Accident Findings Report

Luke Cadagin

11/24/2021

Report of Grand Rapids Accident Analysis

Load necessary packages:

```
library(tidyverse)
library(sf)
library(osmdata)
library(ggpubr)
```

Specify size of all graphs in Knitted Documents:

```
knitr::opts_chunk$set(echo = TRUE, fig.width = 16, fig.height = 8)
```

Upload Grand Rapids Crash Dataset (2008 - 2017):

```
crash_data <- read_csv(here::here("data", "CGR_Crash_Data.csv"))
head(crash_data)
```

```
## # A tibble: 6 x 142
##       X     Y OBJECTID ROADSOFTID BIKE  CITY     COUNTY CRASHDATE  CRASHSEVER
##   <dbl> <dbl>    <dbl>      <dbl> <chr> <chr>    <chr>  <date>     <chr>
## 1 -85.7  42.9        1    2589528 No    Grand R~ Kent   2008-06-16 Property Dam~
## 2 -85.6  42.9        2    2593183 No    Grand R~ Kent   2008-08-30 Property Dam~
## 3 -85.7  43.0        3    2582102 No    Grand R~ Kent   2008-02-13 Property Dam~
## 4 -85.6  42.9        4    2579820 No    Grand R~ Kent   2008-01-25 Property Dam~
## 5 -85.7  43.0        5    2594624 No    Grand R~ Kent   2008-09-26 Property Dam~
## 6 -85.7  43.0        6    2599372 No    Grand R~ Kent   2008-12-13 Property Dam~
## # ... with 133 more variables: CRASHTYPE <chr>, WORKZNEACT <chr>,
## #   WORKZNECLO <chr>, WORKZNETYP <chr>, CTRLMILEPT <dbl>, CTRLSECT <dbl>,
## #   DAYOFMONTH <dbl>, DAYOFWEEK <chr>, ANIMAL <chr>, D1COND <chr>,
## #   D1DRINKIN <chr>, D1HAZACT <chr>, D1INJURY <chr>, D1INTENT <chr>,
## #   D2COND <chr>, D2DRINKIN <chr>, D2HAZACT <chr>, D2INJURY <chr>,
## #   D2INTENT <chr>, D3COND <chr>, D3DRINKIN <chr>, D3HAZACT <chr>,
## #   D3INJURY <chr>, D3INTENT <chr>, DRINKING <chr>, DRIVER1AGE <dbl>,
## #   DRIVER1SEX <chr>, DRIVER2AGE <dbl>, DRIVER2SEX <chr>, DRIVER3AGE <dbl>,
## #   DRIVER3SEX <chr>, EMRGVEH <chr>, FARMEQUIP <chr>, FLEEINGSIT <chr>,
## #   FWSEGID <dbl>, GRTINJSEVE <chr>, HITANDRUN <chr>, HOUR <dbl>,
## #   INTERNAME <chr>, LIGHTING <chr>, MDOTREG <chr>, MILEPOINT <dbl>,
## #   MONTH <chr>, MOTORCYCLE <chr>, NOATYPEINJ <dbl>, NOBTYPEINJ <dbl>,
## #   NOCTYPEINJ <dbl>, NONTRAFFIC <chr>, NUMOFINJ <dbl>, NUMOFKILL <dbl>,
## #   NUMOFOCCUP <dbl>, NUMOFUNINJ <dbl>, NUMOFVEHIC <dbl>, ORV <chr>,
## #   PEDESTRIAN <chr>, PRNAME <chr>, PRNO <dbl>, PUBLICPROP <chr>, REFDIR <chr>,
## #   REFDIST <chr>, ROUTECLASS <chr>, ROUTENUM <dbl>, SCHOOLBUS <chr>,
## #   SNOWMOBILE <chr>, SPDLMTPOST <chr>, SPEEDLIMIT <dbl>, SURFCOND <chr>,
```

```
## #   TRAFCTLDEV <chr>, TRAIN <chr>, TRUCKBUS <chr>, TRUNKLINE <chr>,
## #   UD10NUM <dbl>, V1DEFECT <chr>, V1DAMAGE <chr>, V1HARMEVT1 <chr>,
## #   V1HARMEVT2 <chr>, V1HARMEVT3 <chr>, V1HARMEVT4 <chr>, V1MSTHARME <chr>,
## #   V1SPECCAT <chr>, V1TRAILER <chr>, V1VIOLATOR <chr>, V1WIMPCTPT <chr>,
## #   V2DEFECT <chr>, V2DAMAGE <chr>, V2HARMEVT1 <chr>, V2HARMEVT2 <chr>,
## #   V2HARMEVT3 <chr>, V2HARMEVT4 <chr>, V2MSTHARME <chr>, V2SPECCAT <chr>,
## #   V2TRAILER <chr>, V2VIOLATOR <chr>, V2WIMPCTPT <chr>, V3DEFECT <chr>,
## #   V3DAMAGE <chr>, V3HARMEVT1 <chr>, V3HARMEVT2 <chr>, V3HARMEVT3 <chr>,
## #   V3HARMEVT4 <chr>, ...
```

