

NYC Traffic Accident Review

By Tony Fraser, CUNY MSDS Studnet

2023-09-03

Assignment:

Use Quarto and GT to ask and answer interesting questions about a dataset.

About this dataset

This is an official collection of all 2,021,793 New York City traffic accidents between 1-Jan-2020 and 31-Dec-2021 where reported damage was greater than 1,000 dollars. ¹

My big question, and my finding!

Though ask and answer was the assignment, my big ask was what kind of data is this. And my big finding is, NYC has tank accidents.



¹<https://catalog.data.gov/dataset/motor-vehicle-collisions-crashes>

Prepare part files

This original dataset is a 500Mb of CSV file, and was broken into smaller files using the unix `split` command.

- *It's a good rule of thumb to spend time making sure your part file loading system works correctly. In this example, code could be tested by loading either N data files like `data_files[1:2]` or all `data_files`.*
- *There is a file size limit of 50M to what you put on GIT. Accidentally uploading a massive file breaks your repo and workflow.*
- *Having this in git is not normal. It's almost always going loaded real time, or stored on S3 or in big query.*

Load the entire accidents data set

```
library(stringr)
library(dplyr)
library(gt)
library(data.table)

data_files = c("xaa.csv", "xac.csv", "xae.csv", "xag.csv", "xai.csv", "xak.csv",
               "xam.csv", "xao.csv", "xaq.csv", "xas.csv", "xau.csv", "xab.csv", "xad.csv",
               "xaf.csv", "xah.csv", "xaj.csv", "xal.csv", "xan.csv", "xap.csv", "xar.csv", "xat.csv")

# get all the data files into fully qualified urls.
git_repo <- "https://raw.githubusercontent.com/tonythor/cuny-datascience/"
base_url <- paste(git_repo, "develop/data/nyctraffic/csv_files/", sep="")
data_file_urls <- sapply(data_files, USE.NAMES = FALSE, function(i) str_c(base_url,i))

# load the headers into a vector.
header_url <- str_c(base_url, "../headers.csv")
headers <- colnames(fread(header_url, header=TRUE))

# load all urls into a vector of dataframes, each dataframe has correct headers.
df_list <- lapply(data_file_urls, function(i) fread(i, col.names = headers))

# Combine all data frames into one, most commonly called "unioning", but in R it's rbind.
accidents <- do.call(rbind, df_list)
```

Standardize some of the “human entry” records

This entire data set is based on police officers manually inserting data, so much of it needs to be reviewed and slightly recategorized.

```
vehicle_mappings <- list(
  Ambulance = c("AMBULANCE", "AMBU", "AMBUL", "ambulance", "Ambulance", "ambul", "Ambul",
    "AMB", "FDNY AMBUL"),
  Ebike = c("E-Bike", "Electric Bike", "E-bik", "E-Bik"),
  Motorcycle = c("Moped", "Motorbike", "Motorcycle", "MOPED", "Motorscooter", "moped",
    "MOTORCYCLE", "MOTOR SCOO"),
  Sedan = c("2 dr sedan", "4 dr sedan", "3-Door", "sedan", "Sedan"),
  SUV = c("Station Wagon/Sport Utility Vehicle", "SPORT UTILITY / STATION WAGON"),
  Taxi = c("Taxi", "TAXI", "taxi"),
  Bus = c("BUS", "Bus", "bus", "School", "School Bus", "SCHOO", "schoo"),
  Pickup = c("PICK-UP TRUCK", "PU", "PK", "Pick-up Truck", "TRUCK",
    "Open Body", "Open body", "truck", "Truck"),
  FDNY = c("FIRETRUCK", "FIRE TRUCK", "FDNY", "FIRE", "Fire", "fire", "FIRET",
    "FDNY FIRE", "Fire Truck", "FDNY TRUCK", "fdny"),
  Scooter = c("SCOOT", "E-Sco", "SCOOTER", "E-Scooter", "E-Sco"),
  Van = c("VAN", "van", "Van")
)

by_borough <- accidents %>%
  filter(!is.na(BOROUGH)
    & nchar(BOROUGH) > 0
    & `CONTRIBUTING FACTOR VEHICLE 1` != "unspecified") %>%
  mutate(`VEHICLE TYPE CODE 1` =
    case_when(
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Ambulance ~ "Ambulance_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Ebike ~ "Ebike_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Motorcycle ~ "Motorcycle_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Sedan ~ "Sedan_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$SUV ~ "SUV_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Taxi ~ "Taxi_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Bus ~ "Bus_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Pickup ~ "PickupTruck_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$FDNY ~ "FireTruck_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Scooter ~ "Scooter_",
      `VEHICLE TYPE CODE 1` %in% vehicle_mappings$Van ~ "Van_",
      TRUE ~ `VEHICLE TYPE CODE 1` )) %>%
  group_by(`VEHICLE TYPE CODE 1`, BOROUGH) %>%
  summarise(count = n(), .groups = "drop")
```

Create a match table to move vehicle type blocks around

Create a temporary e groupby/count/sort of the vehicle types. Then, we can join that table back to the data table to get an order of blocks. (Like SVU's, 1, cars, 2, taxis 3).

```
# Create ordering variable table
vehicle_order <- by_borough %>%
  group_by(`VEHICLE TYPE CODE 1`) %>%
  summarise(total_count = sum(count), .groups = "drop") %>%
  arrange(desc(total_count)) %>%
  mutate(order = row_number())

# Consolidate "Other" types and merge with the main data
by_borough_consolidated <- by_borough %>%
  left_join(vehicle_order, by = "VEHICLE TYPE CODE 1") %>%
  mutate(`VEHICLE TYPE CODE 1` = ifelse(count < 5, "Other", `VEHICLE TYPE CODE 1`)) %>%
  group_by(`VEHICLE TYPE CODE 1`, BOROUGH) %>%
  summarise(count = sum(count), order = first(order), .groups = "drop") %>%
  arrange(order, desc(count))
```

Use GT to chart

The purpose of this ask and answer project was to explore, thus a lot of raw data is scrollable in this study. Enjoy scrolling through it!

```
by_borough_consolidated %>%
  gt() %>%
  tab_style(
    style = cell_text(align = "left"),
    locations = cells_body(columns = c(`VEHICLE TYPE CODE 1`, BOROUGH, count))
  ) %>%
  cols_label(
    `VEHICLE TYPE CODE 1` = "Vehicle Code",
    BOROUGH = "Borough",
    count = "Accident Count",
    order = "ChartOrder"
  ) %>%
  tab_style(
    style = cell_text(weight = "bold", size = "large"),
    locations = cells_column_labels(columns = c(`VEHICLE TYPE CODE 1`, BOROUGH, count))
  )
```

Vehicle Code	Borough	Accident Count	ChartOrder
SUV__	BROOKLYN	136801	1
SUV__	QUEENS	124848	1
SUV__	MANHATTAN	72773	1
SUV__	BRONX	59281	1
SUV__	STATEN ISLAND	17436	1
Sedan__	BROOKLYN	129348	2
Sedan__	QUEENS	108068	2
Sedan__	BRONX	64585	2
Sedan__	MANHATTAN	64458	2
Sedan__	STATEN ISLAND	18156	2
PASSENGER VEHICLE	BROOKLYN	99924	3
PASSENGER VEHICLE	QUEENS	89081	3
PASSENGER VEHICLE	MANHATTAN	61731	3
PASSENGER VEHICLE	BRONX	42780	3
PASSENGER VEHICLE	STATEN ISLAND	16841	3
Taxi__	MANHATTAN	42714	4
Taxi__	BROOKLYN	7351	4
Taxi__	QUEENS	5732	4
Taxi__	BRONX	5254	4
Taxi__	STATEN ISLAND	96	4
PickupTruck__	BROOKLYN	9089	5
PickupTruck__	MANHATTAN	8395	5
PickupTruck__	QUEENS	8052	5
PickupTruck__	BRONX	4216	5
PickupTruck__	STATEN ISLAND	1743	5
Van__	MANHATTAN	11024	6
Van__	BROOKLYN	7835	6
Van__	QUEENS	4612	6
Van__	BRONX	2869	6
Van__	STATEN ISLAND	560	6
Bus__	MANHATTAN	8925	7
Bus__	BROOKLYN	7809	7
Bus__	QUEENS	4666	7
Bus__	BRONX	4126	7
Bus__	STATEN ISLAND	659	7
OTHER	BROOKLYN	6574	8
OTHER	MANHATTAN	5599	8
OTHER	QUEENS	2732	8
OTHER	BRONX	2668	8
OTHER	STATEN ISLAND	565	8

UNKNOWN	BROOKLYN	4981	9
UNKNOWN	QUEENS	4501	9
UNKNOWN	MANHATTAN	3142	9
UNKNOWN	BRONX	3093	9
UNKNOWN	STATEN ISLAND	570	9
Box Truck	MANHATTAN	6023	10
Box Truck	BROOKLYN	4739	10
Box Truck	QUEENS	2814	10
Box Truck	BRONX	2031	10
Box Truck	STATEN ISLAND	151	10
SMALL COM VEH(4 TIRES)	MANHATTAN	4763	11
SMALL COM VEH(4 TIRES)	BROOKLYN	2818	11
SMALL COM VEH(4 TIRES)	QUEENS	1911	11
SMALL COM VEH(4 TIRES)	BRONX	1138	11
SMALL COM VEH(4 TIRES)	STATEN ISLAND	129	11
LARGE COM VEH(6 OR MORE TIRES)	MANHATTAN	3613	12
LARGE COM VEH(6 OR MORE TIRES)	BROOKLYN	2878	12
LARGE COM VEH(6 OR MORE TIRES)	QUEENS	2216	12
LARGE COM VEH(6 OR MORE TIRES)	BRONX	1211	12
LARGE COM VEH(6 OR MORE TIRES)	STATEN ISLAND	95	12
Motorcycle__	BROOKLYN	3182	13
Motorcycle__	MANHATTAN	2582	13
Motorcycle__	QUEENS	2223	13
Motorcycle__	BRONX	1491	13
Motorcycle__	STATEN ISLAND	265	13
Bike	MANHATTAN	3502	14
Bike	BROOKLYN	3422	14
Bike	QUEENS	1680	14
Bike	BRONX	978	14
Bike	STATEN ISLAND	125	14
	BROOKLYN	3105	15
	QUEENS	2182	15
	MANHATTAN	2055	15
	BRONX	1847	15
	STATEN ISLAND	357	15
LIVERY VEHICLE	MANHATTAN	4447	16
LIVERY VEHICLE	BRONX	1789	16
LIVERY VEHICLE	BROOKLYN	1304	16
LIVERY VEHICLE	QUEENS	1008	16
LIVERY VEHICLE	STATEN ISLAND	23	16
Ambulance__	BROOKLYN	1698	17
Ambulance__	BRONX	1604	17
Ambulance__	MANHATTAN	1460	17

Ambulance__	QUEENS	913	17
Ambulance__	STATEN ISLAND	85	17
Tractor Truck Diesel	BROOKLYN	1553	18
Tractor Truck Diesel	BRONX	1306	18
Tractor Truck Diesel	QUEENS	1195	18
Tractor Truck Diesel	MANHATTAN	797	18
Tractor Truck Diesel	STATEN ISLAND	39	18
Dump	BROOKLYN	833	19
Dump	MANHATTAN	609	19
Dump	QUEENS	476	19
Dump	BRONX	328	19
Dump	STATEN ISLAND	50	19
Convertible	QUEENS	696	20
Convertible	BROOKLYN	641	20
Convertible	MANHATTAN	489	20
Convertible	BRONX	309	20
Convertible	STATEN ISLAND	87	20
Ebike__	BROOKLYN	748	21
Ebike__	QUEENS	465	21
Ebike__	MANHATTAN	425	21
Ebike__	BRONX	415	21
Ebike__	STATEN ISLAND	9	21
Scooter__	BROOKLYN	574	22
Scooter__	MANHATTAN	506	22
Scooter__	QUEENS	305	22
Scooter__	BRONX	286	22
Scooter__	STATEN ISLAND	9	22
Garbage or Refuse	BROOKLYN	486	23
Garbage or Refuse	MANHATTAN	474	23
Garbage or Refuse	QUEENS	285	23
Garbage or Refuse	BRONX	229	23
Garbage or Refuse	STATEN ISLAND	27	23
Flat Bed	BROOKLYN	477	24
Flat Bed	MANHATTAN	400	24
Flat Bed	QUEENS	284	24
Flat Bed	BRONX	137	24
Flat Bed	STATEN ISLAND	15	24
Carry All	BROOKLYN	519	25
Carry All	MANHATTAN	287	25
Carry All	BRONX	214	25
Carry All	QUEENS	166	25
Carry All	STATEN ISLAND	29	25
FireTruck__	BROOKLYN	383	26

FireTruck__	MANHATTAN	305	26
FireTruck__	QUEENS	226	26
FireTruck__	BRONX	174	26
FireTruck__	STATEN ISLAND	36	26
Tractor Truck Gasoline	BROOKLYN	370	27
Tractor Truck Gasoline	BRONX	175	27
Tractor Truck Gasoline	MANHATTAN	167	27
Tractor Truck Gasoline	QUEENS	157	27
Tractor Truck Gasoline	STATEN ISLAND	17	27
Tow Truck / Wrecker	BROOKLYN	278	28
Tow Truck / Wrecker	QUEENS	197	28
Tow Truck / Wrecker	MANHATTAN	195	28
Tow Truck / Wrecker	BRONX	134	28
Tow Truck / Wrecker	STATEN ISLAND	9	28
BICYCLE	BROOKLYN	224	29
BICYCLE	MANHATTAN	177	29
BICYCLE	QUEENS	83	29
BICYCLE	BRONX	21	29
Tanker	BROOKLYN	165	30
Tanker	QUEENS	103	30
Tanker	BRONX	97	30
Tanker	MANHATTAN	85	30
Tanker	STATEN ISLAND	7	30
Chassis Cab	BROOKLYN	134	31
Chassis Cab	BRONX	96	31
Chassis Cab	QUEENS	94	31
Chassis Cab	MANHATTAN	86	31
Chassis Cab	STATEN ISLAND	10	31
Refrigerated Van	MANHATTAN	111	32
Refrigerated Van	BROOKLYN	103	32
Refrigerated Van	BRONX	48	32
Refrigerated Van	QUEENS	36	32
Refrigerated Van	STATEN ISLAND	5	32
Concrete Mixer	BROOKLYN	124	33
Concrete Mixer	MANHATTAN	69	33
Concrete Mixer	QUEENS	56	33
Concrete Mixer	BRONX	31	33
Concrete Mixer	STATEN ISLAND	6	33
Armored Truck	MANHATTAN	103	34
Armored Truck	BROOKLYN	64	34
Armored Truck	QUEENS	42	34
Armored Truck	BRONX	22	34
Flat Rack	MANHATTAN	89	35

Flat Rack	BROOKLYN	69	35
Flat Rack	QUEENS	45	35
Flat Rack	BRONX	14	35
Beverage Truck	MANHATTAN	59	36
Beverage Truck	BROOKLYN	52	36
Beverage Truck	BRONX	35	36
Beverage Truck	QUEENS	31	36
LIMO	MANHATTAN	58	37
LIMO	BROOKLYN	55	37
LIMO	QUEENS	33	37
LIMO	BRONX	14	37
Lift Boom	BROOKLYN	44	38
Lift Boom	QUEENS	44	38
Lift Boom	MANHATTAN	28	38
Lift Boom	BRONX	24	38
TRAIL	QUEENS	39	39
TRAIL	BROOKLYN	27	39
TRAIL	MANHATTAN	12	39
TRAIL	BRONX	11	39
Stake or Rack	BROOKLYN	35	40
Stake or Rack	MANHATTAN	29	40
Stake or Rack	QUEENS	18	40
Stake or Rack	BRONX	8	40
Multi-Wheeled Vehicle	QUEENS	22	41
Multi-Wheeled Vehicle	MANHATTAN	20	41
Multi-Wheeled Vehicle	BRONX	19	41
Multi-Wheeled Vehicle	BROOKLYN	10	41
Snow Plow	QUEENS	21	42
Snow Plow	BROOKLYN	20	42
Snow Plow	BRONX	18	42
Snow Plow	MANHATTAN	11	42
Snow Plow	STATEN ISLAND	5	42
DUMP	QUEENS	21	43
DUMP	BROOKLYN	19	43
DUMP	MANHATTAN	14	43
DUMP	BRONX	13	43
DUMP	STATEN ISLAND	6	43
USPS	MANHATTAN	25	44
USPS	QUEENS	10	44
USPS	BROOKLYN	9	44
TRACT	BROOKLYN	20	45
TRACT	MANHATTAN	13	45
TRACT	QUEENS	10	45

TRACT	BRONX	6	45
UNKNO	BROOKLYN	14	46
UNKNO	MANHATTAN	11	46
UNKNO	STATEN ISLAND	11	46
UNKNO	QUEENS	9	46
Bulk Agriculture	BROOKLYN	21	47
Bulk Agriculture	MANHATTAN	9	47
Bulk Agriculture	QUEENS	9	47
Bulk Agriculture	BRONX	5	47
COM	MANHATTAN	23	48
COM	BRONX	9	48
COM	BROOKLYN	6	48
Pickup with mounted Camper	BROOKLYN	14	49
Pickup with mounted Camper	MANHATTAN	13	49
Pickup with mounted Camper	BRONX	9	49
Pedicab	MANHATTAN	30	50
Pedicab	BROOKLYN	7	50
TRAILER	QUEENS	13	51
TRAILER	BROOKLYN	12	51
TRAILER	BRONX	7	51
TRAILER	MANHATTAN	6	51
DELIV	MANHATTAN	15	52
DELIV	BROOKLYN	10	52
DELIV	BRONX	7	52
DELIV	QUEENS	5	52
trail	QUEENS	12	53
trail	BROOKLYN	11	53
trail	MANHATTAN	7	53
UNK	BROOKLYN	12	54
UNK	QUEENS	9	54
UNK	MANHATTAN	8	54
UNK	BRONX	6	54
COMME	MANHATTAN	19	55
COMME	BROOKLYN	8	55
DELV	MANHATTAN	12	56
DELV	QUEENS	10	56
DELV	BROOKLYN	5	56
TOW T	BROOKLYN	15	57
TOW T	QUEENS	8	57
Lunch Wagon	MANHATTAN	12	58
Lunch Wagon	QUEENS	8	58
Lunch Wagon	BRONX	5	58
Minibike	BROOKLYN	10	59

Minibike	BRONX	8	59
Minibike	MANHATTAN	6	59
PAS	STATEN ISLAND	12	60
PAS	BROOKLYN	5	60
UTILI	BROOKLYN	13	61
UTILI	QUEENS	7	61
UTILI	MANHATTAN	5	61
ELECT	BRONX	10	62
ELECT	BROOKLYN	9	62
FORKL	BROOKLYN	11	63
FORKL	QUEENS	8	63
FORKL	BRONX	5	63
Pallet	BROOKLYN	10	64
Pallet	MANHATTAN	6	64
Pallet	BRONX	5	64
Pallet	QUEENS	5	64
Trail	QUEENS	11	65
Trail	BROOKLYN	5	65
Trail	MANHATTAN	5	65
UTIL	MANHATTAN	10	66
UTIL	BROOKLYN	8	66
Tow Truck	BRONX	8	67
Tow Truck	BROOKLYN	8	67
BOX T	BROOKLYN	7	68
BOX T	QUEENS	6	68
BOX T	MANHATTAN	5	68
PEDICAB	MANHATTAN	18	69
Motorized Home	MANHATTAN	8	70
Motorized Home	QUEENS	6	70
Van Camper	BROOKLYN	10	71
Van Camper	QUEENS	5	71
Minicycle	BROOKLYN	9	72
Minicycle	QUEENS	5	72
TANK	BROOKLYN	7	73
TANK	QUEENS	6	73
dump	MANHATTAN	8	74
dump	QUEENS	6	74
COMMERCIAL	BROOKLYN	7	75
Trailer	BROOKLYN	5	76
Trailer	MANHATTAN	5	76
unkno	BROOKLYN	5	77
unkno	MANHATTAN	5	77
Firetruck	BROOKLYN	5	78

STREE	BROOKLYN	10	79
FLAT	BROOKLYN	6	80
FORD	BROOKLYN	5	81
Hopper	QUEENS	7	82
SELF	BROOKLYN	6	83
TOW TRUCK	BROOKLYN	5	84
BOX TRUCK	BRONX	5	85
BOX TRUCK	BROOKLYN	5	85
FREIG	BROOKLYN	8	86
Fire truck	QUEENS	6	87
Fire truck	BROOKLYN	5	87
SCHOOL BUS	BRONX	6	88
TRK	MANHATTAN	11	89
ambu	MANHATTAN	5	90
tract	QUEENS	5	91
unk	BROOKLYN	5	92
usps	QUEENS	6	93
MACK	BROOKLYN	6	94
OMR	BROOKLYN	6	96
Scooter	MANHATTAN	5	97
fire truck	BROOKLYN	6	101
scoot	BROOKLYN	5	102
REFG	MANHATTAN	8	111
Vanette	BROOKLYN	5	113
stree	BRONX	5	115
BACKH	MANHATTAN	5	117
PSD	BRONX	5	119
subn	BROOKLYN	5	125
Forklift	BROOKLYN	5	143
MAIL TRUCK	QUEENS	5	144
STAK	MANHATTAN	6	147
Other	QUEENS	512	314
Other	MANHATTAN	614	446
Other	BROOKLYN	659	447
Other	BRONX	450	455
Other	STATEN ISLAND	143	471
