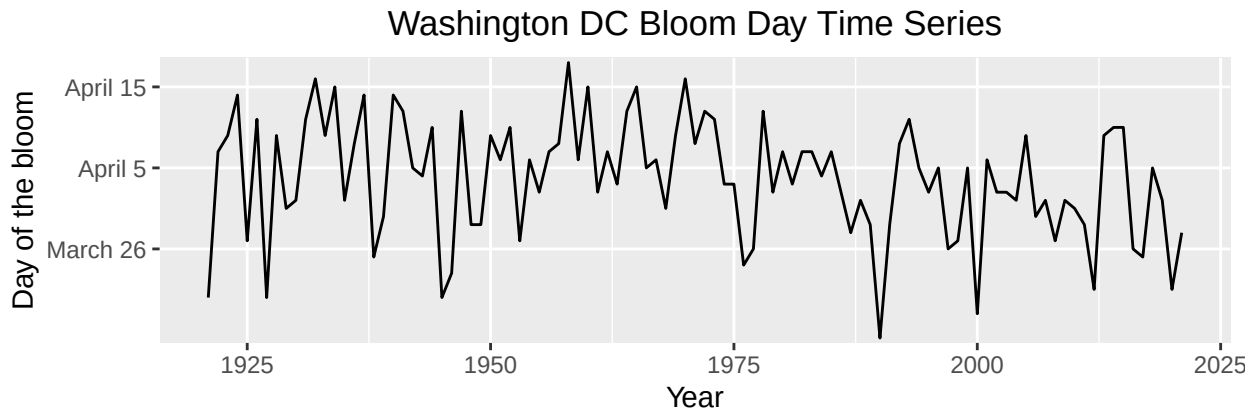
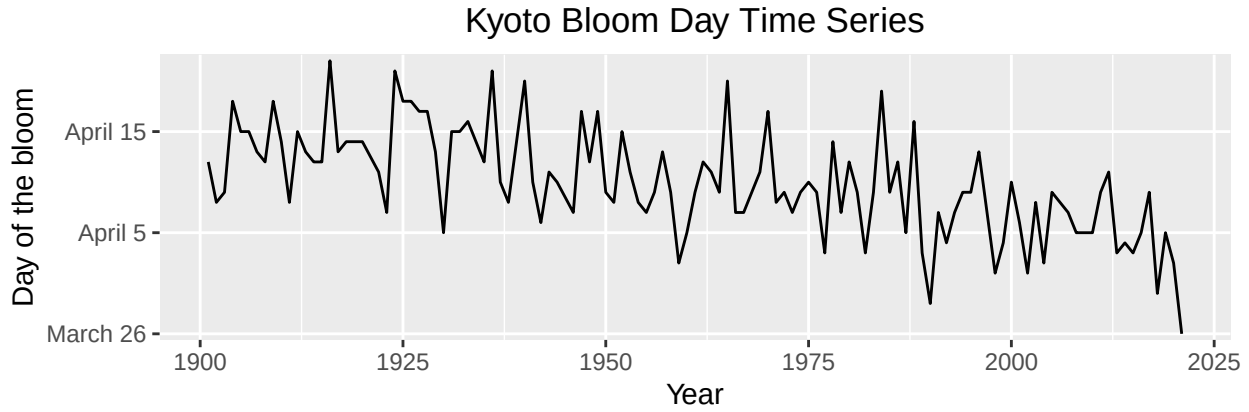
R Notebook

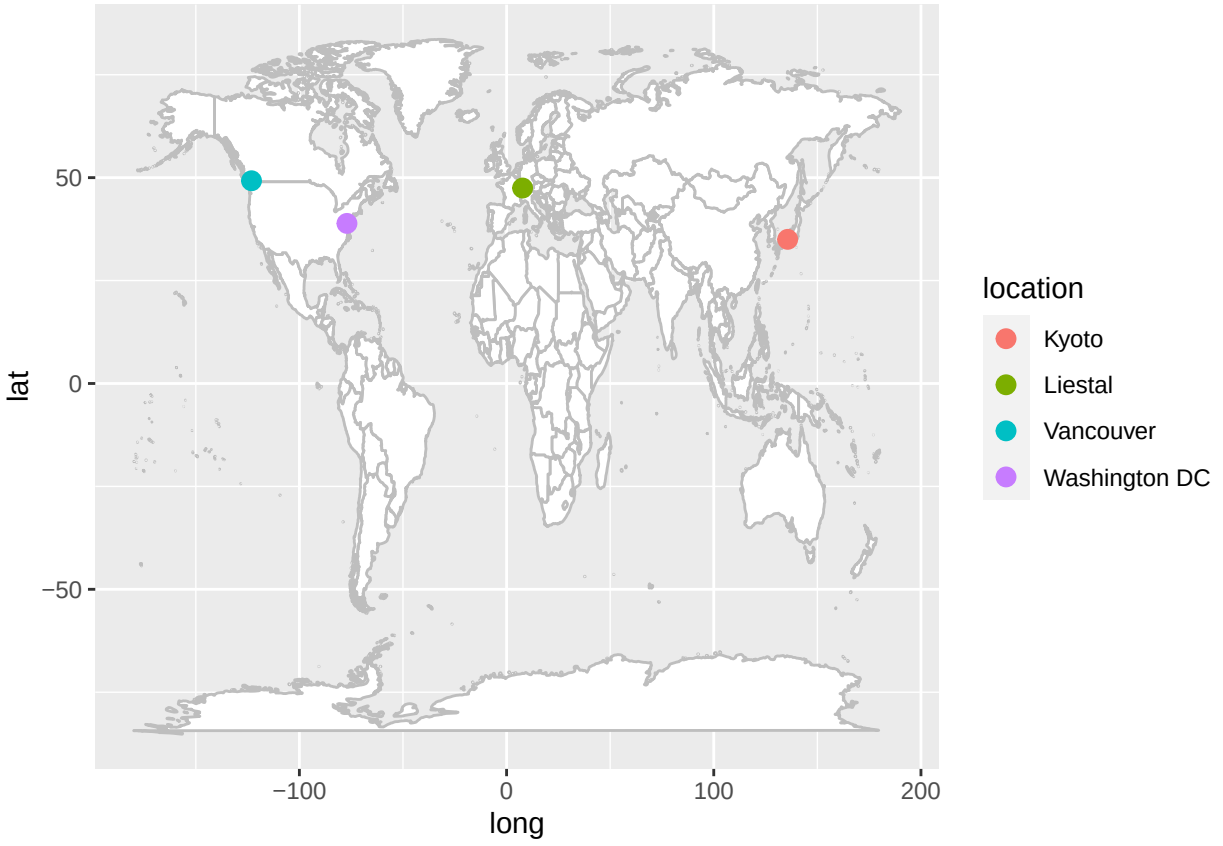
```
library(GGally)
library(rjson)
library(tidyverse)
library(lubridate)
library(cowplot)
library(ggpubr)

USA_individual_phenometrics <- read.csv(file = 'data/raw/USA-NPN_individual_phenometrics_data.csv')
USA_status_intensity <- read.csv(file = 'data/raw/USA-NPN_status_intensity_observations_data.csv')
washingtondc <- read.csv(file = 'data/raw/washingtondc.csv')
south_korea <- read.csv(file = 'data/raw/south_korea.csv')
kyoto <- read.csv(file = 'data/raw/kyoto.csv')
japan <- read.csv(file = 'data/raw/japan.csv')
liestal <- read.csv(file = 'data/raw/liestal.csv')
meteoswiss <- read.csv(file = 'data/raw/meteoswiss.csv')
south_korea <- read.csv(file = 'data/raw/south_korea.csv')
```

```
washingtondc_plot <- washingtondc %>%
ggplot(aes(year, bloom_doy)) +
  geom_line() +
  labs(title="Washington DC Bloom Day Time Series",
        x ="Year", y = "Day of the bloom") +
  scale_y_discrete(limit = seq(85, 105, 10),labels = c('March 26', 'April 5', 'April 15')) +
  theme(plot.title = element_text(hjust = 0.5))

kyoto_plot <- kyoto %>%
  filter(year>1900) %>%
  ggplot(aes(year, bloom_doy)) +
  geom_line() +
  labs(title="Kyoto Bloom Day Time Series",
        x ="Year", y = "Day of the bloom") +
  scale_y_discrete(limit = seq(85, 105, 10),labels = c('March 26', 'April 5', 'April 15')) +
  theme(plot.title = element_text(hjust = 0.5))

ggarrange(washingtondc_plot, kyoto_plot,
          ncol = 1, nrow = 2)
```