Configure features necessary for Grand Rapids map (using openstreetmap api)

```
location_gr <- getbb("Grand Rapids") %>%
    opq()

major_roads_gr <- location_gr %>%
    add_osm_feature(key = "highway", value = c("motorway", "trunk", "primary", "secondary", "tertiary")) %>%
    osmdata_sf()

#minor_roads_gr <- location_gr %>%
    #add_osm_feature(key = "highway", value = c("unclassified", "residential")) %>%
    #osmdata_sf()

water_gr <- location_gr %>%
    add_osm_feature(key = "waterway", value = c("river")) %>%
    osmdata_sf()

boundary_gr <- location_gr %>%
    add_osm_feature(key = "boundary", value = "administrative") %>%
    add_osm_feature(key = "name", value = "Grand Rapids") %>%
    osmdata_sf()
```

```
## Request failed [429]. Retrying in 1 seconds...
```

Grand Rapids accidents associated with trains:

I am interested in studying the impact gates at a railroad crossing have on the number of accidents associated with trains.

Let's start by visualizing the number of Grand Rapids crashes associated with a train from 2008 - 2017.

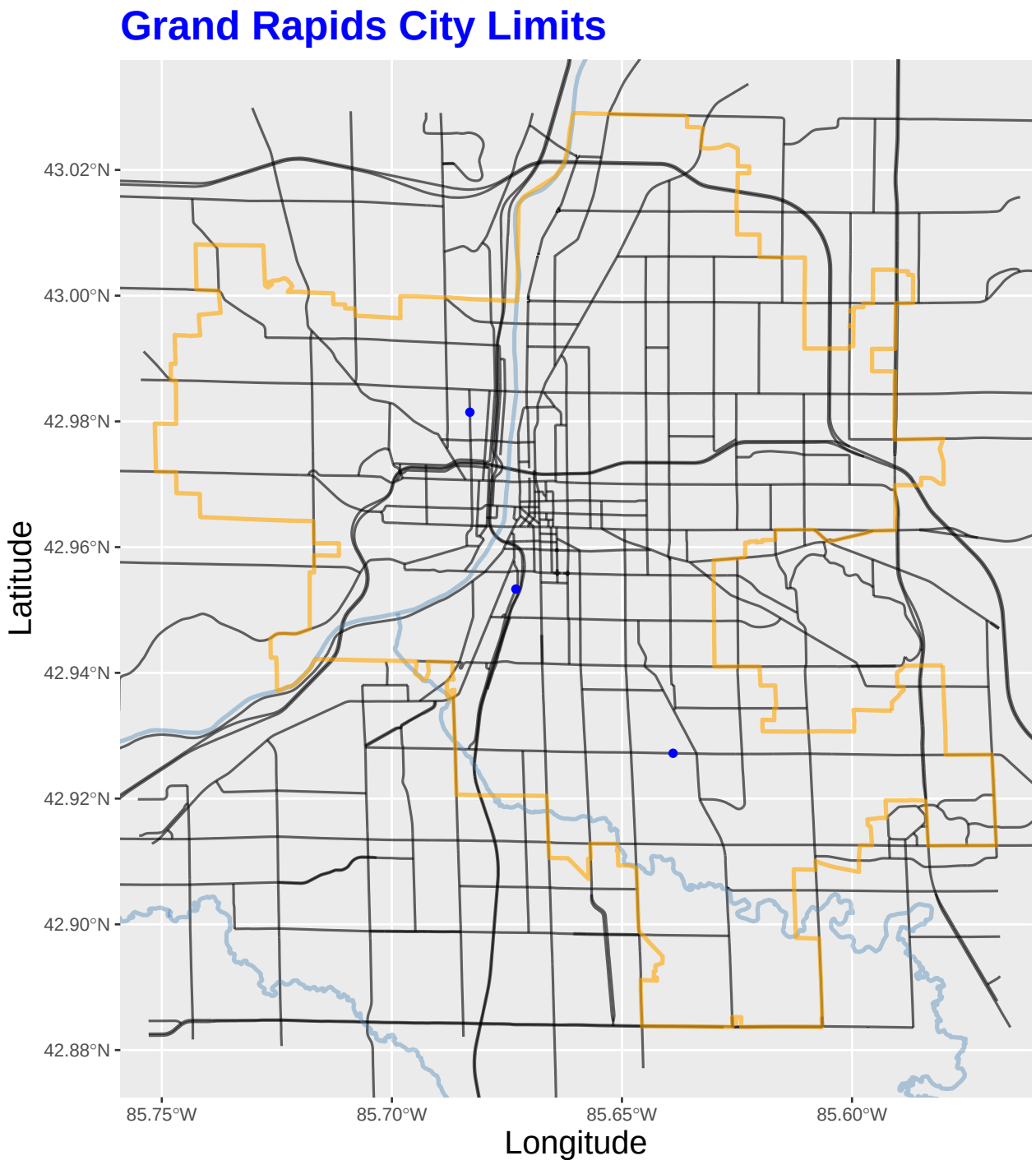
```
crash_data_train <- crash_data %>%
  filter(TRAIN == "Yes")

crash_data_train %>%
  select(`Longitude` = X, `Latitude` = Y, CRASHDATE, TRAIN, `Principal Road` = PRNAME)
