

Univariate descriptive statistics of our data

Group 5 MD

2022-09-29

In this document, we provide the basic statistics and plots of our accidents database.

Dimensions of our dataset:

```
dim(dd)

## [1] 5000  23
n<-dim(dd)[1]
K<-dim(dd)[2]

K #Number of variables (columns)

## [1] 23
n #Number of instances (rows)

## [1] 5000
```

Check of the variables' names:

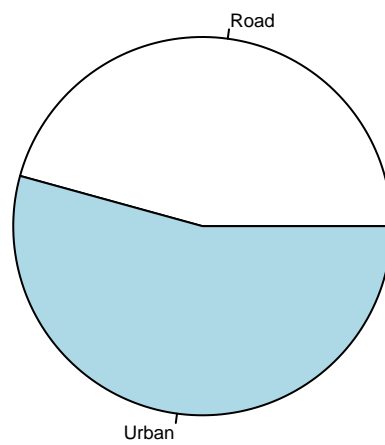
```
names(dd)

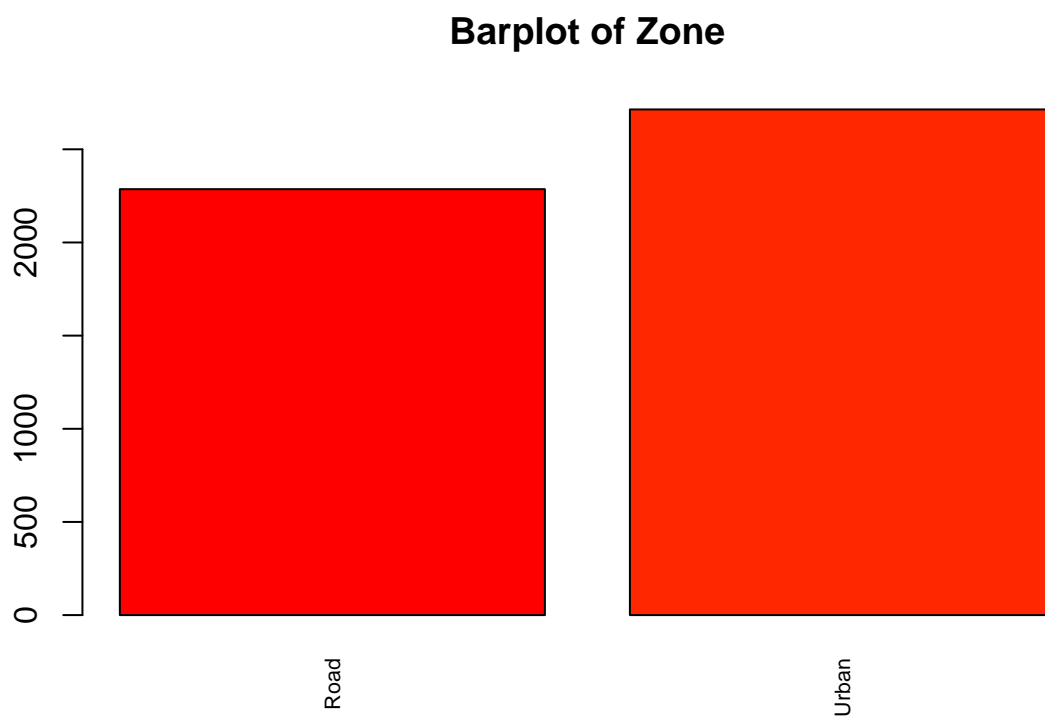
## [1] "Zone"      "Date"      "Region"    "Prov"      "nMortal"
## [6] "nGraveInj" "nMinorInj" "nInv"      "nPedest"   "nBikes"
## [11] "nMotor"    "Vel"       "Escaped"   "Weather"   "TrafficInf"
## [16] "WeatherInf" "LightInf"  "VisionInf" "Intersect" "Surface"
## [21] "DayGroup"  "HourGroup" "AccType"
```

Basic descriptive analysis of our variables:

Variable 1 : Zone

Pie of Zone





Number of modalities: 2

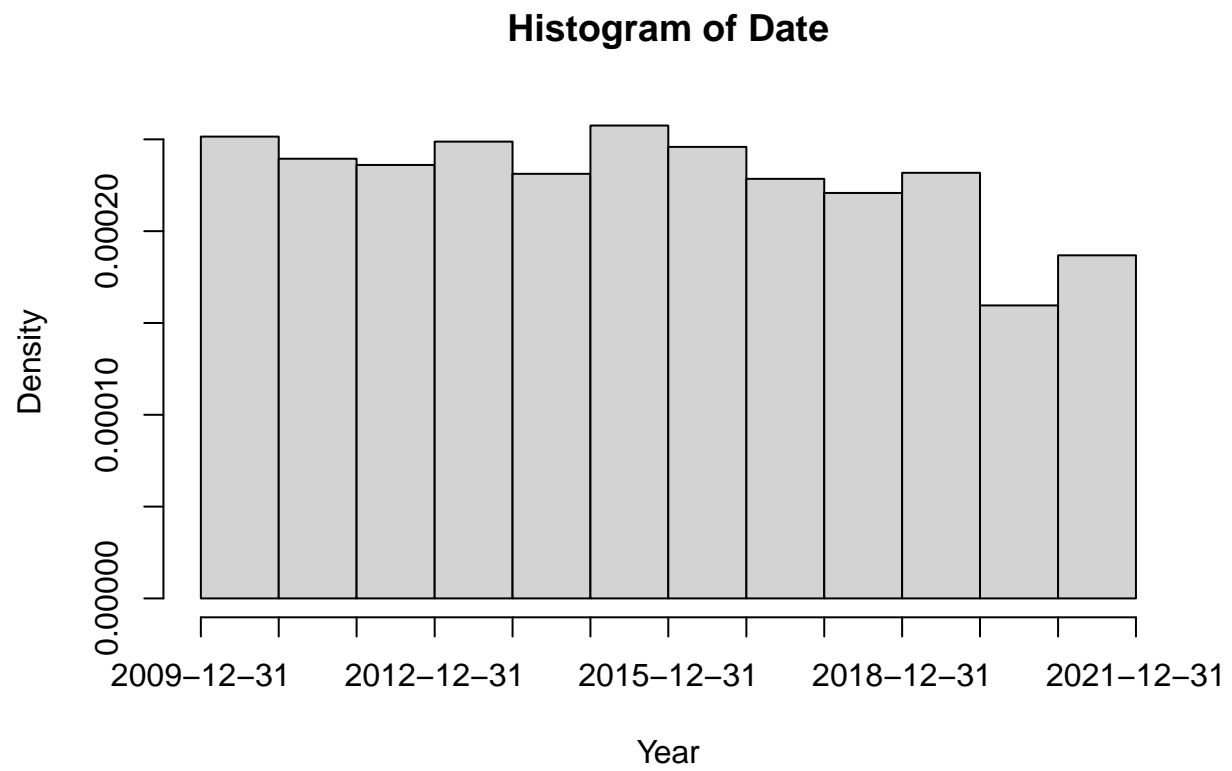
Table 1: Zone frequency table (sorted)

Zone	Frequency
Urban	2714
Road	2286

Table 2: Zone relative frequency table of proportions (sorted)

Zone	Proportion
Urban	0.5428
Road	0.4572

Variable 2 : Date

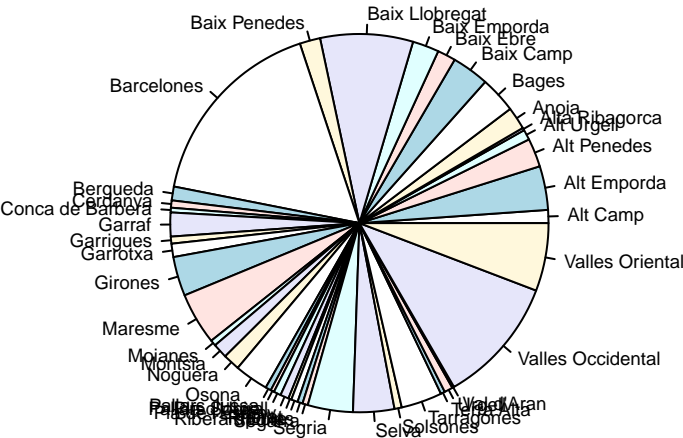


Min. -> 2010-01-01

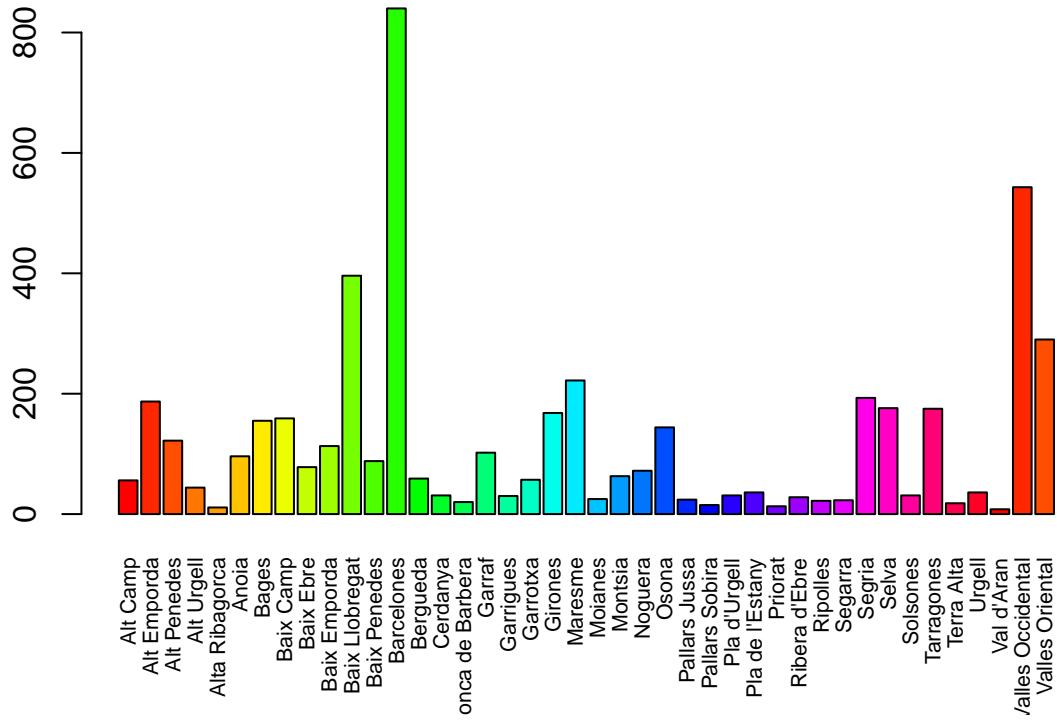
Max. -> 2021-12-30

Variable 3 : Region

Pie of Region



Barplot of Region



Number of modalities: 42

Table 3: Region frequency table (sorted)

Region	Frequency
Barcelones	840
Valles Occidental	543
Baix Llobregat	396
Valles Oriental	290
Maresme	222
Segria	193
Alt Emporda	187
Selva	176
Tarragones	175
Girones	168
Baix Camp	159
Bages	155
Osona	144
Alt Penedes	122
Baix Emporda	113
Garraf	102
Anoia	96
Baix Penedes	88
Baix Ebre	78
Noguera	72
Montsia	63