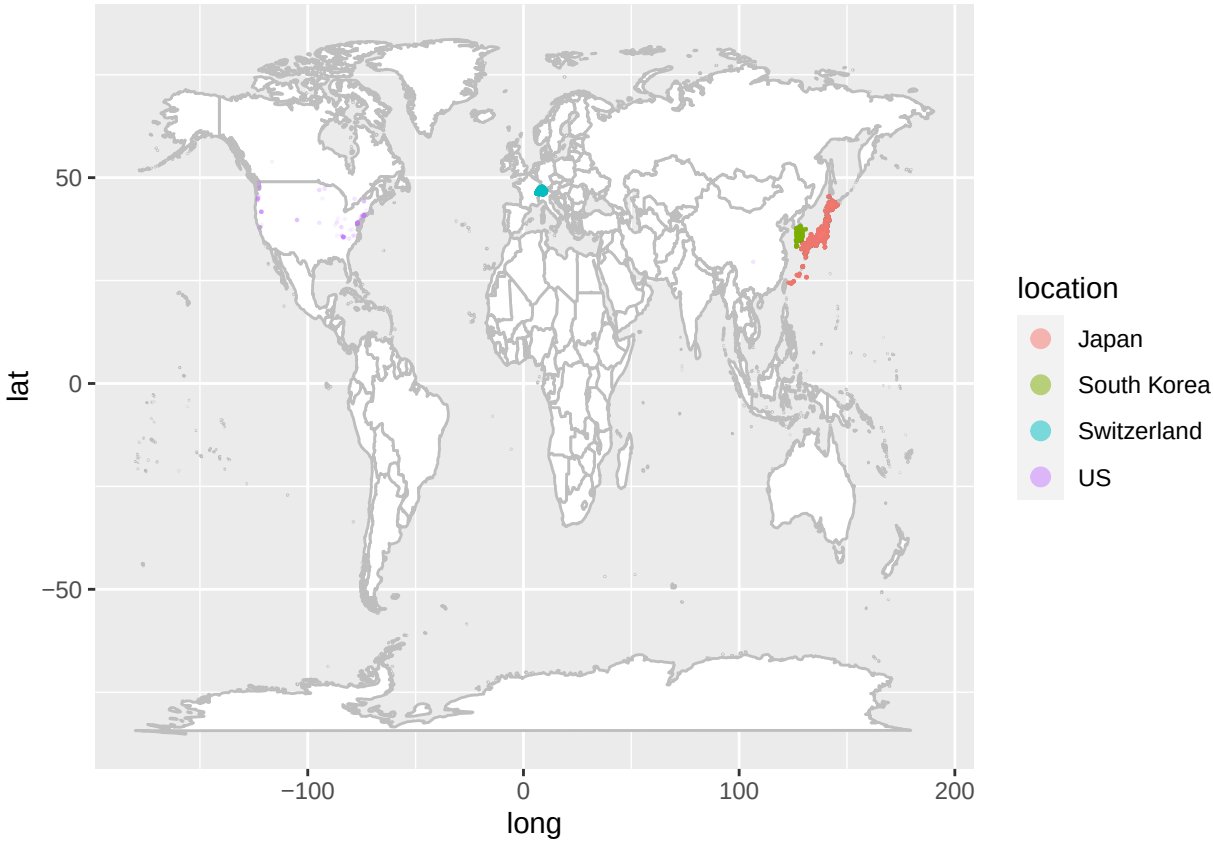




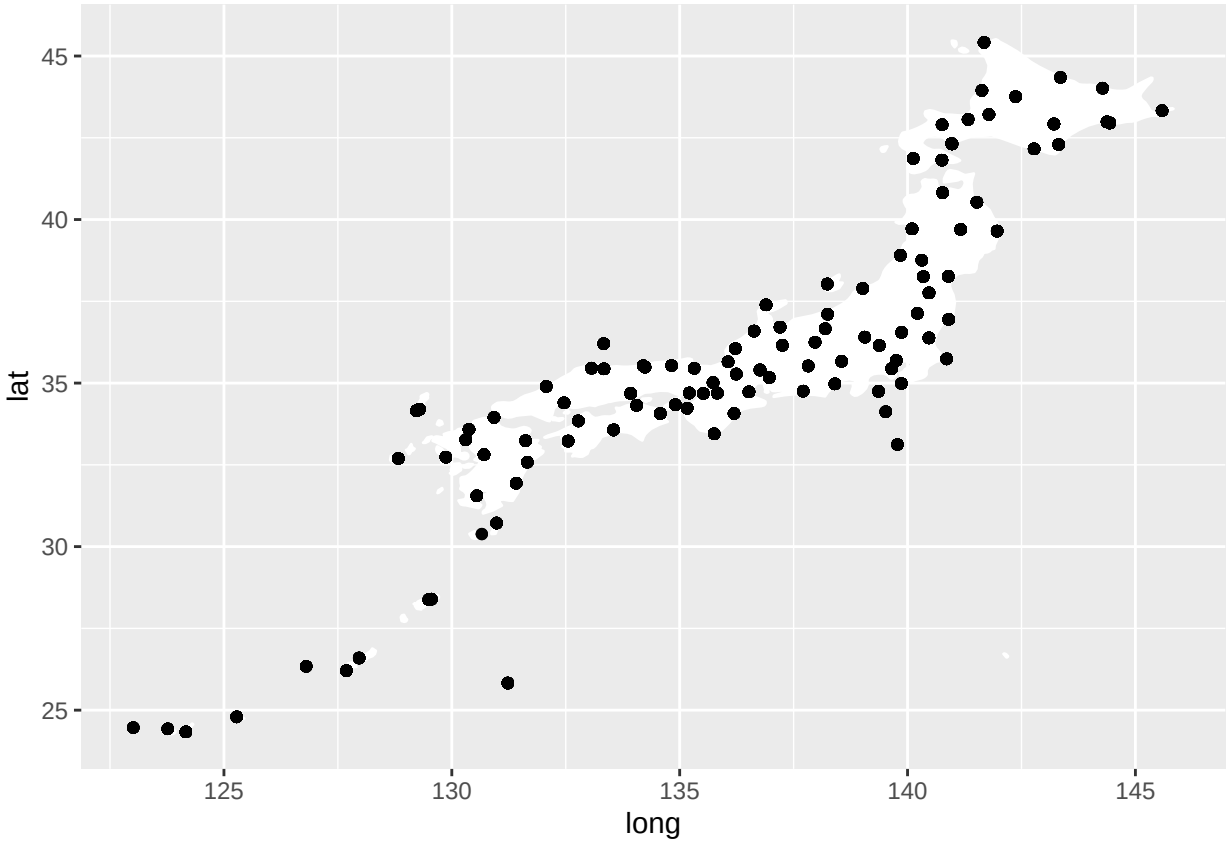

```
# heatmap?

world <- map_data("world")
cities <- rbind(kyoto[1,],washingtondc[1,],liestal[1,]) %>%
  select(lat, long) %>%
  add_row(lat=49.246292, long=-123.116226) %>%
  add_column(location=c('Kyoto','Washington DC', 'Liestal', 'Vancouver'))

ggplot() +
  geom_map(
    data = world, map = world,
    aes(long, lat, map_id = region),
    color = "gray", fill = "white") +
  geom_point(
    data = cities,
    aes(long, lat, color=location),
    size=3
  )
```



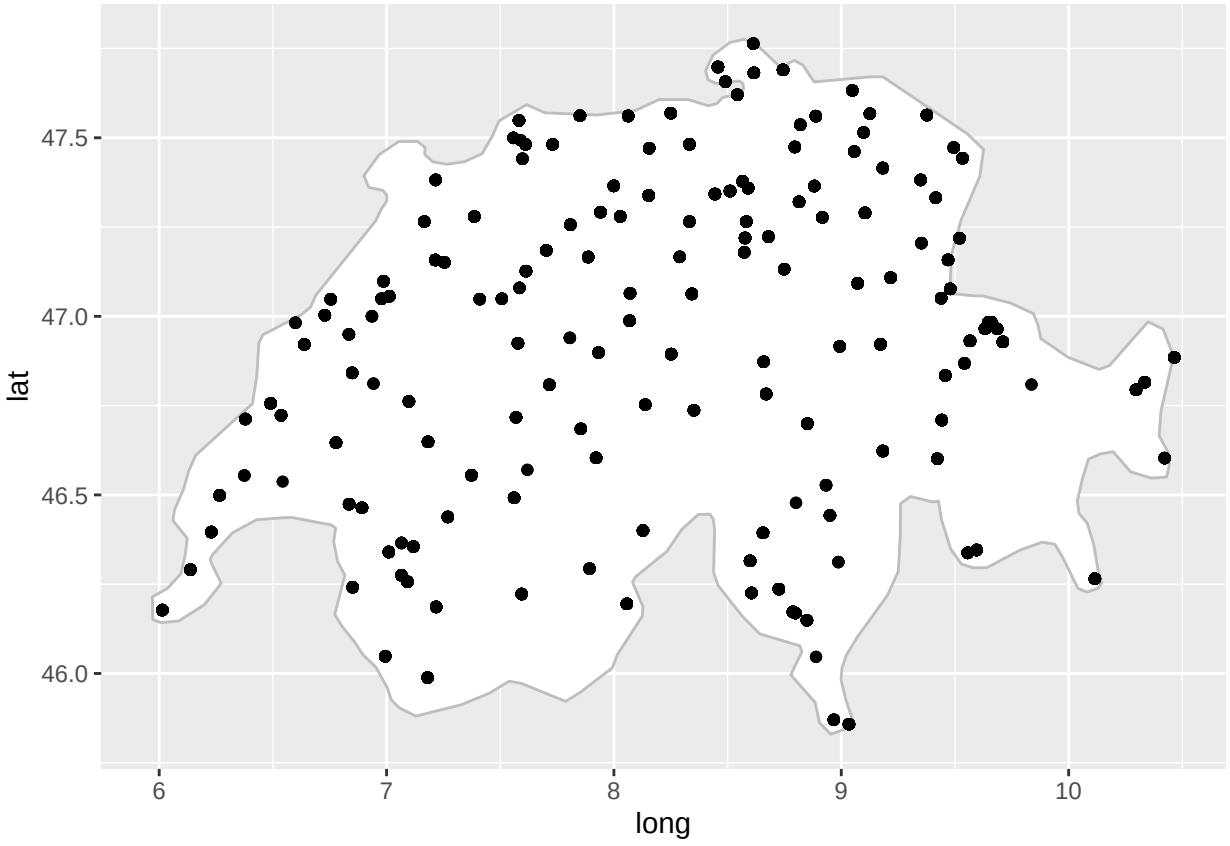

```
ggplot() +
  geom_map(
    data = world, map = world,
    aes(long, lat, map_id = region, fill = region),
    color = "gray", fill = "white") +
  geom_point(
    data = rbind(japan %>% mutate(location=sub("/.*", "",location)) %>% select(location,long,lat),
                 meteoswiss %>% mutate(location=sub("/.*", "",location)) %>% select(location,long,lat),
                 south_korea %>% mutate(location=sub("/.*", "",location)) %>% select(location,long,lat)
                 USA_individual_phenometrics %>% mutate(location='US') %>%
                   select(location,Longitude,Latitude) %>%
                   rename(long=Longitude, lat=Latitude)),
    aes(long, lat, color=location),
    alpha=0.1, size=0.1) +
  guides(colour = guide_legend(override.aes = list(alpha=0.5,size=3)))
```



