MA678 Midterm Project Report

Pruthvi Bharadwaj

12/10/2021

Access the codes at : [github: https://github.com/pruthvibharadwaj93/MA-678-Midterm-Project---Accidents-Analysis]

Abstract

The global epidemic of road crash fatalities and disabilities is gradually being recognized as a major public health concern. Reducing traffic accidents is an important public safety challenge around the world. The first step to being informed about global road safety and to developing effective road safety interventions is to have access to facts. This project aims to build a model that can estimate the 3-year frequency of accidents in a county given its transportation profile. Transportation profile of a county includes variables like no.of commuters, no. of business establishments, no.of bridges, etc. A linear multilevel mixed effects model is used for modeling. The implications and limitations for choosing multilevel mixed effects model are discussed. This report offers only directional inferences to explain fluctuation in the number of accidents across states. Next steps and additional layers of complexity that can be incorporated in accident frequency prediction are also suggested towards the end.

Introduction

Background

More than 38,000 people die every year in crashes on U.S. roadways. The U.S. traffic fatality rate is 12.4 deaths per 100,000 inhabitants. An additional 4.4 million are injured seriously enough to require medical attention. Road crashes are the leading cause of death in the U.S. for people aged 1-54. The economic and societal impact of road crashes costs U.S. citizens $871 billion. Road crashes cost the U.S. more than $380 million in direct medical costs. The U.S. suffers the most road crash deaths of any high-income country, about 50% higher than similar countries in Western Europe, Canada, Australia and Japan.

The $1.2 trillion infrastructure bill signed into law by President Joe Biden in November 2021 contains much more than funding for the repair of roads and bridges. There's also money for study of motor vehicle crash causation as stated below

"The Secretary [of Transportation] shall carry out a comprehensive study (1) to determine the causes of, and contributing factors to, crashes that involve a commercial motor vehicle; and (2) to identify data requirements, data collection procedures, reports, and any other measures that can be used to improve the ability of States and the Secretary (A) to evaluate future crashes involving commercial motor vehicles; (B) to monitor crash trends and identify causes and contributing factors; and (C) to develop effective safety improvement policies and programs."

Thus, analysis of accidents and its prediction is important for optimizing public transportation, enabling safer routes, and cost-effectively improving the transportation infrastructure, all in order to make the roads safer. Given its significance, accident analysis and prediction has been a topic of much research in the past few decades. Analyzing the impact of environmental stimuli (e.g., road-network properties, weather, and traffic) on traffic accident occurrence patterns, predicting frequency of accidents within a geographical region, and predicting risk of accidents are the major related research categories.

Objective

In this project, I am interested in building a model that can estimate 3-year accident frequency of a county given its transportation profile. The model can also be used to understand the association of certain infrastructure related variables on accidents which can be helpful to identify areas that can be prioritised for infrastructure improvements.

The primary motive of this project is to understand the subtleties involved in statitsical modeling and not to answer any specific policy question with accuracy and precision. Hence, the inferences drawn here are only directional.

Method

Dataset

US-Accidents is a dataset that is available on Kaggle. It is a countrywide car accident dataset, which covers 49 states of the USA. It contains car accidents records from February 2016 to December 2020. It was collected using multiple APIs that provide streaming traffic incident (or event) data. Currently, there are about 1.5 million accident records in this dataset.

The dataset offers a wide range of data attributes to describe each accident record. It includes location data(coordinates, Geographical info like State, City, County), time data(timestamp of accident, day/night indicator), natural language description of event, weather data(Temperature, humidity, precipitation, etc), and relevant points-of-interest data (traffic signal, stop sign, etc.).

The other data sets were obtained from the following sources 1.County-level population related numbers were obtained from Census.gov and Economic research service websites 2.County transportation profiles were obtained from Bureau of Transportation Statistics website

Data Cleaning and Preparation

The entire dataset contains accident occurrences across 49 states in the US. During data exploration I found that data from 2016 and 2020 are incomplete. Some of the months were totally missing or had values which were very low compared to previous months. Thus I dropped these 2 years.

The raw data set contains 47 columns. For the scope of this project many columns are irrelevant. These are dropped directly.

The population and transportation profile numbers at a county level were all available in different csvs. All these were imported and stitched together to form one comprehensive dataframe with all the required variables.

After necessary cleaning,transformations and modifications, I combined the data from all these sources to form a final analytical dataset that was used for modeling. Variables used in the model are described in the table below

column names	explanation
n	no. of accidents from 2017-2019
Number.of.Bridges	Number of bridges in a county
poor.bridges	Number of poor bridges in a county

column names	explanation
Number.of.business.establishments	Number of business establishments in a County
res_commuters_dens	No.of commuters who are also residents of the county/ area of county
non_res_commuters_dens	No.of commuters who are not residents of the county/ area of county
trans_comm	Ratio of people who use transit to res_commuters
State	Indicates the state

The final dataset has 2464 records.

