# Analysis of Vehicular Crashes in Iowa

##### Nathan Rethwisch and Zack Swayne

## Introduction

According to the Iowa Department of Transportation, there are over 50,000 crashes per year1. These crashes cause millions in property damage, and unfortunately, the loss of life, with over 300 people dying to vehicular crashes per year. It is important to gain a greater understanding of the causes of these crashes in order to better create prevention strategies and protect the drivers on the road. Furthermore, by informing drivers of potential hazardous practices, they will be better prepared and encouraged to follow safe driving practices. The goal of this report is to explore some of the correlations between crashes and driving conditions to gain a better understanding of how to make the road a safer place.

## The Data

This data set comes from the Iowa Department of Transportation contains data for every recorded vehicle crash since January 2009. It is updated monthly by the Iowa Department of Transportation,

In total there are 728,442 observations in the data set and 37 columns. The columns include information on:

* Date and Time
* Location
* Number of Injuries/Fatalities
* Weather and Road Conditions
* Amounts of Property Damage
* Number and Sobriety of Passengers

## Data Cleaning

When looking at the data set, it became apparent that there was some data cleaning that needed to be done. The following output shows the number of NA values that appeared for each variable:

#NA values are only found in a few values?  
colSums(is.na(crashes))

## Iowa DOT Case Number Law Enforcement Case Number   
## 0 45587   
## Date of Crash Month of Crash   
## 0 0   
## Day of Week Time of Crash   
## 0 0   
## Hour DOT District   
## 0 0   
## City Name County Name   
## 279583 109894   
## Route with System Location Description   
## 451161 50640   
## First Harmful Event Location of First Harmful Event   
## 0 0   
## Manner of Crash/Collision Major Cause   
## 0 0   
## Drug or Alcohol Environmental Conditions   
## 0 0   
## Light Conditions Surface Conditions   
## 0 0   
## Weather Conditions Roadway Contribution   
## 0 0   
## Roadway Type Roadway Surface   
## 0 0   
## Work Zone Crash Severity   
## 0 0   
## Number of Fatalities Number of Injuries   
## 0 0   
## Number of Major Injuries Number of Minor Injuries   
## 0 0   
## Number of Possible Injuries Number of Unknown Injuries   
## 0 0   
## Amount of Property Damage Number of Vehicles Involved   
## 9 0   
## Total Number of Occupants Travel Direction   
## 0 0   
## Location   
## 1

Although there are many NA values in the data set, there were some that did not show up when calling the above function. For example, when looking at the data for Weather Conditions, it appeared that there were NA values, but the below function did not agree:

#It appears that Weather Conditions has NA values, but they are not showing up  
sum(is.na(crashes$`Weather Conditions`))

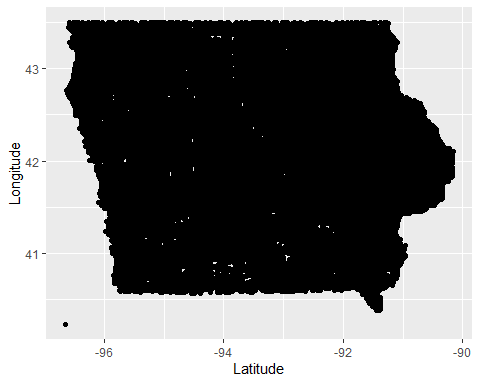
## [1] 0

It turns out, that a lot of NA values were coded as the character “N/A.” We converted those values to NA. In addition to this, some of the columns included blank values, so they were also converted to NA. Finally, under the ‘Total Number of Occupants’ column, there were multiple values listed at 777. Because it is illogical to have 777 passengers in a vehicle, we assumed that this was a coding scheme for when the number of passengers was unknown and also decided to convert those to NA.

## Iowa DOT Case Number Law Enforcement Case Number   
## 0 45587   
## Date of Crash Month of Crash   
## 0 0   
## Day of Week Time of Crash   
## 0 0   
## Hour DOT District   
## 0 4   
## City Name County Name   
## 279583 109894   
## Route with System Location Description   
## 451161 50640   
## First Harmful Event Location of First Harmful Event   
## 2311 23872   
## Manner of Crash/Collision Major Cause   
## 51205 0   
## Drug or Alcohol Environmental Conditions   
## 0 63550   
## Light Conditions Surface Conditions   
## 71992 72352   
## Weather Conditions Roadway Contribution   
## 72784 79089   
## Roadway Type Roadway Surface   
## 68521 50634   
## Work Zone Crash Severity   
## 719474 0   
## Number of Fatalities Number of Injuries   
## 0 0   
## Number of Major Injuries Number of Minor Injuries   
## 0 0   
## Number of Possible Injuries Number of Unknown Injuries   
## 0 0   
## Amount of Property Damage Number of Vehicles Involved   
## 9 0   
## Total Number of Occupants Travel Direction   
## 12325 491785   
## Location   
## 1

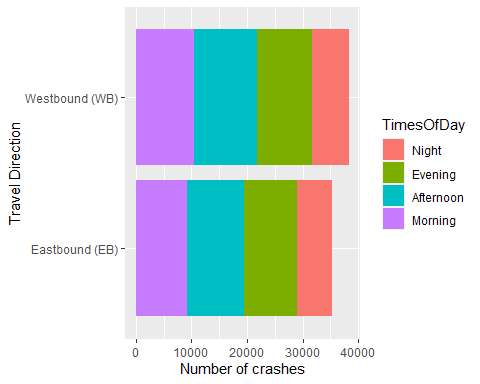
Next, using the lubridate package in R, the ‘Date of Crash’ column was converted into three new columns - Day, Month, and Year.

In addition to this, we cleaned and separated the Position data into longitude and latitude. The resulting longitude and latitude was then used to map every car crash in Iowa which is shown below:



## Sunrise/Sunset Analysis

#Crashes by direction  
crashes$Hour<- gsub("Hour ", "\\1", crashes$Hour)  
crashes$Hour<- parse\_integer(crashes$Hour)  
#View(crashes)  
  
#Creating different times of day  
#9pm-4am is night  
#5am-11am is morning  
#12pm-5pm is the afternoon  
#6pm-9pm is the evening  
crashes<- crashes %>% mutate(TimesOfDay =  
 case\_when((Hour >=21) | (Hour < 5) ~"Night",  
 (Hour >=5) & (Hour <12) ~"Morning",  
 (Hour >=12) & (Hour <17) ~"Afternoon",  
 (Hour >=17) & (Hour <21) ~"Evening")  
)  
  
#Plotting crashes by time of day and travel direction  
#There seems to be little difference, but sunset and sunrise times are changing, so it is hard to know when sunset and sunrise actually is happening  
crashes$TimesOfDay <- factor(crashes$TimesOfDay, levels = c("Night", "Evening", "Afternoon", "Morning"))  
crashes%>%  
 filter(!is.na(`Travel Direction`), `Travel Direction` %in% c('Westbound (WB)', 'Eastbound (EB)'))%>%  
 ggplot(aes(x = `Travel Direction`, fill = `TimesOfDay`)) + geom\_bar() + coord\_flip() + ylab("Number of crashes")



#Using rvest to extract sunrise and sunset data  
#A function that takes in the url and month number  
dfmonth <- function(url, number) {  
   
#Extraction of the table  
page <- read\_html(url)  
tables <- page %>% html\_table(fill = TRUE)  
tables %>% str  
weather<- tables[[2]]  
  
#Setting names to date, sunrise, and sunset  
names(weather)[1] <- "Date"  
names(weather)[2] <- "Sunrise"  
names(weather)[3] <- "Sunset"   
names(weather)[4] <- "Daylight"  
  
