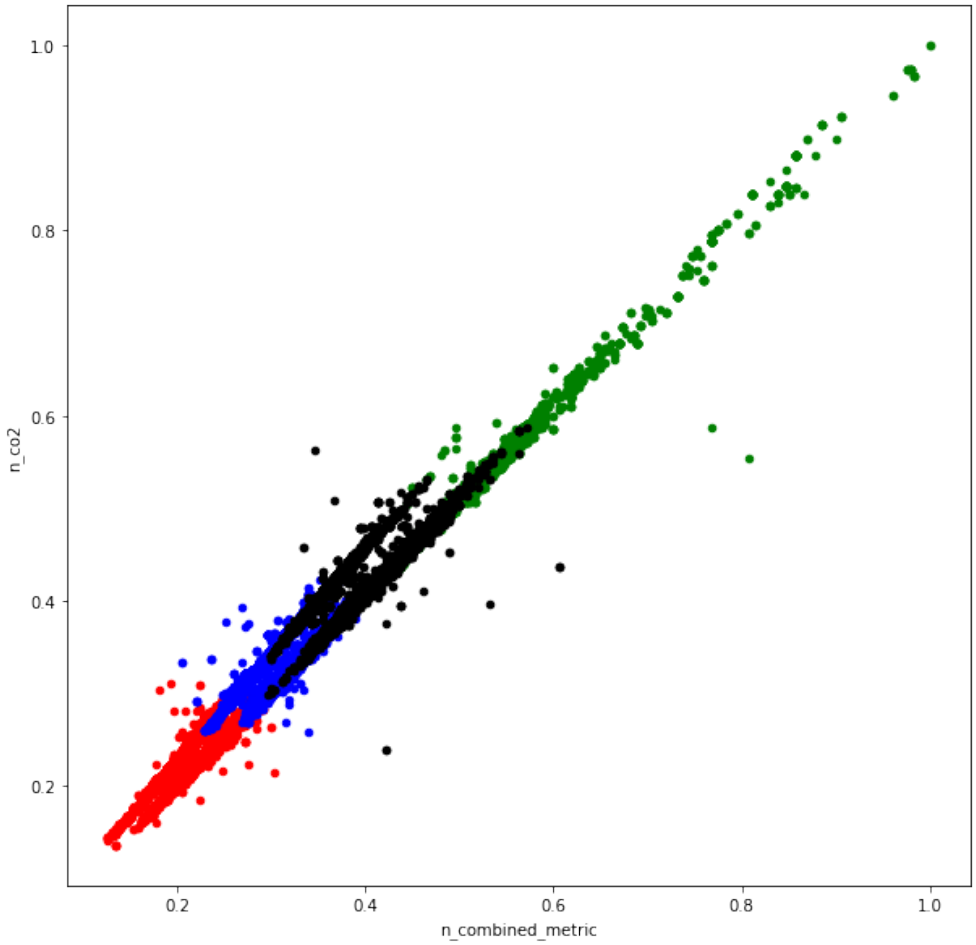
rozmiar silnika a emisje spalin



Klasteryzacja

spalanie a emisje spalin

Emisje CO^2 mocno skorelowane ze spalaniem oraz emisjami CO



Klaster 1 – Samochody Zwyczajne

```
Statistics for fuel_type
Procent of Petrol cars: 41.177760719377126%
Procent of Diesel cars: 58.82223928062288%
```

	max	min	mean	median	std_dev
year	2013.000000	2000.000000	2008.787532	2010.000000	3.628750
euro_standard	6.000000	2.000000	4.274482	4.000000	0.805338
engine_capacity	2993.000000	599.000000	1667.458713	1598.000000	355.681749
combined_metric	7.700000	3.200000	5.636550	5.700000	0.812298
co2	183.000000	80.000000	141.861662	145.000000	17.814914
co_emissions	1733.000000	2.000000	286.359820	252.000000	181.177814
fuel_cost_6000_miles	572.000000	236.000000	427.025806	427.000000	52.604078
n_engine_capacity	0.361255	0.072299	0.201262	0.192879	0.042931
n_combined_metric	0.303150	0.125984	0.221911	0.224409	0.031980
n_co2	0.310169	0.135593	0.240443	0.245763	0.030195
class	2.000000	2.000000	2.000000	2.000000	0.000000