Exploratory Data Analysis

Bivariate plots of different combinations of available variables (with and without transformations) were plotted against each other(often split by State) and were analysed. Those that indicated a strong relationship were chosen for the model. A couple of charts are shown and the reason for choosing them is explained.

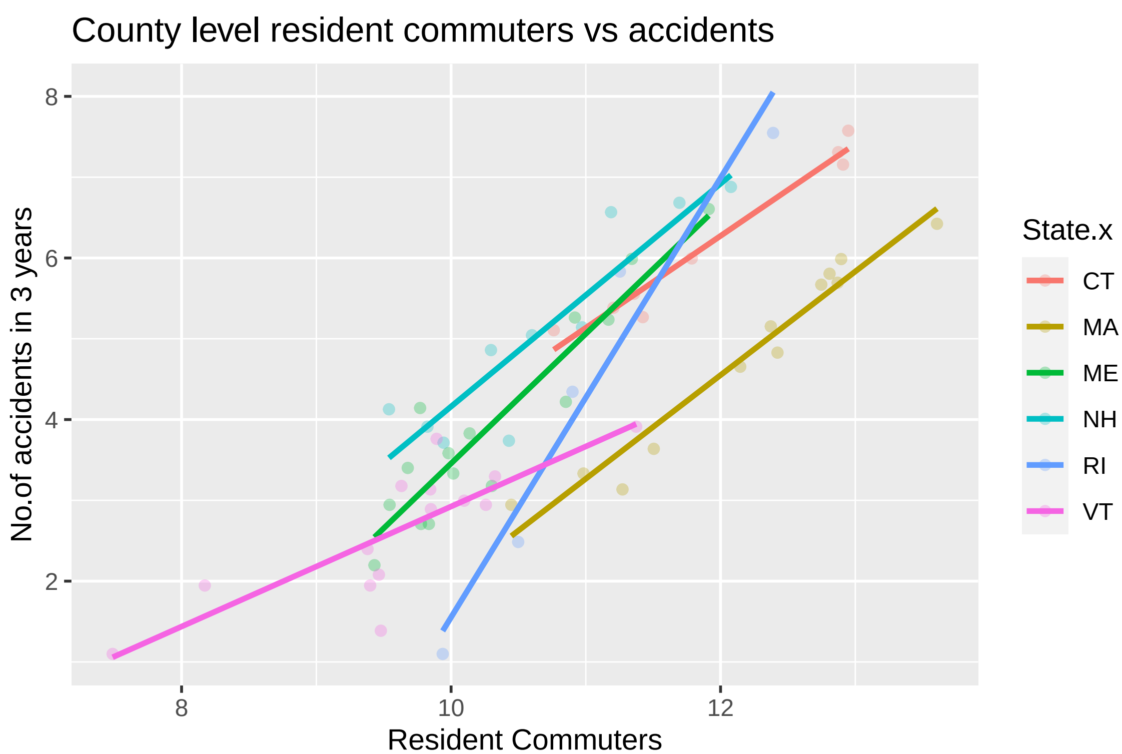

Figure 1: Scatterplot of Resident Commuters vs accidents in Counties of New England

Fig 1 shows a plot of log of accidents against log of resident commuters. Data for 6 states in New England are plotted and points are colored differently for the chosen states. It is observed that there is an approximate linear relationship at State level between logged values of accidents vs resident commuters. The number of accidents go up as the number of commuters goes up. This indicates that a State level model with varying intercept and varying slope model with resident commuter as a predictor would be suitable.

