Firewatch

Tittiwat Tonburinthip

9/3/2021

Retrieve the data from SQLLite

```
con <- dbConnect(drv=RSQLite::SQLite(), dbname="FPA_FOD_20210617.sqlite")

## list all tables
tables <- dbListTables(con)

## exclude sqlite_sequence (contains table information)
tables <- tables[tables != "sqlite_sequence"]

lDataFrames <- vector("list", length=length(tables))

## create a data.frame for each table
for (i in seq(along=tables)) {
  lDataFrames[[i]] <- dbGetQuery(conn=con, statement=paste("SELECT * FROM '", tables[[i]], "'", sep=""))
}
```

```
## Warning in result_fetch(res@ptr, n = n): Column ‘SOURCE_REPORTING_UNIT‘: mixed
## type, first seen values of type integer, coercing other values of type string

## Warning in result_fetch(res@ptr, n = n): Column ‘LOCAL_FIRE_REPORT_ID‘: mixed
## type, first seen values of type integer, coercing other values of type string
```

Store dataframe in fire and place

```
## Rows: 52 Columns: 4

## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): STATE, name
## dbl (2): latitude, longitude

##
## i Use ‘spec()‘ to retrieve the full column specification for this data.
## i Specify the column types or set ‘show_col_types = FALSE‘ to quiet this message.
```

CONT_DATE and CONT_DOY have a lot of NA, but it's the same value as DISCOVERY_DATE, DISCOVERY_DOY

a= store the rows that have NA. It's only have one row that NWCG_CAUSE_CLASSIFICATION== NA, so delete it and store in fire_clean

place_1 has a mutate column named NWCG_REPORTING_UNIT_ID. Then, LEFT JOIN with fire_clean_1