  
#The first day for each month in the dataframe is in it's own column, so I created a new dataframe with the pertinent intformation and renamed those columns  
weather2 <-weather[14:17]  
names(weather2)[1] <- "Date"  
names(weather2)[2] <- "Sunrise"  
names(weather2)[3] <- "Sunset"  
names(weather2)[4] <- "Daylight"  
  
#I then changed the Date to a character so it can be merged with the other data  
weather2$Date <- as.character(weather2$Date)  
  
#I select only the pertinent data from the original weather data set  
weather <- weather[1:4]  
  
#Next, I join the data sets while omitting NA values  
final\_weather <- full\_join(na.omit(weather), na.omit(weather2))  
  
#The first day in the dataset will be at the end, so I move it towards the front  
final\_weather[1,]<- final\_weather[nrow(final\_weather),]  
  
#Next, I remove the labels that stuck onto this dataset that are unnecessary  
final\_weather <- final\_weather[-2,]  
  
#The last two rows are removed because they contain a copy of day 1 and some additional information from the webpage  
final\_weather <- final\_weather[-nrow(final\_weather)+1: -nrow(final\_weather),]  
  
#These loops are to deal with notices for time changes. If there is a time change notice in the dataframe, they are removed.  
if (any(final\_weather == "Note: hours shift because clocks change forward 1 hour. (See the note below this table for details)")){  
 final\_weather <-final\_weather[-which(final\_weather$Sunrise == "Note: hours shift because clocks change forward 1 hour. (See the note below this table for details)"),]  
} else if (any(final\_weather == "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)")){  
 final\_weather <- final\_weather[-which(final\_weather$Sunrise == "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)"),]  
}  
  
#Creating a new column that corresponds to the month number  
final\_weather$Month <- number  
   
#Seperating additional informatoin from the sunset and sunrise times in the graph  
final\_weather<-final\_weather%>%  
 separate(col = Sunrise, into = c("Sunrise Time", NA), remove = TRUE, sep = " ↑")  
  
final\_weather<-final\_weather%>%  
 separate(col = Sunset, into = c("Sunset Time", NA), remove = TRUE, sep = " ↑")  
  
#Making the sunset time in Hour:Minute:Second format  
final\_weather$`Sunset Time` <- format(strptime(final\_weather$`Sunset Time`, "%I:%M %p"), format="%H:%M:%S")  
  
final\_weather$`Sunrise Time` <- format(strptime(final\_weather$`Sunrise Time`, "%I:%M %p"), format="%H:%M:%S")  
  
  
#Making the data type of sunset and sunrise an hms so it can be compared to the original data  
final\_weather$`Sunrise Time` <- as\_hms(final\_weather$`Sunrise Time`)  
final\_weather$`Sunset Time` <- as\_hms(final\_weather$`Sunset Time`)  
final\_weather$`Daylight` <- as\_hms(final\_weather$`Daylight`)  
  
#Returns the altered dataframe  
 return (final\_weather)  
}

#Creating a dataframe using my previous function in the month January  
#I'm choosing 2020 because it is the most recent leap year - so I will have data for February 29th. I'm choosing Ames because it is close to the geographical center of Iowa  
sun<- dfmonth("https://www.timeanddate.com/sun/@4846834?month=1&year=2020", 1)

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:34 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Jan" "Jan" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "7:42 am ↑ (121°)" "7:42 am ↑ (121°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "4:54 pm ↑ (239°)" "4:55 pm ↑ (239°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "9:11:30" "9:12:21" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "+0:46" "+0:50" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "6:02 am" "6:02 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "6:34 pm" "6:35 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "6:36 am" "6:36 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "6:00 pm" "6:01 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "7:11 am" "7:11 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "5:25 pm" "5:26 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "12:18 pm (25.1°)" "12:18 pm (25.2°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "91.400" "91.399" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "7:42 am ↑ (121°)" NA NA ...  
## ..$ : chr [1:33] NA "4:53 pm ↑ (239°)" NA NA ...  
## ..$ : chr [1:33] NA "9:10:43" NA NA ...  
## ..$ : chr [1:33] NA "+0:42" NA NA ...  
## ..$ : chr [1:33] NA "6:01 am" NA NA ...  
## ..$ : chr [1:33] NA "6:34 pm" NA NA ...  
## ..$ : chr [1:33] NA "6:35 am" NA NA ...  
## ..$ : chr [1:33] NA "6:00 pm" NA NA ...  
## ..$ : chr [1:33] NA "7:11 am" NA NA ...  
## ..$ : chr [1:33] NA "5:24 pm" NA NA ...  
## ..$ : chr [1:33] NA "12:17 pm (25.0°)" NA NA ...  
## ..$ : num [1:33] NA 91.4 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")

#For the rest of the months, I run this loop and join them with the dataframe  
for (i in 2:12){  
 url<- paste0("https://www.timeanddate.com/sun/@4846834?month=", i, "&year=2020")  
 dfTemp<- dfmonth(url, i)  
 sun<- full\_join(sun, dfTemp)  
   