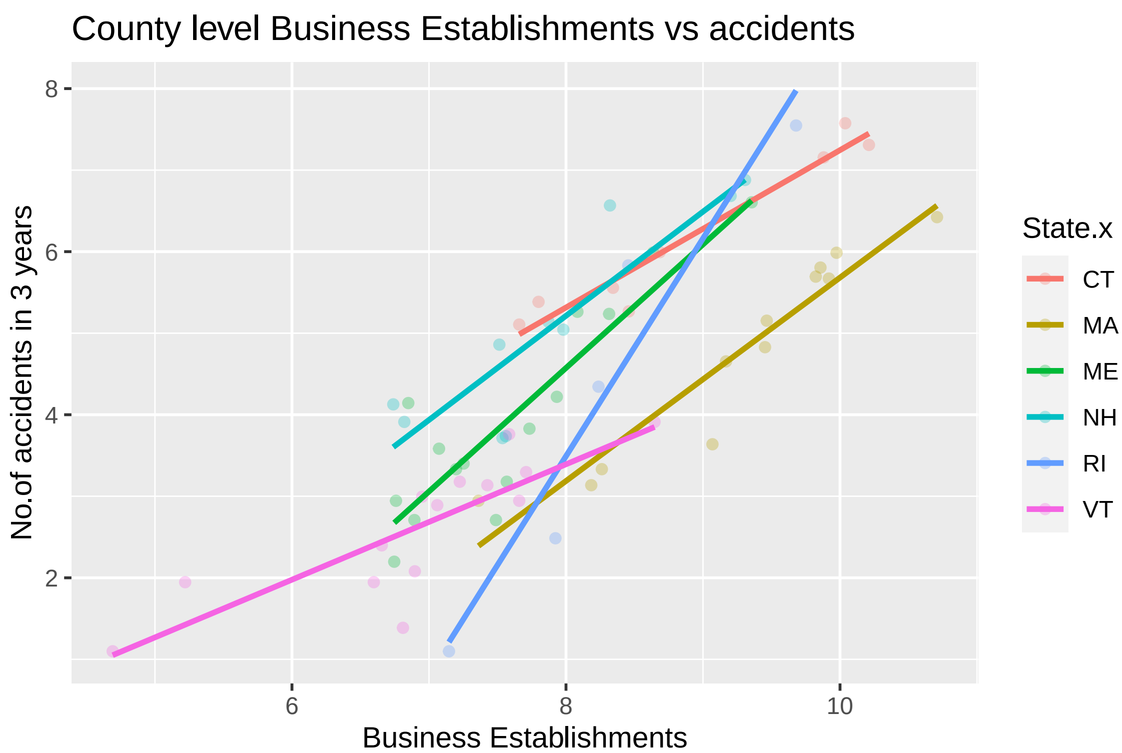

Fig 2 shows a plot of log of accidents against log of business establishments. Data for 6 states in New England are plotted and points are colored differently for the chosen states. It is observed that there is an approximate linear relationship at State level between logged values of accidents vs resident commuters. The number of accidents go up as the number of business establishments go up. This indicates that a State level model with varying intercept and varying slope model with business establishments as a predictor would be suitable.

Model Choice

We need to estimate the 3 year frequency of accidents at a county level. Thus, our outcome variable is chosen as n, the number of accidents from 2017-2019. From the EDA it was understood that 6 variables seem to have a strong relationship with number of accidents and that this relationship varies at a state level. So it would be appropriate to model for random effects at a state level for some of these variables. Most of the variables are log transformed while modeling. This is because of high and non-uniform range of these variables.

Model Fitting

The lmer() function was used to fit a multilevel linear model with a gaussian error distribution, varying intercepts and varying slopes at State level for a few variables. Below is the formula that was used to fit the model.

```
model <- lmer(data = final, log(n) ~ log(Number.of.Bridges)
              + log(poor.bridges+1)  + log(res_commuters) + trans_comm
              + log(Number.of.business.establishments)
             + log(non_res_commuters)
             +(1+log(res_commuters)|State.x)
             +(1+log(Number.of.business.establishments)|State.x)
```

```
            + (1+log(Number.of.Bridges)|State.x)
            + (1+log(non_res_commuters)|State.x))
```

Fixed effects from the above model are shown below, all variables are significant at alpha = 0.05 level.

	Estimate	Std. Error	df	t value	Pr(>|t|)
(Intercept)	-5.83337	0.55659	-54.37423	-10.48	1.16e-14 ***
log(Number.of.Bridges)	0.34112	0.07017	56.99982	4.861	9.53e-06 ***
log(poor.bridges + 1)	-0.04978	0.03065	2412.03928	-1.624	0.10451
log(res_commuters)	0.30802	0.09888	392.59254	3.115	0.00197 **
trans_comm	2.06822	0.68200	1758.15485	3.033	0.00246 **
log(business.estb)	0.21427	0.08165	173.41050	2.624	0.00946 **
log(non_res_comm)	0.33821	0.04324	67.23466	7.823	4.92e-11 ***

Result

Model Coefficients

Let us take the coefficients for Massachusetts and try to interpret each of the coefficients and its relationship with number of accidents.

$$log(n) = -7.48 + 0.29 \cdot log(Number.of.Bridges) - 0.05 \cdot log(poor.bridges + 1)$$
$$+0.34 \cdot log(rescommuters) + 2.07 \cdot transcomm$$
$$+0.16 \cdot log(Number.of.business.establishments) + 0.38 \cdot log(nonrescommuters)$$