#Summary

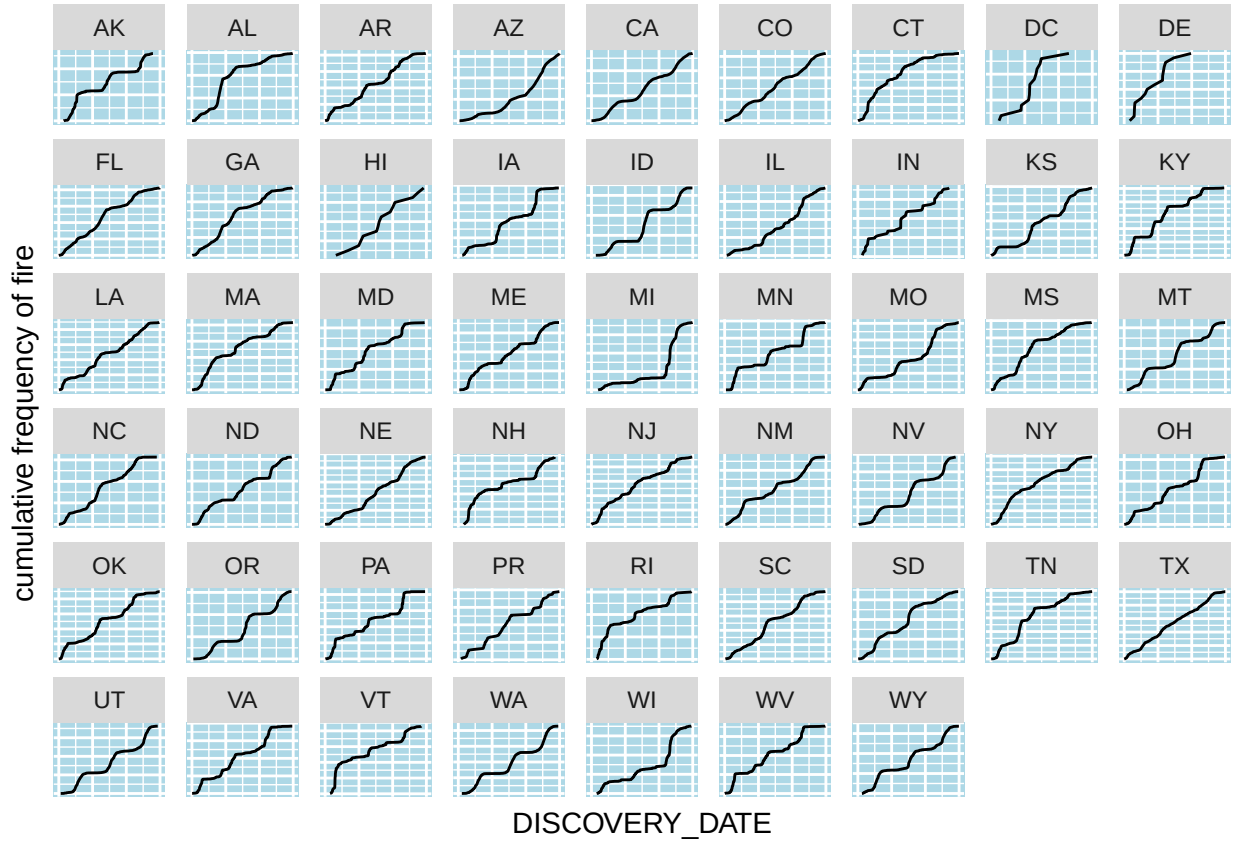
```
##  NWCG_REPORTING_UNIT_ID  FIRE_YEAR         DISCOVERY_DATE       DISCOVERY_DOY
##  Length:2166752         Length:2166752     Min.   :1992-01-01   Min.   :  1
##  Class :character       Class :character   1st Qu.:1999-08-22   1st Qu.: 89
##  Mode  :character       Mode  :character   Median :2006-03-04   Median :165
##                                            Mean   :2005-10-09   Mean   :165
##                                            3rd Qu.:2011-10-20   3rd Qu.:230
##                                            Max.   :2018-12-31   Max.   :366
##  NWCG_CAUSE_CLASSIFICATION NWCG_GENERAL_CAUSE   FIRE_SIZE
##  Length:2166752            Length:2166752     Min.   :     0.0
##  Class :character          Class :character   1st Qu.:     0.1
##  Mode  :character          Mode  :character   Median :     1.0
##                                               Mean   :    76.0
##                                               3rd Qu.:     3.0
##                                               Max.   :662700.0
##  FIRE_SIZE_CLASS       LATITUDE       LONGITUDE          STATE
##  Length:2166752     Min.   :17.94   Min.   :-178.80   Length:2166752
##  Class :character   1st Qu.:32.96   1st Qu.:-110.85   Class :character
##  Mode  :character   Median :35.64   Median : -93.11   Mode  :character
##                     Mean   :36.89   Mean   : -96.19
##                     3rd Qu.:40.81   3rd Qu.: -82.46
##                     Max.   :70.33   Max.   : -65.26
##     UnitId          GeographicArea         Gacc           WildlandRole
##  Length:2166752     Length:2166752     Length:2166752     Length:2166752
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    UnitType          Department           Agency             Parent
##  Length:2166752     Length:2166752     Length:2166752     Length:2166752
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
##    Country             State              Code               Name
##  Length:2166752     Length:2166752     Length:2166752     Length:2166752
##  Class :character   Class :character   Class :character   Class :character
##  Mode  :character   Mode  :character   Mode  :character   Mode  :character
##
##
##
```

Prepare data for sparkline

```
## `summarise()` has grouped output by 'STATE'. You can override using the `.groups` argument.
```

plot- Sparkline_cumulative fire in 5 years [start:01/01/2016]

