

Untitled

Data and package loading

```
library(dplyr)

##
## 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

library(tidyr)
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.0 --
## v ggplot2 3.3.2      v purrr   0.3.4
## v tibble  3.0.4      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.0
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(funModeling)

## Loading required package: Hmisc
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:dplyr':
##
##   src, summarize
## The following objects are masked from 'package:base':
##
##   format.pval, units
## funModeling v.1.9.4 :)
## Examples and tutorials at livebook.datascienceheroes.com
## / Now in Spanish: librovivodecienciadedatos.ai
```

```

library(Hmisc)
library(MASS)

##
## Attaching package: 'MASS'
## The following object is masked from 'package:dplyr':
##
##      select
library(magrittr)

##
## Attaching package: 'magrittr'
## The following object is masked from 'package:purrr':
##
##      set_names
## The following object is masked from 'package:tidyr':
##
##      extract
library(ggplot2)
library(easyGgplot2)
library(reshape2)

##
## Attaching package: 'reshape2'
## The following object is masked from 'package:tidyr':
##
##      smiths
library(Matrix)

##
## Attaching package: 'Matrix'
## The following objects are masked from 'package:tidyr':
##
##      expand, pack, unpack
library(lme4)
library(arm)

##
## arm (Version 1.11-2, built: 2020-7-27)
## Working directory is /Users/nadou/Desktop/678final
library(merTools)

## Registered S3 method overwritten by 'broom.mixed':
##      method      from
##      tidy.gamlss broom
library(sjstats)
library(HLMdiag)

##

```

```
## Attaching package: 'HLMdiag'

## The following object is masked from 'package:stats':
##
##      covratio

us_accident <- read.csv("US_Accidents_June20.csv")
```

Data cleaning

```
#select possible related variables
us_accident1 <- dplyr::select(us_accident,Severity,City,County,State,Weather_Timestamp,Weather_Condition)
#split year month of the data to check whether it contains all the data
us_accident2 <- separate(us_accident1,Weather_Timestamp, c("year","month","day"), sep = "-")
```

```
## Warning: Expected 3 pieces. Missing pieces filled with `NA` in 43323 rows [602,
## 1958, 1969, 1974, 1979, 1990, 2138, 2146, 2343, 2344, 2345, 2361, 2365, 2385,
## 4029, 4621, 4737, 4738, 4751, 5776, ...].
```

```
us_accident3 <- filter(us_accident2, year == "2019")
#check na of variables
df_status(us_accident3)
```

```
##      variable q_zeros p_zeros   q_na p_na q_inf p_inf      type unique
## 1      Severity      0    0.00      0 0.00      0      0    integer      4
## 2         City      0    0.00      0 0.00      0      0  character 9351
## 3        County      0    0.00      0 0.00      0      0  character 1505
## 4         State      0    0.00      0 0.00      0      0  character   49
## 5          year      0    0.00      0 0.00      0      0  character    1
## 6         month      0    0.00      0 0.00      0      0  character   12
## 7          day      0    0.00      0 0.00      0      0  character 34542
## 8 Weather_Condition      0    0.00      0 0.00      0      0  character   102
## 9   Temperature.F.    296    0.03   5495  0.58      0      0    numeric   671
## 10  Wind_Chill.F.    298    0.03 137556 14.60      0      0    numeric   950
## 11    Humidity...      0    0.00   6577  0.70      0      0    numeric   100
## 12  Visibility.mi.    597    0.06   6793  0.72      0      0    numeric    63
## 13  Wind_Speed.mph. 133106 14.13  48401  5.14      0      0    numeric    99
## 14  Precipitation.in. 666397 70.74 207455 22.02      0      0    numeric   172
## 15   Traffic_Signal      0    0.00      0 0.00      0      0  character     2
## 16   Civil_Twilight      0    0.00      0 0.00      0      0  character     3
```

```
#re-select dataset delete na value
new_us_accident <- dplyr::select(us_accident3,Severity,City,County,State,Weather_Condition,Temperature.F.)
#delete na values
new_us_accident <- na.omit(new_us_accident)
#check whether dataset is clean or not
unique(new_us_accident$Temperature.F.)
```

```
## [1] 31.0 28.0 43.0 53.0 54.0 52.0 49.0 46.0 45.0 59.0 58.0 36.0
## [13] 26.0 32.0 33.0 48.0 39.0 17.0 22.0 20.0 30.0 40.0 35.0 34.0
## [25] 21.0 29.0 19.0 38.0 47.0 42.0 41.0 37.0 23.0 24.0 25.0 56.0
## [37] 50.0 55.0 64.0 57.0 60.0 51.0 65.0 78.0 79.0 80.0 76.0 77.0
## [49] 75.0 66.0 63.0 73.0 71.0 62.0 61.0 44.0 68.0 69.0 67.0 70.0
## [61] 84.0 72.0 27.0 74.0 81.0 83.0 15.0 18.0 14.0 12.0 10.0 8.0
## [73] 6.0 16.0 13.0 7.0 11.0 5.0 82.0 85.0 9.0 3.0 1.0 4.0
## [85] 0.0 86.0 -8.0 -3.0 -4.0 -1.0 -2.0 2.0 -5.0 -6.0 -10.0 -7.0
```

```
## [97] -13.0 -12.0 88.0 95.0 89.0 90.0 94.0 92.0 97.0 96.0 93.0 87.0
## [109] 91.0 99.0 98.0 100.0 101.0 102.0 104.0 109.0 103.0 107.0 105.0 106.0
## [121] 108.0 110.0 111.0 112.0 113.0 115.0 -9.0 26.6 33.8 51.1 66.2 64.9
## [133] 69.1 71.6 26.1 24.1 21.9 33.1 39.9 17.1 27.5 27.1 15.1 28.9
## [145] 24.8 37.9 34.2 35.1 35.6 42.1 32.2 30.2 28.4 37.4 39.2 31.3
## [157] 28.2 30.9 30.7 48.9 46.9 44.1 41.5 44.4 47.5 82.9 80.1 84.9
## [169] 78.8 80.6 71.1 75.9 75.2 55.9 57.2 58.5 62.1 60.1 57.9 55.4
## [181] 53.1 78.1 68.9 64.4 66.9 -7.1 -11.9 -2.9 -0.9 -0.4 -14.1 -9.9
## [193] -5.8 8.1 22.1 6.1 10.9 53.6 47.3 21.7 44.6 16.2 29.8 42.6
## [205] 39.6 38.3 36.5 42.8 65.7 36.7 48.2 1.9 6.8 -0.8 -5.1 3.9
## [217] 1.4 69.8 51.8 46.6 46.4 60.8 58.3 51.4 54.3 56.3 52.9 62.6
## [229] 50.2 9.3 8.4 11.1 12.9 13.6 15.4 29.1 19.9 32.4 32.9 33.3
## [241] 35.2 12.2 10.4 38.8 23.5 34.7 35.8 19.4 8.6 12.4 13.5 13.1
## [253] 41.2 41.9 27.3 40.1 48.4 87.1 89.1 74.8 77.2 84.2 82.4 85.8
## [265] 73.9 84.4 35.4 46.8 44.8 55.8 -10.3 -11.0 -9.4 -8.1 -1.7 6.3
## [277] -11.2 3.2 -2.2 -2.4 -2.6 0.9 1.2 2.3 11.3 15.6 15.8 15.3
## [289] -13.2 -7.6 -16.2 5.4 4.6 0.1 31.8 32.5 40.6 29.3 18.7 21.2
## [301] 20.1 30.6 38.5 43.2 43.5 41.4 29.7 36.3 6.6 17.6 17.8 -3.5
## [313] -4.5 54.1 60.6 67.3 70.5 72.7 73.4 49.5 52.3 18.9 49.1 49.3
## [325] 45.5 9.5 14.9 22.8 24.6 25.7 25.5 23.9 24.3 25.2 25.3 27.7
## [337] 34.3 11.7 14.5 11.8 36.1 34.9 47.1 40.3 42.3 66.7 62.2 56.8
## [349] 26.4 37.6 46.2 42.4 52.7 2.8 2.5 4.3 5.7 14.2 18.5 20.8
## [361] 21.6 20.5 4.1 40.5 31.1 23.4 23.7 33.4 49.6 43.7 7.9 17.4
## [373] 28.8 47.7 77.4 76.6 60.3 78.6 78.4 63.1 18.1 45.1 57.4 70.2
## [385] 7.7 7.3 7.2 9.1 14.7 13.3 16.5 24.4 22.6 19.8 20.3 22.3
## [397] 32.7 10.6 27.9 26.8 37.8 44.2 10.8 11.5 28.6 47.8 53.4 55.6
## [409] 23.2 29.5 48.7 43.3 40.8 52.2 45.3 26.2 51.6 81.1 65.3 72.5
## [421] 76.5 83.8 83.1 77.5 59.5 55.2 57.7 58.1 58.6 53.8 14.4 39.4
## [433] 19.2 21.4 61.5 62.4 19.6 -12.3 56.7 59.7 65.5 50.9 48.6 67.1
## [445] 59.9 61.9 50.5 36.9 75.6 77.9 31.6 56.5 41.7 54.9 61.7 37.2
## [457] 30.4 64.6 65.1 79.9 63.7 63.5 68.7 59.2 54.5 73.2 69.4 61.2
## [469] 57.6 61.3 54.7 77.7 68.2 67.5 75.7 31.5 66.4 80.2 70.3 67.8
## [481] 63.9 58.8 87.8 52.5 51.3 62.8 45.9 75.4 79.3 64.8 59.4 71.2
## [493] 76.3 49.8 72.9 73.6 39.7 60.4 38.1 50.7 33.6 38.7 71.4 70.9
## [505] 56.1 53.2 74.1 71.8 65.8 74.7 69.3 70.7 78.3 25.9 69.6 73.8
## [517] 66.6 72.3 63.3 87.6 64.2 50.4 72.1 17.2 43.9 45.7 34.5 81.7
## [529] 76.1 74.5 67.6 68.5 74.3 68.4 84.6 83.7 81.3 22.5 80.8 83.5
## [541] 80.4 79.2 20.7 18.3 76.8 -1.3 9.9 16.7 -3.1 8.8 83.3 16.9
## [553] -15.2 4.5 -4.7 -4.4 0.5 -16.1 -17.0 12.6 6.4 8.2 3.7 -5.3
## [565] -1.8 13.8 -20.0 -23.1 -27.0 -22.0 -18.0 -18.9 -16.6 -15.9 -5.6 16.3
## [577] -25.6 -26.0 -24.0 -18.4 -18.2 -20.9 -23.8 -20.2 -14.8 -27.9 -25.1 -22.4
## [589] -29.0 -15.0 0.7 -1.5 12.7 4.8 3.4 7.5 0.3 10.2 5.9 -4.9
## [601] -7.2 -12.5 -11.6 -8.5 -8.7 -8.9 -7.8 -6.7 -10.1 -11.4 -10.7 -5.4
## [613] -4.2 -2.7 5.5 -12.8 -13.4 -11.7 1.8 86.2 81.5 2.1 82.2 -0.6
## [625] 9.7 86.9 89.6 -3.3 3.6 -6.5 79.5 2.7 5.2 -7.4 -6.2 -12.6
## [637] -10.8 -10.5 -6.9 -9.6 -13.9 -19.3 -15.3 -26.5 -21.1 -24.9 -24.5 -23.4
## [649] -21.3 -0.2 -27.4 -16.8 -24.2 -21.5 -14.0 -19.0 -1.1 170.6 81.9 -9.8
## [661] -23.3 82.8 -3.6 -9.2 -6.3 -13.5 -14.4 -17.9
```

```
unique(new_us_accident$Civil_Twilight)
```

```
## [1] "Day" "Night" ""
```

```
unique(new_us_accident$Traffic_Signal)
```

```
## [1] "False" "True"
```

```
unique(new_us_accident$Humidity...)
```

```
## [1] 85 92 53 86 83 97 93 71 82 38 37 39 59 55 72 81 100 57
## [19] 78 63 89 88 54 66 49 64 56 70 52 76 50 48 61 60 46 58
## [37] 62 65 73 84 79 90 94 74 51 96 75 77 45 41 44 47 99 67
## [55] 69 68 87 42 34 80 29 30 21 95 91 36 43 40 33 22 32 25
## [73] 28 35 98 27 24 26 16 31 23 20 19 17 18 15 13 12 10 11
## [91] 8 9 7 5 14 6 2 3 4 1
```

```
#find index for "" row of Civil_Twilight
space_index <- which(new_us_accident$Civil_Twilight == "")
#delete the observations rows where Civil_Twilight == ""
new_data <- new_us_accident[-c(space_index),,drop=F]
#counts of weather condition
freq(new_data$Weather_Condition)
```

```
## Warning in freq_logic(data = data, input = input, plot, na.rm, path_out =
## path_out): Skipping plot for variable 'var' (more than 100 categories)
```