The formula can be used for calculating the number of accidents in a county in Massachusetts given the variables listed in the formula.
1.Log(n) increases by 0.29 units if log(Number.of.Bridges) increases by one unit, keeping everything else constant. This also means that for a 1% increase in the Number.of.Bridges in a county, the number of accidents can be expected to increase by 0.29%.
2.The coefficient for poor.bridges is small and is slightly negative. It is also not significant at alpha = 0.05. However, the negative effect comes as a bit of a surprise as one would expect higher number of poor bridges to have a positive effect on accidents. But it could be that bad bridges are not used or are used very little and hence do not contribute to accidents.
3.Coefficient for log(res_commuters), log(Number.of.business.establishments) and log(non_res_commuters) all have a positive coeeficient and can be interpreted similar to the first coefficient.
4.The intercept is negative, meaning a county which has no bridges or business establishments or commuters is likely to have 0 accidents. This is totally not surprising. But this will not be the case as any county would have high non-zero values for the predictors and since they all have a positive coefficient, this would make up for the negative intercept and the prediction for accidents would be above 0 for most of the counties.
5.The coefficient for trans_comm, that is ratio of people who use transit to total commuters in a county is 2.07. Since this ratio is always below 1 changes in trans_comm value would be small and the effect of this variable is always small. The positive effect of this variable also comes as a surprise as an increase in people who take transit is expected to reduce vehicles on road and hence the number of accidents. But this effect is small and hence it should be okay.

Model Validation

From the Residual plots in Figure 3 we can see that the mean of residuals is almost 0. The residuals on the left side seem to have a pattern. This is because of the nature of the output where data is counts.

And for the Q-Q plot in Figure 3, all the dots in the middle are as expected on the normal line. But the dots tend to deviate from the normal line at the ends.

Figure 4 shows that the output of pp_check which is obtained by fitting the same model using stan that the posterior predictive checks using simulation match closely with the posterior predictive distribution.

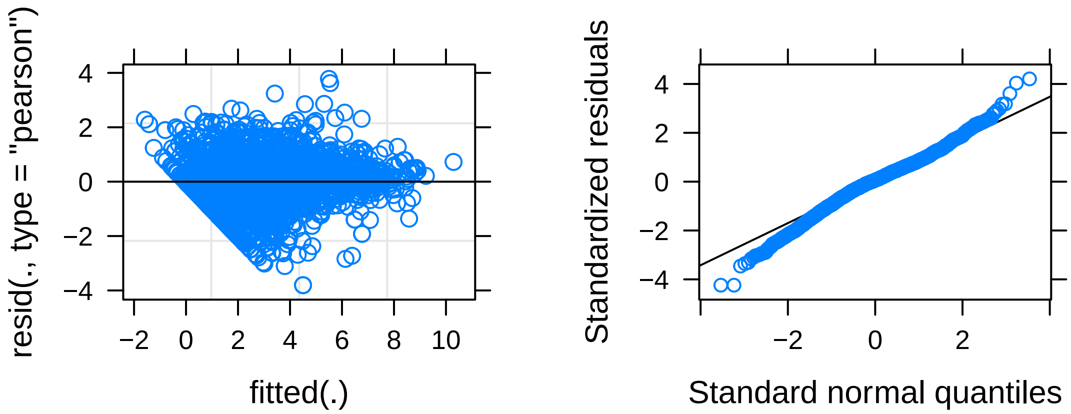

Figure 2: Residual plot and Q-Q plot.

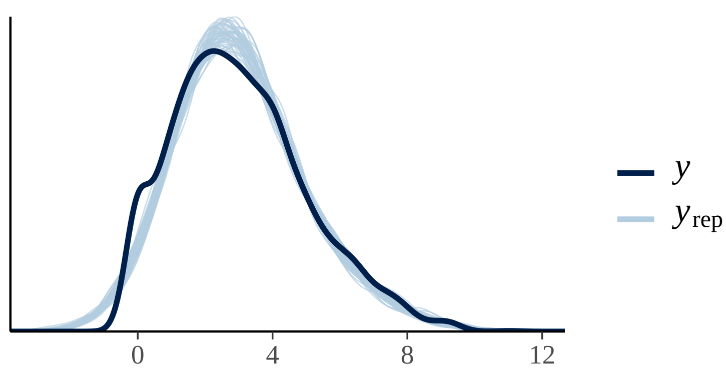

Figure 3: Posterior Predictive check

Discussion

The model gives reasonable estimates for the coefficients of different variables. As expected, the number of accidents increase as the commuters in a county increase. Random effects indicate that not all States are affected by the different variables in the same way. For example, bridges have the highest effect in West Virginia. This might indicate that accidents happen on bridges in West Virginia and bridges here would need to be prioritized when it comes to infrastructure repairs. More variables like length of roads, condition of roads, no.of registered automobiles, etc can be mined and incorporated into this model to make the estimations better. Additionally, the QQ plot of the residuals of this model seems a little off and a different model that would result in a more normal QQ plot can be explored.

Future Direction

This model is useful only to calculate the 3 year accident frequency in a county given transportation profiles. For the ultimate task of predicting accidents and preventing them if possible, this is only just a small directional step. A more sophisticated model that predicts the number of accidents at a county level on any given day. This would involve count regression models. Effects of weather like temperature, precipitation could be modeled and incorporated into a model that predicts daily accidents. Ultimately, spatial datasets with coordinates and road networks can be merged with previous accident records and comple models involving neural networks can be built to predict accidents and their locations.

