Data analysis of Bilbasen data set

```
clear;
```

Load data set

```
opts = detectImportOptions("bilbasen_scrape.csv");
source = readtable("bilbasen_scrape.csv", opts);
preview("bilbasen_scrape.csv", opts)
```

ans = 8×18 table

	Make	Model	Configuration	Rating	Price	Registered
1	'Ford'	'Kuga'	'2,5 PHEV Vignale CVT 5d'	4.4000	'226.300 kr.'	'7/2020'
2	'Peugeot'	'3008'	'1,6 Hybrid First Selection EAT8 5d'	4.4000	'199.900 kr.'	'2/2021'
3	'VW'	'Passat'	'1,4 GTE Variant DSG 5d'	4.3000	'239.000 kr.'	'5/2020'
4	'Ford'	'Kuga'	'2,5 PHEV Vignale CVT 5d'	NaN	'259.900 kr.'	'11/2021'
5	'VW'	'Golf VIII'	'1,4 eHybrid DSG 5d'	4.6000	'209.900 kr.'	'2/2021'
6	'Skoda'	'Octavia'	'1,4 TSi iV Plus DSG 5d'	4.5000	'213.900 kr.'	'6/2021'
7	'Ford'	'Kuga'	'2,5 PHEV Vignale CVT 5d'	4.4000	'239.800 kr.'	'12/2020'
8	'Peugeot'	'208'	'1,5 BlueHDi 100 Active 5d'	4.2000	'79.900 kr.'	'5/2020'

Some fiddling to learn syntax

```
source(1:3, 1:6)
```

ans = 3×6 table

	Make	Model	Configuration	Rating	Price	Registered
1	'Ford'	'Kuga'	'2,5 PHEV Vignale CVT 5d'	4.4000	'226.300 kr.'	'7/2020'
2	'Peugeot'	'3008'	'1,6 Hybrid First Selection EAT8 5d'	4.4000	'199.900 kr.'	'2/2021'
3	'VW'	'Passat'	'1,4 GTE Variant DSG 5d'	4.3000	'239.000 kr.'	'5/2020'

```
source(1:6, "Model")
```

ans = 6×1 table

	Model
1	'Kuga'
2	'3008'
3	'Passat'
4	'Kuga'
5	'Golf VIII'
6	'Octavia'

```
source.Model(1:6)
```

```
ans = 6×1 cell
'Kuga'
'3008'
'Passat'
'Kuga'
'Golf VIII'
'Octavia'
```

```
source.Properties.VariableNames
```

```
ans = 1×18 cell
'Make'      'Model'      'Configuration''Rating'      'Price'      'Registered''Kilo ···
```

```
source.Properties.VariableTypes
```

```
ans = 1×18 string
"cell"      "cell"      "cell"      "double"    "cell"      "cell"      "cell"   ···
```

Check for empty cells in selected columns

```
% Get column as cell matrix (height) or table (b)
height = source.("Height");
b = source(:, "Height");

% Convert to string and look for non-empty strings
indexes_of_non_empty_cells = find(string(height) ~= "")
```

```
indexes_of_non_empty_cells = 1118×1
     1
     2
     3
     4
     5
     6
     7
     9
    10
    12
     :
     :
```

```
height = height(indexes_of_non_empty_cells)
```

```
height = 1118×1 cell
'166 cm'
'162 cm'
'152 cm'
'166 cm'
'146 cm'
'149 cm'
'166 cm'
'148 cm'
'151 cm'
'166 cm'
     :
     :
```

Remove characters from cells and convert to number

The height column content is formatted like '168 cm'. Removing the last three characters allows us to convert to doubles.

```
height = double(erase(string(height), " cm"))
```

```
height = 1118×1
    166
    162
    152
    166
    146
    149
    166
    148
    151
    166
     :
     :
```

Check for and remove 'NaN' values

```
height = height(find(not(isnan(height))))
```

```
height = 1116×1
    166
    162
    152
    166
    146
    149
    166
    148
    151
    166
     :
     :
```

Column is now cleaned.

```
min(height)
```

```
ans =
139
```

```
max(height)
```

```
ans =
172
```

```
mean(height)
```

```
ans =
154.3217
```

Add a column with age of car

This can be calculated from the registration month of the car. MATLABs datetime type is useful for this metric.

```matlab
registered = string(source.("Registered"));
registered = datetime(registered, "InputFormat","MM/yyyy");
source.Age = calmonths(between(registered, datetime("today"))) / 12;
```

