

D. MTO DRC Dataset: Crosstabs

Summary

MTODRC dataset is the MTO dataset with the added variable, `driver`, which has two values, `sober` or `drunk`. The database given has 26,026 records. Each record contains the variable `injury` which is the driver injury in one of four categories: `minimal`, `minor`, `major`, `killed`. Another variable, `prov`, indicates the license plate on the vehicle. The [frequency tabulation](#) indicates the number of accidents involving each value of the `prov` variable. After Ontario the next two are USA and “unknown” code = -1.

[Driver+Injury](#): Drunk drivers tend not to be involved in minimal injury accidents and are more likely to be involved in major injury or fatal accidents. Overall, for about 30% of all accidents, the driver is drunk. Overall about 48%, 40%, 10% and 2% of accidents are respectively minimal injury, minor injury, major injury and fatal.

[Driver+Wkgrp](#): As expected drunk drivers are relatively more common in the ThuSat weekgroup. Overall, about 43% and 57% of all accidents occur respectively in the SunWed and ThuSat weekgroups.

[Driver + Hour](#): Drunk drivers are more prevalent at 1AM and 2AM. Overall about 32%, 22%, 25% and 21% of accidents in the respective hour windows beginning at 11PM, 12AM, 1AM and 2AM.

[Driver + Province](#): USA and Quebec drivers involved in accidents tend to be relatively more sober. Drivers in the “unknown” vehicle province category tend to be relatively more drunk.

Frequencies: Province

After Ontario the next two are USA and “unknown” code = -1 .

| Vehicle Province or State | | | |
|---------------------------|-------------|---------|-------|
| | Frequency | Percent | |
| Valid | -1 | 230 | .9 |
| | Alberta | 38 | .1 |
| | British | 33 | .1 |
| | Columbia | | |
| | Manitoba | 38 | .1 |
| | New | 18 | .1 |
| | Brunswick | | |
| | Newfoundl | 5 | .0 |
| | and | | |
| | Nova | 19 | .1 |
| | Scotia | | |
| | Ontario | 25072 | 96.3 |
| | Prince | 3 | .0 |
| | Edward | | |
| | Island | | |
| | Quebec | 270 | 1.0 |
| | Saskatche | 13 | .0 |
| | wan | | |
| | Yukon and | 2 | .0 |
| | North | | |
| | West | | |
| | Territories | | |
| | U.S.A. | 284 | 1.1 |
| | Other | 1 | .0 |
| | Foreign | | |
| | Total | 26026 | 100.0 |

Driver + Injury

Drunk drivers tend not to be involved in minimal injury accidents and are more likely to be involved in major injury or fatal accidents.

Call:

```
crosstabs(formula = ~ driver + injury, data = mto.df, na.action =
na.exclude)
```

26026 cases in table

```
+-----+
|N      |
|N/RowTotal|
|N/ColTotal|
|N/Total |
+-----+
driver |injury
      |minimal|minor  |major  |killed |RowTotl|
-----+-----+-----+-----+-----+
sober  |9628   |6912   |1284   |197    |18021  |
      |0.53   |0.38   |0.071  |0.011  |0.69   |
      |0.77   |0.67   |0.51   |0.33   |        |
      |0.37   |0.27   |0.049  |0.0076 |        |
-----+-----+-----+-----+-----+
drunk  |2948   |3436   |1228   |393    |8005   |
      |0.37   |0.43   |0.15   |0.049  |0.31   |
      |0.23   |0.33   |0.49   |0.67   |        |
      |0.11   |0.13   |0.047  |0.015  |        |
-----+-----+-----+-----+-----+
ColTotl|12576  |10348  |2512   |590    |26026  |
      |0.48   |0.4    |0.097  |0.023  |        |
-----+-----+-----+-----+-----+
```

Test for independence of all factors

Chi² = 1088.853 d.f.= 3 (p=0)

Yates' correction not used

Chi-sq decomposition: (obs-exp)/sqrt(exp)

```
minimal minor  major killed
sober    9.86 -2.99 -10.92 -10.47
drunk   -14.79  4.49  16.38  15.70
```

Driver + Wkgrp

As expected drunk drivers are relatively more common in the ThuSat weekgroup.

