Create Pandas datasets from raw Scania Truck Dataset

Get the data from [https://archive.ics.uci.edu/ml/machine-learning-databases/00421/ (https://archive.ics.uci.edu/ml/machine-learning-databases/00421/)](https://archive.ics.uci.edu/ml/machine-learning-databases/00421/).

Convert datasets to Pandas dataframe and create new CSV files

In [ ]:
```
import pandas as pd
import numpy as np
```

In [220]:
```
data = pd.read_csv("aps_failure_training_set.csv", sep = "delimiter", header = None)
```

```
C:\Users\GauravP\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: ParserWarning: Falling back to the 'python' engin
e because the 'c' engine does not support regex separators (separators > 1 char and different from '\s+' are interprete
d as regex); you can avoid this warning by specifying engine='python'.
  """Entry point for launching an IPython kernel.
```

In [222]:
```
data
```

Out[222]:

	0
0	This file is part of APS Failure and Operation...
1	Copyright (c) <2016> <Scania CV AB>
2	This program (APS Failure and Operational Data...
3	free software: you can redistribute it and/or ...
4	it under the terms of the GNU General Public L...
5	the Free Software Foundation, either version 3...
6	(at your option) any later version.
7	This program is distributed in the hope that i...
8	but WITHOUT ANY WARRANTY; without even the imp...
9	MERCHANTABILITY or FITNESS FOR A PARTICULAR PU...
10	GNU General Public License for more details.
11	You should have received a copy of the GNU Gen...
12	along with this program. If not, see <http://...
13	----------...
14	class,aa_000,ab_000,ac_000,ad_000,ae_000,af_00...
15	neg,76698,na,2130706438,280,0,0,0,0,0,37250,...
16	neg,33058,na,0,na,0,0,0,0,0,0,18254,653294,172...
17	neg,41040,na,228,100,0,0,0,0,0,0,1648,370592,1...
18	neg,12,0,70,66,0,10,0,0,0,318,2212,3232,1872,0...
19	neg,60874,na,1368,458,0,0,0,0,0,0,43752,196661...
20	neg,38312,na,2130706432,218,0,0,0,0,0,0,9128,7...
21	neg,14,0,6,na,0,0,0,0,0,0,1202,3766,1150,0,0,0...
22	neg,102960,na,2130706432,116,0,0,0,0,0,0,2130,...
23	neg,78696,na,0,na,0,0,0,0,0,0,458,440704,43988...

	0
24	pos,153204,0,182,na,0,0,0,0,0,11804,684444,326...
25	neg,39196,na,204,170,0,0,0,0,0,0,4352,713882,1...
26	neg,45912,na,0,454,0,0,0,0,0,0,2106,418940,221...
27	neg,2104,na,36,26,0,0,0,0,9744,13148,98310,222...
28	neg,118950,na,1390,1298,0,0,0,0,0,0,40932,1657...
29	neg,24416,na,0,na,0,0,0,0,0,0,556,448484,88779...
...	...
59985	neg,16,0,8,na,0,0,0,0,0,0,108,5176,2234,0,0,0,...
59986	neg,39026,na,202,168,0,0,0,0,0,0,2150,1019858,...
59987	neg,14,na,24,20,0,0,0,0,0,0,44,396,5598,0,0,0,...
59988	neg,3248,8,16,10,0,0,0,0,0,9046,123142,100580,...
59989	neg,83818,na,552,532,0,0,0,0,0,0,3088,213034,3...
59990	neg,40274,na,98,94,0,0,0,0,0,0,338,373034,1810...
59991	neg,16978,na,2130706434,1750,0,0,0,0,0,0,582,6...
59992	neg,30320,na,1838,1278,0,0,0,0,0,0,22386,76403...
59993	neg,38414,na,888,758,0,0,0,0,0,0,41546,1732152...
59994	neg,18,0,2130706432,18,0,0,0,0,0,0,3656,1728,2...
59995	neg,562,0,4,4,0,0,0,0,0,0,11538,34194,2252,0,0...
59996	neg,16,na,20,20,0,0,0,0,0,0,90,1304,7306,696,0...
59997	neg,10628,na,2130706434,98,0,0,0,0,0,4868,1167...
59998	neg,39004,na,90,84,0,0,0,0,0,9360,208852,13644...
59999	neg,33386,na,812,744,0,0,0,0,0,0,11470,650532,...
60000	neg,10792,na,784,608,0,0,0,0,0,0,31894,661618,...
60001	neg,644,na,12,na,0,0,0,0,0,18118,43536,41808,2...
60002	neg,41330,na,2130706432,744,0,0,0,0,0,0,32154,...
60003	neg,6078,na,52,46,0,0,0,0,0,0,76,22588,173848,...

