Alpine Phenology

M. DePrenger-Levin and RA Hufft

August 28, 2015

If you need these packages install.packages("RCurl")  
install.packages("ggplot2")  
install.packages("data.table") 'install.packages("curl")' install.packages("Rcpp")

# Try a rm(list=ls()) before running

library(RCurl)

## Loading required package: bitops

library(ggplot2)  
library(data.table)  
library(raster)

## Loading required package: sp  
##   
## Attaching package: 'raster'  
##   
## The following object is masked from 'package:data.table':  
##   
## shift

require(rgdal)

## Loading required package: rgdal  
## rgdal: version: 1.1-3, (SVN revision 594)  
## Geospatial Data Abstraction Library extensions to R successfully loaded  
## Loaded GDAL runtime: GDAL 2.0.1, released 2015/09/15  
## Path to GDAL shared files: C:/Users/deprengm/Documents/R/win-library/3.2/rgdal/gdal  
## GDAL does not use iconv for recoding strings.  
## Loaded PROJ.4 runtime: Rel. 4.9.1, 04 March 2015, [PJ\_VERSION: 491]  
## Path to PROJ.4 shared files: C:/Users/deprengm/Documents/R/win-library/3.2/rgdal/proj  
## Linking to sp version: 1.2-1

library(devtools)  
# devtools::install\_github("ropensci/prism")  
library(prism)

Bring csv in from the website I get an error: Error in str.default(obj, ...) : invalid multibyte string 2 but will see if it matters...  
looks like it's the lat long that didn't get in correctly

## plantID institutionCode catalogNumber  
## 1 21862 RM   
## 2 21863 RM 276570  
## 3 21864 CU 197557  
## 4 21865 CU 423185  
## 5 21866 CU 520710  
## 6 21867 RM   
## Scientific.Name  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 3 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 4 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 5 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 6 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## sppListID eventDate earliestBloomDate year\_plant Month Day  
## 1 2 6/11/1998 0:00 1 1998 6 11  
## 2 2 6/13/1963 0:00 1 1963 6 13  
## 3 2 6/13/1963 0:00 0 1963 6 13  
## 4 2 6/13/1979 0:00 1 1979 6 13  
## 5 2 6/15/2001 0:00 1 2001 6 15  
## 6 2 6/15/2004 0:00 1 2004 6 15  
## startDayOfYear reproductiveCondition county  
## 1 162 Phenology: Flowering. Huerfano  
## 2 164 Phenology: Flowering. Larimer  
## 3 164 Phenology: Flowering. Larimer  
## 4 164 Phenology: Flowering and Fruiting. Park  
## 5 166 Phenology: Flowering. Jackson  
## 6 167 Phenology: Flowering. Saguache  
## decimalLatitude decimalLongitude minimumElevationInMeters  
## 1 37.5948 -105.4949 3810  
## 2 3445  
## 3 3444  
## 4 39 12'N 105 25'W 3383  
## 5 3200  
## 6 37.904 -106.5857 3414  
## verbatimElevation ID Year Raw\_Precip Avg\_Hi Med\_Hi Avg\_Lo Med\_Lo Av\_Temp  
## 1 12500 ft. 49 1998 79954 8.65 8.57 -5.55 -5.53 1.55  
## 2 11300 ft. 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 3 11300 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 4 11100 30 1979 81308 7.59 7.57 -7.45 -7.45 0.07  
## 5 10500 52 2001 73220 9.53 9.44 -5.37 -5.34 2.08  
## 6 11200 ft. 55 2004 77968 8.96 8.91 -5.36 -5.36 1.80  
## Med\_Temp GDD  
## 1 1.52 -8.45  
## 2 1.38 -8.59  
## 3 1.38 -8.59  
## 4 0.06 -9.93  
## 5 2.05 -7.92  
## 6 1.78 -8.20

If not working to load from github:

#bloom<-read.csv(path.expand("P:/alpine-phenology/QR\_final\_R\_plant\_climate\_data.csv"))  
  
#Bring in Q/All.Projects\_by\_species/aa\_Spapefiles\_Maps/aa\_QGIS Projects/AlpinePhenologyproject/  
  
#bloom<-read.csv(path.expand("P:/alpine-phenology/TB\_tempPrecipData.csv"),   
# header = TRUE, as.is=TRUE)

# Add growing degree days

White etal 2015  
Cumulative Growing Degree Days   
 should be the lowest air temperature when flowering will occur  
Temperature Threshold and GDD models...

In the csv Tbase is 10, not negative 10. Which is correct?  
 McMaster and Wilhelm 1997:  
Tbase = -10  
Corn   
Winter wheat   
temperature below which the process of interest does not progress

# not really GDD...

Tbase <- -10  
bloom$GDD2 <- (bloom$Avg\_Hi + bloom$Avg\_Lo)/2 - Tbase   
  
#cumsum(bloom$GDD2) # need by day not by species, same data per species

#created new column that divides Raw\_precip by 100 (see PRISM documentation) to get mm  
head(bloom)

## plantID institutionCode catalogNumber  
## 1 21862 RM   
## 2 21863 RM 276570  
## 3 21864 CU 197557  
## 4 21865 CU 423185  
## 5 21866 CU 520710  
## 6 21867 RM   
## Scientific.Name  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 3 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 4 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 5 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 6 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## sppListID eventDate earliestBloomDate year\_plant Month Day  
## 1 2 6/11/1998 0:00 1 1998 6 11  
## 2 2 6/13/1963 0:00 1 1963 6 13  
## 3 2 6/13/1963 0:00 0 1963 6 13  
## 4 2 6/13/1979 0:00 1 1979 6 13  
## 5 2 6/15/2001 0:00 1 2001 6 15  
## 6 2 6/15/2004 0:00 1 2004 6 15  
## startDayOfYear reproductiveCondition county  
## 1 162 Phenology: Flowering. Huerfano  
## 2 164 Phenology: Flowering. Larimer  
## 3 164 Phenology: Flowering. Larimer  
## 4 164 Phenology: Flowering and Fruiting. Park  
## 5 166 Phenology: Flowering. Jackson  
## 6 167 Phenology: Flowering. Saguache  
## decimalLatitude decimalLongitude minimumElevationInMeters  
## 1 37.5948 -105.4949 3810  
## 2 3445  
## 3 3444  
## 4 39 12'N 105 25'W 3383  
## 5 3200  
## 6 37.904 -106.5857 3414  
## verbatimElevation ID Year Raw\_Precip Avg\_Hi Med\_Hi Avg\_Lo Med\_Lo Av\_Temp  
## 1 12500 ft. 49 1998 79954 8.65 8.57 -5.55 -5.53 1.55  
## 2 11300 ft. 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 3 11300 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 4 11100 30 1979 81308 7.59 7.57 -7.45 -7.45 0.07  
## 5 10500 52 2001 73220 9.53 9.44 -5.37 -5.34 2.08  
## 6 11200 ft. 55 2004 77968 8.96 8.91 -5.36 -5.36 1.80  
## Med\_Temp GDD GDD2  
## 1 1.52 -8.45 11.55  
## 2 1.38 -8.59 11.41  
## 3 1.38 -8.59 11.41  
## 4 0.06 -9.93 10.07  
## 5 2.05 -7.92 12.08  
## 6 1.78 -8.20 11.80

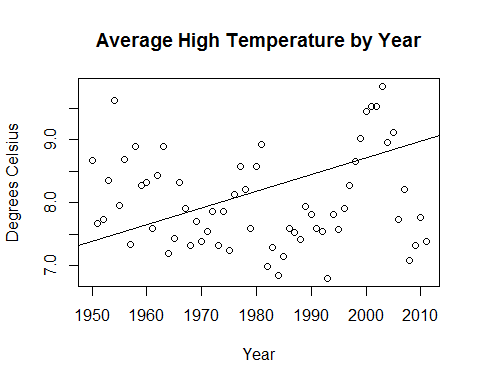
names(bloom)

## [1] "plantID" "institutionCode"   
## [3] "catalogNumber" "Scientific.Name"   
## [5] "sppListID" "eventDate"   
## [7] "earliestBloomDate" "year\_plant"   
## [9] "Month" "Day"   
## [11] "startDayOfYear" "reproductiveCondition"   
## [13] "county" "decimalLatitude"   
## [15] "decimalLongitude" "minimumElevationInMeters"  
## [17] "verbatimElevation" "ID"   
## [19] "Year" "Raw\_Precip"   
## [21] "Avg\_Hi" "Med\_Hi"   
## [23] "Avg\_Lo" "Med\_Lo"   
## [25] "Av\_Temp" "Med\_Temp"   
## [27] "GDD" "GDD2"

str(bloom) #tells me column name, data type and examples, errors in lat long when brought in from web

