```
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9949
regression error: 0.1091
*********************
MODEL 4 SVRegression with all features
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
      'total_link_length_miles', 'pedal_cycles', 'two_wheeled_motor_vehicles',
      'cars_and_taxis', 'buses_and_coaches', 'vans', 'lorries'],
     dtype='object')
score: 0.9998
regression_error: 0.0388
*******************
MODEL SVRegression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9998
regression error: 0.0388
*****************
MODEL Random forest regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score model forest: 0.9941
regression_error: 0.1130
***********************
MODEL Random forest regression with cross validation
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
Traceback (most recent call last):
 File "C:\Users\Utente\Desktop\DataScience\code\main.py", line 57, in <module>
   model_forest_cross(frame)
 File "C:/Users/Utente/Desktop/DataScience/code\model.py", line 84, in
model_forest_cross
   y_pred=model_forest.predict(X)
 File "C:\Users\Utente\Anaconda3\lib\site-packages\sklearn\ensemble\_forest.py", line
764, in predict
   check_is_fitted(self)
 File "C:\Users\Utente\Anaconda3\lib\site-packages\sklearn\utils\validation.py", line
967, in check_is_fitted
   raise NotFittedError(msg % {'name': type(estimator).__name__})
NotFittedError: This RandomForestRegressor instance is not fitted yet. Call 'fit' with
appropriate arguments before using this estimator.
<Figure size 432x288 with 0 Axes>
In [50]: runfile('C:/Users/Utente/Desktop/DataScience/code/main.py', wdir='C:/Users/
Utente/Desktop/DataScience/code')
Reloaded modules: init, model, util, visualization, preprocessing
************************
UNIQUE DISTRIBUTION OF FEATURES
Total number of regions: 11
Total number of road categories: 6
```

	year	region_id	 lorries	all_motor_vehicles
count	1547.000000	1547.000000	 1547.000000	1547.000000
mean	2005.475760	6.212023	 0.283026	5.028188
std	7.517803	3.187790	 0.292345	3.616128
min	1993.000000	1.000000	 0.000648	0.010438
25%	1999.000000	3.000000	 0.067333	2.238600
50%	2005.000000	6.000000	 0.168881	4.270186
75%	2012.000000	9.000000	 0.414246	7.396709
max	2018.000000	11.000000	 1.386260	15.845616

[8 rows x 12 columns]

GROUP DISTRIBUTION ON TRAFFIC BY YEAR							
	year	region_id	road_category_id		vans	lorries	all_motor_vehicles
0	1993	373	219		11.087950	10.849047	16.140504
1	1994	373	219		11.553254	11.082147	16.501645
2	1995	373	219		11.867358	11.382132	16.822704
3	1996	373	219		12.317925	11.736691	17.270079
4	1997	373	219		12.949367	12.014887	17.631146
5	1998	373	219		13.548895	12.411285	17.951606
6	1999	373	219		13.760718	12.586221	18.283915
7	2000	373	219		13.925848	12.627528	18.252833
8	2001	367	216		14.257194	12.544484	18.507065
9	2002	367	216		14.585115	12.653575	18.941853
10	2003	367	216		15.315447	12.717592	19.057797
11	2004	367	216		16.064099	13.102573	19.343582
12	2005	368	218		16.498701	12.951695	19.339939
13	2006	368	218		17.158052	12.996429	19.622870
14	2007	367	216		17.981100	13.093525	19.794497
15	2008	367	216		17.865668	12.799325	19.603296
16	2009	367	216		17.480556	11.743582	19.417265
17	2010	367	216		17.630435	11.783215	19.107573
18	2011	367	216		17.775096	11.466826	19.144793
19	2012	367	216		17.732849	11.185769	19.072997
20	2013	367	216		18.292121	11.284625	19.140180
21	2014	367	216		19.337688	11.576409	19.640368
22	2015	367	216		20.154004	12.003396	19.962072
23	2016	373	219		21.142740	12.115465	20.354558
24	2017	373	219		21.704466	12.258545	20.617241
25	2018	373	219	• • •	21.908306	12.300529	20.681070

[26 rows x 12 columns]

