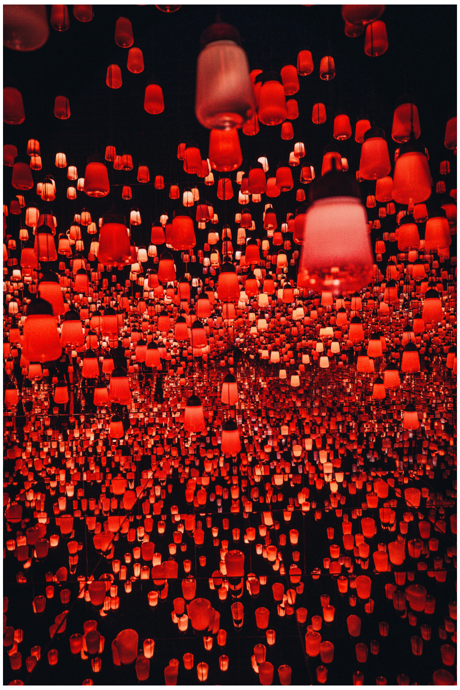

Assignment 4 ggplot2

Siyi Li
29018102

Yusen Wang
27496538

Alexi Tsakis
29682525

Report for
Australian Government

25 May 2021

Our consultancy
Alexi &
Siyi &
Yusen

(03) 9905 2478
questions@company.com

ABN: 12 377 614 630

Table 1: *Drivers killed before and after law*

law	DriversKilled_Average
0	125.8698
1	100.2609

Table 2: *Back seat passengers killed before and after law*

law	Rear_Average
0	400.3195
1	407.7391

{Section 1}

####Compulsory Seatbelts 1983

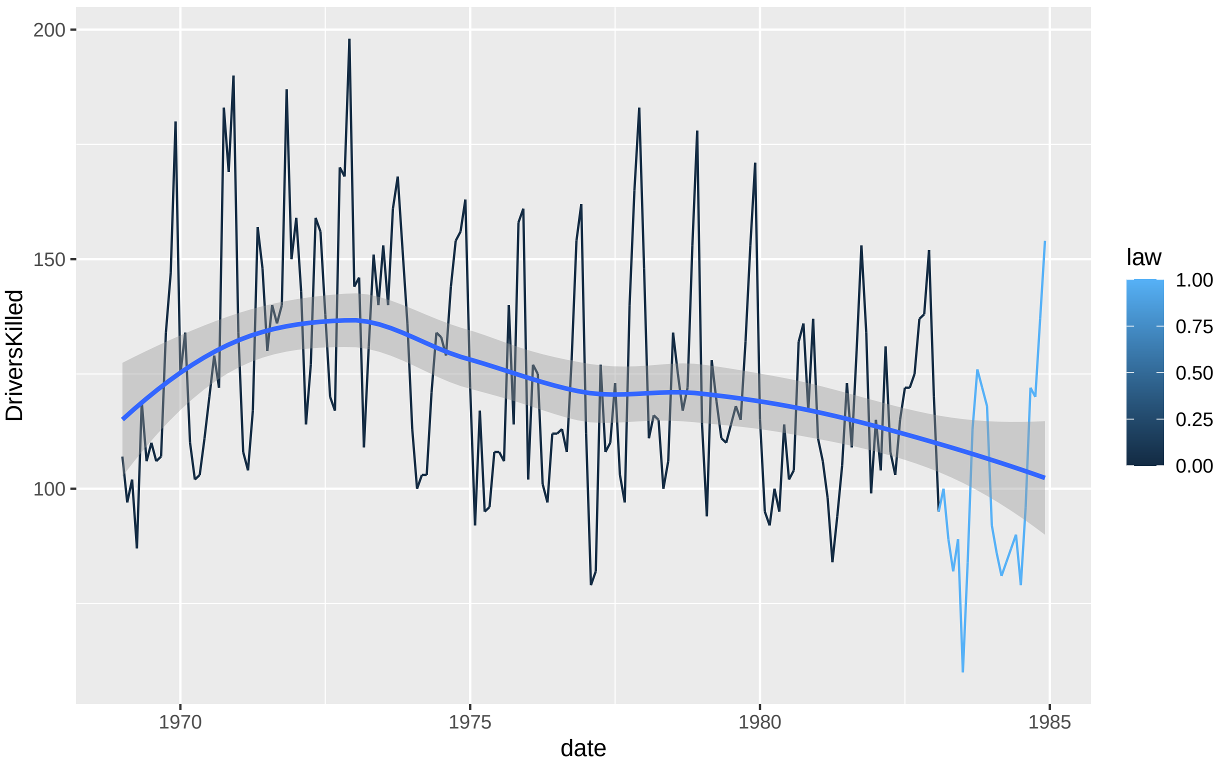

Figure 1: *Drivers killed in accidents over time*

In this section we look at the effect of seatbelts on driver deaths in the UK. Front seatbelts were compulsory on all new cars registered in the UK from 1968, however it was not required for them to be worn until 1983. In this section, we will answer the question - did mandatory wearing of frontrow seatbelts prevent deaths for drivers and passengers?

First, to answer if any relationships shown are causation and not simply correlation, we ask - why do seatbelts prevent deaths? According to the National Highway Traffic Safety Administration, the primary reasons why seatbelts prevent death is that “Buckling up helps keep you safe and secure

