

VARIABLE INVESTIGATION FOR KELOWNA WEATHER-CRASH PROJECT

Contents

1	Summary Statistics	2
2	Accidents over Time	4
2.1	Throughout the Day	4
2.2	Day of the Week	5
2.3	Month of the Year	6
2.4	Over the Years	7
3	Distributions of Weather Variables	8
3.1	Temperature	8
3.2	Relative Humidity	10
3.3	Precipitation	12
3.4	Wind Speed	13
3.5	Wind Direction	15
3.6	Visibility	17
3.7	Air Pressure	18
4	Crash Subtypes	19
4.1	All Combined	19
4.2	Cyclist Crashes	21
4.3	Animal Crashes	23
4.4	Heavy Vehicle Crashes	25
4.5	Motorcycle Crashes	27
4.6	Parked Vehicle Crashes	29
4.7	Pedestrian Crashes	31
5	Regression/Neural Network Dataset	33
5.1	Weather Variables	34
5.1.1	Temperature	34
5.1.2	Relative Humidity	35
5.1.3	Precipitation	36
5.1.4	Wind Direction	37
5.1.5	Wind Speed	38
5.1.6	Visibility	39
5.1.7	Air Pressure	40
5.2	Crash Variables	41
5.2.1	Number of Crashes	41
5.2.2	Number of Deaths	42
5.2.3	Locations	43

1 Summary Statistics

Need to remove dew point temp and wind chill,
since they are so correlated with temperature.

	Temp...C.	Dew.Point.Temp...C.	Wind.Chill
Temp...C.	1.0000000	0.9209662	0.9517406
Dew.Point.Temp...C.	0.9209662	1.0000000	0.9312014
Wind.Chill	0.9517406	0.9312014	1.0000000

Summary of all variables:

linker	Date.Of.Loss.Year	Animal.Flag	Crash.Severity
Length:56136	Min. :2017	No :54118	CASUALTY CRASH :11473
Class :character	1st Qu.:2018	Yes: 2018	PROPERTY DAMAGE ONLY:44663
Mode :character	Median :2019		
	Mean :2019		
	3rd Qu.:2020		
	Max. :2021		

Cyclist.Flag	Day.Of.Week	Derived.Crash.Configuration
No :55725	FRIDAY :9316	REAR END :13024
Yes: 411	MONDAY :8024	SINGLE VEHICLE :12495
	SATURDAY :6753	UNDETERMINED :11453
	SUNDAY :5464	SIDE IMPACT :11369
	THURSDAY :9061	CONFLICTED : 3068
	TUESDAY :8728	SIDE SWIPE - SAME DIRECTION: 1833
	WEDNESDAY:8790	(Other) : 2894

Heavy.Veh.Flag	Intersection.Crash	Month.Of.Year	Motorcycle.Flag
No :54085	No :31677	JULY : 5152	No :55673
Yes: 2051	Yes:24459	DECEMBER: 5095	Yes: 463
		JANUARY : 5063	
		AUGUST : 4907	
		OCTOBER : 4836	
		JUNE : 4735	
		(Other) :26348	

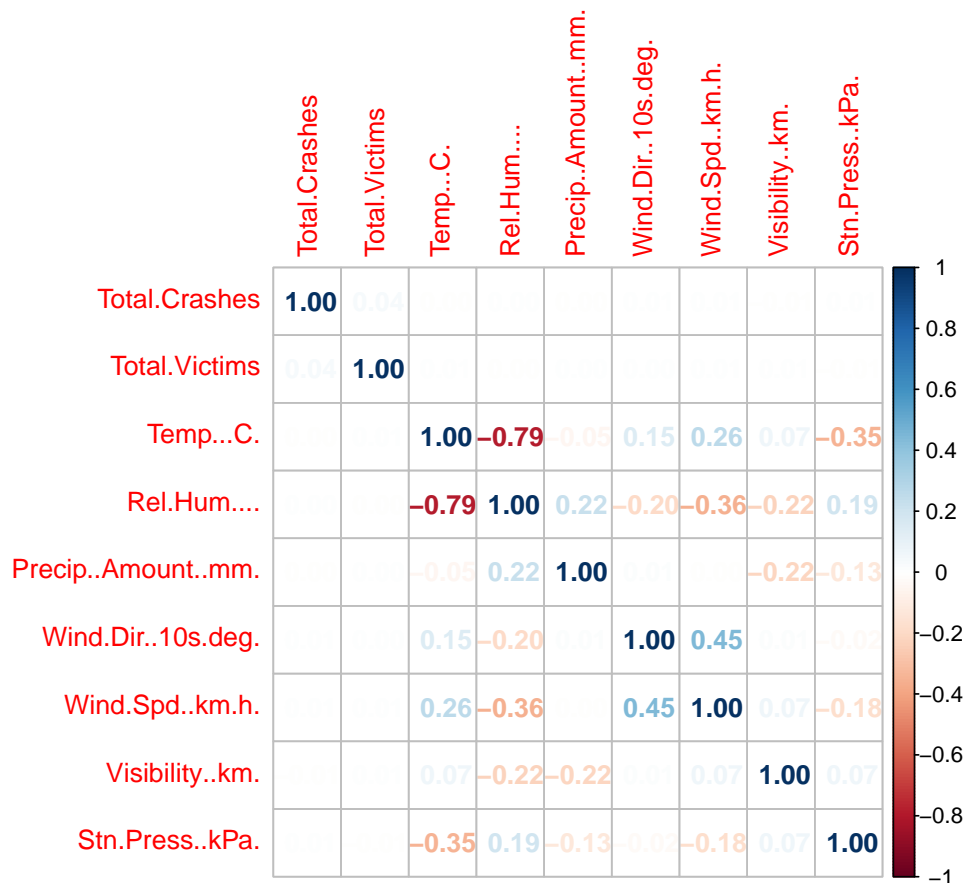
Parked.Vehicle.Flag	Parking.Lot.Flag	Pedestrian.Flag	Time.Category
No :38567	No :37189	No :55790	12:00-14:59:14870
Yes:17569	Yes:18947	Yes: 346	15:00-17:59:14473
			09:00-11:59:11021
			18:00-20:59: 5805
			06:00-08:59: 5489
			21:00-23:59: 2684
			(Other) : 1794

Municipality.Name	Road.Location.Description	Street.Full.Name
KELOWNA :45943	UNKNOWN : 2211	HWY 97 : 6019
WEST KELOWNA:10193	HWY 97 : 1961	HARVEY AVE : 3168
	HARVEY AVE : 1670	HWY 33 : 2064
	LAKESHORE RD: 932	GORDON DR : 1921
	LOUIE DR : 751	LAKESHORE RD : 1372
	HWY 33 : 715	SPRINGFIELD RD: 1326
	(Other) :47896	(Other) :40266

Total.Crashes	Total.Victims	Temp...C.	Rel.Hum....
Min. :1.000	Min. :0.0000	Min. : -22.43	Min. : 12.67
1st Qu.:1.000	1st Qu.:0.0000	1st Qu.: 1.60	1st Qu.: 38.00
Median :1.000	Median :0.0000	Median : 10.37	Median : 62.00
Mean :1.006	Mean :0.2917	Mean : 10.92	Mean : 60.51
3rd Qu.:1.000	3rd Qu.:0.0000	3rd Qu.: 20.00	3rd Qu.: 82.33
Max. :3.000	Max. :9.0000	Max. : 37.73	Max. :100.00
	NA's :35	NA's :35	

Precip..Amount..mm.	Wind.Dir..10s.deg.	Wind.Spd..km.h.	Visibility..km.
Min. : 0.00000	Min. : 1.00	Min. : 0.000	Min. : 0.20
1st Qu.: 0.00000	1st Qu.:15.67	1st Qu.: 4.667	1st Qu.:16.10
Median : 0.00000	Median :19.50	Median : 8.000	Median :16.10
Mean : 0.09327	Mean :21.16	Mean :10.277	Mean :15.15
3rd Qu.: 0.00000	3rd Qu.:31.00	3rd Qu.:13.667	3rd Qu.:16.10
Max. :12.00000	Max. :36.00	Max. :51.333	Max. :16.10
	NA's :4020	NA's :35	NA's :36

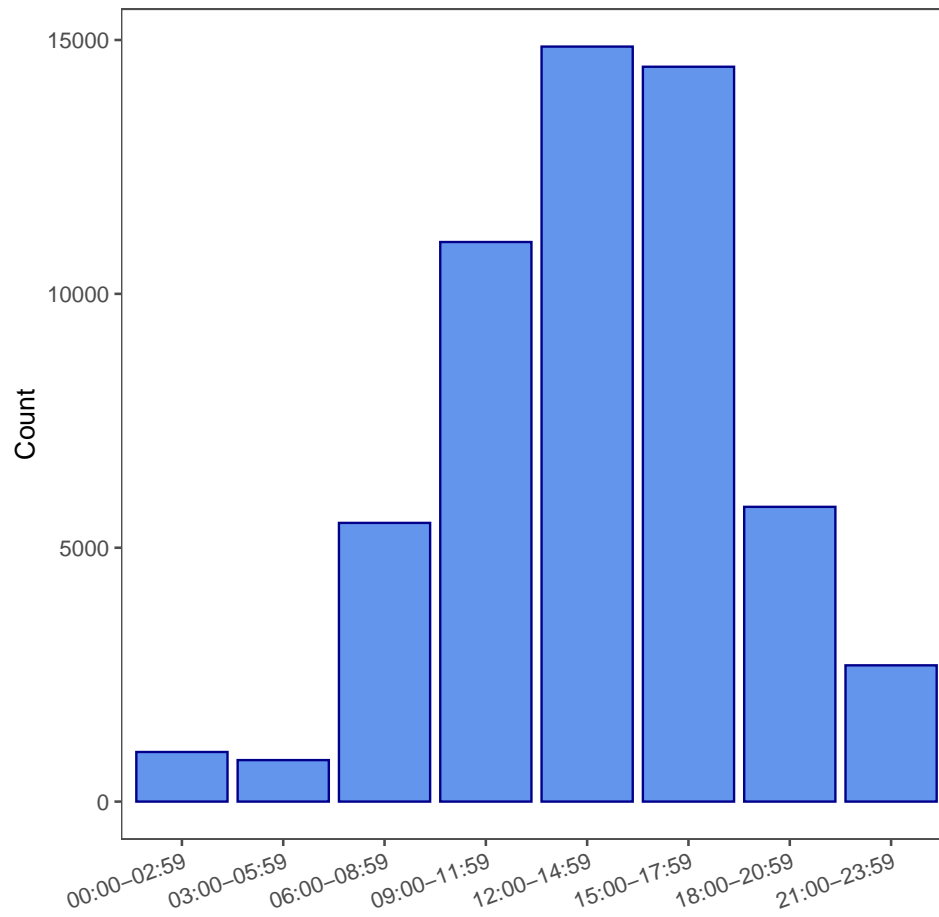
Stn.Press..kPa.	Fog	Freezing.Rain	Snow	Rain	Thunderstorms
Min. :94.08	0:52106	0:56134	0:52325	0:50075	0:56002
1st Qu.:96.10	1: 4030	1: 2	1: 3811	1: 6061	1: 134
Median :96.51					
Mean :96.56					
3rd Qu.:96.90					
Max. :99.33					
NA's :35					