}

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:34 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [31 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:31] "Feb" "Feb" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:31] "Sunrise" "Sunrise" "7:26 am ↑ (112°)" "7:25 am ↑ (112°)" ...  
## ..$ Sunrise/Sunset : chr [1:31] "Sunset" "Sunset" "5:30 pm ↑ (248°)" "5:31 pm ↑ (248°)" ...  
## ..$ Daylength : chr [1:31] "Length" "Length" "10:03:22" "10:05:43" ...  
## ..$ Daylength : chr [1:31] "Diff." "Diff." "+2:19" "+2:20" ...  
## ..$ Astronomical Twilight: chr [1:31] "Start" "Start" "5:51 am" "5:50 am" ...  
## ..$ Astronomical Twilight: chr [1:31] "End" "End" "7:05 pm" "7:06 pm" ...  
## ..$ Nautical Twilight : chr [1:31] "Start" "Start" "6:23 am" "6:22 am" ...  
## ..$ Nautical Twilight : chr [1:31] "End" "End" "6:33 pm" "6:34 pm" ...  
## ..$ Civil Twilight : chr [1:31] "Start" "Start" "6:57 am" "6:56 am" ...  
## ..$ Civil Twilight : chr [1:31] "End" "End" "5:59 pm" "6:00 pm" ...  
## ..$ Solar Noon : chr [1:31] "Time" "Time" "12:28 pm (31.2°)" "12:28 pm (31.5°)" ...  
## ..$ Solar Noon : chr [1:31] "Mil. mi" "Mil. mi" "91.602" "91.615" ...  
## ..$ : int [1:31] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:31] NA "7:27 am ↑ (113°)" NA NA ...  
## ..$ : chr [1:31] NA "5:28 pm ↑ (248°)" NA NA ...  
## ..$ : chr [1:31] NA "10:01:03" NA NA ...  
## ..$ : chr [1:31] NA "+2:17" NA NA ...  
## ..$ : chr [1:31] NA "5:51 am" NA NA ...  
## ..$ : chr [1:31] NA "7:04 pm" NA NA ...  
## ..$ : chr [1:31] NA "6:24 am" NA NA ...  
## ..$ : chr [1:31] NA "6:31 pm" NA NA ...  
## ..$ : chr [1:31] NA "6:58 am" NA NA ...  
## ..$ : chr [1:31] NA "5:58 pm" NA NA ...  
## ..$ : chr [1:31] NA "12:28 pm (30.9°)" NA NA ...  
## ..$ : num [1:31] NA 91.6 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:35 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [34 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:34] "Mar" "Mar" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:34] "Sunrise" "Sunrise" "6:47 am ↑ (99°)" "6:45 am ↑ (98°)" ...  
## ..$ Sunrise/Sunset : chr [1:34] "Sunset" "Sunset" "6:06 pm ↑ (262°)" "6:07 pm ↑ (262°)" ...  
## ..$ Daylength : chr [1:34] "Length" "Length" "11:19:29" "11:22:18" ...  
## ..$ Daylength : chr [1:34] "Diff." "Diff." "+2:48" "+2:48" ...  
## ..$ Astronomical Twilight: chr [1:34] "Start" "Start" "5:14 am" "5:12 am" ...  
## ..$ Astronomical Twilight: chr [1:34] "End" "End" "7:39 pm" "7:40 pm" ...  
## ..$ Nautical Twilight : chr [1:34] "Start" "Start" "5:46 am" "5:45 am" ...  
## ..$ Nautical Twilight : chr [1:34] "End" "End" "7:06 pm" "7:08 pm" ...  
## ..$ Civil Twilight : chr [1:34] "Start" "Start" "6:19 am" "6:17 am" ...  
## ..$ Civil Twilight : chr [1:34] "End" "End" "6:34 pm" "6:35 pm" ...  
## ..$ Solar Noon : chr [1:34] "Time" "Time" "12:26 pm (41.2°)" "12:26 pm (41.6°)" ...  
## ..$ Solar Noon : chr [1:34] "Mil. mi" "Mil. mi" "92.144" "92.166" ...  
## ..$ : int [1:34] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:34] NA "6:48 am ↑ (99°)" NA NA ...  
## ..$ : chr [1:34] NA "6:05 pm ↑ (261°)" NA NA ...  
## ..$ : chr [1:34] NA "11:16:41" NA NA ...  
## ..$ : chr [1:34] NA "+2:47" NA NA ...  
## ..$ : chr [1:34] NA "5:16 am" NA NA ...  
## ..$ : chr [1:34] NA "7:38 pm" NA NA ...  
## ..$ : chr [1:34] NA "5:48 am" NA NA ...  
## ..$ : chr [1:34] NA "7:05 pm" NA NA ...  
## ..$ : chr [1:34] NA "6:20 am" NA NA ...  
## ..$ : chr [1:34] NA "6:33 pm" NA NA ...  
## ..$ : chr [1:34] NA "12:26 pm (40.8°)" NA NA ...  
## ..$ : num [1:34] NA 92.1 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:35 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [32 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:32] "Apr" "Apr" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunrise" "Sunrise" "6:54 am ↑ (82°)" "6:52 am ↑ (82°)" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunset" "Sunset" "7:41 pm ↑ (278°)" "7:43 pm ↑ (279°)" ...  
## ..$ Daylength : chr [1:32] "Length" "Length" "12:47:21" "12:50:09" ...  
## ..$ Daylength : chr [1:32] "Diff." "Diff." "+2:48" "+2:48" ...  
## ..$ Astronomical Twilight: chr [1:32] "Start" "Start" "5:18 am" "5:16 am" ...  
## ..$ Astronomical Twilight: chr [1:32] "End" "End" "9:18 pm" "9:20 pm" ...  
## ..$ Nautical Twilight : chr [1:32] "Start" "Start" "5:52 am" "5:50 am" ...  
## ..$ Nautical Twilight : chr [1:32] "End" "End" "8:43 pm" "8:45 pm" ...  
## ..$ Civil Twilight : chr [1:32] "Start" "Start" "6:26 am" "6:24 am" ...  
## ..$ Civil Twilight : chr [1:32] "End" "End" "8:10 pm" "8:11 pm" ...  
## ..$ Solar Noon : chr [1:32] "Time" "Time" "1:17 pm (53.3°)" "1:17 pm (53.7°)" ...  
## ..$ Solar Noon : chr [1:32] "Mil. mi" "Mil. mi" "92.935" "92.961" ...  
## ..$ : int [1:32] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:32] NA "6:56 am ↑ (83°)" NA NA ...  
## ..$ : chr [1:32] NA "7:40 pm ↑ (278°)" NA NA ...  
## ..$ : chr [1:32] NA "12:44:32" NA NA ...  
## ..$ : chr [1:32] NA "+2:48" NA NA ...  
## ..$ : chr [1:32] NA "5:20 am" NA NA ...  
## ..$ : chr [1:32] NA "9:17 pm" NA NA ...  
## ..$ : chr [1:32] NA "5:54 am" NA NA ...  
## ..$ : chr [1:32] NA "8:42 pm" NA NA ...  
## ..$ : chr [1:32] NA "6:28 am" NA NA ...  
## ..$ : chr [1:32] NA "8:09 pm" NA NA ...  
## ..$ : chr [1:32] NA "1:18 pm (52.9°)" NA NA ...  
## ..$ : num [1:32] NA 92.9 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:35 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "May" "May" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "6:08 am ↑ (68°)" "6:07 am ↑ (68°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "8:15 pm ↑ (292°)" "8:16 pm ↑ (293°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "14:06:47" "14:09:09" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "+2:24" "+2:22" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "4:19 am" "4:17 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "10:04 pm" "10:06 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "5:00 am" "4:58 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "9:23 pm" "9:24 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "5:37 am" "5:36 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "8:45 pm" "8:47 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "1:11 pm (63.6°)" "1:11 pm (63.9°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "93.702" "93.725" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "6:09 am ↑ (68°)" NA NA ...  
## ..$ : chr [1:33] NA "8:14 pm ↑ (292°)" NA NA ...  
## ..$ : chr [1:33] NA "14:04:22" NA NA ...  
## ..$ : chr [1:33] NA "+2:25" NA NA ...  
## ..$ : chr [1:33] NA "4:21 am" NA NA ...  
## ..$ : chr [1:33] NA "10:03 pm" NA NA ...  
## ..$ : chr [1:33] NA "5:01 am" NA NA ...  
## ..$ : chr [1:33] NA "9:22 pm" NA NA ...  
## ..$ : chr [1:33] NA "5:39 am" NA NA ...  
## ..$ : chr [1:33] NA "8:44 pm" NA NA ...  
## ..$ : chr [1:33] NA "1:11 pm (63.4°)" NA NA ...  
## ..$ : num [1:33] NA 93.7 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:36 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [32 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:32] "Jun" "Jun" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunrise" "Sunrise" "5:40 am ↑ (58°)" "5:40 am ↑ (58°)" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunset" "Sunset" "8:44 pm ↑ (302°)" "8:45 pm ↑ (302°)" ...  
## ..$ Daylength : chr [1:32] "Length" "Length" "15:04:12" "15:05:19" ...  
## ..$ Daylength : chr [1:32] "Diff." "Diff." "+1:09" "+1:06" ...  
## ..$ Astronomical Twilight: chr [1:32] "Start" "Start" "3:32 am" "3:31 am" ...  
## ..$ Astronomical Twilight: chr [1:32] "End" "End" "10:53 pm" "10:54 pm" ...  
## ..$ Nautical Twilight : chr [1:32] "Start" "Start" "4:23 am" "4:23 am" ...  
## ..$ Nautical Twilight : chr [1:32] "End" "End" "10:01 pm" "10:02 pm" ...  
## ..$ Civil Twilight : chr [1:32] "Start" "Start" "5:06 am" "5:06 am" ...  
## ..$ Civil Twilight : chr [1:32] "End" "End" "9:18 pm" "9:19 pm" ...  
## ..$ Solar Noon : chr [1:32] "Time" "Time" "1:12 pm (70.3°)" "1:12 pm (70.4°)" ...  
## ..$ Solar Noon : chr [1:32] "Mil. mi" "Mil. mi" "94.286" "94.299" ...  
## ..$ : int [1:32] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:32] NA "5:41 am ↑ (59°)" NA NA ...  
## ..$ : chr [1:32] NA "8:44 pm ↑ (301°)" NA NA ...  
## ..$ : chr [1:32] NA "15:03:02" NA NA ...  
## ..$ : chr [1:32] NA "+1:13" NA NA ...  
## ..$ : chr [1:32] NA "3:33 am" NA NA ...  
## ..$ : chr [1:32] NA "10:52 pm" NA NA ...  
## ..$ : chr [1:32] NA "4:24 am" NA NA ...  
## ..$ : chr [1:32] NA "10:01 pm" NA NA ...  
## ..$ : chr [1:32] NA "5:07 am" NA NA ...  
## ..$ : chr [1:32] NA "9:17 pm" NA NA ...  
## ..$ : chr [1:32] NA "1:12 pm (70.2°)" NA NA ...  
## ..$ : num [1:32] NA 94.3 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:36 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Jul" "Jul" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "5:43 am ↑ (57°)" "5:44 am ↑ (58°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "8:53 pm ↑ (303°)" "8:53 pm ↑ (302°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "15:10:11" "15:09:24" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "−0:43" "−0:47" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "3:32 am" "3:33 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "11:04 pm" "11:03 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "4:25 am" "4:25 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "10:11 pm" "10:11 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "5:09 am" "5:09 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "9:27 pm" "9:27 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "1:18 pm (70.9°)" "1:18 pm (70.8°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "94.507" "94.508" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "5:42 am ↑ (57°)" NA NA ...  
## ..$ : chr [1:33] NA "8:53 pm ↑ (303°)" NA NA ...  
## ..$ : chr [1:33] NA "15:10:55" NA NA ...  
## ..$ : chr [1:33] NA "−0:40" NA NA ...  
## ..$ : chr [1:33] NA "3:31 am" NA NA ...  
## ..$ : chr [1:33] NA "11:04 pm" NA NA ...  
## ..$ : chr [1:33] NA "4:24 am" NA NA ...  
## ..$ : chr [1:33] NA "10:12 pm" NA NA ...  
## ..$ : chr [1:33] NA "5:08 am" NA NA ...  
## ..$ : chr [1:33] NA "9:28 pm" NA NA ...  
## ..$ : chr [1:33] NA "1:18 pm (71.0°)" NA NA ...  
## ..$ : num [1:33] NA 94.5 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:36 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Aug" "Aug" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "6:09 am ↑ (65°)" "6:10 am ↑ (66°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "8:31 pm ↑ (295°)" "8:30 pm ↑ (294°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "14:21:47" "14:19:34" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "−2:10" "−2:12" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "4:15 am" "4:17 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "10:24 pm" "10:22 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "4:59 am" "5:00 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "9:41 pm" "9:39 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "5:38 am" "5:39 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "9:02 pm" "9:01 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "1:20 pm (65.5°)" "1:20 pm (65.2°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "94.324" "94.312" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "6:08 am ↑ (65°)" NA NA ...  
## ..$ : chr [1:33] NA "8:32 pm ↑ (295°)" NA NA ...  
## ..$ : chr [1:33] NA "14:23:57" NA NA ...  
## ..$ : chr [1:33] NA "−2:08" NA NA ...  
## ..$ : chr [1:33] NA "4:14 am" NA NA ...  
## ..$ : chr [1:33] NA "10:26 pm" NA NA ...  
## ..$ : chr [1:33] NA "4:57 am" NA NA ...  
## ..$ : chr [1:33] NA "9:42 pm" NA NA ...  
## ..$ : chr [1:33] NA "5:36 am" NA NA ...  
## ..$ : chr [1:33] NA "9:03 pm" NA NA ...  
## ..$ : chr [1:33] NA "1:20 pm (65.7°)" NA NA ...  
## ..$ : num [1:33] NA 94.3 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:37 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [32 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:32] "Sep" "Sep" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunrise" "Sunrise" "6:41 am ↑ (79°)" "6:42 am ↑ (79°)" ...  
## ..$ Sunrise/Sunset : chr [1:32] "Sunset" "Sunset" "7:45 pm ↑ (281°)" "7:43 pm ↑ (280°)" ...  
## ..$ Daylength : chr [1:32] "Length" "Length" "13:03:45" "13:01:01" ...  
## ..$ Daylength : chr [1:32] "Diff." "Diff." "−2:43" "−2:43" ...  
## ..$ Astronomical Twilight: chr [1:32] "Start" "Start" "5:02 am" "5:04 am" ...  
## ..$ Astronomical Twilight: chr [1:32] "End" "End" "9:23 pm" "9:21 pm" ...  
## ..$ Nautical Twilight : chr [1:32] "Start" "Start" "5:38 am" "5:40 am" ...  
## ..$ Nautical Twilight : chr [1:32] "End" "End" "8:48 pm" "8:46 pm" ...  
## ..$ Civil Twilight : chr [1:32] "Start" "Start" "6:13 am" "6:14 am" ...  
## ..$ Civil Twilight : chr [1:32] "End" "End" "8:14 pm" "8:12 pm" ...  
## ..$ Solar Noon : chr [1:32] "Time" "Time" "1:13 pm (55.5°)" "1:13 pm (55.2°)" ...  
## ..$ Solar Noon : chr [1:32] "Mil. mi" "Mil. mi" "93.771" "93.748" ...  
## ..$ : int [1:32] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:32] NA "6:40 am ↑ (78°)" NA NA ...  
## ..$ : chr [1:32] NA "7:47 pm ↑ (281°)" NA NA ...  
## ..$ : chr [1:32] NA "13:06:28" NA NA ...  
## ..$ : chr [1:32] NA "−2:42" NA NA ...  
## ..$ : chr [1:32] NA "5:01 am" NA NA ...  
## ..$ : chr [1:32] NA "9:25 pm" NA NA ...  
## ..$ : chr [1:32] NA "5:37 am" NA NA ...  
## ..$ : chr [1:32] NA "8:49 pm" NA NA ...  
## ..$ : chr [1:32] NA "6:11 am" NA NA ...  
## ..$ : chr [1:32] NA "8:15 pm" NA NA ...  
## ..$ : chr [1:32] NA "1:14 pm (55.9°)" NA NA ...  
## ..$ : num [1:32] NA 93.8 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:37 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Oct" "Oct" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "7:13 am ↑ (94°)" "7:14 am ↑ (95°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "6:53 pm ↑ (265°)" "6:51 pm ↑ (265°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "11:40:12" "11:37:24" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "−2:47" "−2:47" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "5:40 am" "5:41 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "8:26 pm" "8:24 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "6:12 am" "6:13 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "7:53 pm" "7:51 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "6:45 am" "6:46 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "7:21 pm" "7:19 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "1:03 pm (44.0°)" "1:03 pm (43.6°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "93.016" "92.990" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "7:12 am ↑ (94°)" NA NA ...  
## ..$ : chr [1:33] NA "6:55 pm ↑ (266°)" NA NA ...  
## ..$ : chr [1:33] NA "11:42:59" NA NA ...  
## ..$ : chr [1:33] NA "−2:47" NA NA ...  
## ..$ : chr [1:33] NA "5:39 am" NA NA ...  
## ..$ : chr [1:33] NA "8:27 pm" NA NA ...  
## ..$ : chr [1:33] NA "6:11 am" NA NA ...  
## ..$ : chr [1:33] NA "7:55 pm" NA NA ...  
## ..$ : chr [1:33] NA "6:44 am" NA NA ...  
## ..$ : chr [1:33] NA "7:22 pm" NA NA ...  
## ..$ : chr [1:33] NA "1:03 pm (44.4°)" NA NA ...  
## ..$ : num [1:33] NA 93 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:37 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Nov" "Nov" "1" "2" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "6:47 am ↑ (109°)" "6:49 am ↑ (110°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "5:07 pm ↑ (251°)" "5:06 pm ↑ (250°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "10:19:44" "10:17:16" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "−2:29" "−2:28" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "5:13 am" "5:14 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "6:42 pm" "6:41 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "5:45 am" "5:46 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "6:09 pm" "6:08 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "6:18 am" "6:19 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "5:36 pm" "5:35 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "11:58 am (33.2°)" "11:58 am (32.9°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "92.240" "92.216" ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...  
## ..$ : chr [1:33] NA "Note: hours shift because clocks change backward 1 hour. (See the note below this table for details)" NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