##		var	frequency	percentage	cumulative_perc
## 1		Fair	324784	34.88	34.88
## 2		Mostly Cloudy	128698	13.82	48.70
## 3		Cloudy	112081	12.04	60.74
## 4		Partly Cloudy	100059	10.75	71.49
## 5		Clear	64881	6.97	78.46
## 6		Overcast	53034	5.70	84.16
## 7		Light Rain	46907	5.04	89.20
## 8		Light Snow	16091	1.73	90.93
## 9		Scattered Clouds	14499	1.56	92.49
## 10		Fog	11948	1.28	93.77
## 11		Rain	11178	1.20	94.97
## 12		Haze	8727	0.94	95.91
## 13		Heavy Rain	4055	0.44	96.35
## 14		Fair / Windy	3580	0.38	96.73
## 15		Light Drizzle	3459	0.37	97.10
## 16		T-Storm	2107	0.23	97.33
## 17		Thunder in the Vicinity	2104	0.23	97.56
## 18		Cloudy / Windy	2021	0.22	97.78
## 19		Mostly Cloudy / Windy	1959	0.21	97.99
## 20		Snow	1936	0.21	98.20
## 21		Light Rain with Thunder	1895	0.20	98.40
## 22		Thunder	1601	0.17	98.57
## 23			1476	0.16	98.73
## 24		Partly Cloudy / Windy	1266	0.14	98.87
## 25		Heavy T-Storm	1235	0.13	99.00
## 26		Light Rain / Windy	997	0.11	99.11
## 27		Light Freezing Rain	800	0.09	99.20
## 28		Wintry Mix	792	0.09	99.29
## 29		Patches of Fog	694	0.07	99.36
## 30		Drizzle	666	0.07	99.43
## 31		Mist	570	0.06	99.49
## 32		Smoke	553	0.06	99.55

## 33	Light Snow / Windy	442	0.05	99.60
## 34	Light Freezing Drizzle	379	0.04	99.64
## 35	Heavy Snow	376	0.04	99.68
## 36	Shallow Fog	371	0.04	99.72
## 37	Light Freezing Fog	308	0.03	99.75
## 38	Rain / Windy	282	0.03	99.78
## 39	Showers in the Vicinity	251	0.03	99.81
## 40	N/A Precipitation	159	0.02	99.83
## 41	Heavy Rain / Windy	141	0.02	99.85
## 42	Heavy T-Storm / Windy	138	0.01	99.86
## 43	Light Thunderstorms and Rain	132	0.01	99.87
## 44	T-Storm / Windy	126	0.01	99.88
## 45	Blowing Snow	121	0.01	99.89
## 46	Thunderstorms and Rain	110	0.01	99.90
## 47	Heavy Drizzle	89	0.01	99.91
## 48	Thunder / Windy	80	0.01	99.92
## 49	Haze / Windy	65	0.01	99.93
## 50	Drizzle and Fog	64	0.01	99.94
## 51	Blowing Dust / Windy	62	0.01	99.95
## 52	Fog / Windy	58	0.01	99.96
## 53	Heavy Thunderstorms and Rain	55	0.01	99.97
## 54	Light Rain Shower	52	0.01	99.98
## 55	Light Ice Pellets	51	0.01	99.99
## 56	Ice Pellets	46	0.00	99.99
## 57	Snow / Windy	44	0.00	99.99
## 58	Blowing Dust	43	0.00	99.99
## 59	Rain Showers	38	0.00	99.99
## 60	Thunderstorm	37	0.00	99.99
## 61	Sand / Dust Whirlwinds	27	0.00	99.99
## 62	Light Drizzle / Windy	25	0.00	99.99
## 63	Widespread Dust	25	0.00	99.99
## 64	Heavy Snow / Windy	21	0.00	99.99
## 65	Wintry Mix / Windy	19	0.00	99.99
## 66	Freezing Rain	17	0.00	99.99
## 67	Blowing Snow / Windy	10	0.00	99.99
## 68	Partial Fog	10	0.00	99.99
## 69	Light Snow with Thunder	9	0.00	99.99
## 70	Snow and Sleet	9	0.00	99.99
## 71	Snow and Sleet / Windy	9	0.00	99.99
## 72	Light Snow and Sleet	8	0.00	99.99
## 73	Small Hail	7	0.00	99.99
## 74	Heavy Sleet	6	0.00	99.99
## 75	Heavy Snow with Thunder	6	0.00	99.99
## 76	Rain Shower	6	0.00	99.99
## 77	Light Thunderstorms and Snow	5	0.00	99.99
## 78	Squalls / Windy	5	0.00	99.99
## 79	Drizzle / Windy	4	0.00	99.99
## 80	Light Freezing Rain / Windy	4	0.00	99.99
## 81	Light Snow and Sleet / Windy	4	0.00	99.99
## 82	Light Snow Showers	4	0.00	99.99
## 83	Low Drifting Snow	4	0.00	99.99
## 84	Squalls	4	0.00	99.99
## 85	Light Rain Showers	3	0.00	99.99
## 86	Light Sleet	3	0.00	99.99

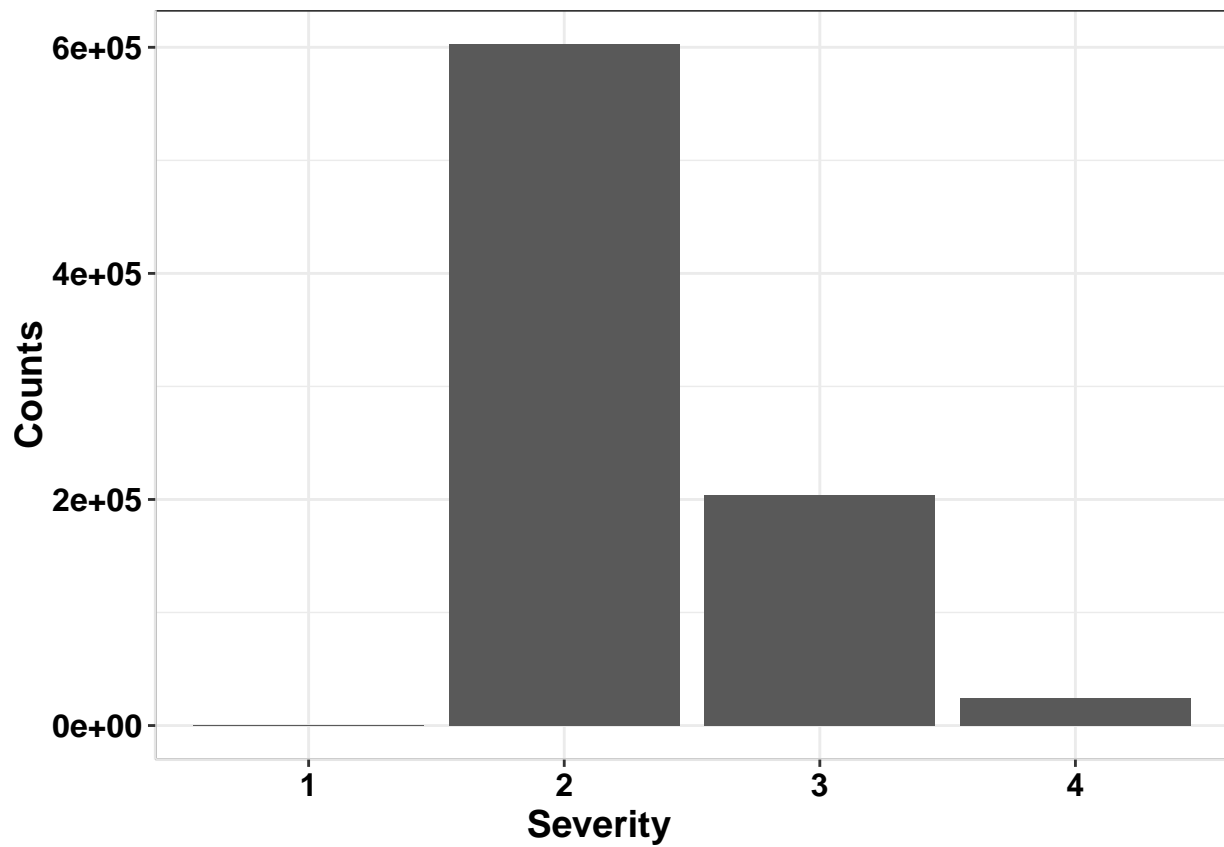
## 87	Light Snow Grains	3	0.00	99.99
## 88	Sleet	3	0.00	99.99
## 89	Heavy Blowing Snow	2	0.00	99.99
## 90	Sand / Dust Whirlwinds / Windy	2	0.00	99.99
## 91	Smoke / Windy	2	0.00	99.99
## 92	Tornado	2	0.00	99.99
## 93	Funnel Cloud	1	0.00	99.99
## 94	Hail	1	0.00	99.99
## 95	Heavy Freezing Drizzle	1	0.00	99.99
## 96	Light Rain Shower / Windy	1	0.00	99.99
## 97	Light Snow Shower	1	0.00	99.99
## 98	Partial Fog / Windy	1	0.00	99.99
## 99	Snow and Thunder	1	0.00	99.99
## 100	Thunder / Wintry Mix / Windy	1	0.00	99.99
## 101	Thunder and Hail / Windy	1	0.00	99.99
## 102	Widespread Dust / Windy	1	0.00	100.00

```
#select top 6 weather conditions
index1 <- which(new_data$Weather_Condition == "Fair")
index2 <- which(new_data$Weather_Condition == "Mostly Cloudy")
index3 <- which(new_data$Weather_Condition == "Cloudy")
index4 <- which(new_data$Weather_Condition == "Partly Cloudy")
index5 <- which(new_data$Weather_Condition == "Clear")
index6 <- which(new_data$Weather_Condition == "Overcast")
index7 <- which(new_data$Weather_Condition == "Light Rain")
index8 <- c(index1,index2,index3,index4,index5,index6,index7)
severity_data <- new_data[index8,]
```