Citations

https://www2.census.gov/programs-surveys/popest/datasets/2010-2020/counties/asrh/

https://www.ers.usda.gov/data-products/county-level-data-sets/download-data/

https://data.bts.gov/Research-and-Statistics/County-Transportation-Profiles/qdmf-cxm3/data

Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. "A Countrywide Traffic Accident Dataset.", arXiv preprint arXiv:1906.05409 (2019).

Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. "Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights." In proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2019.

Appendix

Full Results

Random effects of the model

```
## $State.x
##    (Intercept) log(res_commuters)  (Intercept)
## AL -1.13483966        0.094922742 -0.174203575
## AR  0.47470868       -0.039706622 -0.114277578
## AZ  0.85852338       -0.071810491  0.187180001
## CA  1.71289569       -0.143273769  0.502623264
## CO  2.52513775       -0.211213096  0.616279417
## CT -0.10276334        0.008595556  0.091237814
## DE -0.05495514        0.004596678  0.050350830
## FL -0.52570728        0.043972359 -0.094738731
## GA  1.29730713       -0.108512201  0.289298198
## IA -0.12265773        0.010259606 -0.070620522
## ID  1.43495833       -0.120025924  0.271602915
## IL -1.39499115        0.116682901 -0.290636113
## IN  0.61325984       -0.051295621  0.013264651
## KS  0.21161804       -0.017700619 -0.367400482
## KY -1.64880116        0.137912634 -0.253031422
## LA -2.38085149        0.199144389 -0.267283023
## MA -0.41041190        0.034328570 -0.206313401
## MD  1.62949592       -0.136297863  0.333221583
## ME -0.77800995        0.065076010 -0.007679323
## MI -0.93724583        0.078395166 -0.050233953
## MN -1.87158583        0.156547276  0.270793407
## MO -0.35924809        0.030049015 -0.217646736
## MS -1.39521599        0.116701708 -0.232127914
## MT  1.25708023       -0.105147455  0.127082137
## NC -0.49140923        0.041103526 -0.075719699
## ND  1.12914892       -0.094446744 -0.157234874
## NE -0.40151779        0.033584629 -0.001511477
## NH -0.11703401        0.009789215  0.145872558
## NJ  0.17648378       -0.014761842  0.096021767
## NM  1.23682676       -0.103453370 -0.114025766
## NV  0.72239067       -0.060423781  0.249527569
## NY -0.36416701        0.030460454 -0.217392438
## OH -1.25483369        0.104959545 -0.231821185
## OK -0.50780716        0.042475117 -0.371360602
## OR  2.93154635       -0.245206813  0.840086187
## PA  0.34427849       -0.028796895  0.026224664
## RI -0.84551351        0.070722291 -0.092170898
## SC -0.17911599        0.014982012 -0.057287379
## SD -0.07170981        0.005998109 -0.352201313
## TN -2.25337843        0.188482008 -0.317239168
## TX -1.16200860        0.097195265 -0.469268459
## UT  2.14585948       -0.179488673  0.871133669
## VA  0.83884918       -0.070164858  0.254012601
## VT -0.21321535        0.017834224 -0.099990419
## WA  0.80937740       -0.067699715 -0.001854099
## WI -0.02154318        0.001801962  0.106421504
```

```
## WV -1.62379501        0.135821015 -0.416569171
## WY  0.27458726       -0.022967628 -0.020395013
##    log(Number.of.business.establishments) (Intercept) log(Number.of.Bridges)
## AL                           -0.0999914533 -1.48527360            0.346640730
## AR                           -0.0961508903 -0.77167306            0.180087988
## AZ                            0.0002469796 -0.90965578            0.212329910
## CA                            0.1427663617 -0.02401432            0.005662341
## CO                           -0.0489453535 -1.28612524            0.300219607
## CT                            0.0492273028 -0.40651368            0.094896318
## DE                            0.0325722877 -0.28608813            0.066782980
## FL                            0.0806152479 -1.00275072            0.234053503
## GA                            0.0114056077 -0.83880801            0.195802383
## IA                           -0.0204206124 -0.01639177            0.003817796
## ID                           -0.0850038216 -0.29173978            0.068096982
## IL                            0.0944399755  1.76121189           -0.411077605
## IN                           -0.0336972989 -1.07969758            0.252001151
## KS                           -0.0790389442  0.66739933           -0.155811425
## KY                            0.0521262439  1.55429066           -0.362785946
## LA                            0.0971597990  1.83449809           -0.428180906
## MA                           -0.0501528524  0.22737378           -0.053092091