## List of 2  
## $ : tibble [7 × 2] (S3: tbl\_df/tbl/data.frame)  
## ..$ X1: chr [1:7] "Current Time:" "Sun Direction:" "Sun Altitude:" "Sun Distance:" ...  
## ..$ X2: chr [1:7] "Dec 5, 2022 at 11:38:38 pm" "↑ 342° North" "-69.7°" "91.595 million mi" ...  
## $ : tibble [33 × 26] (S3: tbl\_df/tbl/data.frame)  
## ..$ 2020 : chr [1:33] "Dec" "Dec" "2" "3" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunrise" "Sunrise" "7:24 am ↑ (120°)" "7:25 am ↑ (120°)" ...  
## ..$ Sunrise/Sunset : chr [1:33] "Sunset" "Sunset" "4:43 pm ↑ (240°)" "4:43 pm ↑ (240°)" ...  
## ..$ Daylength : chr [1:33] "Length" "Length" "9:18:48" "9:17:34" ...  
## ..$ Daylength : chr [1:33] "Diff." "Diff." "−1:16" "−1:13" ...  
## ..$ Astronomical Twilight: chr [1:33] "Start" "Start" "5:44 am" "5:45 am" ...  
## ..$ Astronomical Twilight: chr [1:33] "End" "End" "6:23 pm" "6:23 pm" ...  
## ..$ Nautical Twilight : chr [1:33] "Start" "Start" "6:18 am" "6:19 am" ...  
## ..$ Nautical Twilight : chr [1:33] "End" "End" "5:49 pm" "5:49 pm" ...  
## ..$ Civil Twilight : chr [1:33] "Start" "Start" "6:53 am" "6:54 am" ...  
## ..$ Civil Twilight : chr [1:33] "End" "End" "5:14 pm" "5:14 pm" ...  
## ..$ Solar Noon : chr [1:33] "Time" "Time" "12:04 pm (25.9°)" "12:04 pm (25.8°)" ...  
## ..$ Solar Noon : chr [1:33] "Mil. mi" "Mil. mi" "91.632" "91.618" ...  
## ..$ : int [1:33] NA 1 NA NA NA NA NA NA NA NA ...  
## ..$ : chr [1:33] NA "7:23 am ↑ (119°)" NA NA ...  
## ..$ : chr [1:33] NA "4:43 pm ↑ (241°)" NA NA ...  
## ..$ : chr [1:33] NA "9:20:04" NA NA ...  
## ..$ : chr [1:33] NA "−1:20" NA NA ...  
## ..$ : chr [1:33] NA "5:43 am" NA NA ...  
## ..$ : chr [1:33] NA "6:23 pm" NA NA ...  
## ..$ : chr [1:33] NA "6:17 am" NA NA ...  
## ..$ : chr [1:33] NA "5:49 pm" NA NA ...  
## ..$ : chr [1:33] NA "6:52 am" NA NA ...  
## ..$ : chr [1:33] NA "5:14 pm" NA NA ...  
## ..$ : chr [1:33] NA "12:03 pm (26.0°)" NA NA ...  
## ..$ : num [1:33] NA 91.6 NA NA NA ...

