

Crime Data EDA

Aya Klos

2023-01-04

Notes on the dataset: - each row has data for a given crime - NIBRS stands for “National Incident-Based-Reporting System, which is an incident-based reporting system where law enforcement collects data on each crime occurrence - NIBRS codes correspond to a given offense (offenses with the same code number followed by a letter are all categorized under the same broad offense category) - there are group “A” and group “B” offenses, the latter of which must have an arrest before it is NIBRS reportable - this link has descriptions of each of the variables in the data set (<https://www.minneapolismn.gov/media/-www-content-assets/documents/CrimeDashboardInfo.pdf>)

Strange notes about the dataset: - the X column seems to be the same as the Longitude column and Y the same as Latitude

```
data <- read.csv("/Users/ayaklos/Documents/GitHub/carleton_comps_22_23/data/original/Crime_Data.csv")

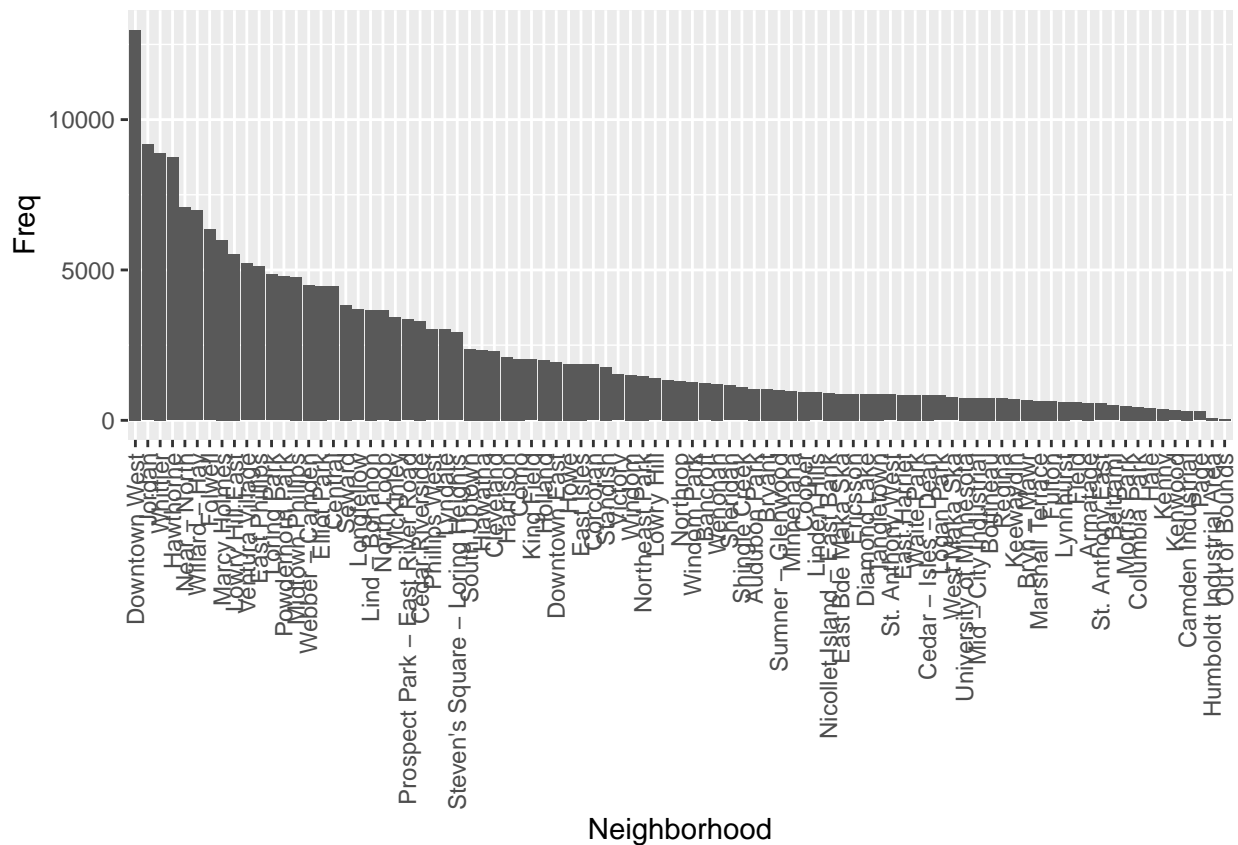
data <- data %>% arrange(NIBRS_Crime_Against)

#41269 crimes were not part of the NIBRS data
sum(data$NIBRS_Crime_Against == "Non NIBRS Data")