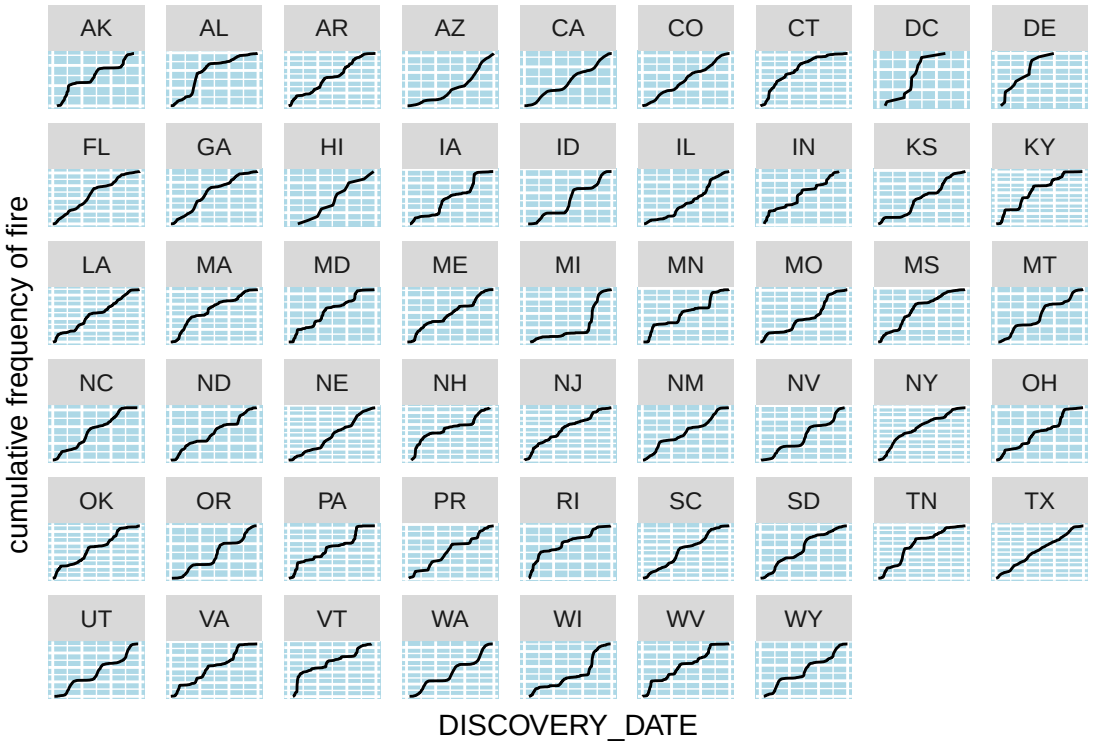
```
g <- fire_clean_2 %>%
  ggplot(aes(x = DISCOVERY_DATE, y = cum))+
  facet_wrap(~ STATE, ncol = 9, scales ="free_y") +
  geom_line()+ theme(axis.ticks.y = element_blank(), axis.text.y = element_blank())+
  theme(axis.ticks.x = element_blank(), axis.text.x = element_blank())+theme(panel.background = element_
                                colour = "lightblue",
                                size = 0.5, linetype = "solid"),
  panel.grid.major = element_line(size = 0.5, linetype = 'solid',
                                colour = "white"),
  panel.grid.minor = element_line(size = 0.25, linetype = 'solid',colour = "white"))+ylab('cumulative fi
g
```