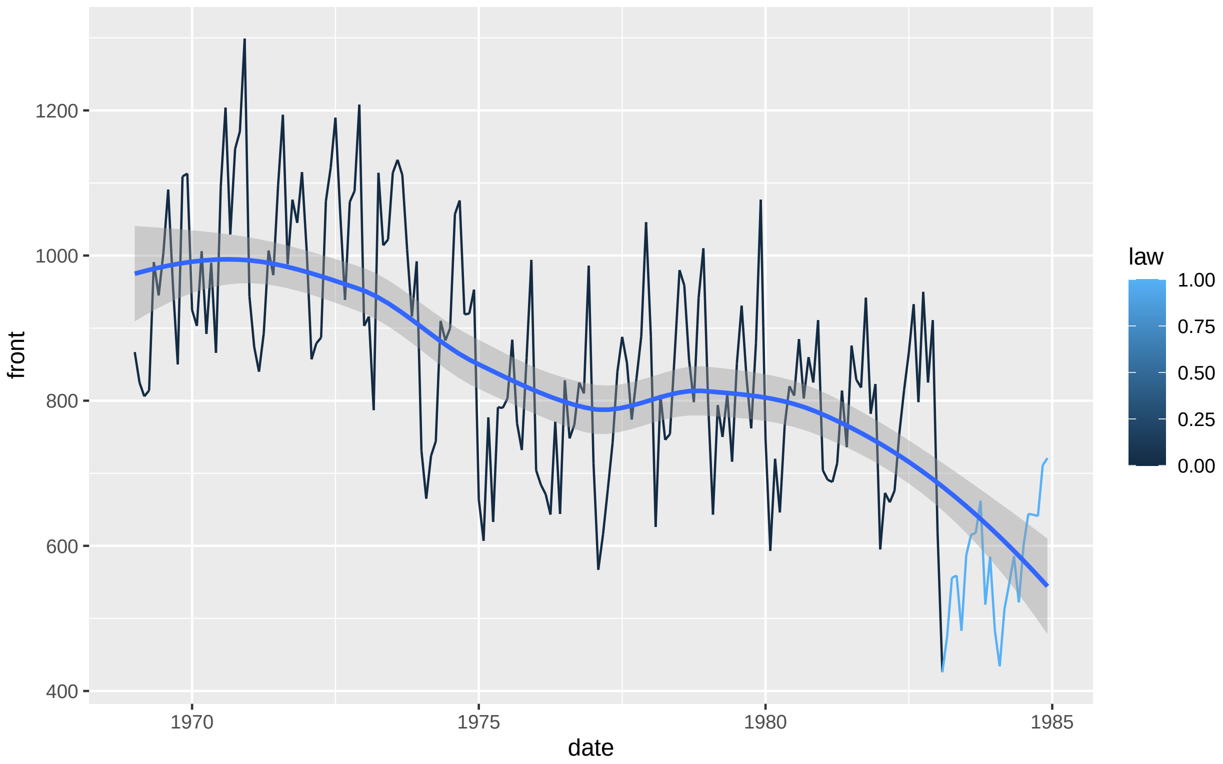

Figure 2: *Frontseat passengers killed in accidents over time*

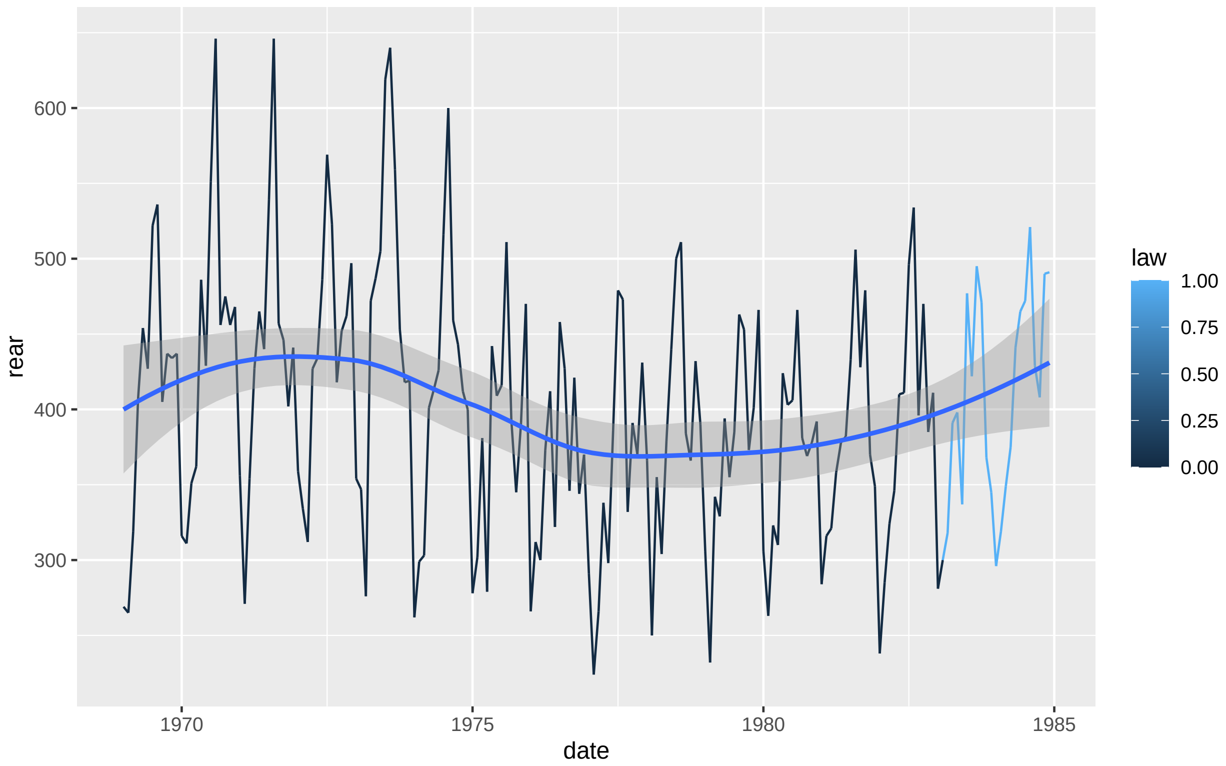

Figure 3: *Backseat passengers killed in accidents over time*

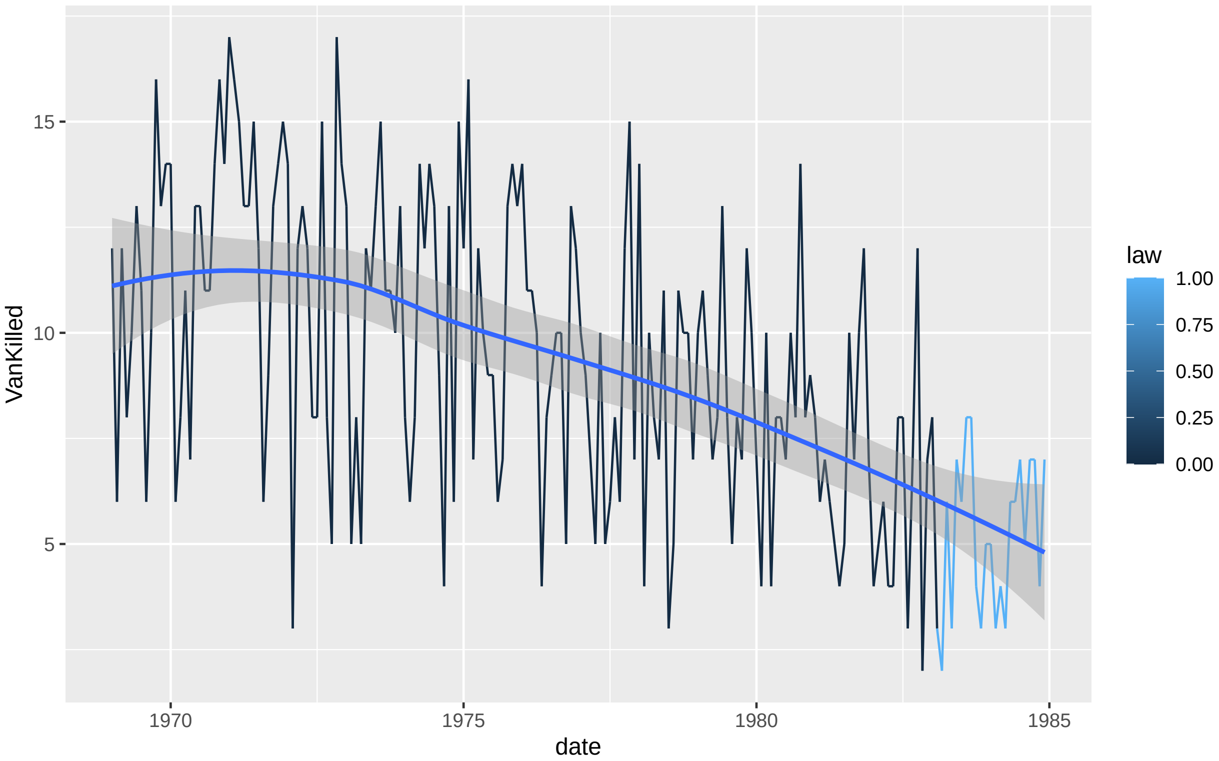

Figure 4: *Light goods vehicle drivers killed in accidents over time*

Table 3: *Front seat passengers killed before and after law*

law	Front_Average
0	873.4556
1	570.9565

inside your vehicle, whereas not buckling up can result in being totally ejected from the vehicle in a crash, which is almost always deadly." (NHTSA 2021)

These scenarios are referred to as "secondary collisions" - i.e. a crash event that results from the impact of an initial collison, such as ejection from a vehicle, or even a high speed impact with objects inside the vehicle.

Now that we understand the theory behind the impact of seatbelts on preventing death, let us look at the data:

Table 4: *Light good vehicle drivers killed before and after law*

law	Van_Average
0	9.585799
1	5.173913

There is a clear negative correlation between frontrow passengers and light good vehicle drivers killed since the introduction of the law (figure 2, figure 4), data from the introduction of the law shown in blue).

There is also a clear reduction in number of drivers killed (figure 1) since the introduction of compulsory seatbelts - the lack of change since the introduction of the law, but the large change since approximately 1973 can be described as " drivers that are least likely to use seat belts might be those that are more likely to be involved in an accident" (Cohen, A. and Einav. L 2003) - i.e. that careful drivers adopted seatbelts already without the law, and those likely to get involved in accidents would be unlikely to wear seatbelts even with the law in place.

There is little change in number of rear seat passengers killed (figure 3) - this is because the regulations and law were only for front seat seatbelts, and not backseat seatbelts, suggesting the total number of accidents didn't change.

We can also see this in tabular format, so it is easier to see the effect of the law numerically (table 1, table 2, table 3, table 4).