Klaster 2 – Super Samochody

```
Statistics for fuel_type
Procent of Petrol cars: 98.87842081650965%
Procent of Diesel cars: 1.1215791834903546%
```

	max	min	mean	median	std_dev
year	2013.0	2000.000000	2006.887393	2007.000000	3.662901
euro_standard	6.0	2.000000	3.808434	4.000000	0.816460
engine_capacity	8285.0	2790.000000	5064.292059	4966.000000	872.015620
combined_metric	25.4	9.800000	14.349529	13.900000	2.704700
co2	590.0	230.000000	340.557649	328.000000	62.382161
co_emissions	2000.0	2.000000	514.608476	417.000000	338.928814
fuel_cost_6000_miles	1934.0	706.000000	1062.254019	988.500000	241.288836
n_engine_capacity	1.0	0.336753	0.611260	0.599396	0.105252
n_combined_metric	1.0	0.385827	0.564942	0.547244	0.106484
n_co2	1.0	0.389831	0.577216	0.555932	0.105732
class	1.0	1.000000	1.000000	1.000000	0.000000

Klaster 3 – Samochody Zwyczajne

```
Statistics for fuel_type
Procent of Petrol cars: 70.18827834503925%
Procent of Diesel cars: 29.81172165496075%
```

	max	min	mean	median	std_dev
year	2013.000000	2000.000000	2006.344626	2006.000000	3.676095
euro_standard	6.000000	2.000000	3.776663	4.000000	0.804122
engine_capacity	3498.000000	1199.000000	2026.467214	1991.000000	370.670411
combined_metric	10.400000	5.200000	7.930189	7.900000	0.856268
co2	250.000000	152.000000	194.761967	192.000000	18.065221
co_emissions	1909.000000	3.000000	404.813732	359.000000	275.093633
fuel_cost_6000_miles	860.000000	417.000000	588.699763	584.000000	63.107304
n_engine_capacity	0.422209	0.144719	0.244595	0.240314	0.044740
n_combined_metric	0.409449	0.204724	0.312212	0.311024	0.033711
n_co2	0.423729	0.257627	0.330105	0.325424	0.030619
class	0.000000	0.000000	0.000000	0.000000	0.000000

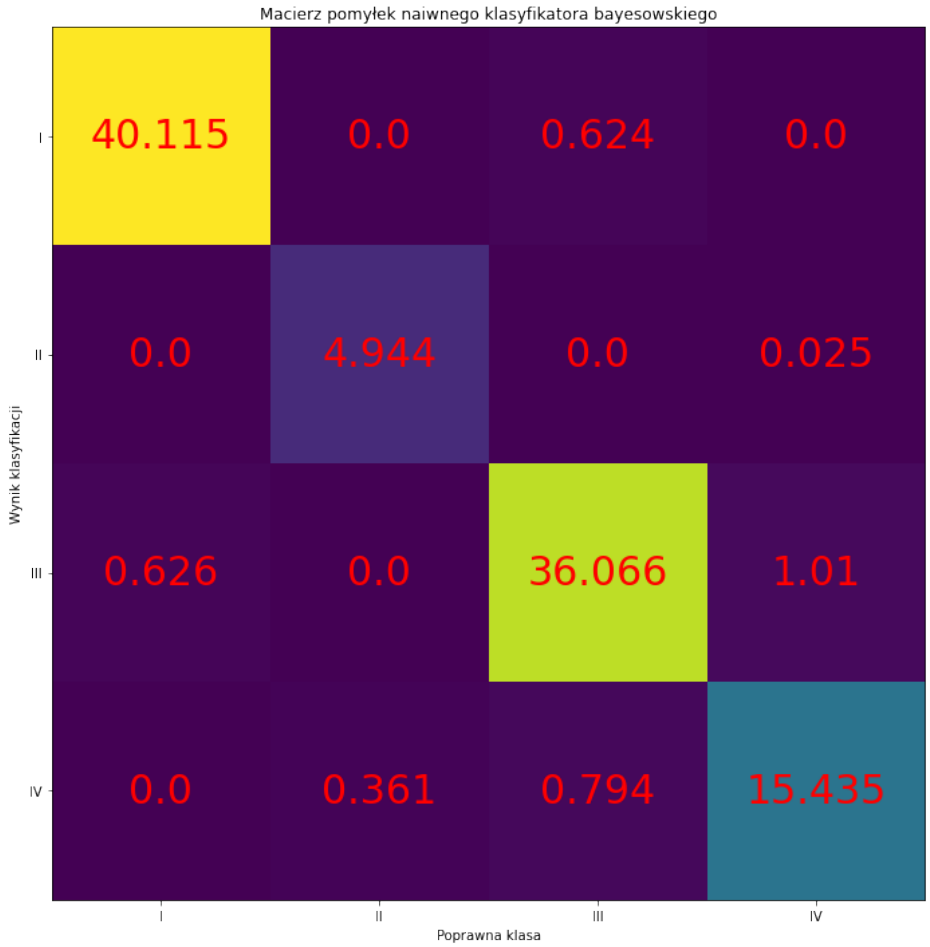
Klaster 4 – Samochody klasy premium

```
Statistics for fuel_type
Procent of Petrol cars: 78.17156665324204%
Procent of Diesel cars: 21.828433346757954%
```