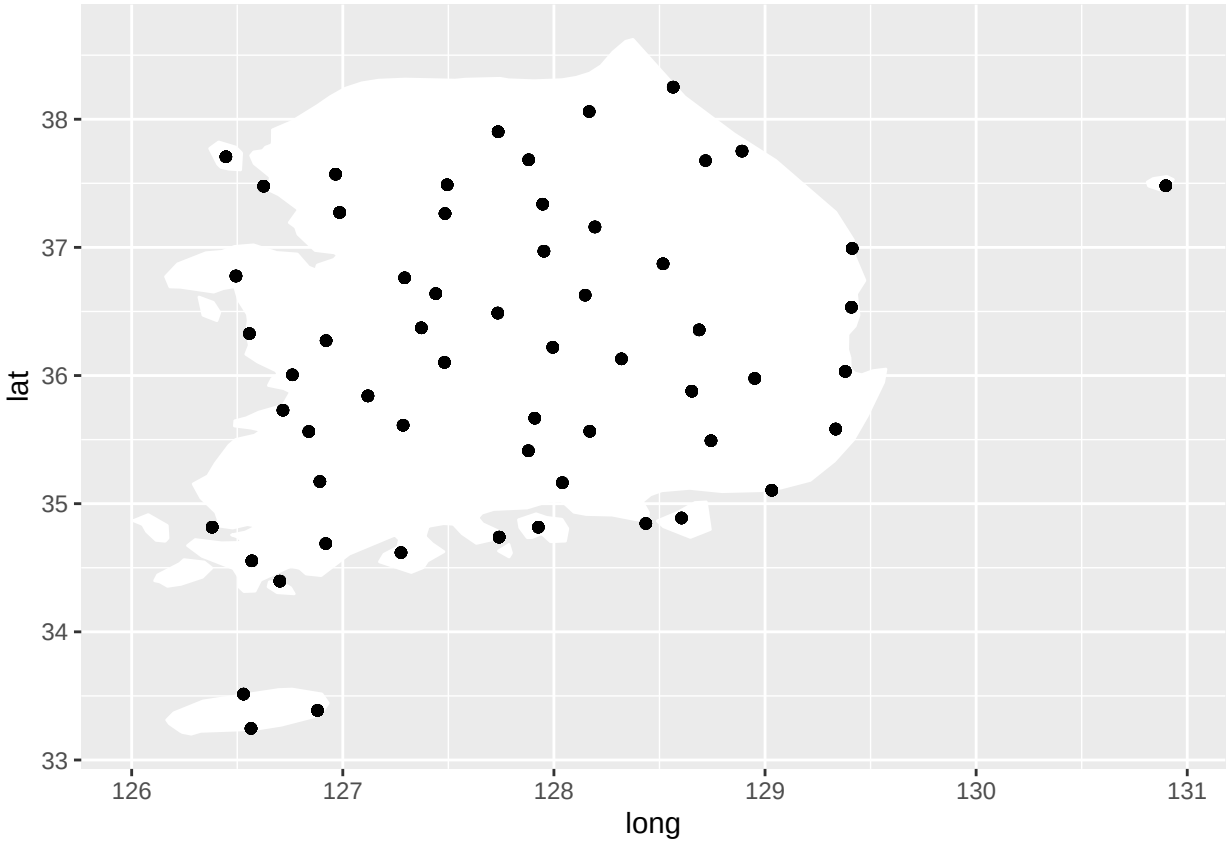

```
jp <- map_data('world', 'Japan')

ggplot() +
  geom_map(
    data = jp, map = jp,
    aes(long, lat, map_id = region),
    color = "white", fill = "white"
  ) +
  geom_point(
    data = japan,
    aes(long, lat))
```




```
swiss <- map_data('world', 'Switzerland')

ggplot() +
  geom_map(
    data = swiss, map = swiss,
    aes(long, lat, map_id = region),
    color = "gray", fill = "white"
  ) +
  geom_point(
    data = meteoswiss,
    aes(long, lat))
```

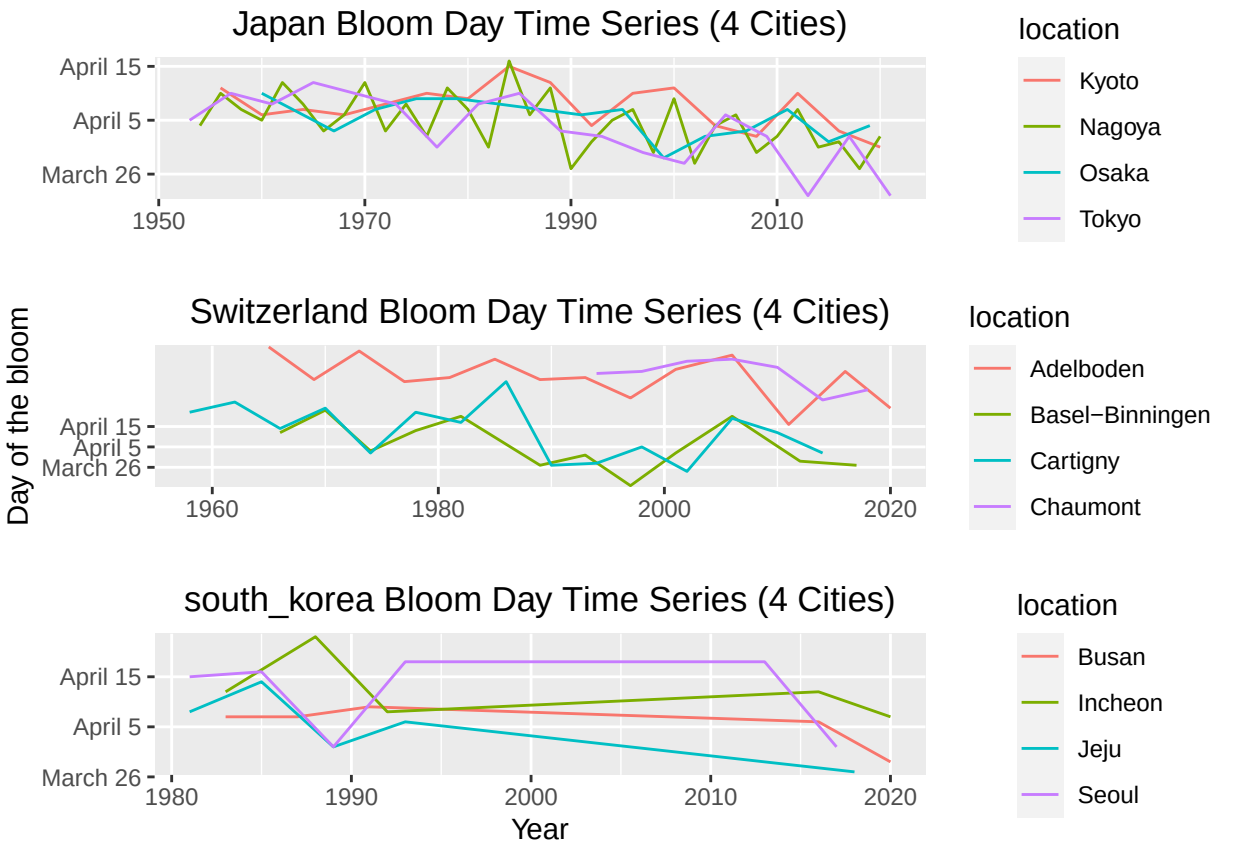



```
sk <- map_data('world', 'South Korea')

ggplot() +
  geom_map(
    data = sk, map = sk,
    aes(long, lat, map_id = region),
    color = "white", fill = "white"
  ) +
  geom_point(
    data = south_korea,
    aes(long, lat))
```




```
#unique(japan$location)
japan_plot <- japan %>% mutate(location=gsub("Japan/","",location)) %>%
  filter(location==c('Nagoya', 'Tokyo', 'Kyoto', 'Osaka'))%>%
  ggplot(aes(year, bloom_doy)) +
  aes(x = year, color=location) +
  geom_line() +
  labs(title="Japan Bloom Day Time Series (4 Cities)",
       x =" ", y = " ") +
  scale_y_discrete(limit = seq(85, 105, 10),labels = c('March 26', 'April 5', 'April 15')) +
  theme(plot.title = element_text(hjust = 0.5))


#unique(meteoswiss$location)
meteoswiss_plot <- meteoswiss %>% mutate(location=gsub("Switzerland/","",location)) %>% arrange(locatio
  filter(location==c('Adelboden', 'Basel-Binningen', 'Cartigny', 'Chaumont')) %>%
  ggplot(aes(year, bloom_doy)) +
  aes(x = year, color=location) +
  geom_line() +
  labs(title="Switzerland Bloom Day Time Series (4 Cities)",
       x =" ", y = "Day of the bloom") +
  scale_y_discrete(limit = seq(85, 105, 10),labels = c('March 26', 'April 5', 'April 15')) +
  theme(plot.title = element_text(hjust = 0.5))



#unique(south_korea$location)
```