Region	Frequency
Bergueda	59
Garrotxa	57
Alt Camp	56
Alt Urgell	44
Pla de l'Estany	36
Urgell	36
Cerdanya	31
Pla d'Urgell	31
Solsones	31
Garrigues	30
Ribera d'Ebre	28
Moianes	25
Pallars Jussà	24
Segarra	23
Ripolles	22
Conca de Barbera	20
Terra Alta	18
Pallars Sobirà	15
Priorat	13
Alta Ribagorça	11
Val d'Aran	8

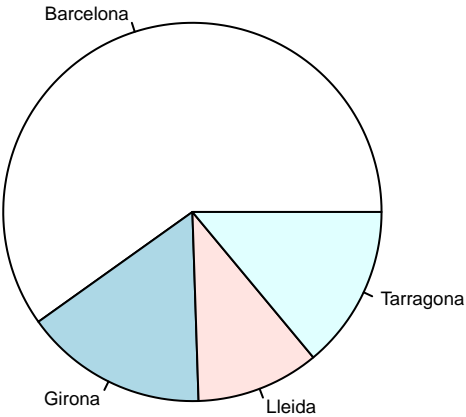
Table 4: Region relative frequency table of proportions (sorted)

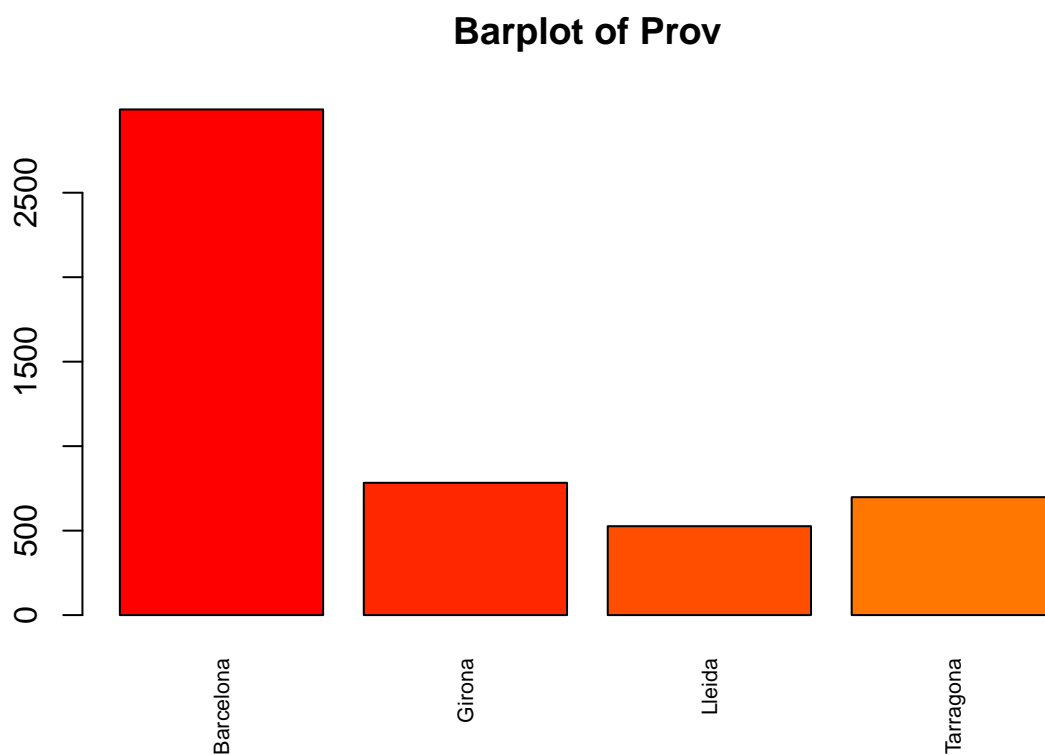
Region	Proportion
Barcelones	0.1680
Valles Occidental	0.1086
Baix Llobregat	0.0792
Valles Oriental	0.0580
Maresme	0.0444
Segria	0.0386
Alt Empordà	0.0374
Selva	0.0352
Tarragones	0.0350
Girones	0.0336
Baix Camp	0.0318
Bages	0.0310
Osona	0.0288
Alt Penedes	0.0244
Baix Empordà	0.0226
Garraf	0.0204
Anoia	0.0192
Baix Penedes	0.0176
Baix Ebre	0.0156
Noguera	0.0144
Montsia	0.0126
Bergueda	0.0118
Garrotxa	0.0114
Alt Camp	0.0112
Alt Urgell	0.0088

Region	Proportion
Pla de l'Estany	0.0072
Urgell	0.0072
Cerdanya	0.0062
Pla d'Urgell	0.0062
Solsones	0.0062
Garrigues	0.0060
Ribera d'Ebre	0.0056
Moianes	0.0050
Pallars Jussà	0.0048
Segarra	0.0046
Ripollès	0.0044
Conca de Barbera	0.0040
Terra Alta	0.0036
Pallars Sobirà	0.0030
Priorat	0.0026
Alta Ribagorça	0.0022
Val d'Aran	0.0016

Variable 4 : Prov

Pie of Prov





Number of modalities: 4

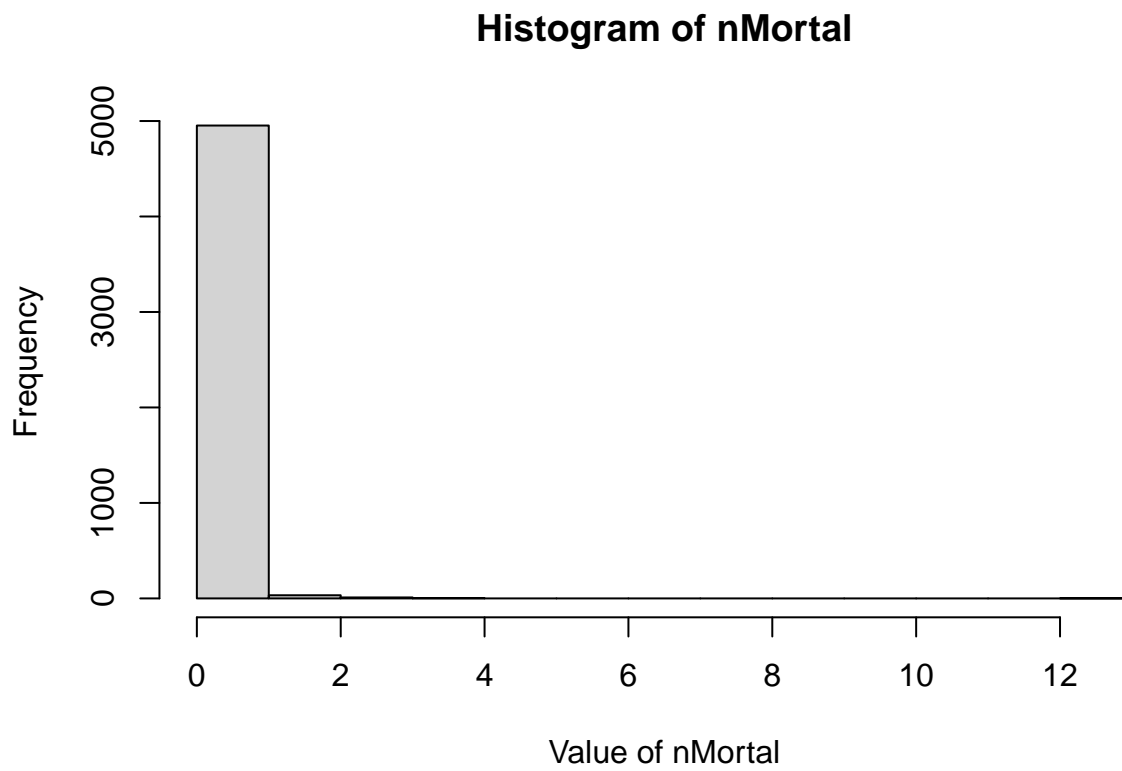
Table 5: Prov frequency table (sorted)

Prov	Frequency
Barcelona	2993
Girona	783
Tarragona	698
Lleida	526

Table 6: Prov relative frequency table of proportions (sorted)

Prov	Proportion
Barcelona	0.5986
Girona	0.1566
Tarragona	0.1396
Lleida	0.1052

Variable 5 : nMortal



Boxplot of nMortal

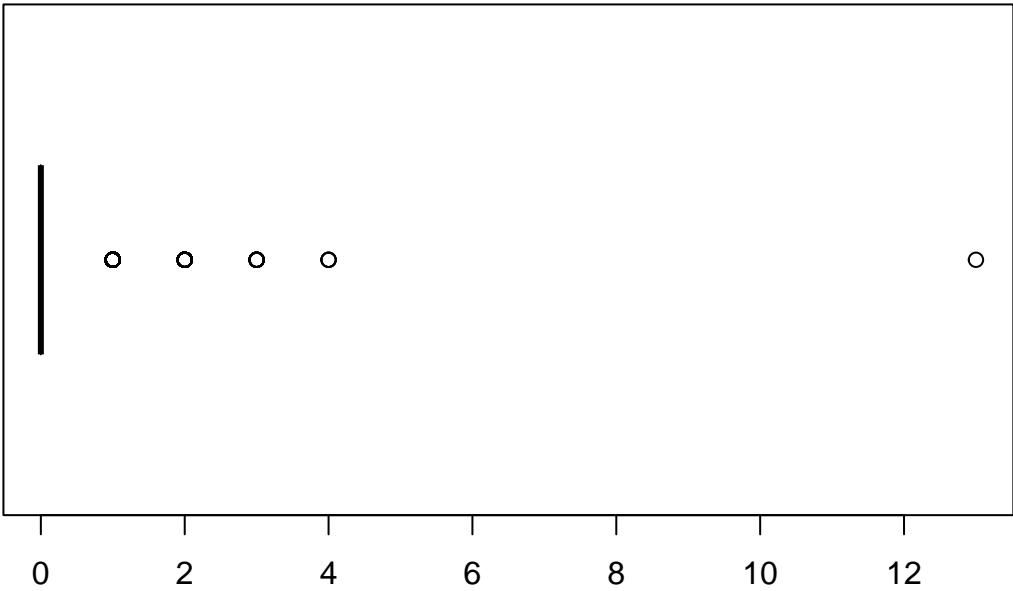
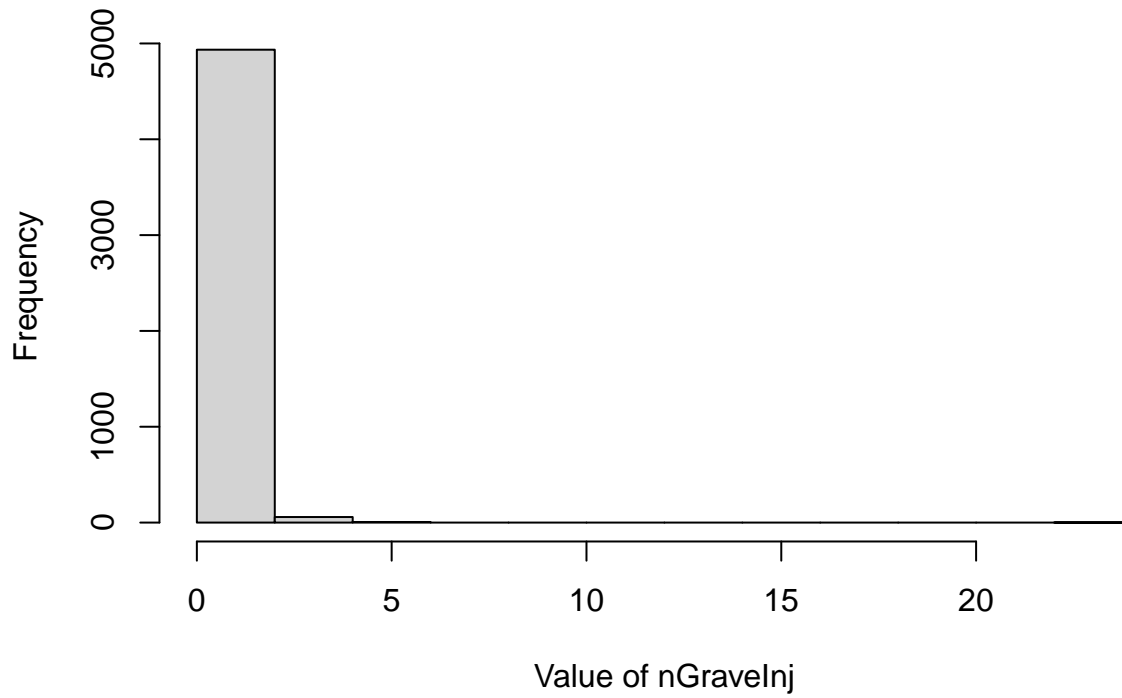