```

```
## # A tibble: 3 x 5
##   Longitude Latitude CRASHDATE  TRAIN `Principal Road`
##       <dbl>    <dbl> <date>     <chr> <chr>
## 1     -85.6     42.9 2008-02-19 Yes   BURTON
## 2     -85.7     43.0 2014-05-03 Yes   11TH
## 3     -85.7     43.0 2017-12-27 Yes   CENTURY
```

```
  ggplot()+
            geom_sf(data = major_roads_gr$osm_lines, size = .6, alpha = .6, color = 'black') +
            #geom_sf(data = minor_roads_gr$osm_lines, size = .3, alpha = .3, color = 'black') +
```

```
  geom_sf(data = water_gr$osm_lines, size = 1, alpha = .4, color = 'steelblue') +
  geom_sf(data = boundary_gr$osm_lines, size = 1, alpha = .6, color = "orange") +
  geom_point(data = crash_data_train, mapping = aes(x = X, y = Y), color = "blue") +
  coord_sf(xlim = c(-85.57, -85.75), ylim = c(42.88, 43.03)) +
  labs(title = "Grand Rapids City Limits", x = "Longitude", y = "Latitude") +
  font("title", size = 20, color = "blue", face = "bold") +
  font("x", size = 16) +
  font("y", size = 16)
```



As you can see, from 2008 to 2017 there were only three accidents that occurred in Grand Rapids involving a train (In 17, 14, and 08).

Although we have a very limited number of accidents directly involving a train, this is not a dead-end for our analysis.