## MD                            0.0017492790 -0.02388833            0.005598000
## ME                            0.0894332317 -0.35245918            0.082279980
## MI                            0.1253204108  0.40781851           -0.095168801
## MN                            0.1734815008  1.07815669           -0.251599180
## MO                           -0.0116186232 -0.02181785            0.005076009
## MS                            0.0447420463  0.74456170           -0.173791671
## MT                           -0.1494472385 -0.10197561            0.023784820
## NC                            0.0306701279 -0.59835293            0.139658223
## ND                           -0.2307184922  1.66417983           -0.388475760
## NE                           -0.0304532659  2.07037253           -0.483238846
## NH                            0.0815055148 -0.53989579            0.126037307
## NJ                            0.0362839725 -0.04717615            0.011023573
## NM                           -0.1406010345 -0.27177572            0.063402396
## NV                           -0.0256881817  0.13597732           -0.031725470
## NY                            0.0175979775 -0.82903967            0.193489894
## OH                           -0.0097847472 -0.72089401            0.168242683
## OK                           -0.0661882141  0.98383990           -0.229668081
## OR                            0.1150586484 -0.46840781            0.109403337
## PA                           -0.0274141477  0.08090222           -0.018885790
## RI                            0.0374049839 -1.75120334            0.408738383
## SC                           -0.0431834284 -1.99194707            0.464917737
## SD                           -0.1632327119  1.54660183           -0.361034125
## TN                            0.0800471610  1.11945815           -0.261293774
## TX                           -0.0904572130  0.91358532           -0.213280934
## UT                            0.1200974003 -0.18525425            0.043317013
## VA                           -0.0038551792  1.10784020           -0.258558371
## VT                           -0.0176428146  0.92358942           -0.215579231
## WA                            0.0395730898 -0.30962393            0.072274055
## WI                            0.0158311541  0.01094390           -0.002544640
## WV                            0.0272691995 -2.48281675            0.579476829
## WY                           -0.0729389853  0.26265880           -0.061319282
##    (Intercept) log(non_res_commuters)
## AL -0.07622899            0.006359696
## AR -0.33674623            0.028321540
```

```
## AZ  1.26649835           -0.106688929
## CA  1.37583792           -0.115801622
## CO  0.95238223           -0.080176010
## CT  0.22489196           -0.018917632
## DE  0.09008656           -0.007571085
## FL  0.18134356           -0.015263942
## GA -0.32207046            0.027185245
## IA -1.13295350            0.095445955
## ID  0.98964147           -0.083380186
## IL -1.84273894            0.155263998
## IN -1.35441318            0.114113868
## KS -0.76249021            0.064164190
## KY -2.82794228            0.238268139
## LA -2.00654994            0.169071408
## MA -0.52459244            0.044153370
## MD  1.46539464           -0.123425971
## ME -0.53080875            0.044759145
## MI -1.72627591            0.145497934
## MN  0.20204982           -0.016919543
## MO -2.39487950            0.201758970
## MS -2.08912481            0.176014755
## MT  1.05089994           -0.088587580
## NC -2.43938660            0.205546150
## ND -0.24829563            0.020811565
## NE -0.66614192            0.056118283
## NH -0.01355145            0.001194162
## NJ  0.54024219           -0.045493608
## NM  1.94849445           -0.164252452
## NV  2.59859507           -0.218933620
## NY  0.11857893           -0.010017214
## OH -0.24866228            0.020914555
## OK -0.61620062            0.051841892
## OR  4.53463498           -0.381926302
## PA  0.54166918           -0.045648167
## RI -0.85564648            0.072098280
## SC -0.36493473            0.030725122
## SD  0.71148998           -0.060064698
## TN -0.16372550            0.013778821
## TX  0.87136229           -0.073525707
## UT  2.38007049           -0.200373447
## VA  1.17567819           -0.099027943
## VT -0.10510000            0.008834398
## WA -0.08900475            0.007514273
## WI -0.72925170            0.061469413
## WV  0.52070824           -0.043927028
## WY  0.72716637           -0.061302442
##
## with conditional variances for "State.x"
```