Table 7: nMortal extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	0	0	0.14	0	13	0.4289951	3.064251

Variable 6 : nGraveInj

Histogram of nGravelnj



Boxplot of nGravelInj

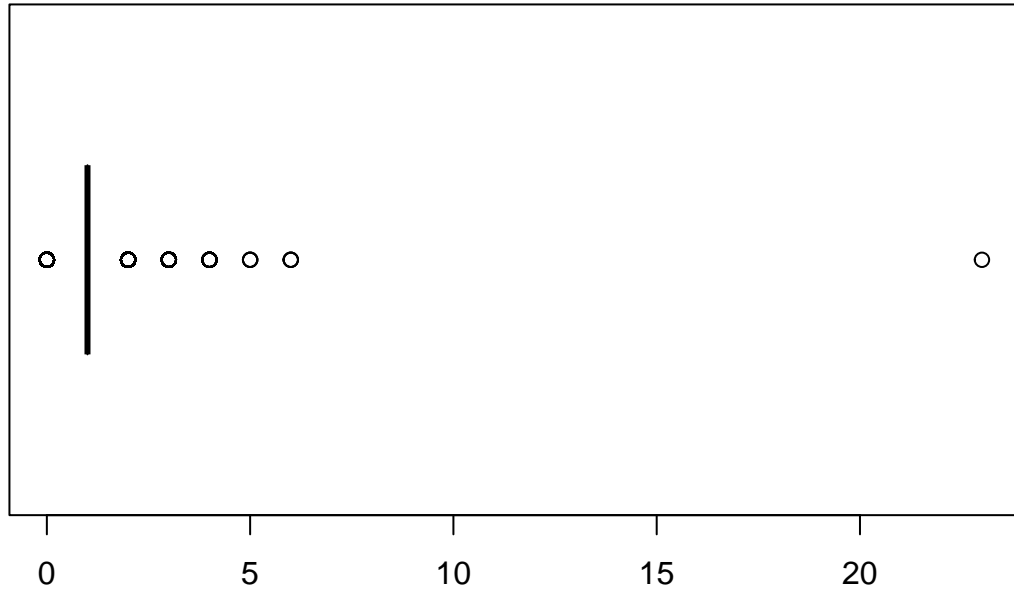


Table 8: nGraveInj extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	1	1	1.0034	1	23	0.5904728	0.588472

Variable 7 : nMinorInj



Boxplot of nMinorInj

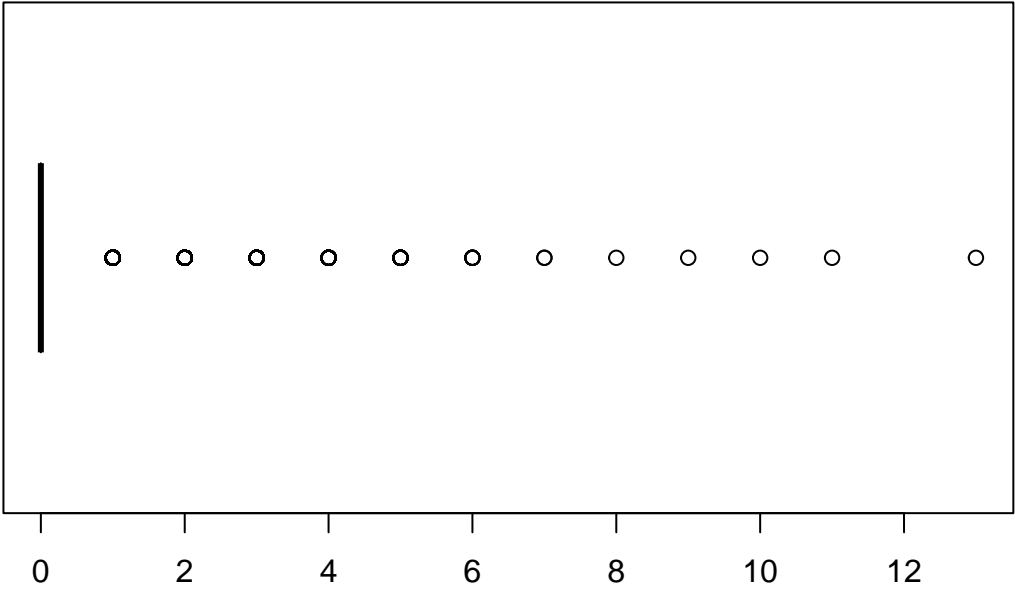
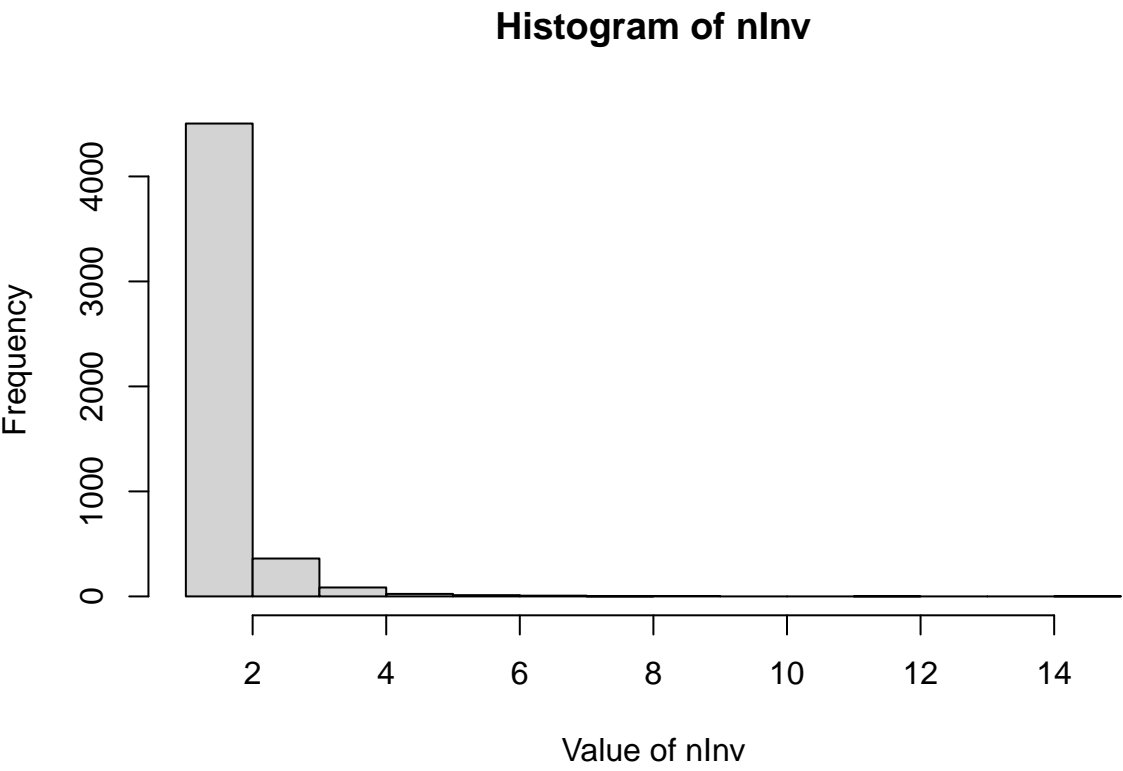


Table 9: nMinorInj extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	0	0	0.3924	0	13	0.899213	2.291572

Variable 8 : nInv



Boxplot of nInv

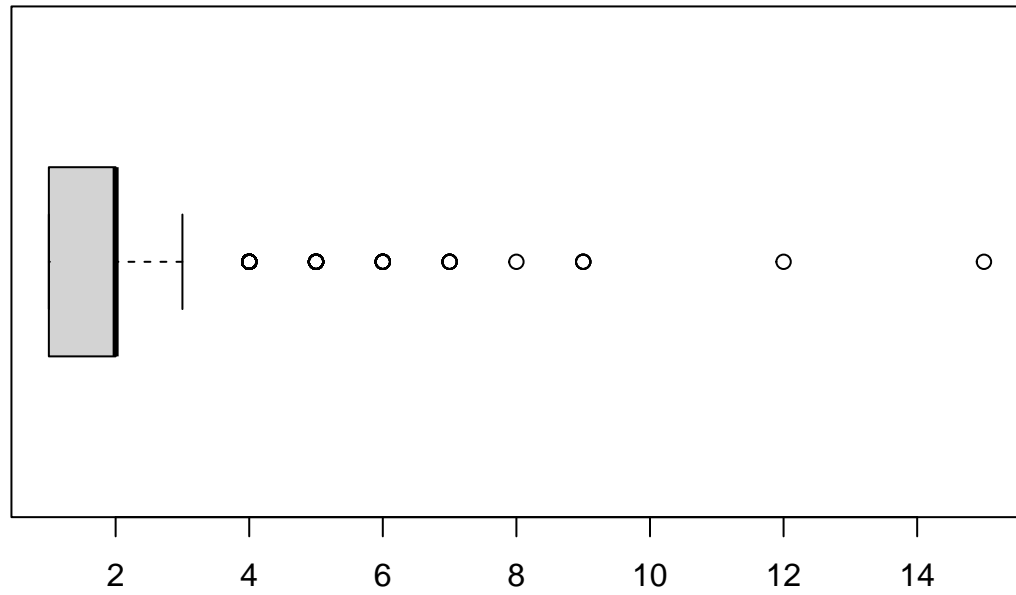
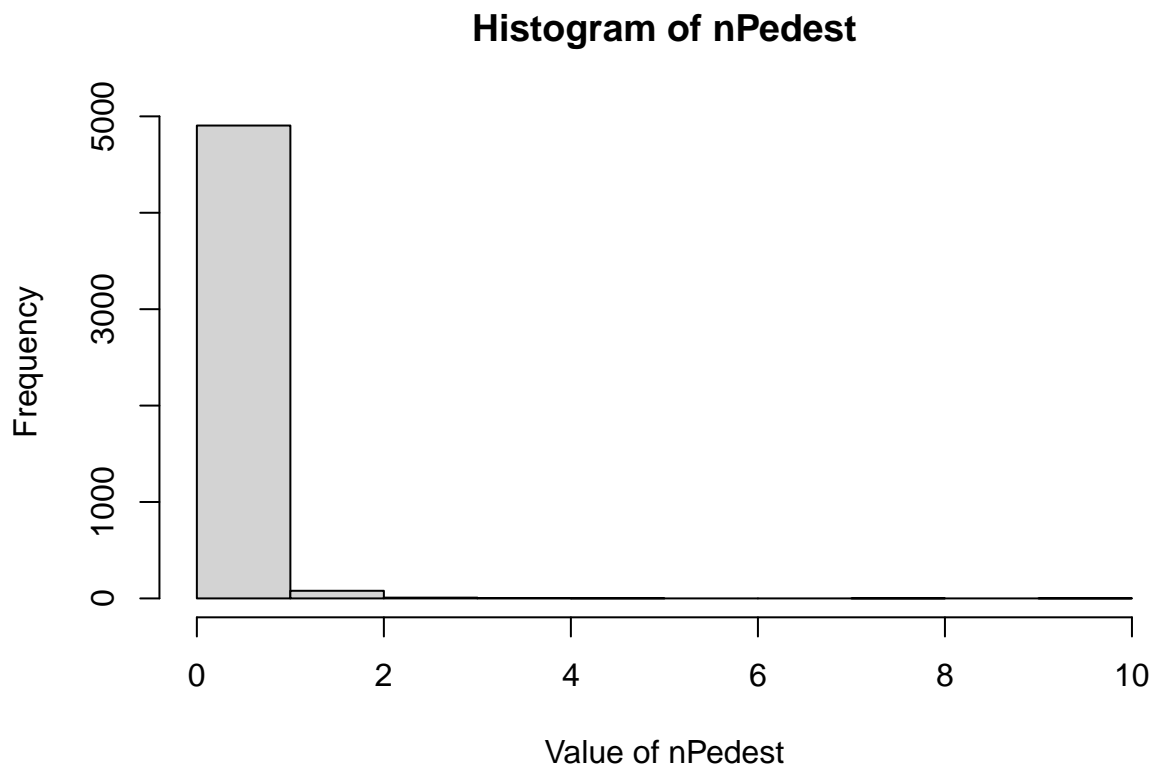


