Traffic_Crashes_Linear_Regression

Divya Sharma (ds655)

```r
library(dplyr, quietly = T)
```

```
Attaching package: 'dplyr'
```

```
The following objects are masked from 'package:stats':

    filter, lag
```

```
The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```r
base_data <- readRDS(file='C:/Users/divya/OneDrive/Documents/Duke/MIDS/Semester 1/IDS 702
head(base_data)
```

A data.frame: 6 × 18

	crash_date	day_of_week	month	hour	weather_condition	road_type	lighting_condition	first_crash_type	street_temp	precipitation	snow_depth	speed	day	damage	police_notified_tim
	<dttm>	<fct>	<fct>	<int>	<fct>	<fct>	<fct>	<int>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<fct>	<dbl>
1	2023-11-07-29 11:40:00	Saturday	July	30	CLEAR	NOT DIVIDED	DRY	2	7180.245	17.100	2	0	33.3	15.6	Major	3

	crash_date	hr	day_of_week	month_of_year	speed_limit	weather_condition	trafficway_type	roadway_surface_cond	num_units	dist_to_dt	temp	precipitation	snow	snowdepth	windspeed	visibility	damage	police_notified_time
2	2023-07-29 09:58:00	9	Saturday	Jul	45	CLEAR	NOT DIVIDED	DRY	2	24147.7	24.6	17.1	0.2	0	33.3	15.6	Minor	42
3	2023-07-10 16:30:00	16	Monday	Jul	30	CLEAR	ONE-WAY	DRY	2	9038.22	34.7	0.000	0	0	22.4	16.0	Minor	49
4	2019-09-29 03:45:00	3	Sunday	Sep	30	RAIN	NOT DIVIDED	WET	2	13069.73	19.6	17.98	0.7	0	22.1	13.1	Major	6
5	2023-09-05 16:36:00	16	Tuesday	Sep	30	CLEAR	FOUR-WAY	WET	2	20479.3	28.2	0.000	0	0	29.0	16.0	Major	10
6	2023-07-29 01:10:00	1	Saturday	Jul	30	RAIN	NOT DIVIDED	WET	1	20798.63	24.6	17.1	0.2	0	33.3	15.6	Major	5

```r
model <- lm(police_notified_time ~
            #crash_date +
            hr_of_day +
            day_of_week +
            month_of_year +
            speed_limit +
            weather_condition +
            trafficway_type +
            roadway_surface_cond +
            num_units +
            dist_to_dt +
            #temp +
            precipitation +
            snow +
            snowdepth +
            windspeed +
            visibility +
```

```
                damage_class
              , data = base_data)
  summary(model)
```

Call:
lm(formula = police_notified_time ~ hr_of_day + day_of_week +
    month_of_year + speed_limit + weather_condition + trafficway_type +
    roadway_surface_cond + num_units + dist_to_dt + precipitation +
    snow + snowdepth + windspeed + visibility + damage_class,
    data = base_data)

Residuals:
   Min   1Q Median   3Q   Max
-423.0 -100.2 -66.2 -22.6 1448.3

Coefficients:

	Estimate	Std. Error	t value
(Intercept)	5.460e+01	1.201e+02	0.454
hr_of_day	4.777e+00	6.686e-02	71.446
day_of_weekMonday	-1.653e+01	1.420e+00	-11.640
day_of_weekTuesday	-1.861e+01	1.408e+00	-13.212
day_of_weekWednesday	-1.853e+01	1.410e+00	-13.149
day_of_weekThursday	-1.661e+01	1.406e+00	-11.815
day_of_weekFriday	-1.949e+01	1.375e+00	-14.179
day_of_weekSaturday	-1.629e+01	1.407e+00	-11.571
month_of_yearFeb	3.769e+00	1.932e+00	1.951
month_of_yearMar	-1.393e+00	1.898e+00	-0.734
month_of_yearApr	1.861e-01	1.927e+00	0.097
month_of_yearMay	2.220e+00	1.881e+00	1.180
month_of_yearJun	3.559e+00	1.883e+00	1.890
month_of_yearJul	4.622e+00	1.897e+00	2.436
month_of_yearAug	4.979e+00	1.880e+00	2.648
month_of_yearSep	3.649e+00	1.880e+00	1.941
month_of_yearOct	2.542e+00	1.880e+00	1.352
month_of_yearNov	8.463e-01	1.875e+00	0.451
month_of_yearDec	3.095e-01	1.852e+00	0.167
speed_limit	-7.925e-01	7.023e-02	-11.285
weather_conditionBLOWING SNOW	3.654e+01	1.211e+02	0.302
weather_conditionCLEAR	7.557e+01	1.200e+02	0.630
weather_conditionCLOUDY/OVERCAST	4.337e+01	1.201e+02	0.361
weather_conditionFOG/SMOKE/HAZE	5.169e+01	1.204e+02	0.429

```
weather_conditionFREEZING RAIN/DRIZZLE                  6.769e+01  1.203e+02   0.562
weather_conditionOTHER                                 7.794e+01  1.202e+02   0.648
weather_conditionRAIN                                  7.884e+01  1.201e+02   0.657
weather_conditionSEVERE CROSS WIND GATE                7.679e+01  1.228e+02   0.626
weather_conditionSLEET/HAIL                            6.484e+01  1.205e+02   0.538
weather_conditionSNOW                                  7.742e+01  1.201e+02   0.645
weather_conditionUNKNOWN                               2.619e+02  1.201e+02   2.181
trafficway_typeCENTER TURN LANE                       -3.895e+01  5.064e+00  -7.692
trafficway_typeDIVIDED - W/MEDIAN (NOT RAISED)        -3.223e+00  3.096e+00  -1.041