	0
60004	neg,61478,na,134,0,0,0,0,0,0,0,2050,243970,279...
60005	neg,81354,na,2130706432,156,0,0,0,0,0,0,62040,...
60006	neg,39308,na,452,374,0,0,0,0,0,0,3564,1275494,...
60007	neg,14,0,2,2,0,0,0,0,0,408,3274,614,2336,0,0,0...
60008	neg,0,2,0,0,0,0,0,0,0,0,14,1772,1036,0,0,0,554...
60009	neg,32,0,0,na,0,0,0,0,0,0,14,4032,4960,0,0,0,1...
60010	neg,153002,na,664,186,0,0,0,0,0,2564,59100,160...
60011	neg,2286,na,2130706538,224,0,0,0,0,0,0,104,991...
60012	neg,112,0,2130706432,18,0,0,0,0,0,0,28,11592,1...
60013	neg,80292,na,2130706432,494,0,0,0,0,0,0,330,20...
60014	neg,40222,na,698,628,0,0,0,0,0,0,1226,46284,19...

60015 rows × 1 columns

In [223]:
```
for i in range(15):
    data.drop(i, inplace = True)
```

In [224]:
```
data
```

Out[224]:

	0
15	neg,76698,na,2130706438,280,0,0,0,0,0,37250,...
16	neg,33058,na,0,na,0,0,0,0,0,18254,653294,172...
17	neg,41040,na,228,100,0,0,0,0,0,1648,370592,1...
18	neg,12,0,70,66,0,10,0,0,318,2212,3232,1872,0...
19	neg,60874,na,1368,458,0,0,0,0,0,43752,196661...
20	neg,38312,na,2130706432,218,0,0,0,0,0,9128,7...
21	neg,14,0,6,na,0,0,0,0,0,1202,3766,1150,0,0,0...
22	neg,102960,na,2130706432,116,0,0,0,0,0,2130,...
23	neg,78696,na,0,na,0,0,0,0,0,458,440704,43988...
24	pos,153204,0,182,na,0,0,0,0,0,11804,684444,326...
25	neg,39196,na,204,170,0,0,0,0,0,4352,713882,1...
26	neg,45912,na,0,454,0,0,0,0,0,2106,418940,221...
27	neg,2104,na,36,26,0,0,0,0,9744,13148,98310,222...
28	neg,118950,na,1390,1298,0,0,0,0,0,40932,1657...
29	neg,24416,na,0,na,0,0,0,0,0,556,448484,88779...
30	neg,14,0,62,34,0,0,0,0,0,6,4256,2250,0,0,0,2...
31	neg,31300,0,784,740,0,0,0,0,0,20386,847078,240...
32	neg,736,2,24,22,16,20,0,0,0,1144,39526,13380,4...
33	neg,332,na,2130706432,20,0,0,0,0,0,82,8182,1...
34	neg,1432,na,2130706440,82,0,0,0,0,0,78,17784...
35	neg,41212,0,2130706434,104,104,172,0,0,0,926...
36	neg,14,na,6,6,0,0,0,0,0,66,184,4352,936,0,0,...
37	neg,157128,na,2130706456,424,0,0,0,0,0,26568...
38	pos,453236,na,2926,na,0,0,0,222,323436,29992...

	0
39	neg,58246,na,2130706432,2416,0,0,0,0,0,12714...
40	neg,29394,na,0,na,0,0,0,0,0,952,334198,13618...
41	neg,8690,na,476,364,0,0,0,0,0,602,173208,440...
42	neg,46978,na,334,322,0,0,0,0,0,55002,1274996...
43	neg,1870,na,na,na,0,0,0,0,0,140,68816,42816,...
44	neg,12516,0,120,na,0,0,0,0,21886,631842,5713...
...	...
59985	neg,16,0,8,na,0,0,0,0,0,108,5176,2234,0,0,0,...
59986	neg,39026,na,202,168,0,0,0,0,0,2150,1019858,...
59987	neg,14,na,24,20,0,0,0,0,0,44,396,5598,0,0,0,...
59988	neg,3248,8,16,10,0,0,0,0,0,9046,123142,100580,...
59989	neg,83818,na,552,532,0,0,0,0,0,3088,213034,3...
59990	neg,40274,na,98,94,0,0,0,0,0,338,373034,1810...
59991	neg,16978,na,2130706434,1750,0,0,0,0,0,582,6...
59992	neg,30320,na,1838,1278,0,0,0,0,0,22386,76403...
59993	neg,38414,na,888,758,0,0,0,0,0,41546,1732152...
59994	neg,18,0,2130706432,18,0,0,0,0,0,3656,1728,2...
59995	neg,562,0,4,4,0,0,0,0,0,11538,34194,2252,0,0...
59996	neg,16,na,20,20,0,0,0,0,0,90,1304,7306,696,0...
59997	neg,10628,na,2130706434,98,0,0,0,0,4868,1167...
59998	neg,39004,na,90,84,0,0,0,0,9360,208852,13644...
59999	neg,33386,na,812,744,0,0,0,0,0,11470,650532,...
60000	neg,10792,na,784,608,0,0,0,0,0,31894,661618,...
60001	neg,644,na,12,na,0,0,0,0,18118,43536,41808,2...
60002	neg,41330,na,2130706432,744,0,0,0,0,0,32154,...
60003	neg,6078,na,52,46,0,0,0,0,0,76,22588,173848,...