Table 10: nInv extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
1	1	2	1.8806	2	15	0.7815791	0.4156009

Variable 9 : nPedest



Boxplot of nPedest

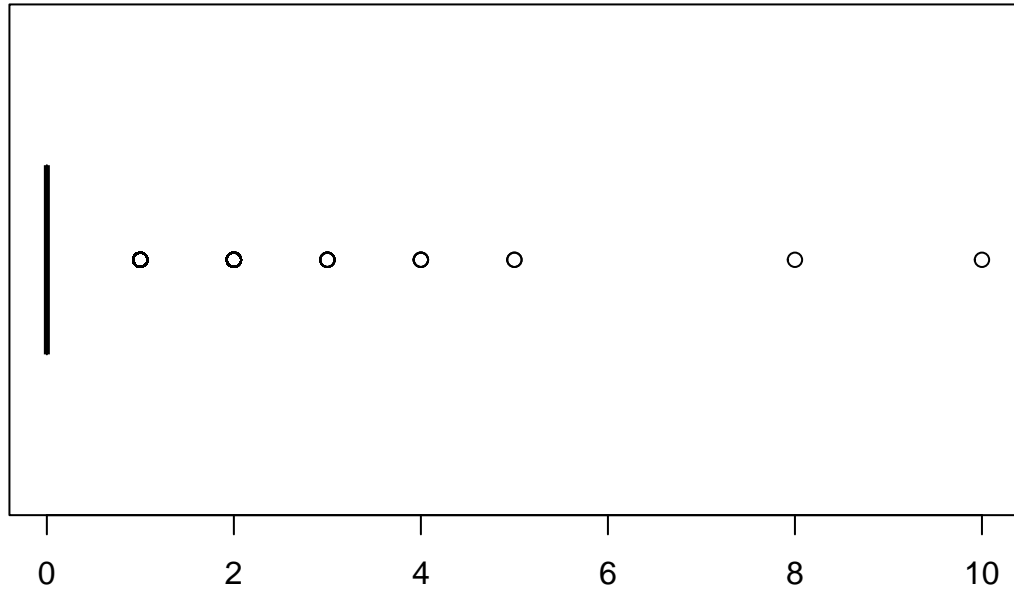
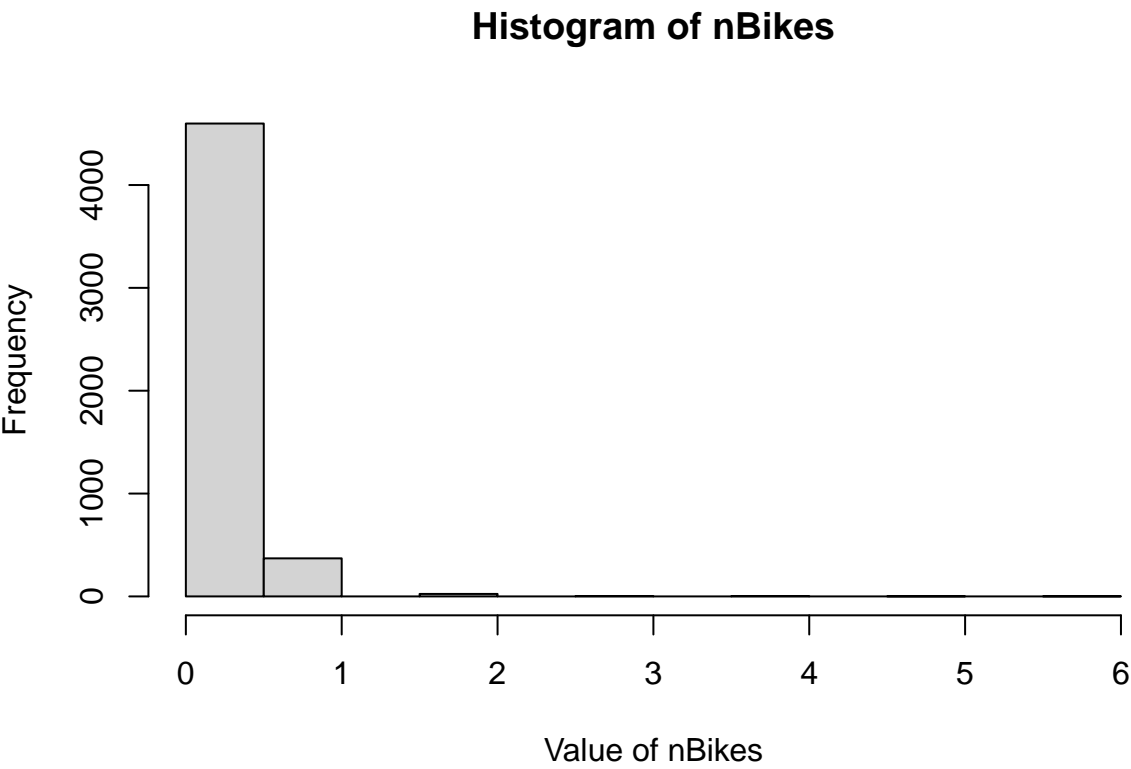


Table 11: nPedest extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	0	0	0.2456	0	10	0.5228148	2.128725

Variable 10 : nBikes



Boxplot of nBikes

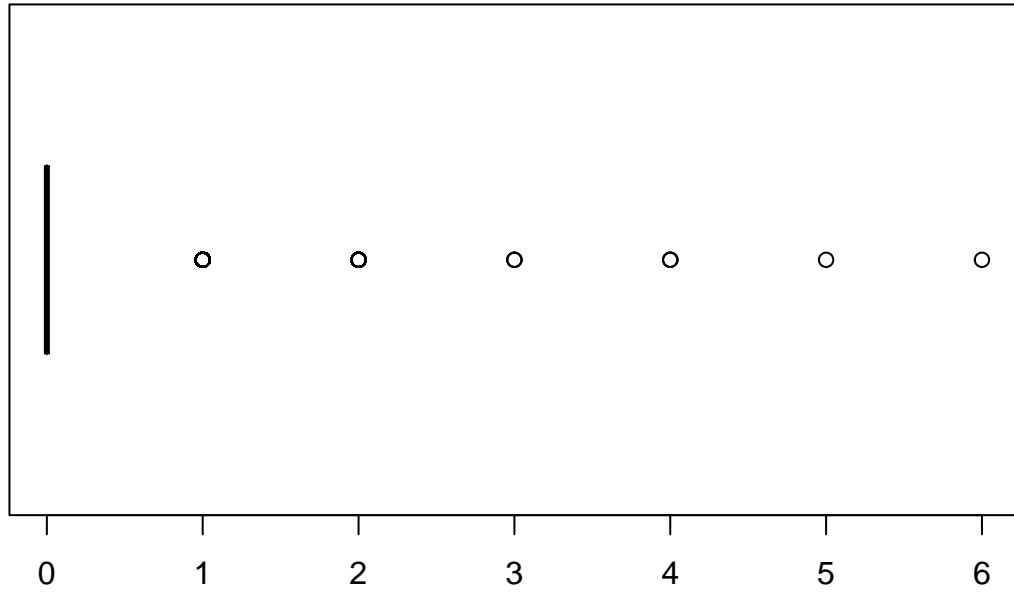
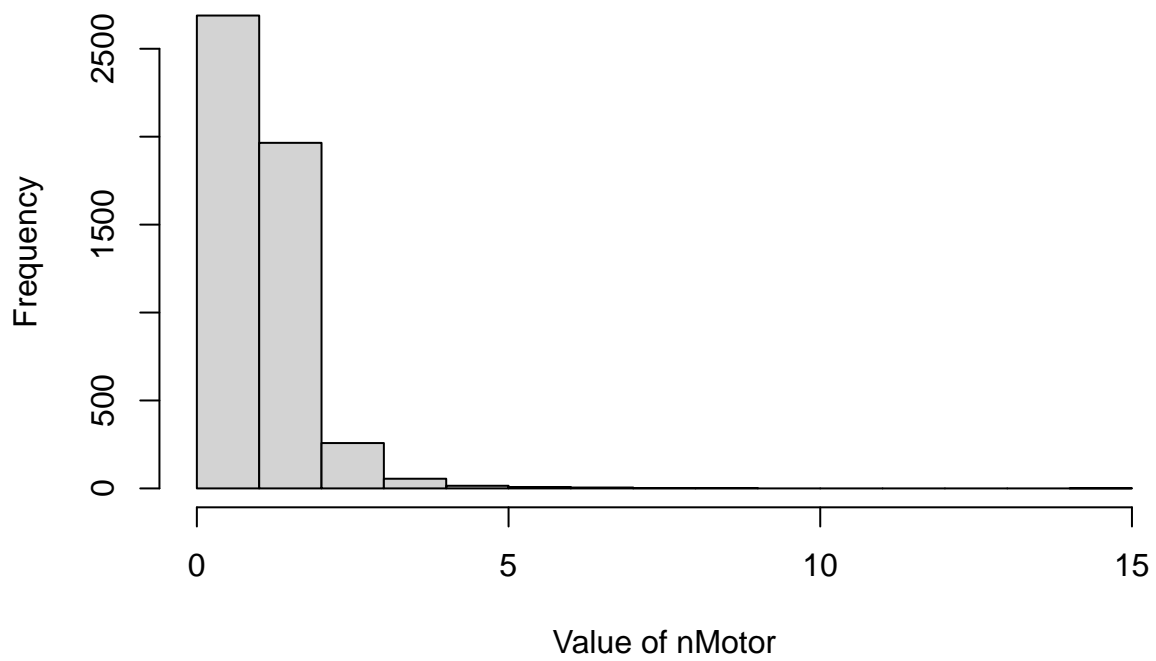