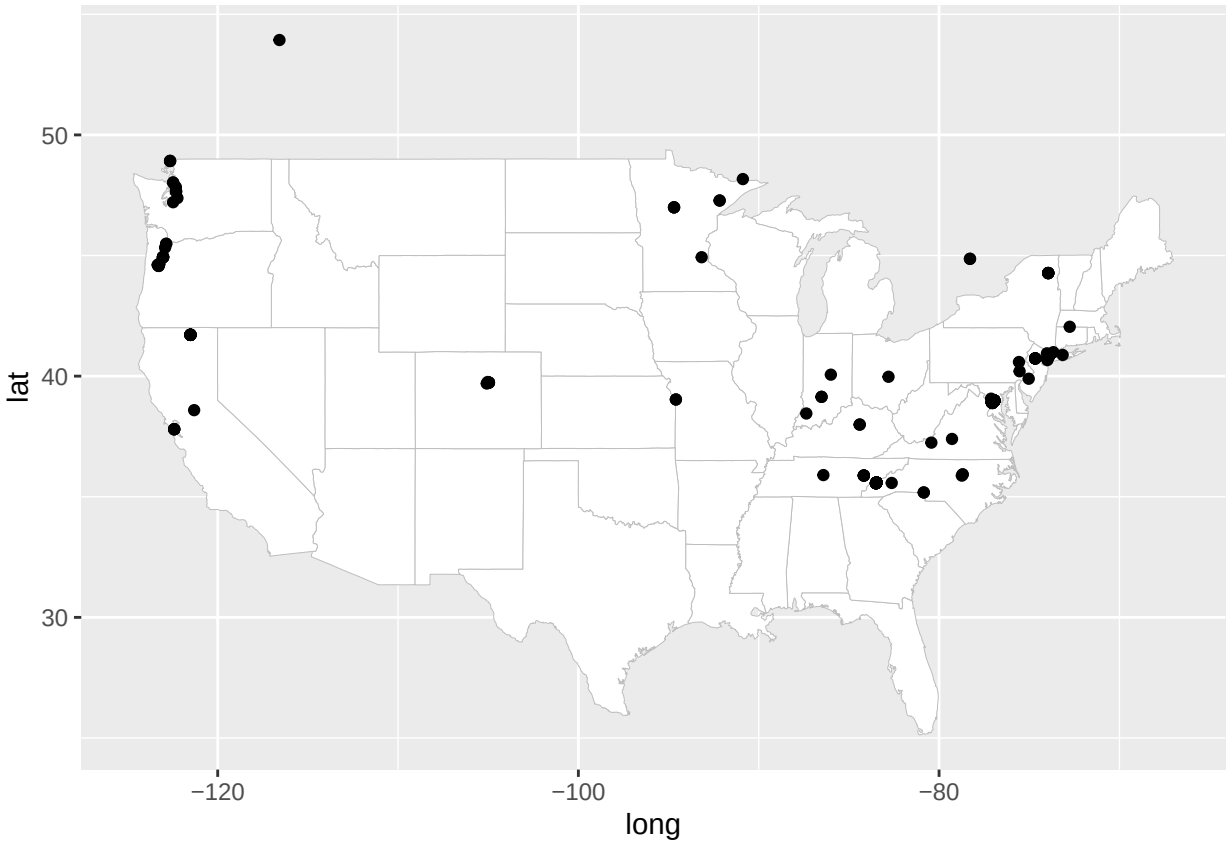
```
south_korea_plot <- south_korea %>% mutate(location=gsub("South Korea/","",location)) %>% arrange(locati
ggplot(aes(year, bloom_doy)) +
    aes(x = year, color=location) +
    geom_line(na.rm = FALSE) +
  labs(title="south_korea Bloom Day Time Series (4 Cities)",
       x ="Year", y = " ") +
  scale_y_discrete(limit = seq(85, 105, 10),labels = c('March 26', 'April 5', 'April 15')) +
  theme(plot.title = element_text(hjust = 0.5))

ggarrange(japan_plot, meteoswiss_plot, south_korea_plot,
          align = "v",
          ncol = 1, nrow = 3)
```




```
USA_individual_phenometrics <- USA_individual_phenometrics %>%
  select(-c(Genus,Kingdom,Species_ID,Phenophase_ID,Phenophase_Description)) %>%
  mutate(First_Yes_Date = make_date(First_Yes_Year,First_Yes_Month,First_Yes_Day),
         First_Yes_MonthDay = format(First_Yes_Date,'%m-%d'),
         Last_Yes_Date = make_date(Last_Yes_Year,Last_Yes_Month,Last_Yes_Day),
         Last_Yes_MonthDay = format(Last_Yes_Date,'%m-%d')) %>%
  filter(State != "Chongqing Shi",
         State != "Hyogo",
         First_Yes_DOY < 180)
```

```
usa <- map_data("state")

ggplot() +
```

```
  geom_map(
    data = usa, map = usa,
    aes(long, lat, map_id = region),
    color = "gray", fill = "white", size = 0.1
  ) +
  geom_point(
    data = USA_individual_phenometrics,
    aes(Longitude, Latitude))
```




```
unique(USA_individual_phenometrics%>%select(Common_Name))
```

```
##                   Common_Name
## 1              Yoshino cherry
## 2               bitter cherry
## 10  Japanese flowering cherry
## 22                 pin cherry
## 154               sour cherry
```

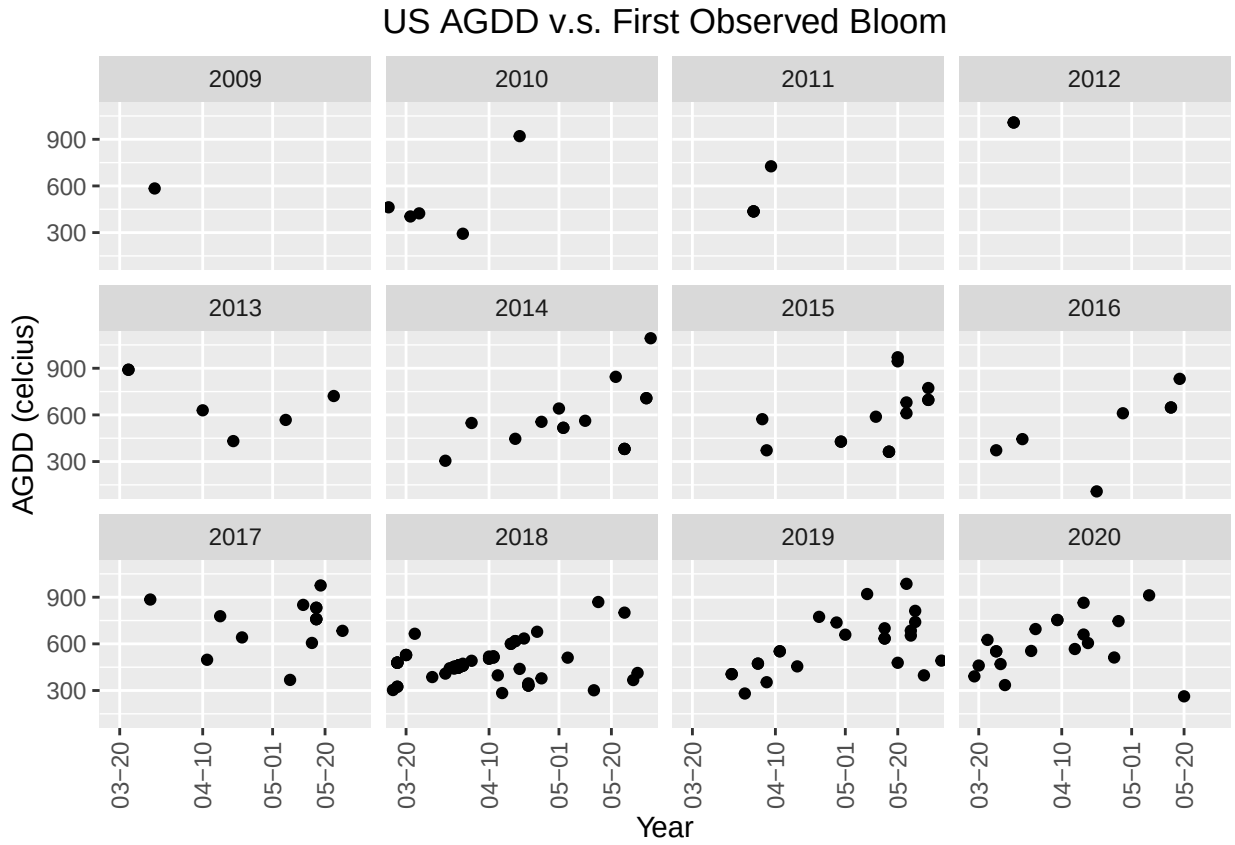
```
unique(USA_individual_phenometrics%>%filter(State=='DC')%>%select(Common_Name))
```

```
##                 Common_Name
## 1            Yoshino cherry
## 4 Japanese flowering cherry
```