Clean up km column and PriceNew column

```matlab
km = string(source.Kilometers);
km = double(erase(km,[" km", "."]));
source.("Kilometers") = km;

price = string(source.Price);
price = double(erase(price, [" kr.", "."]));
source.Price = price;

pricenew = string(source.PriceNew);
pricenew = double(erase(pricenew, [" kr.", "."]));
source.PriceNew = pricenew;
```

Find every car less than 50.000 kr

Clean price column

```matlab
df = source;
price = df.("Price");
% Convert to string and look for non-empty strings
indexes_of_non_empty_cells = find(string(price) ~= "");
df = df(indexes_of_non_empty_cells, :);

price = erase(string(price), " kr.");
price = double(erase(string(price), "."));
df = df(find(not(isnan(price))),:);
```

Cleaned, ready for logical indexing.

```matlab
lessthan50k = df(price <= 50000, :)
```

lessthan50k = 59×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Fiat'	'Punto'	'0,9 TwinAir 100 Lounge 5d'	3.9000	39000
2	'Citroën'	'Grand C4 Picasso'	'1,6 BlueHDi 120 Intensive 7prs 5d'	4.1000	39900
3	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	44900
4	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
5	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	46900
6	'Opel'	'Astra'	'1,3 CDTi 95 Sport eco 5d'	4	44800

	Make	Model	Configuration	Rating	Price
7	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	24900
8	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	49900
9	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
10	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900
11	'Hyundai'	'i30'	'1,6 CRDi 110 Comfort Eco 5d'	NaN	44900
12	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	22500
13	'Nissan'	'Juke'	'1,5 dCi 110 Tekna 5d'	4.2000	44900
14	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
15	'Hyundai'	'i30'	'1,6 CRDi 110 EM-Edition CW 5d'	NaN	45000
16	'Skoda'	'Octavia'	'1,6 TDi 110 Active Combi 5d'	4.5000	47000
17	'Skoda'	'Rapid'	'1,6 TDi 105 Ambition GreenTec 5d'	4.5000	44900
18	'Seat'	'Leon'	'1,6 TDi Style 5d'	4.4000	44900
19	'Ford'	'Focus'	'1,5 TDCi 120 Trend stc. 5d'	4.3000	44900
20	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
21	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
22	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
23	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar 5d'	3.9000	34900
24	'Citroën'	'Grand C4 Picasso'	'1,6 e-HDi 115 Attraction 5d'	4.1000	39900
25	'Skoda'	'Octavia'	'1,6 TDi 105 Elegance Combi DSG 5d'	4.5000	49900
26	'VW'	'Golf VI'	'1,6 TDi 105 BlueMotion 5d'	4.4000	47900
27	'Volvo'	'V40'	'1,6 D2 115 Drive-E 5d'	NaN	44900
28	'Renault'	'Scenic III'	'1,5 dCi 110 Expression aut. 5d'	4.2000	37500
29	'Peugeot'	'308'	'1,6 e-HDi 114 Style 5d'	4.3000	39400
30	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
31	'Opel'	'Meriva'	'1,6 CDTi 95 Enjoy 5d'	4.1000	44900
32	'Opel'	'Astra'	'1,6 CDTi 136 Enjoy 5d'	4	49700
33	'Skoda'	'Octavia'	'1,6 TDi 105 Ambiente 5d'	4.5000	29400
34	'VW'	'Golf VII'	'1,6 TDi 105 BlueMotion 5d'	4.6000	49999
35	'Opel'	'Corsa'	'1,3 CDTi 95 Sport 5d'	4.1000	42000
36	'Hyundai'	'i30'	'1,6 CDRi 110 Comfort 5d'	NaN	44900
37	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
38	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction 5d'	4.1000	39900
39	'Audi'	'A3'	'1,6 TDi Attraction Sportback 5d'	4.5000	44900

	Make	Model	Configuration	Rating	Price
40	'Opel'	'Astra'	'1,6 CDTi 110 Innovation Sports Tourer 5d'	4	49900
41	'Opel'	'Corsa'	'1,3 CDTi 95 Cosmo 5d'	4.1000	19400
42	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	39800
43	'Nissan'	'Juke'	'1,5 dCi 110 Tekna 5d'	4.2000	49800
44	'Citroën'	'C3'	'1,6 BlueHDi 100 Cool Comfort 5d'	4.1000	29800
45	'Citroën'	'C3'	'1,6 BlueHDi 100 Challenge 5d'	4.1000	49700
46	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
47	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	43900
48	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
49	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	39990
50	'Hyundai'	'i30'	'1,6 CRDi 110 Life+ CW 5d'	4.4000	44900
51	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	49800
52	'Hyundai'	'i30'	'1,6 CRDi 110 Style Eco 5d'	4.4000	26900
53	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	39900
54	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	29900
55	'Hyundai'	'i30'	'1,6 CRDi 110 Style Eco 5d'	4.4000	47900
56	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel Complet 5d'	4.1000	39900
57	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel 5d'	4.1000	49000
58	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	35000
59	'Hyundai'	'i30'	'1,6 CRDi 110 Comfort Eco 5d'	4.4000	39900