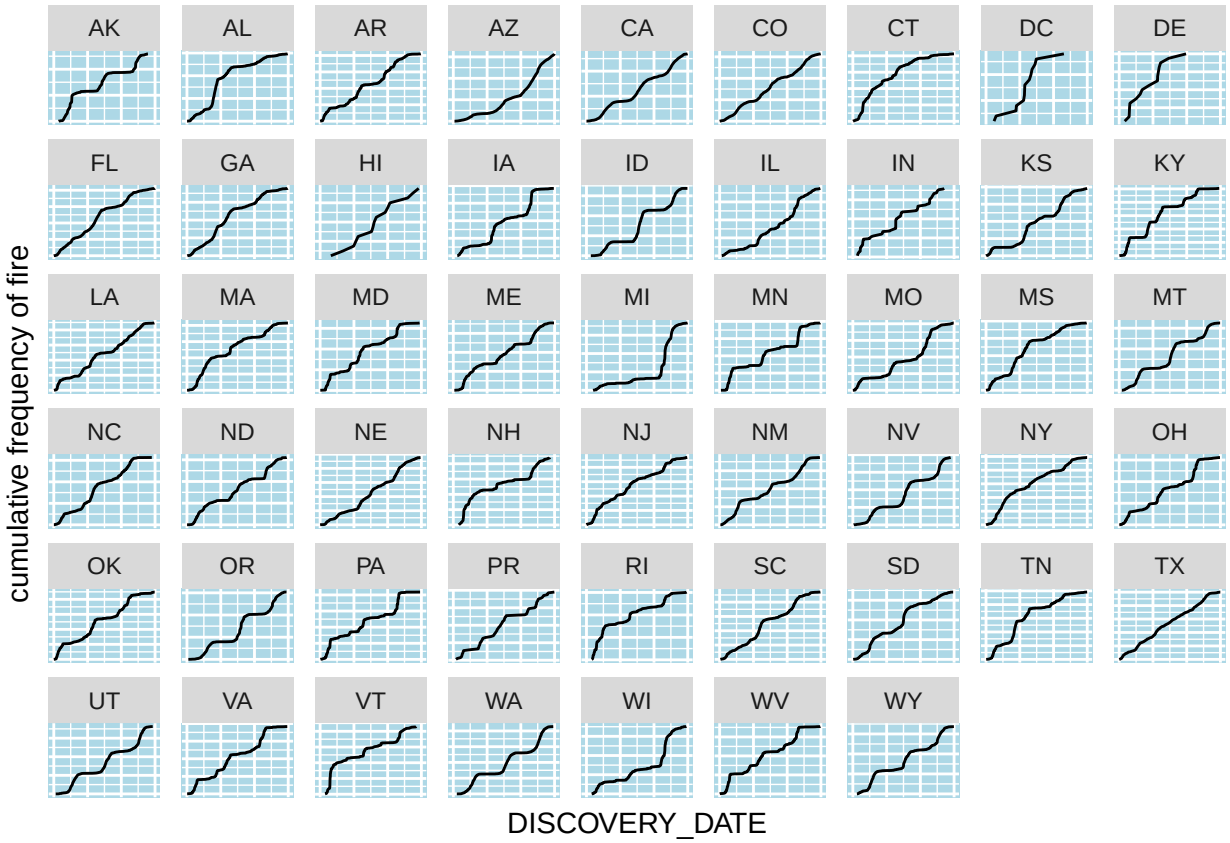

```
library(ggh4x)
```

```
## Warning: package 'ggh4x' was built under R version 4.1.1
```

```
g + force_panelsizes(rows = unit(0.3, "in"),
                     cols = unit(0.5, "in"))
```


AK
AL
AR
AZ
CA
CO
CT
DC
DE
FL
GA
HI
IA
ID
IL
IN
KS
KY
LA
MA
MD
ME
MI
MN
MO
MS
MT
NC
ND
NE
NH
NJ
NM
NV
NY
OH
OK
OR
PA
PR
RI
SC
SD
TN
TX
UT
VA
VT
WA
WI
WV
WY
cumulative frequency of fire
DISCOVERY_DATE

```
g
```