## Joining, by = c("Date", "Sunrise", "Sunset", "Daylight")  
## Joining, by = c("Date", "Sunrise Time", "Sunset Time", "Daylight", "Month")

#Checking if there are any null values   
sum(is.na(sun$`Sunrise Time`))

## [1] 0

sum(is.na(sun$`Date`))

## [1] 0

sum(is.na(sun$`Sunset Time`))

## [1] 0

sum(is.na(sun$`Daylight`))

## [1] 0

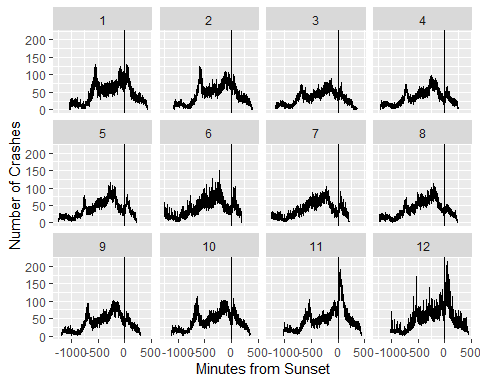
crashes$Month<- gsub("Month ", "\\1", crashes$Month)  
crashes$Month<- parse\_integer(crashes$Month)  
  
crashes$Day <- sub("^0+", "", crashes$Day)   
crashes$Day<- parse\_integer(crashes$Day)

#I needed to convert the Date back to integer so I could join with original data  
sun$Date <- as.integer(sun$Date)  
   
#Joining the original data and scraped data by day and month  
df<- left\_join(crashes, sun, by = c("Day" = "Date", "Month" = "Month"))

#Finding the difference between the sunset time and the time of the crash  
df$sunsetDiff = difftime(df$`Time of Crash`, df$`Sunset Time`, unit = 'mins')  
df$sunriseDiff = difftime(df$`Time of Crash`, df$`Sunrise Time`, unit = 'mins')  
df$sunDiff <- with(df,pmin(abs(sunsetDiff), abs(sunriseDiff)))

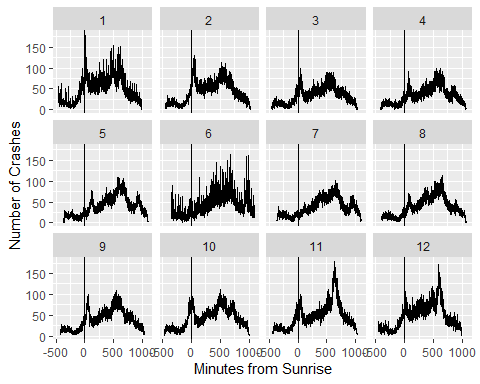
#Looking at the difference between sunset and time of the crash by month  
df%>%  
 group\_by(Month, sunsetDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunsetDiff, y= n)) + geom\_line() + facet\_wrap(~Month)+geom\_vline(xintercept = 0) + xlab("Minutes from Sunset")+ylab("Number of Crashes")

## `summarise()` has grouped output by 'Month'. You can override using the  
## `.groups` argument.  
## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



#Looking at the difference between sunrise and time of the crash by month  
df%>%  
 group\_by(Month, sunriseDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunriseDiff, y= n)) + geom\_line() + facet\_wrap(~Month)+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunrise")+ylab("Number of Crashes")

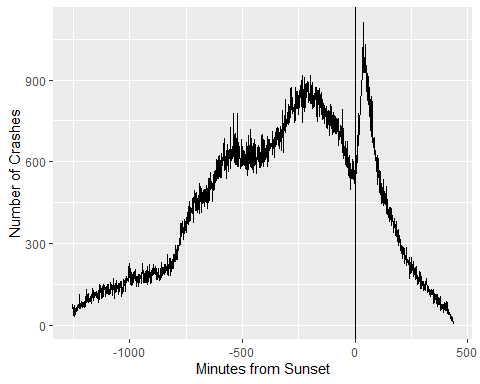
## `summarise()` has grouped output by 'Month'. You can override using the  
## `.groups` argument.  
## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



We do have some skewed left graphs - meaning crashes are more likely to be close to sunset