## 'data.frame': 22549 obs. of 28 variables:  
## $ plantID : int 21862 21863 21864 21865 21866 21867 21868 21869 21870 21873 ...  
## $ institutionCode : Factor w/ 11 levels "ARIZ","ASC","CU",..: 8 8 3 3 3 8 8 3 3 3 ...  
## $ catalogNumber : Factor w/ 13752 levels "","103961","10436",..: 1 2408 1294 3708 6830 1 1 631 4139 1319 ...  
## $ Scientific.Name : Factor w/ 290 levels "Achillea lanulosa Nuttall",..: 2 2 2 2 2 2 2 2 2 2 ...  
## $ sppListID : int 2 2 2 2 2 2 2 2 2 2 ...  
## $ eventDate : Factor w/ 1683 levels "","10/1/1969 0:00",..: 15 19 19 21 29 30 30 32 36 46 ...  
## $ earliestBloomDate : int 1 1 0 1 1 1 0 1 1 0 ...  
## $ year\_plant : int 1998 1963 1963 1979 2001 2004 2004 1961 1987 1963 ...  
## $ Month : int 6 6 6 6 6 6 6 6 6 6 ...  
## $ Day : int 11 13 13 13 15 15 15 16 17 19 ...  
## $ startDayOfYear : int 162 164 164 164 166 167 167 167 168 170 ...  
## $ reproductiveCondition : Factor w/ 14 levels "Flowering","Flowering and Fruiting",..: 13 13 13 12 13 13 13 13 13 13 ...  
## $ county : Factor w/ 76 levels "","Alamosa","Alamosa County",..: 37 46 46 56 38 67 67 6 23 46 ...  
## $ decimalLatitude : Factor w/ 1039 levels "","\"36 59' 45.5\"\" N\"",..: 199 1 1 605 1 335 366 1 1 1 ...  
## $ decimalLongitude : Factor w/ 1079 levels "","-104.5844",..: 85 1 1 1012 1 565 570 1 1 1 ...  
## $ minimumElevationInMeters: int 3810 3445 3444 3383 3200 3414 3762 3353 3500 3475 ...  
## $ verbatimElevation : Factor w/ 1306 levels "","- 11500ft.",..: 931 296 289 218 9 264 860 178 369 323 ...  
## $ ID : int 49 14 14 30 52 55 55 12 38 14 ...  
## $ Year : int 1998 1963 1963 1979 2001 2004 2004 1961 1987 1963 ...  
## $ Raw\_Precip : int 79954 67301 67301 81308 73220 77968 77968 85358 74396 67301 ...  
## $ Avg\_Hi : num 8.65 8.89 8.89 7.59 9.53 8.96 8.96 7.6 7.53 8.89 ...  
## $ Med\_Hi : num 8.57 8.83 8.83 7.57 9.44 8.91 8.91 7.55 7.59 8.83 ...  
## $ Avg\_Lo : num -5.55 -6.07 -6.07 -7.45 -5.37 -5.36 -5.36 -6.83 -7.01 -6.07 ...  
## $ Med\_Lo : num -5.53 -6.07 -6.07 -7.45 -5.34 -5.36 -5.36 -6.84 -7.01 -6.07 ...  
## $ Av\_Temp : num 1.55 1.41 1.41 0.07 2.08 1.8 1.8 0.38 0.26 1.41 ...  
## $ Med\_Temp : num 1.52 1.38 1.38 0.06 2.05 1.78 1.78 0.36 0.29 1.38 ...  
## $ GDD : num -8.45 -8.59 -8.59 -9.93 -7.92 ...  
## $ GDD2 : num 11.6 11.4 11.4 10.1 12.1 ...

#Climate data figures  
#This section runs linear regression on Avg High temp by Year  
#hist(bloom$Avg\_Hi)  
plot(bloom$Year,bloom$Avg\_Hi, main="Average High Temperature by Year",xlab="Year", ylab="Degrees Celsius")  
abline(lm(bloom$Avg\_Hi~bloom$Year))

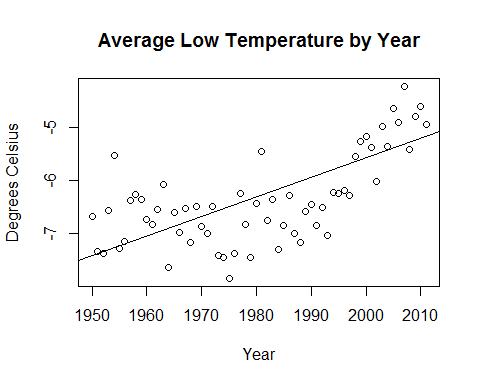


summary(lm(bloom$Avg\_Hi~bloom$Year))

##   
## Call:  
## lm(formula = bloom$Avg\_Hi ~ bloom$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.83744 -0.56444 -0.01312 0.73402 2.11988   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.415e+01 6.932e-01 -63.68 <2e-16 \*\*\*  
## bloom$Year 2.643e-02 3.481e-04 75.94 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.7368 on 22547 degrees of freedom  
## Multiple R-squared: 0.2037, Adjusted R-squared: 0.2036   
## F-statistic: 5767 on 1 and 22547 DF, p-value: < 2.2e-16

#plot(lm(bloom$Avg\_Hi~bloom$Year)) #check residuals for heteroscedasticity and nonlinearity

#This section runs linear regression on Avg Low temp by Year  
#hist(bloom$Avg\_Lo)  
plot(bloom$Year,bloom$Avg\_Lo, main="Average Low Temperature by Year",xlab="Year", ylab="Degrees Celsius")  
abline(lm(bloom$Avg\_Lo~bloom$Year))



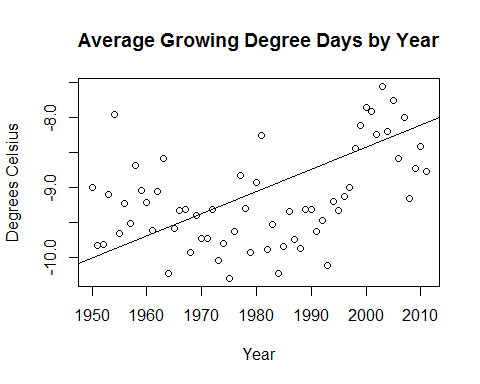
summary(lm(bloom$Avg\_Lo~bloom$Year))

##   
## Call:  
## lm(formula = bloom$Avg\_Lo ~ bloom$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.35073 -0.44165 0.09886 0.35189 1.74572   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.952e+01 4.715e-01 -168.7 <2e-16 \*\*\*  
## bloom$Year 3.697e-02 2.367e-04 156.2 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5011 on 22547 degrees of freedom  
## Multiple R-squared: 0.5197, Adjusted R-squared: 0.5197   
## F-statistic: 2.44e+04 on 1 and 22547 DF, p-value: < 2.2e-16

#plot(lm(bloom$Avg\_Lo~bloom$Year)) #check residuals for heteroscedasticity and nonlinearity

# This section runs linear regression on GDD by Year

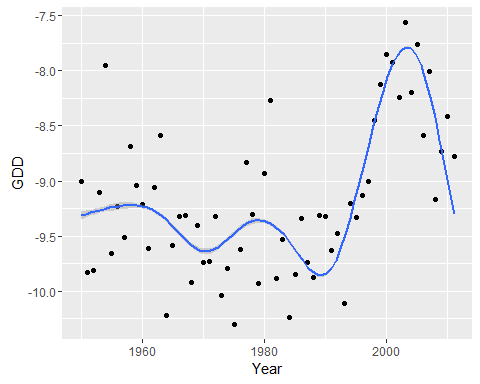
#hist(bloom$GDD)  
#hist(bloom$GDD2) #if Tbase is -10  
plot(bloom$Year,bloom$GDD, main="Average Growing Degree Days by Year",xlab="Year", ylab="Degrees Celsius")  
abline(lm(bloom$GDD~bloom$Year))



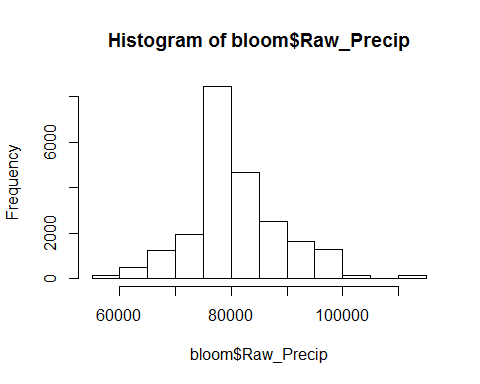
summary(lm(bloom$GDD~bloom$Year))