Table 12: nBikes extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	0	0	0.09	0	6	0.3351454	3.723838

Variable 11 : nMotor

Histogram of nMotor



Boxplot of nMotor

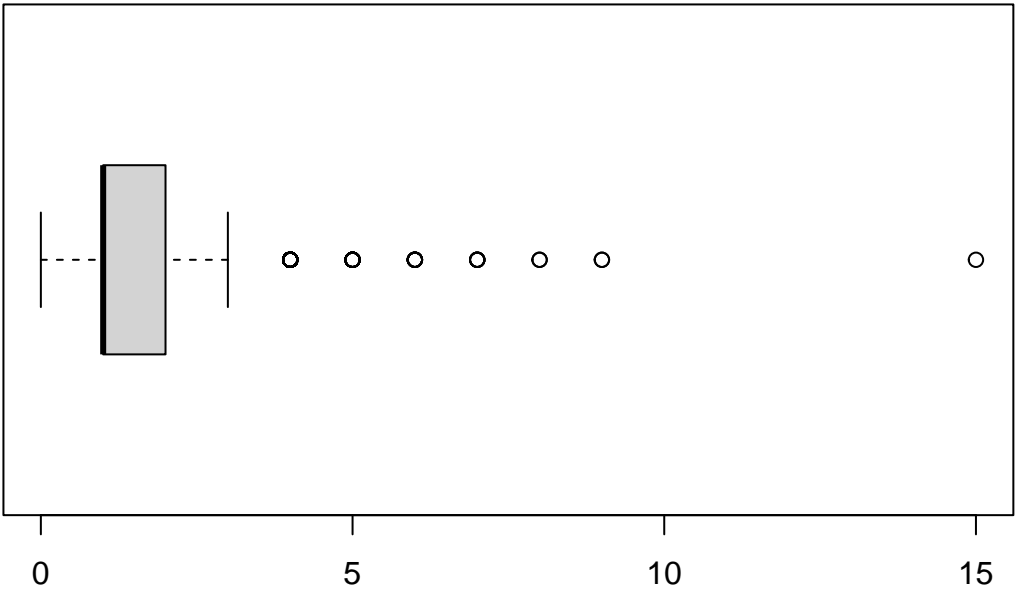
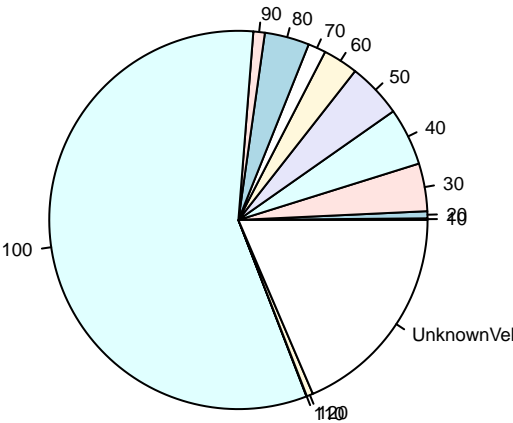


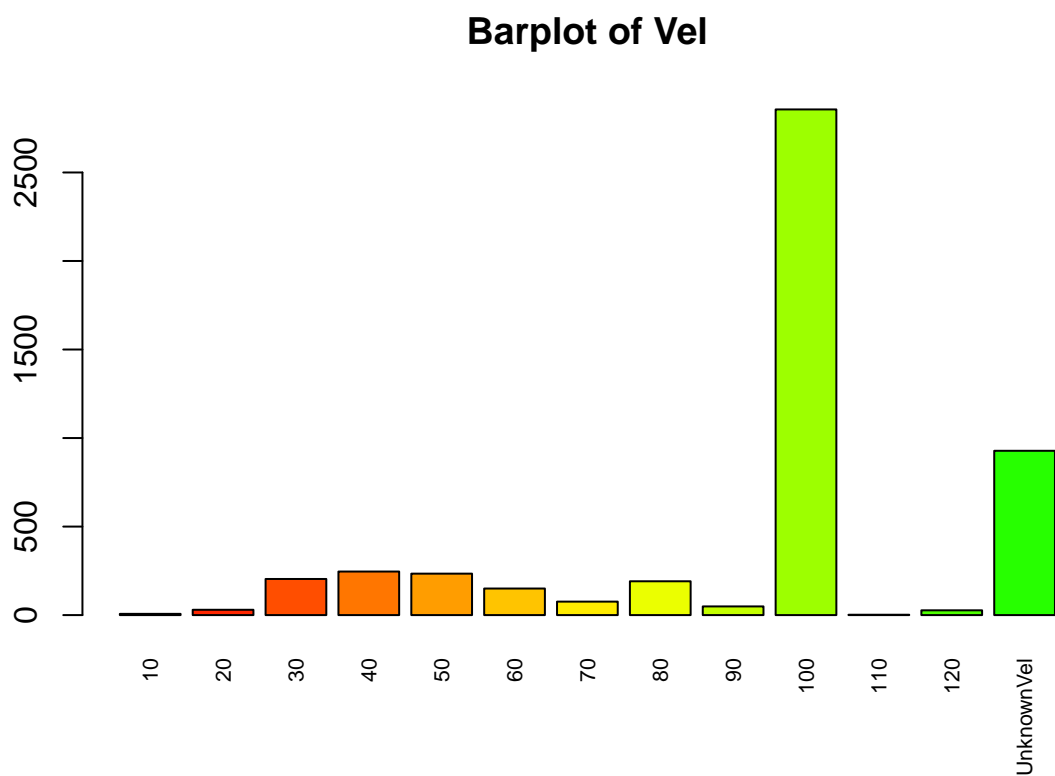
Table 13: nMotor extended Summary Statistics

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	sd	vc.
0	1	1	1.5214	2	15	0.8232117	0.5410883

Variable 12 : Vel

Pie of Vel





Number of modalities: 13

Table 14: Vel frequency table (sorted)

Vel	Frequency
100	2856
UnknownVel	928
40	246
50	234
30	204
80	191
60	150
70	76
90	49
20	30
120	27
10	7
110	2

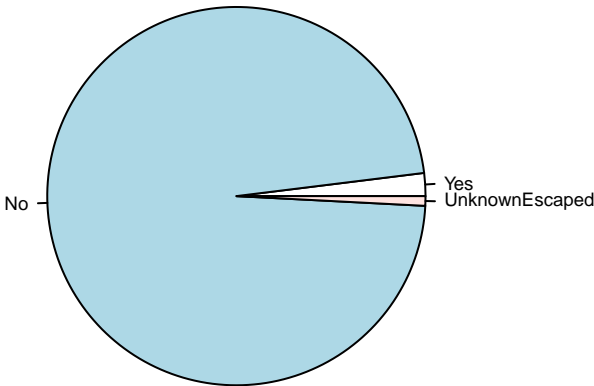
Table 15: Vel relative frequency table of proportions (sorted)

Vel	Proportion
100	0.5712
UnknownVel	0.1856

Vel	Proportion
40	0.0492
50	0.0468
30	0.0408
80	0.0382
60	0.0300
70	0.0152
90	0.0098
20	0.0060
120	0.0054
10	0.0014
110	0.0004

Variable 13 : Escaped

Pie of Escaped





Number of modalities: 3

Table 16: Escaped frequency table (sorted)

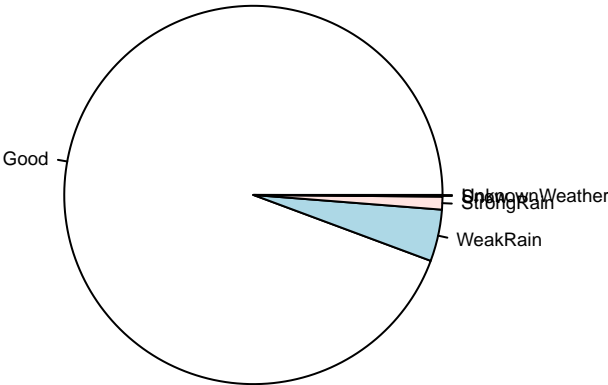
Escaped	Frequency
No	4861
Yes	98
UnknownEscaped	41

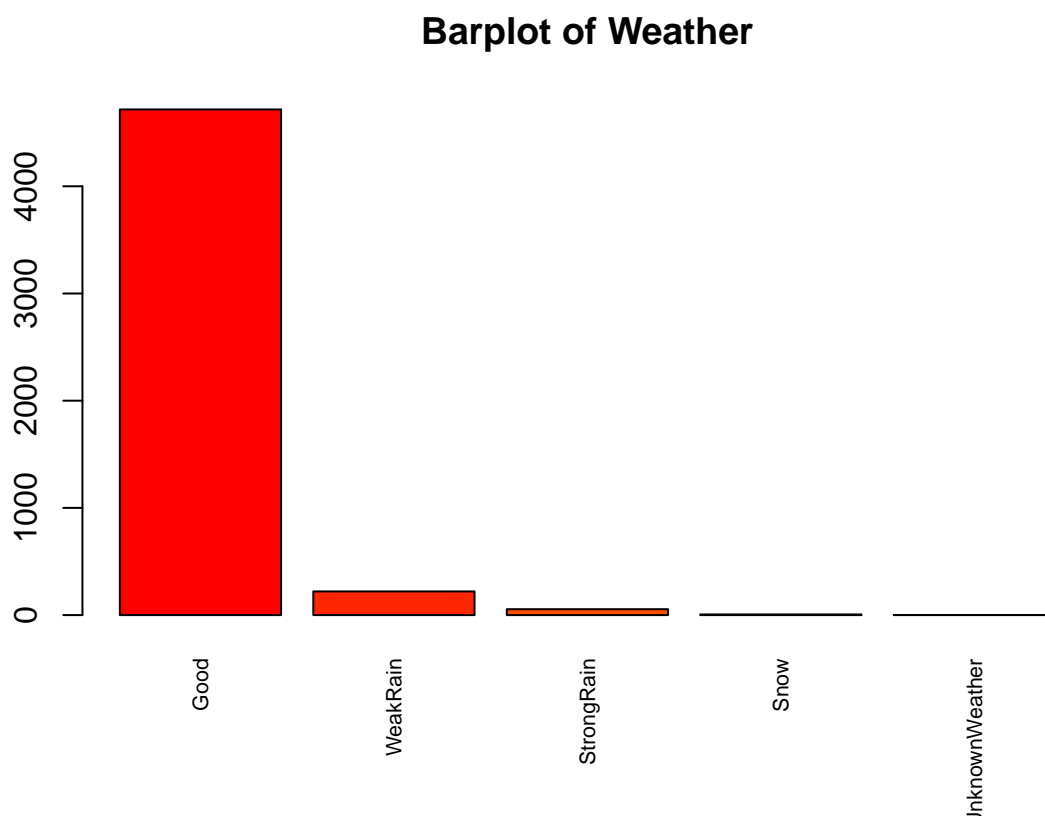
Table 17: Escaped relative frequency table of proportions (sorted)

Escaped	Proportion
No	0.9722
Yes	0.0196
UnknownEscaped	0.0082

Variable 14 : Weather

Pie of Weather





Number of modalities: 5

Table 18: Weather frequency table (sorted)