2 Accidents over Time

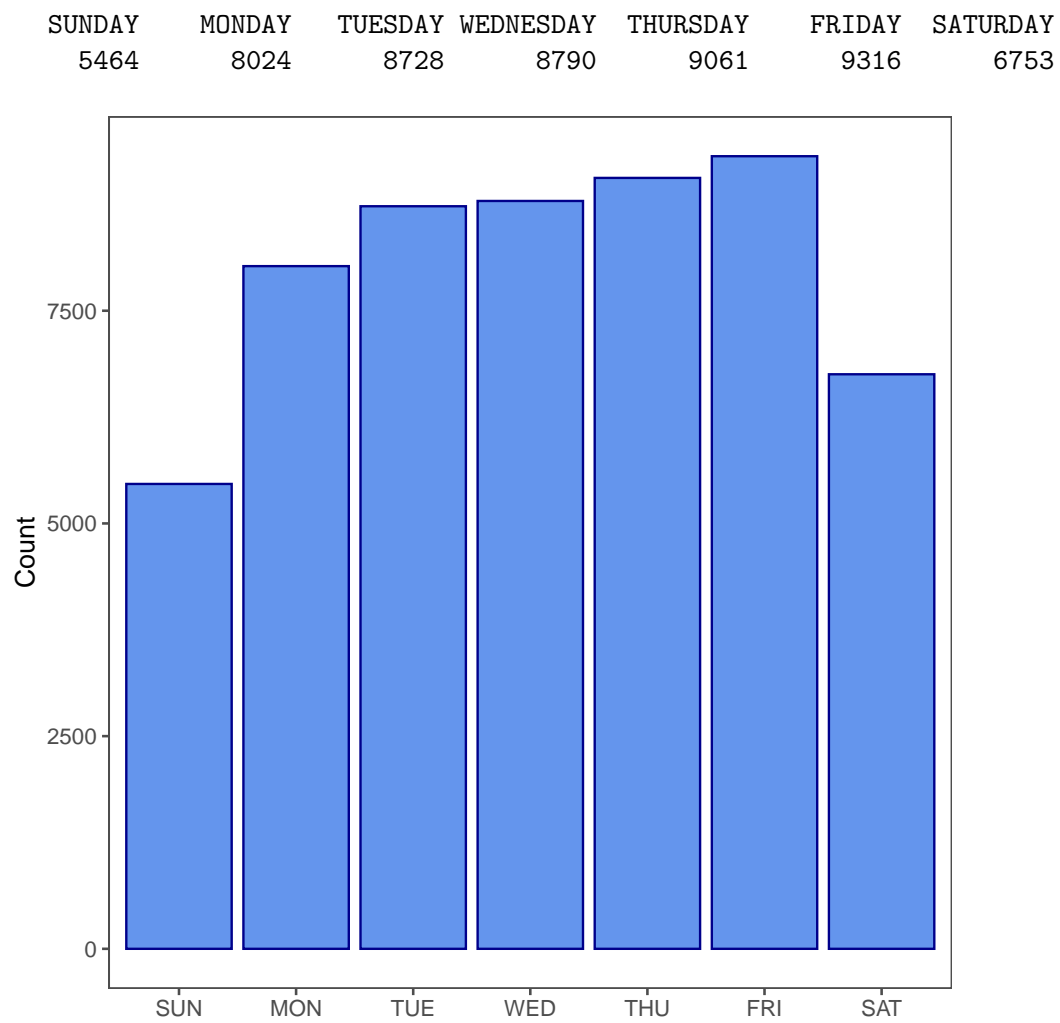
2.1 Throughout the Day

00:00-02:59	03:00-05:59	06:00-08:59	09:00-11:59	12:00-14:59	15:00-17:59
977	817	5489	11021	14870	14473
18:00-20:59	21:00-23:59				
5805	2684				



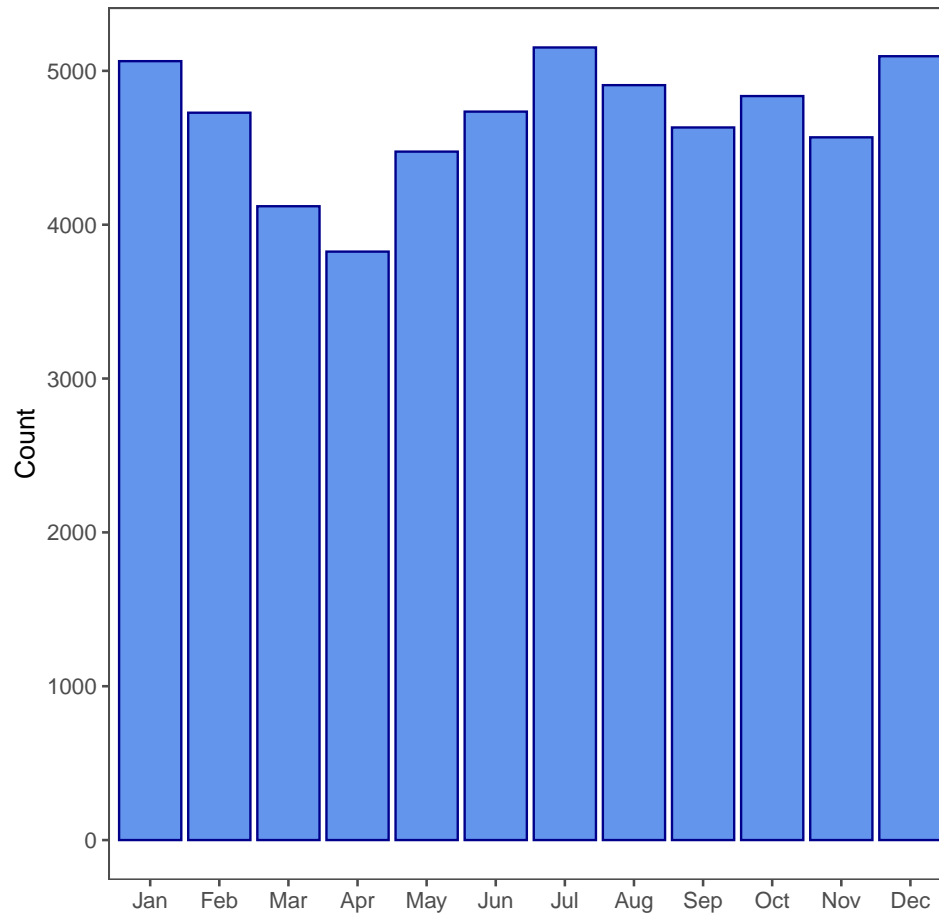
2.2 Day of the Week

Distribution of accidents throughout the week:



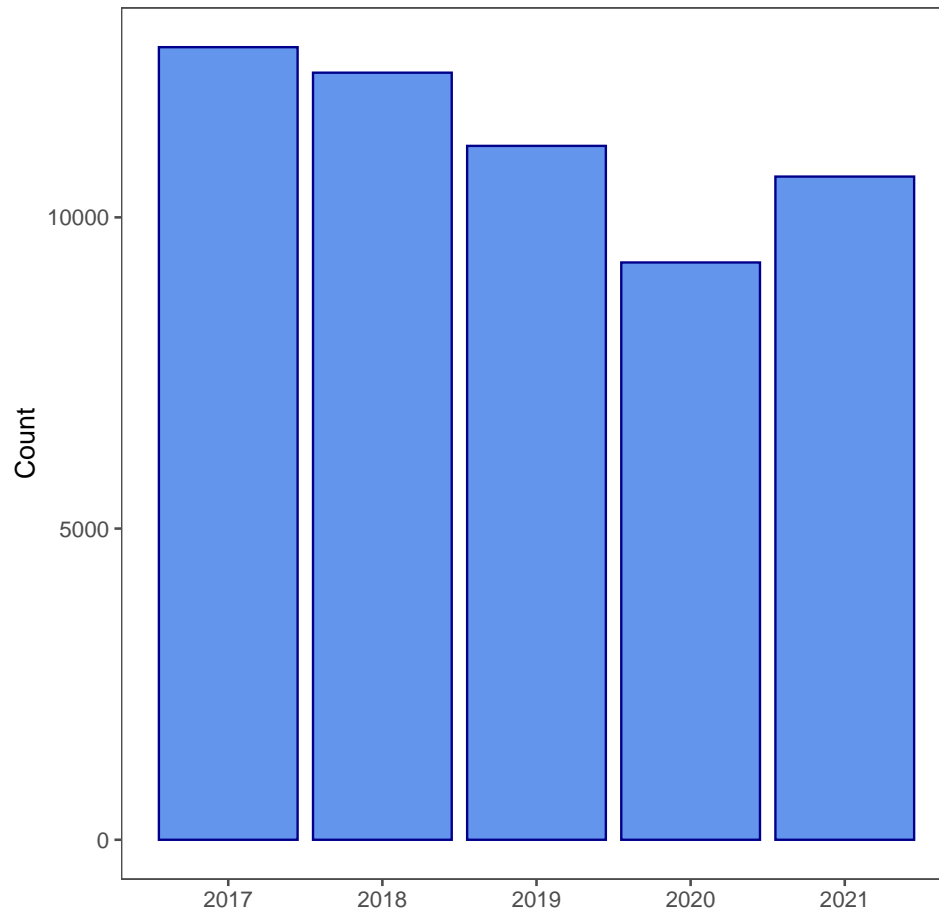
2.3 Month of the Year

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
5063	4728	4120	3825	4475	4735	5152	4907
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
4632	4836	4568	5095				