	0
60004	neg,61478,na,134,0,0,0,0,0,0,2050,243970,279...
60005	neg,81354,na,2130706432,156,0,0,0,0,0,62040,...
60006	neg,39308,na,452,374,0,0,0,0,0,3564,1275494,...
60007	neg,14,0,2,2,0,0,0,0,408,3274,614,2336,0,0,0...
60008	neg,0,2,0,0,0,0,0,0,0,14,1772,1036,0,0,0,554...
60009	neg,32,0,0,na,0,0,0,0,0,14,4032,4960,0,0,0,1...
60010	neg,153002,na,664,186,0,0,0,0,0,2564,59100,160...
60011	neg,2286,na,2130706538,224,0,0,0,0,0,104,991...
60012	neg,112,0,2130706432,18,0,0,0,0,0,28,11592,1...
60013	neg,80292,na,2130706432,494,0,0,0,0,0,330,20...
60014	neg,40222,na,698,628,0,0,0,0,0,1226,46284,19...

60000 rows × 1 columns

In [186]:
```
columnNames = data.loc[14]
```

In [187]:
```
print(columnNames)
print(type(columnNames))
```

```
0    class,aa_000,ab_000,ac_000,ad_000,ae_000,af_00...
Name: 14, dtype: object
<class 'pandas.core.series.Series'>
```

In [188]:
```
columnNamesSplit = columnNames.str.split(",")
```

In [189]:
```
print(type(columnNamesSplit))
print(columnNamesSplit)
```

```
<class 'pandas.core.series.Series'>
0    [class, aa_000, ab_000, ac_000, ad_000, ae_000...
Name: 14, dtype: object
```

In [250]:
```
columnNames = []
for j in columnNamesSplit:
    for l in j:
        columnNames.append(l)
```

In [246]:
```
rows = []
for i in range(60000):
    subRows = []
    row_i = data.loc[i + 15]
    row_i_Split = row_i.str.split(",")
    for j in row_i_Split:
        for l in j:
            subRows.append(l)
    rows.append(subRows)
```

In [247]:
```
df = pd.DataFrame(rows, columns=columnNames)
```

In [248]:
```
df
```

Out[248]:

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	ee
0	neg	76698	na	2130706438	280	0	0	0	0	0	...	1240520	493384	721044	469792	339156	15
1	neg	33058	na	0	na	0	0	0	0	0	...	421400	178064	293306	245416	133654	8
2	neg	41040	na	228	100	0	0	0	0	0	...	277378	159812	423992	409564	320746	15
3	neg	12	0	70	66	0	10	0	0	0	...	240	46	58	44	10	
4	neg	60874	na	1368	458	0	0	0	0	0	...	622012	229790	405298	347188	286954	31
5	neg	38312	na	2130706432	218	0	0	0	0	0	...	388574	288278	900430	300412	1534	
6	neg	14	0	6	na	0	0	0	0	0	...	168	48	60	28	0	
7	neg	102960	na	2130706432	116	0	0	0	0	0	...	715518	384948	915978	1052166	1108672	34
8	neg	78696	na	0	na	0	0	0	0	0	...	699290	362510	1190028	1012704	160090	6
9	pos	153204	0	182	na	0	0	0	0	0	...	129862	26872	34044	22472	34362	
10	neg	39196	na	204	170	0	0	0	0	0	...	198386	99614	215734	189966	219996	20
11	neg	45912	na	0	454	0	0	0	0	0	...	495400	278660	544710	438256	249386	7
12	neg	2104	na	36	26	0	0	0	0	9744	...	4776	2330	4634	6534	29980	3
13	neg	118950	na	1390	1298	0	0	0	0	0	...	1677646	1380302	2531322	426284	17996	1
14	neg	24416	na	0	na	0	0	0	0	0	...	190606	95674	207476	232470	154962	9
15	neg	14	0	62	34	0	0	0	0	0	...	248	42	48	28	32	
16	neg	31300	0	784	740	0	0	0	0	0	...	798872	112724	51736	7054	6628	2
17	neg	736	2	24	22	16	20	0	0	0	...	1370	812	2488	2340	2566	2
18	neg	332	na	2130706432	20	0	0	0	0	0	...	1260	296	878	11774	8	
19	neg	1432	na	2130706440	82	0	0	0	0	0	...	2714	1480	4822	53694	1116	
20	neg	41212	0	2130706434	104	104	172	0	0	0	...	171390	95844	177206	173184	404690	73
21	neg	14	na	6	6	0	0	0	0	0	...	80	0	0	0	0	
22	neg	157128	na	2130706456	424	0	0	0	0	0	...	1784108	683976	1228736	971366	838200	55
23	pos	453236	na	2926	na	0	0	0	0	222	...	7908038	3026002	5025350	2025766	1160638	53