Simple EDA

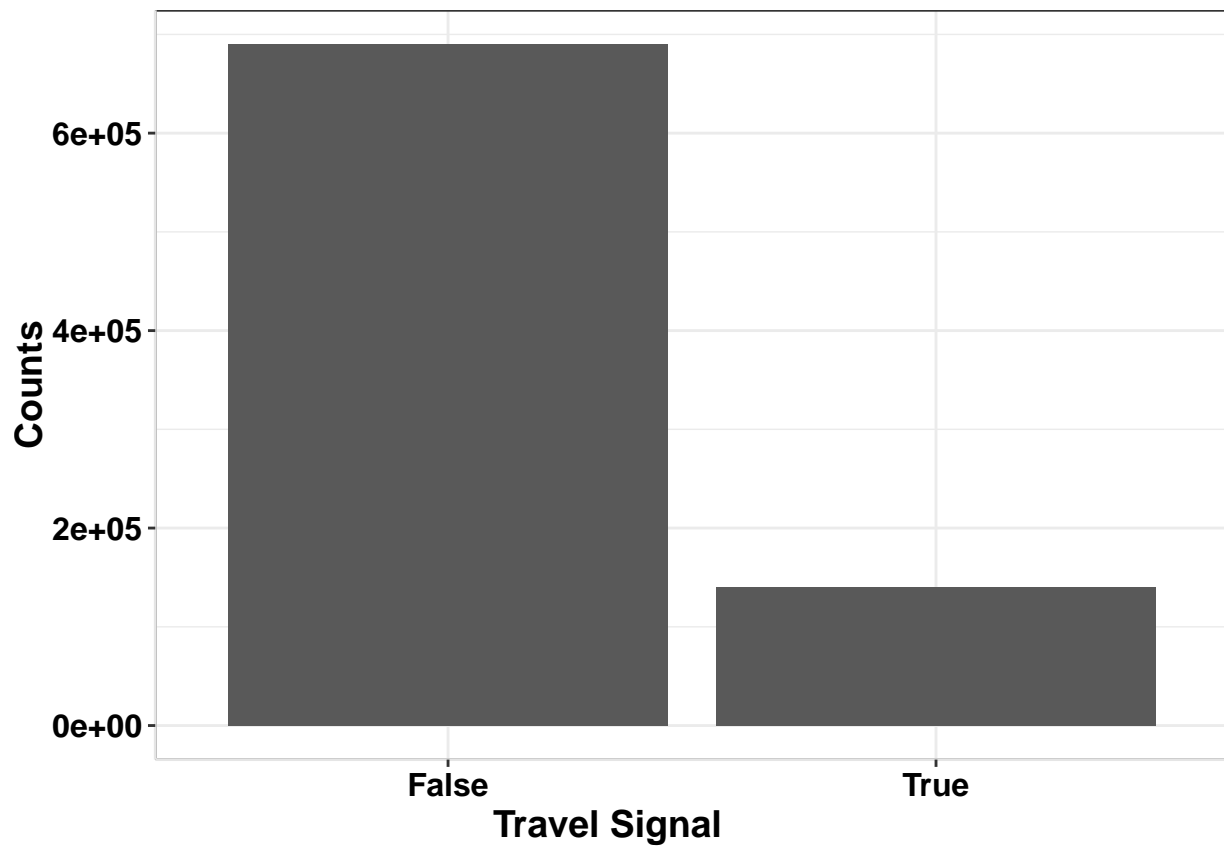
```
#count for severity of total observations
ggplot2.barplot(data = severity_data, xName = "Severity", yNmae = "frequency",
  backgroundColor="white", fill='lightblue', color="lightblue",
  xtitle = "Severity", ytitle= "Counts")
```

```
## Warning: Ignoring unknown parameters: binwidth
```



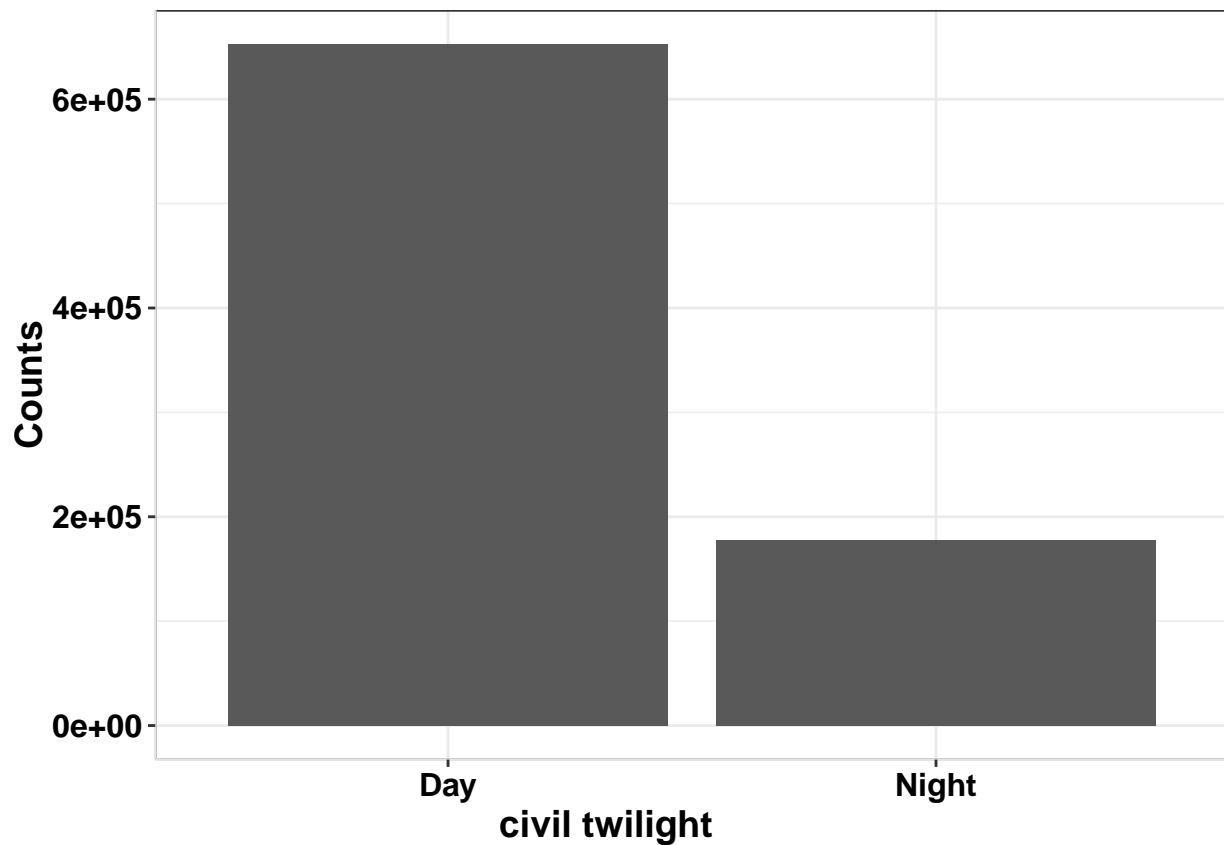
```
#count for travel signal  
ggplot2.barplot(data = severity_data, xName = "Traffic_Signal", yName = "frequency",  
  backgroundColor="white", fill='lightblue', color="lightblue",  
  xtitle = "Travel Signal", ytitle= "Counts")
```

```
## Warning: Ignoring unknown parameters: binwidth
```

```
#count for civil twilight  
ggplot2.barplot(data = severity_data, xName = "Civil_Twilight", yNmae = "frequency",  
                backgroundColor="white", fill='lightblue', color="lightblue",  
                xtitle = "civil twilight", ytitle= "Counts")
```

```
## Warning: Ignoring unknown parameters: binwidth
```



```
#range of continuous variable
range(severity_data$Temperature.F.)
```

```
## [1] -29.0 170.6
```

```
range(severity_data$Humidity...)
```

```
## [1] 1 100
```

```
range(severity_data$Visibility.mi.)
```

```
## [1] 0 110
```

```
#severity and temperature
```

```
sev <- severity_data$Severity[1:4000]
```

```
tem <- new_data$Temperature.F.[1:4000]
```

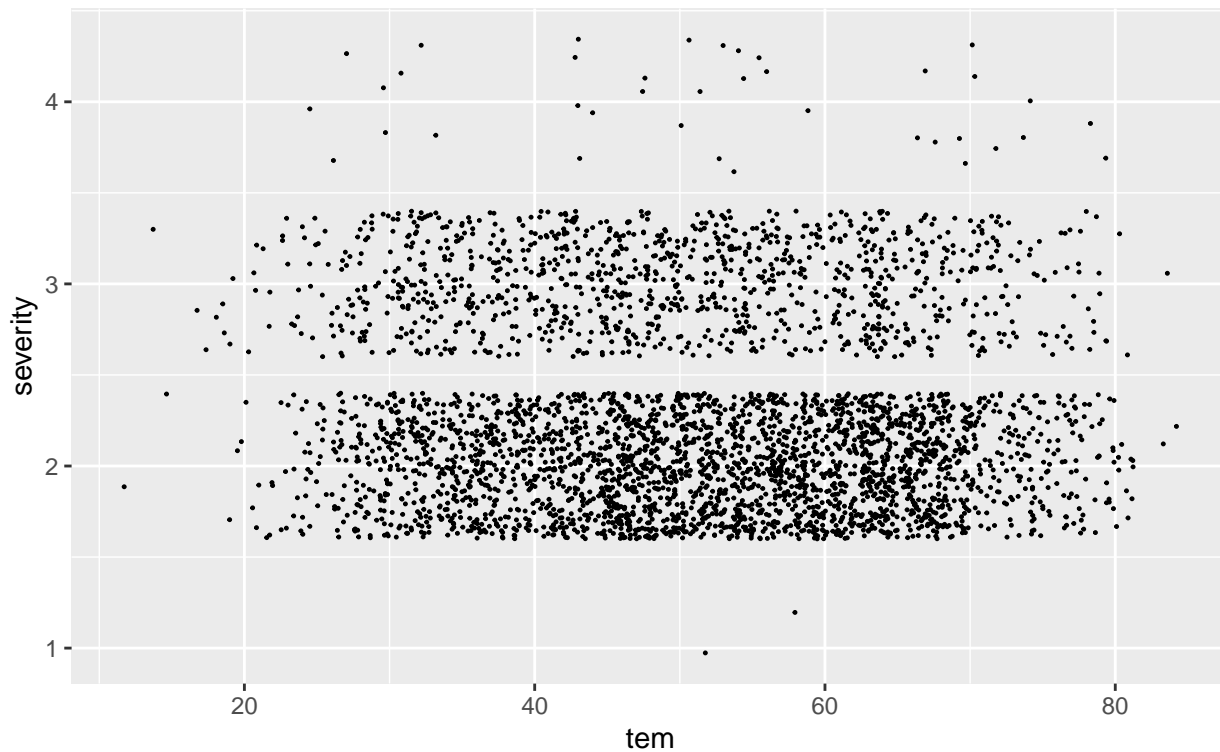
```
df1 <- data.frame(sev,tem)
```

```
g1 <- ggplot(df1, aes(tem,sev))
```

```
g1 + geom_jitter(width = 0.5, size =0.2) +
```

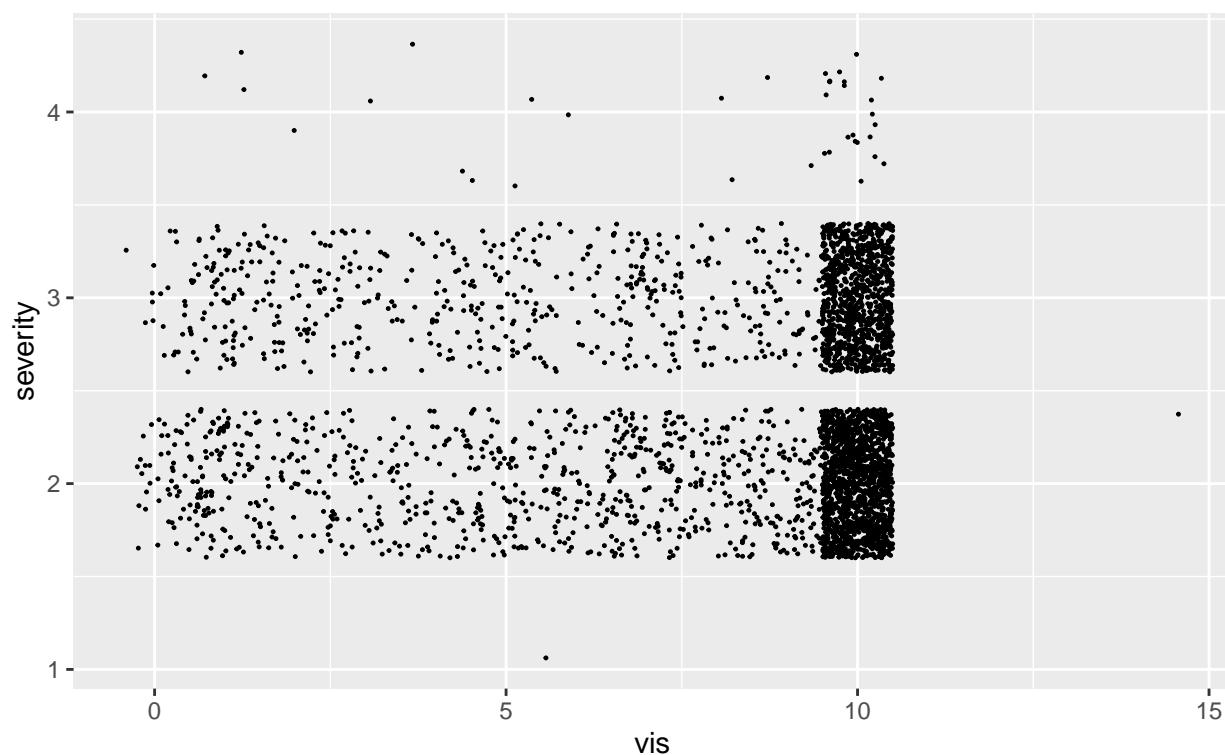
```
labs(subtitle = "Severity with tem", x ="tem", y = "severity", title = "jittered points")
```

jittered points
Severity with tem



```
#severity and visibility
sev <- new_data$Severity[1:4000]
vis <- new_data$Visibility.mi[1:4000]
df2 <- data.frame(sev,vis)
g2 <- ggplot(df2, aes(vis,sev))
g2 + geom_jitter(width = 0.5, size =0.2) +
  labs(subtitle = "Severity with tem", x ="vis", y = "severity", title = "jittered points")
```