##   
## Call:  
## lm(formula = bloom$GDD ~ bloom$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.46361 -0.48043 0.04287 0.47776 1.93280   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -7.184e+01 5.499e-01 -130.6 <2e-16 \*\*\*  
## bloom$Year 3.170e-02 2.761e-04 114.8 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.5845 on 22547 degrees of freedom  
## Multiple R-squared: 0.369, Adjusted R-squared: 0.3689   
## F-statistic: 1.318e+04 on 1 and 22547 DF, p-value: < 2.2e-16

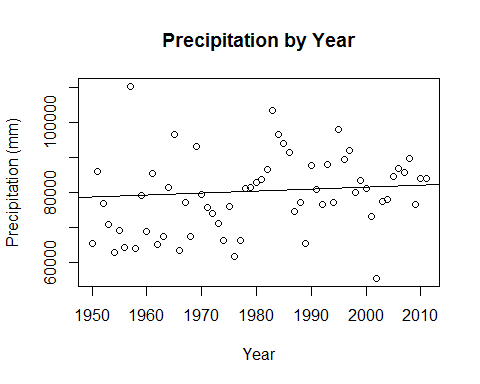
#plot(lm(bloom$Avg\_Lo~bloom$Year)) #check residuals for heteroscedasticity and nonlinearity  
  
  
ggplot(bloom, aes(Year,GDD)) +  
 geom\_point()+  
 stat\_smooth() #uses loess()



#This section runs linear regression on Precipitation by Year  
hist(bloom$Raw\_Precip)



plot(bloom$Year,bloom$Raw\_Precip, main="Precipitation by Year",xlab="Year", ylab="Precipitation (mm)")  
abline(lm(bloom$Raw\_Precip~bloom$Year))



summary(lm(bloom$Raw\_Precip~bloom$Year))

##   
## Call:  
## lm(formula = bloom$Raw\_Precip ~ bloom$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -26175 -4180 -1443 3834 31100   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -29090.211 7602.753 -3.826 0.00013 \*\*\*  
## bloom$Year 55.299 3.817 14.487 < 2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 8080 on 22547 degrees of freedom  
## Multiple R-squared: 0.009222, Adjusted R-squared: 0.009178   
## F-statistic: 209.9 on 1 and 22547 DF, p-value: < 2.2e-16

#plot(lm(bloom$Raw\_Precip ~bloom$Year)) #check residuals for heteroscedasticity and nonlinearity

# Graph just the significant spp by year to see how many days earlier for these species that are showing a signal by year

# re-run code that runs regressions on each species to create table to subset from

# Limit all data to after 1949, 7220 obs. from table (7217 obs. from query in Access)

bl50 <- subset(bloom, bloom$year\_plant > 1949) #Our dataset already has this subset but re-running as a double check  
  
nrow(bl50)

## [1] 22549

nrow(subset(bloom, bloom$year\_plant > 1949))

## [1] 22549

str(bl50)

## 'data.frame': 22549 obs. of 28 variables:  
## $ plantID : int 21862 21863 21864 21865 21866 21867 21868 21869 21870 21873 ...  
## $ institutionCode : Factor w/ 11 levels "ARIZ","ASC","CU",..: 8 8 3 3 3 8 8 3 3 3 ...  
## $ catalogNumber : Factor w/ 13752 levels "","103961","10436",..: 1 2408 1294 3708 6830 1 1 631 4139 1319 ...  
## $ Scientific.Name : Factor w/ 290 levels "Achillea lanulosa Nuttall",..: 2 2 2 2 2 2 2 2 2 2 ...  
## $ sppListID : int 2 2 2 2 2 2 2 2 2 2 ...  
## $ eventDate : Factor w/ 1683 levels "","10/1/1969 0:00",..: 15 19 19 21 29 30 30 32 36 46 ...  
## $ earliestBloomDate : int 1 1 0 1 1 1 0 1 1 0 ...  
## $ year\_plant : int 1998 1963 1963 1979 2001 2004 2004 1961 1987 1963 ...  
## $ Month : int 6 6 6 6 6 6 6 6 6 6 ...  
## $ Day : int 11 13 13 13 15 15 15 16 17 19 ...  
## $ startDayOfYear : int 162 164 164 164 166 167 167 167 168 170 ...  
## $ reproductiveCondition : Factor w/ 14 levels "Flowering","Flowering and Fruiting",..: 13 13 13 12 13 13 13 13 13 13 ...  
## $ county : Factor w/ 76 levels "","Alamosa","Alamosa County",..: 37 46 46 56 38 67 67 6 23 46 ...  
## $ decimalLatitude : Factor w/ 1039 levels "","\"36 59' 45.5\"\" N\"",..: 199 1 1 605 1 335 366 1 1 1 ...  
## $ decimalLongitude : Factor w/ 1079 levels "","-104.5844",..: 85 1 1 1012 1 565 570 1 1 1 ...  
## $ minimumElevationInMeters: int 3810 3445 3444 3383 3200 3414 3762 3353 3500 3475 ...  
## $ verbatimElevation : Factor w/ 1306 levels "","- 11500ft.",..: 931 296 289 218 9 264 860 178 369 323 ...  
## $ ID : int 49 14 14 30 52 55 55 12 38 14 ...  
## $ Year : int 1998 1963 1963 1979 2001 2004 2004 1961 1987 1963 ...  
## $ Raw\_Precip : int 79954 67301 67301 81308 73220 77968 77968 85358 74396 67301 ...  
## $ Avg\_Hi : num 8.65 8.89 8.89 7.59 9.53 8.96 8.96 7.6 7.53 8.89 ...  
## $ Med\_Hi : num 8.57 8.83 8.83 7.57 9.44 8.91 8.91 7.55 7.59 8.83 ...  
## $ Avg\_Lo : num -5.55 -6.07 -6.07 -7.45 -5.37 -5.36 -5.36 -6.83 -7.01 -6.07 ...  
## $ Med\_Lo : num -5.53 -6.07 -6.07 -7.45 -5.34 -5.36 -5.36 -6.84 -7.01 -6.07 ...  
## $ Av\_Temp : num 1.55 1.41 1.41 0.07 2.08 1.8 1.8 0.38 0.26 1.41 ...  
## $ Med\_Temp : num 1.52 1.38 1.38 0.06 2.05 1.78 1.78 0.36 0.29 1.38 ...  
## $ GDD : num -8.45 -8.59 -8.59 -9.93 -7.92 ...  
## $ GDD2 : num 11.6 11.4 11.4 10.1 12.1 ...

head(bl50)

## plantID institutionCode catalogNumber  
## 1 21862 RM   
## 2 21863 RM 276570  
## 3 21864 CU 197557  
## 4 21865 CU 423185  
## 5 21866 CU 520710  
## 6 21867 RM   
## Scientific.Name  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 3 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 4 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 5 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 6 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## sppListID eventDate earliestBloomDate year\_plant Month Day  
## 1 2 6/11/1998 0:00 1 1998 6 11  
## 2 2 6/13/1963 0:00 1 1963 6 13  
## 3 2 6/13/1963 0:00 0 1963 6 13  
## 4 2 6/13/1979 0:00 1 1979 6 13  
## 5 2 6/15/2001 0:00 1 2001 6 15  
## 6 2 6/15/2004 0:00 1 2004 6 15  
## startDayOfYear reproductiveCondition county  
## 1 162 Phenology: Flowering. Huerfano  
## 2 164 Phenology: Flowering. Larimer  
## 3 164 Phenology: Flowering. Larimer  
## 4 164 Phenology: Flowering and Fruiting. Park  
## 5 166 Phenology: Flowering. Jackson  
## 6 167 Phenology: Flowering. Saguache  
## decimalLatitude decimalLongitude minimumElevationInMeters  
## 1 37.5948 -105.4949 3810  
## 2 3445  
## 3 3444  
## 4 39 12'N 105 25'W 3383  
## 5 3200  
## 6 37.904 -106.5857 3414  
## verbatimElevation ID Year Raw\_Precip Avg\_Hi Med\_Hi Avg\_Lo Med\_Lo Av\_Temp  
## 1 12500 ft. 49 1998 79954 8.65 8.57 -5.55 -5.53 1.55  
## 2 11300 ft. 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 3 11300 14 1963 67301 8.89 8.83 -6.07 -6.07 1.41  
## 4 11100 30 1979 81308 7.59 7.57 -7.45 -7.45 0.07  
## 5 10500 52 2001 73220 9.53 9.44 -5.37 -5.34 2.08  
## 6 11200 ft. 55 2004 77968 8.96 8.91 -5.36 -5.36 1.80  
## Med\_Temp GDD GDD2  
## 1 1.52 -8.45 11.55  
## 2 1.38 -8.59 11.41  
## 3 1.38 -8.59 11.41  
## 4 0.06 -9.93 10.07  
## 5 2.05 -7.92 12.08  
## 6 1.78 -8.20 11.80

names <- unique(bl50$Scientific.Name) # Names includes synonyms, sspListID contains unique taxa  
 namesID <- unique(bl50$sppListID)  
length(namesID) # 290