	max	min	mean	median	std_dev
year	2013.000000	2000.000000	2005.993825	2006.000000	3.505284
euro_standard	6.000000	2.000000	3.688415	4.000000	0.768944
engine_capacity	5461.000000	1781.000000	2932.328769	2967.000000	518.338990
combined_metric	15.400000	7.500000	10.281487	10.200000	1.128751
co2	346.000000	141.000000	250.675258	248.000000	25.222856
co_emissions	1985.000000	2.000000	394.270788	314.000000	281.356247
fuel_cost_6000_miles	3329.000000	517.000000	759.086370	750.000000	104.742701
n_engine_capacity	0.659143	0.214967	0.353932	0.358117	0.062564
n_combined_metric	0.606299	0.295276	0.404783	0.401575	0.044439
n_co2	0.586441	0.238983	0.424873	0.420339	0.042751
class	3.000000	3.000000	3.000000	3.000000	0.000000

Klasyfikacja (Klasyfikator Bayesowski)

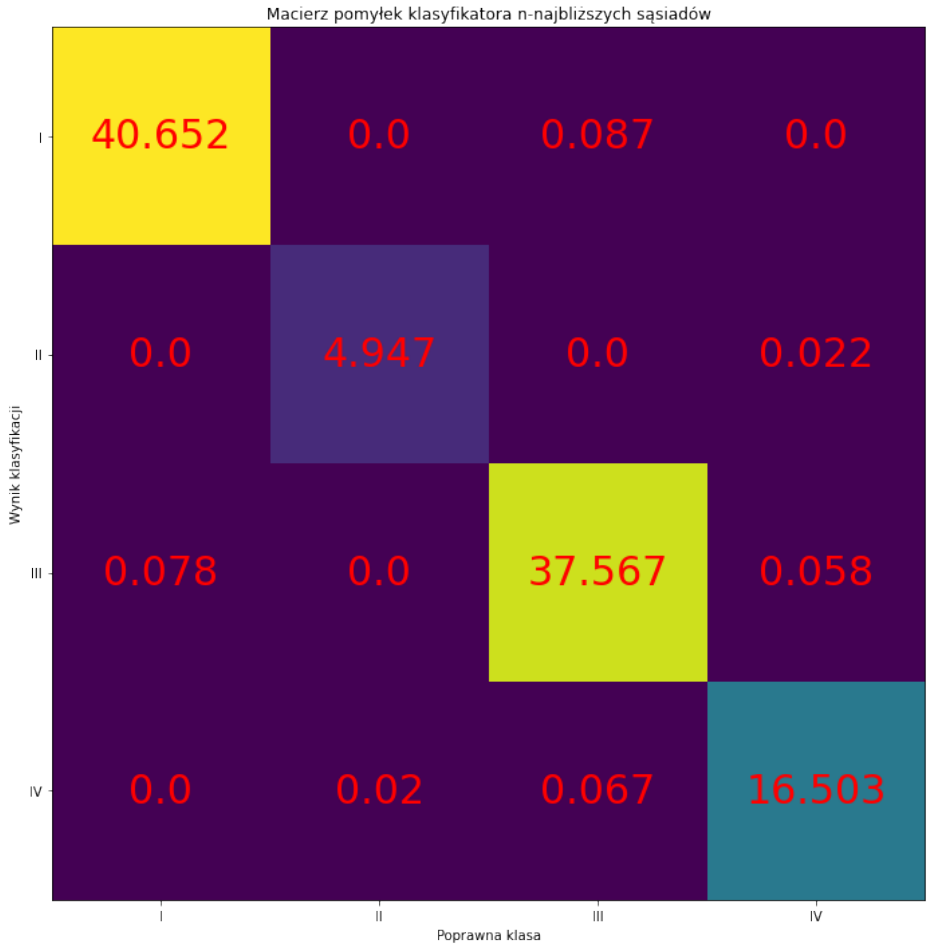
Confusion Matrix



Prawidłowo zklasyfikowane = ~96.56%

Klasyfikacja (Klasyfikator KNN)