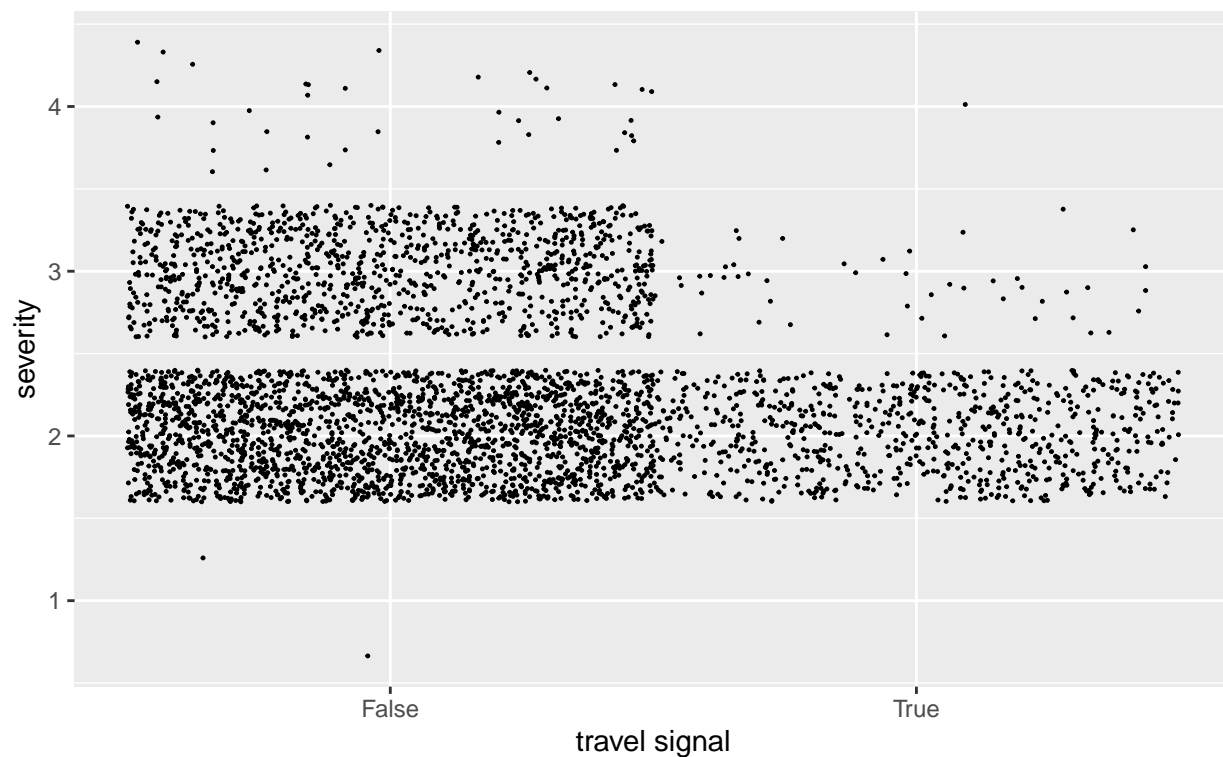
jittered points
Severity with tem



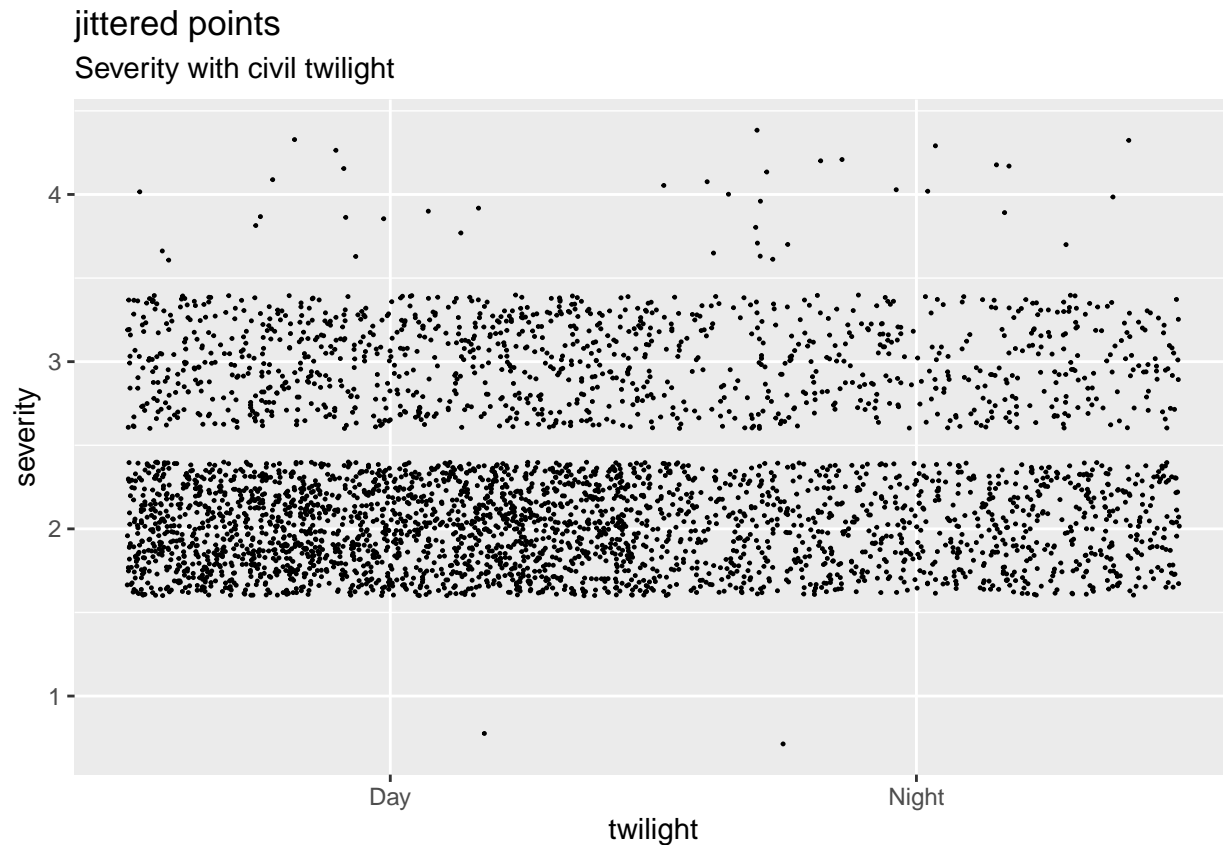
```
#severity and traffic signal
traffic <- severity_data$Traffic_Signal[1:4000]
sev <- severity_data$Severity[1:4000]
df3 <- data.frame(sev,traffic)
g3 <- ggplot(df3, aes(traffic,sev))
g3 + geom_jitter(width = 0.5, size =0.2) +
  labs(subtitle = "Severity with traffic signal", x ="travel signal", y = "severity", title = "jittered
```

jittered points

Severity with traffic signal



```
#severity and civil twilight
light <- severity_data$Civil_Twilight[1:4000]
df4 <- data.frame(sev,light)
g4 <- ggplot(df4, aes(light,sev))
g4 + geom_jitter(width = 0.5, size =0.2) +
  labs(subtitle = "Severity with civil twilight", x ="twilight", y = "severity", title = "jittered poin
```



Fit model

```
##Model fit
##HLM
#fit the empty model
fit_empty <- lmer(Severity ~ (1|State), data = severity_data)
#intraclass correlation coefficient
sjstats::icc(fit_empty)
```

```
## Warning: 'sjstats::icc' is deprecated.
## Use 'performance::icc()' instead.
## See help("Deprecated")
```

```
## # Intraclass Correlation Coefficient
##
##      Adjusted ICC: 0.219
##      Conditional ICC: 0.219
```

```
#random intercept models
```

```
fita <- lmer(Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal + Civil_Twilight
summary(fita)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal +
##      Civil_Twilight + Weather_Condition + (1 | State)
##      Data: severity_data
##
```

```

## REML criterion at convergence: 1170819
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.5704 -0.6354 -0.4131  0.7624  4.4720
##
## Random effects:
##   Groups   Name      Variance Std.Dev.
##   State    (Intercept) 0.06677  0.2584
##   Residual                0.23965  0.4895
## Number of obs: 830444, groups: State, 49
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)      2.461e+00  3.739e-02  65.834
## Humidity...      -5.128e-04  3.089e-05 -16.602
## Temperature.F.    1.651e-03  3.829e-05  43.121
## Visibility.mi.     -1.736e-05  2.657e-04  -0.065
## Traffic_SignalTrue -2.609e-01  1.489e-03 -175.167
## Civil_TwilightNight 8.926e-02  1.388e-03  64.321
## Weather_ConditionCloudy -6.487e-02  2.605e-03 -24.901
## Weather_ConditionFair -9.362e-02  2.280e-03 -41.063
## Weather_ConditionLight Rain -1.394e-02  3.247e-03  -4.293
## Weather_ConditionMostly Cloudy -4.378e-02  2.515e-03 -17.405
## Weather_ConditionOvercast 3.403e-04  2.955e-03   0.115
## Weather_ConditionPartly Cloudy -6.032e-02  2.619e-03 -23.033
##
## Correlation of Fixed Effects:
##      (Intr) Hmd... Tmp.F. Vsbl... Trf_ST Cvl_TN Wth_CC Wth_CF Wt_CLR
## Humidity... -0.072
## Tempertr.F. -0.057  0.307
## Visiblty.m. -0.079  0.186  0.002
## Trffc_SgnlT -0.006  0.007 -0.001  0.001
## Cvl_TwlghtN -0.008 -0.152  0.191 -0.031  0.024
## Wthr_CndtnC -0.014 -0.280 -0.297  0.021 -0.006 -0.002
## Wthr_CndtnF -0.021 -0.119 -0.373 -0.004 -0.008 -0.031  0.736
## Wthr_CndtLR -0.023 -0.276 -0.224  0.200  0.001  0.020  0.567  0.584
## Wthr_CndtMC -0.016 -0.183 -0.318 -0.016 -0.002  0.009  0.681  0.767  0.548
## Wthr_CndtnO -0.029 -0.184 -0.072  0.095  0.003  0.016  0.549  0.574  0.475
## Wthr_CndtPC -0.018 -0.116 -0.318 -0.012 -0.004 -0.007  0.642  0.743  0.511
##      Wt_CMC Wth_CO
## Humidity...
## Tempertr.F.
## Visiblty.m.
## Trffc_SgnlT
## Cvl_TwlghtN
## Wthr_CndtnC
## Wthr_CndtnF
## Wthr_CndtLR
## Wthr_CndtMC
## Wthr_CndtnO 0.539
## Wthr_CndtPC 0.674  0.504