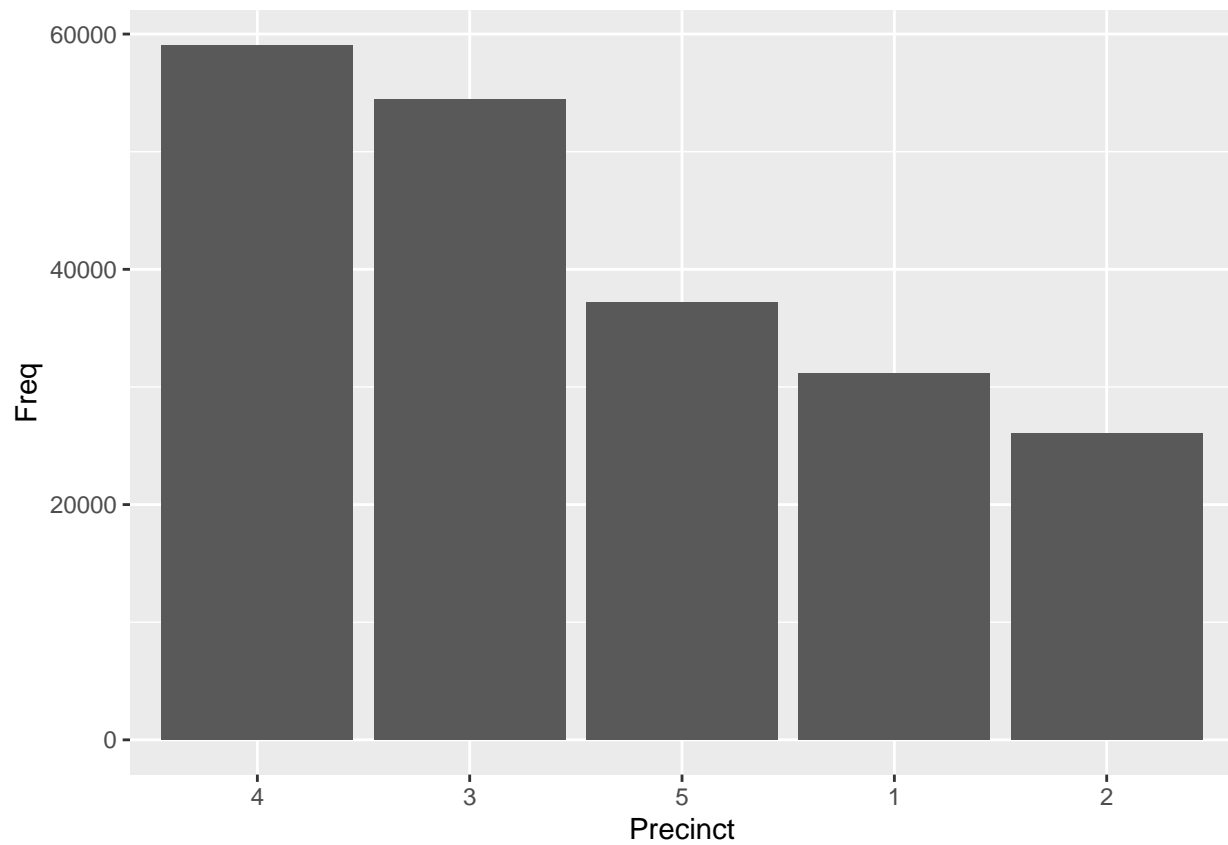
## [1] 41269

neighborhood.df <- table(data$Neighborhood) %>% as.data.frame()
precinct.df <- table(data$Precinct) %>% as.data.frame()
ward.df <- table(data$Ward) %>% as.data.frame()

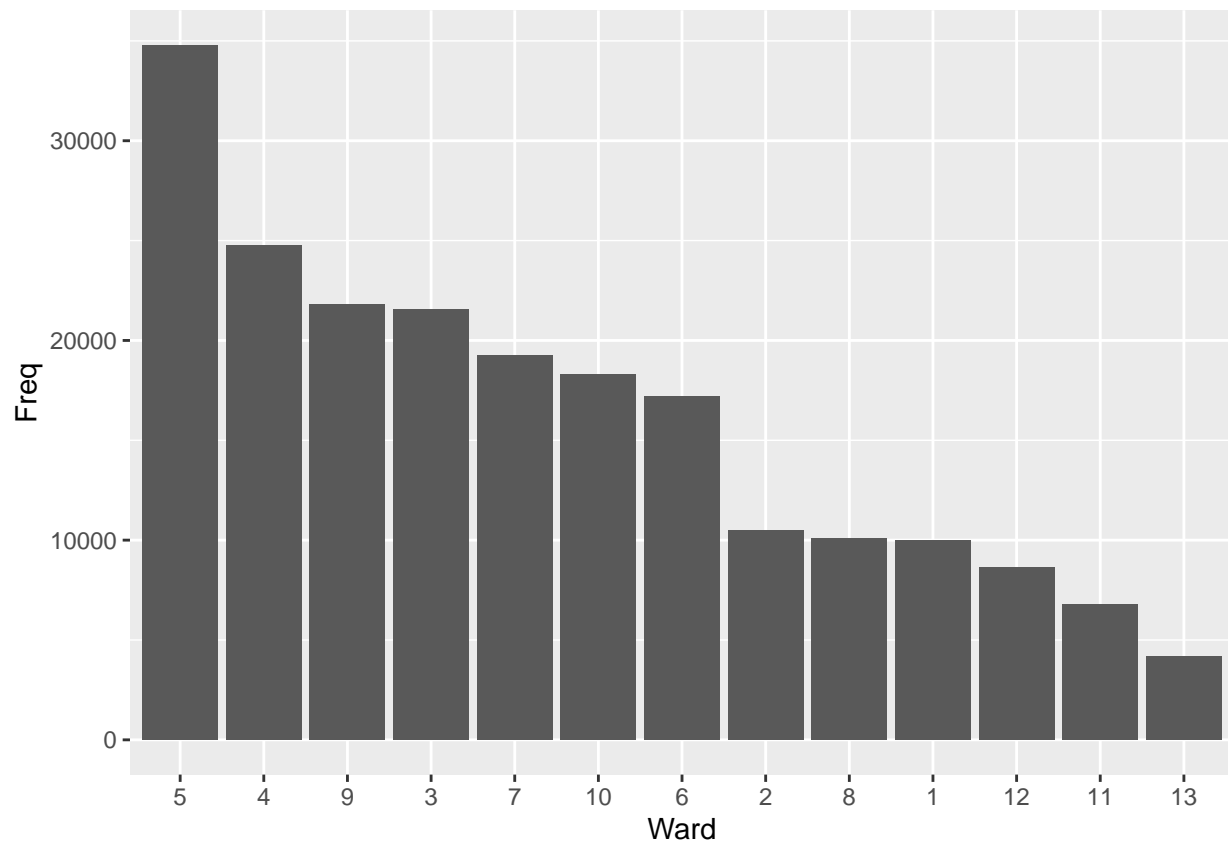
ggplot(neighborhood.df, aes(x = reorder(Var1,-Freq), y = Freq)) +
  geom_bar(stat = "identity") +
  labs(x = "Neighborhood") +
  theme(axis.text.x = element_text(angle = 90, vjust = 0.5, hjust=1))
```



```
ggplot(precinct.df, aes(x = reorder(Var1,-Freq), y = Freq)) +