trafficway_typeDIVIDED - W/MEDIAN BARRIER             -2.087e+01  3.360e+00  -6.209
trafficway_typeDRIVEWAY                               -7.503e+00  6.977e+00  -1.075
trafficway_typeFIVE POINT, OR MORE                    -3.644e+01  1.028e+01  -3.543
trafficway_typeFOUR WAY                               -3.122e+01  3.321e+00  -9.399
trafficway_typeL-INTERSECTION                         -4.305e+01  2.587e+01  -1.664
trafficway_typeNOT DIVIDED                             8.177e+00  3.012e+00   2.715
trafficway_typeNOT REPORTED                           -2.704e+01  1.443e+01  -1.874
trafficway_typeONE-WAY                                 4.311e+01  3.119e+00  13.818
trafficway_typeOTHER                                   3.376e+00  3.659e+00   0.923
trafficway_typePARKING LOT                             2.762e+01  3.226e+00   8.563
trafficway_typeRAMP                                   -8.441e+00  7.223e+00  -1.169
trafficway_typeROUNDABOUT                             -2.548e+01  2.196e+01  -1.160
trafficway_typeT-INTERSECTION                         -3.389e+01  4.423e+00  -7.663
trafficway_typeTRAFFIC ROUTE                          -3.731e+01  1.132e+01  -3.297
trafficway_typeUNKNOWN                                 1.842e+01  4.665e+00   3.947
trafficway_typeUNKNOWN INTERSECTION TYPE              -2.008e+01  7.426e+00  -2.704
trafficway_typeY-INTERSECTION                         -2.519e+01  1.019e+01  -2.473
roadway_surface_condICE                                1.489e+00  4.533e+00   0.329
roadway_surface_condOTHER                             -7.204e+00  7.657e+00  -0.941
roadway_surface_condSAND, MUD, DIRT                   -4.535e+01  1.876e+01  -2.417
roadway_surface_condSNOW OR SLUSH                     -4.784e-02  2.948e+00  -0.016
roadway_surface_condUNKNOWN                            3.892e+01  1.929e+00  20.171
roadway_surface_condWET                               -2.593e+01  1.746e+00 -14.850