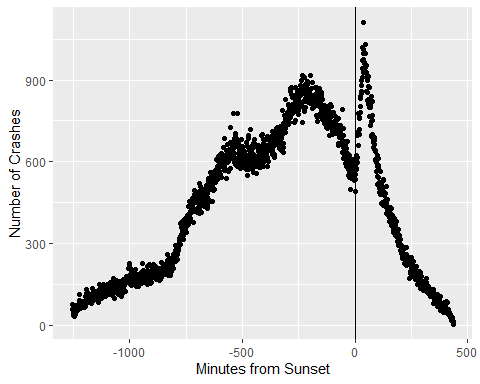
#A graph of all crashes by the difference between crash time and sunset  
df%>%  
 group\_by(sunsetDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunsetDiff, y = n)) + geom\_line()+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunset")+ylab("Number of Crashes")

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



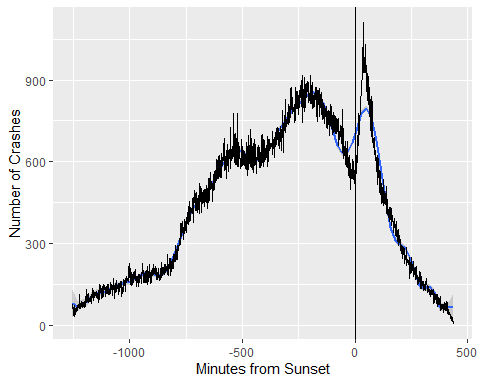
#Shows the same in a scatterplot, just in case there are outliers  
df%>%  
 group\_by(sunsetDiff)%>%  
 summarise("Number of Crashes" = n())%>%  
 ggplot(aes(x = sunsetDiff, y = `Number of Crashes`)) + geom\_point()+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunset")+ylab("Number of Crashes")

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



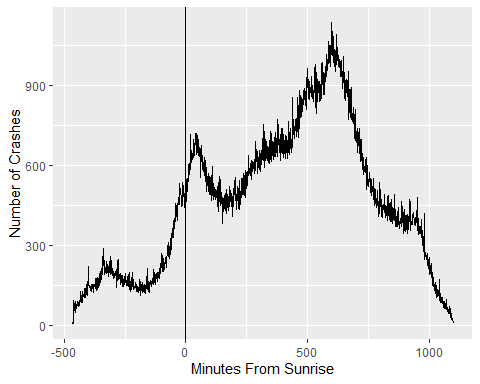
#A fitted line for the difference between sunset and crash time  
df%>%  
 group\_by(sunsetDiff)%>%  
 summarise("Number of Crashes" = n())%>%  
 ggplot(aes(x = sunsetDiff, y = `Number of Crashes`)) + geom\_smooth(method = "gam", formula = y ~ poly(x, 27)) + geom\_line() +geom\_vline(xintercept = 0) + xlab("Minutes from Sunset")

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



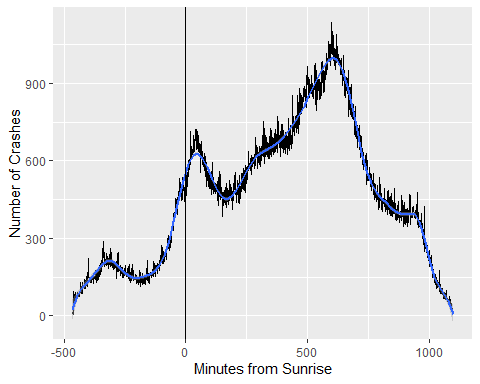
#A graph of all crashes by the difference between crash time and sunrise, with a fitted line  
df%>%  
 group\_by(sunriseDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunriseDiff, y = n)) + geom\_line() + ylab("Number of Crashes") + xlab("Minutes From Sunrise") + geom\_vline(xintercept = 0)

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



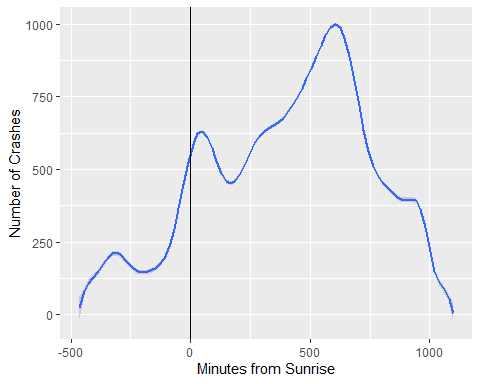
#A graph of all crashes by the difference between crash time and sunrise, with a fitted line  
df%>%  
 group\_by(sunriseDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunriseDiff, y = n)) + geom\_line() + geom\_smooth(method = "gam", formula = y ~ poly(x, 25))+geom\_vline(xintercept = 0) + xlab("Minutes from Sunrise")+ylab("Number of Crashes")

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



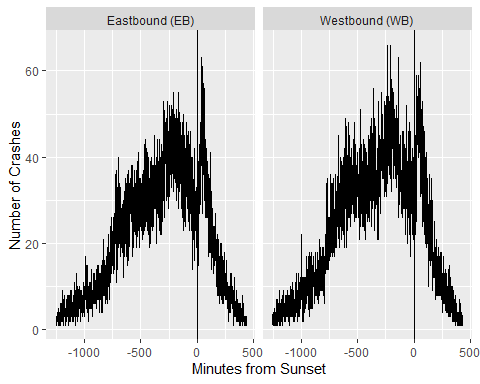
#Just the fitted line from the previous data  
df%>%  
 group\_by(sunriseDiff)%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = sunriseDiff, y = n))+ geom\_smooth(method = "gam", formula = y ~ poly(x, 25))+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunrise")+ylab("Number of Crashes")

## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



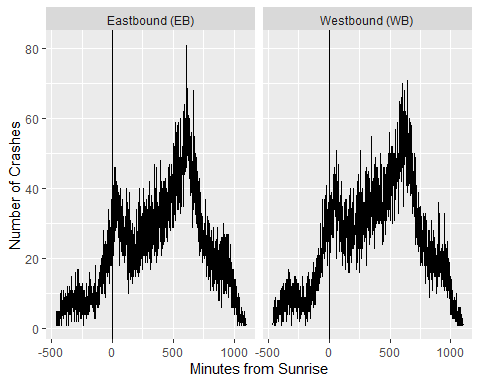
#Looking at sunset vs. crash time going different directions  
df%>%  
 group\_by(sunsetDiff, `Travel Direction`)%>%  
 filter(!is.na(`Travel Direction`), `Travel Direction` %in% c('Westbound (WB)', 'Eastbound (EB)'))%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = `sunsetDiff`, y = n)) + geom\_line() + facet\_wrap(~`Travel Direction`)+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunset")+ylab("Number of Crashes")

## `summarise()` has grouped output by 'sunsetDiff'. You can override using the  
## `.groups` argument.  
## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



#Looking at sunrise vs. crash time going different directions  
df%>%  
 group\_by(sunriseDiff, `Travel Direction`)%>%  
 filter(!is.na(`Travel Direction`), `Travel Direction` %in% c('Westbound (WB)', 'Eastbound (EB)'))%>%  
 summarise(n = n())%>%  
 ggplot(aes(x = `sunriseDiff`, y = n)) + geom\_line() + facet\_wrap(~`Travel Direction`)+geom\_vline(xintercept = 0)+ xlab("Minutes from Sunrise")+ylab("Number of Crashes")

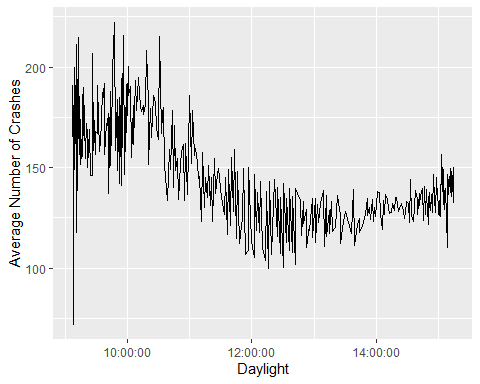
## `summarise()` has grouped output by 'sunriseDiff'. You can override using the  
## `.groups` argument.  
## Don't know how to automatically pick scale for object of type <difftime>.  
## Defaulting to continuous.