2.4 Over the Years

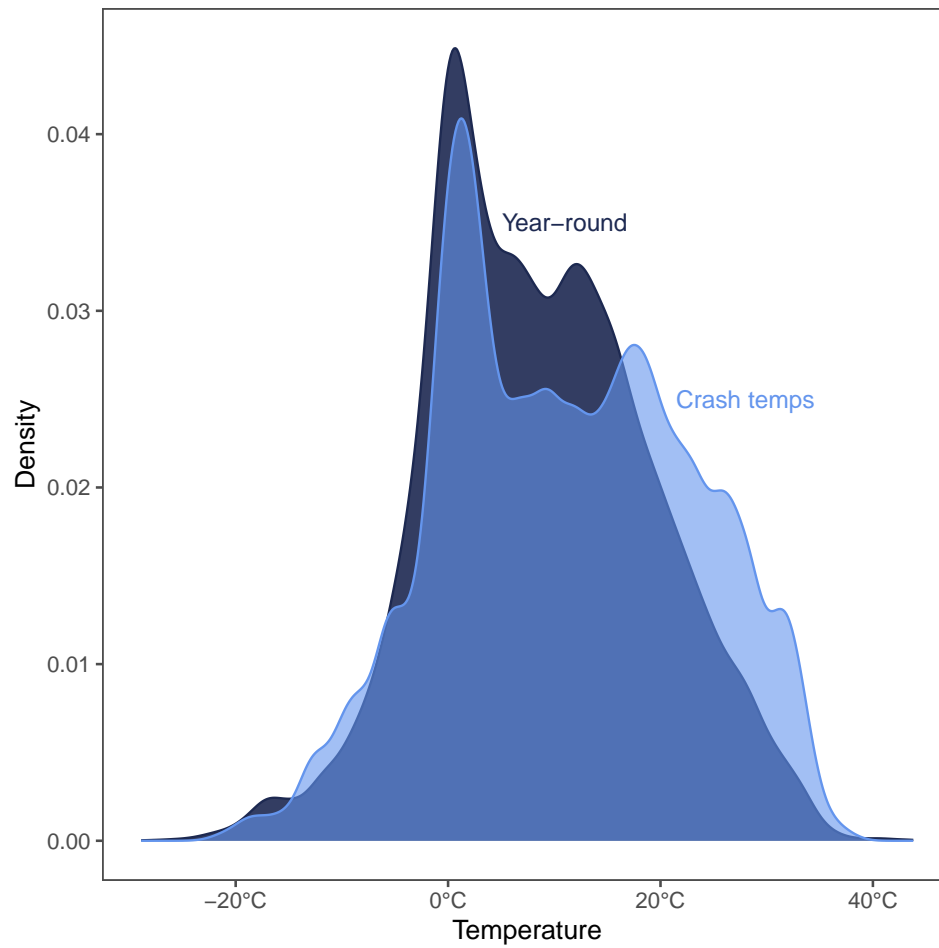
2017	2018	2019	2020	2021
12734	12324	11148	9275	10655

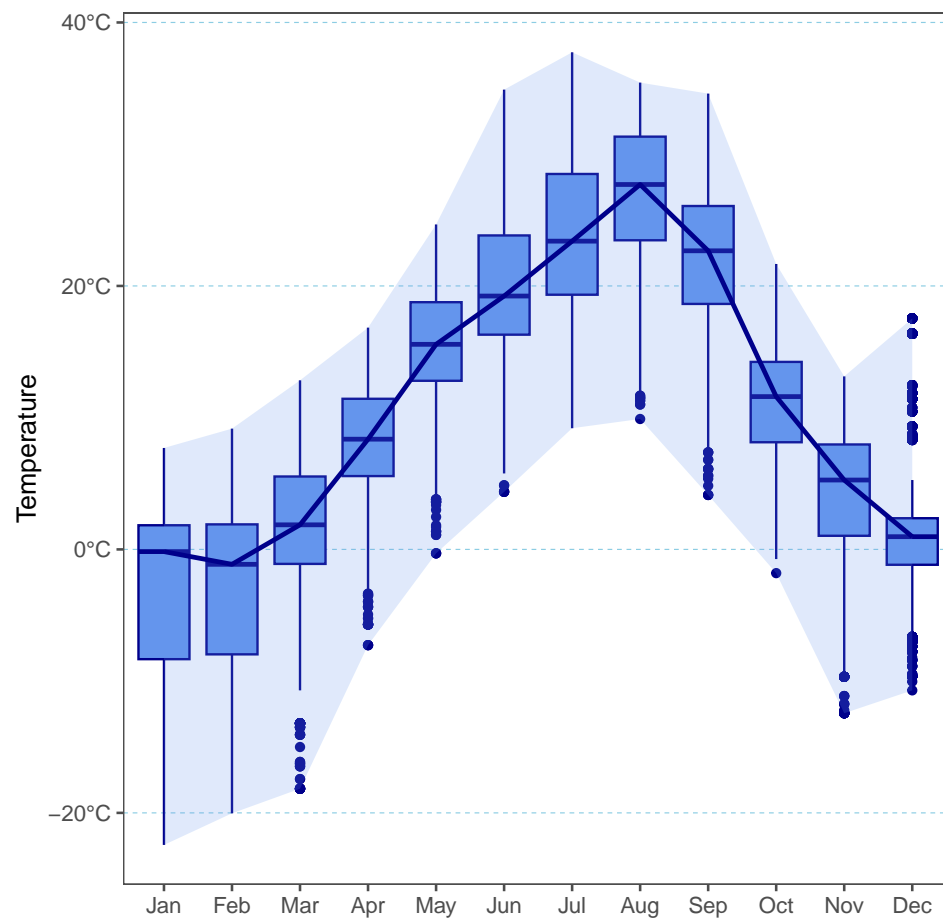


3 Distributions of Weather Variables

3.1 Temperature

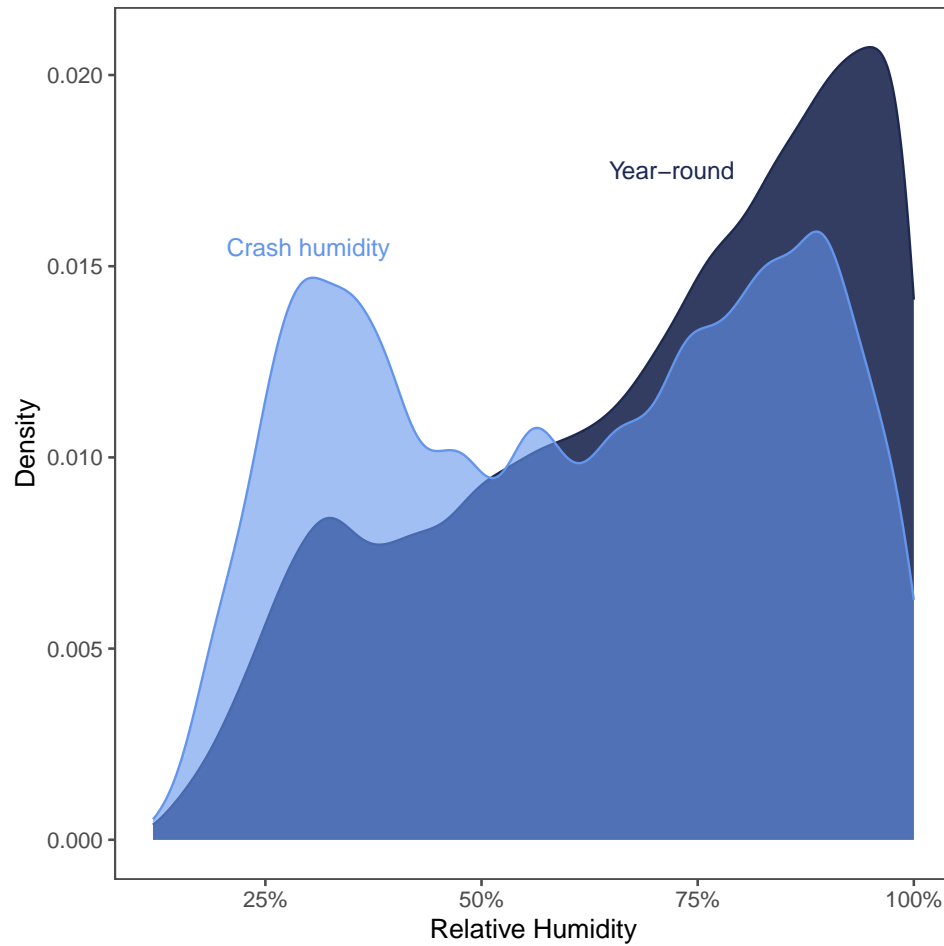
	Avg.temp.	Avg.crash.temp.
1	8.511442	10.91872