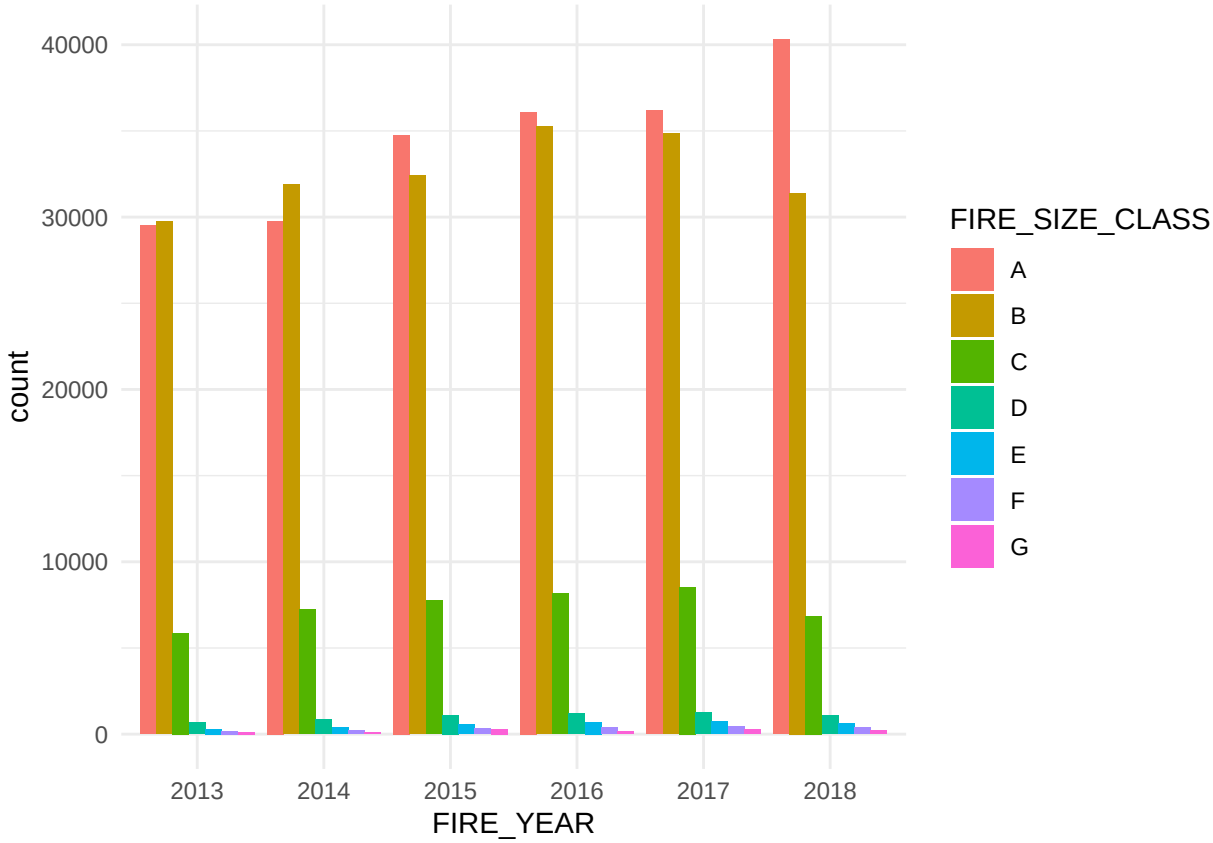

```
fire_clean_3<-fire_clean_1%>%
  filter(Country=="US", DISCOVERY_DATE>='2013-01-01')%>%
  group_by(FIRE_YEAR, FIRE_SIZE_CLASS)%>%
  summarise(count=n())%>%
  mutate(cum=cumsum(count))%>%
  arrange(FIRE_YEAR)
```

```
## `summarise()` has grouped output by 'FIRE_YEAR'. You can override using the `.groups` argument.
```

```
#View(fire_clean_1)
```

stack bar

```
ggplot(fire_clean_3, aes(fill=FIRE_SIZE_CLASS, y=count, x=FIRE_YEAR))+geom_bar(position="dodge", stat="
```



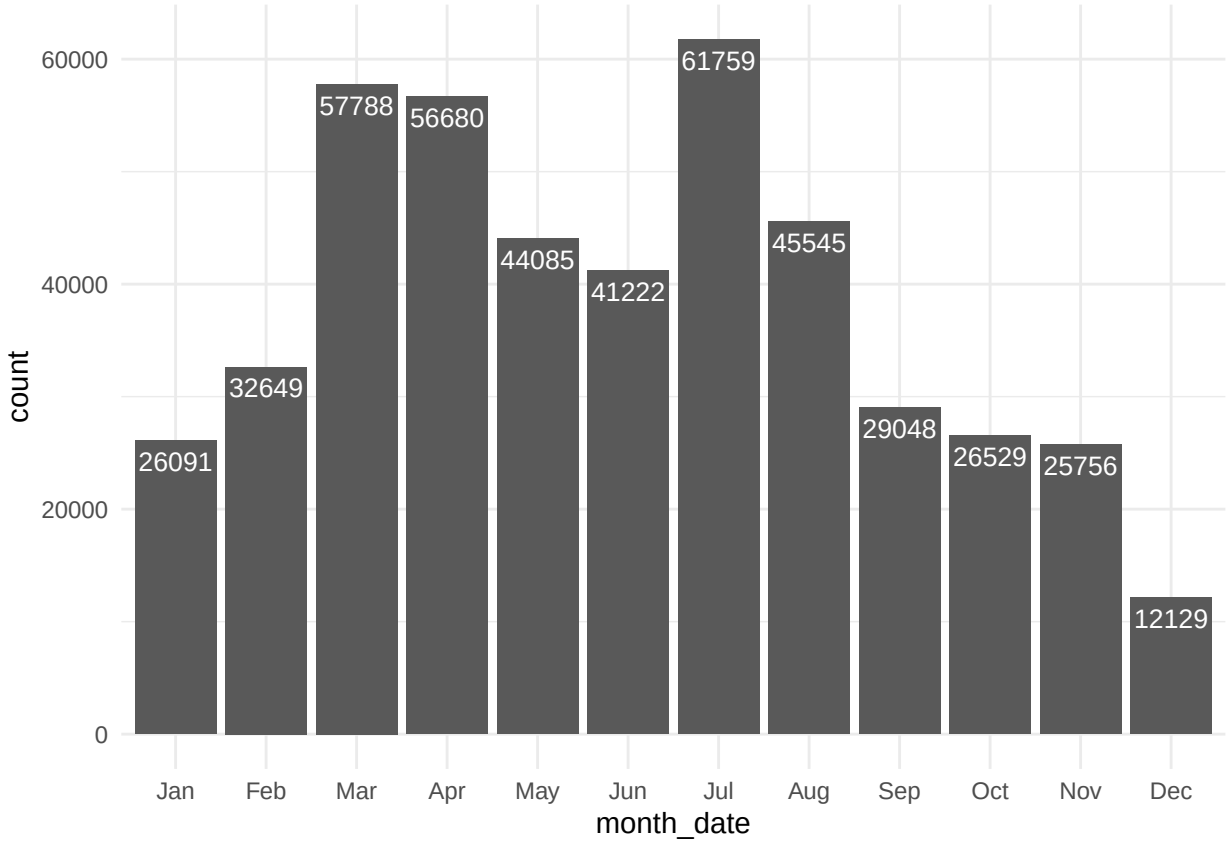

```
fire_clean_4<-fire_clean_1%>%
  filter(Country=="US", DISCOVERY_DATE>='2013-01-01')%>%
  mutate(month_date=factor(months(DISCOVERY_DATE, abbreviate = TRUE), levels = month.abb))%>%
  group_by(month_date)%>%
  summarise(count=n())




#View(fire_clean_4)
```