dfDaylight<-df%>%  
group\_by(Daylight, `Date of Crash`)%>%  
 summarise(n = n()/length(unique(`Date of Crash`)))%>%  
 mutate(uniqueDate = length(unique(`Date of Crash`)))%>%  
 mutate(total = sum(n)/uniqueDate)

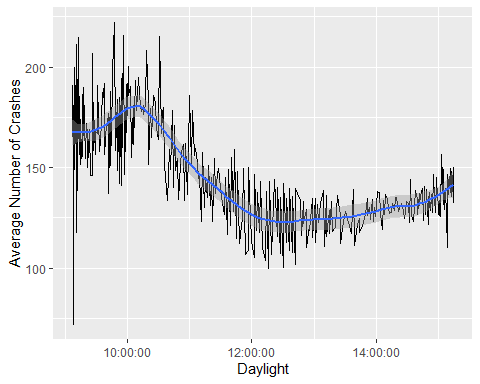
## `summarise()` has grouped output by 'Daylight'. You can override using the  
## `.groups` argument.

dfDaylight <- dfDaylight[!duplicated(dfDaylight$total), ]  
  
dfDaylight%>%  
 ggplot(aes(x = Daylight, y = total)) + geom\_line() + ylab("Average Number of Crashes")



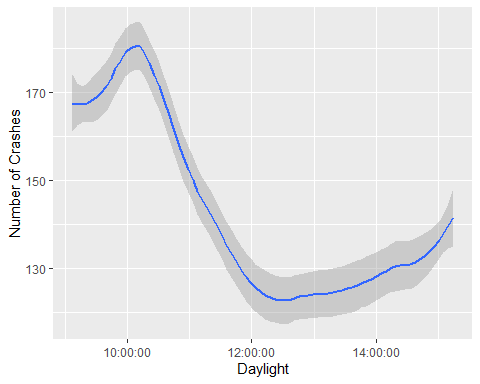
dfDaylight%>%  
 ggplot(aes(x = Daylight, y = total)) + geom\_line() + ylab("Average Number of Crashes") + geom\_smooth(span = 0.3)

## `geom\_smooth()` using method = 'loess' and formula = 'y ~ x'



dfDaylight%>%  
 ggplot(aes(x = Daylight, y = total))+ geom\_smooth(span= 0.3)+ylab("Number of Crashes")

## `geom\_smooth()` using method = 'loess' and formula = 'y ~ x'



# Creates new data set to do modifications on  
c <- crashes  
  
# Changes the drug or alcohol column to a factor so that all the levels can be  
# found  
c$`Drug or Alcohol` <- as.factor(c$`Drug or Alcohol`)  
  
# Finds the levels of the column  
levels(c$`Drug or Alcohol`)

## [1] "Alcohol (< Statutory)"   
## [2] "Alcohol (Statutory)"   
## [3] "Drug"   
## [4] "Drug/Alcohol (< Statutory)"   
## [5] "Drug/Alcohol (Statutory)"   
## [6] "None Indicated"   
## [7] "Refused"   
## [8] "Under Influence of Alcohol/Drugs/Medications"

c %>%  
 filter( (`Drug or Alcohol` == "None Indicated") | (`Drug or Alcohol` == "Refused")) %>%  
 select(`Drug or Alcohol`)

## # A tibble: 698,423 × 1  
## `Drug or Alcohol`  
## <fct>   
## 1 None Indicated   
## 2 None Indicated   
## 3 None Indicated   
## 4 None Indicated   
## 5 None Indicated   
## 6 None Indicated   
## 7 None Indicated   
## 8 None Indicated   
## 9 None Indicated   
## 10 None Indicated   
## # … with 698,413 more rows

# Creates new column that combines all the levels of substance abuse  
c <- c %>%  
 mutate(Drug\_Usage =   
 (`Drug or Alcohol` != "None Indicated") &  
 (`Drug or Alcohol` != "Refused"))

# Total number of crashes without substance abuse  
(false\_rows <- c %>%  
 filter(Drug\_Usage == FALSE) %>%  
 select(`Amount of Property Damage`) %>%  
 nrow())

## [1] 698423

# Total amount of crashes with substance abuse  
(true\_rows <- nrow(c) - false\_rows)

## [1] 30019

# Total cost of crashes without substance abuse  
(false\_total\_cost <- c %>%  
 filter(Drug\_Usage == FALSE) %>%  
 select(`Amount of Property Damage`) %>%  
 sum(na.rm=TRUE))

## [1] 4348396923

# Total cost of crashes with substance abuse  
(true\_total\_cost <- c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 select(`Amount of Property Damage`) %>%  
 sum(na.rm=TRUE))

## [1] 270017936

sum(c$`Amount of Property Damage`, na.rm = TRUE) / nrow(c)

## [1] 6340.127

# Average property damage of crash without substance abuse  
false\_total\_cost / false\_rows

## [1] 6226.022

# Average property damage of crash with substance abuse  
true\_total\_cost / true\_rows

## [1] 8994.901

# Total number of injuries with and without substance abuse  
(false\_total\_inj <- c %>%  
 filter(Drug\_Usage == FALSE) %>%  
 select(`Number of Injuries`) %>%  
 sum(na.rm=TRUE))

## [1] 246078

(true\_total\_inj <- c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 select(`Number of Injuries`) %>%  
 sum(na.rm=TRUE))

## [1] 18043

# Average injuries per accident without substance abuse  
false\_total\_inj / false\_rows

## [1] 0.3523338

# Average injuries per accident with substance abuse  
true\_total\_inj / true\_rows

## [1] 0.6010527

# Total number of fatalities with and without substance abuse  
(false\_total\_fatal <- c %>%  
 filter(Drug\_Usage == FALSE) %>%  
 select(`Number of Fatalities`) %>%  
 sum(na.rm=TRUE))

## [1] 3214

(true\_total\_fatal <- c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 select(`Number of Fatalities`) %>%  
 sum(na.rm=TRUE))

## [1] 1557

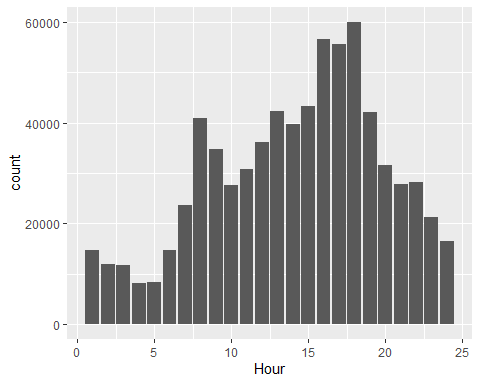
# Average fatalities per accident without substance abuse  
false\_total\_fatal / false\_rows

## [1] 0.004601796

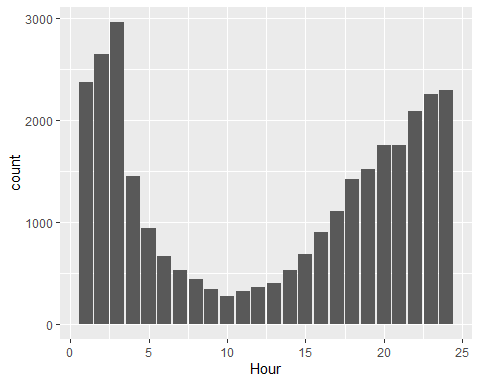
# Average fatalities per accident with substance abuse  
true\_total\_fatal / true\_rows