## [1] 290

head(namesID)

## [1] 2 1 3 4 5 6

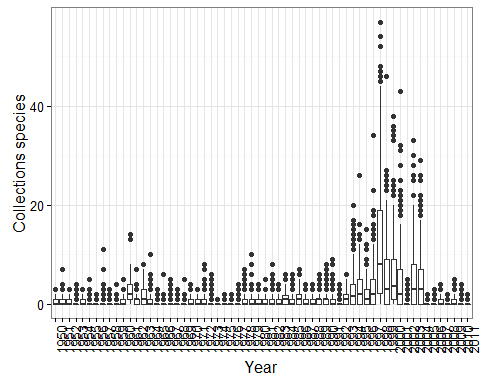
names(bl50)

## [1] "plantID" "institutionCode"   
## [3] "catalogNumber" "Scientific.Name"   
## [5] "sppListID" "eventDate"   
## [7] "earliestBloomDate" "year\_plant"   
## [9] "Month" "Day"   
## [11] "startDayOfYear" "reproductiveCondition"   
## [13] "county" "decimalLatitude"   
## [15] "decimalLongitude" "minimumElevationInMeters"  
## [17] "verbatimElevation" "ID"   
## [19] "Year" "Raw\_Precip"   
## [21] "Avg\_Hi" "Med\_Hi"   
## [23] "Avg\_Lo" "Med\_Lo"   
## [25] "Av\_Temp" "Med\_Temp"   
## [27] "GDD" "GDD2"

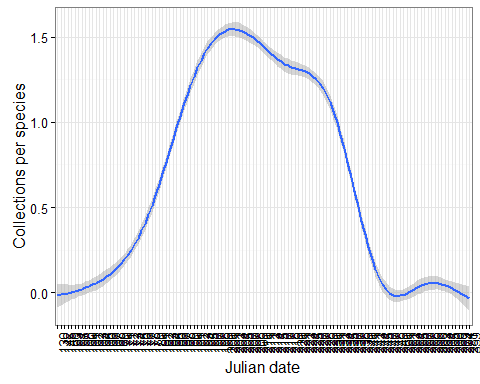
#There are multiple bloom dates per year for many species. Want to narrow down to one, earliest, per year  
head(table(bl50$sppListID, bl50$Year))

##   
## 1950 1951 1952 1953 1954 1955 1956 1957 1958 1959 1960 1961 1962 1963  
## 1 0 1 0 1 0 0 0 0 0 0 0 2 1 3  
## 2 0 0 1 0 2 0 0 4 0 4 0 7 1 4  
## 3 1 0 0 0 2 0 0 0 1 0 1 2 2 0  
## 4 1 0 1 0 0 0 0 0 0 0 0 4 0 2  
## 5 0 0 0 0 0 0 0 0 0 0 0 1 0 1  
## 6 0 0 0 0 0 0 0 0 0 0 0 4 1 1  
##   
## 1964 1965 1966 1967 1968 1969 1970 1971 1972 1973 1974 1975 1976 1977  
## 1 0 0 0 0 1 0 1 1 1 0 0 0 0 0  
## 2 2 1 0 1 2 2 1 1 1 0 0 0 0 1  
## 3 0 1 0 1 0 0 1 0 1 0 0 0 0 0  
## 4 0 0 0 0 1 0 0 1 0 0 0 0 0 2  
## 5 0 0 0 0 0 0 0 2 0 0 0 0 0 0  
## 6 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## 1978 1979 1980 1981 1982 1983 1984 1985 1986 1987 1988 1989 1990 1991  
## 1 2 0 0 0 2 0 2 0 0 0 0 2 2 0  
## 2 5 1 0 5 2 1 0 3 2 3 3 6 7 8  
## 3 0 1 0 0 0 0 0 0 1 0 0 0 0 0  
## 4 0 0 0 1 1 1 0 3 1 0 0 1 0 1  
## 5 0 0 1 0 0 0 0 2 1 2 0 0 1 0  
## 6 1 0 0 0 1 0 2 0 0 0 0 0 0 0  
##   
## 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005  
## 1 0 2 15 6 8 8 40 11 33 19 0 19 19 0  
## 2 0 5 11 26 9 23 52 46 34 43 4 25 22 0  
## 3 0 0 0 1 0 0 2 2 0 1 0 0 3 0  
## 4 0 1 7 4 2 4 32 17 9 10 0 6 7 0  
## 5 0 2 6 6 5 2 10 7 10 9 0 6 7 0  
## 6 0 2 1 0 1 0 13 5 2 4 0 3 5 1  
##   
## 2006 2007 2008 2009 2010 2011  
## 1 0 0 0 1 0 0  
## 2 0 2 0 1 1 1  
## 3 0 0 1 1 0 0  
## 4 0 1 0 0 0 1  
## 5 0 0 0 0 0 0  
## 6 0 0 0 0 0 0

collections<-data.frame(table(bl50$sppListID,bl50$Year))  
  
ggplot(collections, aes(Var2, Freq))+  
 geom\_boxplot()+  
 stat\_smooth()+  
 xlab("Year")+  
 ylab("Collections species")+  
 theme\_bw()+  
 theme(axis.text.x = element\_text(angle=90, hjust=1))



# collections per month  
collJ <- data.frame(table(bl50$sppListID, bl50$startDayOfYear))  
  
ggplot(collJ, aes(Var2, Freq, group =1))+  
 stat\_smooth()+  
 xlab("Julian date")+  
 ylab("Collections per species")+  
 theme\_bw()+  
 theme(axis.text.x = element\_text(angle=90, hjust=1))



# collections getting

Don't worry about bloom vs. bl50, they are identical... and other checks

#bl50 and bloom are the same  
identical(bl50, bloom)

## [1] TRUE

#There should be 290 species  
length(unique(bl50$sppListID))

## [1] 290

#Only need some columns of bloom:bl50 for minimum julian dates  
names(bl50[,c(4:5,11,19:28)])

## [1] "Scientific.Name" "sppListID" "startDayOfYear"   
## [4] "Year" "Raw\_Precip" "Avg\_Hi"   
## [7] "Med\_Hi" "Avg\_Lo" "Med\_Lo"   
## [10] "Av\_Temp" "Med\_Temp" "GDD"   
## [13] "GDD2"

Pull out the earliest bloom date per year per species ## Break up the function into gathering the earliest bloom day per year seperate from regression per species ## Earliest bloom date held in min.dates Aggregate instead of apply? Much faster and more corrector!

minJul.dates <- lapply(unique(bl50$sppListID), function(x){  
 this.species <- bl50[bl50$sppListID == x,c(4:5,11,19:28)]  
 these.rows <- aggregate(startDayOfYear~Year, this.species, FUN=min)  
 unique(merge(this.species, these.rows, by = c("startDayOfYear","Year")))  
 })  
  
names(minJul.dates) <- unique(bl50$sppListID)  
  
head(minJul.dates[[219]])

## startDayOfYear Year Scientific.Name sppListID  
## 1 180 1998 Ligularia amplectens (A. Gray) W. A. Weber 155  
## 2 192 1950 Ligularia amplectens (A. Gray) W. A. Weber 155  
## 3 196 2002 Ligularia amplectens (A. Gray) W. A. Weber 155  
## 4 197 1972 Ligularia amplectens (A. Gray) W. A. Weber 155  
## 5 204 1986 Ligularia amplectens (A. Gray) W. A. Weber 155  
## 6 204 1992 Ligularia amplectens (A. Gray) W. A. Weber 155  
## Raw\_Precip Avg\_Hi Med\_Hi Avg\_Lo Med\_Lo Av\_Temp Med\_Temp GDD GDD2  
## 1 79954 8.65 8.57 -5.55 -5.53 1.55 1.52 -8.45 11.55  
## 2 65511 8.68 8.68 -6.68 -6.61 1.00 1.04 -9.00 11.00  
## 3 55443 9.53 9.46 -6.01 -6.07 1.76 1.70 -8.24 11.76  
## 4 74014 7.86 7.86 -6.50 -6.51 0.68 0.68 -9.32 10.68  
## 5 91333 7.60 7.65 -6.28 -6.31 0.66 0.67 -9.34 10.66  
## 6 76457 7.55 7.47 -6.51 -6.58 0.52 0.45 -9.48 10.52

min.dates <- do.call(rbind,minJul.dates)  
  
head(table(min.dates$Year, min.dates$sppListID))