```

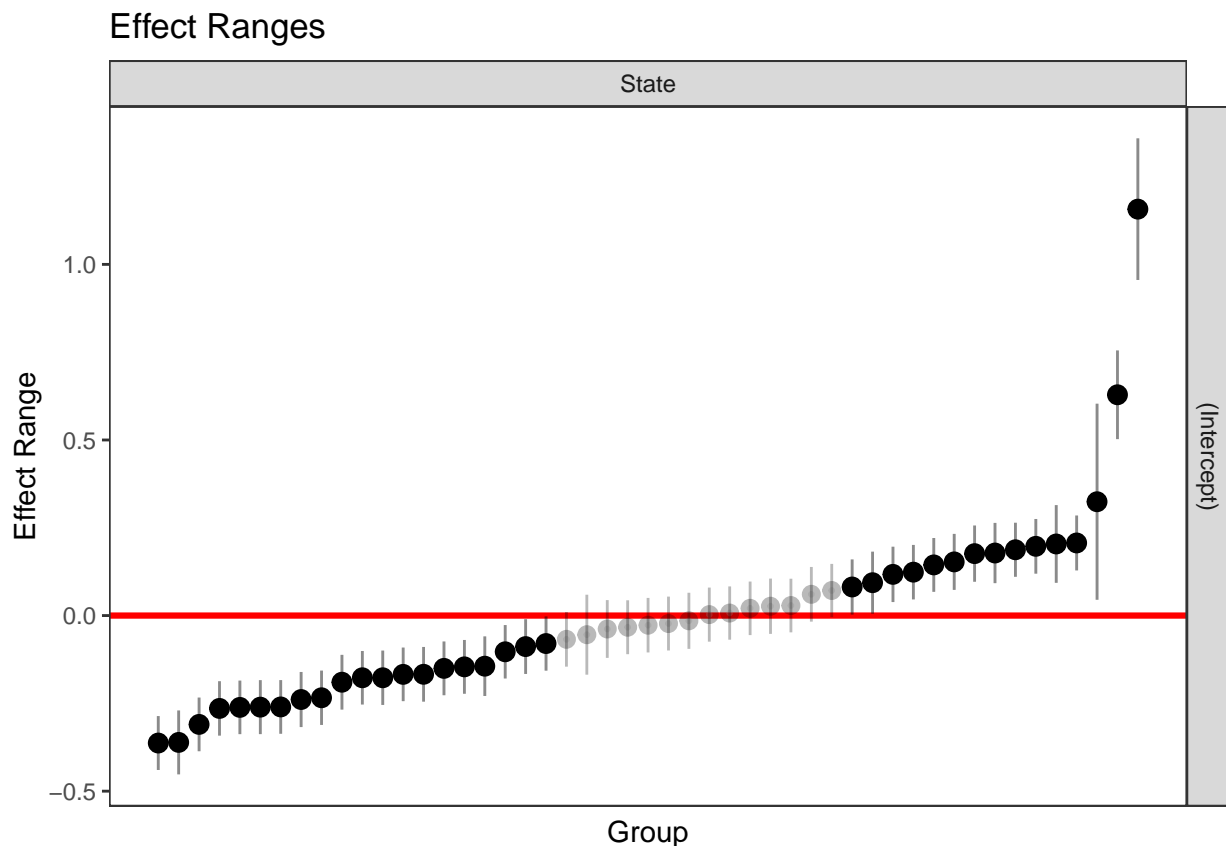
```
confint(fita)
```

```
## Computing profile confidence intervals ...

##                2.5 %        97.5 %
## .sig01          0.2079305987  0.3204700409
## .sigma          0.4887904491  0.4902795691
## (Intercept)     2.3875243680  2.5356721654
## Humidity...     -0.0005733165 -0.0004522430
## Temperature.F.  0.0015762331  0.0017263441
## Visibility.mi.   -0.0005381159  0.0005034625
## Traffic_SignalTrue -0.2637731667 -0.2579357297
## Civil_TwilightNight 0.0865442202  0.0919842036
## Weather_ConditionCloudy -0.0699721455 -0.0597609309
## Weather_ConditionFair -0.0980840207 -0.0891474704
## Weather_ConditionLight Rain -0.0203048477 -0.0075760037
## Weather_ConditionMostly Cloudy -0.0487104424 -0.0388504921
## Weather_ConditionOvercast -0.0054520449  0.0061323298
## Weather_ConditionPartly Cloudy -0.0654497496 -0.0551847619
```

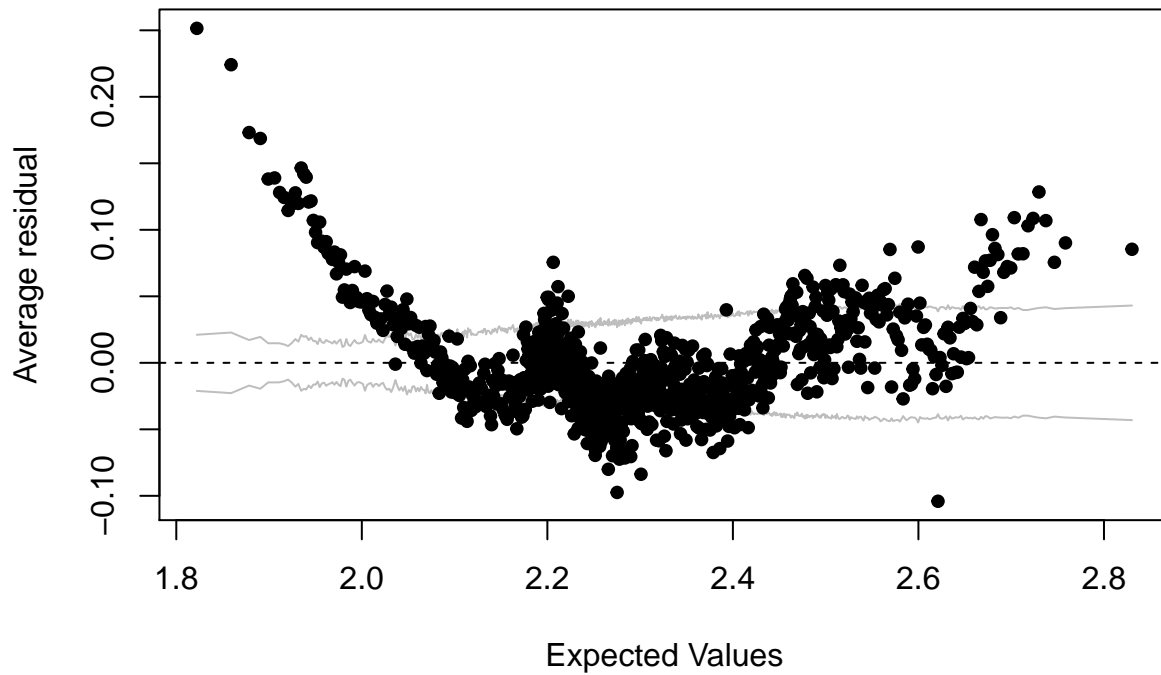
```
#plot the interval estimates
```

```
plotREsim(REsim(fita))
```



```
binnedplot(predict(fita),resid(fita))
```


Binned residual plot



#with interaction

```
fitd <- lmer(Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal + Civil_Twilight
summary(fitd)
```

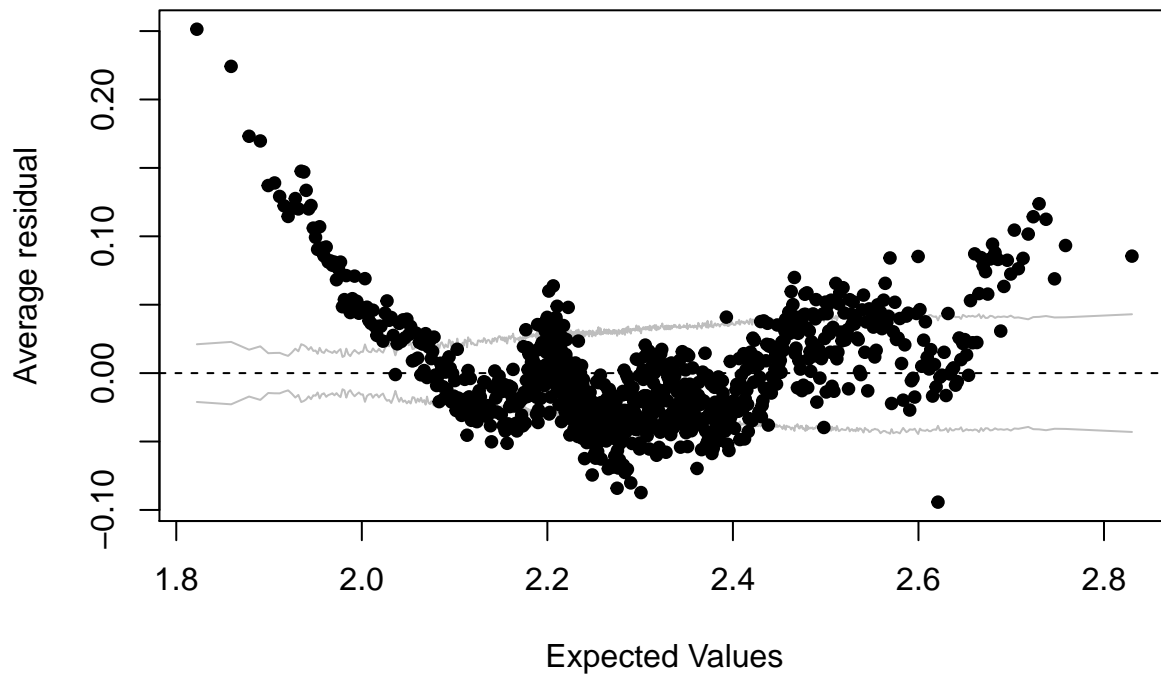
```
## Linear mixed model fit by REML ['lmerMod']
## Formula:
## Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal +
## Civil_Twilight + Weather_Condition + Humidity...:Visibility.mi. +
## (1 | State)
## Data: severity_data
##
## REML criterion at convergence: 1170840
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.5718 -0.6354 -0.4131  0.7624  4.4720
##
## Random effects:
##  Groups   Name                Variance Std.Dev.
##  State    (Intercept)  0.06677  0.2584
##  Residual                    0.23965  0.4895
## Number of obs: 830444, groups: State, 49
##
## Fixed effects:
##
##              Estimate Std. Error t value
## (Intercept)    2.460e+00  3.772e-02  65.213
## Humidity...   -4.889e-04  8.519e-05  -5.739
## Temperature.F.  1.651e-03  3.830e-05  43.114
## Visibility.mi.  1.400e-04  5.874e-04   0.238
## Traffic_SignalTrue -2.609e-01  1.489e-03 -175.160
```

```
## Civil_TwilightNight      8.927e-02  1.388e-03  64.310
## Weather_ConditionCloudy -6.488e-02  2.605e-03 -24.903
## Weather_ConditionFair   -9.361e-02  2.280e-03 -41.057
## Weather_ConditionLight Rain -1.412e-02  3.304e-03  -4.275
## Weather_ConditionMostly Cloudy -4.377e-02  2.515e-03 -17.402
## Weather_ConditionOvercast  2.519e-04  2.970e-03   0.085
## Weather_ConditionPartly Cloudy -6.031e-02  2.619e-03 -23.032
## Humidity...:Visibility.mi. -2.528e-06  8.418e-06  -0.300
```

```
##
## Correlation matrix not shown by default, as p = 13 > 12.
## Use print(x, correlation=TRUE) or
##     vcov(x)         if you need it
```

```
binnedplot(predict(fitd),resid(fitd))
```