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	ee
24	neg	58246	na	2130706432	2416	0	0	0	0	0	...	670578	293816	464286	336782	229078	16
25	neg	29394	na	0	na	0	0	0	0	0	...	234998	136388	470258	548540	8384	
26	neg	8690	na	476	364	0	0	0	0	0	...	52700	28890	55588	85370	140892	1
27	neg	46978	na	334	322	0	0	0	0	0	...	407008	221836	534336	606554	325542	8
28	neg	1870	na	na	na	0	0	0	0	0	...	7688	5050	33502	40466	1436	
29	neg	12516	0	120	na	0	0	0	0	0	...	133648	53830	92568	72416	70266	11
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
59970	neg	16	0	8	na	0	0	0	0	0	...	78	30	56	28	34	
59971	neg	39026	na	202	168	0	0	0	0	0	...	277756	155262	419110	597950	306722	1
59972	neg	14	na	24	20	0	0	0	0	0	...	388	56	30	0	0	
59973	neg	3248	8	16	10	0	0	0	0	0	...	10464	6202	11372	24842	25718	5
59974	neg	83818	na	552	532	0	0	0	0	0	...	713474	343926	681198	532938	376156	26
59975	neg	40274	na	98	94	0	0	0	0	0	...	313326	182178	466482	587110	271706	3
59976	neg	16978	na	2130706434	1750	0	0	0	0	0	...	167080	87242	141522	147958	103198	6
59977	neg	30320	na	1838	1278	0	0	0	0	0	...	153770	86110	246966	274254	220696	14
59978	neg	38414	na	888	758	0	0	0	0	0	...	391604	150648	269002	200140	149768	11
59979	neg	18	0	2130706432	18	0	0	0	0	0	...	56	16	34	22	86	
59980	neg	562	0	4	4	0	0	0	0	0	...	2880	644	1110	526	1162	1
59981	neg	16	na	20	20	0	0	0	0	0	...	256	10	30	48	30	
59982	neg	10628	na	2130706434	98	0	0	0	0	0	...	553156	99122	21720	7124	1464	
59983	neg	39004	na	90	84	0	0	0	0	0	...	337238	179550	349658	263388	164592	11
59984	neg	33386	na	812	744	0	0	0	0	0	...	265982	135322	255766	208478	188012	24
59985	neg	10792	na	784	608	0	0	0	0	0	...	155506	49046	98602	55236	15122	1
59986	neg	644	na	12	na	0	0	0	0	0	...	5466	2164	6338	5008	5328	
59987	neg	41330	na	2130706432	744	0	0	0	0	0	...	434454	165964	497548	392222	256944	8
59988	neg	6078	na	52	46	0	0	0	0	0	...	20474	12750	52018	190226	3504	

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	ee
59989	neg	61478	na	134	0	0	0	0	0	0	...	632856	410326	1130874	535990	205766	3
59990	neg	81354	na	2130706432	156	0	0	0	0	0	...	627610	376028	931828	1035236	674626	16
59991	neg	39308	na	452	374	0	0	0	0	0	...	312426	161694	323438	292350	250112	15
59992	neg	14	0	2	2	0	0	0	0	0	...	132	44	80	52	54	
59993	neg	0	2	0	0	0	0	0	0	0	...	0	0	0	0	0	
59994	neg	32	0	0	na	0	0	0	0	0	...	258	98	224	166	36	
59995	neg	153002	na	664	186	0	0	0	0	0	...	998500	566884	1290398	1218244	1019768	71
59996	neg	2286	na	2130706538	224	0	0	0	0	0	...	10578	6760	21126	68424	136	
59997	neg	112	0	2130706432	18	0	0	0	0	0	...	792	386	452	144	146	
59998	neg	80292	na	2130706432	494	0	0	0	0	0	...	699352	222654	347378	225724	194440	16
59999	neg	40222	na	698	628	0	0	0	0	0	...	440066	183200	344546	254068	225148	15