num_units                                             -2.303e+01  8.205e-01 -28.064
dist_to_dt                                            -2.905e-04  6.806e-05  -4.268
precipitation                                          2.409e-01  6.125e-02   3.932
snow                                                  -8.512e-02  7.780e-01  -0.109
snowdepth                                              5.206e-03  2.683e-01   0.019
windspeed                                              1.904e-01  5.154e-02   3.694
visibility                                            -1.263e+00  2.114e-01  -5.975
damage_classMinor                                      1.102e+01  7.490e-01  14.710
                                                      Pr(>|t|)
(Intercept)                                           0.649530
hr_of_day                                              < 2e-16 ***
```

```
day_of_weekMonday                                       < 2e-16 ***
day_of_weekTuesday                                      < 2e-16 ***
day_of_weekWednesday                                    < 2e-16 ***
day_of_weekThursday                                     < 2e-16 ***
day_of_weekFriday                                       < 2e-16 ***
day_of_weekSaturday                                     < 2e-16 ***
month_of_yearFeb                                        0.051039 .
month_of_yearMar                                        0.462989
month_of_yearApr                                        0.923044
month_of_yearMay                                        0.237805
month_of_yearJun                                        0.058758 .
month_of_yearJul                                        0.014847 *
month_of_yearAug                                        0.008091 **
month_of_yearSep                                        0.052272 .
month_of_yearOct                                        0.176378