  geom_bar(stat = "identity") +
  labs(x = "Precinct")
```



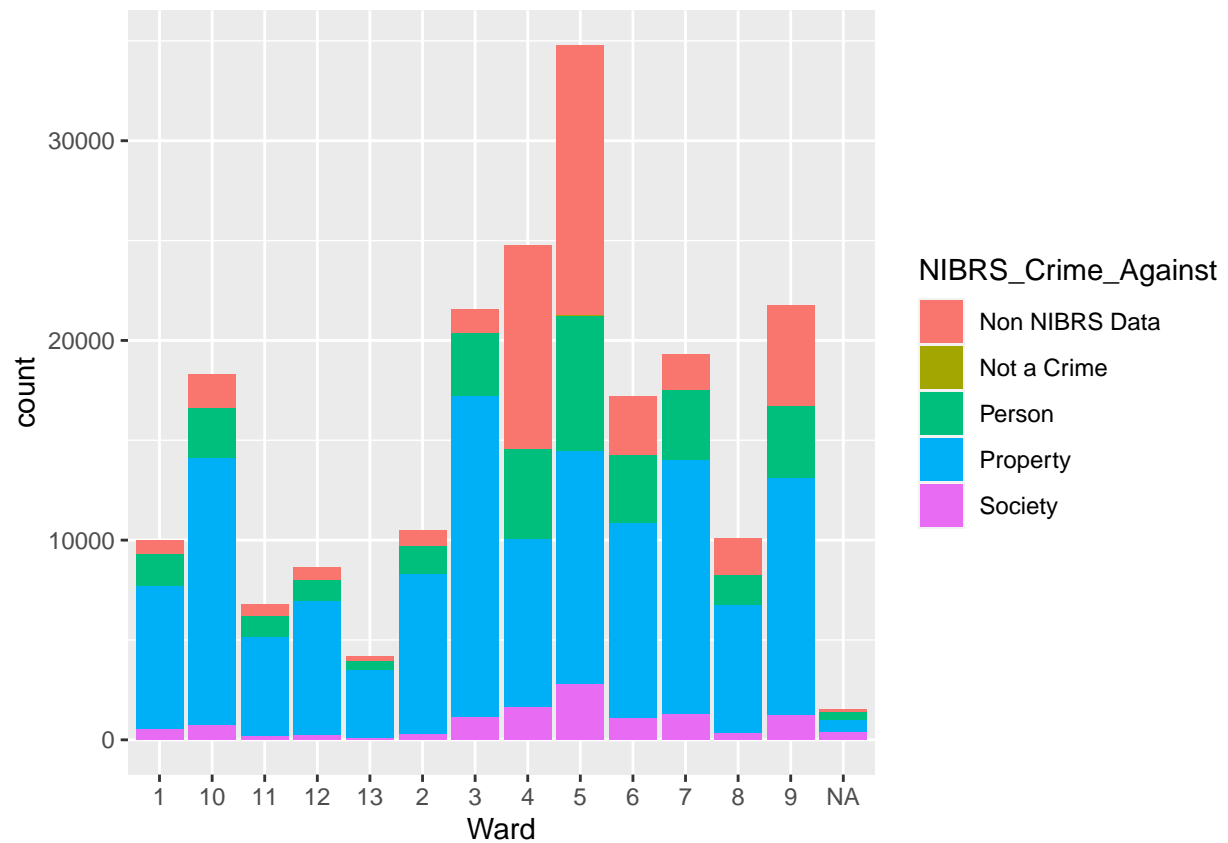
```
ggplot(ward.df, aes(x = reorder(Var1,-Freq), y = Freq)) +  
  geom_bar(stat = "identity") +  
  labs(x = "Ward")
```



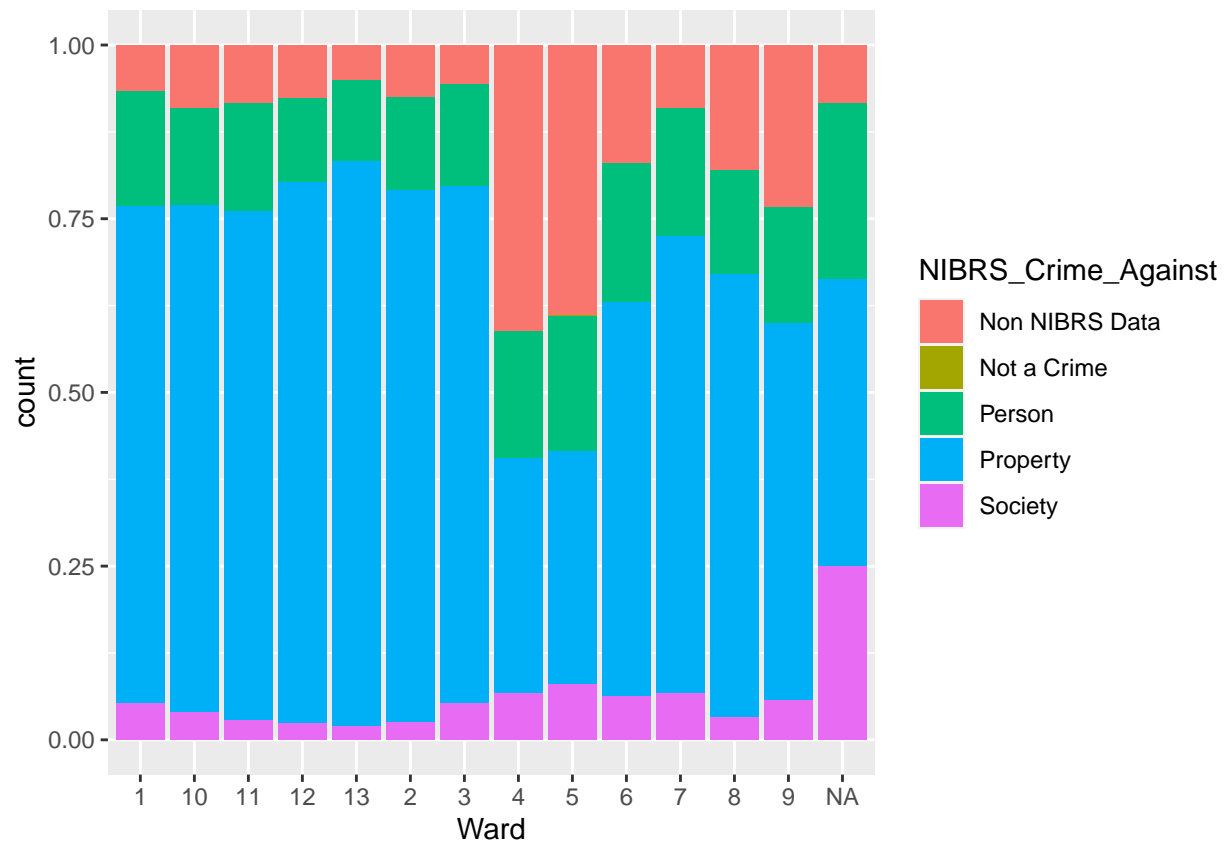
```
data.1 <- data
data.1$Ward <- as.character(data.1$Ward)
ggplot(data.1, aes(x = Ward, fill = Offense)) +
  geom_bar()
```

ated Assault	Gambling Equipment Violations	Purse–snatching
or Larceny	Gunshot Wound Victims	Robbery
Cruelty	Hacking/Computer Invasion	Sex Offenses