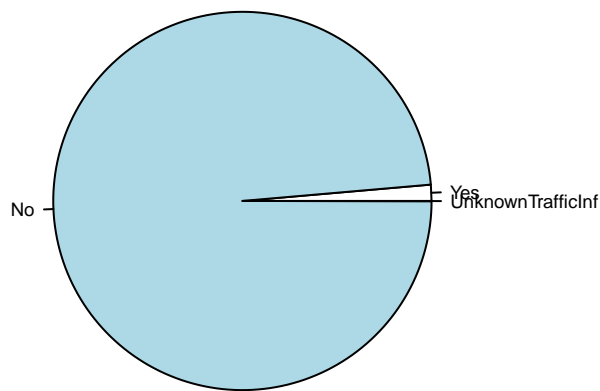
Weather	Frequency
Good	4717
WeakRain	221
StrongRain	55
Snow	6
UnknownWeather	1

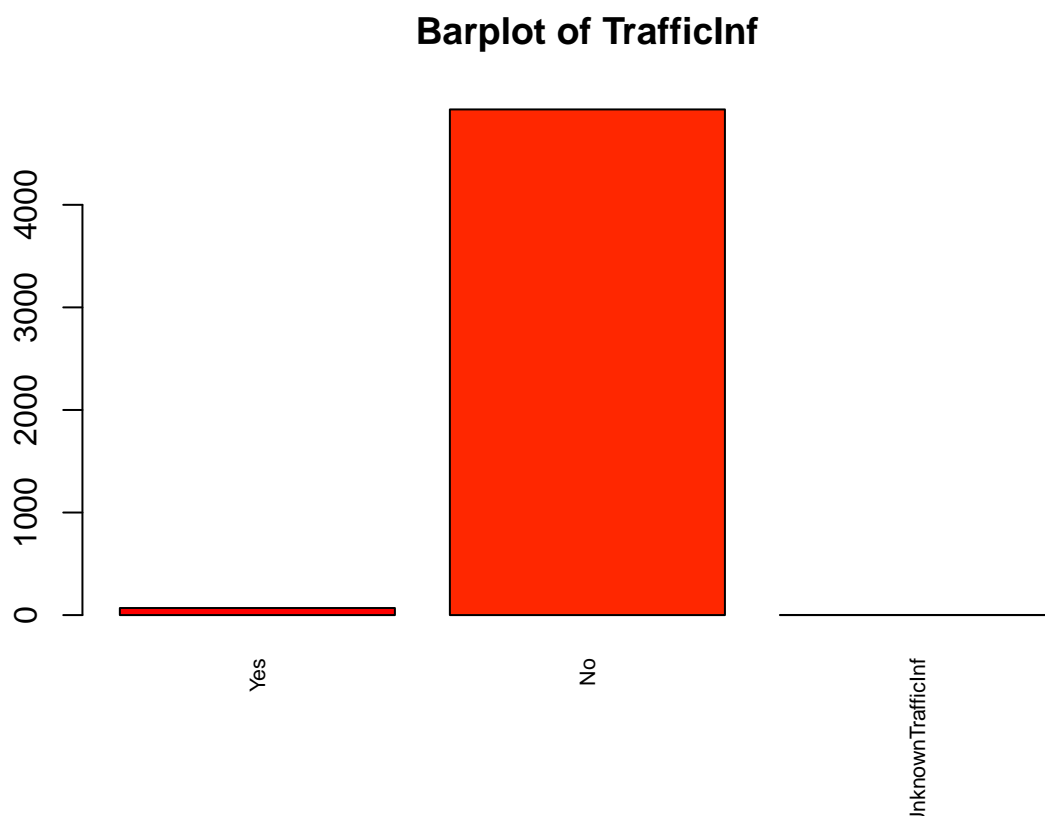
Table 19: Weather relative frequency table of proportions (sorted)

Weather	Proportion
Good	0.9434
WeakRain	0.0442
StrongRain	0.0110
Snow	0.0012
UnknownWeather	0.0002

Variable 15 : TrafficInf

Pie of TrafficInf





Number of modalities: 3

Table 20: TrafficInf frequency table (sorted)

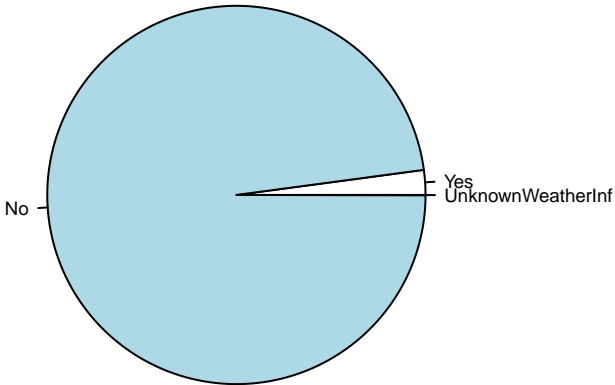
TrafficInf	Frequency
No	4930
Yes	69
UnknownTrafficInf	1

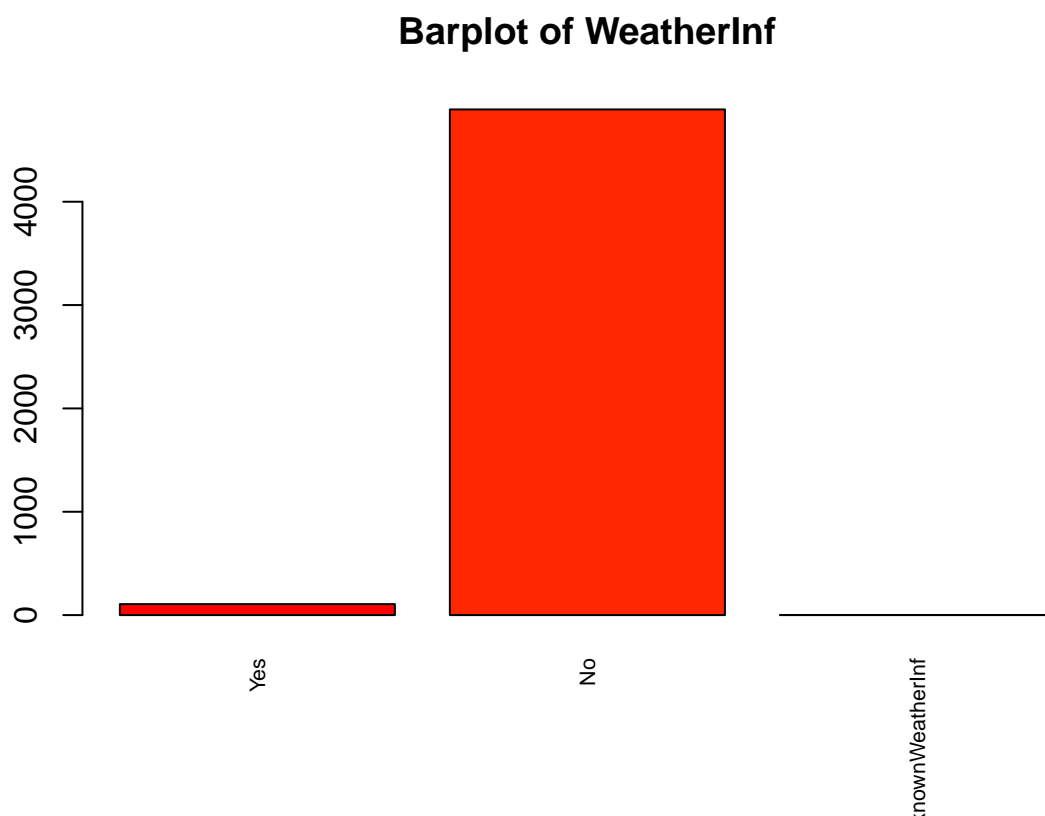
Table 21: TrafficInf relative frequency table of proportions (sorted)

TrafficInf	Proportion
No	0.9860
Yes	0.0138
UnknownTrafficInf	0.0002

Variable 16 : WeatherInf

Pie of WeatherInf





Number of modalities: 3

Table 22: WeatherInf frequency table (sorted)

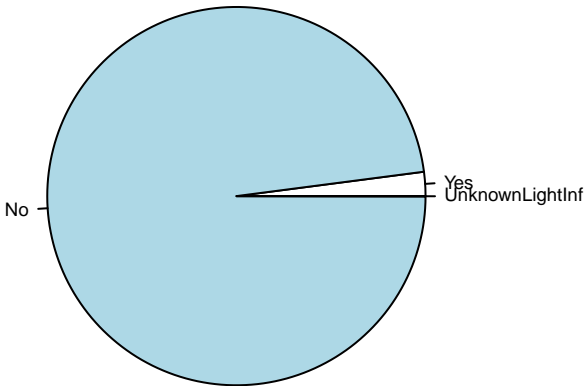
WeatherInf	Frequency
No	4893
Yes	106
UnknownWeatherInf	1

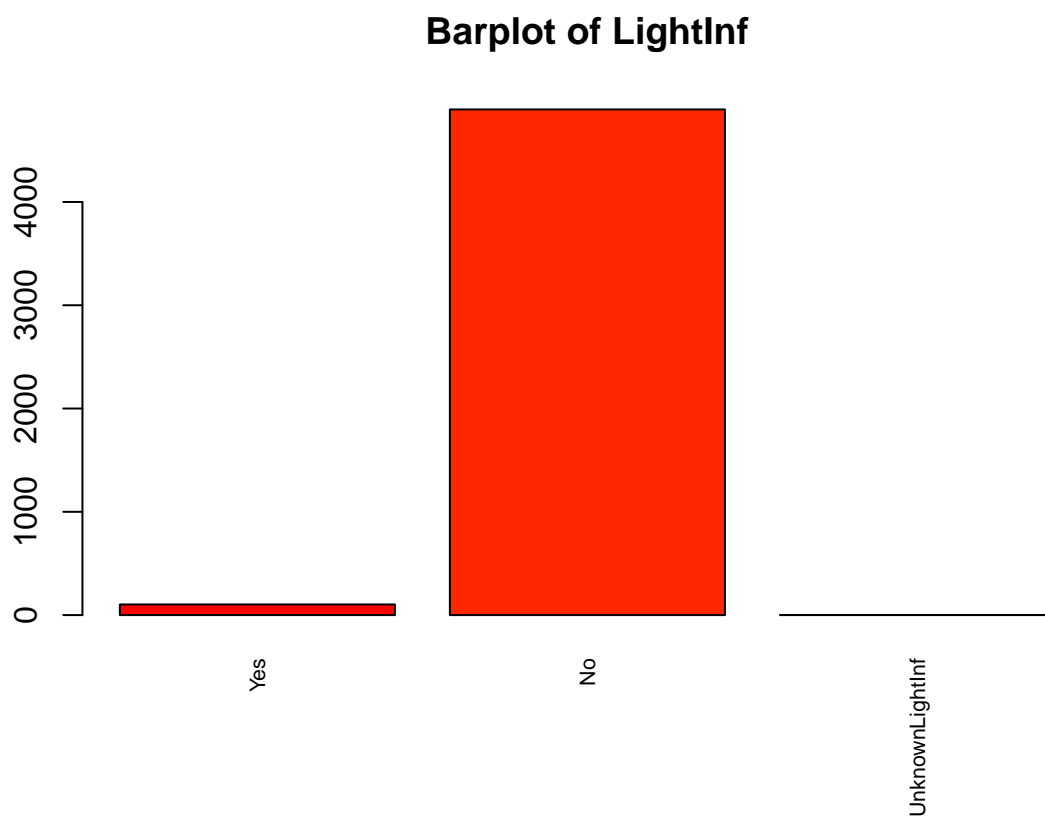
Table 23: WeatherInf relative frequency table of proportions (sorted)