Confusion Matrix

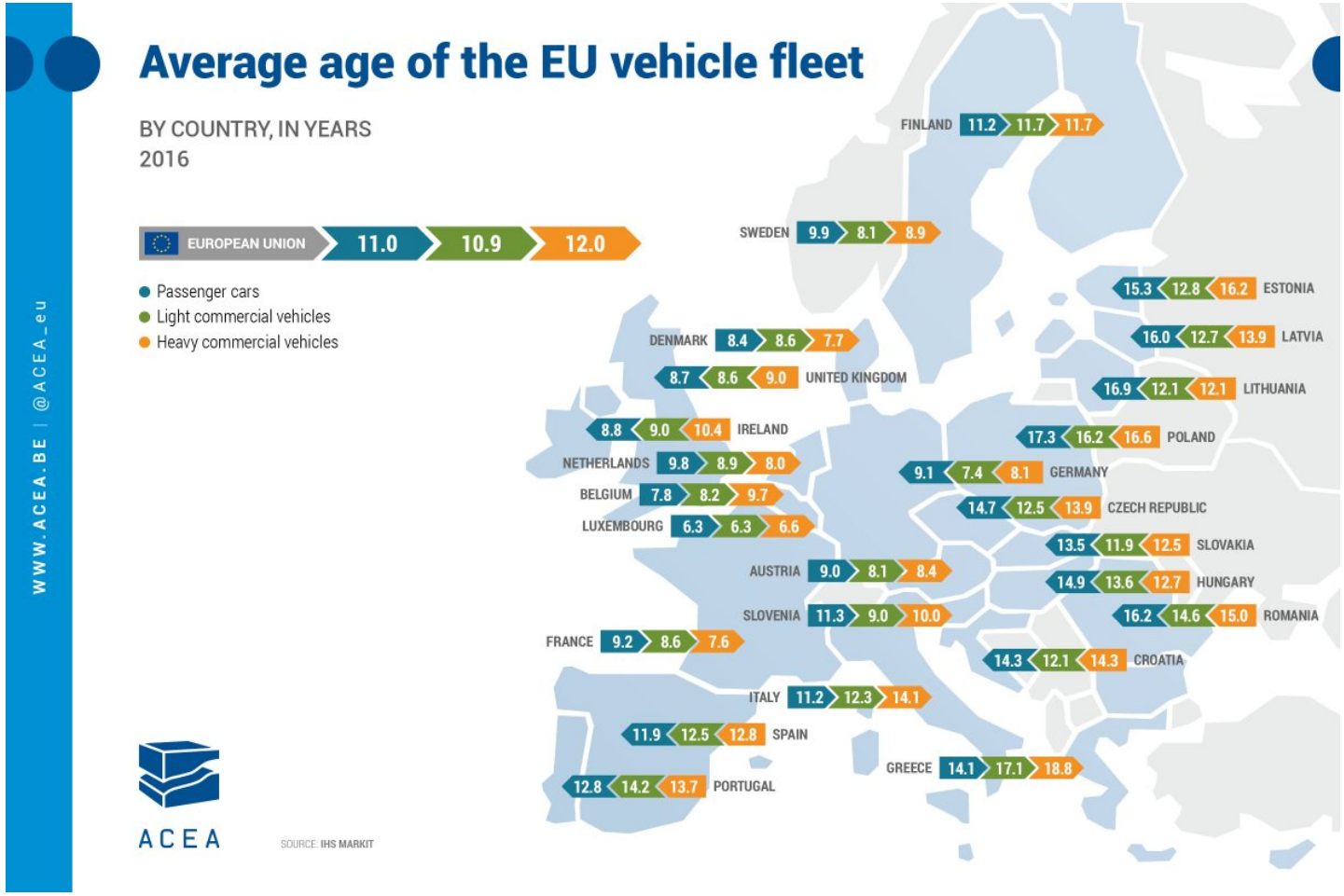


Prawidłowo zklasyfikowane = ~99.68%

Przykład nieprawidłowego doboru klasyfikatora do klasteryzacji. Klasyfikator działa na tym samym aparacie matematycznym co mechanizm klasteryzacji

Dziękujemy za uwagę!

(i życzymy wielu ekonomicznie przejechanych kilometrów)



Źródło: https://twitter.com/ACEA_eu/status/1149201912673255429?s=20