**Target Variable**

Removing N,U and X

Columns to keep

C\_MNTH (removing UU and XX)

**C\_MNTH** count

*<chr>* *<int>*

1 1 24703

2 10 24617

3 11 23497

4 12 23255

5 2 22286

6 3 19581

7 4 17493

8 5 20259

9 6 23849

10 7 24804

11 8 24627

12 9 23328

**C\_WDAY**

|  |
| --- |
| C\_WDAY count  *<chr>* *<int>*  1 1 37786  2 2 39399  3 3 39684  4 4 41057  5 5 45610  6 6 37163  7 7 31600  8 U 2 |
|  |
| |  | | --- | | > | |

**C\_HOUR**

C\_HOUR count

*<chr>* *<int>*

1 0 3353

2 1 2543

3 10 12235

4 11 14413

5 12 17601

6 13 17179

7 14 18929

8 15 23124

9 16 24904

10 17 23709

**C\_VEHS**

C\_VEHS count

*<dbl>* *<int>*

1 1 58234

2 2 165012

3 3 34636

4 4 9175

5 5 2522

6 6 904

7 7 342

8 8 223

9 9 76

10 10 174

**C\_CONF** (requires grouping)

C\_CONF count

*<chr>* *<int>*

1 1 2836

2 2 11421

3 21 79451

4 22 8964

5 23 3336

6 24 2548

7 25 871

8 3 6646

9 31 8589

10 32 2994

11 33 20271

12 34 1590

13 35 37133

14 36 23896

15 4 9646

16 41 3771

17 5 826

18 6 23555

19 QQ 15579

20 UU 8378

**C\_RCFG** (group a few, ie 11,12 and and drop QQ UU XX)

C\_RCFG count

*<chr>* *<int>*

1 1 113317

2 10 81

3 2 121015

4 3 9829

5 4 794

6 5 2258

7 6 355

8 7 42

9 8 887

10 9 297

11 QQ 8608

12 UU 14818

**C\_WTHR**

C\_WTHR count

*<chr>* *<int>*

1 1 193680

2 2 24822

3 3 22824

4 4 18369

5 5 1860

6 6 5617

7 7 969

8 Q 616

9 U 3544

C\_**RSUR**

|  |
| --- |
| C\_RSUR count  *<chr>* *<int>*  1 1 180433  2 2 42929  3 3 13111  4 4 4411  5 5 17730  6 6 847  7 7 176  8 8 43  9 9 62  10 Q 10241  11 U 2318 |
|  |
| |  | | --- | | > | |

**C\_RALN**

*<chr>* *<int>*

1 1 77410

2 10 107

3 11 291

4 12 547

5 13 151

6 15 147

7 16 78

8 17 176

9 18 135069

10 2 993

11 3 31351

12 4 4687

13 5 100

14 6 2208

15 7 52

16 8 873

17 9 23

18 QQ 4332

19 UU 13706

**C\_TRAF (consider grouping)**

C\_TRAF count

*<chr>* *<int>*

1 1 77410

2 10 107

3 11 291

4 12 547

5 13 151

6 15 147

7 16 78

8 17 176

9 18 135069

10 2 993

11 3 31351

12 4 4687

13 5 100

14 6 2208

15 7 52

16 8 873

17 9 23

18 QQ 4332

19 UU 13706

V\_TYPE

V\_TYPE count

*<chr>* *<int>*

1 1 225381

2 10 39

3 11 2770

4 14 6032

5 16 412

6 17 4687

7 18 133

8 19 118

9 20 406

10 21 73

11 22 152

12 23 142

13 5 2384

14 6 4492

15 7 4621

16 8 3346

17 9 1364

18 NN 13093

19 QQ 1037

20 UU 1619

**V\_YEAR**

V\_YEAR count

*<chr>* *<int>*

1 1918 1

2 1920 24

3 1927 2

4 1938 2

5 1947 3

6 1950 2

7 1951 2

8 1954 2

9 1955 2

10 1956 2

**P\_SEX**

|  |
| --- |
| P\_SEX count  *<chr>* *<int>*  1 F 115007  2 M 140920  3 N 582  4 U 15792 |
|  |
| |  | | --- | | > | |

**P\_AGE**

P\_AGE count

*<chr>* *<int>*

1 1 2899

2 10 1129

3 11 1147

4 12 1161

5 13 1161

6 14 1570

7 15 1836

8 16 3002

9 17 5104

10 18 5626

**P\_PSN**

|  |
| --- |
| P\_PSN count  *<chr>* *<int>*  1 11 185435  2 12 2696  3 13 33280  4 21 8889  5 22 2771  6 23 10607  7 31 340  8 32 939  9 33 429  10 96 2990  11 97 51  12 98 201  13 99 10779  14 NN 630  15 QQ 5942  16 UU 6322 |
|  |
| |  | | --- | | > | |

**P\_USER**

|  |
| --- |
| P\_USER count  *<chr>* *<int>*  1 1 173210  2 2 60918  3 3 12685  4 4 4687  5 5 6032  6 U 14769 |
|  |
| |  | | --- | | > | |

Columns to drop

C\_YEAR

C\_SEV

V\_ID

P\_ID

P\_SAFE