WeatherInf	Proportion
No	0.9786
Yes	0.0212
UnknownWeatherInf	0.0002

Variable 17 : LightInf

Pie of LightInf





Number of modalities: 3

Table 24: LightInf frequency table (sorted)

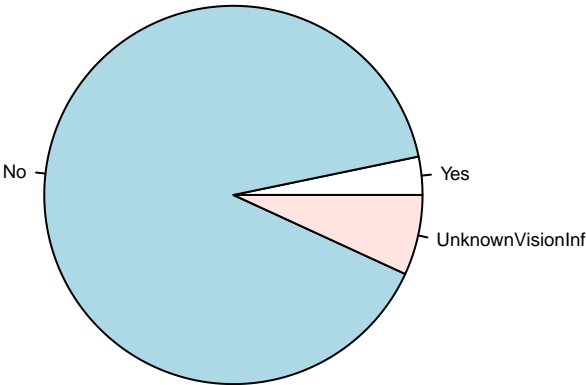
LightInf	Frequency
No	4896
Yes	103
UnknownLightInf	1

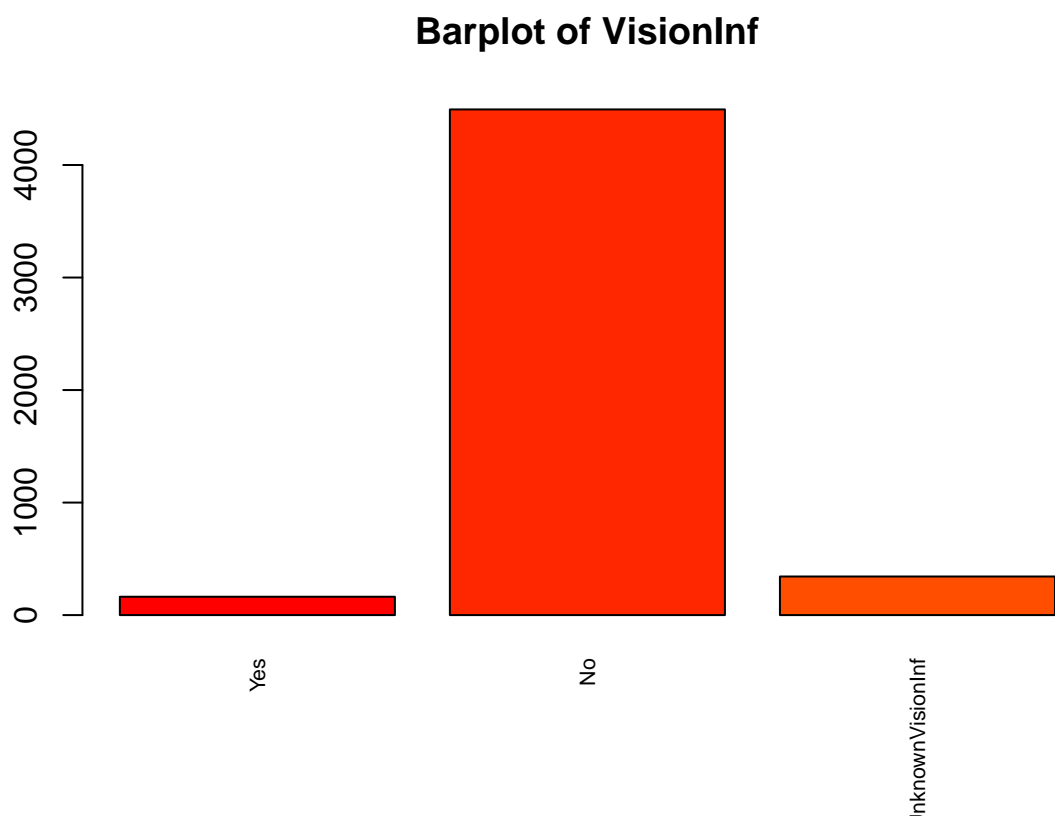
Table 25: LightInf relative frequency table of proportions (sorted)

LightInf	Proportion
No	0.9792
Yes	0.0206
UnknownLightInf	0.0002

Variable 18 : VisionInf

Pie of VisionInf





Number of modalities: 3

Table 26: VisionInf frequency table (sorted)

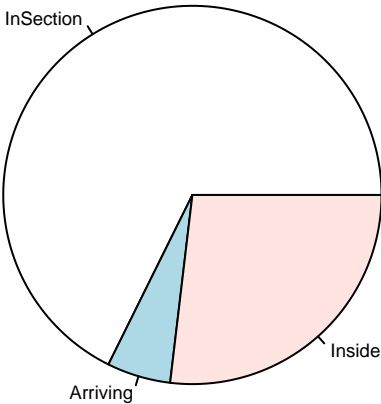
VisionInf	Frequency
No	4494
UnknownVisionInf	343
Yes	163

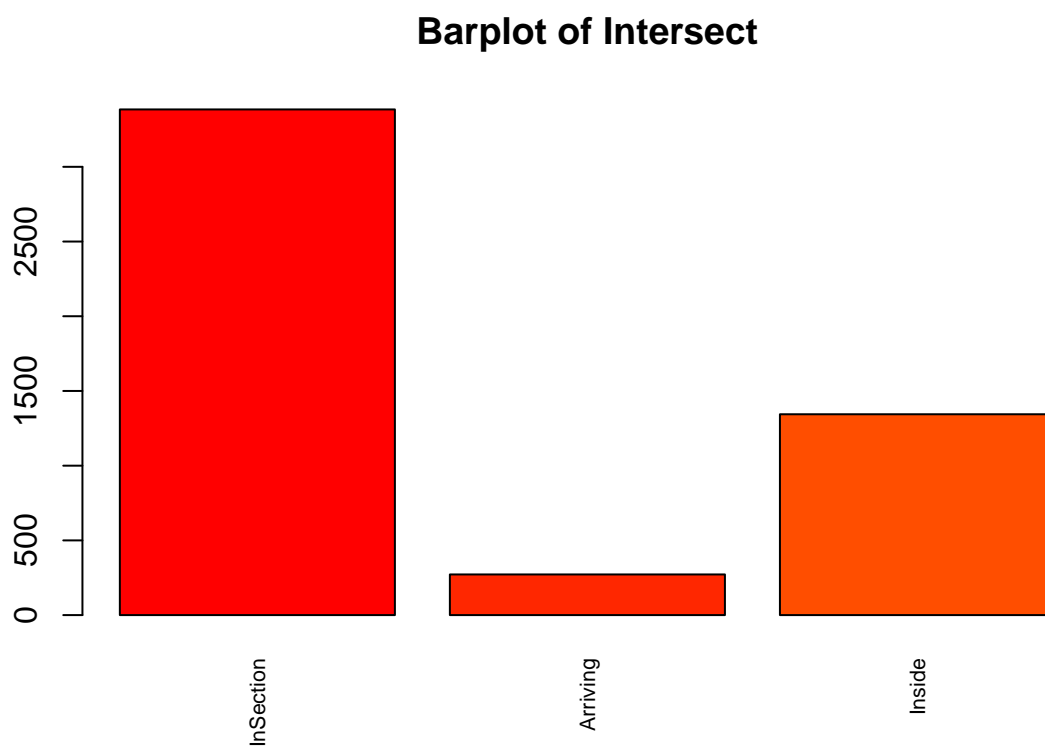
Table 27: VisionInf relative frequency table of proportions (sorted)

VisionInf	Proportion
No	0.8988
UnknownVisionInf	0.0686
Yes	0.0326

Variable 19 : Intersect

Pie of Intersect





Number of modalities: 3

Table 28: Intersect frequency table (sorted)

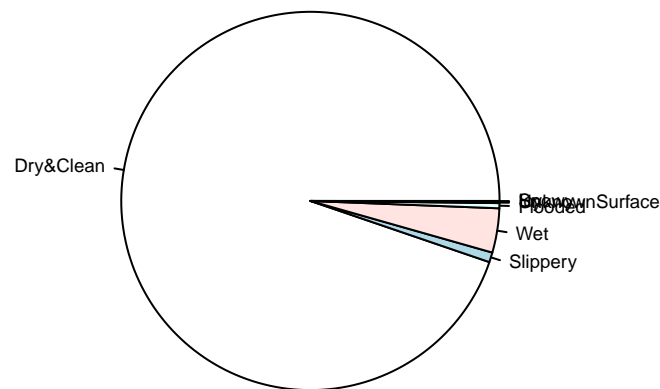
Intersect	Frequency
InSection	3384
Inside	1344
Arriving	272

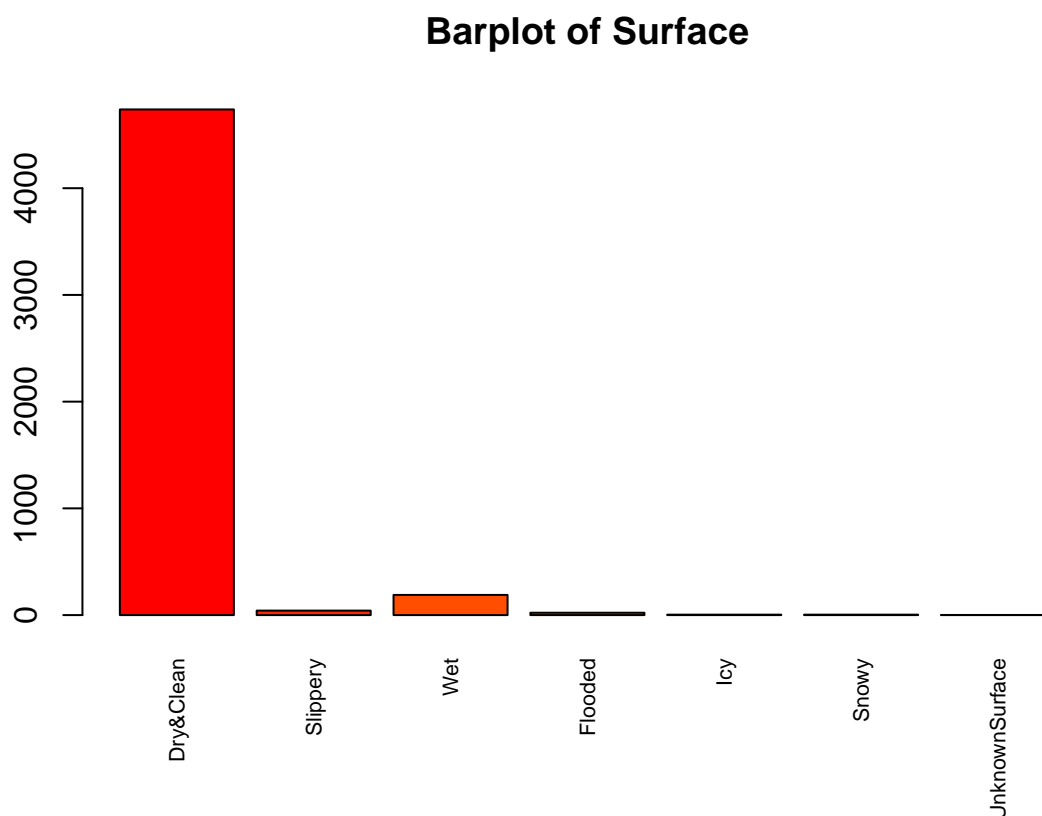
Table 29: Intersect relative frequency table of proportions (sorted)

Intersect	Proportion
InSection	0.6768
Inside	0.2688
Arriving	0.0544

Variable 20 : Surface

Pie of Surface





Number of modalities: 7

Table 30: Surface frequency table (sorted)