TRAFFIC BY REGION IN 2018

	name	year	 all_motor_vehicles	other_vehicles
0	North East	12108	 12.274201	0.221726
1	Wales	10090	 18.260918	0.309984
2	London	10090	 18.354733	0.744356
3	East Midlands	10090	 27.647983	0.365226
4	Yorkshire and The Humber	12108	 27.760255	0.374354
5	Scotland	10090	 29.716436	0.507561
6	West Midlands	12108	 31.554014	0.372814
7	South West	10090	 33.238699	0.513533
8	North West	12108	 35.697093	0.399917
9	East of England	10090	 38.701718	0.512950
10	South East	12108	 54.908644	0.713793

[11 rows x 14 columns]

```
VEHICLE TYPE COMPARISON
  year region_id ... all_motor_vehicles other_vehicles
            373 ...
   1993
                          256.214013
                                        5.211234
25 2018
            373 ...
                          328.114695
                                         5.036214
[2 rows x 13 columns]
+----+
| Vehicle type | Billion vehicle miles |
| all_motor_vehicles
                                   328.115
cars_and_taxis
                                   255.013
vans
                                    50.9832
lorries
                                    17.0826
| pedal cycles
                                    3.32911
                                    2.73903
| two wheeled motor vehicles |
                                    2.29719 İ
buses and coaches
+----+
******************
MODEL Decision tree regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9954
regression_error: 0.1057
**********************
MODEL 4 SVRegression with all features
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
     'total_link_length_miles', 'pedal_cycles', 'two_wheeled_motor_vehicles',
     'cars_and_taxis', 'buses_and_coaches', 'vans', 'lorries'],
    dtype='object')
score: 0.9998
regression_error: 0.0349
*********************
MODEL SVRegression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9998
regression_error: 0.0349
**********************
MODEL Random forest regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score model forest: 0.9976
regression error: 0.1046
******************
MODEL Random forest regression with cross validation
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
cross validation scores [0.98780831 0.99852015 0.9913102 0.98870579 0.99885281
0.99862432
0.99954502 0.99839595 0.99866884 0.98766006]
regression error: 0.0288
************************
MODEL Linear regression 1
```

```
features= Index(['pedal_cycles', 'two_wheeled_motor_vehicles', 'cars_and_taxis',
      'buses_and_coaches', 'vans'],
     dtype='object')
score linear: 0.9983
regression error: 0.1121
******************
MODEL Linear regression 2
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
      'total_link_length_miles'],
     dtype='object')
score linear2: 0.2173
regression error: 2.5691
<Figure size 432x288 with 0 Axes>
In [51]: runfile('C:/Users/Utente/Desktop/DataScience/code/main.py', wdir='C:/Users/
Utente/Desktop/DataScience/code')
Reloaded modules: init, model, util, visualization, preprocessing
******************
UNIQUE DISTRIBUTION OF FEATURES
Total number of regions: 11
Total number of road categories: 6
            year
                    region_id ...
                                       lorries all_motor_vehicles
count 1547.000000
                  1547.000000 ... 1547.000000
                                                      1547.000000
                     6.212023 ...
      2005.475760
                                      0.283026
                                                         5.028188
mean
                     3.187790 ...
         7.517803
                                      0.292345
                                                         3.616128
std
                     1.000000 ...
      1993.000000
                                      0.000648
                                                         0.010438
min
25%
      1999.000000
                     3.000000 ... 0.067333
                                                         2,238600
50%
      2005.000000
                     6.000000 ... 0.168881
                                                         4.270186

      9.000000
      ...
      0.414246

      11.000000
      ...
      1.386260

75%
      2012.000000
                                                         7.396709
      2018.000000
                                                        15.845616
max
[8 rows x 12 columns]
**********************
GROUP DISTRIBUTION ON TRAFFIC BY YEAR
   year region_id road_category_id ...
                                                     lorries all_motor_vehicles
                                              vans
                                219 ... 11.087950 10.849047
                                                                      16.140504
   1993
               373
1
   1994
               373
                                219 ... 11.553254 11.082147
                                                                      16.501645
   1995
              373
                                219 ... 11.867358 11.382132
                                                                      16.822704
3
   1996
              373
                               219 ... 12.317925 11.736691
                                                                      17.270079
4
   1997
              373
                               219 ... 12.949367 12.014887
                                                                      17.631146
   1998
              373
                               219 ... 13.548895 12.411285
                                                                      17.951606
6
   1999
              373
                               219 ... 13.760718 12.586221
                                                                      18.283915
7
   2000
              373
                               219 ... 13.925848 12.627528
                                                                      18.252833
8
   2001
              367
                               216 ... 14.257194 12.544484
                                                                      18.507065
9
   2002
              367
                               216 ... 14.585115 12.653575
                                                                      18.941853
10 2003
              367
                               216 ... 15.315447 12.717592
                                                                      19.057797
11 2004
              367
                               216 ... 16.064099 13.102573
                                                                      19.343582
12 2005
              368
                               218 ... 16.498701 12.951695
                                                                      19.339939
13 2006
              368
                               218 ... 17.158052 12.996429
                                                                      19.622870
14 2007
                               216 ... 17.981100 13.093525
                                                                      19.794497
              367
15 2008
                                216 ... 17.865668 12.799325
              367
                                                                      19.603296
16 2009
                                216 ... 17.480556 11.743582
              367
                                                                      19.417265
17 2010
                                216 ... 17.630435 11.783215
              367
                                                                      19.107573
18 2011
                                216 ... 17.775096 11.466826
              367
                                                                      19.144793
19 2012
              367
                                216 ... 17.732849 11.185769
                                                                      19.072997
                                                                      19.140180
20 2013
              367
                               216 ... 18.292121 11.284625
                               216 ... 19.337688 11.576409
21 2014
               367
                                                                      19.640368
```