##   
## 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25  
## 1950 0 0 1 1 0 0 0 1 0 1 0 0 0 1 0 1 0 0 1 1 1 0 0 1 0  
## 1951 1 0 0 0 0 0 0 1 0 1 1 1 1 0 1 0 0 0 0 0 0 1 0 1 1  
## 1952 0 1 0 1 0 0 0 0 0 0 1 1 0 1 0 0 0 0 1 1 0 1 0 1 0  
## 1953 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 0 1 1 0 0 0 1 1 0 0 1 0 1 0 0 0 0 0 0 1 0 0 1 1 0  
## 1955 0 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0  
##   
## 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47  
## 1950 0 1 1 0 0 0 0 0 1 0 1 0 1 1 1 0 0 0 0 0 0 0  
## 1951 1 0 0 1 1 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1  
## 1952 0 0 1 0 1 0 0 1 1 0 0 0 0 0 0 1 0 0 0 0 0 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 1 0  
## 1954 0 0 0 0 0 0 0 1 0 0 0 0 1 0 1 0 1 1 0 0 1 1  
## 1955 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 0 1 0 0  
##   
## 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69  
## 1950 0 0 0 0 0 0 0 0 0 0 0 1 1 1 1 0 0 1 1 0 0 0  
## 1951 0 1 1 0 0 0 0 0 0 1 0 1 0 0 1 0 0 0 0 0 0 0  
## 1952 0 1 0 0 0 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0 1 1  
## 1953 0 1 1 0 0 0 0 1 1 1 0 0 0 0 1 0 0 1 0 0 0 0  
## 1954 1 1 0 1 0 1 0 1 1 1 0 1 1 1 1 1 0 1 1 1 0 1  
## 1955 1 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 1 0 0  
##   
## 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91  
## 1950 1 1 1 0 1 0 1 0 1 1 0 1 1 0 1 1 0 0 0 1 0 0  
## 1951 1 1 1 0 0 1 1 0 0 1 0 0 0 0 1 0 0 0 0 0 0 0  
## 1952 0 0 1 1 1 1 0 1 0 0 0 0 1 0 1 1 0 1 0 1 0 0  
## 1953 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 1 0 0 1 0  
## 1955 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## 92 93 94 95 96 97 98 99 100 101 104 105 106 107 108 109 110 111 112  
## 1950 0 0 1 1 0 1 1 0 0 0 1 0 1 0 0 0 0 0 0  
## 1951 1 1 1 1 0 1 0 1 1 0 1 0 0 1 0 1 0 1 1  
## 1952 0 0 0 0 0 0 0 0 0 0 0 1 0 0 1 0 1 0 0  
## 1953 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0 0 0  
## 1955 0 0 1 1 0 0 0 1 0 0 0 0 0 1 0 0 0 0 0  
##   
## 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129  
## 1950 1 1 0 0 1 1 0 0 1 0 0 1 0 1 1 0 0  
## 1951 0 0 0 1 1 0 0 1 1 1 1 1 0 1 1 0 0  
## 1952 1 0 0 1 0 0 1 0 0 0 0 0 0 0 1 1 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 0 1 1 1 0 0 0 0 1 0 0 0 0 0 1 0  
## 1955 0 0 1 0 1 0 0 0 0 0 0 0 0 0 1 0 1  
##   
## 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 147  
## 1950 0 1 0 0 0 1 1 1 0 1 1 0 0 1 1 1 1  
## 1951 0 0 0 0 0 0 0 0 0 0 0 0 1 0 1 1 0  
## 1952 0 0 0 0 1 0 1 0 1 0 0 0 0 1 0 1 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 1 1 0 0 0 0 1 0 0 1 0 0 1 0 1 1  
## 1955 0 1 0 0 0 0 0 0 0 0 0 0 0 0 1 0 0  
##   
## 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164  
## 1950 1 1 0 0 0 0 1 1 0 0 1 0 0 0 1 0 0  
## 1951 1 0 0 1 1 0 0 0 1 1 1 0 0 1 0 1 0  
## 1952 0 0 1 0 0 0 0 0 1 0 1 0 0 0 1 0 0  
## 1953 0 1 0 0 1 1 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 1 0 0 1 1 0 0 0 0 0 0 1 0 0 1 0  
## 1955 0 0 0 0 0 0 1 0 0 1 0 0 0 1 0 0 0  
##   
## 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181  
## 1950 0 0 0 0 0 1 0 0 1 0 1 1 0 0 0 0 0  
## 1951 0 0 0 0 1 1 0 1 0 0 0 0 1 1 0 1 1  
## 1952 0 1 0 0 0 1 1 1 1 0 0 1 0 0 0 0 0  
## 1953 1 0 0 0 0 0 0 0 0 0 0 0 1 0 0 0 0  
## 1954 0 1 0 0 0 0 0 0 0 1 0 1 1 0 1 0 0  
## 1955 0 0 0 0 0 0 0 0 0 1 0 0 1 0 0 0 0  
##   
## 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198  
## 1950 0 1 0 0 1 0 1 0 1 1 1 0 0 1 1 0 1  
## 1951 0 0 1 0 1 1 1 0 1 1 1 0 0 0 0 0 0  
## 1952 0 1 0 0 0 0 0 1 0 0 0 1 0 1 1 0 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 1955 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215  
## 1950 0 1 1 0 1 0 0 1 0 1 1 0 0 0 0 1 1  
## 1951 0 0 1 1 0 0 0 0 1 1 0 0 0 0 0 0 0  
## 1952 1 1 0 1 1 0 0 1 0 0 0 0 0 0 0 0 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0  
## 1954 0 0 0 1 0 1 0 1 0 0 1 0 0 0 0 0 1  
## 1955 0 0 0 1 0 0 0 0 0 0 0 1 0 0 0 0 0  
##   
## 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232  
## 1950 1 0 1 0 0 1 1 1 1 0 1 0 0 1 0 1 1  
## 1951 1 0 0 0 0 0 0 0 1 0 0 0 1 0 0 0 0  
## 1952 0 0 0 0 1 0 0 0 0 0 1 0 0 0 1 0 0  
## 1953 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 0 1 0 0 0 0 1 1 0 0 0 0 0 0 0 0  
## 1955 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249  
## 1950 1 0 0 0 0 1 1 0 0 1 0 0 0 0 0 0 0  
## 1951 1 1 1 0 1 0 1 0 0 1 0 1 0 1 1 0 0  
## 1952 0 0 0 1 1 0 1 0 0 1 0 0 1 0 1 0 0  
## 1953 0 0 0 0 0 0 1 0 1 0 0 0 0 0 0 0 0  
## 1954 0 0 1 1 0 0 0 1 1 0 1 0 1 0 0 0 0  
## 1955 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
##   
## 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266  
## 1950 0 0 1 1 0 0 0 0 0 0 0 1 1 1 0 1 1  
## 1951 1 1 1 0 1 1 1 1 1 0 0 0 0 0 1 1 0  
## 1952 1 0 0 1 0 0 0 0 0 0 0 0 1 0 1 1 0  
## 1953 0 0 1 0 0 0 0 0 0 1 0 0 0 0 0 0 0  
## 1954 0 1 1 1 0 1 0 0 1 0 0 0 1 0 0 0 0  
## 1955 0 0 0 0 0 0 0 0 0 1 0 0 0 0 0 0 0  
##   
## 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 284  
## 1950 1 0 0 0 0 0 1 0 0 1 1 1 1 0 0 1 0  
## 1951 1 0 0 0 0 1 0 1 0 1 1 1 1 0 1 1 0  
## 1952 0 0 0 0 1 0 1 0 0 0 0 1 0 0 0 0 0  
## 1953 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1954 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0  
## 1955 0 0 0 1 0 0 0 1 0 0 0 0 1 0 0 0 0  
##   
## 285 286 287 288 289 290 291 292 293 294  
## 1950 0 0 0 0 0 0 1 1 1 0  
## 1951 0 0 1 0 0 0 0 0 1 0  
## 1952 0 0 0 0 0 0 1 1 1 1  
## 1953 0 0 0 0 0 0 0 0 0 0  
## 1954 1 1 1 0 1 0 0 1 1 0  
## 1955 0 0 0 1 0 0 0 0 0 0

# Check that there is only one minimum date per year

minpyear <- table(min.dates$Year, min.dates$sppListID)  
mpy.df <- data.frame(minpyear)  
mpy.df[mpy.df$Freq > 1,] #when we have ealiestBloomDate in the lapply, that's not unique so it picks a few duplicates, now fixed

## [1] Var1 Var2 Freq  
## <0 rows> (or 0-length row.names)

library(psych)

##   
## Attaching package: 'psych'  
##   
## The following objects are masked from 'package:ggplot2':  
##   
## %+%, alpha