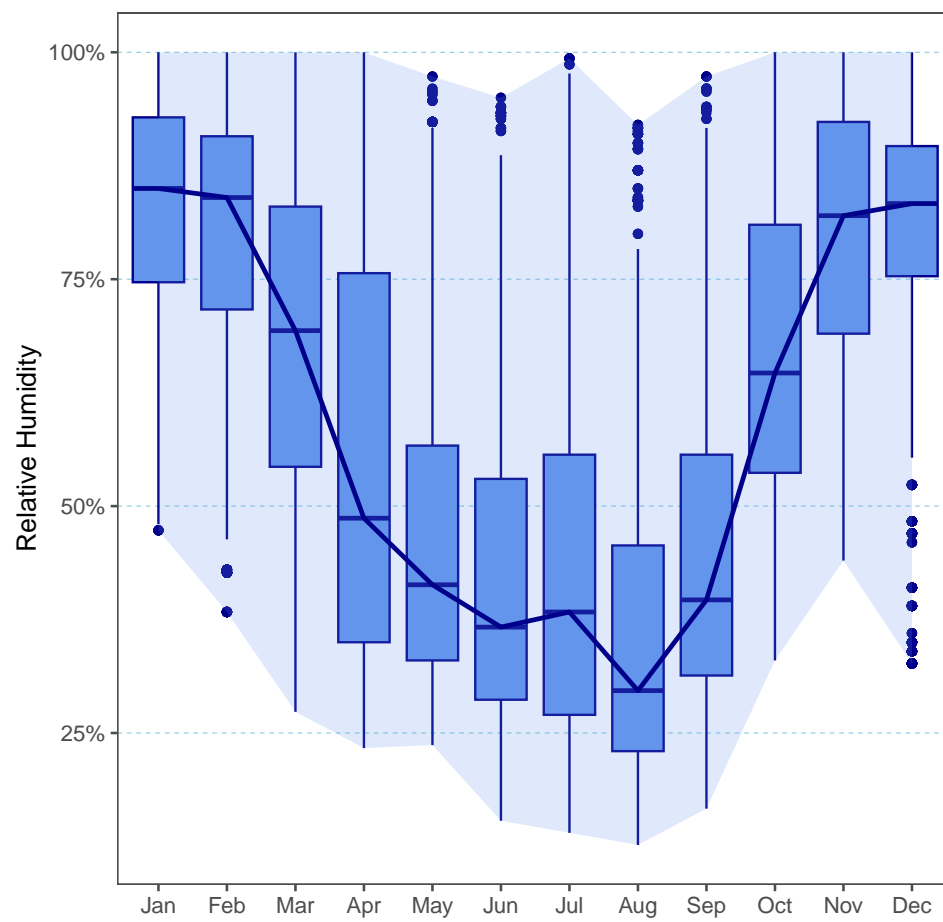




3.2 Relative Humidity

	Avg.humidity	Avg.crash.humidity
1	69.32555	60.50743





3.3 Precipitation

```

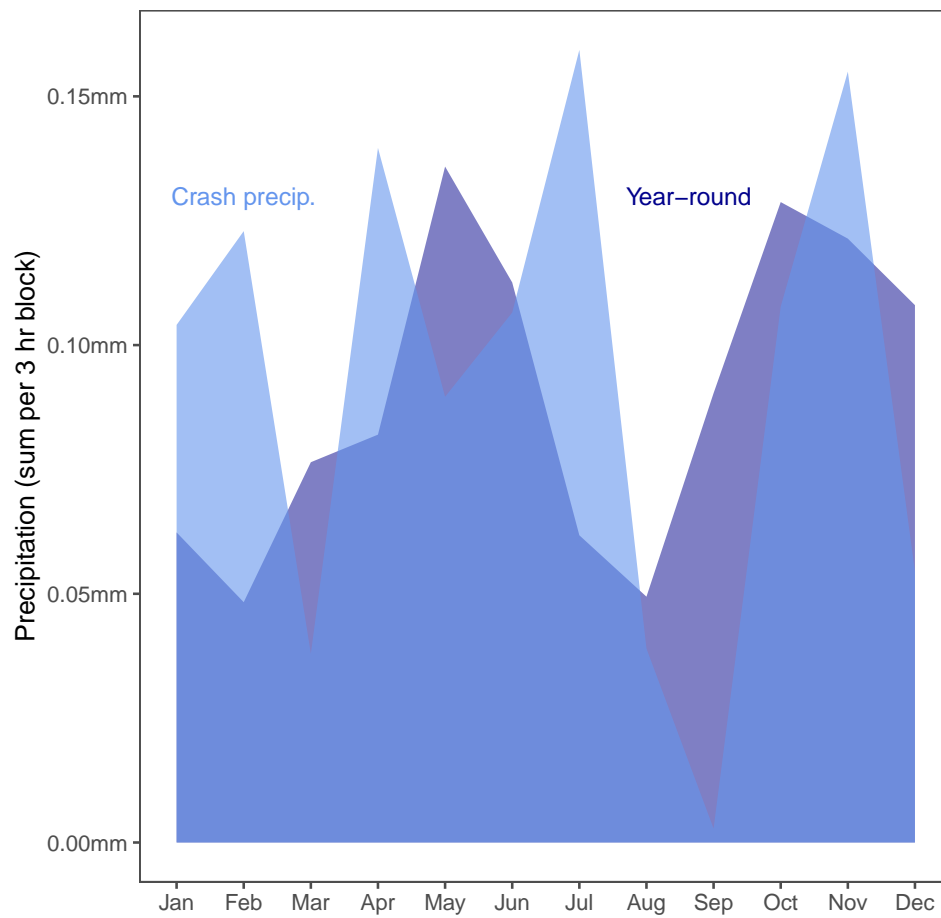
Avg.precipitation Avg.crash.precipitation
1      0.08996575      0.09326635

```

Rounded table of precipitation values:

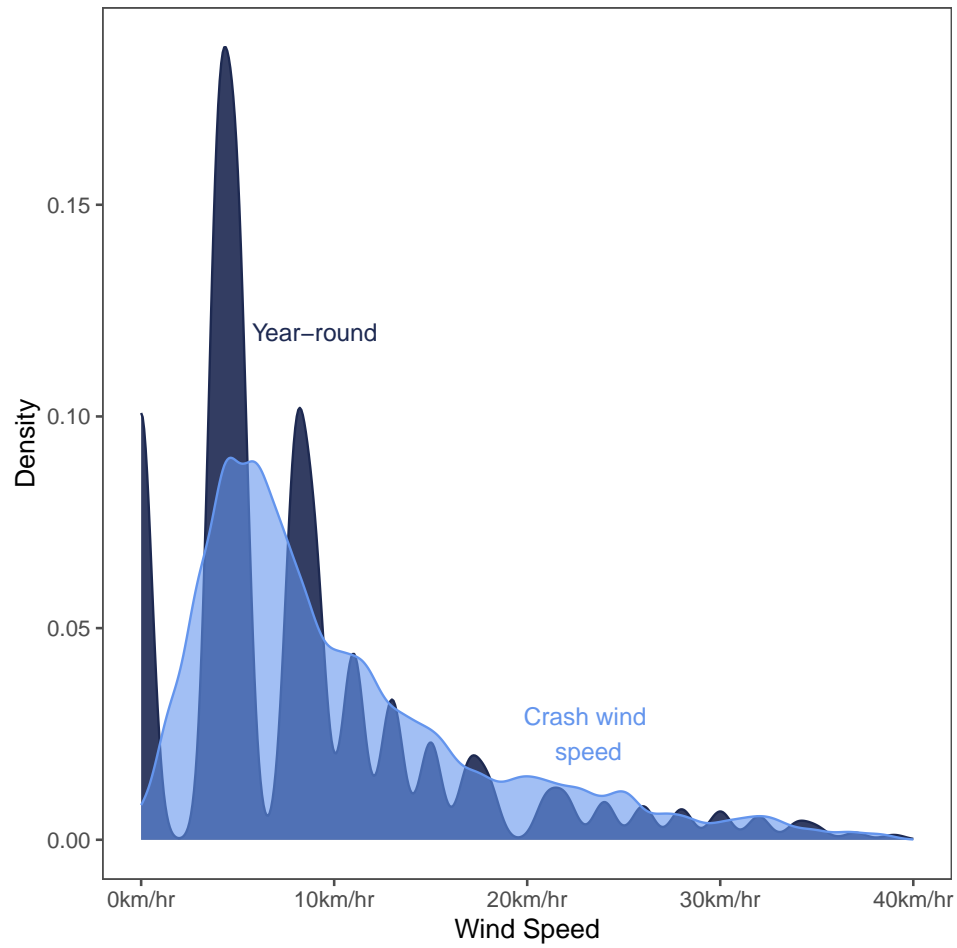
	0	1	2	3	4	5	6	7	8	12
53658	1174	781	263	206	11	34	1	1	1	7

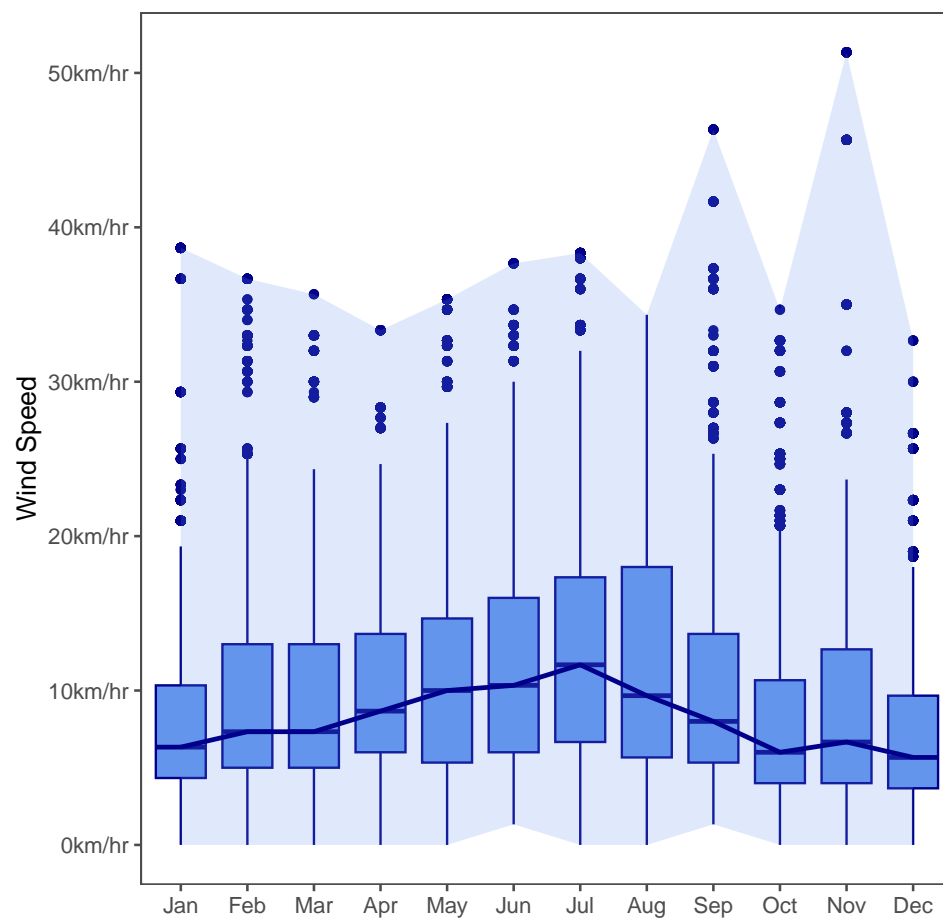
Many are 0...



3.4 Wind Speed

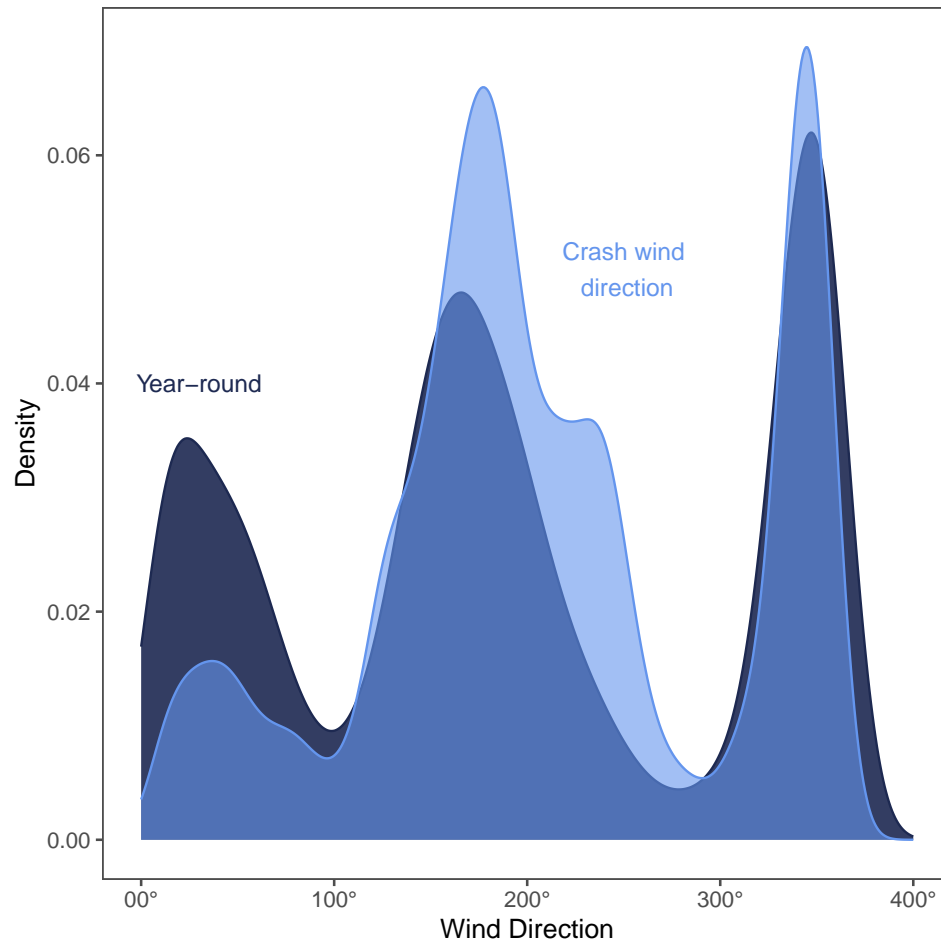
	Avg.wind.speed	Avg.crash.wind.speed
1	8.41538	10.27689