23 24	2015 2016 2017 2018	367 373 373 373	219 219	•••	20.154004 21.142740 21.704466 21.908306	12.115465 12.258545		19.962072 20.354558 20.617241 20.681070
[26 rows x 12 columns]								

TRAFFIC BY REGION IN 2018

	name	year	 all_motor_vehicles	other_vehicles
0	North East	12108	 12.274201	0.221726
1	Wales	10090	 18.260918	0.309984
2	London	10090	 18.354733	0.744356
3	East Midlands	10090	 27.647983	0.365226
4	Yorkshire and The Humber	12108	 27.760255	0.374354
5	Scotland	10090	 29.716436	0.507561
6	West Midlands	12108	 31.554014	0.372814
7	South West	10090	 33.238699	0.513533
8	North West	12108	 35.697093	0.399917
9	East of England	10090	 38.701718	0.512950
10	South East	12108	 54.908644	0.713793

[11 rows x 14 columns]

VEHICLE TYPE COMPARISON

	year	region_id	 all_motor_vehicles	other_vehicles
0	1993	373	 256.214013	5.211234
25	2018	373	 328.114695	5.036214

[2 rows x 13 columns]

Vehicle type	Billion vehicle miles
all_motor_vehicles cars_and_taxis vans lorries pedal_cycles two_wheeled_motor_vehicles buses_and_coaches	328.115 255.013 50.9832 17.0826 3.32911 2.73903 2.29719

```
MODEL Decision tree regression
```

features= Index(['year', 'region_id', 'road_category_id'], dtype='object')

score: 0.9976

regression_error: 0.1087

```
MODEL 4 SVRegression with all features
```

score: 0.9998

regression_error: 0.0365

```
**********************
MODEL SVRegression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9998
regression error: 0.0365
*******************
MODEL Random forest regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score model forest: 0.9970
regression_error: 0.1083
******************
MODEL Random forest regression with cross validation
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
cross validation scores [0.9869853 0.99816707 0.99102495 0.98868196 0.99882561
0.99862245
0.99958982 0.99839038 0.998564 0.98951221]
Average cross validation score: %.4f
regression_error: 0.0285
***********************
MODEL Linear regression 1
features= Index(['pedal_cycles', 'two_wheeled_motor_vehicles', 'cars_and_taxis',
      'buses_and_coaches', 'vans'],
     dtype='object')
score linear: 0.9976
regression_error: 0.1271
**********************
MODEL Linear regression 2
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
      'total_link_length_miles'],
     dtype='object')
score linear2: 0.1629
regression_error: 2.5806
<Figure size 432x288 with 0 Axes>
In [52]: runfile('C:/Users/Utente/Desktop/DataScience/code/main.py', wdir='C:/Users/
Utente/Desktop/DataScience/code')
Reloaded modules: init, model, util, visualization, preprocessing
******************
UNIQUE DISTRIBUTION OF FEATURES
Total number of regions: 11
Total number of road categories: 6
                   region_id ...
                                    lorries all motor vehicles
            year
count 1547.000000 1547.000000 ... 1547.000000
                                                   1547.000000
     2005.475760
                   6.212023 ...
                                   0.283026
                                                     5.028188
mean
        7.517803
                    3.187790 ...
                                   0.292345
                                                     3,616128
std
      1993.000000
                    1.000000 ...
                                   0.000648
                                                     0.010438
min
     1999.000000
                    3.000000 ...
                                   0.067333
25%
                                                     2.238600
      2005.000000
                   6.000000 ...
                                   0.168881
                                                     4.270186
50%
75%
      2012.000000
                   9.000000 ...
                                  0.414246
                                                     7.396709
      2018.000000
                   11.000000 ...
                                   1.386260
                                                    15.845616
max
```