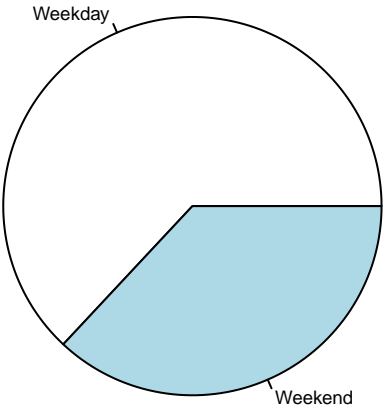
Surface	Frequency
Dry&Clean	4738
Wet	189
Slippery	42
Flooded	22
Icy	4
Snowy	4
UnknownSurface	1

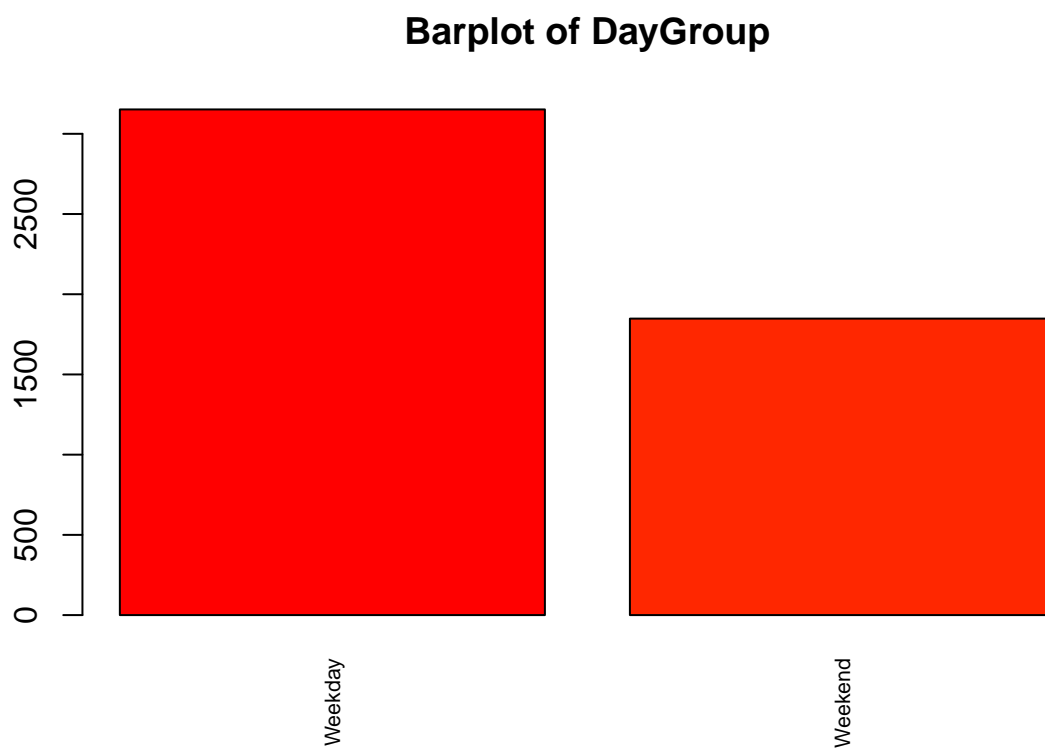
Table 31: Surface relative frequency table of proportions (sorted)

Surface	Proportion
Dry&Clean	0.9476
Wet	0.0378
Slippery	0.0084
Flooded	0.0044
Icy	0.0008
Snowy	0.0008
UnknownSurface	0.0002

Variable 21 : DayGroup

Pie of DayGroup





Number of modalities: 2

Table 32: DayGroup frequency table (sorted)

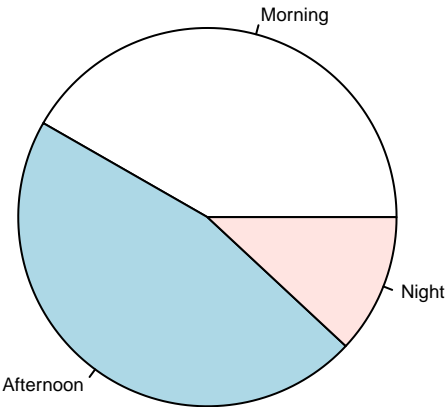
DayGroup	Frequency
Weekday	3152
Weekend	1848

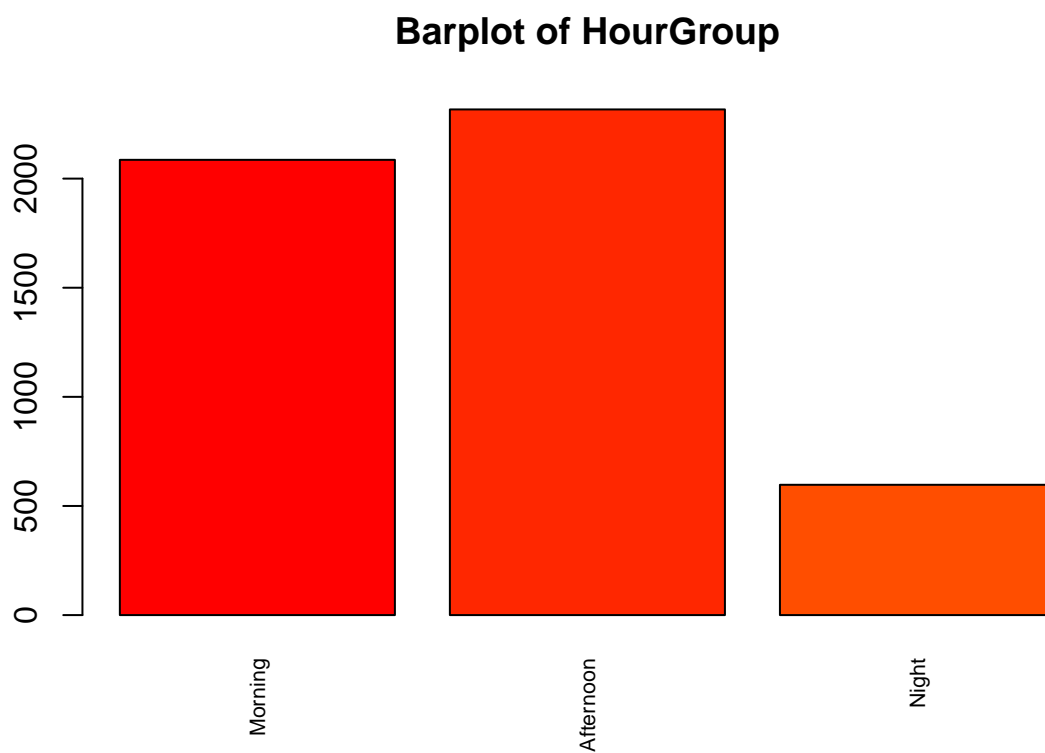
Table 33: DayGroup relative frequency table of proportions (sorted)

DayGroup	Proportion
Weekday	0.6304
Weekend	0.3696

Variable 22 : HourGroup

Pie of HourGroup





Number of modalities: 3

Table 34: HourGroup frequency table (sorted)

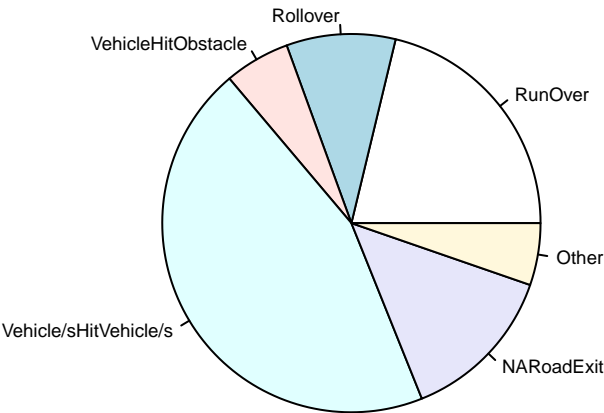
HourGroup	Frequency
Afternoon	2317
Morning	2086
Night	597

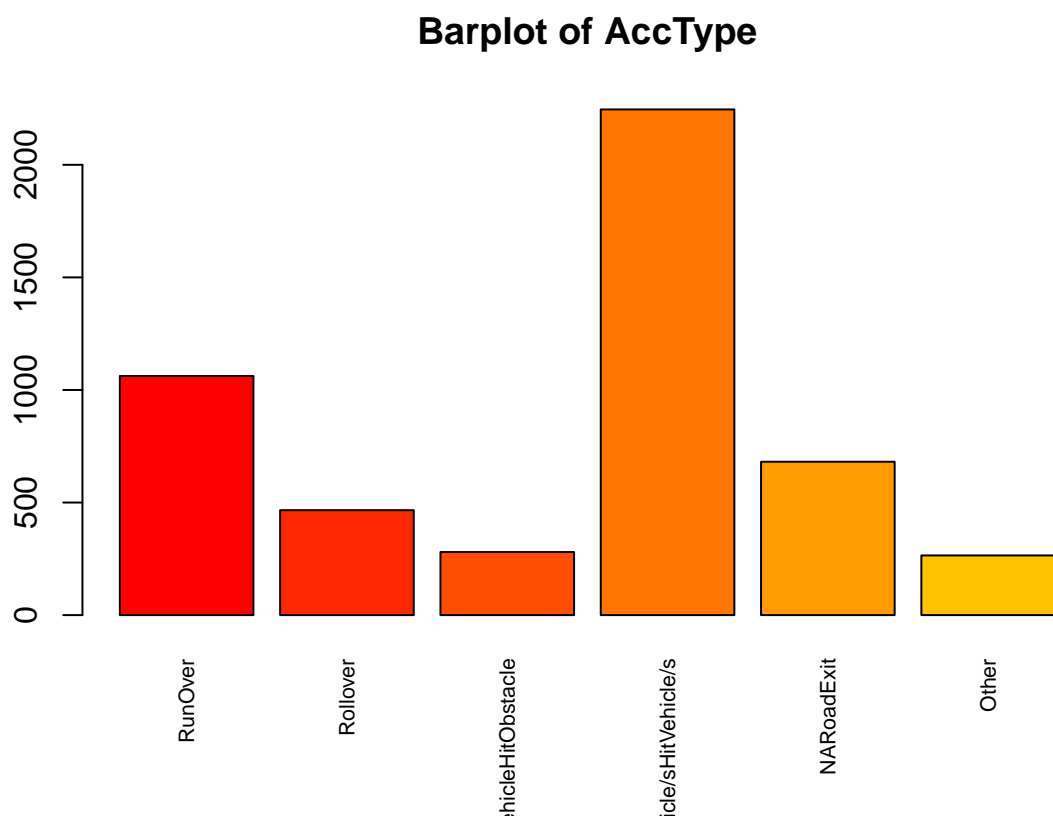
Table 35: HourGroup relative frequency table of proportions (sorted)

HourGroup	Proportion
Afternoon	0.4634
Morning	0.4172
Night	0.1194

Variable 23 : AccType

Pie of AccType





Number of modalities: 6

Table 36: AccType frequency table (sorted)

AccType	Frequency
Vehicle/sHitVehicle/s	2246
RunOver	1062
NARoadExit	681
Rollover	466
VehicleHitObstacle	280
Other	265

Table 37: AccType relative frequency table of proportions (sorted)

AccType	Proportion
Vehicle/sHitVehicle/s	0.4492
RunOver	0.2124
NARoadExit	0.1362
Rollover	0.0932
VehicleHitObstacle	0.0560
Other	0.0530