Find every car in tax bracket less than 1860 kr/year

Remove characters and convert to numbers

```
tax = lessthan50k.("Tax");
tax = erase(string(tax), " kr. / år");
tax = erase(string(tax), ".");
tax = double(tax);
lessthan50k.Tax = tax
```

lessthan50k = 59×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Fiat'	'Punto'	'0,9 TwinAir 100 Lounge 5d'	3.9000	39000
2	'Citroën'	'Grand C4 Picasso'	'1,6 BlueHDi 120 Intensive 7prs 5d'	4.1000	39900
3	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	44900
4	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900

	Make	Model	Configuration	Rating	Price
5	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	46900
6	'Opel'	'Astra'	'1,3 CDTi 95 Sport eco 5d'	4	44800
7	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	24900
8	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	49900
9	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
10	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900
11	'Hyundai'	'i30'	'1,6 CRDi 110 Comfort Eco 5d'	NaN	44900
12	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	22500
13	'Nissan'	'Juke'	'1,5 dCi 110 Tekna 5d'	4.2000	44900
14	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
15	'Hyundai'	'i30'	'1,6 CRDi 110 EM-Edition CW 5d'	NaN	45000
16	'Skoda'	'Octavia'	'1,6 TDi 110 Active Combi 5d'	4.5000	47000
17	'Skoda'	'Rapid'	'1,6 TDi 105 Ambition GreenTec 5d'	4.5000	44900
18	'Seat'	'Leon'	'1,6 TDi Style 5d'	4.4000	44900
19	'Ford'	'Focus'	'1,5 TDCi 120 Trend stc. 5d'	4.3000	44900
20	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
21	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
22	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
23	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar 5d'	3.9000	34900
24	'Citroën'	'Grand C4 Picasso'	'1,6 e-HDi 115 Attraction 5d'	4.1000	39900
25	'Skoda'	'Octavia'	'1,6 TDi 105 Elegance Combi DSG 5d'	4.5000	49900
26	'VW'	'Golf VI'	'1,6 TDi 105 BlueMotion 5d'	4.4000	47900
27	'Volvo'	'V40'	'1,6 D2 115 Drive-E 5d'	NaN	44900
28	'Renault'	'Scenic III'	'1,5 dCi 110 Expression aut. 5d'	4.2000	37500
29	'Peugeot'	'308'	'1,6 e-HDi 114 Style 5d'	4.3000	39400
30	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
31	'Opel'	'Meriva'	'1,6 CDTi 95 Enjoy 5d'	4.1000	44900
32	'Opel'	'Astra'	'1,6 CDTi 136 Enjoy 5d'	4	49700
33	'Skoda'	'Octavia'	'1,6 TDi 105 Ambiente 5d'	4.5000	29400
34	'VW'	'Golf VII'	'1,6 TDi 105 BlueMotion 5d'	4.6000	49999
35	'Opel'	'Corsa'	'1,3 CDTi 95 Sport 5d'	4.1000	42000
36	'Hyundai'	'i30'	'1,6 CDRi 110 Comfort 5d'	NaN	44900
37	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900

	Make	Model	Configuration	Rating	Price
38	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction 5d'	4.1000	39900
39	'Audi'	'A3'	'1,6 TDi Attraction Sportback 5d'	4.5000	44900
40	'Opel'	'Astra'	'1,6 CDTi 110 Innovation Sports Tourer 5d'	4	49900
41	'Opel'	'Corsa'	'1,3 CDTi 95 Cosmo 5d'	4.1000	19400
42	'Volvo'	'V50'	'1,6 D DRIVe 5d'	4.4000	39800
43	'Nissan'	'Juke'	'1,5 dCi 110 Tekna 5d'	4.2000	49800
44	'Citroën'	'C3'	'1,6 BlueHDi 100 Cool Comfort 5d'	4.1000	29800
45	'Citroën'	'C3'	'1,6 BlueHDi 100 Challenge 5d'	4.1000	49700
46	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
47	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	43900
48	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
49	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	39990
50	'Hyundai'	'i30'	'1,6 CRDi 110 Life+ CW 5d'	4.4000	44900
51	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	49800
52	'Hyundai'	'i30'	'1,6 CRDi 110 Style Eco 5d'	4.4000	26900
53	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	39900
54	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	29900
55	'Hyundai'	'i30'	'1,6 CRDi 110 Style Eco 5d'	4.4000	47900
56	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel Complet 5d'	4.1000	39900
57	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel 5d'	4.1000	49000
58	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	35000
59	'Hyundai'	'i30'	'1,6 CRDi 110 Comfort Eco 5d'	4.4000	39900

```
lessthan50k = lessthan50k(lessthan50k.("Tax") < 1860, :)
```

lessthan50k = 31×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Fiat'	'Punto'	'0,9 TwinAir 100 Lounge 5d'	3.9000	39000
2	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	44900
3	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
4	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	46900
5	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	49900
6	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
7	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900