[8 rows x 12 columns]

GROUP DISTRIBUTION ON TRAFFIC BY VEAR						
year				vans	lorries	all_motor_vehicles
1993	373	219		11.087950	10.849047	 16.140504
1994	373	219		11.553254	11.082147	16.501645
1995	373	219		11.867358	11.382132	16.822704
1996	373	219		12.317925	11.736691	17.270079
1997	373	219		12.949367	12.014887	17.631146
1998	373	219		13.548895	12.411285	17.951606
1999	373	219		13.760718	12.586221	18.283915
2000	373	219		13.925848	12.627528	18.252833
2001	367	216		14.257194	12.544484	18.507065
2002	367	216		14.585115	12.653575	18.941853
2003	367	216		15.315447	12.717592	19.057797
2004	367	216		16.064099	13.102573	19.343582
2005	368	218		16.498701	12.951695	19.339939
2006	368	218		17.158052	12.996429	19.622870
2007	367	216		17.981100	13.093525	19.794497
2008	367	216		17.865668	12.799325	19.603296
2009	367	216		17.480556	11.743582	19.417265
2010	367	216		17.630435	11.783215	19.107573
2011	367	216		17.775096	11.466826	19.144793
		216	• • •	17.732849	11.185769	19.072997
2013	367	216	• • •	18.292121	11.284625	19.140180
2014	367	216		19.337688	11.576409	19.640368
			• • •			19.962072
2016	373		• • •	21.142740	12.115465	20.354558
2017	373	219		21.704466	12.258545	20.617241
2018	373	219	• • •	21.908306	12.300529	20.681070
	year 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012 2013 2014 2015 2016 2017	year region_id 1993 373 1994 373 1995 373 1996 373 1997 373 1998 373 1999 373 2000 373 2001 367 2002 367 2003 367 2004 367 2005 368 2006 368 2007 367 2008 367 2009 367 2010 367 2011 367 2011 367 2011 367 2012 367 2013 367 2014 367 2015 367 2014 367 2015 367 2016 373 2017 373	1993 373 219 1994 373 219 1995 373 219 1996 373 219 1997 373 219 1998 373 219 2000 373 219 2001 367 216 2002 367 216 2003 367 216 2004 367 216 2005 368 218 2007 367 216 2008 367 216 2009 367 216 2010 367 216 2011 367 216 2012 367 216 2013 367 216 2014 367 216 2015 367 216 2016 373 219 2017 373 219	year region_id road_category_id 1993 373 219 1994 373 219 1995 373 219 1996 373 219 1997 373 219 1998 373 219 2000 373 219 2001 367 216 2002 367 216 2003 367 216 2004 367 216 2005 368 218 2006 368 218 2007 367 216 2008 367 216 2010 367 216 2011 367 216 2012 367 216 2013 367 <td>year region_id road_category_id vans 1993 373 219 11.087950 1994 373 219 11.553254 1995 373 219 12.317925 1996 373 219 12.949367 1998 373 219 13.548895 1999 373 219 13.760718 2000 373 219 13.760718 2001 367 216 14.257194 2002 367 216 14.585115 2003 367 216 15.315447 2004 367 216 15.315447 2004 367 216 17.158052 2007 368 218 17.158052 2007 367 216 17.480556 2010 367 2</td> <td>year region_id road_category_id vans lorries 1993 373 219 11.087950 10.849047 1994 373 219 11.553254 11.082147 1995 373 219 11.867358 11.382132 1996 373 219 12.317925 11.736691 1997 373 219 12.949367 12.014887 1998 373 219 13.548895 12.411285 1999 373 219 13.760718 12.586221 2000 373 219 13.760718 12.586221 2000 373 219 13.760718 12.586221 2000 373 219 13.7560718 12.586221 2001 367 216 14.257194 12.544484 2002 367 216 15.315447 12</td>	year region_id road_category_id vans 1993 373 219 11.087950 1994 373 219 11.553254 1995 373 219 12.317925 1996 373 219 12.949367 1998 373 219 13.548895 1999 373 219 13.760718 2000 373 219 13.760718 2001 367 216 14.257194 2002 367 216 14.585115 2003 367 216 15.315447 2004 367 216 15.315447 2004 367 216 17.158052 2007 368 218 17.158052 2007 367 216 17.480556 2010 367 2	year region_id road_category_id vans lorries 1993 373 219 11.087950 10.849047 1994 373 219 11.553254 11.082147 1995 373 219 11.867358 11.382132 1996 373 219 12.317925 11.736691 1997 373 219 12.949367 12.014887 1998 373 219 13.548895 12.411285 1999 373 219 13.760718 12.586221 2000 373 219 13.760718 12.586221 2000 373 219 13.760718 12.586221 2000 373 219 13.7560718 12.586221 2001 367 216 14.257194 12.544484 2002 367 216 15.315447 12