Binned residual plot

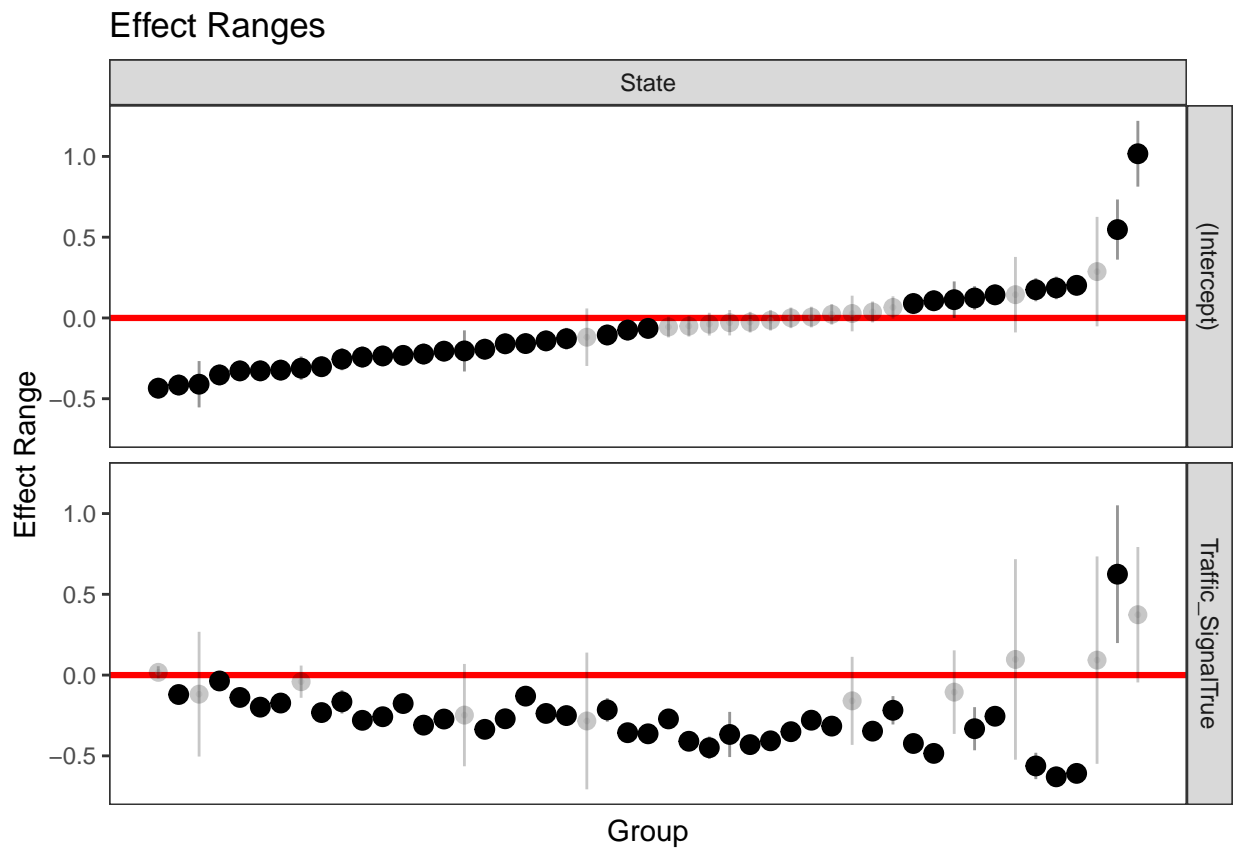


```
AIC(fita,fitd)
```

```
##      df      AIC
## fita 14 1170847
## fitd 15 1170870
```

```
#random slope models
```

```
fita1 <- lmer(Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Civil_Twilight + Weather_Condi
plotRESim(RESim(fita1))
```

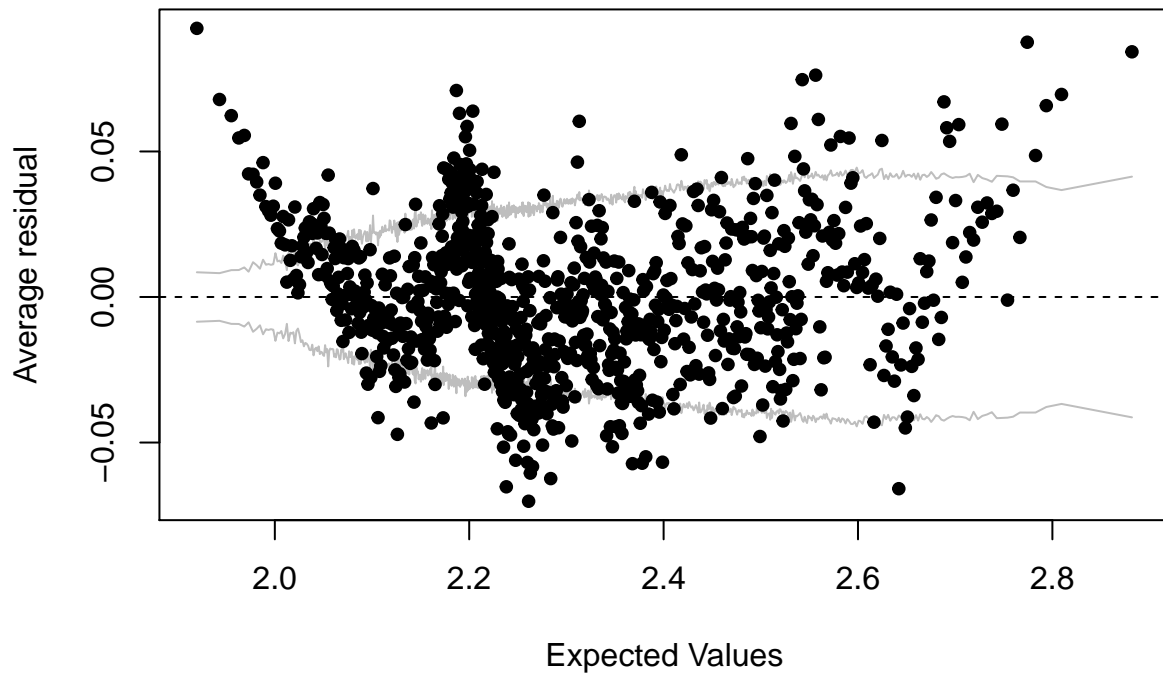


```
AIC(fita1)
```

```
## [1] 1165196
```

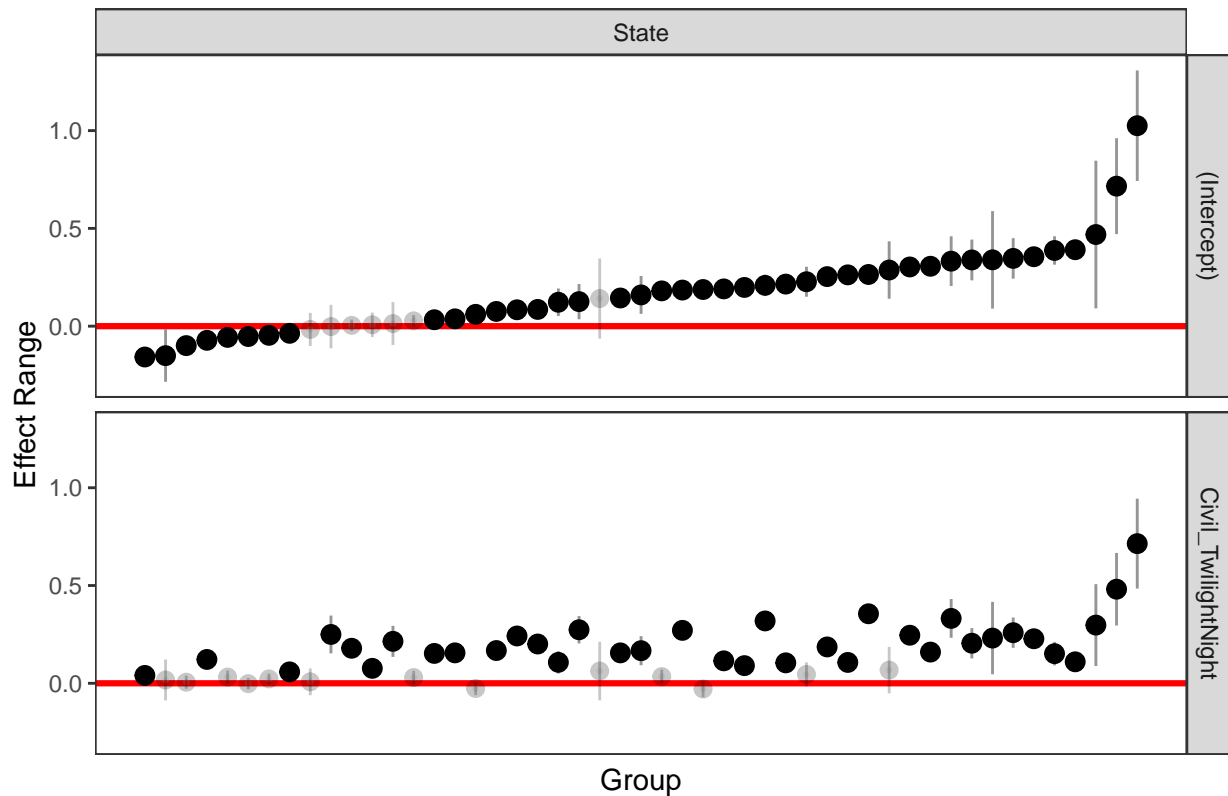
```
binnedplot(predict(fita1),resid(fita1))
```

Binned residual plot



```
fita2 <- lmer(Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal + Weather_Condi
plotREsim(REsim(fita2))
```

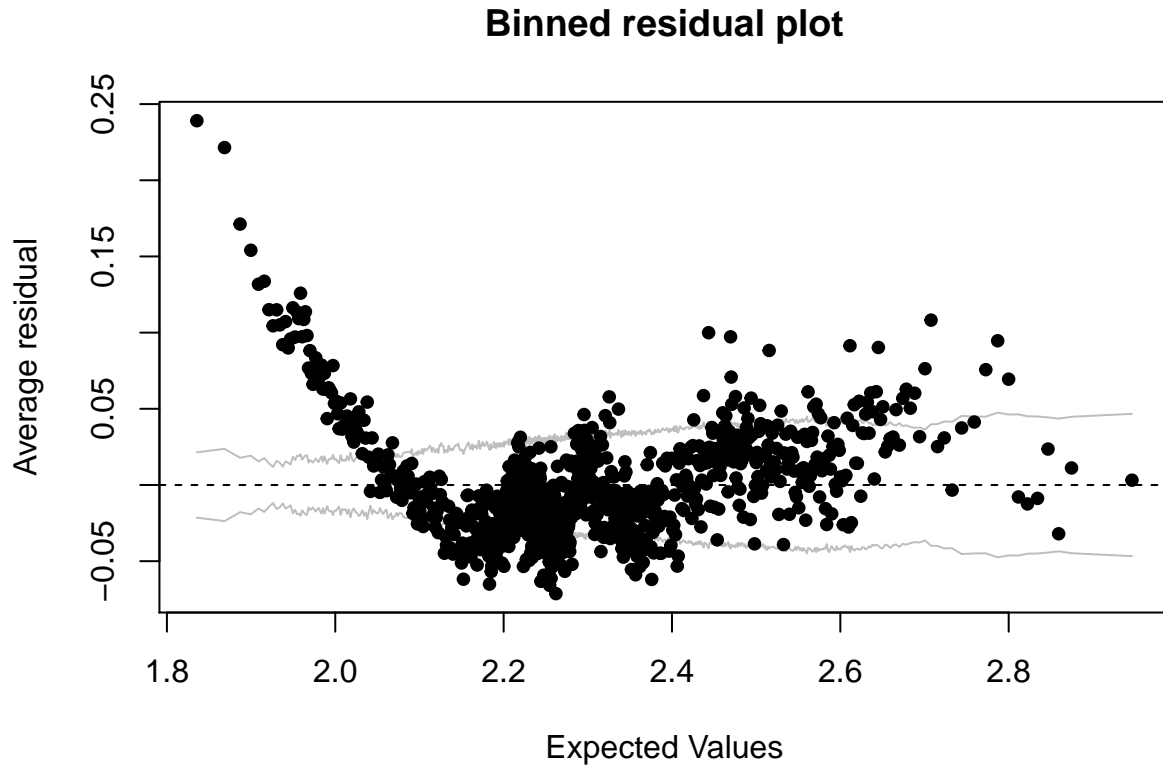
Effect Ranges



```

binnedplot(predict(fita2),resid(fita2))

```



```

AIC(fita1,fita2)

```

```

##      df      AIC
## fita1 15 1165196
## fita2 15 1166695

```

```

anova(fita,fitd,fita1,fita2)