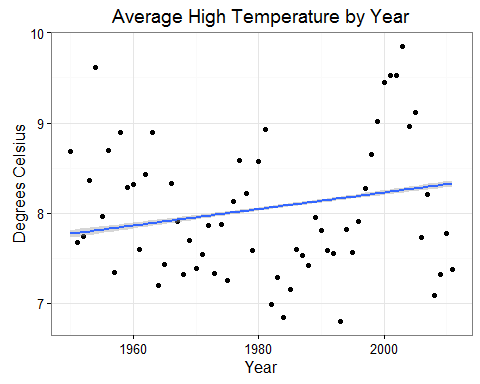
#cor.plot(cor(min.dates[,-(1:4)]), numbers = TRUE)  
# should keep Raw\_Precip and Avg\_Lo  
  
cor(min.dates[,-(1:4)])

## Raw\_Precip Avg\_Hi Med\_Hi Avg\_Lo Med\_Lo  
## Raw\_Precip 1.00000000 -0.4594490 -0.4736374 -0.03219269 -0.04271174  
## Avg\_Hi -0.45944900 1.0000000 0.9981899 0.63539351 0.64221543  
## Med\_Hi -0.47363739 0.9981899 1.0000000 0.61730201 0.62483374  
## Avg\_Lo -0.03219269 0.6353935 0.6173020 1.00000000 0.99864645  
## Med\_Lo -0.04271174 0.6422154 0.6248337 0.99864645 1.00000000  
## Av\_Temp -0.27182137 0.9045386 0.8935072 0.90397288 0.90704069  
## Med\_Temp -0.28261853 0.9068056 0.8979689 0.89977060 0.90464990  
## GDD -0.27201374 0.9044092 0.8934113 0.90412246 0.90714932  
## GDD2 -0.27201374 0.9044092 0.8934113 0.90412246 0.90714932  
## Av\_Temp Med\_Temp GDD GDD2  
## Raw\_Precip -0.2718214 -0.2826185 -0.2720137 -0.2720137  
## Avg\_Hi 0.9045386 0.9068056 0.9044092 0.9044092  
## Med\_Hi 0.8935072 0.8979689 0.8934113 0.8934113  
## Avg\_Lo 0.9039729 0.8997706 0.9041225 0.9041225  
## Med\_Lo 0.9070407 0.9046499 0.9071493 0.9071493  
## Av\_Temp 1.0000000 0.9989101 0.9999889 0.9999889  
## Med\_Temp 0.9989101 1.0000000 0.9989214 0.9989214  
## GDD 0.9999889 0.9989214 1.0000000 1.0000000  
## GDD2 0.9999889 0.9989214 1.0000000 1.0000000

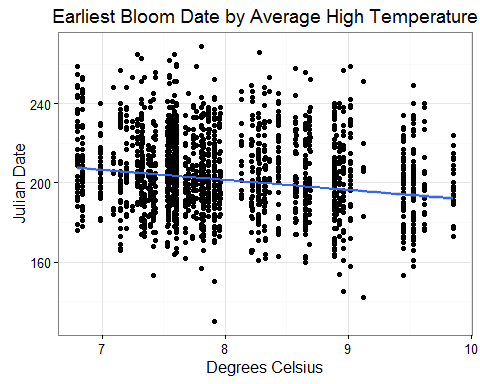
#getting a ggplot error of "Error in .Call.graphics(C\_palette2, .Call(C\_palette2, NULL)) :   
# invalid graphics state"  
  
#fix error by running dev.off()  
# dev.off()

ggplot option after subset to earliest observation per year per species  
Avg high temperature

ggplot(min.dates, aes(Year,Avg\_Hi))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Average High Temperature by Year")+  
 xlab("Year")+  
 ylab("Degrees Celsius")

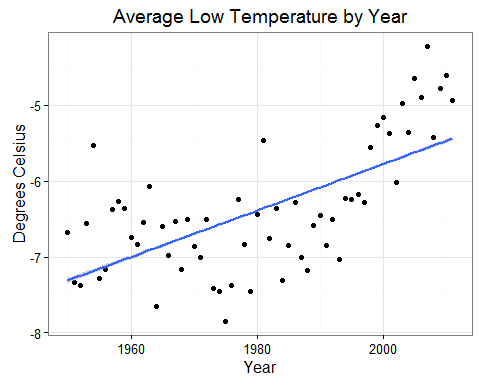


ggplot(min.dates, aes(Avg\_Hi, startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Average High Temperature") +  
 xlab("Degrees Celsius")+  
 ylab("Julian Date")

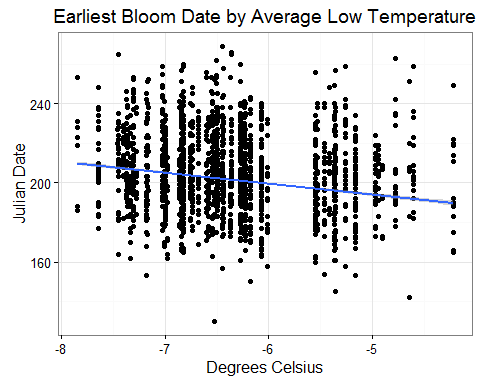


Average low temperature

ggplot(min.dates, aes(Year,Avg\_Lo))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Average Low Temperature by Year")+  
 xlab("Year")+  
 ylab("Degrees Celsius")



ggplot(min.dates, aes(Avg\_Lo, startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Average Low Temperature") +  
 xlab("Degrees Celsius")+  
 ylab("Julian Date")

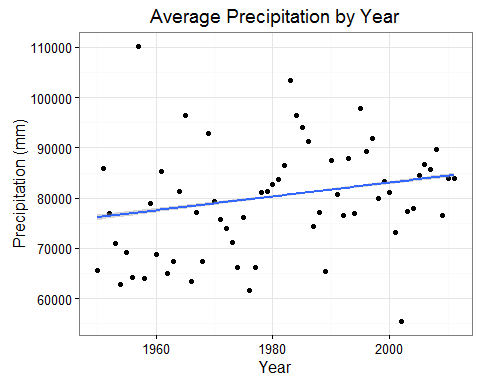


summary(lm(startDayOfYear~Avg\_Lo, data = min.dates))

##   
## Call:  
## lm(formula = startDayOfYear ~ Avg\_Lo, data = min.dates)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -72.491 -13.084 -1.553 12.109 70.164   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 166.4639 1.7387 95.74 <2e-16 \*\*\*  
## Avg\_Lo -5.5171 0.2758 -20.00 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.12 on 6790 degrees of freedom  
## Multiple R-squared: 0.05565, Adjusted R-squared: 0.05551   
## F-statistic: 400.1 on 1 and 6790 DF, p-value: < 2.2e-16

Average precipitation

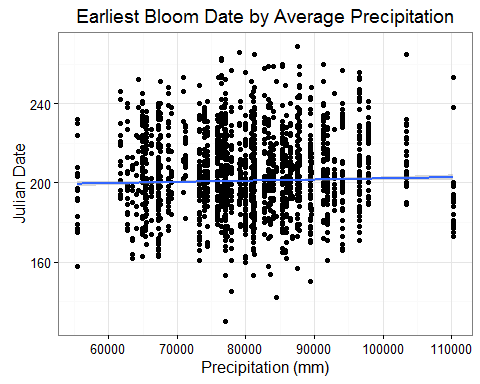
ggplot(min.dates, aes(Year,Raw\_Precip))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Average Precipitation by Year")+  
 xlab("Year")+  
 ylab("Precipitation (mm)")



summary(lm(Raw\_Precip~Year, data = min.dates))

##   
## Call:  
## lm(formula = Raw\_Precip ~ Year, data = min.dates)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -27894 -6215 -597 5919 33098   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -1.927e+05 1.407e+04 -13.70 <2e-16 \*\*\*  
## Year 1.379e+02 7.088e+00 19.46 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9740 on 6790 degrees of freedom  
## Multiple R-squared: 0.0528, Adjusted R-squared: 0.05266   
## F-statistic: 378.5 on 1 and 6790 DF, p-value: < 2.2e-16

ggplot(min.dates, aes(Raw\_Precip, startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Average Precipitation") +  
 xlab("Precipitation (mm)")+  
 ylab("Julian Date")



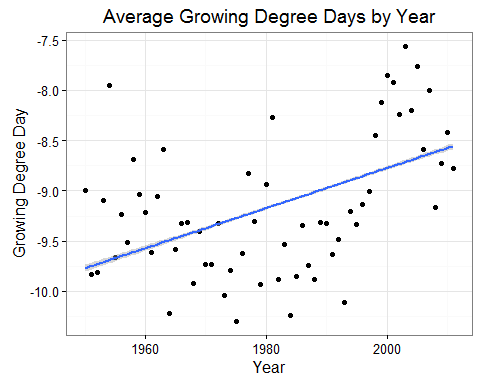
summary(lm(startDayOfYear~Raw\_Precip, data = min.dates))