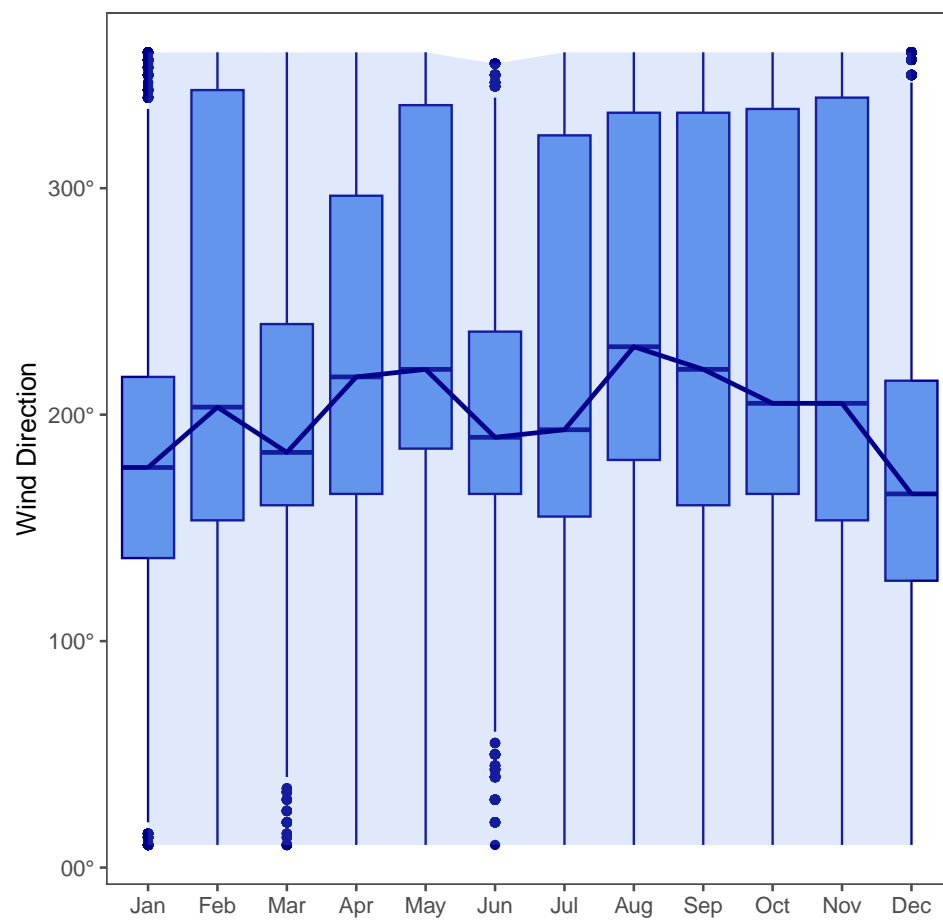




3.5 Wind Direction

	Avg.wind.dir.	Avg.crash.wind.dir.
1	19.1302	21.163



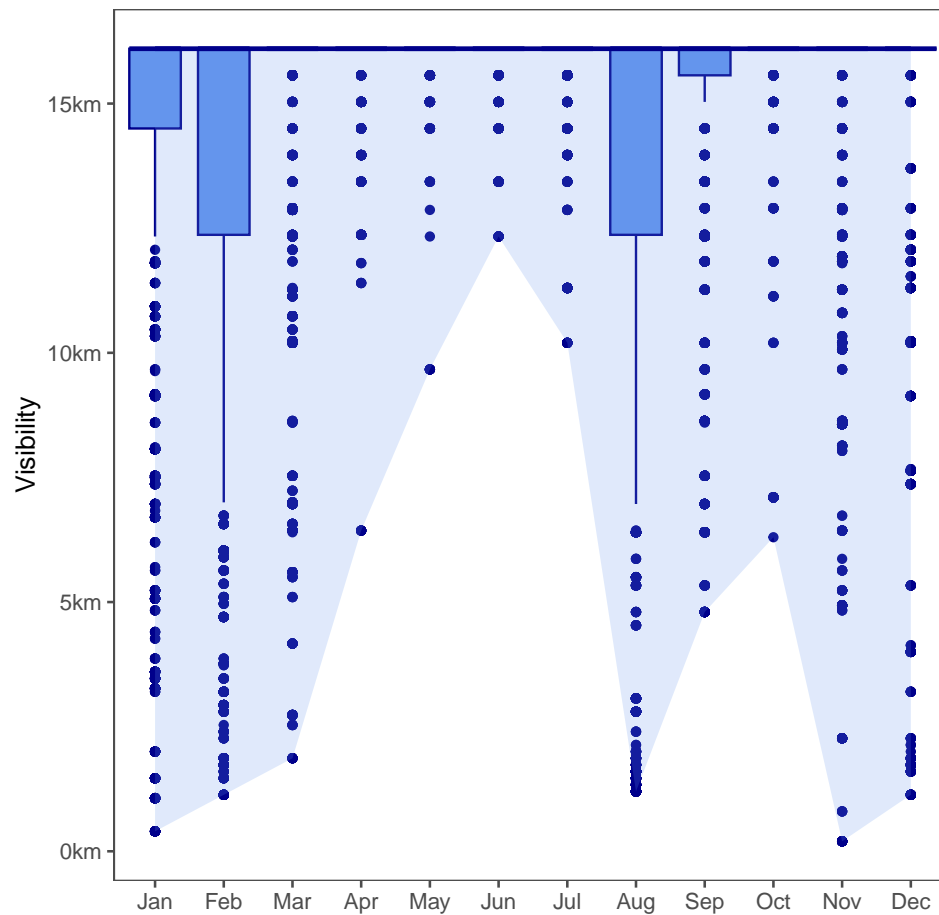


3.6 Visibility

	Avg.visibility	Avg.crash.visibility
1	15.10429	15.14964

Very little differences in visibility...

Density plot isn't helpful because cases are so spread out.



3.7 Air Pressure

	Avg.air.pressure	Avg.crash.air.pressure
1	96.55096	96.56156

4 Crash Subtypes

4.1 All Combined

Number of crashes: 56136

Number of intersection crashes: 24459 (43.6%)

Number of casualty crashes: 11473 (20.4%)

Top 5 most common streets:

HWY 97	HARVEY AVE	HWY 33	GORDON DR	LAKESHORE RD
6019	3168	2064	1921	1372

Avg. victims per year: 3275

Months of the year: (same as above)

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
5063	4728	4120	3825	4475	4735	5152	4907
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
4632	4836	4568	5095				

Days of the week: (same as above)

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
5464	8024	8728	8790	9061	9316	6753

Avg. Temperature: (same as above)

	Avg.temp.	Avg.crash.temp.
1	8.511442	10.91872

Avg Relative Humidity: (same as above)

	Avg.humidity	Avg.crash.humidity
1	69.32555	60.50743

Avg. Precipitation Amount (mm): (same as above)

	Avg.precipitation	Avg.crash.precipitation
1	0.08996575	0.09326635

Avg. Wind Speed (km/h): (same as above)

	Avg.wind.speed	Avg.crash.wind.speed
1	8.41538	10.27689

Avg. Visibility (km): (same as above)

	Avg.visibility	Avg.crash.visibility
1	15.10429	15.14964

Top 5 Most Common Crash Types:

REAR END SINGLE VEHICLE	UNDETERMINED	SIDE IMPACT	CONFLICTED	
13024	12495	11453	11369	3068

4.2 Cyclist Crashes

Number of crashes: 411

Number of intersection crashes: 328 (79.8%)

Number of casualty crashes: 332 (80.8%)

Top 5 most common streets:

GORDON DR	LAKE SHORE RD	HWY 97	BERNARD AVE	HWY 33
40	23	21	19	18

Avg. victims per year: 72

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
2	11	18	24	45	48	53	68
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
60	36	30	16				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
32	52	78	75	65	78	31

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.cyclist.crash.temp.
1	8.511442	10.91872	17.02561

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.cyclist.crash.humidity
1	69.32555	60.50743	50.07967

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.cyclist.crash.precipitation
1	0.02998858	0.2797991	0.08126521

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.cyclist.crash.wind.speed
1	8.41538	10.27689	10.45772

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.cyclist.crash.visibility
1	15.10429	15.14964	15.40862

Top 5 Most Common Crash Types:

SINGLE VEHICLE	UNDETERMINED
318	71
SIDE IMPACT	SIDE SWIPE - SAME DIRECTION
21	1
CONFLICTED	
0	

4.3 Animal Crashes

Number of crashes: 2018

Number of intersection crashes: 686 (34%)

Number of casualty crashes: 130 (6.4%)

Top 5 most common streets:

HWY 97	HWY 33	WESTSIDE RD	LAKESHORE RD	BOUCHERIE RD
366	272	194	71	52

Avg. victims per year: 30

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
152	135	153	126	127	152	143	158
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
182	257	280	153				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
244	307	284	312	303	329	239

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.animal.crash.temp.
1	8.511442	10.91872	8.280433

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.animal.crash.humidity
1	69.32555	60.50743	70.59081

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.animal.crash.precipitation
1	0.08996575	0.09326635	0.08062438

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.animal.crash.wind.speed
1	8.41538	10.27689	8.21608

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.animal.crash.visibility
1	15.10429	15.14964	15.27789

Top 5 Most Common Crash Types:

SINGLE VEHICLE	UNDETERMINED	REAR END	SIDE IMPACT	CONFLICTED
2005	7	3	2	1

4.4 Heavy Vehicle Crashes

Number of crashes: 2051

Number of intersection crashes: 932 (45.4%)

Number of casualty crashes: 365 (17.8%)

Top 5 most common streets:

HWY 97	HWY 33	HARVEY AVE	HWY 97 N	GORDON DR
310	77	73	58	47

Avg. victims per year: 100

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
187	173	145	116	157	197	182	181
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
185	170	184	174				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
56	370	378	365	380	402	100

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.heavy.crash.temp.
1	8.511442	10.91872	10.61056

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.heavy.crash.humidity
1	69.32555	60.50743	61.85107

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.heavy.crash.precipitation
1	0.08996575	0.09326635	0.09780595

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.heavy.crash.wind.speed
1	8.41538	10.27689	9.744303

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.heavy.crash.visibility
1	15.10429	15.14964	15.04907

Top 5 Most Common Crash Types:

UNDETERMINED	SIDE IMPACT	REAR END SINGLE VEHICLE	CONFLICTED
546	422	392	330
			137

4.5 Motorcycle Crashes

Number of crashes: 463

Number of intersection crashes: 288 (62.2%)

Number of casualty crashes: 275 (59.4%)

Top 5 most common streets:

HWY 97	HARVEY AVE	HWY 33	WESTSIDE RD	KLO RD
67	22	20	13	11

Avg. victims per year: 66

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
0	4	17	43	51	71	97	71
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
61	39	9	0				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
65	49	58	57	69	90	75

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.motorcycle.crash.temp.
1	8.511442	10.91872	19.28287

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.motorcycle.crash.humidity
1	69.32555	60.50743	45.53852

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.motorcycle.crash.precipitation
1	0.08996575	0.09326635	0.0762419

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.motorcycle.crash.wind.speed
1	8.41538	10.27689	11.33549

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.motorcycle.crash.visibility
1	15.10429	15.14964	15.54327