Call:

```
crosstabs(formula = ~ driver + wkgrp, data = mto.df, na.action =
na.exclude)
```

26026 cases in table

```
+-----+
|N      |
|N/RowTotal|
|N/ColTotal|
|N/Total |
+-----+
driver |wkgrp
      |SunWed |ThuSat |RowTotl|
+-----+-----+-----+
sober  |8051   |9970   |18021  |
      |0.45   |0.55   |0.69   |
      |0.72   |0.67   |        |
      |0.31   |0.38   |        |
+-----+-----+-----+
drunk  |3191   |4814   |8005   |
      |0.4    |0.6    |0.31   |
      |0.28   |0.33   |        |
      |0.12   |0.18   |        |
+-----+-----+-----+
ColTotl|11242  |14784  |26026  |
      |0.43   |0.57   |        |
+-----+-----+-----+
```

Test for independence of all factors

Chi² = 52.3307 d.f. = 1 (p=4.689582e-013)

Yates' correction not used

Chi-sq decomposition: (obs-exp)/sqrt(exp)

```
SunWed ThuSat
sober   3.02  -2.64
drunk  -4.54   3.96
```

Driver + Hour

Drunk drivers are more prevalent at 1AM and 2AM.

Call:

```
crosstabs(formula = ~ driver + hour, data = mto.df, na.action =
na.exclude)
```

26026 cases in table

```
+-----+
|N      |
|N/RowTotal|
|N/ColTotal|
|N/Total |
+-----+
driver |hour
      |11PM  |12AM  |1AM   |2AM   |RowTotl|
+-----+-----+-----+-----+-----+
sober  |6581  |4192  |4069  |3179  |18021  |
      |0.37  |0.23  |0.23  |0.18  |0.69   |
      |0.8   |0.72  |0.62  |0.59  |       |
      |0.25  |0.16  |0.16  |0.12  |       |
+-----+-----+-----+-----+-----+
drunk  |1634  |1662  |2537  |2172  |8005   |
      |0.2   |0.21  |0.32  |0.27  |0.31   |
      |0.2   |0.28  |0.38  |0.41  |       |
      |0.063 |0.064 |0.097 |0.083 |       |
+-----+-----+-----+-----+-----+
ColTotl|8215  |5854  |6606  |5351  |26026  |
      |0.32  |0.22  |0.25  |0.21  |       |
+-----+-----+-----+-----+-----+
```

Test for independence of all factors

Chi^2 = 895.2281 d.f.= 3 (p=0)

Yates' correction not used

Chi-sq decomposition: (obs-exp)/sqrt(exp)

```
      11PM 12AM  1AM  2AM
sober 11.84 2.18 -7.47 -8.64
drunk -17.76 -3.27 11.21 12.97
```

Driver + Province

USA and Quebec drivers involved in accidents tend to be relatively more sober. Drivers in the “unknown” vehicle province category tend to be relatively more drunk.

Call:

```
crosstabs(formula = ~ prov + driver, data = mto.df, na.action = na.exclude)
```

26026 cases in table

```

+-----+
|N      |
|N/RowTotal|
|N/ColTotal|
|N/Total |
+-----+

```

| prov | driver | | RowTotal |
|----------|-------------------|---------------|---------------|
| | sober | drunk | |
| Al | 24 0.63 | 14 0.37 | 38 0.0015 |
| BC | 24 0.73 | 9 0.27 | 33 0.0013 |
| Man | 31 0.82 | 7 0.18 | 38 0.0015 |
| NB | 10 0.56 | 8 0.44 | 18 6.9e-4 |
| NWFLD | 4 0.8 | 1 0.2 | 5 1.9e-4 |
| NS | 10 0.53 | 9 0.47 | 19 7.3e-4 |
| Ont | 173397733 0.69 | 25072 0.31 | 0.96 |
| PEI | 1 0.33 | 2 0.67 | 3 1.2e-4 |
| Que | 212 0.79 | 58 0.21 | 270 0.01 |
| Sask | 11 0.85 | 2 0.15 | 13 5.e-4 |
| YNWT | 1 0.5 | 1 0.5 | 2 7.7e-5 |
| USA | 225 0.79 | 59 0.21 | 284 0.011 |
| Other | 0 0 | 1 1 | 1 3.8e-5 |
| Unknown | 129 0.56 | 101 0.44 | 230 0.0088 |
| ColTotal | 18021 0.69 | 8005 0.31 | 26026 |

Test for independence of all factors

Chi² = 56.71319 d.f.= 13 (p=2.017336e-007)

Yates' correction not used

Some expected values are less than 5, don't trust stated p-value

Chi-sq decomposition: (obs-exp)/sqrt(exp)

| | sober | drunk |
|---------|-------|-------|
| Al | -0.45 | 0.68 |
| BC | 0.24 | -0.36 |
| Man | 0.91 | -1.37 |
| NB | -0.70 | 1.05 |
| NWFLD | 0.29 | -0.43 |
| NS | -0.87 | 1.31 |
| Ont | -0.16 | 0.24 |
| PEI | -0.75 | 1.12 |
| Que | 1.83 | -2.75 |
| Sask | 0.67 | -1.00 |
| YNWT | -0.33 | 0.49 |
| USA | 2.02 | -3.03 |
| Other | -0.83 | 1.25 |
| Unknown | -2.40 | 3.60 |