[26 rows x 12 columns]

```
TRAFFIC BY REGION IN 2018
```

	0 _ 2			
	name	year	 all_motor_vehicles	other_vehicles
0	North East	12108	 12.274201	0.221726
1	Wales	10090	 18.260918	0.309984
2	London	10090	 18.354733	0.744356
3	East Midlands	10090	 27.647983	0.365226
4	Yorkshire and The Humber	12108	 27.760255	0.374354
5	Scotland	10090	 29.716436	0.507561
6	West Midlands	12108	 31.554014	0.372814
7	South West	10090	 33.238699	0.513533
8	North West	12108	 35.697093	0.399917
9	East of England	10090	 38.701718	0.512950
10	South East	12108	 54.908644	0.713793

[11 rows x 14 columns]

VEHICLE TYPE COMPARISON

	year	region_id	 all_motor_vehicles	other_vehicles
0	1993	373	 256.214013	5.211234
25	2018	373	 328.114695	5.036214

[2 rows x 13 columns]

+	-++
Vehicle type	Billion vehicle miles
	-+
all_motor_vehicles	328.115

```
cars and taxis
                                       255.013
                                       50.9832
 vans
                                       17.0826
lorries
| pedal cycles
                                        3.32911
two_wheeled_motor_vehicles
                                        2.73903
buses and coaches
                                        2.29719
*******************
MODEL Decision tree regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9956
regression error: 0.1183
*****************
MODEL 4 SVRegression with all features
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
      'total_link_length_miles', 'pedal_cycles', 'two_wheeled_motor_vehicles',
      'cars_and_taxis', 'buses_and_coaches', 'vans', 'lorries'],
     dtype='object')
score: 0.9998
regression error: 0.0364
******************
MODEL SVRegression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score: 0.9998
regression_error: 0.0364
*******************
MODEL Random forest regression
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
score model_forest: 0.9982
regression_error: 0.0922
**********************
MODEL Random forest regression with cross validation
features= Index(['year', 'region_id', 'road_category_id'], dtype='object')
cross validation scores [0.98695588 0.99812871 0.99246222 0.98883811 0.99883923
0.99860616
0.99958042 0.99833726 0.99873595 0.98792697]
Average cross validation score: 0.9948
regression_error: 0.0285
*******************
MODEL Linear regression 1
features= Index(['pedal_cycles', 'two_wheeled_motor_vehicles', 'cars_and_taxis',
      'buses_and_coaches', 'vans'],
     dtype='object')
score linear: 0.9975
regression_error: 0.1265
************************
MODEL Linear regression 2
features= Index(['year', 'region_id', 'road_category_id', 'total_link_length_km',
      'total_link_length_miles'],
```

dtype='object')
score linear2: 0.1920
regression_error: 2.4400
<Figure size 432x288 with 0 Axes>

In [53]: