Data cleaning lab report

Riley Payung

2019-10-09

Exercises

Exercise 1

```
towed_renamed <- towed %>%
  rename(
    `tow_date` = `Tow Date`,
    towed_to_address = `Towed to Address`,
    tow_facility_phone = `Tow Facility Phone`,
    inventory_number = `Inventory Number`,
    make = Make,
    style = Style,
    model = Model,
    color = Color,
    plate = Plate,
    state = State
  )
```

Exercise 2

```
towed_renamed_2 <- towed_renamed %>%
  mutate(
    tow_date = mdy(tow_date)
  )
```

Exercise 3

```
tow_make_count <- towed_renamed_2 %>% count(make)
head(tow_make_count)
```

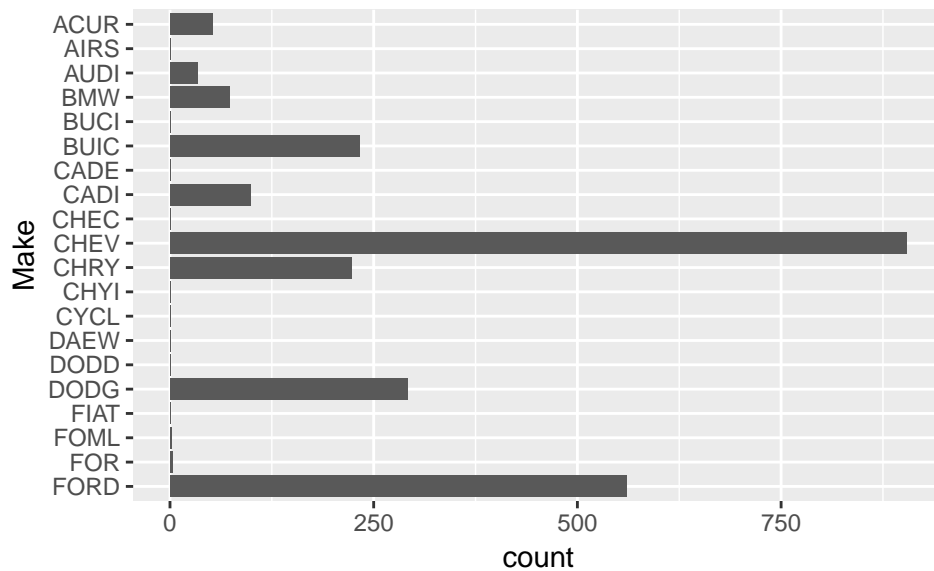| make	n
ACUR	52
AIRS	1
AUDI	34
BMW	73
BUCI	1
BUIC	232

There are 78 unique values. There are many rows with only 1 count.

Exercise 4

```
tow_make_count_1 <- tow_make_count %>% slice(1:20)
tow_make_count_2 <- tow_make_count %>% slice(21:40)
tow_make_count_3 <- tow_make_count %>% slice(41:60)
tow_make_count_4 <- tow_make_count %>% slice(61:80)
```
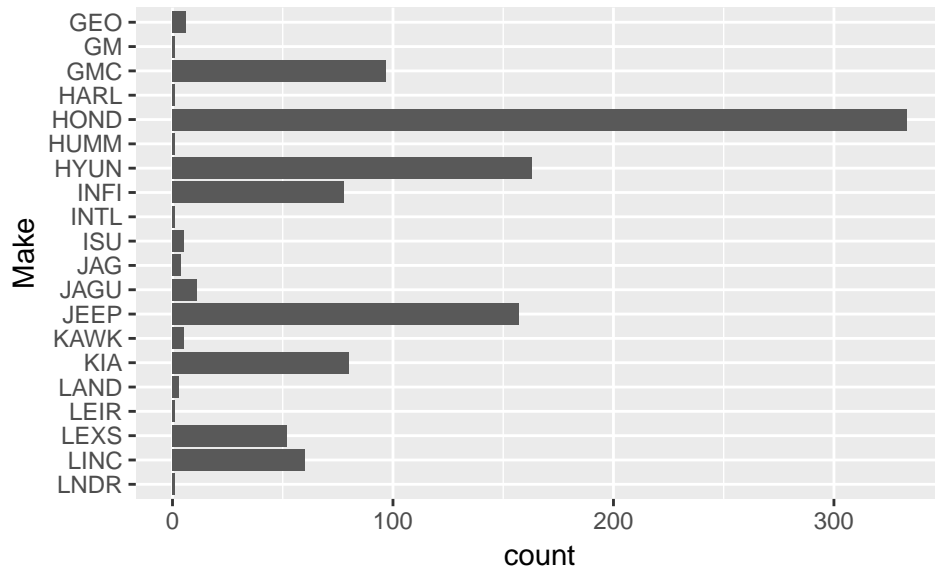
Exercise 5

```
ggplot(data = tow_make_count_1) +
geom_col(mapping = aes(x = fct_rev(make), y = n)) +
coord_flip() +
xlab("Make") +
ylab("count")
```



```
ggplot(data = tow_make_count_2) +
geom_col(mapping = aes(x = fct_rev(make), y = n)) +
coord_flip() +
xlab("Make") +
ylab("count")
```
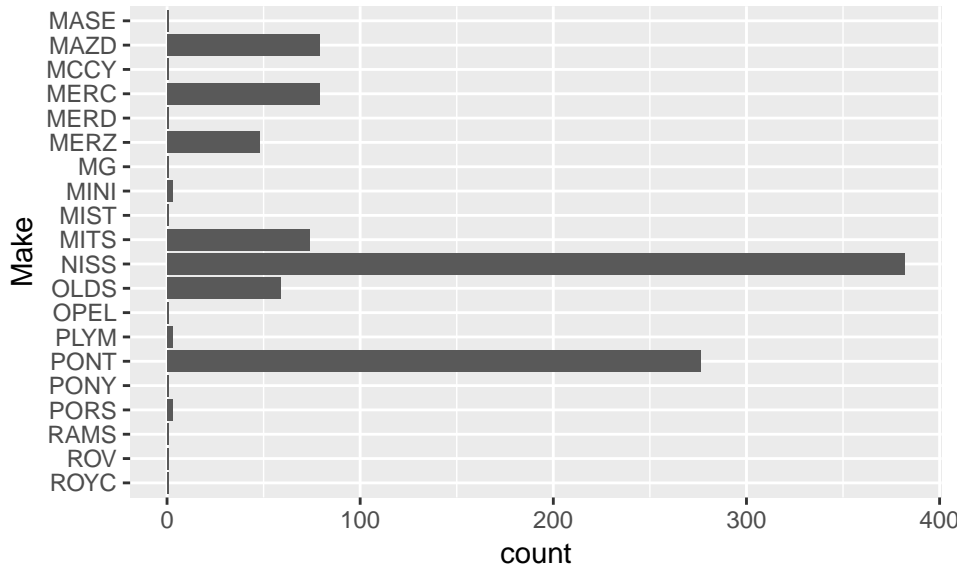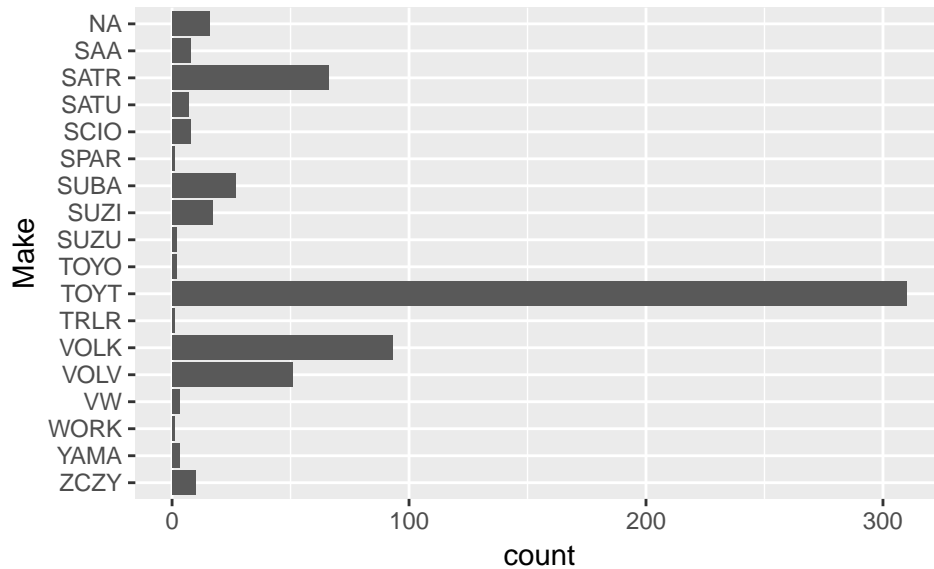
```r
ggplot(data = tow_make_count_3) +
geom_col(mapping = aes(x = fct_rev(make), y = n)) +
coord_flip() +
xlab("Make") +
ylab("count")
```



```r
ggplot(data = tow_make_count_4) +
geom_col(mapping = aes(x = fct_rev(make), y = n)) +
coord_flip() +
xlab("Make") +
ylab("count")
```