60000 rows × 171 columns

In [251]:
```
df.to_csv("ScaniaTrucksTrainingSet.csv")
```

In [262]:
```
data = pd.read_csv("aps_failure_test_set.csv", sep = "delimiter", header = None)
```

```
C:\Users\GauravP\Anaconda3\lib\site-packages\ipykernel_launcher.py:1: ParserWarning: Falling back to the 'python' engin
e because the 'c' engine does not support regex separators (separators > 1 char and different from '\s+' are interprete
d as regex); you can avoid this warning by specifying engine='python'.
  """Entry point for launching an IPython kernel.
```

In [263]:
```
data
```

Out[263]:

	0
0	This file is part of APS Failure and Operation...
1	Copyright (c) <2016> <Scania CV AB>
2	This program (APS Failure and Operational Data...
3	free software: you can redistribute it and/or ...
4	it under the terms of the GNU General Public L...
5	the Free Software Foundation, either version 3...
6	(at your option) any later version.
7	This program is distributed in the hope that i...
8	but WITHOUT ANY WARRANTY; without even the imp...
9	MERCHANTABILITY or FITNESS FOR A PARTICULAR PU...
10	GNU General Public License for more details.
11	You should have received a copy of the GNU Gen...
12	along with this program. If not, see <http://...
13	----------...
14	class,aa_000,ab_000,ac_000,ad_000,ae_000,af_00...
15	neg,60,0,20,12,0,0,0,0,0,2682,4736,3862,1846,0...
16	neg,82,0,68,40,0,0,0,0,0,0,748,12594,3636,0,0,...
17	neg,66002,2,212,112,0,0,0,0,0,199486,1358536,1...
18	neg,59816,na,1010,936,0,0,0,0,0,0,123922,98431...
19	neg,1814,na,156,140,0,0,0,0,0,0,72,17926,82834...
20	neg,174,na,26,24,0,0,na,na,na,na,na,na,na,na,n...
21	neg,40202,na,640,606,0,0,0,0,0,0,156074,677848...
22	neg,120278,na,2130706432,256,0,0,0,0,0,0,208,6...
23	neg,14592,0,2130706438,162,0,0,0,0,0,0,90276,7...

	0
24	neg,28338,na,2130706432,238,0,0,0,0,0,0,582,31...
25	neg,30758,na,450,402,0,0,0,0,0,0,44326,813956,...
26	neg,40358,na,1672,1378,0,0,0,0,0,0,38954,348954,...
27	neg,33072,na,na,na,na,na,0,0,0,0,278,375848,33...
28	neg,30664,na,530,514,0,0,0,0,0,0,362,139478,12...
29	neg,41628,na,2130706432,380,0,0,0,0,0,0,6980,7...
...	...
15985	neg,41416,na,102,98,0,0,0,0,0,0,578,406590,165...
15986	neg,30346,na,0,na,0,0,0,0,0,0,72736,1859756,24...
15987	neg,61584,na,860,398,0,0,0,0,0,0,2880,283468,4...
15988	neg,39310,na,498,434,0,0,0,0,0,0,24620,558788,...
15989	neg,2,na,4,4,0,0,0,0,0,0,18,0,4978,0,0,0,1286,...
15990	neg,724,0,10,8,0,0,0,0,0,2796,39516,6124,3072,...
15991	neg,32540,na,730,1520,0,0,0,0,0,0,256,169356,1...
15992	neg,1084610,na,na,na,na,na,0,0,13320,2388752,1...
15993	neg,1138,0,44,na,0,0,0,0,962,26932,36318,45282...
15994	neg,8,0,12,4,0,0,0,0,0,1710,8542,6816,1976,0,0...
15995	neg,38636,na,118,112,0,0,0,0,0,0,100054,344966...
15996	neg,61998,na,2816,2590,0,0,0,0,0,0,1814,412610...
15997	neg,1104,na,88,82,0,0,0,0,0,0,136,68,46698,947...
15998	neg,40350,na,2130706432,430,0,0,0,0,0,0,356,30...
15999	neg,136186,na,2130706434,2226,0,0,0,0,0,0,8504...
16000	neg,40542,na,50,48,0,0,0,0,0,0,5292,262572,161...
16001	neg,186,na,14,14,0,0,0,0,0,8498,13652,14832,20...
16002	neg,40586,na,578,476,0,0,0,0,0,0,266,148520,19...
16003	neg,38532,na,0,na,0,0,0,0,0,0,10082,1112476,11...