	Make	Model	Configuration	Rating	Price
8	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
9	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
10	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
11	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
12	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar 5d'	3.9000	34900
13	'Volvo'	'V40'	'1,6 D2 115 Drive-E 5d'	NaN	44900
14	'Renault'	'Scenic III'	'1,5 dCi 110 Expression aut. 5d'	4.2000	37500
15	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
16	'Opel'	'Corsa'	'1,3 CDTi 95 Sport 5d'	4.1000	42000
17	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
18	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction 5d'	4.1000	39900
19	'Opel'	'Astra'	'1,6 CDTi 110 Innovation Sports Tourer 5d'	4	49900
20	'Citroën'	'C3'	'1,6 BlueHDi 100 Cool Comfort 5d'	4.1000	29800
21	'Citroën'	'C3'	'1,6 BlueHDi 100 Challenge 5d'	4.1000	49700
22	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
23	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	43900
24	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
25	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	39990
26	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	49800
27	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	39900
28	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	29900
29	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel Complet 5d'	4.1000	39900
30	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel 5d'	4.1000	49000
31	'Fiat'	'Punto'	'0,9 TwinAir 100 Popstar Edition 5d'	3.9000	35000

Find every car that has a fuel economy of >28km/l

Remove characters and convert to numbers

```
consumption = lessthan50k.("Consumption");
consumption = erase(string(consumption), "(NEDC) ");
consumption = erase(string(consumption), " km/l");
consumption = double(replace(string(consumption), ",", "."))
```

consumption = 31×1
    26.3000
    33.3000
    33.3000
    33.3000

```
26.3000
30.3000
28.6000
31.3000
33.3000
31.3000
    :
    :
```

```
lessthan50k.Consumption = consumption;
lessthan50k = lessthan50k(lessthan50k.("Consumption") > 28, :)
```

lessthan50k = 26×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	44900
2	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
3	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	46900
4	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
5	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900
6	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
7	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
8	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
9	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
10	'Volvo'	'V40'	'1,6 D2 115 Drive-E 5d'	NaN	44900
11	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	49900
12	'Opel'	'Corsa'	'1,3 CDTi 95 Sport 5d'	4.1000	42000
13	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
14	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction 5d'	4.1000	39900
15	'Opel'	'Astra'	'1,6 CDTi 110 Innovation Sports Tourer 5d'	4	49900
16	'Citroën'	'C3'	'1,6 BlueHDi 100 Cool Comfort 5d'	4.1000	29800
17	'Citroën'	'C3'	'1,6 BlueHDi 100 Challenge 5d'	4.1000	49700
18	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
19	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	43900
20	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
21	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	39990
22	'Citroën'	'C3'	'1,6 BlueHDi 100 Attraction 5d'	4.1000	49800
23	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	39900
24	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	29900

	Make	Model	Configuration	Rating	Price
25	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel Complet 5d'	4.1000	39900
26	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel 5d'	4.1000	49000

Sort by PriceNew

The more expensive-as-new priced cars should be more desirable

```
pricenew = lessthan50k.("PriceNew");
pricenew = double(erase(string(pricenew), [" kr.", "."]));
lessthan50k.("PriceNew") = pricenew;
sorted_price_new = sortrows(lessthan50k, "PriceNew", 'descend');
```

Get unique cars

```
[C,ia] = unique(lessthan50k.Configuration);
unique_cars = lessthan50k(ia, :)
```

unique_cars = 14×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Opel'	'Corsa'	'1,3 CDTi 95 Sport 5d'	4.1000	42000
2	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900
3	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
4	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
5	'Citroën'	'C3'	'1,6 BlueHDi 100 Cool Comfort 5d'	4.1000	29800
6	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel 5d'	4.1000	49000
7	'Citroën'	'C3'	'1,6 BlueHDi 100 Feel Complet 5d'	4.1000	39900
8	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction 5d'	4.1000	39900
9	'Citroën'	'C3'	'1,6 BlueHDi 100 Seduction Upgrade 5d'	4.1000	44900
10	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
11	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
12	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
13	'Opel'	'Astra'	'1,6 CDTi 110 Innovation Sports Tourer 5d'	4	49900
14	'Volvo'	'V40'	'1,6 D2 115 Drive-E 5d'	NaN	44900