##   
## Call:  
## lm(formula = startDayOfYear ~ Raw\_Precip, data = min.dates)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -70.734 -12.731 -2.146 12.045 67.641   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.961e+02 1.842e+00 106.482 < 2e-16 \*\*\*  
## Raw\_Precip 5.961e-05 2.259e-05 2.638 0.00835 \*\*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.63 on 6790 degrees of freedom  
## Multiple R-squared: 0.001024, Adjusted R-squared: 0.0008771   
## F-statistic: 6.961 on 1 and 6790 DF, p-value: 0.008348

## not real GDD!

GDD

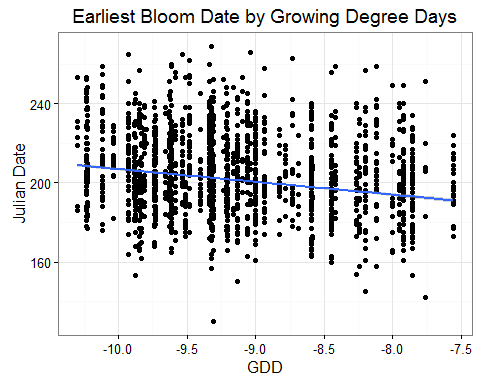
ggplot(min.dates, aes(Year,GDD))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Average Growing Degree Days by Year")+  
 xlab("Year")+  
 ylab("Growing Degree Day")



summary(lm(GDD~Year, data = min.dates))

##   
## Call:  
## lm(formula = GDD ~ Year, data = min.dates)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -1.20219 -0.42298 -0.09338 0.49368 1.73472   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -4.861e+01 9.245e-01 -52.58 <2e-16 \*\*\*  
## Year 1.992e-02 4.659e-04 42.76 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.6402 on 6790 degrees of freedom  
## Multiple R-squared: 0.2121, Adjusted R-squared: 0.212   
## F-statistic: 1828 on 1 and 6790 DF, p-value: < 2.2e-16

ggplot(min.dates, aes(GDD, startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Growing Degree Days") +  
 xlab("GDD")+  
 ylab("Julian Date")

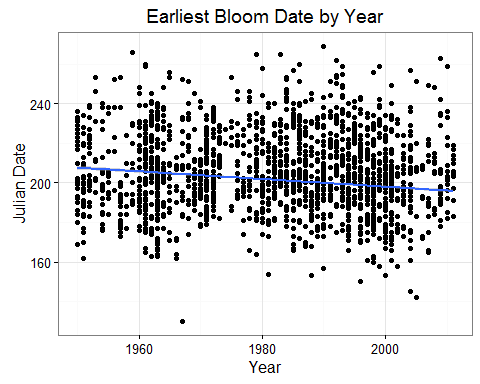


summary(lm(startDayOfYear~GDD, data = min.dates))

##   
## Call:  
## lm(formula = startDayOfYear ~ GDD, data = min.dates)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -72.422 -13.744 -1.487 11.864 66.513   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 142.2879 2.7675 51.41 <2e-16 \*\*\*  
## GDD -6.4591 0.3037 -21.27 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.05 on 6790 degrees of freedom  
## Multiple R-squared: 0.06246, Adjusted R-squared: 0.06232   
## F-statistic: 452.3 on 1 and 6790 DF, p-value: < 2.2e-16

### Minimum Julian bloom date by year

ggplot(min.dates, aes(Year,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Year")+  
 xlab("Year")+  
 ylab("Julian Date")



summary(lm(min.dates$startDayOfYear~min.dates$Year))

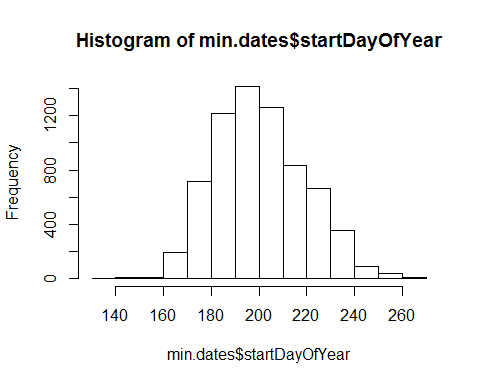
##   
## Call:  
## lm(formula = min.dates$startDayOfYear ~ min.dates$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -74.330 -13.985 -1.363 12.669 69.121   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 584.98455 26.51355 22.06 <2e-16 \*\*\*  
## min.dates$Year -0.19352 0.01336 -14.48 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.36 on 6790 degrees of freedom  
## Multiple R-squared: 0.02997, Adjusted R-squared: 0.02983   
## F-statistic: 209.8 on 1 and 6790 DF, p-value: < 2.2e-16

max(min.dates$startDayOfYear)

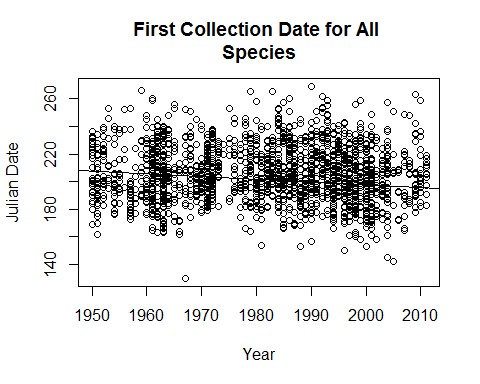
## [1] 269

Graph of all spp data by year using plot instead of ggplot

#bloom<-read.csv(path.expand("P:/alpine-phenology/QR\_final\_R\_plant\_climate\_data.csv"),   
# header = TRUE, as.is=TRUE)   
hist(min.dates$startDayOfYear)#need to log transform data



plot(min.dates$Year,min.dates$startDayOfYear, main="First Collection Date for All   
Species",xlab="Year", ylab="Julian Date")  
abline(lm(min.dates$startDayOfYear~min.dates$Year))



summary(lm(min.dates$startDayOfYear~min.dates$Year))

##   
## Call:  
## lm(formula = min.dates$startDayOfYear ~ min.dates$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -74.330 -13.985 -1.363 12.669 69.121   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 584.98455 26.51355 22.06 <2e-16 \*\*\*  
## min.dates$Year -0.19352 0.01336 -14.48 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.36 on 6790 degrees of freedom  
## Multiple R-squared: 0.02997, Adjusted R-squared: 0.02983   
## F-statistic: 209.8 on 1 and 6790 DF, p-value: < 2.2e-16

#both intercept and slope significant. If include all data, the date of first collection is statistically earlier by 0.062563/year, or 3.75 days earlier over the course of the study period.   
#plot(lm(bloom$startDayOfYear ~bloom$Year)) #check residuals for heteroscedasticity and nonlinearity

# Regression per species

# updated regression loop to match pulling minimum dates above

The variable can be set to which variable you want to use. Choose from "Year", "Raw\_Precip", "Avg\_Hi", or "Avg\_Lo"...

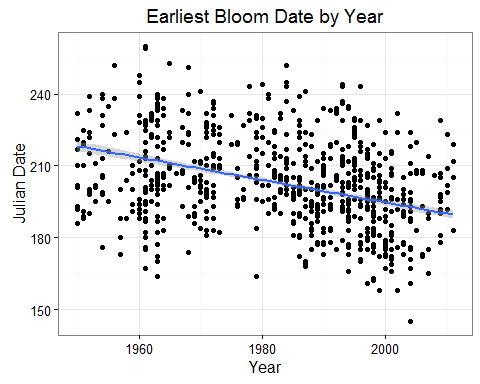
JulBloomDate <- function(variable){  
 namesID <- unique(min.dates$sppListID)  
   
 #Which column number has the variable of interest  
 colnum <- match(variable,names(min.dates))  
   
 JYC <- lapply(namesID, function(x){  
 blnew <- min.dates[min.dates$sppListID == x,]  
 foo1 <- summary(lm(blnew$startDayOfYear~blnew[,colnum]))  
 jul.yr.coef <- data.frame(cbind(data.frame(t(foo1$coefficients[2,])),  
 foo1$coefficients[1,1],  
 t(foo1$fstatistic),  
 Species = blnew$Scientific.Name[1],  
 foo1$adj.r.squared))  
 })  
   
 JYC <- do.call(rbind, JYC)  
 colnames(JYC) <- c("Slope", "StErr", "t.value", "p.value", "Intercept",   
 "F Stat", "numDF","denDF", "Species", "adj.r.squ")  
 JYC  
}

# "Year" regression per species

yr.bl <- JulBloomDate("Year")  
  
head(yr.bl)