month_of_yearNov                                        0.651745
month_of_yearDec                                        0.867311
speed_limit                                             < 2e-16 ***
weather_conditionBLOWING SNOW                           0.762899
weather_conditionCLEAR                                  0.528959
weather_conditionCLOUDY/OVERCAST                        0.717889
weather_conditionFOG/SMOKE/HAZE                         0.667718
weather_conditionFREEZING RAIN/DRIZZLE                  0.573791
weather_conditionOTHER                                  0.516758
weather_conditionRAIN                                   0.511370
weather_conditionSEVERE CROSS WIND GATE                 0.531611
weather_conditionSLEET/HAIL                             0.590489
weather_conditionSNOW                                   0.519050
weather_conditionUNKNOWN                                0.029159 *
trafficway_typeCENTER TURN LANE                         1.45e-14 ***
trafficway_typeDIVIDED - W/MEDIAN (NOT RAISED) 0.297954
trafficway_typeDIVIDED - W/MEDIAN BARRIER               5.33e-10 ***
trafficway_typeDRIVEWAY                                 0.282215
trafficway_typeFIVE POINT, OR MORE                      0.000396 ***
trafficway_typeFOUR WAY                                 < 2e-16 ***
trafficway_typeL-INTERSECTION                           0.096032 .
trafficway_typeNOT DIVIDED                              0.006631 **
trafficway_typeNOT REPORTED                             0.060966 .
trafficway_typeONE-WAY                                  < 2e-16 ***
trafficway_typeOTHER                                    0.356222
trafficway_typePARKING LOT                              < 2e-16 ***
trafficway_typeRAMP                                     0.242583
trafficway_typeROUNDABOUT                               0.245867
```

```