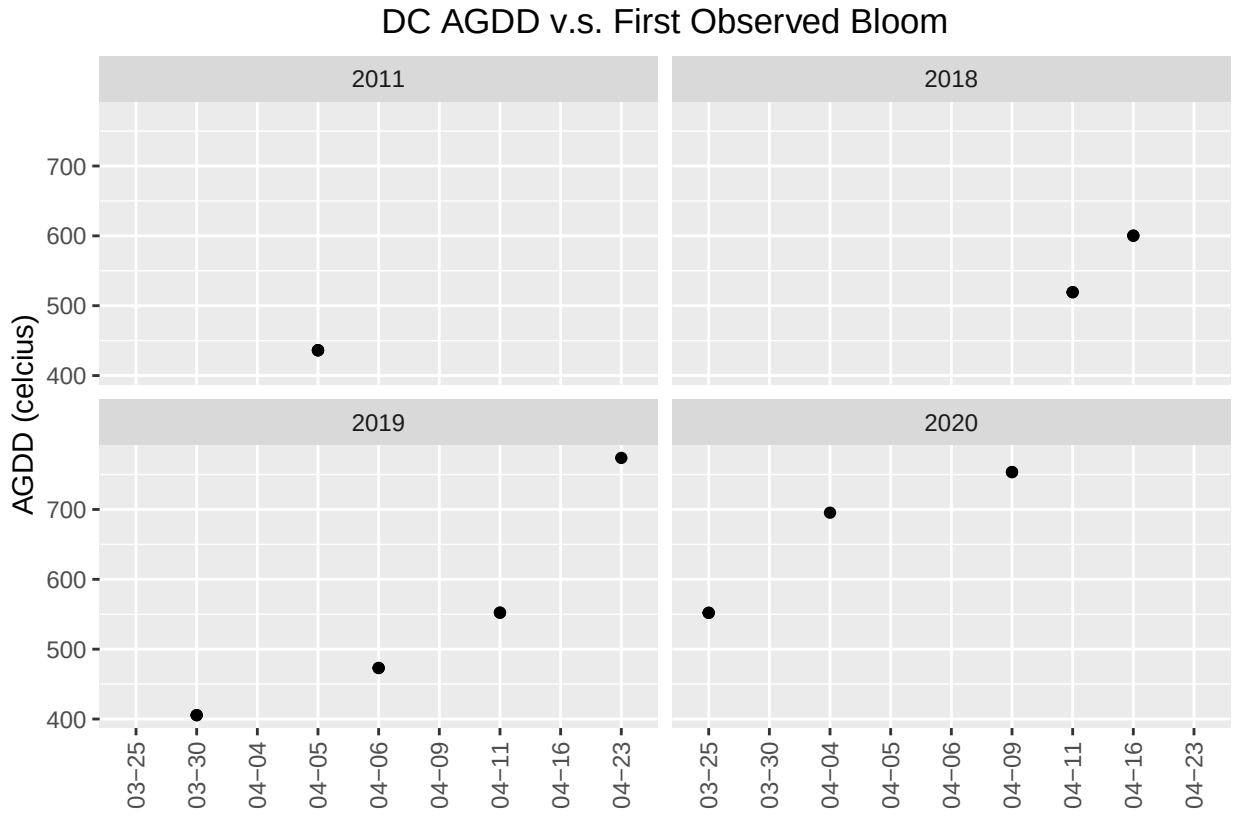
```
USA_individual_phenometrics %>%
  filter(State == 'DC') %>%
  select(First_Yes_Date)
```

```
##    First_Yes_Date
## 1      2011-04-05
## 2      2011-04-05
## 3      2011-04-05
## 4      2018-04-16
## 5      2018-04-16
## 6      2018-04-11
## 7      2018-04-11
## 8      2019-04-11
## 9      2019-04-11
## 10     2019-03-30
## 11     2019-04-06
## 12     2019-03-30
## 13     2019-04-06
## 14     2019-04-23
## 15     2020-04-09
## 16     2020-04-09
## 17     2020-03-25
## 18     2020-03-25
## 19     2020-04-04
```

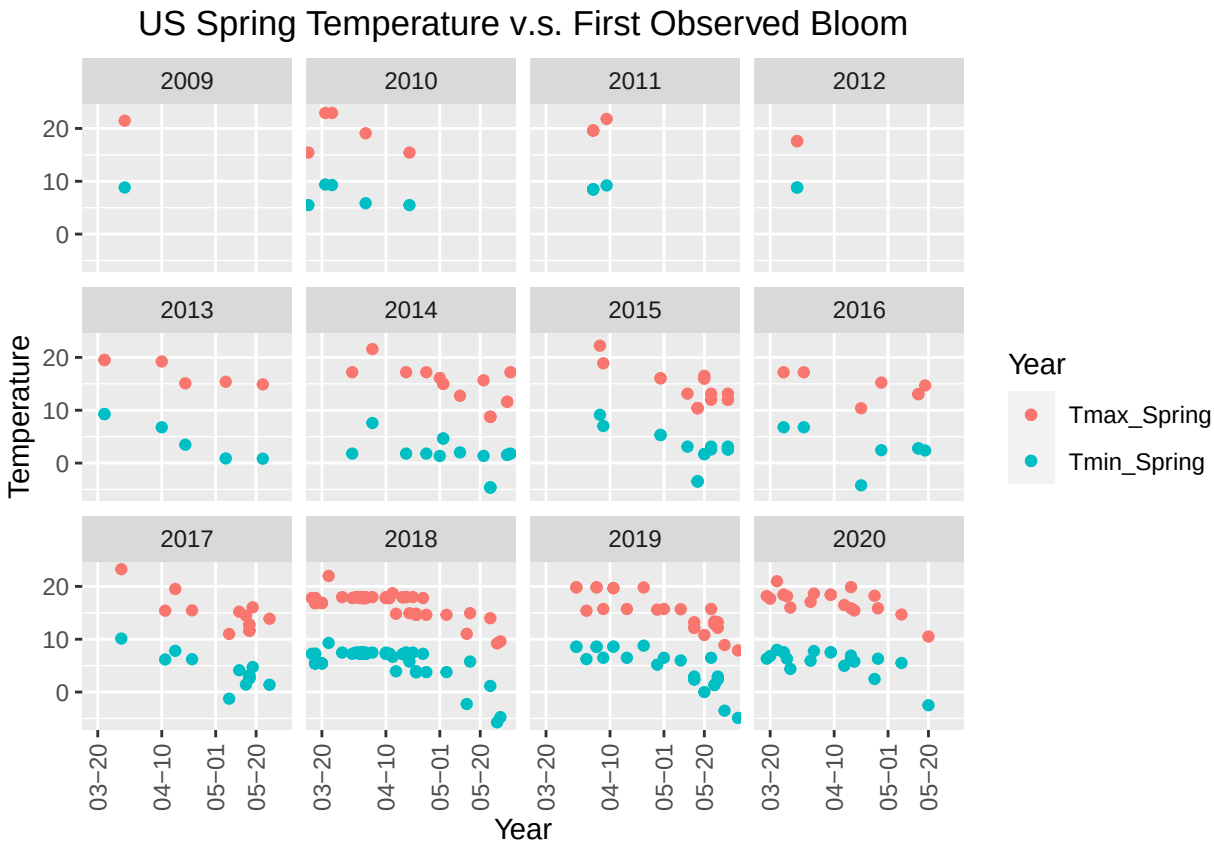
```
USA_individual_phenometrics %>%
  filter(AGDD > 0) %>%
  ggplot() +
  geom_point(aes(x = First_Yes_MonthDay, y = AGDD)) +
  facet_wrap(~First_Yes_Year) +
  labs(title="US AGDD v.s. First Observed Bloom",
       x ="Year", y = "AGDD (celcius)") +
  scale_x_discrete(breaks = c('03-20','04-10','05-01','05-20')) +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```



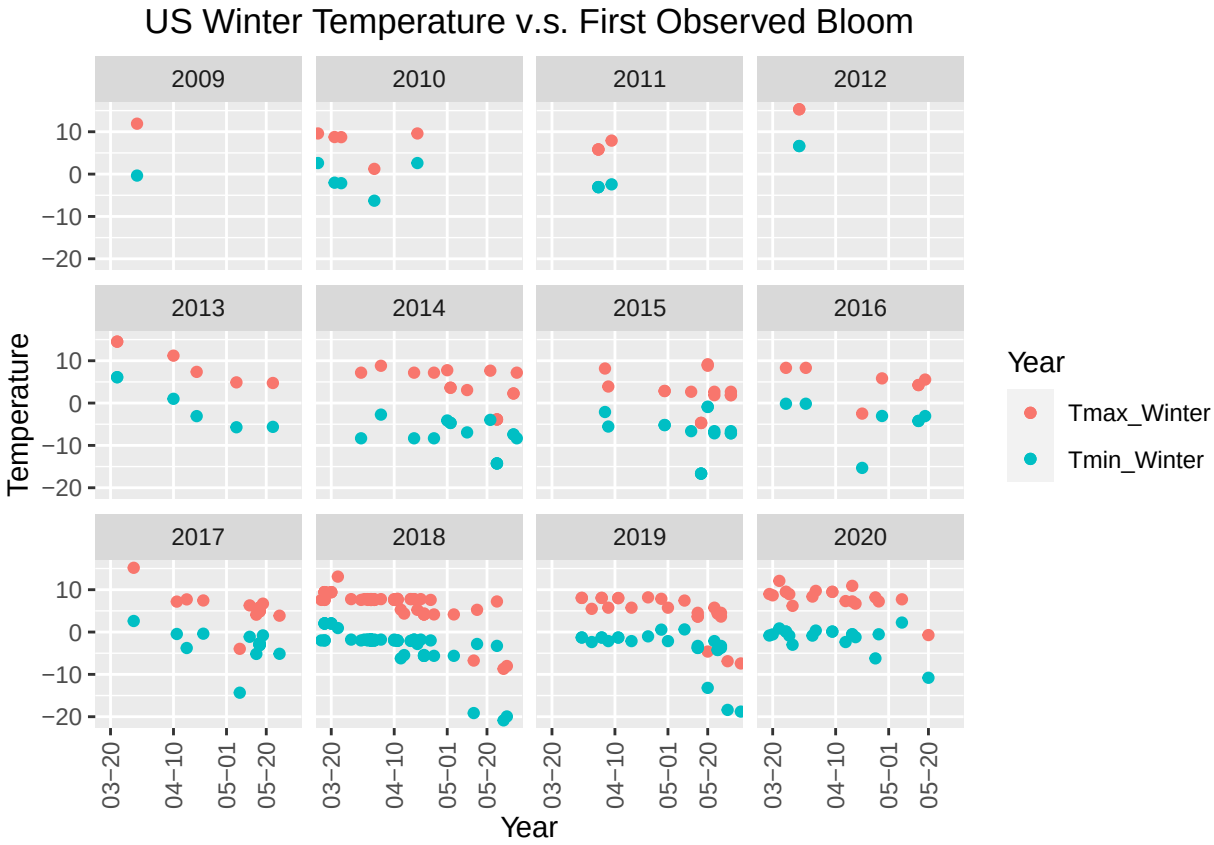