Manual selection

At this point, I would like to narrow down the scope of models for further analysis. **Opel Corsa** is dropped for its weak engine and small size. The **Citroën C3** is deemed too small as well, and the **Volvo V40** is dropped as it is too long, and predumably more expensive to repair in case of issues.

```
c4indexes = string(source.("Model")) == 'C4';
c4 = source(c4indexes, :)
```

c4 = 24×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel EAT6 5d'	4.2000	89800
2	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel 5d'	4.2000	64500
3	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel 5d'	4.2000	74900
4	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	49700
5	'Citroën'	'C4'	'1,6 BlueHDi 100 Seduction 5d'	NaN	54700
6	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel 5d'	4.2000	94800
7	'Citroën'	'C4'	'1,6 HDi 112 Seduction 5d'	4.2000	57500
8	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	79900
9	'Citroën'	'C4'	'1,6 BlueHDi 120 Challenge EAT6 5d'	4.2000	56500
10	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel Complet EAT6 5d'	4.2000	54999
11	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	89900
12	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel Complet EAT6 5d'	4.2000	59700
13	'Citroën'	'C4'	'1,6 BlueHDi 100 Feel Complet 5d'	4.2000	99990
14	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	66000
15	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel Complet EAT6 5d'	4.2000	84900
16	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel 5d'	NaN	69700
17	'Citroën'	'C4'	'1,6 BlueHDi 100 Feel Complet 5d'	4.2000	89900
18	'Citroën'	'C4'	'1,6 BlueHDi 100 Feel 5d'	4.2000	62000
19	'Citroën'	'C4'	'1,6 BlueHDi 120 Challenge EAT6 5d'	4.2000	99990
20	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	69800
21	'Citroën'	'C4'	'1,6 BlueHDi 120 Feel Complet EAT6 5d'	4.2000	54700
22	'Citroën'	'C4'	'1,6 BlueHDi 100 Attraction 5d'	4.2000	38900
23	'Citroën'	'C4'	'1,6 e-HDi 115 Seduction E6G 5d'	4.2000	59900
24	'Citroën'	'C4'	'1,6 BlueHDi 100 Challenge 5d'	4.2000	39900

```
p308indexes = string(source.("Model")) == '308';
p308 = source(p308indexes, :)
```

p308 = 85×19 table

...

	Make	Model	Configuration	Rating	Price
1	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	69900

	Make	Model	Configuration	Rating	Price
2	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	67900
3	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	84900
4	'Peugeot'	'308'	'1,6 BlueHDi 120 Style Sky SW EAT6 5d'	4.3000	114900
5	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	59800
6	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	99000
7	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	58900
8	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	68800
9	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	53999
10	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	69900
11	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	75600
12	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	64900
13	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	84900
14	'Peugeot'	'308'	'1,6 BlueHDi 120 Style Limited Sky 5d'	4.3000	74995
15	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	69900
16	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure Sky 5d'	4.3000	99900
17	'Peugeot'	'308'	'1,6 BlueHDi 120 GT Line SW EAT6 5d'	4.3000	99800
18	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure Sky 5d'	4.3000	89500
19	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	72800
20	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	69000
21	'Peugeot'	'308'	'1,6 BlueHDi 120 Desire Sky SW 5d'	4.3000	65000
22	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	45000
23	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	70000
24	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	63900
25	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	64900
26	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	89900
27	'Peugeot'	'308'	'1,6 BlueHDi 120 Desire Sky SW 5d'	4.3000	94900
28	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	99800
29	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	69000
30	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	64900
31	'Peugeot'	'308'	'1,2 e-THP 110 Active 5d'	4.3000	99500
32	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	80600
33	'Peugeot'	'308'	'1,6 BlueHDi 120 Chili Allure SW 5d'	4.3000	64900
34	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	89900