trafficway_typeT-INTERSECTION   1.82e-14 ***
trafficway_typeTRAFFIC ROUTE   0.000977 ***
trafficway_typeUNKNOWN   7.91e-05 ***
trafficway_typeUNKNOWN INTERSECTION TYPE   0.006842 **
trafficway_typeY-INTERSECTION   0.013407 *
roadway_surface_condICE   0.742533
roadway_surface_condOTHER   0.346812
roadway_surface_condSAND, MUD, DIRT   0.015632 *
roadway_surface_condSNOW OR SLUSH   0.987054
roadway_surface_condUNKNOWN   < 2e-16 ***
roadway_surface_condWET   < 2e-16 ***
num_units   < 2e-16 ***
dist_to_dt   1.97e-05 ***
precipitation   8.41e-05 ***
snow   0.912874
snowdepth   0.984519
windspeed   0.000221 ***
visibility   2.31e-09 ***
damage_classMinor   < 2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 268.3 on 555292 degrees of freedom
Multiple R-squared: 0.04479,   Adjusted R-squared: 0.04468
F-statistic: 413.3 on 63 and 555292 DF, p-value: < 2.2e-16
```
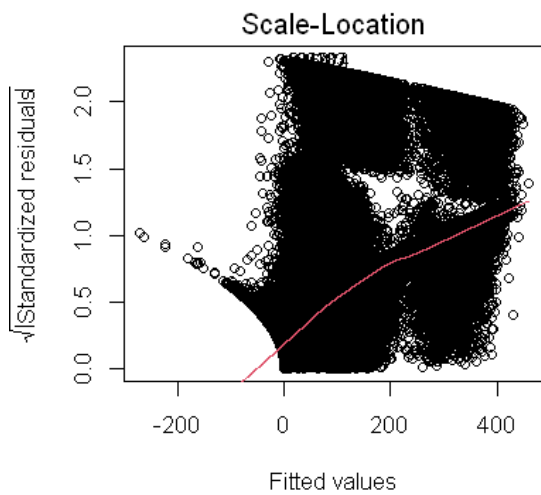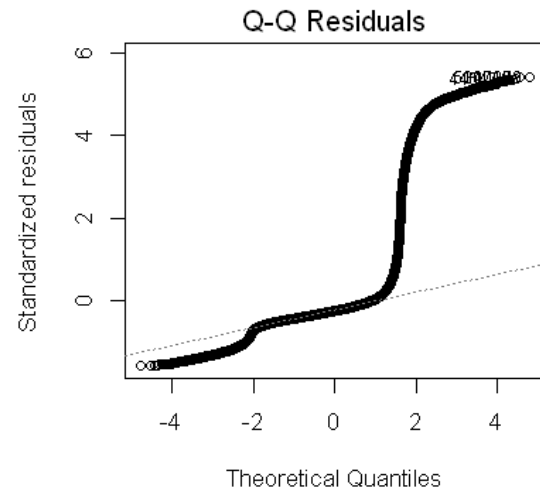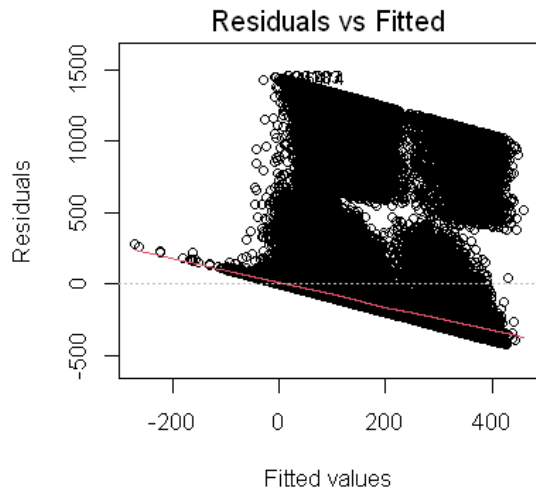
```r
  # Set up a 2x2 grid for the plots
  par(mfrow = c(2, 2))

  # Generate the four diagnostic plots
  plot(model)
```
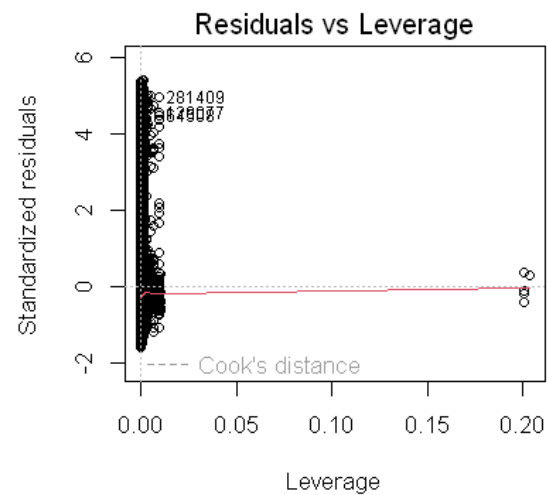
```
# Set up a 2x2 grid for the plots
par(mfrow = c(2, 2))

# Generate the Residuals vs Fitted plot with a title
plot(model, which = 1, main = "Residuals vs Fitted")

# Generate the Normal Q-Q plot with a title
```
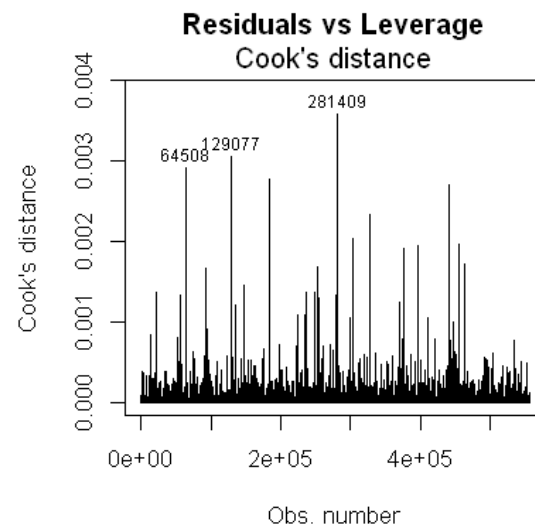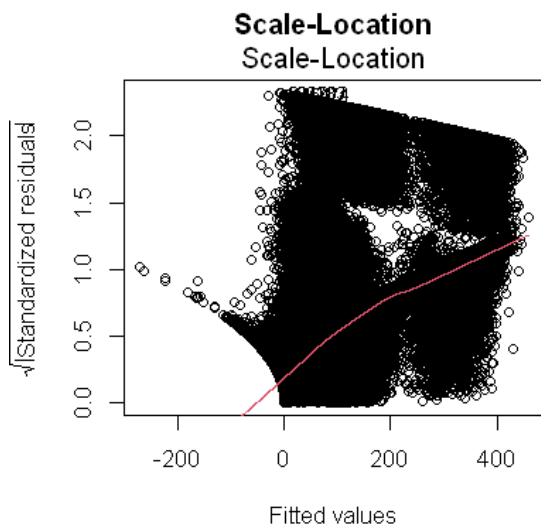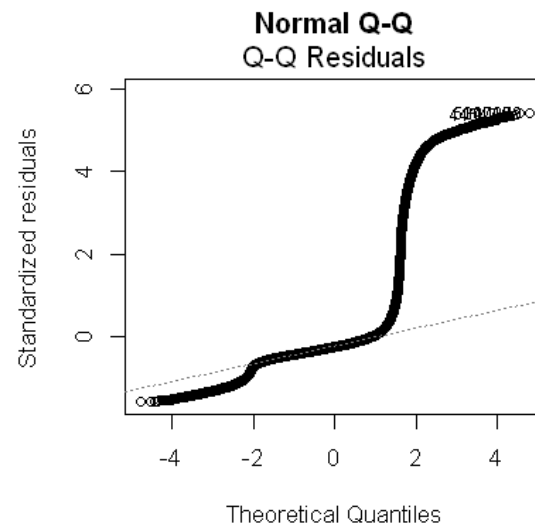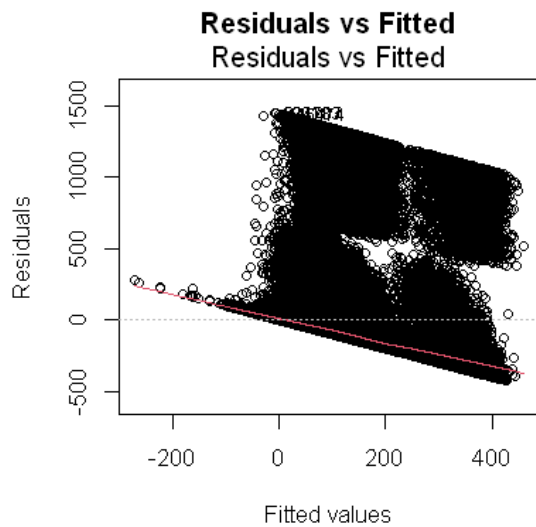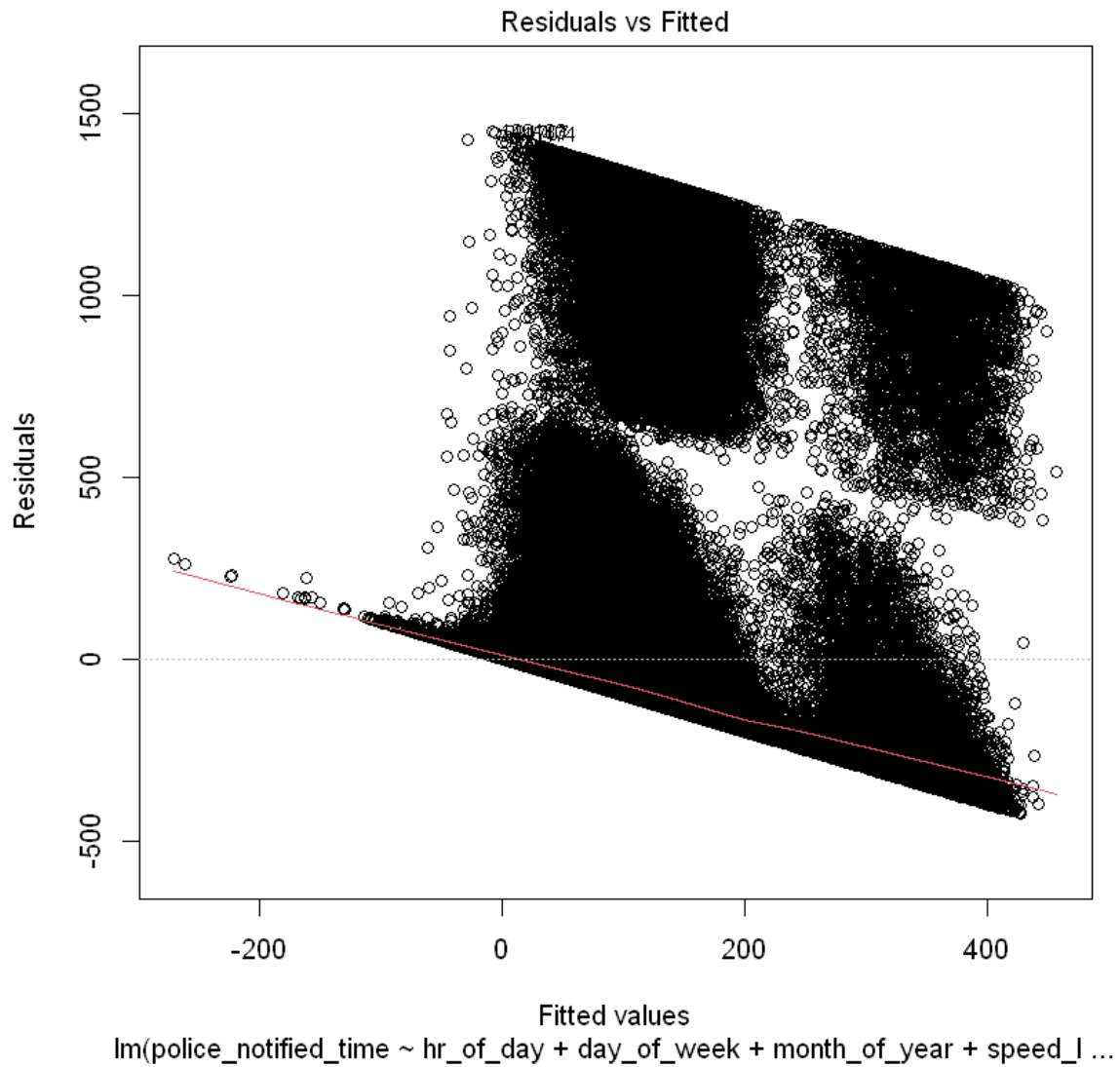
```r
plot(model, which = 2, main = "Normal Q-Q")

# Generate the Scale-Location plot with a title
plot(model, which = 3, main = "Scale-Location")

# Generate the Residuals vs Leverage plot with a title
plot(model, which = 4, main = "Residuals vs Leverage")
```
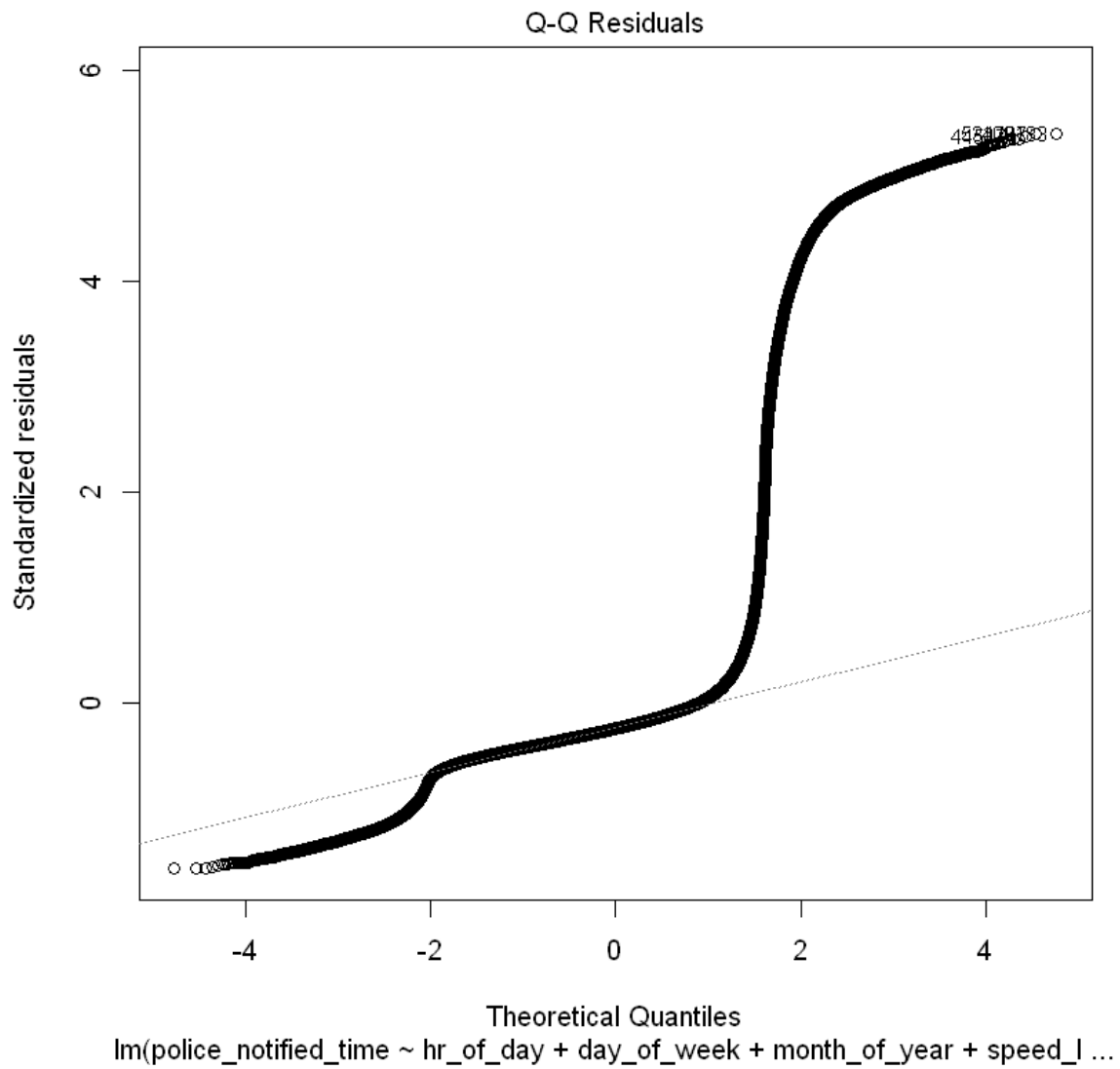
Residuals vs Fitted
Residuals vs Fitted