	0
16004	neg,374,na,0,na,0,0,0,0,0,0,7122,18184,3002,0,...
16005	neg,7256,na,280,258,0,0,0,0,0,0,336,69936,8614...
16006	neg,40262,na,1802,1720,0,0,0,0,0,0,10808,12136...
16007	neg,1298,na,80,70,0,0,0,0,0,0,88,30196,52746,0...
16008	neg,2146,na,2130706432,300,0,0,0,0,0,0,37834,5...
16009	pos,215220,na,414,0,0,0,0,0,149300,2650616,493...
16010	neg,81852,na,2130706432,892,0,0,0,0,0,0,5486,7...
16011	neg,18,0,52,46,8,26,0,0,0,0,6444,3308,2720,0,0...
16012	neg,79636,na,1670,1518,0,0,0,0,0,0,15062,19912...
16013	neg,110,na,36,32,0,0,0,0,0,0,198,3730,27418,97...
16014	neg,8,0,6,4,2,2,0,0,0,0,1350,3328,1126,0,0,0,2...

16015 rows × 1 columns

In [259]:
```
for i in range(15):
    data.drop(i, inplace = True)
```

In [260]:
```
data
```

Out[260]:

	0
15	neg,60,0,20,12,0,0,0,0,0,2682,4736,3862,1846,0...
16	neg,82,0,68,40,0,0,0,0,0,0,748,12594,3636,0,0,...
17	neg,66002,2,212,112,0,0,0,0,0,199486,1358536,1...
18	neg,59816,na,1010,936,0,0,0,0,0,0,123922,98431...
19	neg,1814,na,156,140,0,0,0,0,0,0,72,17926,82834...
20	neg,174,na,26,24,0,0,na,na,na,na,na,na,na,na,n...
21	neg,40202,na,640,606,0,0,0,0,0,0,156074,677848...
22	neg,120278,na,2130706432,256,0,0,0,0,0,0,208,6...
23	neg,14592,0,2130706438,162,0,0,0,0,0,0,90276,7...
24	neg,28338,na,2130706432,238,0,0,0,0,0,0,582,31...
25	neg,30758,na,450,402,0,0,0,0,0,0,44326,813956,...
26	neg,40358,na,1672,1378,0,0,0,0,0,38954,348954,...
27	neg,33072,na,na,na,na,na,0,0,0,0,278,375848,33...
28	neg,30664,na,530,514,0,0,0,0,0,0,362,139478,12...
29	neg,41628,na,2130706432,380,0,0,0,0,0,0,6980,7...
30	neg,1204,0,2130706432,na,0,0,0,0,0,0,0,0,0,0,0...
31	neg,1434,na,32,26,0,0,0,0,0,0,112,7108,75972,3...
32	neg,31576,na,378,350,0,0,0,0,0,0,10740,1306844...
33	neg,30812,na,2130706432,144,0,0,0,0,0,0,262,19...
34	neg,846,na,100,52,0,0,0,0,0,12988,57428,26286,...
35	neg,42676,na,696,548,0,0,0,0,0,0,2374,312058,1...
36	neg,30,na,34,32,0,0,0,0,0,0,46,870,6806,2750,0...
37	neg,121576,na,1218,na,0,0,0,0,0,0,884,308914,6...
38	neg,24272,na,176,168,0,0,0,0,0,0,5618,520860,8...

	0
39	neg,98470,na,2130706432,284,0,0,0,0,4956,773...
40	neg,970,0,2130706432,52,0,0,0,0,0,1376,6430,...
41	neg,8,0,10,10,0,0,0,0,0,0,2620,3486,1994,0,0,0...
42	neg,63496,na,1094,0,0,0,0,0,0,0,1036,155646,32...
43	neg,844,0,26,26,0,0,0,0,340,39950,20464,3926,1...
44	neg,32200,na,564,0,0,0,0,0,0,0,1626,129788,159...
...	...
15985	neg,41416,na,102,98,0,0,0,0,0,0,578,406590,165...
15986	neg,30346,na,0,na,0,0,0,0,0,0,72736,1859756,24...
15987	neg,61584,na,860,398,0,0,0,0,0,0,2880,283468,4...
15988	neg,39310,na,498,434,0,0,0,0,0,0,24620,558788,...
15989	neg,2,na,4,4,0,0,0,0,0,0,18,0,4978,0,0,0,1286,...
15990	neg,724,0,10,8,0,0,0,0,0,2796,39516,6124,3072,...
15991	neg,32540,na,730,1520,0,0,0,0,0,0,256,169356,1...
15992	neg,1084610,na,na,na,na,na,0,0,13320,2388752,1...
15993	neg,1138,0,44,na,0,0,0,0,962,26932,36318,45282...
15994	neg,8,0,12,4,0,0,0,0,0,1710,8542,6816,1976,0,0...
15995	neg,38636,na,118,112,0,0,0,0,0,0,100054,344966...
15996	neg,61998,na,2816,2590,0,0,0,0,0,0,1814,412610...
15997	neg,1104,na,88,82,0,0,0,0,0,0,136,68,46698,947...
15998	neg,40350,na,2130706432,430,0,0,0,0,0,0,356,30...
15999	neg,136186,na,2130706434,2226,0,0,0,0,0,0,8504...
16000	neg,40542,na,50,48,0,0,0,0,0,0,5292,262572,161...
16001	neg,186,na,14,14,0,0,0,0,0,8498,13652,14832,20...
16002	neg,40586,na,578,476,0,0,0,0,0,0,266,148520,19...
16003	neg,38532,na,0,na,0,0,0,0,0,0,10082,1112476,11...