Fixed effects of the model

```
##                        (Intercept)                log(Number.of.Bridges)
##                         -5.8333694                             0.3411162
##                 log(poor.bridges + 1)                      log(res_commuters)
```

```
##                            -0.0497798                             0.3080244
##                            trans_comm log(Number.of.business.establishments)
##                             2.0682166                             0.2142685
##               log(non_res_commuters)
##                             0.3382125
```

Coefficients of the model

```
## $State.x
##     (Intercept) log(Number.of.Bridges) log(poor.bridges + 1) log(res_commuters)
## AL -10.37272800             0.68775689            -0.0497798         0.40294716
## AR  -3.93453462             0.52120415            -0.0497798         0.26831779
## AZ  -2.39927584             0.55344607            -0.0497798         0.23621392
## CA   1.01821341             0.34677850            -0.0497798         0.16475065
## CO   4.26718166             0.64133576            -0.0497798         0.09681132
## CT  -6.24442270             0.43601248            -0.0497798         0.31661997
## DE  -6.05318992             0.40789914            -0.0497798         0.31262109
## FL  -7.93619849             0.57516966            -0.0497798         0.35199677
## GA  -0.64414083             0.53691854            -0.0497798         0.19951221
## IA  -6.32400026             0.34493395            -0.0497798         0.31828402
## ID  -0.09353605             0.40921314            -0.0497798         0.18799849
## IL -11.41333395            -0.06996145            -0.0497798         0.42470732
## IN  -3.38033000             0.59311731            -0.0497798         0.25672879
## KS  -4.98689719             0.18530473            -0.0497798         0.29032380
## KY -12.42857401            -0.02166979            -0.0497798         0.44593705
## LA -15.35677531            -0.08706475            -0.0497798         0.50716880
## MA  -7.47501697             0.28802407            -0.0497798         0.34235299
## MD   0.68461433             0.34671416            -0.0497798         0.17172655
## ME  -8.94540915             0.42339614            -0.0497798         0.37310043
## MI  -9.58235268             0.24594736            -0.0497798         0.38641958
## MN -13.31971267             0.08951698            -0.0497798         0.46457169
## MO  -7.27036171             0.34619217            -0.0497798         0.33807343
## MS -11.41423333             0.16732449            -0.0497798         0.42472612
## MT  -0.80504842             0.36490098            -0.0497798         0.20287696
## NC  -7.79900626             0.48077438            -0.0497798         0.34912794
## ND  -1.31677368            -0.04735960            -0.0497798         0.21357767
## NE  -7.43944050            -0.14212269            -0.0497798         0.34160904
## NH  -6.30150539             0.46715346            -0.0497798         0.31781363
## NJ  -5.12743424             0.35213973            -0.0497798         0.29326257
## NM  -0.88606231             0.40451855            -0.0497798         0.20457105
## NV  -2.94380667             0.30939069            -0.0497798         0.24760063
## NY  -7.29003741             0.53460605            -0.0497798         0.33848487
## OH -10.85270411             0.50935884            -0.0497798         0.41298396
## OK  -7.86459800             0.11144808            -0.0497798         0.35049953
## OR   5.89281604             0.45051949            -0.0497798         0.06281760
## PA  -4.45625538             0.32223037            -0.0497798         0.27922752
## RI  -9.21542340             0.74985454            -0.0497798         0.37874671
## SC  -6.54983332             0.80603389            -0.0497798         0.32300643
## SD  -6.12020859            -0.01991797            -0.0497798         0.31402252
## TN -14.84688306             0.07982238            -0.0497798         0.49650642
## TX -10.48140375             0.12783522            -0.0497798         0.40521968
## UT   2.75006856             0.38443317            -0.0497798         0.12853574
## VA  -2.47797263             0.08255779            -0.0497798         0.23785956
## VT  -6.68623074             0.12553693            -0.0497798         0.32585864
```

```
## WA  -2.59585975            0.41339021           -0.0497798         0.24032470
## WI  -5.91954208            0.33857152           -0.0497798         0.30982638
## WV -12.32854940            0.92059299           -0.0497798         0.44384543
## WY  -4.73502031            0.27979688           -0.0497798         0.28505679
##    trans_comm log(Number.of.business.establishments) log(non_res_commuters)
## AL   2.068217                             0.11427704             0.34457219
## AR   2.068217                             0.11811761             0.36653403
## AZ   2.068217                             0.21451548             0.23152357
## CA   2.068217                             0.35703486             0.22241087
## CO   2.068217                             0.16532314             0.25803648
## CT   2.068217                             0.26349580             0.31929486
## DE   2.068217                             0.24684078             0.33064141
## FL   2.068217                             0.29488374             0.32294855
## GA   2.068217                             0.22567410             0.36539774
## IA   2.068217                             0.19384788             0.43365845
## ID   2.068217                             0.12926468             0.25483231
## IL   2.068217                             0.30870847             0.49347649
## IN   2.068217                             0.18057120             0.45232636
## KS   2.068217                             0.13522955             0.40237669
## KY   2.068217                             0.26639474             0.57648063
## LA   2.068217                             0.31142830             0.50728390
## MA   2.068217                             0.16411564             0.38236587
## MD   2.068217                             0.21601778             0.21478652
## ME   2.068217                             0.30370173             0.38297164
## MI   2.068217                             0.33958891             0.48371043
## MN   2.068217                             0.38775000             0.32129295
## MO   2.068217                             0.20264987             0.53997147
## MS   2.068217                             0.25901054             0.51422725
## MT   2.068217                             0.06482126             0.24962491
## NC   2.068217                             0.24493862             0.54375864
## ND   2.068217                            -0.01645000             0.35902406
## NE   2.068217                             0.18381523             0.39433078
## NH   2.068217                             0.29577401             0.33940666
## NJ   2.068217                             0.25055247             0.29271889
## NM   2.068217                             0.07366746             0.17396004
## NV   2.068217                             0.18858032             0.11927888
## NY   2.068217                             0.23186647             0.32819528
## OH   2.068217                             0.20448375             0.35912705
## OK   2.068217                             0.14808028             0.39005439
## OR   2.068217                             0.32932715            -0.04371381
## PA   2.068217                             0.18685435             0.29256433
## RI   2.068217                             0.25167348             0.41031078
## SC   2.068217                             0.17108507             0.36893762
## SD   2.068217                             0.05103579             0.27814780
## TN   2.068217                             0.29431566             0.35199132
## TX   2.068217                             0.12381128             0.26468679
## UT   2.068217                             0.33436590             0.13783905
## VA   2.068217                             0.21041332             0.23918455
## VT   2.068217                             0.19662568             0.34704689
## WA   2.068217                             0.25384159             0.34572677
## WI   2.068217                             0.23009965             0.39968191
## WV   2.068217                             0.24153770             0.29428547
## WY   2.068217                             0.14132951             0.27691005
##
```

```
## attr(,"class")
## [1] "coef.mer"
```