```

```

## refitting model(s) with ML (instead of REML)
## Data: severity_data
## Models:
## fita: Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal +
## fita:      Civil_Twilight + Weather_Condition + (1 | State)
## fitd: Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal +
## fitd:      Civil_Twilight + Weather_Condition + Humidity...:Visibility.mi. +
## fitd:      (1 | State)
## fita1: Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Civil_Twilight +
## fita1:      Weather_Condition + (Traffic_Signal | State)
## fita2: Severity ~ Humidity... + Temperature.F. + Visibility.mi. + Traffic_Signal +
## fita2:      Weather_Condition + (Civil_Twilight | State)
##      npar      AIC      BIC logLik deviance      Chisq Df Pr(>Chisq)
## fita    14 1170703 1170866 -585337  1170675
## fitd    15 1170705 1170879 -585337  1170675    0.0902  1    0.7639
## fita1   15 1165063 1165238 -582517  1165033 5641.3720  0    <2e-16 ***
## fita2   15 1166562 1166737 -583266  1166532    0.0000  0    1.0000
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

```
#ordinal logistic regression
```

```
#fit1
```

```
fit1 <- polr(as.factor(Severity) ~ Humidity... + Temperature.F. + Visibility.mi. + as.factor(Weather_Condition), data = severity_data, Hess = TRUE)
```

```
## Call:
```

```
## polr(formula = as.factor(Severity) ~ Humidity... + Temperature.F. +  
##     Visibility.mi. + as.factor(Weather_Condition), data = severity_data,  
##     Hess = TRUE)
```

```
##
```

```
## Coefficients:
```

```
##                               Value Std. Error t value  
## Humidity...                   0.001382  0.0001304  10.595  
## Temperature.F.                0.001657  0.0001511  10.960  
## Visibility.mi.                 0.024319  0.0010881  22.349  
## as.factor(Weather_Condition)Cloudy -0.247437  0.0113183 -21.862  
## as.factor(Weather_Condition)Fair   -0.401182  0.0099726 -40.228  
## as.factor(Weather_Condition)Light Rain  0.009694  0.0139484   0.695  
## as.factor(Weather_Condition)Mostly Cloudy -0.104507  0.0108648  -9.619  
## as.factor(Weather_Condition)Overcast    0.020245  0.0128319   1.578  
## as.factor(Weather_Condition)Partly Cloudy -0.207235  0.0114629 -18.079
```

```
##
```

```
## Intercepts:
```

```
##      Value      Std. Error t value  
## 1|2   -8.2391    0.0763  -108.0432  
## 2|3    1.1725    0.0188   62.3279  
## 3|4    3.7205    0.0198  187.9102
```

```
##
```

```
## Residual Deviance: 1127757.12
```

```
## AIC: 1127781.12
```

```
coef_fit1 <- coef(summary(fit1))
```

```
pval <- pnorm(abs(coef_fit1[, "t value"]), lower.tail = FALSE) * 2
```

```
coef_fit1 <- cbind(coef_fit1, "p value" = pval)
```

```
coef_fit1
```

```
##                               Value      Std. Error  
## Humidity...                   0.001381793  0.0001304222  
## Temperature.F.                0.001656532  0.0001511379  
## Visibility.mi.                 0.024318637  0.0010881449  
## as.factor(Weather_Condition)Cloudy -0.247437334  0.0113183151  
## as.factor(Weather_Condition)Fair   -0.401181794  0.0099726107  
## as.factor(Weather_Condition)Light Rain  0.009694470  0.0139483827  
## as.factor(Weather_Condition)Mostly Cloudy -0.104506903  0.0108647920  
## as.factor(Weather_Condition)Overcast    0.020245185  0.0128319035  
## as.factor(Weather_Condition)Partly Cloudy -0.207235313  0.0114629047  
## 1|2   -8.239142445  0.0762578557  
## 2|3    1.172464765  0.0188112258  
## 3|4    3.720516118  0.0197994417  
##                               t value      p value  
## Humidity...                   10.5947684  3.151201e-26  
## Temperature.F.                10.9604032  5.923488e-28  
## Visibility.mi.                 22.3487126  1.242756e-110  
## as.factor(Weather_Condition)Cloudy -21.8616757  6.019348e-106
```

```
## as.factor(Weather_Condition)Fair          -40.2283622  0.000000e+00
## as.factor(Weather_Condition)Light Rain      0.6950247  4.870398e-01
## as.factor(Weather_Condition)Mostly Cloudy   -9.6188590  6.656539e-22
## as.factor(Weather_Condition)Overcast        1.5777227  1.146293e-01
## as.factor(Weather_Condition)Partly Cloudy  -18.0787783  4.683436e-73
## 1|2          -108.0431959  0.000000e+00
## 2|3           62.3279300  0.000000e+00
## 3|4          187.9101532  0.000000e+00
```

```
head(predict(fit1,severity_data,type = "p"))
```

```
##           1           2           3           4
## 1  0.0002612029  0.7613722  0.2144671  0.02389953
## 2  0.0002612029  0.7613722  0.2144671  0.02389953
## 14 0.0002660951  0.7647207  0.2115429  0.02347033
## 15 0.0002669046  0.7652656  0.2110667  0.02340080
## 16 0.0002669046  0.7652656  0.2110667  0.02340080
## 17 0.0002669046  0.7652656  0.2110667  0.02340080
```

```
#fit2
```

```
fit2 <- polr(as.factor(Severity) ~ as.factor(Traffic_Signal) + as.factor(Civil_Twilight), data = severity_data)
summary(fit2)
```

```
## Call:
```

```
## polr(formula = as.factor(Severity) ~ as.factor(Traffic_Signal) +
##       as.factor(Civil_Twilight), data = severity_data, Hess = TRUE)
##
```

```
## Coefficients:
```

```
##                               Value Std. Error t value
## as.factor(Traffic_Signal)True  -1.6271   0.010113 -160.90
## as.factor(Civil_Twilight)Night  0.1849   0.005933  31.17
##
```

```
## Intercepts:
```

```
##      Value      Std. Error t value
## 1|2  -8.9111    0.0743   -119.8904
## 2|3   0.8282    0.0029   282.1636
## 3|4   3.4101    0.0067   505.9757
##
```

```
## Residual Deviance: 1093123.77
```

```
## AIC: 1093133.77
```

```
coef_fit2 <- coef(summary(fit2))
```

```
pval2 <- pnorm(abs(coef_fit2[, "t value"]),lower.tail = FALSE)* 2
```

```
coef_fit2 <- cbind(coef_fit2, "p value" = pval2)
```

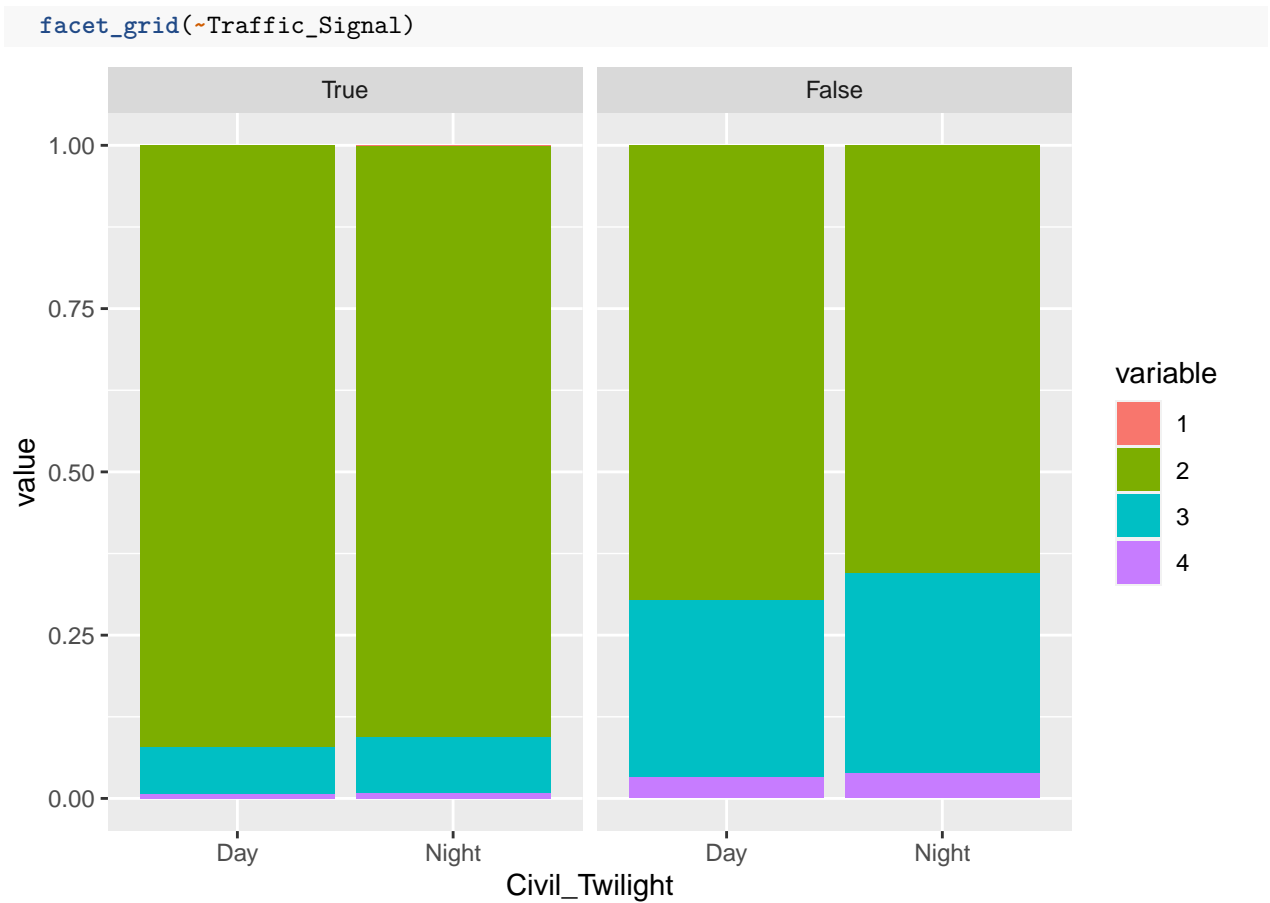
```
coef_fit2
```

```
##                               Value Std. Error  t value      p value
## as.factor(Traffic_Signal)True  -1.6270688  0.010112562 -160.89580  0.000000e+00
## as.factor(Civil_Twilight)Night  0.1849081  0.005932916  31.16648  3.033286e-213
## 1|2          -8.9110693  0.074326807 -119.89038  0.000000e+00
## 2|3           0.8281964  0.002935163  282.16364  0.000000e+00
## 3|4           3.4101413  0.006739734  505.97566  0.000000e+00
```

```
predx<- expand.grid(Traffic_Signal=c("True","False"),Civil_Twilight=c("Day","Night"))
```

```
predy<- predict (fit2,newdata=predx,type = "p")
```

```
ggplot(melt(cbind(predx,predy),id.vars = c("Traffic_Signal","Civil_Twilight")))+
  geom_bar(stat="identity")+aes(Civil_Twilight,y=value, fill=variable)+
```



```
#fit3
fit3 <- polr(as.factor(Severity) ~ Visibility.mi. + as.factor(Traffic_Signal) + as.factor(Civil_Twilight) + Humidity... + Temperature.F.,
summary(fit3)
```

```
## Call:
## polr(formula = as.factor(Severity) ~ Visibility.mi. + as.factor(Traffic_Signal) +
##       as.factor(Civil_Twilight) + Humidity... + Temperature.F.,
##       data = severity_data, Hess = TRUE)
##
## Coefficients:
##
##              Value Std. Error t value
## Visibility.mi.    0.018618  0.0010895  17.089
## as.factor(Traffic_Signal)True -1.630325  0.0101190 -161.115
## as.factor(Civil_Twilight)Night  0.183725  0.0061878  29.692
## Humidity...        0.001780  0.0001250  14.236
## Temperature.F.     0.001331  0.0001442   9.234
##
## Intercepts:
##      Value      Std. Error t value
## 1|2  -8.5236    0.0758  -112.5048
## 2|3   1.2009    0.0187   64.3308
## 3|4   3.7837    0.0197  192.2538
##
## Residual Deviance: 1092699.44
## AIC: 1092715.44
```