	Make	Model	Configuration	Rating	Price
35	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure Sky 5d'	4.3000	99900
36	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	47900
37	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	44900
38	'Peugeot'	'308'	'1,6 BlueHDi 120 Envy Sky SW 5d'	4.3000	64900
39	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure Sky 5d'	4.3000	89900
40	'Peugeot'	'308'	'1,6 BlueHDi 120 Desire Sky 5d'	4.3000	114400
41	'Peugeot'	'308'	'2,0 BlueHDi 150 Active 5d'	4.3000	54900
42	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	67498
43	'Peugeot'	'308'	'1,6 e-HDi 114 Style 5d'	4.3000	39400
44	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	76900
45	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	59900
46	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	54900
47	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	74900
48	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure LTD SW 5d'	4.3000	69900
49	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
50	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure 5d'	4.3000	94700
51	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	62900
52	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	76900
53	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	71900
54	'Peugeot'	'308'	'1,6 BlueHDi 120 Style Sky SW 5d'	4.3000	79500
55	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	94800
56	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW EAT6 5d'	4.3000	114900
57	'Peugeot'	'308'	'1,6 BlueHDi 120 Strike SW 5d'	4.3000	89800
58	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	NaN	54900
59	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	49900
60	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	79500
61	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	43900
62	'Peugeot'	'308'	'1,6 BlueHDi 120 Chili Allure SW 5d'	4.3000	104900
63	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	52900
64	'Peugeot'	'308'	'1,6 BlueHDi 120 Style Limited SW 5d'	4.3000	69700
65	'Peugeot'	'308'	'1,6 BlueHDi 120 Businessline SW 5d'	4.3000	79900
66	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure 5d'	4.3000	86900
67	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW EAT6 5d'	4.3000	99900

	Make	Model	Configuration	Rating	Price
68	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	69900
69	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	66900
70	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	89900
71	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure 5d'	4.3000	99900
72	'Peugeot'	'308'	'1,6 BlueHDi 120 Collection SW 5d'	4.3000	79900
73	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW EAT6 5d'	4.3000	114900
74	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure Sky 5d'	4.3000	89900
75	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	72900
76	'Peugeot'	'308'	'1,6 BlueHDi 120 Style Sky SW 5d'	4.3000	104900
77	'Peugeot'	'308'	'1,6 BlueHDi 120 Active 5d'	4.3000	60000
78	'Peugeot'	'308'	'1,6 BlueHDi 120 Envy Sky SW 5d'	4.3000	74900
79	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	NaN	124900
80	'Peugeot'	'308'	'1,6 BlueHDi 120 Style 5d'	4.3000	59900
81	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	99900
82	'Peugeot'	'308'	'1,6 BlueHDi 120 Allure SW 5d'	4.3000	74900
83	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	69900
84	'Peugeot'	'308'	'1,6 BlueHDi 120 Active SW 5d'	4.3000	124900
85	'Peugeot'	'308'	'1,6 BlueHDi 120 Style SW 5d'	4.3000	104400

The Renault Megane is listed as Megane III. To make this analysis consistent, I want to consider all Megane models.

```
meganeindexes = contains(string(source.("Model")), "Megane");
megane = source(meganeindexes, :)
```

megane = 31×19 table

. . .

	Make	Model	Configuration	Rating	Price
1	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	99900
2	'Renault'	'Megane III'	'1,5 dCi 110 Limited Navi Style Sport Tourer 5d'	4.4000	74900
3	'Renault'	'Megane III'	'1,5 dCi 110 Limited Navi Style Sport Tourer 5d'	4.4000	89900
4	'Renault'	'Megane III'	'1,5 dCi 110 Dynamique Sport Tourer 5d'	4.4000	34900
5	'Renault'	'Megane III'	'1,5 dCi 110 Limited Edition Sport Tourer 5d'	4.4000	57000
6	'Renault'	'Megane IV'	'1,5 dCi 110 Zen EDC 5d'	4.4000	99900
7	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	89500
8	'Renault'	'Megane III'	'1,6 dCi 130 GT-Line Sport Tourer 5d'	4.4000	69400

	Make	Model	Configuration	Rating	Price
9	'Renault'	'Megane IV'	'1,6 dCi 130 Intens Sport Tourer 5d'	4.4000	79990
10	'Renault'	'Megane III'	'1,5 dCi 110 Limited Edition 5d'	4.4000	57700
11	'Renault'	'Megane IV'	'1,5 dCi 110 Zen 5d'	4.4000	79700
12	'Renault'	'Megane IV'	'1,5 dCi 110 Intens Sport Tourer 5d'	4.4000	126900
13	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	89900
14	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	94800
15	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	84900
16	'Renault'	'Megane IV'	'1,5 dCi 110 Zen 5d'	4.4000	99900
17	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer EDC 5d'	4.4000	99900
18	'Renault'	'Megane IV'	'1,5 dCi 110 Zen 5d'	4.4000	69700
19	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	124900
20	'Renault'	'Megane III'	'1,5 dCi 110 Limited Edition Sport Tourer 5d'	4.4000	79700
21	'Renault'	'Megane IV'	'1,5 dCi 110 Zen EDC 5d'	4.4000	99900
22	'Renault'	'Megane III'	'1,5 dCi 110 Expression Sport Tourer 5d'	4.4000	79900
23	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	89900
24	'Renault'	'Megane III'	'1,5 dCi 110 Expression Sport Tourer 5d'	4.4000	74990
25	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	NaN	124900
26	'Renault'	'Megane IV'	'1,6 dCi 130 Bose Edition Sport Tourer 5d'	4.4000	129900
27	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	4.4000	99700
28	'Renault'	'Megane III'	'1,5 dCi 110 Expression Sport Tourer 5d'	4.4000	54900
29	'Renault'	'Megane III'	'1,5 dCi 110 Expression Sport Tourer 5d'	NaN	99900
30	'Renault'	'Megane IV'	'1,5 dCi 110 Zen Sport Tourer 5d'	NaN	139900
31	'Renault'	'Megane III'	'1,6 dCi 130 Bose Edition Sport Tourer 5d'	4.4000	59400