	0
16004	neg,374,na,0,na,0,0,0,0,0,0,7122,18184,3002,0,...
16005	neg,7256,na,280,258,0,0,0,0,0,0,336,69936,8614...
16006	neg,40262,na,1802,1720,0,0,0,0,0,0,10808,12136...
16007	neg,1298,na,80,70,0,0,0,0,0,0,88,30196,52746,0...
16008	neg,2146,na,2130706432,300,0,0,0,0,0,0,37834,5...
16009	pos,215220,na,414,0,0,0,0,0,149300,2650616,493...
16010	neg,81852,na,2130706432,892,0,0,0,0,0,0,5486,7...
16011	neg,18,0,52,46,8,26,0,0,0,0,6444,3308,2720,0,0...
16012	neg,79636,na,1670,1518,0,0,0,0,0,0,15062,19912...
16013	neg,110,na,36,32,0,0,0,0,0,0,198,3730,27418,97...
16014	neg,8,0,6,4,2,2,0,0,0,0,1350,3328,1126,0,0,0,2...

16000 rows × 1 columns

In [275]:
```
columnsNames = data.loc[14]
```

In [276]:
```
columnsNames
```

Out[276]:
```
0    class,aa_000,ab_000,ac_000,ad_000,ae_000,af_00...
Name: 14, dtype: object
```

In [279]:
```
type(columnsNames)
```

Out[279]:
```
pandas.core.series.Series
```

In [280]:
```
columnNamesSplit = columnsNames.str.split(",")
```

In [281]:
```
columnNames = []
for j in columnNamesSplit:
    for l in j:
        columnNames.append(l)
```

In [283]:
```
rows = []
for i in range(16000):
    subRows = []
    row_i = data.loc[i + 15]
    row_i_Split = row_i.str.split(",")
    for j in row_i_Split:
        for l in j:
            subRows.append(l)
    rows.append(subRows)
```

In [284]:
```
df = pd.DataFrame(rows, columns=columnNames)
```

In [285]:
```
df
```

Out[285]:

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006
0	neg	60	0	20	12	0	0	0	0	0	...	1098	138	412	654	78
1	neg	82	0	68	40	0	0	0	0	0	...	1068	276	1620	116	86
2	neg	66002	2	212	112	0	0	0	0	0	...	495076	380368	440134	269556	1315022
3	neg	59816	na	1010	936	0	0	0	0	0	...	540820	243270	483302	485332	431376
4	neg	1814	na	156	140	0	0	0	0	0	...	7646	4144	18466	49782	3176
5	neg	174	na	26	24	0	0	na	na	na	...	na	na	na	na	na
6	neg	40202	na	640	606	0	0	0	0	0	...	526218	239734	439556	374248	169096
7	neg	120278	na	2130706432	256	0	0	0	0	0	...	1006722	459658	876356	689532	568540
8	neg	14592	0	2130706438	162	0	0	0	0	0	...	105024	57398	49152	38256	243268
9	neg	28338	na	2130706432	238	0	0	0	0	0	...	357006	150056	264776	239282	140548
10	neg	30758	na	450	402	0	0	0	0	0	...	278338	123232	221032	184386	164702
11	neg	40358	na	1672	1378	0	0	0	0	0	...	318762	165798	329940	276276	200222
12	neg	33072	na	na	na	na	na	0	0	0	...	533668	188530	308002	355434	31382
13	neg	30664	na	530	514	0	0	0	0	0	...	248048	125438	262020	273194	166606
14	neg	41628	na	2130706432	380	0	0	0	0	0	...	415556	209392	651410	298746	163608
15	neg	1204	0	2130706432	na	0	0	0	0	0	...	0	0	0	0	0
16	neg	1434	na	32	26	0	0	0	0	0	...	3532	1852	5064	54170	410
17	neg	31576	na	378	350	0	0	0	0	0	...	445546	195810	285304	204240	180044
18	neg	30812	na	2130706432	144	0	0	0	0	0	...	362262	163308	307224	245010	130652
19	neg	846	na	100	52	0	0	0	0	0	...	3978	966	3558	15876	2510
20	neg	42676	na	696	548	0	0	0	0	0	...	428068	205904	390528	342716	231334
21	neg	30	na	34	32	0	0	0	0	0	...	604	20	58	10	0
22	neg	121576	na	1218	na	0	0	0	0	0	...	979374	547552	1253854	1055204	834420
23	neg	24272	na	176	168	0	0	0	0	0	...	182716	88570	157644	193214	111462