Which month do forest fires occur mostly? [IN US]

```
ggplot(fire_clean_4, aes(y=count, x=month_date))+geom_bar(position="dodge", stat="identity")+theme_mini
```




```
#View(fire_clean_3)
```

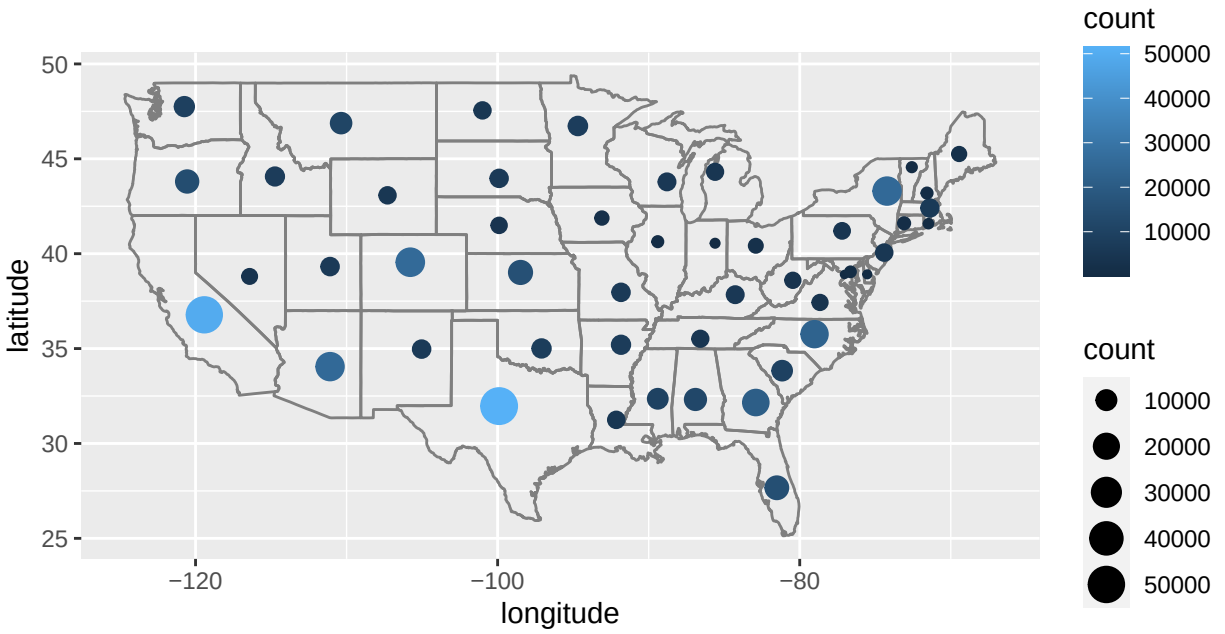
prepare data for the map chart

```
fire_clean_5<-fire_clean_1%>%
  filter(Country=="US", DISCOVERY_DATE>='2013-01-01')%>%
  left_join(latitude, by=c('STATE'))%>%
  group_by(STATE, latitude, longitude)%>%
  summarise(count=n())%>%
  ungroup()
```

```
## `summarise()` has grouped output by 'STATE', 'latitude'. You can override using the `.groups` argumer
```