```
USA_individual_phenometrics %>%
  filter(AGDD > 0, State=='DC') %>%
  ggplot(aes(x = First_Yes_MonthDay, y = AGDD)) + geom_point() +
  facet_wrap(~First_Yes_Year) +
  labs(title="DC AGDD v.s. First Observed Bloom",
       x ="", y = "AGDD (celcius)") +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hjus
```



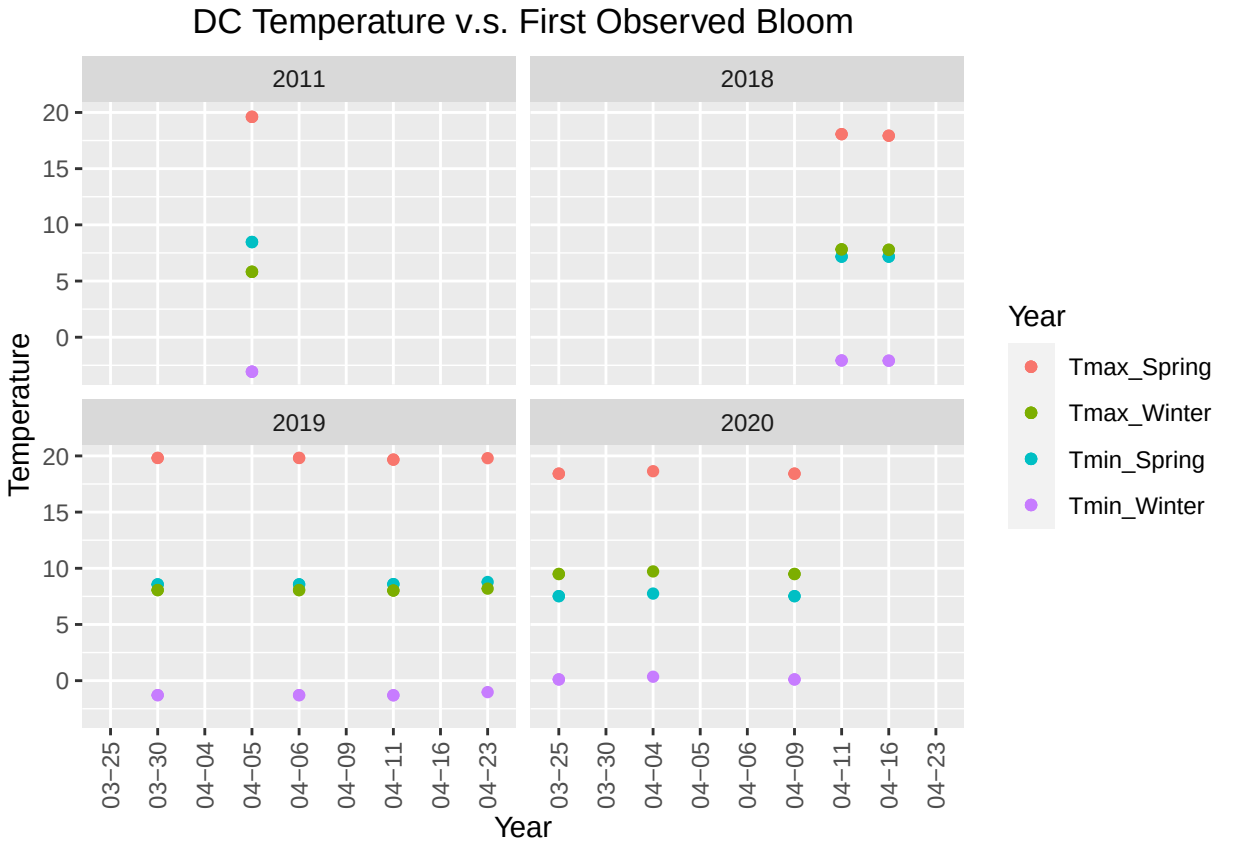

```
USA_individual_phenometrics %>%
  filter(Tmin_Winter !=-9999) %>%
  ggplot() +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Spring, color='Tmax_Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Spring, color='Tmin_Spring')) +
#  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Winter, color='Tmax_Winter')) +
#  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Winter, color='Tmin_Winter')) +
  facet_wrap(~First_Yes_Year) +
  labs(title="US Spring Temperature v.s. First Observed Bloom",
       x ="Year", y = "Temperature", color='Year') +
  scale_x_discrete(breaks = c('03-20','04-10','05-01','05-20')) +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```



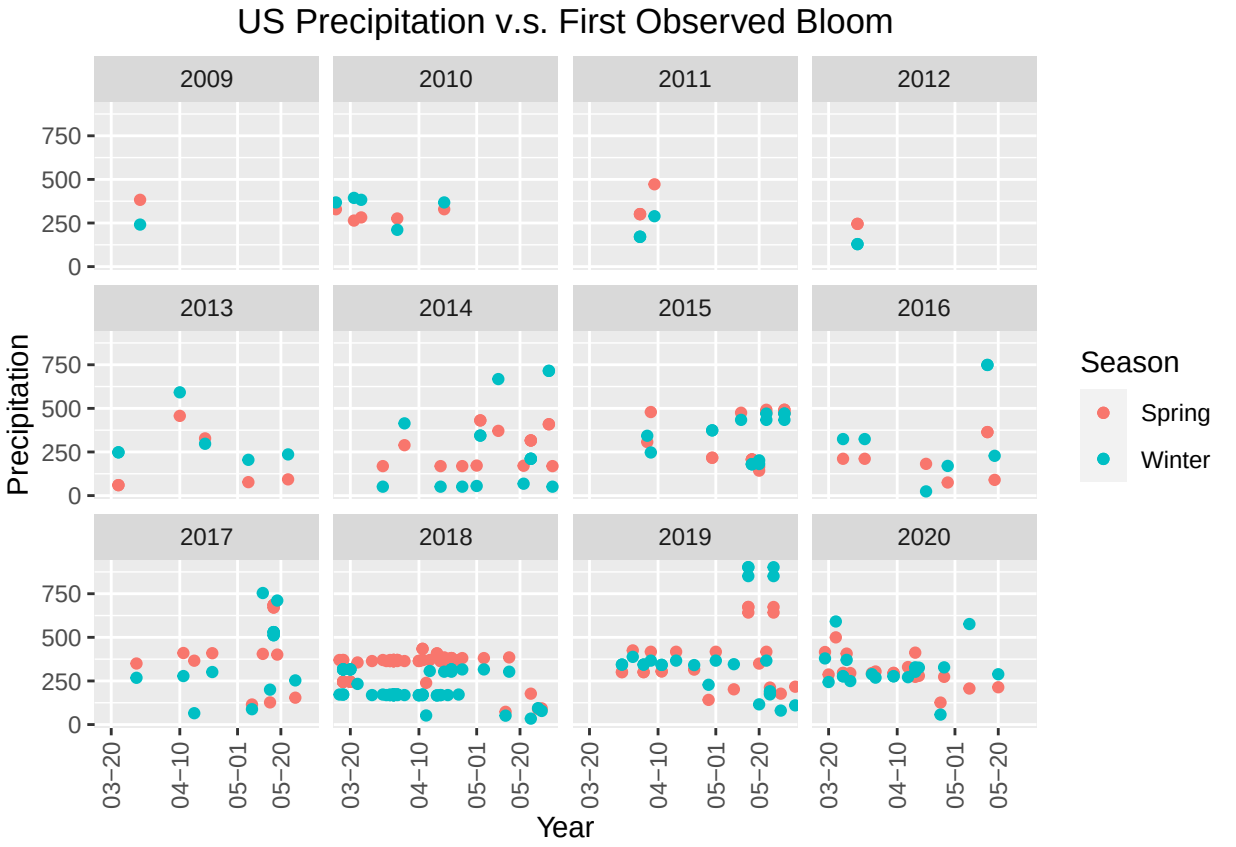

```
USA_individual_phenometrics %>%
  filter(Tmin_Winter !=-9999) %>%
  ggplot() +
#  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Spring, color='Tmax_Spring')) +
#  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Spring, color='Tmin_Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Winter, color='Tmax_Winter')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Winter, color='Tmin_Winter')) +
  facet_wrap(~First_Yes_Year) +
  labs(title="US Winter Temperature v.s. First Observed Bloom",
       x ="Year", y = "Temperature", color='Year') +
  scale_x_discrete(breaks = c('03-20','04-10','05-01','05-20')) +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```



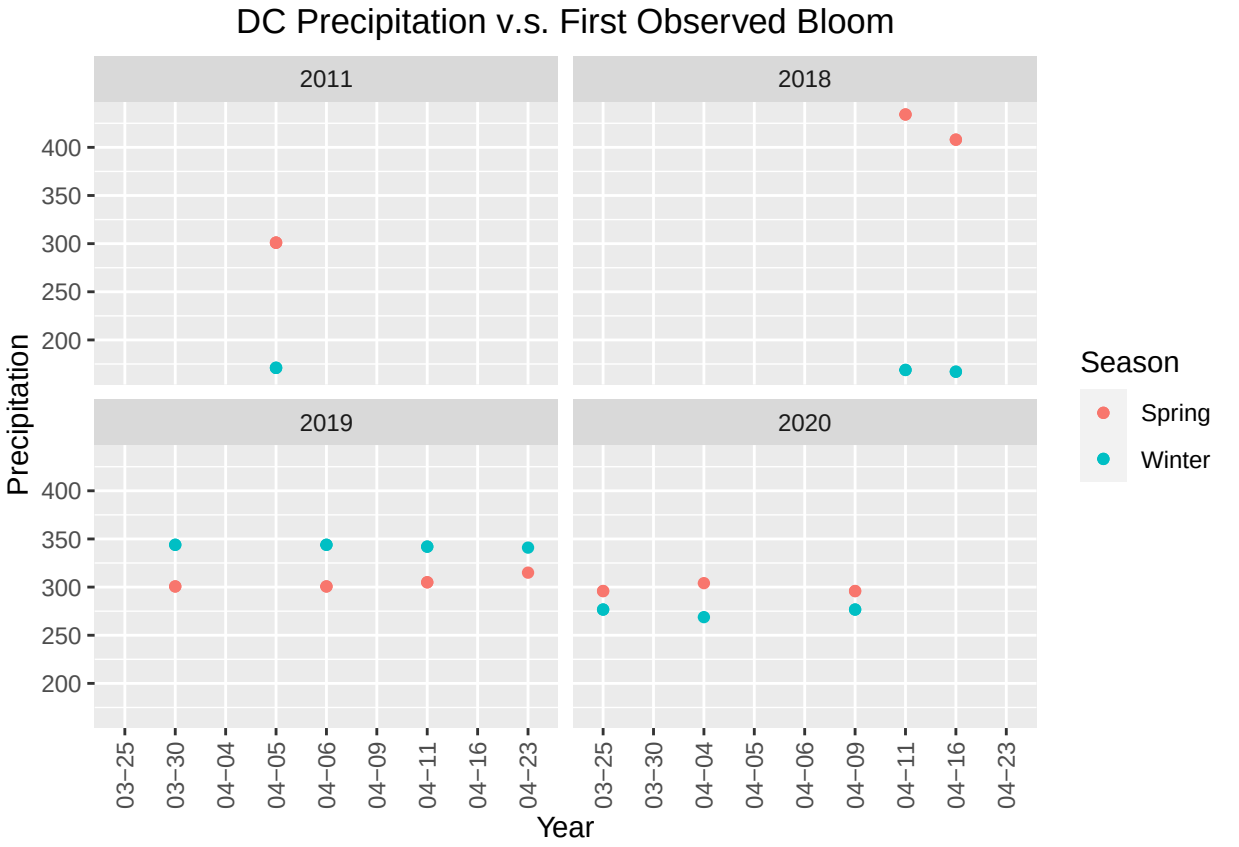