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006	
24	neg	98470	na	2130706432	284	0	0	0	0	0	...	526748	293572	674316	724476	577864	
25	neg	970	0	2130706432	52	0	0	0	0	0	...	506	318	342	736	970	
26	neg	8	0	10	10	0	0	0	0	0	...	180	22	36	0	0	
27	neg	63496	na	1094	0	0	0	0	0	0	...	614600	626038	1167352	661122	52690	
28	neg	844	0	26	26	0	0	0	0	340	...	1004	416	2660	3444	4304	
29	neg	32200	na	564	0	0	0	0	0	0	...	105080	58614	182246	284746	342304	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	
15970	neg	41416	na	102	98	0	0	0	0	0	...	296678	176368	401124	363816	251226	
15971	neg	30346	na	0	na	0	0	0	0	0	...	417120	179630	262138	138668	61056	
15972	neg	61584	na	860	398	0	0	0	0	0	...	1057368	481984	715574	617086	117658	
15973	neg	39310	na	498	434	0	0	0	0	0	...	262254	144558	374246	517732	322342	
15974	neg	2	na	4	4	0	0	0	0	0	...	342	6	0	0	0	
15975	neg	724	0	10	8	0	0	0	0	0	...	2294	1040	3888	7808	7508	
15976	neg	32540	na	730	1520	0	0	0	0	0	...	189944	80290	169822	203434	236202	
15977	neg	1084610	na	na	na	na	na	0	0	13320	...	4558940	1967118	3751872	3556346	4354632	2
15978	neg	1138	0	44	na	0	0	0	0	962	...	5052	1950	5690	9174	15966	
15979	neg	8	0	12	4	0	0	0	0	0	...	34	10	20	22	28	
15980	neg	38636	na	118	112	0	0	0	0	0	...	397730	179888	339858	271610	221140	
15981	neg	61998	na	2816	2590	0	0	0	0	0	...	582144	258570	716652	693328	331398	
15982	neg	1104	na	88	82	0	0	0	0	0	...	5556	3278	25764	12534	452	
15983	neg	40350	na	2130706432	430	0	0	0	0	0	...	284768	215712	672766	526172	168674	
15984	neg	136186	na	2130706434	2226	0	0	0	0	0	...	1618760	723430	1252324	882944	630434	
15985	neg	40542	na	50	48	0	0	0	0	0	...	304478	160170	364694	410482	253518	
15986	neg	186	na	14	14	0	0	0	0	0	...	1050	1298	2758	982	1112	
15987	neg	40586	na	578	476	0	0	0	0	0	...	421748	203364	409294	345044	228806	
15988	neg	38532	na	0	na	0	0	0	0	0	...	235342	129744	298800	298142	345866	

	class	aa_000	ab_000	ac_000	ad_000	ae_000	af_000	ag_000	ag_001	ag_002	...	ee_002	ee_003	ee_004	ee_005	ee_006
15989	neg	374	na	0	na	0	0	0	0	0	...	1104	268	556	13724	506
15990	neg	7256	na	280	258	0	0	0	0	0	...	90254	51562	108716	44928	146
15991	neg	40262	na	1802	1720	0	0	0	0	0	...	484302	201836	305924	282268	111826
15992	neg	1298	na	80	70	0	0	0	0	0	...	3716	2884	11802	38696	1376
15993	neg	2146	na	2130706432	300	0	0	0	0	0	...	11102	5980	16934	65806	268
15994	pos	215220	na	414	0	0	0	0	0	149300	...	1127092	495454	1345792	3078658	3278768
15995	neg	81852	na	2130706432	892	0	0	0	0	0	...	632658	273242	510354	373918	349840
15996	neg	18	0	52	46	8	26	0	0	0	...	266	44	46	14	2
15997	neg	79636	na	1670	1518	0	0	0	0	0	...	806832	449962	778826	581558	375498
15998	neg	110	na	36	32	0	0	0	0	0	...	588	210	180	544	1004
15999	neg	8	0	6	4	2	2	0	0	0	...	46	10	48	14	42

16000 rows × 171 columns

In [286]:
```
df.to_csv("ScaniaTrucksTestSet.csv")
```