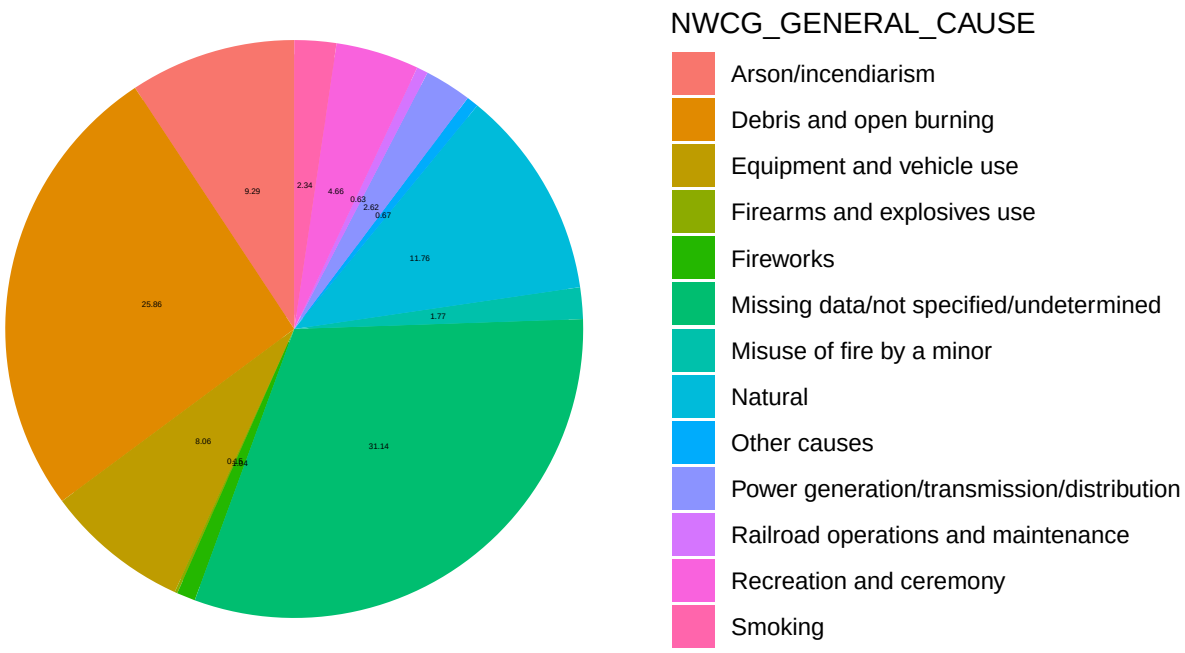
```
#View(fire_clean_5)
```

```
fire_clean_5%>%
  filter(latitude < 50, latitude > 20, longitude > -150, longitude < -50) %>%
  ggplot(aes(longitude, latitude)) +
    borders("state") +
    geom_point(aes(color = count,
                   size = count)) +coord_quickmap()
```