Normal Q-Q
Q-Q Residuals

Scale-Location
Scale-Location

Residuals vs Leverage
Cook's distance

```
plot(model, which = 1)
```

Residuals vs Fitted

Fitted values
lm(police_notified_time ~ hr_of_day + day_of_week + month_of_year + speed_l ...

```
plot(model, which = 2)
```

Q-Q Residuals

Theoretical Quantiles
lm(police_notified_time ~ hr_of_day + day_of_week + month_of_year + speed_l ...

```
plot(model, which = 3)
```

Scale-Location

√|Standardized residuals|

Fitted values
lm(police_notified_time ~ hr_of_day + day_of_week + month_of_year + speed_l ...

```
plot(model, which = 4)
```

Cook's distance

281409

129077

64508

Cook's distance

0.004

0.003

0.002

0.001

0.000

0e+00   1e+05   2e+05   3e+05   4e+05   5e+05

Obs. number
lm(police_notified_time ~ hr_of_day + day_of_week + month_of_year + speed_l ...

```
car::vif(model)
```

A matrix: 15 × 3 of type dbl

	GVIF	Df	GVIF^(1/(2*Df))
hr_of_day	1.015173	1	1.007558
day_of_week	1.018192	6	1.001504
month_of_year	1.733831	11	1.025331

	GVIF	Df	GVIF^(1/(2*Df))
speed_limit	1.386287	1	1.177407
weather_condition	10.057917	11	1.110628
trafficway_type	1.532143	19	1.011291
roadway_surface_cond	10.258189	6	1.214104
num_units	1.025228	1	1.012535
dist_to_dt	1.021129	1	1.010509
precipitation	1.457018	1	1.207070
snow	1.372103	1	1.171368
snowdepth	1.567669	1	1.252066
windspeed	1.124771	1	1.060552
visibility	1.679774	1	1.296061
damage_class	1.023165	1	1.011516