Top 5 Most Common Crash Types:

SIDE IMPACT	SINGLE VEHICLE	REAR END	UNDETERMINED	CONFLICTED
111	102	101	86	38

4.6 Parked Vehicle Crashes

Number of crashes: 17569

Number of intersection crashes: 2790 (15.9%)

Number of casualty crashes: 353 (2%)

Top 5 most common streets:

HARVEY AVE	HWY 97 LAKESHORE RD	LOUIE DR	GORDON DR
1221	760	520	473
			446

Avg. victims per year: 92

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
1589	1415	1369	1271	1439	1566	1629	1607
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
1449	1487	1316	1432				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
1948	2518	2615	2649	2664	2780	2395

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.parked.crash.temp.
1	8.511442	10.91872	11.32006

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.parked.crash.humidity
1	69.32555	60.50743	59.20128

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.parked.crash.precipitation
1	0.08996575	0.09326635	0.08685184

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.parked.crash.wind.speed
1	8.41538	10.27689	10.38783

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.parked.crash.visibility
1	15.10429	15.14964	15.1956

Top 5 Most Common Crash Types:

UNDETERMINED	SIDE IMPACT
7156	5519
REAR END	CONFLICTED
1501	852
SIDE SWIPE - SAME DIRECTION	
692	

4.7 Pedestrian Crashes

Number of crashes: 346

Number of intersection crashes: 220 (63.6%)

Number of casualty crashes: 289 (83.5%)

Top 5 most common streets:

HARVEY AVE	GORDON DR	HWY 97	HWY 33	BANKS RD
26	22	22	12	11

Avg. victims per year: 69

Months of the year:

JANUARY	FEBRUARY	MARCH	APRIL	MAY	JUNE	JULY	AUGUST
35	13	21	19	27	28	28	32
SEPTEMBER	OCTOBER	NOVEMBER	DECEMBER				
34	40	33	36				

Days of the week:

SUNDAY	MONDAY	TUESDAY	WEDNESDAY	THURSDAY	FRIDAY	SATURDAY
33	44	39	66	57	65	42

Avg. Temperature:

	Avg.temp.	Avg.crash.temp.	Avg.pedestrian.crash.temp.
1	8.511442	10.91872	10.77013

Avg Relative Humidity:

	Avg.humidity	Avg.crash.humidity	Avg.pedestrian.crash.humidity
1	69.32555	60.50743	63.38247

Avg. Precipitation Amount (mm):

	Avg.precipitation	Avg.crash.precipitation	Avg.pedestrian.crash.precipitation
1	0.08996575	0.09326635	0.09364162

Avg. Wind Speed (km/h):

	Avg.wind.speed	Avg.crash.wind.speed	Avg.pedestrian.crash.wind.speed
1	8.41538	10.27689	9.628131

Avg. Visibility (km):

	Avg.visibility	Avg.crash.visibility	Avg.pedestrian.crash.visibility
1	15.10429	15.14964	15.20434

Top 5 Most Common Crash Types:

SINGLE VEHICLE	UNDETERMINED	SIDE IMPACT	CONFLICTED	REAR END
273	53	8	5	4

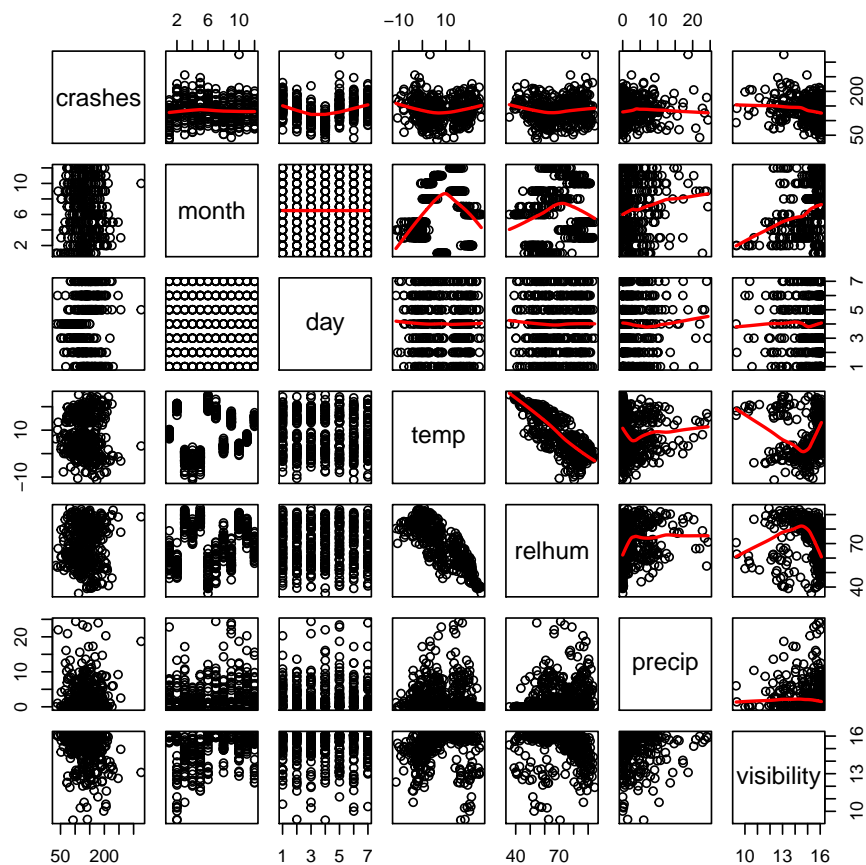
5 Regression/Neural Network Dataset

As opposed to linking the weather to each crash, this dataset links (some) crash data to the weather on each individual day (ie. the reverse). This is desirable if we want to predict the # of crashes and # of victims per day based on weather conditions.

Notice that rows are a bit weird/unintuitive for this dataset. Each row is a unique weekday in each month, and so actually represents the sum of all of these weekdays (ex. Fridays in October, Mondays in December, etc.). This is the case because the crash dataset was anonymized.

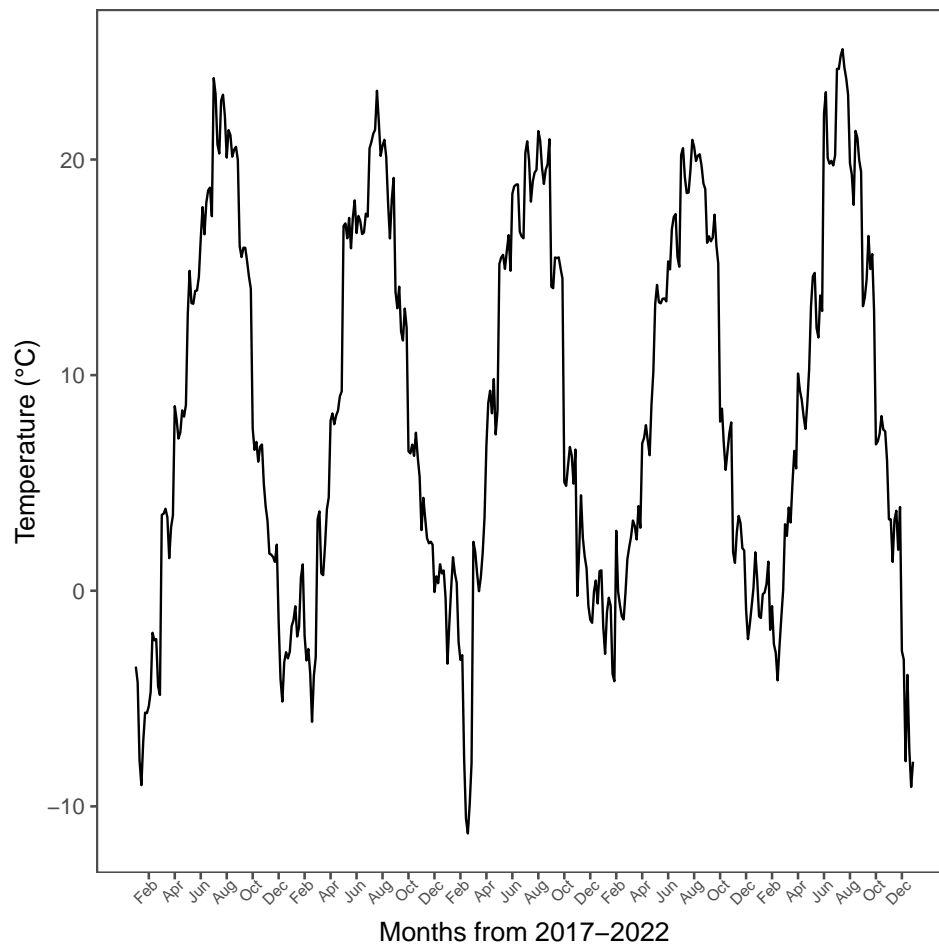
Variables:

[1]	"linker"	"month"	"day"	"temp"	"relhum"
[6]	"precip"	"wind.dir"	"wind.spd"	"visibility"	"pressure"
[11]	"crashes"	"victims"	"parked"	"HWY97"	"HARVEY"
[16]	"HWY33"	"GORDON"			