```
coef_fit3 <- coef(summary(fit3))
pval3 <- pnorm(abs(coef_fit3[, "t value"]), lower.tail = FALSE) * 2
coef_fit3 <- cbind(coef_fit3, "p value" = pval3)
coef_fit3
```

```
##              Value   Std. Error   t value
## Visibility.mi.      0.018618310 0.0010895018  17.088829
## as.factor(Traffic_Signal)True -1.630325054 0.0101189920 -161.115362
## as.factor(Civil_Twilight)Night 0.183725489 0.0061877570  29.691775
## Humidity...         0.001779991 0.0001250337  14.236089
## Temperature.F.      0.001331136 0.0001441570   9.233933
## 1|2                 -8.523601502 0.0757620883 -112.504838
## 2|3                  1.200868933 0.0186670893  64.330808
## 3|4                  3.783682020 0.0196806651 192.253768
##              p value
## Visibility.mi.      1.797480e-65
## as.factor(Traffic_Signal)True 0.000000e+00
## as.factor(Civil_Twilight)Night 9.805716e-194
## Humidity...         5.470419e-46
## Temperature.F.      2.608738e-20
## 1|2                 0.000000e+00
## 2|3                 0.000000e+00
## 3|4                 0.000000e+00
```

```
#fit4
fit4 <- polr(as.factor(Severity) ~ as.factor(Weather_Condition) + Visibility.mi. + as.factor(Traffic_Si,
summary(fit4)
```

```
## Call:
## polr(formula = as.factor(Severity) ~ as.factor(Weather_Condition) +
##      Visibility.mi. + as.factor(Traffic_Signal) + as.factor(Civil_Twilight) +
##      Humidity... + Temperature.F., data = severity_data, Hess = TRUE)
##
## Coefficients:
##              Value Std. Error t value
## as.factor(Weather_Condition)Cloudy      -0.2847280 0.0115397 -24.674
## as.factor(Weather_Condition)Fair        -0.4560951 0.0101869 -44.773
## as.factor(Weather_Condition)Light Rain  -0.0151657 0.0142273  -1.066
## as.factor(Weather_Condition)Mostly Cloudy -0.1082874 0.0110985  -9.757
## as.factor(Weather_Condition)Overcast      0.0268230 0.0130978   2.048
## as.factor(Weather_Condition)Partly Cloudy -0.2279688 0.0117038 -19.478
## Visibility.mi.         0.0252173 0.0011135  22.646
## as.factor(Traffic_Signal)True           -1.6439512 0.0101378 -162.161
## as.factor(Civil_Twilight)Night          0.2191403 0.0062327  35.160
## Humidity...             0.0007702 0.0001345   5.727
## Temperature.F.          0.0042847 0.0001586  27.019
##
## Intercepts:
##      Value      Std. Error t value
## 1|2    -8.6299    0.0765 -112.7488
## 2|3     1.1350    0.0194  58.6414
## 3|4     3.7258    0.0203 183.2223
##
## Residual Deviance: 1088187.36
```

```
## AIC: 1088215.36
```

```
coef_fit4 <- coef(summary(fit4))
pval4 <- pnorm(abs(coef_fit4[, "t value"]), lower.tail = FALSE) * 2
coef_fit4 <- cbind(coef_fit4, "p value" = pval4)
coef_fit4
```

```
##                               Value   Std. Error
## as.factor(Weather_Condition)Cloudy   -0.2847280087 0.0115396550
## as.factor(Weather_Condition)Fair     -0.4560951467 0.0101869039
## as.factor(Weather_Condition)Light Rain -0.0151656958 0.0142273295
## as.factor(Weather_Condition)Mostly Cloudy -0.1082873944 0.0110985371
## as.factor(Weather_Condition)Overcast    0.0268229869 0.0130977558
## as.factor(Weather_Condition)Partly Cloudy -0.2279687682 0.0117038275
## Visibility.mi.                        0.0252172825 0.0011135474
## as.factor(Traffic_Signal)True         -1.6439511769 0.0101377870
## as.factor(Civil_Twilight)Night         0.2191402966 0.0062326658
## Humidity...                          0.0007701575 0.0001344799
## Temperature.F.                       0.0042847291 0.0001585813
## 1|2                                   -8.6298861816 0.0765407995
## 2|3                                    1.1349536763 0.0193541523
## 3|4                                    3.7257946775 0.0203348310
##                               t value    p value
## as.factor(Weather_Condition)Cloudy   -24.673875 2.040449e-134
## as.factor(Weather_Condition)Fair     -44.772696 0.000000e+00
## as.factor(Weather_Condition)Light Rain  -1.065955 2.864439e-01
## as.factor(Weather_Condition)Mostly Cloudy  -9.756907 1.723309e-22
## as.factor(Weather_Condition)Overcast    2.047907 4.056910e-02
## as.factor(Weather_Condition)Partly Cloudy -19.478138 1.682890e-84
## Visibility.mi.                        22.645900 1.531145e-113
## as.factor(Traffic_Signal)True         -162.160753 0.000000e+00
## as.factor(Civil_Twilight)Night         35.159963 8.187110e-271
## Humidity...                          5.726935 1.022614e-08
## Temperature.F.                       27.019130 8.809281e-161
## 1|2                                   -112.748838 0.000000e+00
## 2|3                                    58.641353 0.000000e+00
## 3|4                                    183.222308 0.000000e+00
```

```
residuals(fit1)
```

```
## NULL
```

```
#fit5
```

```
fit5 <- polr(as.factor(Severity) ~ as.factor(Weather_Condition) + Visibility.mi. + as.factor(Traffic_Si,
summary(fit5)
```

```
## Warning in sqrt(diag(vc)): NaNs produced
```

```
## Call:
```

```
## polr(formula = as.factor(Severity) ~ as.factor(Weather_Condition) +
##       Visibility.mi. + as.factor(Traffic_Signal) + as.factor(Civil_Twilight) +
##       Humidity... * Temperature.F. + Temperature.F., data = severity_data,
##       Hess = TRUE)
```

```
##
```

```
## Coefficients:
```

```
##                               Value Std. Error   t value
## as.factor(Weather_Condition)Cloudy   -0.2637832 7.388e-04  -357.05
```

```
## as.factor(Weather_Condition)Fair      -0.4268028  4.261e-03  -100.17
## as.factor(Weather_Condition)Light Rain  0.0276793  1.555e-04   178.05
## as.factor(Weather_Condition)Mostly Cloudy -0.0975569  1.847e-03   -52.83
## as.factor(Weather_Condition)Overcast    0.0947423  1.755e-04   539.73
## as.factor(Weather_Condition)Partly Cloudy -0.2060611  1.083e-03  -190.34
## Visibility.mi.                        0.0206924  1.057e-03    19.59
## as.factor(Traffic_Signal)True          -1.6489776  9.055e-05 -18211.29
## as.factor(Civil_Twilight)Night         0.2240829  8.214e-04   272.82
## Humidity...                          -0.0150242      NaN      NaN
## Temperature.F.                      -0.0105750  3.638e-05  -290.70
## Humidity...:Temperature.F.           0.0002401      NaN      NaN
##
## Intercepts:
##      Value      Std. Error  t value
## 1|2      -9.6668      0.0001 -72016.6155
## 2|3       0.1054      0.0147   7.1720
## 3|4       2.6984      0.0173  155.7010
##
## Residual Deviance: 1086854.91
## AIC: 1086884.91
```

```
coef_fit5 <- coef(summary(fit5))
```

```
## Warning in sqrt(diag(vc)): NaNs produced
```

```
pval5 <- pnorm(abs(coef_fit5[, "t value"]),lower.tail = FALSE)* 2
coef_fit5 <- cbind(coef_fit5, "p value" = pval5)
coef_fit5
```

```
##                               Value  Std. Error
## as.factor(Weather_Condition)Cloudy  -0.2637832437  7.387787e-04
## as.factor(Weather_Condition)Fair    -0.4268027538  4.260746e-03
## as.factor(Weather_Condition)Light Rain  0.0276793010  1.554556e-04
## as.factor(Weather_Condition)Mostly Cloudy -0.0975569329  1.846781e-03
## as.factor(Weather_Condition)Overcast    0.0947422669  1.755352e-04
## as.factor(Weather_Condition)Partly Cloudy -0.2060611409  1.082620e-03
## Visibility.mi.                      0.0206924406  1.056541e-03
## as.factor(Traffic_Signal)True        -1.6489776095  9.054701e-05
## as.factor(Civil_Twilight)Night       0.2240829466  8.213691e-04
## Humidity...                        -0.0150242379      NaN
## Temperature.F.                    -0.0105750460  3.637777e-05
## Humidity...:Temperature.F.          0.0002400564      NaN
## 1|2                        -9.6668133870  1.342303e-04
## 2|3                         0.1054374693  1.470129e-02
## 3|4                         2.6984482434  1.733096e-02
##                               t value      p value
## as.factor(Weather_Condition)Cloudy    -357.053118  0.000000e+00
## as.factor(Weather_Condition)Fair      -100.170895  0.000000e+00
## as.factor(Weather_Condition)Light Rain   178.052737  0.000000e+00
## as.factor(Weather_Condition)Mostly Cloudy -52.825407  0.000000e+00
## as.factor(Weather_Condition)Overcast     539.733783  0.000000e+00
## as.factor(Weather_Condition)Partly Cloudy -190.335613  0.000000e+00
## Visibility.mi.                        19.585085  2.072611e-85
## as.factor(Traffic_Signal)True        -18211.286861  0.000000e+00
## as.factor(Civil_Twilight)Night         272.816377  0.000000e+00
```

```
## Humidity...           NaN           NaN
## Temperature.F.      -290.700764 0.000000e+00
## Humidity...:Temperature.F.      NaN           NaN
## 1|2                 -72016.615508 0.000000e+00
## 2|3                   7.171987 7.391711e-13
## 3|4                  155.700996 0.000000e+00
```

```
head(predict(fit5,severity_data,type = "p"))
```

```
##           1           2           3           4
## 1  0.0002086598 0.7852312 0.1945375 0.02002266
## 2  0.0002086598 0.7852312 0.1945375 0.02002266
## 14 0.0001216336 0.6807706 0.2852416 0.03386616
## 15 0.0001213234 0.6802158 0.2857131 0.03394982
## 16 0.0001213234 0.6802158 0.2857131 0.03394982
## 17 0.0001213234 0.6802158 0.2857131 0.03394982
```

```
fit6 <- polr(as.factor(Severity) ~ as.factor(Weather_Condition) + Visibility.mi. + as.factor(Traffic_Si,
AIC(fit1,fit2,fit3,fit4,fit5,fit6))
```

```
##      df      AIC
## fit1 12 1127781
## fit2  5 1093134
## fit3  8 1092715
## fit4 14 1088215
## fit5 15 1086885
## fit6 15 1088149
```

##Package reference dplyr: <https://cran.r-project.org/web/packages/dplyr/index.html> tidyr: <https://cran.r-project.org/web/packages/tidyr/index.html> funModeling: <https://cran.r-project.org/web/packages/funModeling/index.html> Hmisc: <https://cran.r-project.org/web/packages/Hmisc/index.html> MASS: <https://cran.r-project.org/web/packages/MASS/index.html> magrittr: <https://cran.r-project.org/web/packages/magrittr/index.html> ggplot2: <https://cran.r-project.org/web/packages/ggplot2/index.html> easyGgplot2: <https://github.com/kassambara/easyGgplot2> reshape2: <https://cran.r-project.org/web/packages/reshape2/index.html> Matrix: <https://cran.r-project.org/web/packages/Matrix/index.html> lme4: <https://cran.r-project.org/web/packages/lme4/index.html> arm: <https://cran.r-project.org/web/packages/arm/index.html> sjstats: <https://cran.r-project.org/web/packages/sjstats/index.html> HLMdiag: <https://cran.r-project.org/web/packages/HLMdiag/index.html>