{Section 2}

0.1 Can we build an model among these variables?

```
##
## Call:
## lm(formula = DriversKilled ~ ., data = Seatbelt)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -29.489  -7.620  -0.531   7.676  34.700
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.330e+01  1.472e+01  -1.583    0.115
## drivers      8.327e-02  5.544e-03  15.018   <2e-16 ***
## front       -3.838e-03  1.723e-02  -0.223    0.824
## rear         5.189e-03  2.474e-02   0.210    0.834
## kms          5.875e-04  5.025e-04   1.169    0.244
```

```
## PetrolPrice -1.629e+01  8.643e+01  -0.188    0.851
## VanKilled    5.928e-02  2.911e-01   0.204    0.839
## law1         4.070e+00  4.076e+00   0.998    0.319
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.58 on 184 degrees of freedom
## Multiple R-squared:  0.7993, Adjusted R-squared:  0.7917
## F-statistic: 104.7 on 7 and 184 DF,  p-value: < 2.2e-16
```

The linear model shows that just the drivers variable are significant which means that the drivers are fit the model. Other variables include the Intercept are not significant since their p-value are larger than 0.05, which means that the model can not fit all the variables very well.

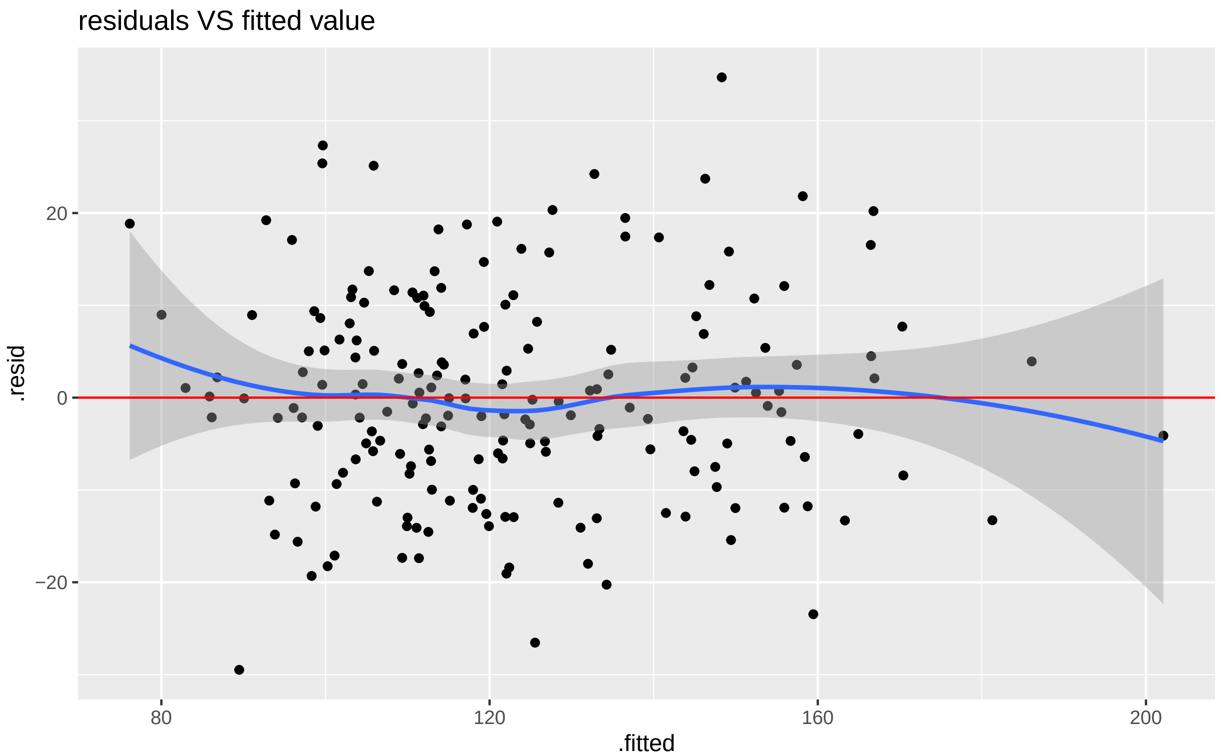

Figure 5: *residuals VS fitted value*

In order to analyze the residuals, I use the **broom** package(/cite{11}).

From the figure 5, it indicates that the residuals does not include obvious pattern of the variables since the residuals are randomly around the red line. It means that the model can capture most information about the response variable DriversKilled. But we can see that the tail at the beginning and end is not like a straingt since the obervations are not enough at the end and the beginning.

{Section 3}