Calculate mean new price values

new price values are missing for some entries. We either scrap those value for modelling, or give them an average value. As the dataset is not that large, I prefer to give them an average value to keep the associated data points for that entry. A better way would be to find an entry with a matching configuration, and use the price_new from there.

```
mean_price_new_c4 = mean(c4.PriceNew, "omitnan")
```

```
mean_price_new_c4 =
2.2218e+05
```

```
i = find(isnan(c4.PriceNew));
c4.PriceNew(i) = mean_price_new_c4;
```

```
mean_price_new_p308 = mean(p308.PriceNew, "omitnan")
```

```
mean_price_new_p308 =
2.7036e+05
```

```
i = find(isnan(p308.PriceNew));
p308.PriceNew(i) = mean_price_new_p308;

mean_price_new_megane = mean(megane.PriceNew, "omitnan")
```

```
mean_price_new_megane =
2.6142e+05
```

```
i = find(isnan(megane.PriceNew));
megane.PriceNew(i) = mean_price_new_megane;
```

Fitted models

```
[c4fitted, c4gof] = fitcar(c4)
```

```
c4fitted =
     General model:
     c4fitted(km,age) = exp(b*km + c*age)
     Coefficients (with 95% confidence bounds):
       b =  -3.773e-06  (-5.022e-06, -2.523e-06)
       c =    -0.06236  (-0.08362, -0.0411)
c4gof = struct with fields:
           sse: 0.0485
       rsquare: 0.7374
           dfe: 22
     adjrsquare: 0.7254
          rmse: 0.0470
```

```
[p308fitted, p308gof] = fitcar(p308)
```

```
p308fitted =
     General model:
     p308fitted(km,age) = exp(b*km + c*age)
     Coefficients (with 95% confidence bounds):
       b =  -3.275e-06  (-3.946e-06, -2.605e-06)
       c =    -0.07268  (-0.08587, -0.05949)
p308gof = struct with fields:
           sse: 0.1582
       rsquare: 0.7148
           dfe: 83
     adjrsquare: 0.7114
          rmse: 0.0437
```

```
[meganefitted, meganegof] = fitcar(megane)
```

```
meganefitted =
     General model:
     meganefitted(km,age) = exp(b*km + c*age)
     Coefficients (with 95% confidence bounds):
       b =  -2.595e-06  (-4.128e-06, -1.062e-06)
       c =    -0.07607  (-0.1025, -0.04965)
```

```
meganegof = struct with fields:
          sse: 0.0982
      rsquare: 0.6583
          dfe: 29
   adjrsquare: 0.6465
         rmse: 0.0582
```
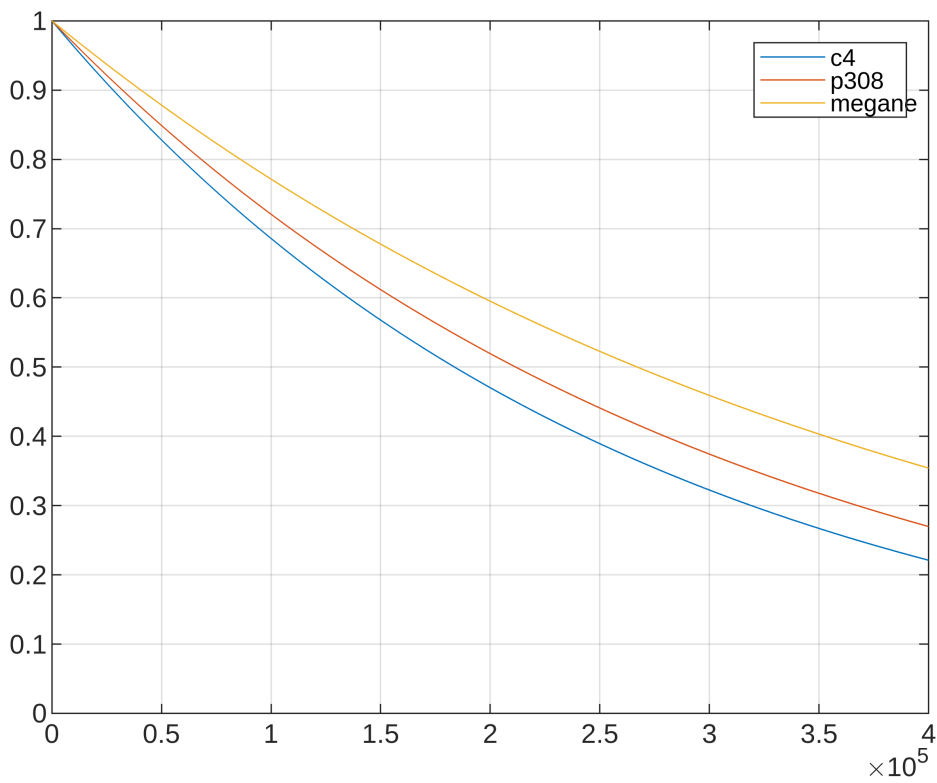
```
x_age = 0:25
```

```
x_age = 1×26
     0    1    2    3    4    5    6    7    8    9   10   11   12 ···
```

```matlab
ys = [exp(c4fitted.c * x_age); exp(p308fitted.c * x_age); exp(meganefitted.c
* x_age)];
p1 = plot(x_age,ys);
legend("c4", "p308", "megane")
ylim([0, 1])
grid on
```
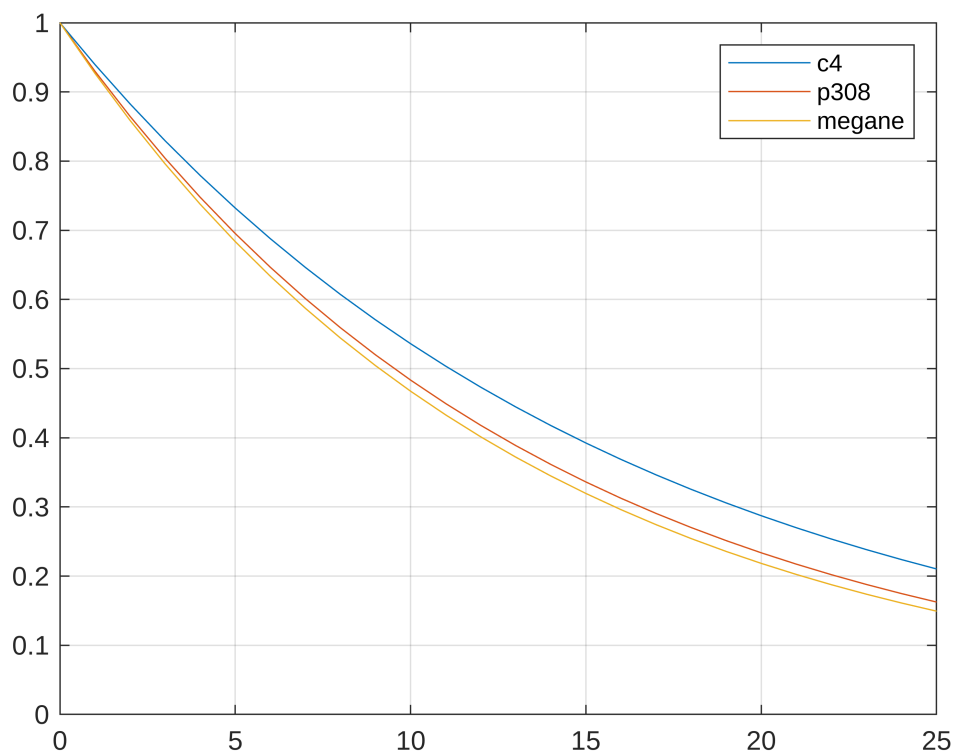


```matlab
x_km = 0:400000;
ys = [exp(c4fitted.b * x_km); exp(p308fitted.b * x_km); exp(meganefitted.b *
x_km)];
p1 = plot(x_km,ys);
legend("c4", "p308", "megane")
ylim([0, 1])
grid on
```

Find best deal

```
%valueofdeal
c4.vod = c4.PriceNew .* exp(c4fitted.b * c4.Kilometers + c4fitted.c *
c4.Age) - c4.Price;
```