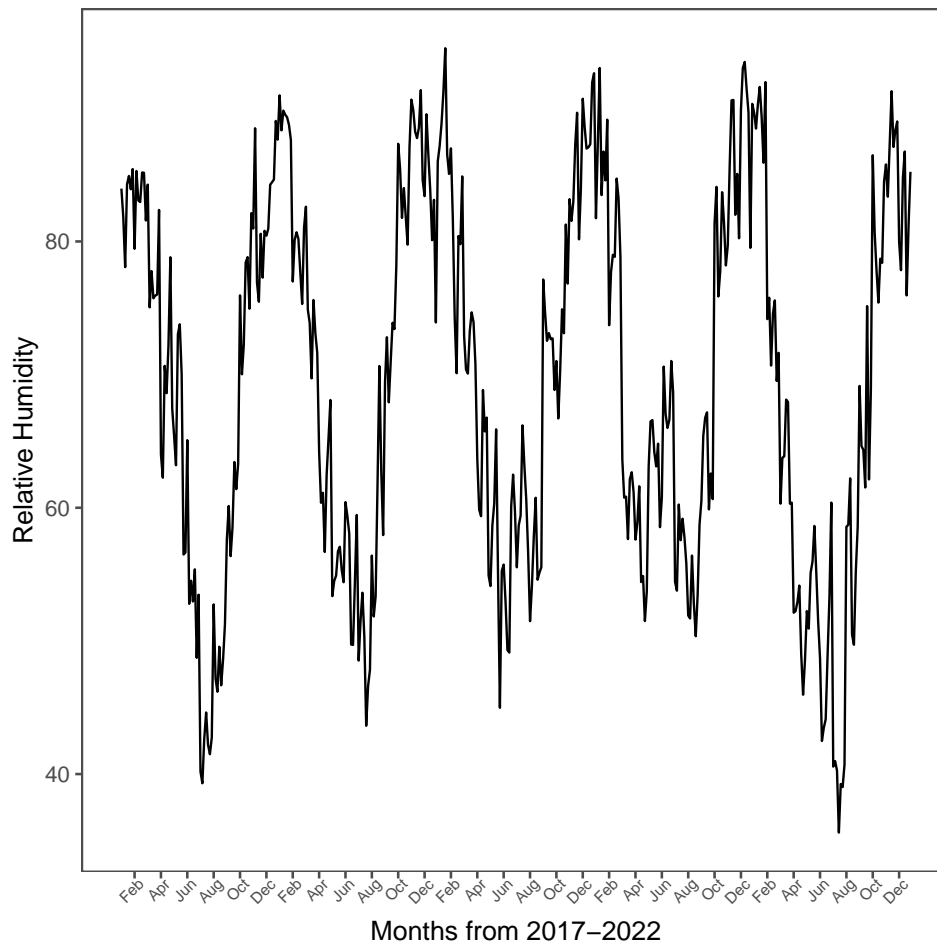


5.1 Weather Variables

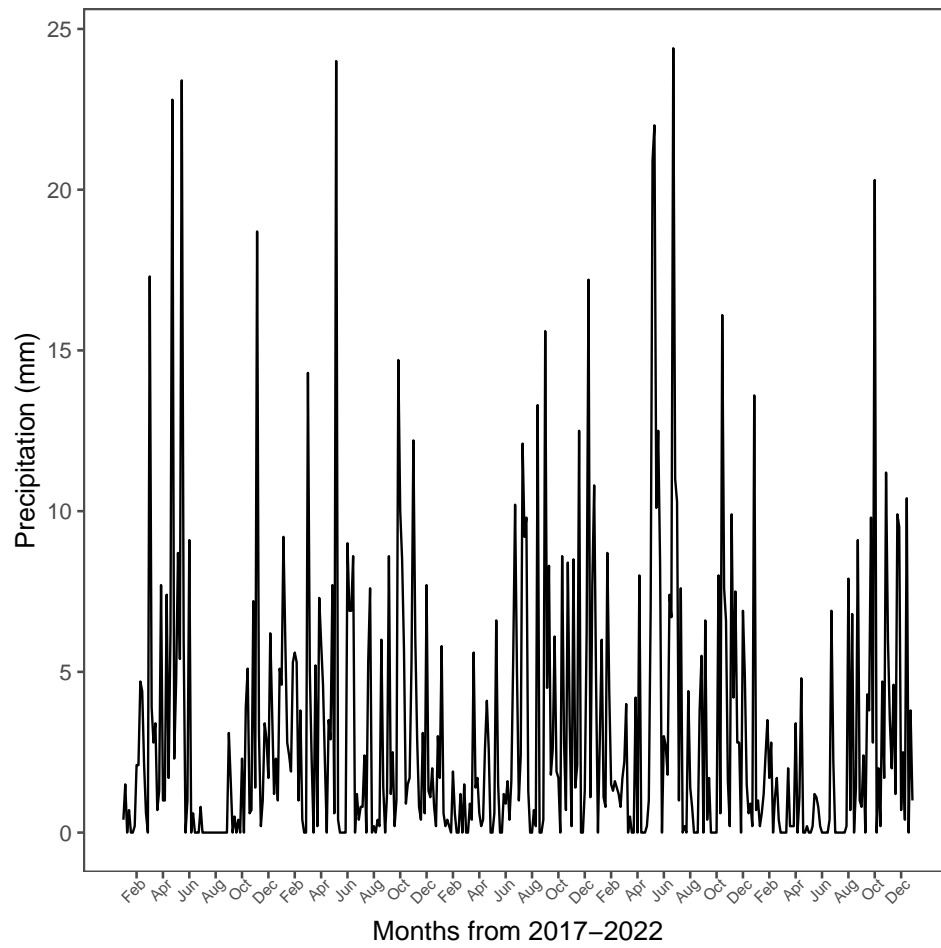
5.1.1 Temperature



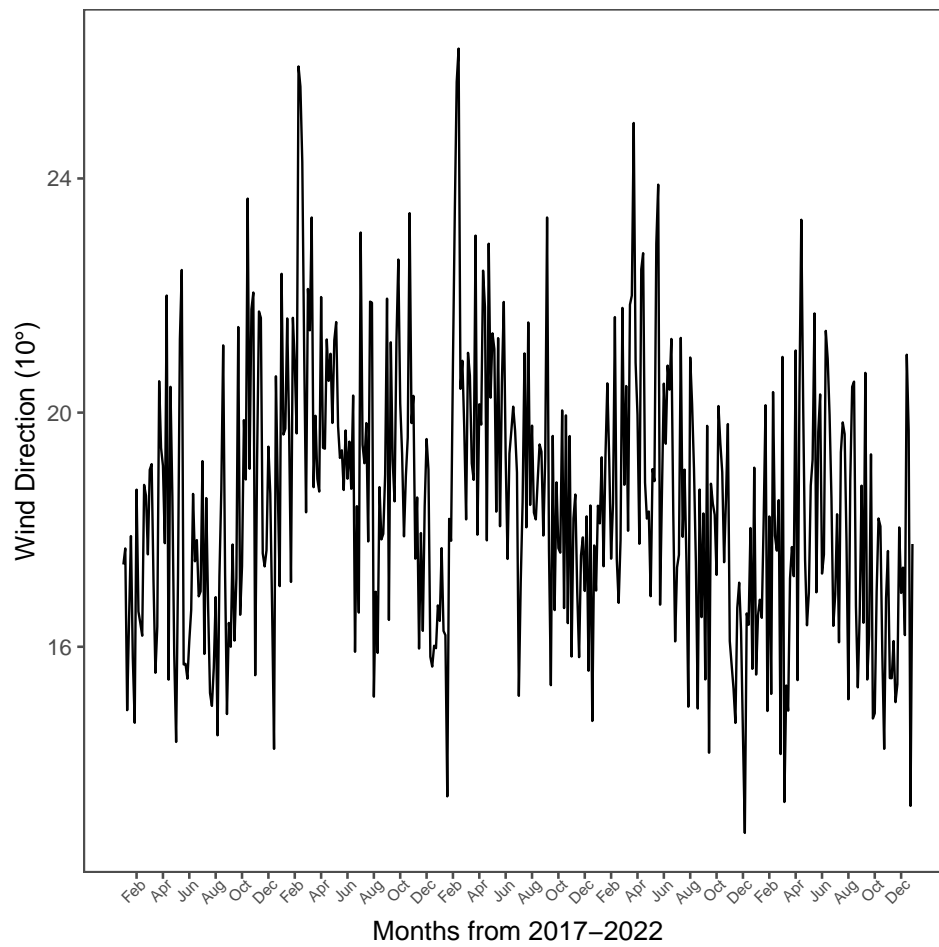
5.1.2 Relative Humidity



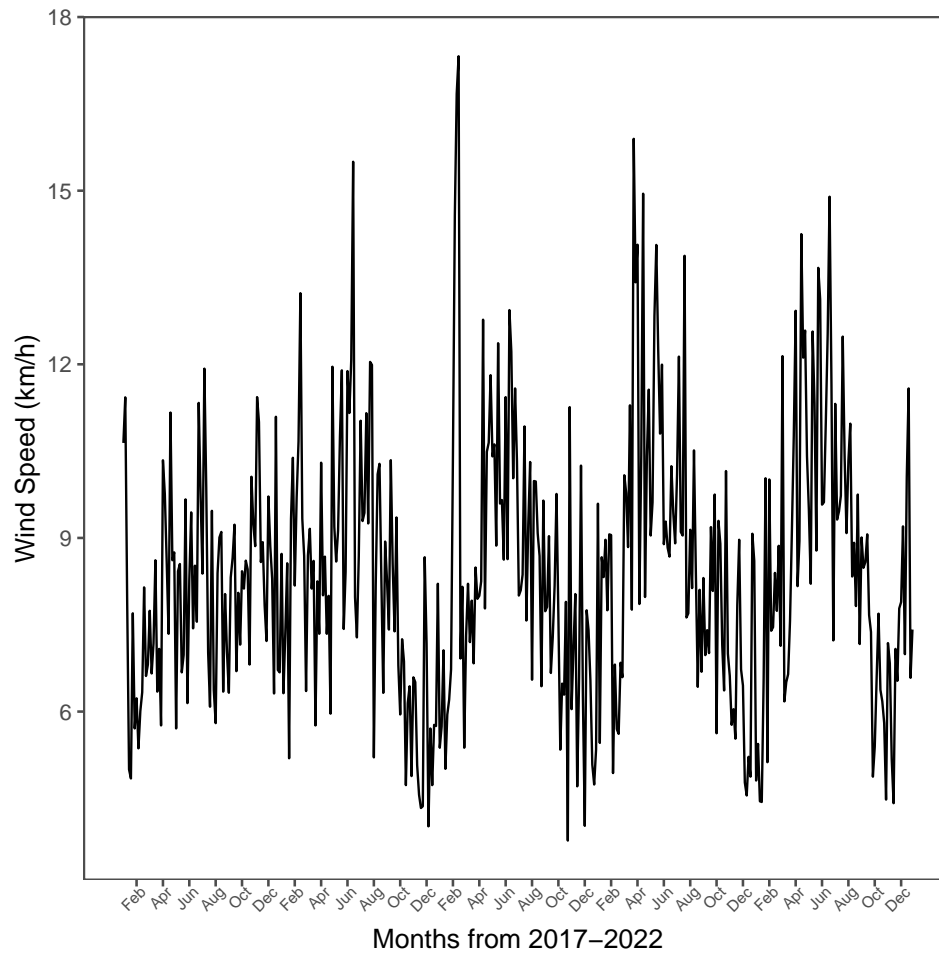
5.1.3 Precipitation



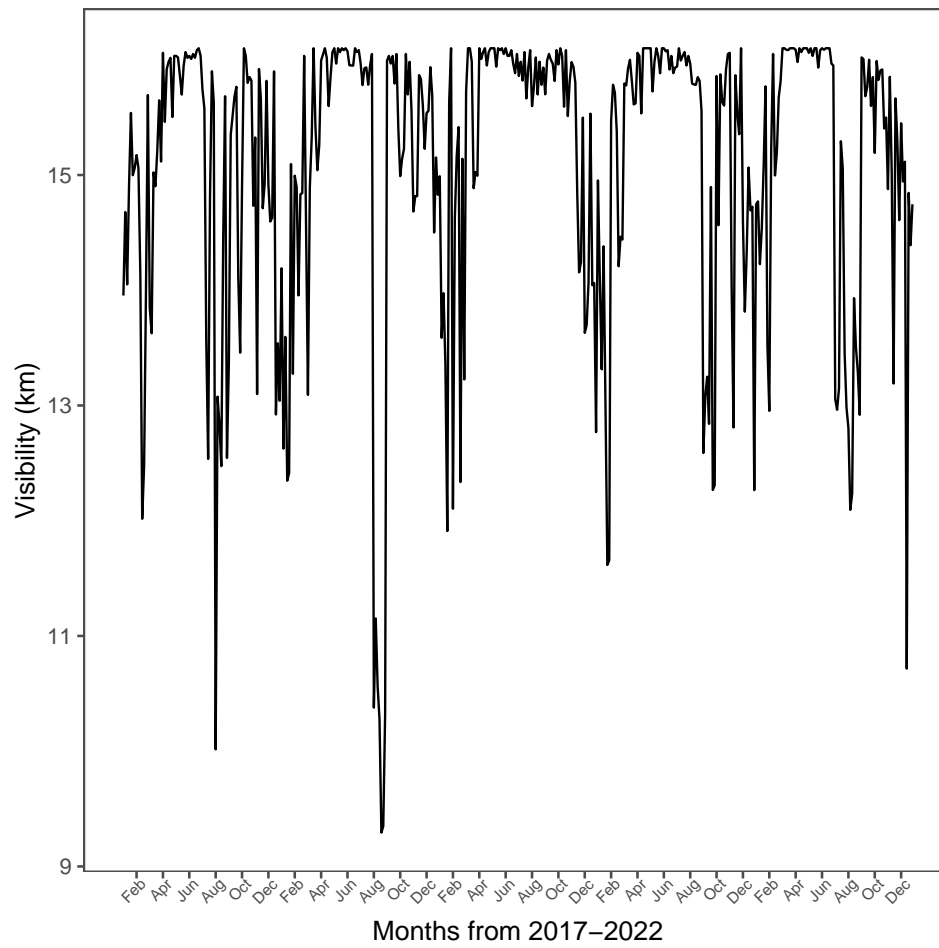
5.1.4 Wind Direction



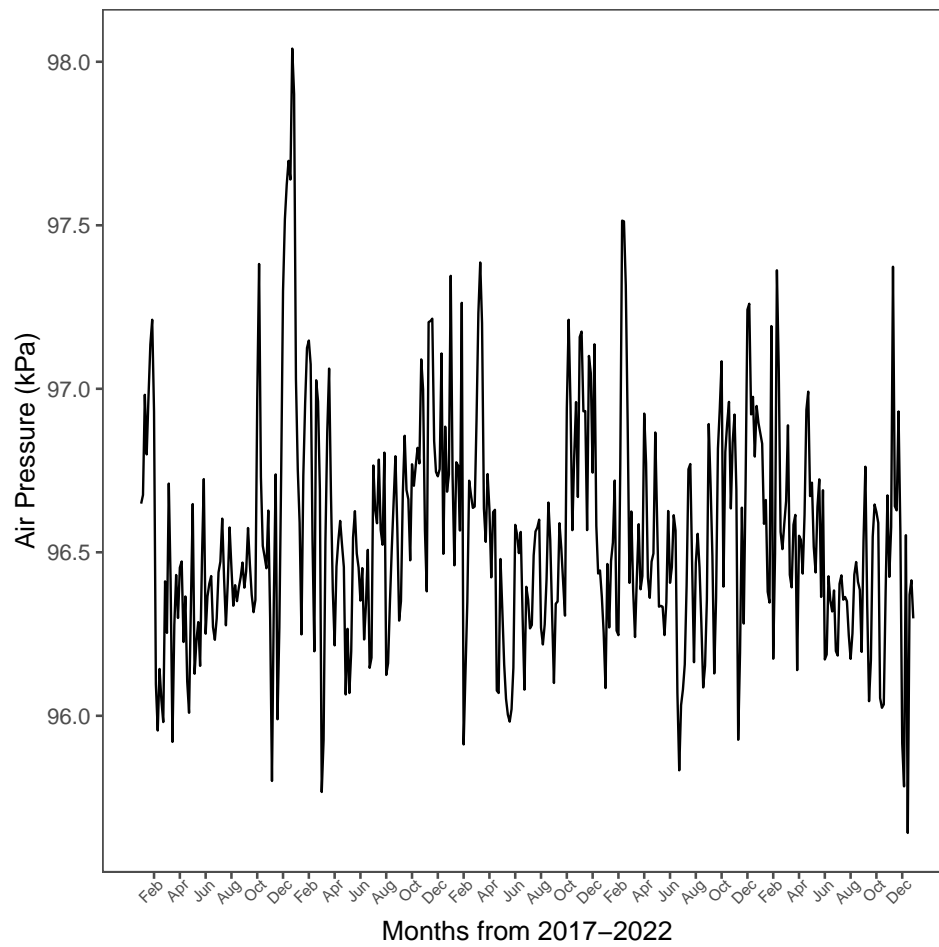
5.1.5 Wind Speed



5.1.6 Visibility



5.1.7 Air Pressure

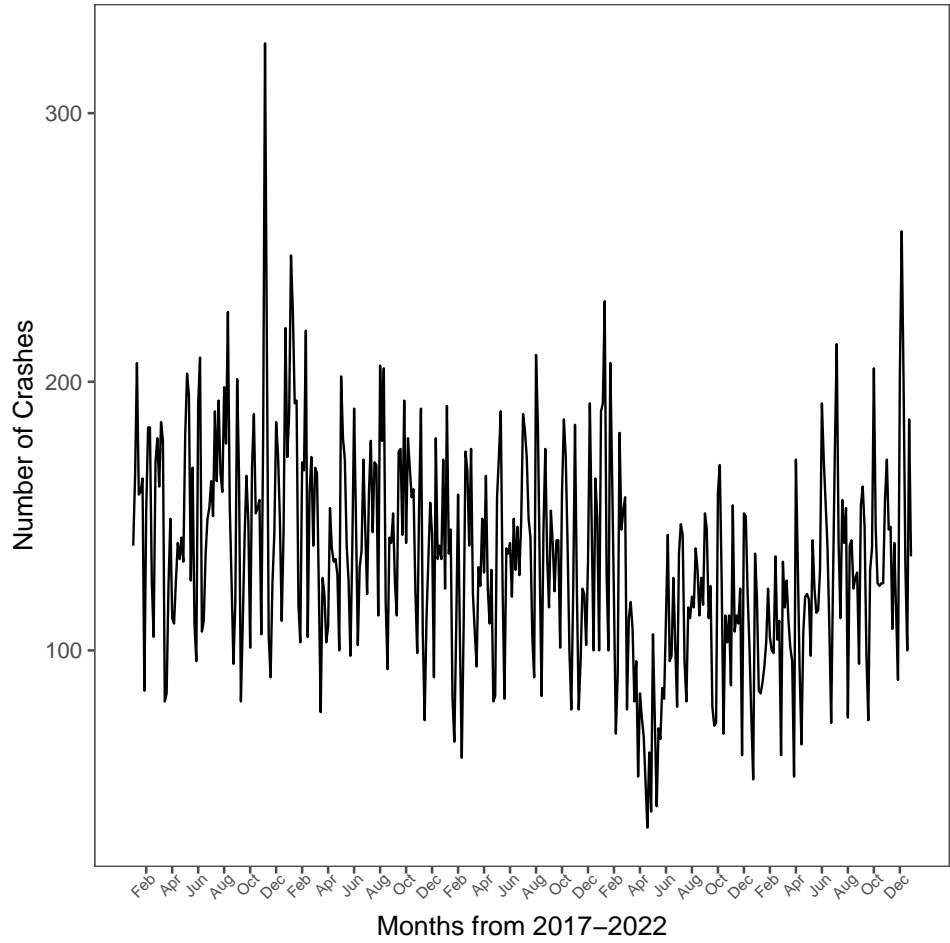


5.2 Crash Variables

5.2.1 Number of Crashes

Row with abnormally high number of crashes:

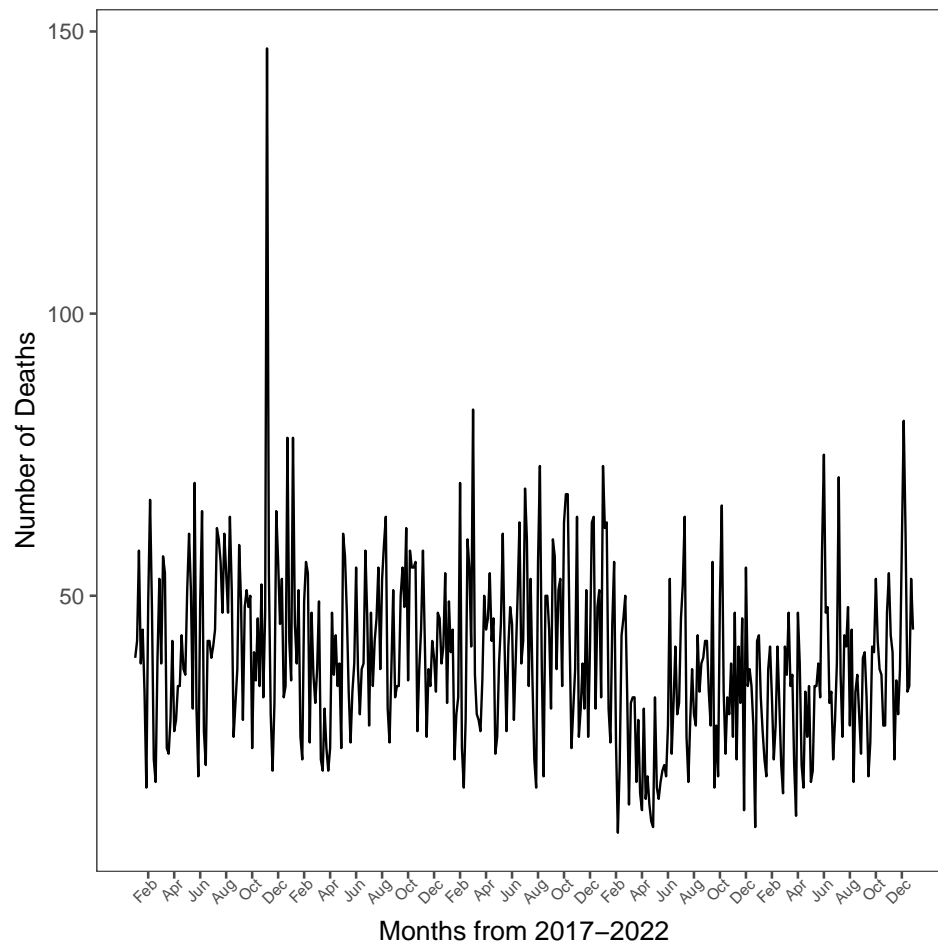
	linker	month	day	temp	relhum	precip	wind.dir				
72	2017	NOVEMBER	THURSDAY	NOVEMBER	THURSDAY	3.244167	88.5	18.7	19.20085		
	wind.spd	visibility	pressure	crashes	victims	parked	HWY97	HARVEY	HWY33		
72	11.43333	13.09917	95.80075	326	147	70	33	10	21		
	GORDON										
72	14										



5.2.2 Number of Deaths

Row with abnormally high number of deaths:

	linker	month	day	temp	relhum	precip	wind.dir
72	2017	NOVEMBER	THURSDAY	NOVEMBER	THURSDAY	3.244167	88.5 18.7 19.20085
	wind.spd	visibility	pressure	crashes	victims	parked	HWY97 HARVEY HWY33
72	11.43333	13.09917	95.80075	326	147	70	33 10 21
	GORDON						
72	14						



5.2.3 Locations