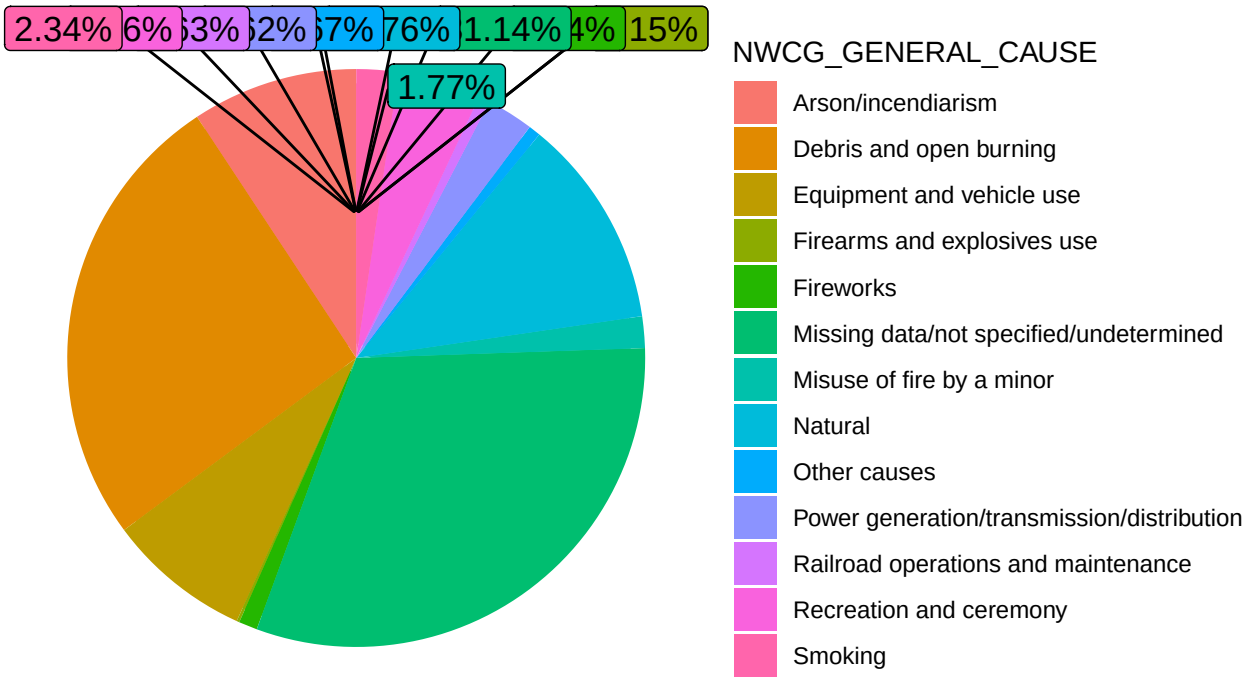

```
fire_clean_6<-fire_clean_1%>%
  filter(Country=="US", DISCOVERY_DATE>='2013-01-01')%>%
  left_join(latitude, by=c('STATE'))%>%
  group_by(NWCG_GENERAL_CAUSE)%>%
  summarise(count=n())%>%
  ungroup()%>%
  mutate(percent_1=round(count/sum(count)*100,2))%>%
  mutate(csum = rev(cumsum(rev(percent_1))),
         pos = percent_1/2 + lead(csum, 1),
         pos = if_else(is.na(pos), percent_1/2, pos))
View(fire_clean_6)
```

```
ggplot(fire_clean_6, aes(x="", y=count, fill=NWCG_GENERAL_CAUSE)) +
  geom_bar(stat="identity", width=1) +
  geom_col()+geom_text(size= 1,aes(label=percent_1, x=1),
  position = position_stack(vjust = 0.5))+coord_polar(theta="y", start=0)+theme_void()
```




```
ggplot(fire_clean_6, aes(x="", y=count, fill=NWCG_GENERAL_CAUSE)) +
  geom_bar(stat="identity", width=1) +
  geom_col()+geom_label_repel(aes(y = pos, label = paste0(percent_1, "%")),size = 4.5, nudge_x = 10,
                              show.legend = FALSE)+coord_polar(theta="y", start=0)+theme_void()
```

```
## Warning: ggrepel: 2 unlabeled data points (too many overlaps). Consider
## increasing max.overlaps
```


2.34%
6%
3%
2%
7%
76%
1.14%
4%
15%
1.77%
NWCG_GENERAL_CAUSE
Arson/incendiarism
Debris and open burning
Equipment and vehicle use
Firearms and explosives use
Fireworks
Missing data/not specified/undetermined
Misuse of fire by a minor
Natural
Other causes
Power generation/transmission/distribution
Railroad operations and maintenance
Recreation and ceremony
Smoking