Exercice 1

1. The dimension of the data (num. of rows, num. of columns) is (***67856***,***11***).

2. The first six lines of this data are

X	veh_value	exposure	clm	numclaims	claimcst0	veh_body	veh_age	gender	area	agecat
1	1.06	0.3039014	0	0	0	HBACK	3	F	C	2
2	1.03	0.6488706	0	0	0	HBACK	2	F	A	4
3	3.26	0.5694730	0	0	0	UTE	2	F	E	2
4	4.14	0.3175907	0	0	0	STNWG	2	F	D	2
5	0.72	0.6488706	0	0	0	HBACK	4	F	C	2
6	2.01	0.8542094	0	0	0	HDTOP	3	M	C	4

3. Using the R sunction str(), we get the structure of the data

```
'data.frame':   67856 obs. of  11 variables:
 $ X        : int  1 2 3 4 5 6 7 8 9 10 ...
 $ veh_value: num  1.06 1.03 3.26 4.14 0.72 2.01 1.6 1.47 0.52 0.38 ...
 $ exposure : num  0.304 0.649 0.569 0.318 0.649 ...
 $ clm      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ numclaims: int  0 0 0 0 0 0 0 0 0 0 ...
 $ claimcst0: num  0 0 0 0 0 0 0 0 0 0 ...
 $ veh_body : chr  "HBACK" "HBACK" "UTE" "STNWG" ...
 $ veh_age  : int  3 2 2 2 4 3 3 2 4 4 ...
 $ gender   : chr  "F" "F" "F" "F" ...
 $ area     : chr  "C" "A" "E" "D" ...
 $ agecat   : int  2 4 2 2 2 4 4 6 3 4 ...
```

4. We made use of the function `subset()` to delete the first column of `dataCar` and the function `transform()` to transform the variables clm, numclaims, claimcst0, veh_body, veh_age, gender, area, agecat to a factor. The summary of the resulting data is

```
  veh_value          exposure            clm             numclaims
 Min.   : 0.000   Min.   :0.002738   Min.   :0.00000   Min.   :0.00000
 1st Qu.: 1.010   1st Qu.:0.219028   1st Qu.:0.00000   1st Qu.:0.00000
 Median : 1.500   Median :0.446270   Median :0.00000   Median :0.00000
 Mean   : 1.777   Mean   :0.468651   Mean   :0.06814   Mean   :0.07276
 3rd Qu.: 2.150   3rd Qu.:0.709103   3rd Qu.:0.00000   3rd Qu.:0.00000
 Max.   :34.560   Max.   :0.999316   Max.   :1.00000   Max.   :4.00000
   claimcst0          veh_body           veh_age           gender
 Min.   :   0.0   Length:67856       Min.   :1.000    Length:67856
 1st Qu.:   0.0   Class :character   1st Qu.:2.000    Class :character
 Median :   0.0   Mode  :character   Median :3.000    Mode  :character
 Mean   : 137.3                      Mean   :2.674
 3rd Qu.:   0.0                      3rd Qu.:4.000
```

```
Max.   :55922.1                    Max.   :4.000
      area                 agecat
Length:67856      Min.   :1.000
Class :character  1st Qu.:2.000
Mode  :character  Median :3.000
                  Mean   :3.485
                  3rd Qu.:5.000
                  Max.   :6.000
```