Exercise 6

```
towed_partial_clean <- towed_renamed_2 %>%
  mutate(
    make = recode(make, DODD = "DODG"),
    make = recode(make, BUCI = "BUIC")
  )
  #Proof that it was completed:
  filter(towed_partial_clean,make=='DODD' | make=='BUCI')
```

tow_date make style model color plate state towed_to_address tow_facility_phone inventory_number ———— —— —— —— —— —— —— —————— —————- ————

Exercise 7

We should be using JAGU

```
towed_partial_clean <- towed_partial_clean %>%
  mutate(
    make = recode(make, JAG = "JAGU")
  )
  filter(towed_partial_clean,make=='JAG' | make=='JAGU')
```

tow_date	make	style	model	color	plate	state	towed_to_address	tow_facility_phone	inven
2018-03-08	JAGU	4D	NA	SIL	246T709	IL	701 N. Sacramento	(773) 265-7605	
2018-03-09	JAGU	4D	NA	SIL	AP47282	IL	10300 S. Doty	(773) 568-8495	
2018-05-02	JAGU	4T	NA	SIL	AG20729	IL	701 N. Sacramento	(773) 265-7605	
2018-03-04	JAGU	4D	NA	BLK	E102757	IL	701 N. Sacramento	(773) 265-1846	
2018-05-02	JAGU	4D	NA	BLK	AN75625	IL	701 N. Sacramento	(773) 265-7605	
2018-03-07	JAGU	4D	NA	BLK	NA	NA	10300 S. Doty	(773) 568-8495	
2018-03-20	JAGU	4D	NA	BLK	AL13808	IL	701 N. Sacramento	(773) 265-7605	
2018-04-22	JAGU	4D	NA	GRN	AJ90522	IL	701 N. Sacramento	(773) 265-7605	

tow_date	make	style	model	color	plate	state	towed_to_address	tow_facility_phone	inven
2018-04-22	JAGU	4D	NA	LBL	845293	IL	701 N. Sacramento	(773) 265-7605	
2018-03-01	JAGU	4D	NA	GRN	ZV89485	IL	701 N. Sacramento	(773) 265-7605	
2018-05-27	JAGU	2D	NA	SIL	AP44162	IL	10300 S. Doty	(773) 568-8495	
2018-04-26	JAGU	4T	NA	WHI	7X1K058	CA	10300 S. Doty	(773) 568-8495	
2018-05-28	JAGU	4D	NA	BLK	ZX23445	IL	10300 S. Doty	(773) 568-8495	
2018-03-04	JAGU	4T	NA	BLK	KDT7481	TX	701 N. Sacramento	(773) 265-7605	
2018-03-28	JAGU	4D	NA	BLK	NA	IL	701 N. Sacramento	(773) 265-7605	

Exercise 8

```r
towed_clean <- towed_partial_clean %>%
  mutate(
    make = recode(make, CADE = "CADI"),
    make = recode(make, CHYI = "CHRY"),
    make = recode(make, FOR = "FORD"),
    make = recode(make, GM = "GMC"),
    make = recode(make, HARL = "HR"),
    make = recode(make, LAND = "LNDR"),
    make = recode(make, LEIR = "LEIS"),
    make = recode(make, MCCY = "MC"),
    make = recode(make, MERD = "MERZ"),
    make = recode(make, MINI = "MNNI"),
    make = recode(make, MIST = "MITS"),
    make = recode(make, PONY = "PONT"),
    make = recode(make, RAMS = "RAM"),
    make = recode(make, ROYC = "ROL"),
    make = recode(make, NASH = "NA"),
    make = recode(make, SATR = "STRN"),
    make = recode(make, SATU = "STRN"),
    make = recode(make, SCIO = "TOYT"),
    make = recode(make, SPAR = "SPRT"),
    make = recode(make, SUZU = "ISU"),
    make = recode(make, TOYO = "TOYT"),
    make = recode(make, TRLR = "TVLC"),
    make = recode(make, VW = "VOLK"),
    make = recode(make, WORK = "KW")

  )
tow_make_count_clean <- towed_clean %>% count(make)
```

Additional questions

Question 1

Chevrolet (CHEV) was the most impounded vehicle make at 904 total counts.

Question 2

```
tow_by_day <- towed_clean %>% count(tow_date)
```

The highest number of vehicles impounded in one day, is on March 4, 2018 with 166 total impounds.