More EDA

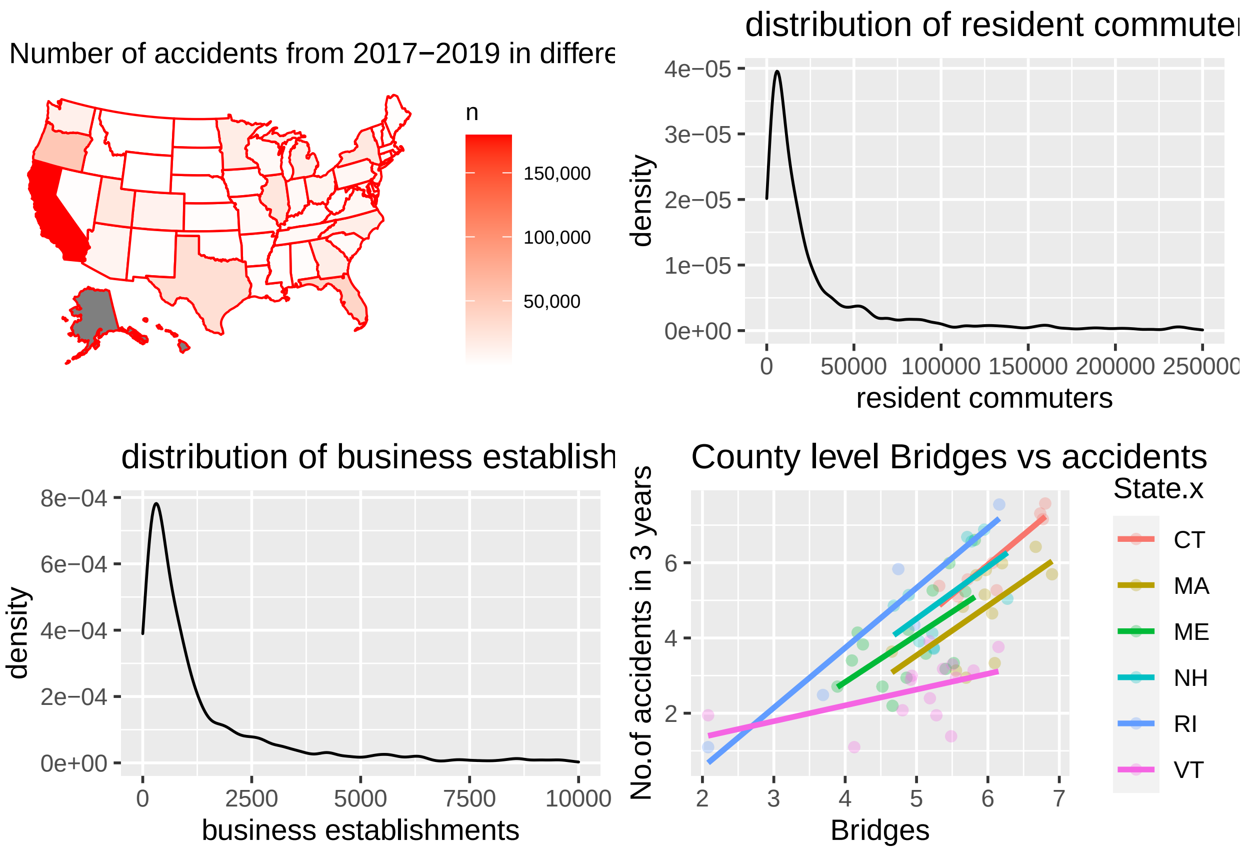

Figure 4: Other EDA plots