```
USA_individual_phenometrics %>%
  filter(Tmin_Winter !=-9999, State=='DC') %>%
  ggplot() +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Spring, color='Tmax_Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Spring, color='Tmin_Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmax_Winter, color='Tmax_Winter')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Tmin_Winter, color='Tmin_Winter')) +
  facet_wrap(~First_Yes_Year) +
  labs(title="DC Temperature v.s. First Observed Bloom",
       x ="Year", y = "Temperature", color='Year') +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```



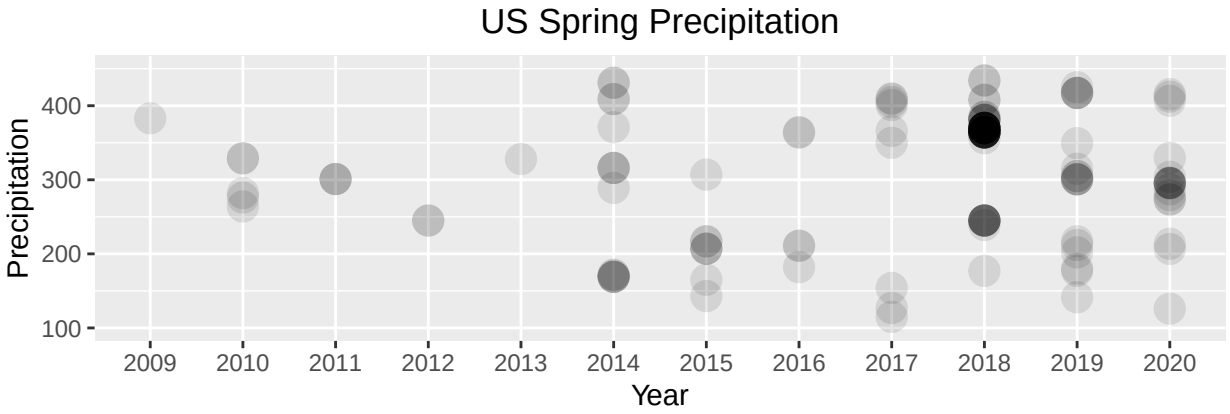

```
USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0) %>%
  ggplot() +
  geom_point(aes(x = First_Yes_MonthDay, y = Prcp_Spring, color = 'Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Prcp_Winter, color = 'Winter')) +
  facet_wrap(~First_Yes_Year) +
  labs(title="US Precipitation v.s. First Observed Bloom",
       x ="Year", y = "Precipitation", color='Season') +
  scale_x_discrete(breaks = c('03-20','04-10','05-01','05-20')) +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```

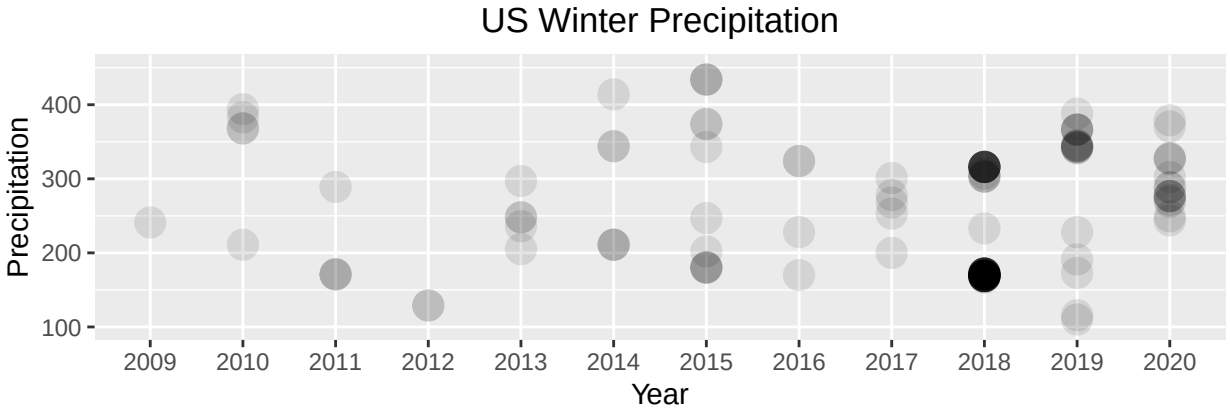



```
USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0, State == 'DC') %>%
  ggplot() +
  geom_point(aes(x = First_Yes_MonthDay, y = Prcp_Spring, color = 'Spring')) +
  geom_point(aes(x = First_Yes_MonthDay, y = Prcp_Winter, color = 'Winter')) +
  facet_wrap(~First_Yes_Year) +
  labs(title="DC Precipitation v.s. First Observed Bloom",
       x ="Year", y = "Precipitation", color='Season') +
  theme(plot.title = element_text(hjust = 0.5), axis.text.x = element_text(angle = 90, vjust = 0.5, hju
```

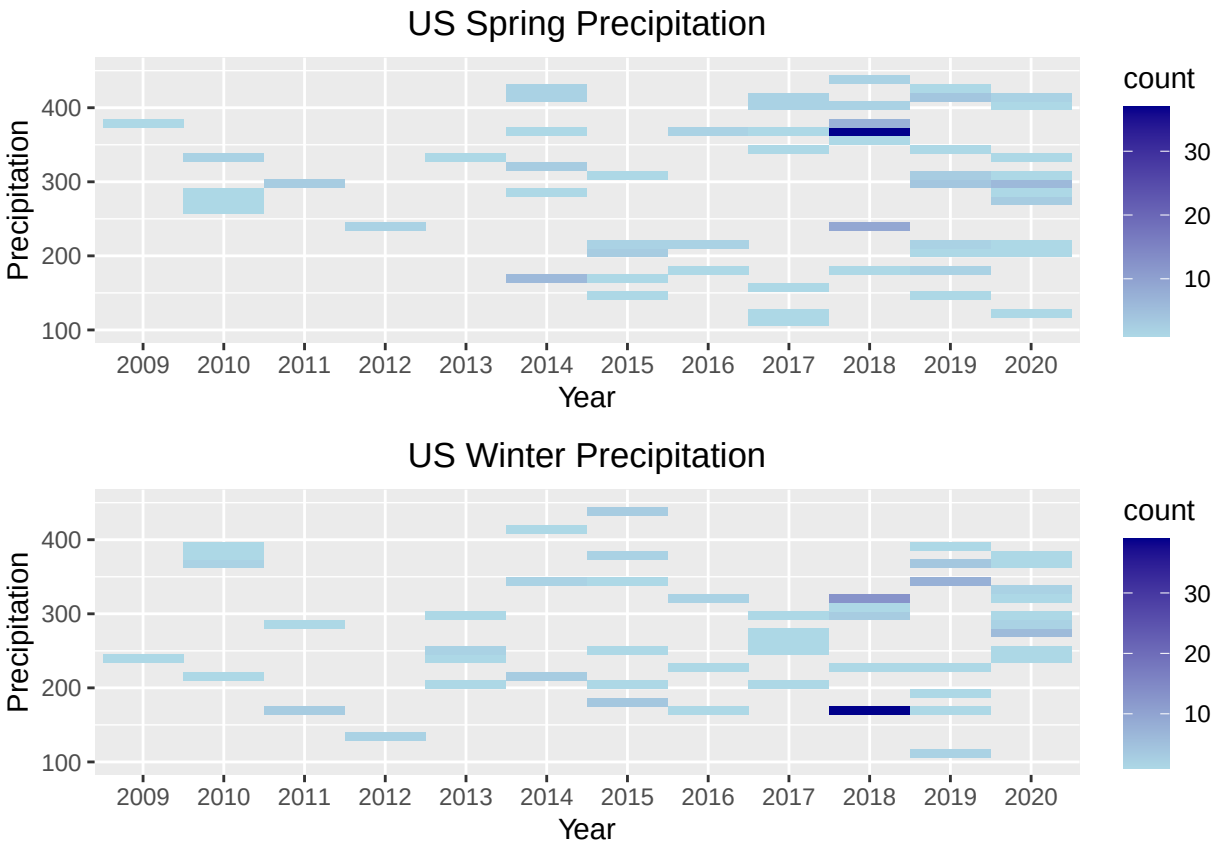