	Human Trafficking, Commercial Sex Acts	Shooting (PFE)
g or Promoting Prostitution	Human Trafficking, Involuntary Servitude	Shooting Report On
Wagering	Identity Theft	Shoplifting
	Impersonation	ShotSpotter Activati
//Breaking & Entering	Intimidation	Simple Assault
ing – Subset of Robbery	Justifiable Homicide	Sound of Shots Fire
feiting/Forgery	Kidnapping/Abduction	Stolen Property Offe
ard/Automated Teller Machine Fraud	Motor Vehicle Theft	Theft From Building
tion/Damage/Vandalism of Property	Murder and Nonnegligent Manslaughter	Theft From Coin–Op
ic Aggravated Assault – Subset of Assault	Negligent Manslaughter	Theft From Motor Ve
quipment Violations	Operating/Promoting/Assisting Gambling	Theft of Motor Vehic
arcotic Violations	Pocket–picking	Weapon Law Violati
lement	Pornography/Obscene Material	Welfare Fraud
n/Blackmail	Prostitution	Wire Fraud
retenses/Swindle/Confidence Game	Purchasing Prostitution	

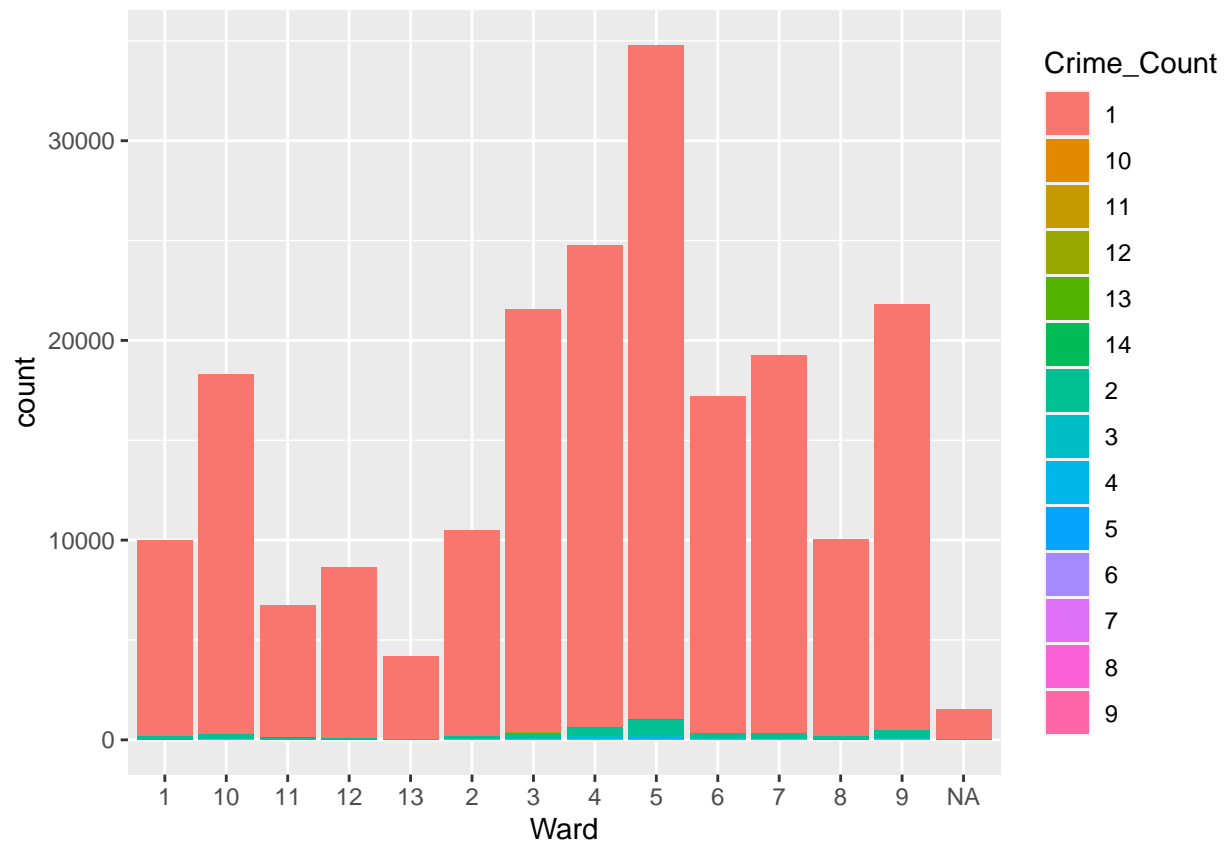
```
ggplot(data.1, aes(x = Ward, fill = NIBRS_Crime_Against)) +
  geom_bar()
```



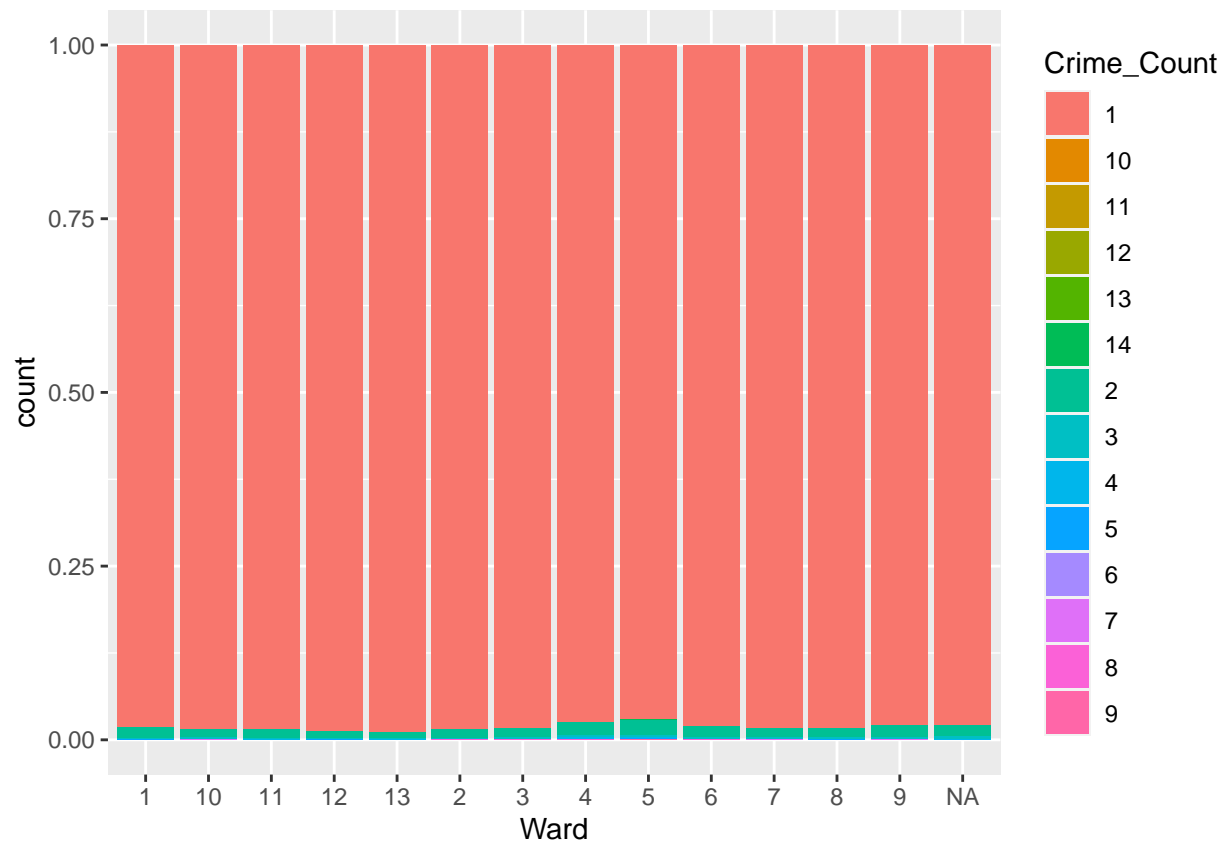
```
ggplot(data.1, aes(x = Ward, fill = NIBRS_Crime_Against)) +  
  geom_bar(position = "fill")
```



```
data.1$Crime_Count <- as.character(data.1$Crime_Count)
ggplot(data.1, aes(x = Ward, fill = Crime_Count)) +
  geom_bar()
```



```
ggplot(data.1, aes(x = Ward, fill = Crime_Count)) +  
  geom_bar(position = "fill")
```

```
table(data$Crime_Count)
```

```
##
##      1      2      3      4      5      6      7      8      9     10     11
## 205188 3209  564  203   89   23   28   11   3    5    3
##      12     13     14
##       2      1      2
```