library("tidyverse")

## ── Attaching packages ─────────────────────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ ggplot2 3.2.1 ✓ purrr 0.3.3  
## ✓ tibble 2.1.3 ✓ dplyr 0.8.4  
## ✓ tidyr 1.0.2 ✓ stringr 1.4.0  
## ✓ readr 1.3.1 ✓ forcats 0.4.0

## ── Conflicts ────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

library("GGally")

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

##   
## Attaching package: 'GGally'

## The following object is masked from 'package:dplyr':  
##   
## nasa

library("MASS")

##   
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':  
##   
## select

library("caret")

## Loading required package: lattice

##   
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':  
##   
## lift

library("lubridate")

##   
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':  
##   
## date

library("dplyr")

chicago = read\_csv("chicago.csv")

## Parsed with column specification:  
## cols(  
## .default = col\_character(),  
## ID = col\_double(),  
## Arrest = col\_logical(),  
## Domestic = col\_logical(),  
## Beat = col\_double(),  
## District = col\_double(),  
## Ward = col\_double(),  
## `Community Area` = col\_double(),  
## `X Coordinate` = col\_double(),  
## `Y Coordinate` = col\_double(),  
## Year = col\_double(),  
## Latitude = col\_double(),  
## Longitude = col\_double()  
## )

## See spec(...) for full column specifications.

chicago = chicago %>%  
 mutate(Date = mdy\_hms(Date))

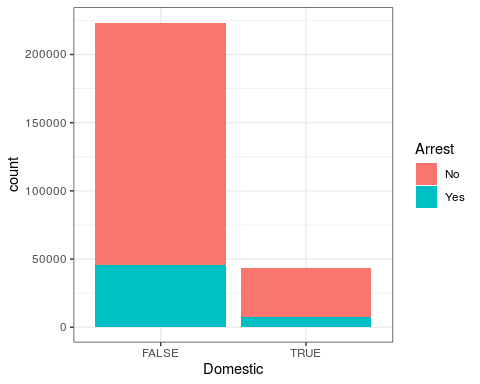
chicago = chicago %>%   
 mutate(Hour = minute(Date))  
summary(chicago$Hour)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00 9.00 14.00 13.08 18.00 23.00

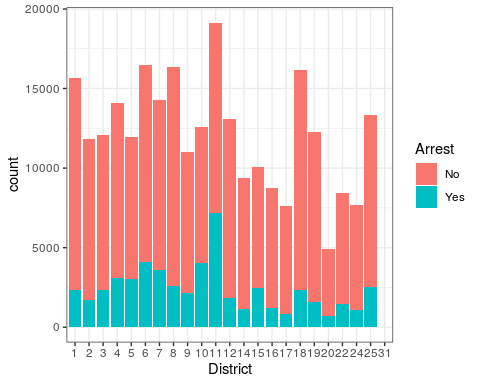
chicago = chicago %>%  
 dplyr::select("Date", "Hour", "Block", "IUCR", 'Primary Type', "Description", "Arrest", "Domestic", "Beat", "District", "Ward", 'Community Area', 'FBI Code','Location Description', "Year", "Latitude", "Longitude")

chicago = chicago %>%  
 mutate(Arrest = as.factor(Arrest)) %>%  
 mutate(Arrest = fct\_recode(Arrest, "No" = "FALSE", "Yes" = "TRUE")) %>%  
 mutate(Hour = as.factor(Hour)) %>%  
 mutate('Primary\_Type' = as.factor(as.character('Primary\_Type')))%>%  
 mutate(Description = as.factor(Description)) %>%  
 mutate('Location\_Description' = as.factor(as.character('Location\_Description'))) %>%  
 mutate(Domestic = as.factor(Domestic)) %>%  
 mutate(District = as.factor(District)) %>%  
 mutate(Year = as.factor(Year)) %>%  
 mutate(Ward = as.factor(Ward)) %>%  
 mutate('Community\_Area' = as.factor(as.character('Community\_Area'))) %>%  
 mutate(Block = as.factor(Block)) %>%  
 mutate(IUCR = as.factor(IUCR))

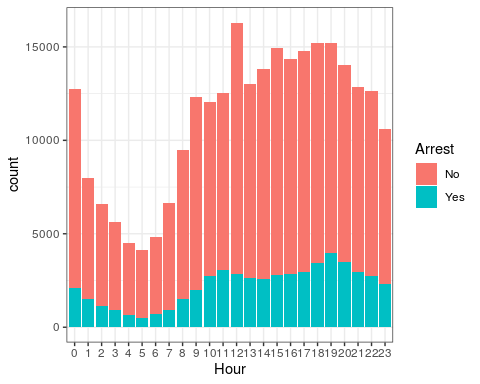
p1 = ggplot(chicago, aes(x = Domestic, fill = Arrest))+geom\_bar()+theme\_bw()  
show(p1)



ggplot(chicago, aes(x = District, fill = Arrest)) + geom\_bar() + theme\_bw()



ggplot(chicago, aes(x = Hour, fill = Arrest)) + geom\_bar()+theme\_bw()



ggplot(chicago, aes(x = `Primary Type`, fill = Arrest)) + geom\_bar()+theme\_bw()+theme(axis.text.x = element\_text(angle = 90))