Next let's upload a a dataset from Transportation.gov (https://data.transportation.gov/Railroads/Crossing-Inventory-Data-Current/m2f8-22s6) that provides information about every railroad crossing in the USA:

```
rr_crossing_data <- read_csv(here::here("data", "Crossing_Inventory_Data_-_Current.csv"))
```

```
## Warning: 6574718 parsing failures.
##  row                                        col           expected              actual
## 1224 Number Other MUTCD 1                       1/0/T/F/TRUE/FALSE 2                  '/home/cadag
## 1349 Highway Traffic Signal Interconnection Code 1/0/T/F/TRUE/FALSE 2                  '/home/cadag
## 1349 Highway Traffic Signal Interconnection      1/0/T/F/TRUE/FALSE For Traffic Signals '/home/cadag
## 1349 Highway Traffic Signal Preemption           1/0/T/F/TRUE/FALSE Simultaneous        '/home/cadag
## 1350 Highway Traffic Signal Interconnection Code 1/0/T/F/TRUE/FALSE 2                  '/home/cadag
## .... ......................................... ................... ................... .....
## See problems(...) for more details.
```

```
head(rr_crossing_data)
```

```
## # A tibble: 6 x 248
##   `Revision Date` `Reporting Agency Type I~ `Reporting Agency Typ~ `Reason Code`
##   <date>                              <dbl> <chr>                            <dbl>
## 1 1970-01-01                              1 Railroad                            15
## 2 1970-01-01                              1 Railroad                            15
## 3 1970-01-01                              1 Railroad                            15
## 4 1970-01-01                              1 Railroad                            15
## 5 1970-01-01                              1 Railroad                            15
## 6 1970-01-01                              1 Railroad                            15
## # ... with 244 more variables: Reason Description <chr>, Crossing ID <chr>,
## #   Crossing ID Suffix <chr>, Reporting Agency Code <chr>,
## #   Reporting Agency Name <chr>, State Code <chr>, State Name <chr>,
## #   County Code <chr>, County Name <chr>, In/Near Code <dbl>, In/Near <chr>,
## #   City Code <chr>, City Name <chr>, City Description <chr>, Street <chr>,
## #   Block Number <lgl>, Highway Name <chr>, Separate Track <chr>,
## #   Separate Track Railroad 1 <chr>, Separate Track Railroad 2 <lgl>,
## #   Separate Track Railroad 3 <lgl>, Separate Track Railroad 4 <lgl>,
## #   Same Track <chr>, Same Track Railroad 1 <chr>, Same Track Railroad 2 <lgl>,
## #   Same Track Railroad 3 <lgl>, Same Track Railroad 4 <lgl>,
## #   Railroad Division <chr>, Railroad Subdivision <chr>, Branch Name <chr>,
## #   Railroad Milepost Prefix <chr>, Railroad Milepost Number <chr>,
## #   Railroad Milepost Suffix <lgl>, Line Segment <chr>,
## #   Nearest Timetable Station <chr>, Timetable Station <dbl>,
## #   Parent Railroad Code <lgl>, Crossing Owner Code <lgl>,
## #   Crossing Type Code <dbl>, Crossing Type <chr>, Crossing Purpose Code <dbl>,
## #   Crossing Purpose <chr>, Crossing Position Code <dbl>,
## #   Crossing Position <chr>, Public Access <lgl>,
## #   Type Of Train Service IDs <lgl>, Type Of Train Service ID 1 <lgl>,
## #   Type Of Train Service 1 <lgl>, Type Of Train Service ID 2 <lgl>,
## #   Type Of Train Service 2 <lgl>, Type Of Train Service ID 3 <lgl>,
## #   Type Of Train Service 3 <lgl>, Type Of Train Service ID 4 <lgl>,
## #   Type Of Train Service 4 <lgl>, Type Of Train Service ID 5 <lgl>,
## #   Type Of Train Service 5 <lgl>, Type Of Train Service ID 6 <lgl>,
```

```
## #   Type Of Train Service 6 <lgl>, Less Than One Passenger Train Per Day <lgl>,
## #   Number Passenger Train Per Day <dbl>, Development Type Code <dbl>,
## #   Development Type <chr>, Adjacent Crossing <lgl>,
## #   Adjacent Crossing Number <lgl>, Whistleban Code <dbl>, Whistle Ban <chr>,
## #   Whistle Date <lgl>, High-Speed Rail Corridor ID Suffix <lgl>,
## #   High-Speed Rail Corridor ID <lgl>, Latitude <dbl>, Longitude <dbl>,
## #   Lat/Long Source Code <dbl>, Lat/Long Source <chr>, Railroad Use <lgl>,
## #   Railroad Narrative <lgl>, State Use <lgl>, State Narrative <lgl>,
## #   Emergency Telephone Number <dbl>, Railroad Contact Telephone Number <lgl>,
## #   State Contact Telephone Number <dbl>, Total Daylight Thru Trains <dbl>,
## #   Total Nighttime Thru Trains <dbl>, Total Switching Trains <dbl>,
## #   Total Transit Trains <lgl>, Movements Per Day Code <lgl>,
## #   Movements Per Day <lgl>, Trains Per Week <lgl>,
## #   Trains Per Week Captured Year <lgl>, Maximum Timetable Speed <dbl>,
## #   Typical Minimum Speed Over Crossing <dbl>,
## #   Typical Maximum Speed Over Crossing <dbl>, Number Of Main Tracks <dbl>,
## #   Number Of Siding Tracks <dbl>, Number Of Yard Tracks <lgl>,
## #   Number Of Transit Tracks <lgl>, Number Of Industry Tracks <lgl>,
## #   Train Detection IDs <dbl>, Train Detection ID 1 <dbl>,
## #   Train Detection 1 <chr>, Train Detection ID 2 <lgl>, ...
```

Now we filter this data for only railroad crossing within the Grand Rapids city limits:

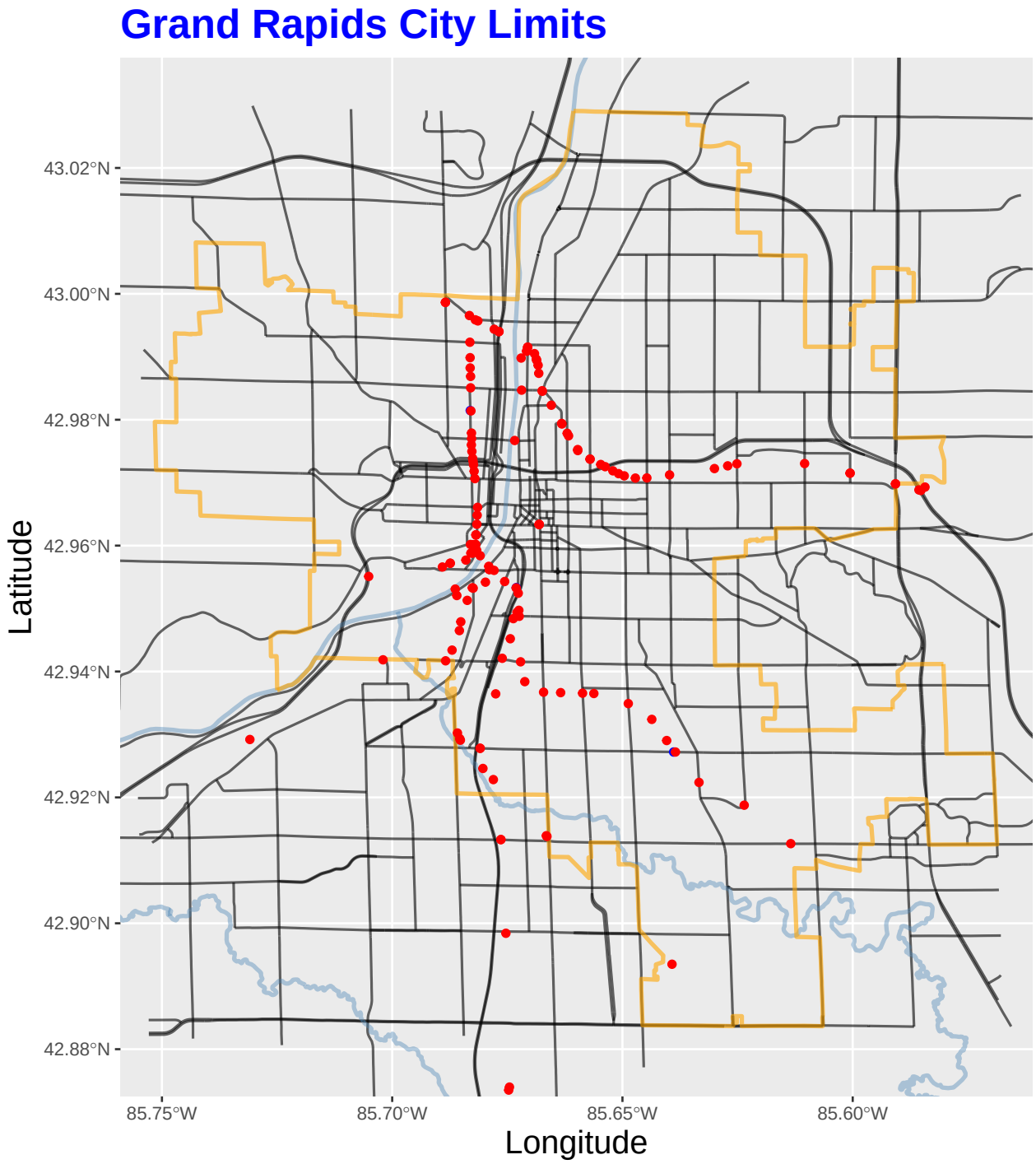
```
rr_crossing_data_gr <- rr_crossing_data %>%
  filter(`State Name` == "MICHIGAN", `City Name` == "GRAND RAPIDS", Latitude < 43)

head(rr_crossing_data_gr)
```

```
## # A tibble: 6 x 248
##   `Revision Date` `Reporting Agency Type I~ `Reporting Agency Typ~ `Reason Code`
##   <date>                              <dbl> <chr>                            <dbl>
## 1 1989-09-14                              2 State                               16
## 2 1991-04-03                              1 Railroad                            16
## 3 1991-04-03                              1 Railroad                            16
## 4 1991-04-03                              1 Railroad                            16
## 5 1991-04-03                              1 Railroad                            16
## 6 1991-04-03                              1 Railroad                            16
## # ... with 244 more variables: Reason Description <chr>, Crossing ID <chr>,
## #   Crossing ID Suffix <chr>, Reporting Agency Code <chr>,
## #   Reporting Agency Name <chr>, State Code <chr>, State Name <chr>,
## #   County Code <chr>, County Name <chr>, In/Near Code <dbl>, In/Near <chr>,
## #   City Code <chr>, City Name <chr>, City Description <chr>, Street <chr>,
## #   Block Number <lgl>, Highway Name <chr>, Separate Track <chr>,
## #   Separate Track Railroad 1 <chr>, Separate Track Railroad 2 <lgl>,
## #   Separate Track Railroad 3 <lgl>, Separate Track Railroad 4 <lgl>,
## #   Same Track <chr>, Same Track Railroad 1 <chr>, Same Track Railroad 2 <lgl>,
## #   Same Track Railroad 3 <lgl>, Same Track Railroad 4 <lgl>,
## #   Railroad Division <chr>, Railroad Subdivision <chr>, Branch Name <chr>,
## #   Railroad Milepost Prefix <chr>, Railroad Milepost Number <chr>,
## #   Railroad Milepost Suffix <lgl>, Line Segment <chr>,
## #   Nearest Timetable Station <chr>, Timetable Station <dbl>,
## #   Parent Railroad Code <lgl>, Crossing Owner Code <lgl>,
## #   Crossing Type Code <dbl>, Crossing Type <chr>, Crossing Purpose Code <dbl>,
## #   Crossing Purpose <chr>, Crossing Position Code <dbl>,
## #   Crossing Position <chr>, Public Access <lgl>,
```

```
## #   Type Of Train Service IDs <lgl>, Type Of Train Service ID 1 <lgl>,
## #   Type Of Train Service 1 <lgl>, Type Of Train Service ID 2 <lgl>,
## #   Type Of Train Service 2 <lgl>, Type Of Train Service ID 3 <lgl>,
## #   Type Of Train Service 3 <lgl>, Type Of Train Service ID 4 <lgl>,
## #   Type Of Train Service 4 <lgl>, Type Of Train Service ID 5 <lgl>,
## #   Type Of Train Service 5 <lgl>, Type Of Train Service ID 6 <lgl>,
## #   Type Of Train Service 6 <lgl>, Less Than One Passenger Train Per Day <lgl>,
## #   Number Passenger Train Per Day <dbl>, Development Type Code <dbl>,
## #   Development Type <chr>, Adjacent Crossing <lgl>,
## #   Adjacent Crossing Number <lgl>, Whistleban Code <dbl>, Whistle Ban <chr>,
## #   Whistle Date <lgl>, High-Speed Rail Corridor ID Suffix <lgl>,
## #   High-Speed Rail Corridor ID <lgl>, Latitude <dbl>, Longitude <dbl>,
## #   Lat/Long Source Code <dbl>, Lat/Long Source <chr>, Railroad Use <lgl>,
## #   Railroad Narrative <lgl>, State Use <lgl>, State Narrative <lgl>,
## #   Emergency Telephone Number <dbl>, Railroad Contact Telephone Number <lgl>,
## #   State Contact Telephone Number <dbl>, Total Daylight Thru Trains <dbl>,
## #   Total Nighttime Thru Trains <dbl>, Total Switching Trains <dbl>,
## #   Total Transit Trains <lgl>, Movements Per Day Code <lgl>,
## #   Movements Per Day <lgl>, Trains Per Week <lgl>,
## #   Trains Per Week Captured Year <lgl>, Maximum Timetable Speed <dbl>,
## #   Typical Minimum Speed Over Crossing <dbl>,
## #   Typical Maximum Speed Over Crossing <dbl>, Number Of Main Tracks <dbl>,
## #   Number Of Siding Tracks <dbl>, Number Of Yard Tracks <lgl>,
## #   Number Of Transit Tracks <lgl>, Number Of Industry Tracks <lgl>,
## #   Train Detection IDs <dbl>, Train Detection ID 1 <dbl>,
## #   Train Detection 1 <chr>, Train Detection ID 2 <lgl>, ...
```

```r
ggplot()+
          geom_sf(data = major_roads_gr$osm_lines, size = .6, alpha = .6, color = 'black') +
          #geom_sf(data = minor_roads_gr$osm_lines, size = .3, alpha = .3, color = 'black') +
          geom_sf(data = water_gr$osm_lines, size = 1, alpha = .4, color = 'steelblue') +
          geom_sf(data = boundary_gr$osm_lines, size = 1, alpha = .6, color = "orange") +
          geom_point(data = crash_data_train, mapping = aes(x = X, y = Y), color = "blue") +
          geom_point(data = rr_crossing_data_gr, mapping = aes(x = Longitude, y = Latitude), color =
          coord_sf(xlim = c(-85.57, -85.75), ylim = c(42.88, 43.03)) +
          labs(title = "Grand Rapids City Limits", x = "Longitude", y = "Latitude") +
          font("title", size = 20, color = "blue", face = "bold") +
          font("x", size = 16) +
          font("y", size = 16)
```


Grand Rapids City Limits
43.02°N
43.00°N
42.98°N
42.96°N
42.94°N
42.92°N
42.90°N
42.88°N
Latitude
85.75°W
85.70°W
85.65°W
85.60°W
Longitude