```
US_prcp_spring <- USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0) %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Spring)) + geom_point(size=5, alpha=0.1) +
  labs(title="US Spring Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450)

US_prcp_winter <- USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0) %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Winter)) + geom_point(size=5, alpha=0.1) +
  labs(title="US Winter Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450)

ggarrange(US_prcp_spring, US_prcp_winter,
          align = "v",
          ncol = 1, nrow = 2)
```

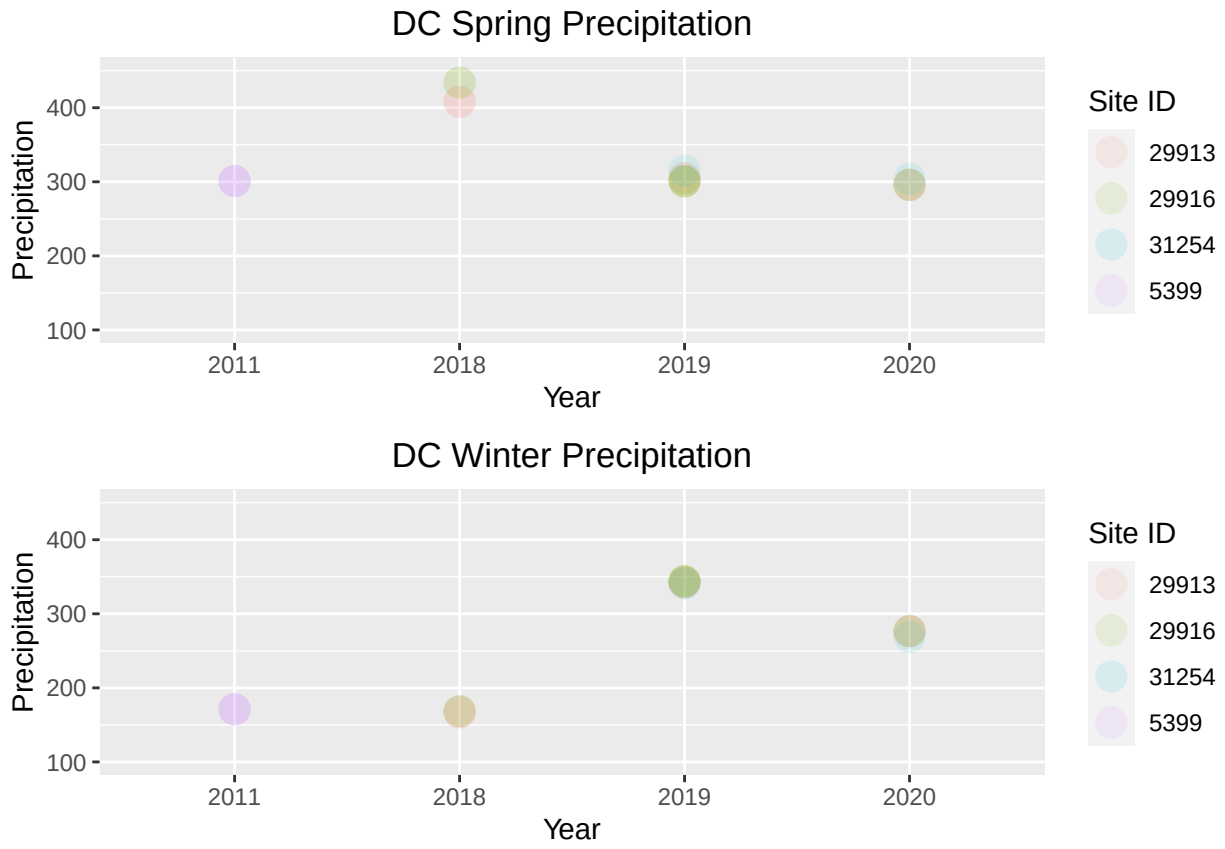




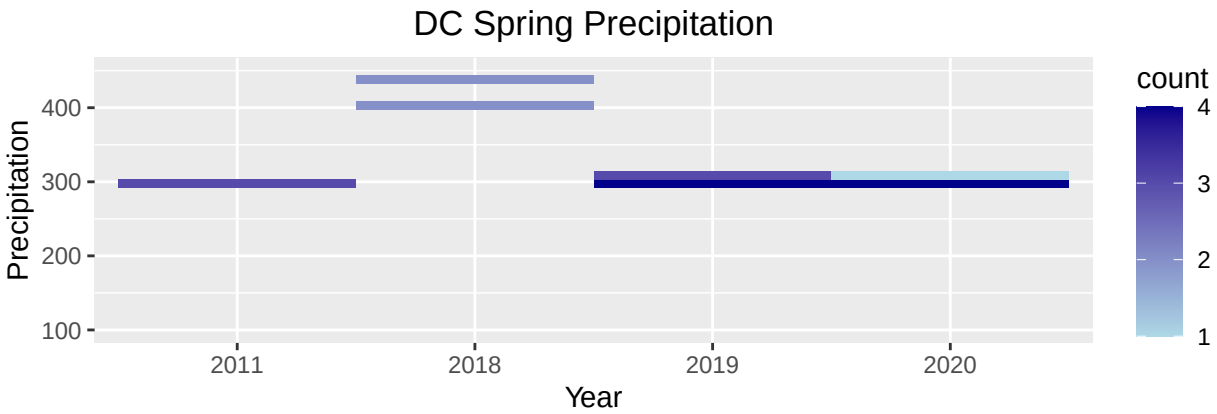

```
US_prcp_spring2 <- USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0) %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Spring)) + geom_bin2d() +
  labs(title="US Spring Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450) +
  scale_fill_continuous(low="lightblue", high = "darkblue")

US_prcp_winter2 <- USA_individual_phenometrics %>%
  filter(Prcp_Spring > 0, Prcp_Winter > 0) %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Winter)) + geom_bin2d() +
  labs(title="US Winter Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450) +
  scale_fill_continuous(low="lightblue", high = "darkblue")

ggarrange(US_prcp_spring2, US_prcp_winter2,
          align = "v",
          ncol = 1, nrow = 2)
```

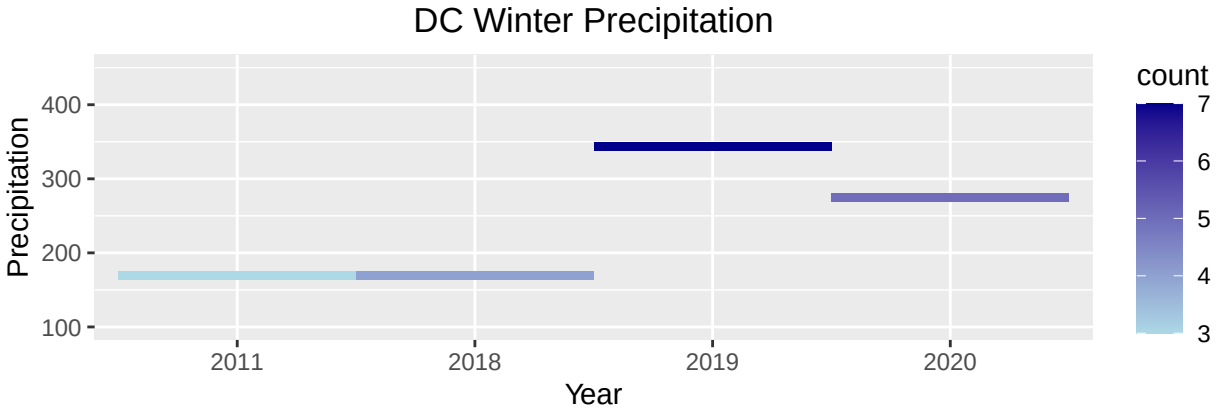



```
DC_prcp_spring <- USA_individual_phenometrics %>%
  filter(State == 'DC') %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Spring, color=as.character(Site_ID))) + geom_po
  labs(title="DC Spring Precipitation",
       x ="Year", y = "Precipitation", color='Site ID') +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450)

DC_prcp_winter <- USA_individual_phenometrics %>%
  filter(State == 'DC') %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Winter, color=as.character(Site_ID))) + geom_po
  labs(title="DC Winter Precipitation",
       x ="Year", y = "Precipitation", color='Site ID') +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450)

ggarrange(DC_prcp_spring, DC_prcp_winter,
          align = "v",
          ncol = 1, nrow = 2)
```




```
DC_prcp_spring2 <- USA_individual_phenometrics %>%
  filter(State == 'DC') %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Spring)) + geom_bin2d() +
  labs(title="DC Spring Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450) +
  scale_fill_continuous(low="lightblue", high = "darkblue")

DC_prcp_winter2 <- USA_individual_phenometrics %>%
  filter(State == 'DC') %>%
  ggplot(aes(x = as.character(First_Yes_Year), y = Prcp_Winter)) + geom_bin2d() +
  labs(title="DC Winter Precipitation",
       x ="Year", y = "Precipitation") +
  theme(plot.title = element_text(hjust = 0.5)) +
  ylim(100, 450) +
  scale_fill_continuous(low="lightblue", high = "darkblue")

ggarrange(DC_prcp_spring2, DC_prcp_winter2,
          align = "v",
          ncol = 1, nrow = 2)
```


DC Spring Precipitation
Precipitation
400
300
200
100
2011
2018
2019
2020
Year
count
4
3
2
1


DC Winter Precipitation
Precipitation
400
300
200
100
2011
2018
2019
2020
Year
count
7
6
5
4
3