## [1] 0.05186715

#c <- c %>%  
# mutate(Hour = parse\_number(Hour))  
  
# Plot of all crashes by the hour  
c %>%  
 ggplot(aes(x=Hour)) + geom\_bar()



# Most crashes happen during "rush hour" from 4 to 6 P.M.  
  
# Plot of all crashes with sustance abuse by the hour  
c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 ggplot(aes(x=Hour)) + geom\_bar()



# Most crashes with substances involved happen from 11 P.M. to 3 A.M.  
  
c\_night <- c %>%  
 filter(Hour >= 22 | Hour <= 3)  
  
(total\_night <- nrow(c\_night))

## [1] 104123

(true\_total\_night <- c\_night %>%  
 filter(Drug\_Usage == TRUE) %>%  
 nrow())

## [1] 14619

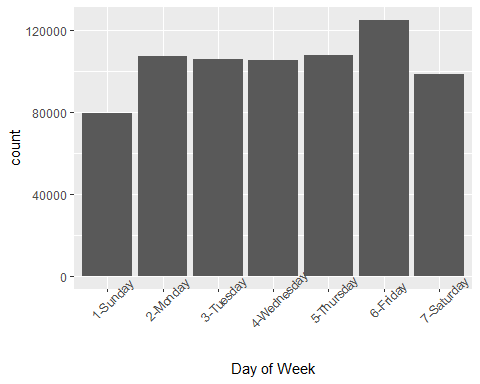
# Crashes with substances are 4% of all crashes  
true\_rows / nrow(c)

## [1] 0.04120987

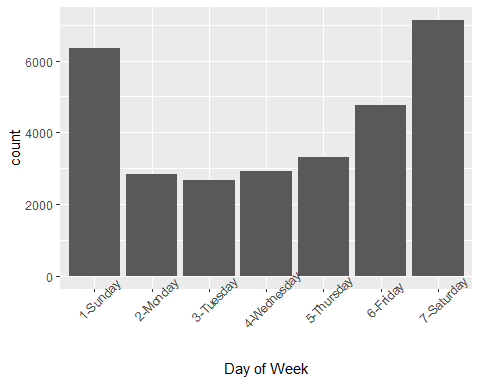
# Crashes with substances are 14% of all crashes from 10 P.M. to 3 A.M.  
true\_total\_night / total\_night

## [1] 0.1404013

# Plot of all crashes by day of the week  
c %>%  
 ggplot(aes(x=`Day of Week`)) + geom\_bar() + theme(axis.text.x = element\_text(angle = 45))

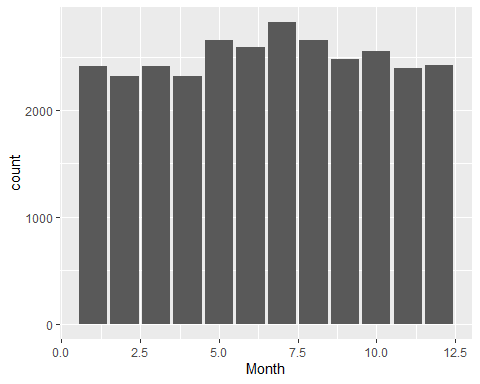


# Crashes occur pretty much uniformly throughout the week  
#  
# Slight uptick on Friday, people driving home more recklessly to get  
# home faster?  
  
# Plot of all crashes with substance abuse by day of the week  
c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 ggplot(aes(x=`Day of Week`)) + geom\_bar() + theme(axis.text.x = element\_text(angle = 45))

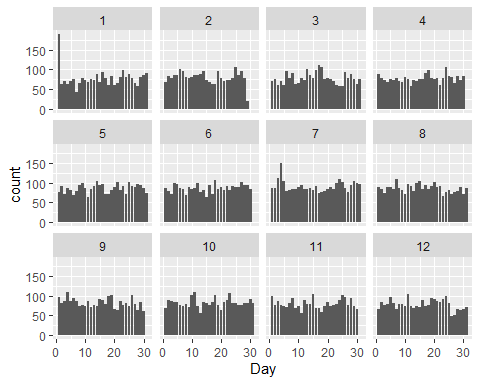


# Majority of these crashes occur on the weekends, Saturday and Sunday are  
# likely higher than Friday because they happen early in the "morning" after  
# Friday and Saturday nights

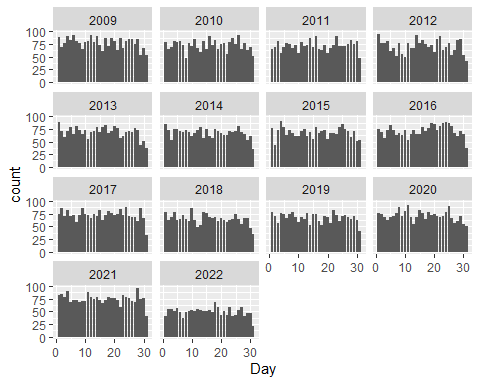
c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 ggplot(aes(x=Month)) + geom\_bar()



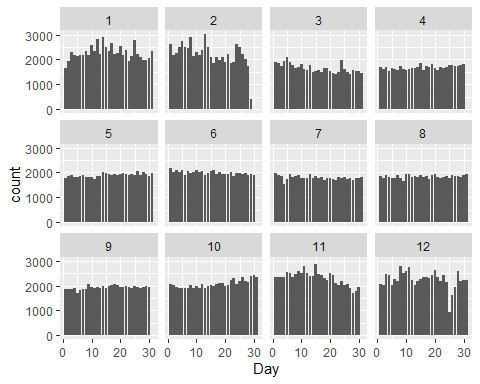
c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 ggplot(aes(x=Day)) + geom\_bar() + facet\_wrap(~Month)



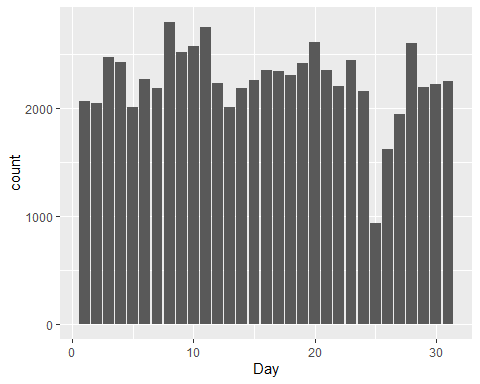
c %>%  
 filter(Drug\_Usage == TRUE) %>%  
 ggplot(aes(x=Day)) + geom\_bar() + facet\_wrap(~Year)



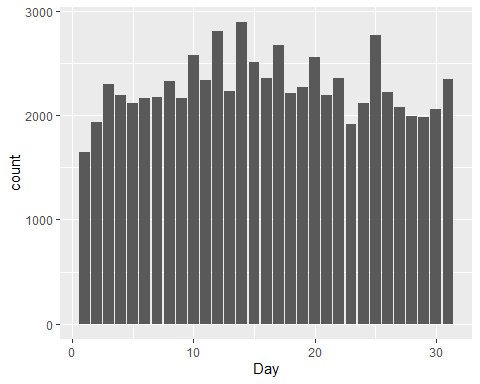
c %>%  
 ggplot(aes(x=Day)) + geom\_bar() + facet\_wrap(~Month)



c %>%  
 filter(Month == 12) %>%  
 ggplot(aes(x=Day)) + geom\_bar()



c %>%  
 filter(Month == 1) %>%  
 ggplot(aes(x=Day)) + geom\_bar()



## Bibliography

1 <https://delawaresafety.org/resources/Documents/Safety%20Documents/Safe%20Driving%20-%20%20Dusk%20and%20Dawn.pdf>

2 <https://iowadot.gov/mvd/stats/crashhistory.pdf>