## Slope StErr t.value p.value Intercept F Stat numDF  
## 1 -0.45016752 0.1828673 -2.4617167 0.018015922 1082.4470 6.06004926 1  
## 2 -0.40659778 0.2228528 -1.8245127 0.080549327 1012.4821 3.32884642 1  
## 3 0.04489947 0.2058786 0.2180871 0.829957647 113.5721 0.04756198 1  
## 4 -0.24516852 0.1912197 -1.2821299 0.211559542 690.4853 1.64385712 1  
## 5 -0.60587373 0.1944068 -3.1165255 0.006278158 1413.0445 9.71273111 1  
## 6 -0.69175166 0.1732347 -3.9931483 0.001333853 1594.6633 15.94523353 1  
## denDF  
## 1 42  
## 2 24  
## 3 17  
## 4 25  
## 5 17  
## 6 14  
## Species  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Achillea lanulosa Nuttall  
## 3 Adoxa moschatellina L.  
## 4 Agoseris aurantiaca (Hooker) Greene  
## 5 Agoseris glauca (Pursh) Rafinesque var dasycephala (Torrey & Gray) Jepson  
## 6 Agrostis scabra Willdenow  
## adj.r.squ  
## 1 0.10528598  
## 2 0.08521569  
## 3 -0.05586946  
## 4 0.02416531  
## 5 0.32616400  
## 6 0.49908556

#subset that signficantchange over time  
yr.sig <- yr.bl[yr.bl$p.value < 0.05,]  
  
#merge the significant species with the min.dates species  
yr.sig.min <- merge(yr.sig, min.dates, by.x = "Species", by.y = "Scientific.Name")  
  
  
ggplot(yr.sig.min, aes(Year,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Year")+  
 xlab("Year")+  
 ylab("Julian Date")



summary(lm(yr.sig.min$startDayOfYear~yr.sig.min$Year))

##   
## Call:  
## lm(formula = yr.sig.min$startDayOfYear ~ yr.sig.min$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -48.138 -12.233 -1.685 12.336 49.723   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1133.9245 55.3910 20.47 <2e-16 \*\*\*  
## yr.sig.min$Year -0.4696 0.0279 -16.83 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 16.93 on 1421 degrees of freedom  
## Multiple R-squared: 0.1662, Adjusted R-squared: 0.1656   
## F-statistic: 283.3 on 1 and 1421 DF, p-value: < 2.2e-16

nrow(yr.sig) #53 species

## [1] 53

nrow(yr.sig[yr.sig$Slope > 0,]) # no species have later bloom dates over time

## [1] 0

nrow(yr.sig[yr.sig$adj.r.squ > 0.25,]) # 18 species have adj.r.squares over 0.25

## [1] 18

yr.sig[yr.sig$adj.r.squ > 0.25,]

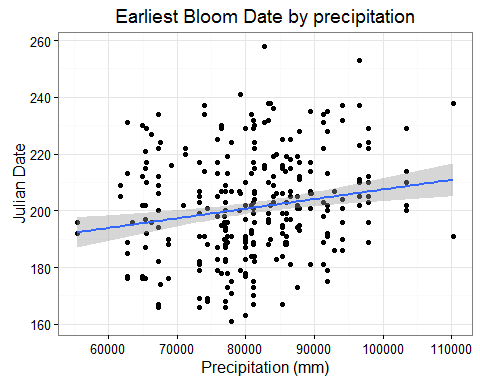
## Slope StErr t.value p.value Intercept F Stat numDF  
## 5 -0.6058737 0.1944068 -3.116525 0.0062781579 1413.044 9.712731 1  
## 6 -0.6917517 0.1732347 -3.993148 0.0013338532 1594.663 15.945234 1  
## 14 -0.7291745 0.2427297 -3.004060 0.0065328500 1652.838 9.024377 1  
## 22 -0.5147996 0.1627065 -3.163978 0.0046781957 1226.888 10.010754 1  
## 36 -0.6766929 0.2028178 -3.336457 0.0032890610 1539.530 11.131948 1  
## 45 -0.6312274 0.1652114 -3.820726 0.0006494715 1463.756 14.597948 1  
## 66 -0.9170587 0.2492922 -3.678650 0.0011252035 2031.147 13.532466 1  
## 74 -0.6911140 0.2130611 -3.243736 0.0037269136 1580.576 10.521826 1  
## 101 -0.8705972 0.3075933 -2.830352 0.0115431489 1932.016 8.010892 1  
## 110 -0.6503618 0.1792717 -3.627800 0.0017914396 1481.241 13.160936 1  
## 122 -0.5002499 0.1748223 -2.861476 0.0211034751 1181.994 8.188047 1  
## 144 -0.6910101 0.2339231 -2.954005 0.0084926009 1579.558 8.726145 1  
## 148 -0.4695939 0.1933940 -2.428172 0.0380966285 1118.489 5.896017 1  
## 172 -0.4402197 0.1150473 -3.826424 0.0006992450 1081.360 14.641520 1  
## 197 -0.6484052 0.2097426 -3.091433 0.0074463865 1494.643 9.556960 1  
## 198 -0.9141139 0.2251490 -4.060039 0.0007347396 2023.864 16.483917 1  
## 225 -1.3373370 0.3496197 -3.825119 0.0033454449 2869.852 14.631533 1  
## 289 -0.7649151 0.3006903 -2.543863 0.0233940944 1710.893 6.471241 1  
## denDF  
## 5 17  
## 6 14  
## 14 22  
## 22 21  
## 36 20  
## 45 29  
## 66 25  
## 74 22  
## 101 17  
## 110 19  
## 122 8  
## 144 18  
## 148 9  
## 172 27  
## 197 15  
## 198 18  
## 225 10  
## 289 14  
## Species  
## 5 Agoseris glauca (Pursh) Rafinesque var dasycephala (Torrey & Gray) Jepson  
## 6 Agrostis scabra Willdenow  
## 14 Antennaria media Greene  
## 22 Arnica latifolia Bongard  
## 36 Boechera drummondii (A. Gray) L ve & L ve (B. stricta)  
## 45 Senecio crassulus A. Gray  
## 66 Carex haydeniana Olney  
## 74 Carex phaeocephala Piper  
## 101 Poa abbreviata R. Brown ssp pattersonii (Vasey) L ve et al.  
## 110 Poa fendleriana (Steudel) Vasey  
## 122 Micranthes oregana (T. J. Howell) Small  
## 144 Poa nemoralis L. ssp interior (Rydberg) W. A. Weber  
## 148 Ribes coloradense Coville  
## 172 Elymus trachycaulus (Link) Gould  
## 197 Festuca saximontana Rydberg  
## 198 Festuca thurberi Vasey in Rothrock  
## 225 Solidago simplex Humboldt  
## 289 Vaccinium myrtillus L. ssp oreophilum (Rydberg) L ve et al.  
## adj.r.squ  
## 5 0.3261640  
## 6 0.4990856  
## 14 0.2586475  
## 22 0.2905687  
## 36 0.3254518  
## 45 0.3118942  
## 66 0.3252443  
## 74 0.2927826  
## 101 0.2803135  
## 110 0.3781275  
## 122 0.4440342  
## 144 0.2890856  
## 148 0.3286796  
## 172 0.3275942  
## 197 0.3484535  
## 198 0.4490185  
## 225 0.5534180  
## 289 0.2672648

# "Raw\_Precip" regression per species

precip.bl <- JulBloomDate("Raw\_Precip")  
  
head(precip.bl)

## Slope StErr t.value p.value Intercept F Stat  
## 1 -1.258871e-04 0.0003039296 -0.41419826 0.6808367 199.7928 0.171560197  
## 2 2.099863e-04 0.0004495595 0.46709343 0.6446459 188.9991 0.218176274  
## 3 -1.098059e-04 0.0003895035 -0.28191249 0.7814137 211.0280 0.079474652  
## 4 2.914071e-04 0.0003214772 0.90646291 0.3733453 179.7434 0.821675000  
## 5 5.470743e-04 0.0003654109 1.49714828 0.1526910 162.2246 2.241452969  
## 6 2.991075e-05 0.0004879483 0.06129901 0.9519876 216.3949 0.003757569  
## numDF denDF  
## 1 1 42  
## 2 1 24  
## 3 1 17  
## 4 1 25  
## 5 1 17  
## 6 1 14  
## Species  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Achillea lanulosa Nuttall  
## 3 Adoxa moschatellina L.  
## 4 Agoseris aurantiaca (Hooker) Greene  
## 5 Agoseris glauca (Pursh) Rafinesque var dasycephala (Torrey & Gray) Jepson  
## 6 Agrostis scabra Willdenow  
## adj.r.squ  
## 1 -0.01964451  
## 2 -0.03228252  
## 3 -0.05389658  
## 4 -0.00690602  
## 5 0.06451971  
## 6 -0.07114108

#subset that signficant change by precipitation  
precip.sig <- precip.bl[precip.bl$p.value < 0.05,]  
  
#merge the significant species with the min.dates species  
precip.sig.min <- merge(precip.sig, min.dates, by.x = "Species", by.y = "Scientific.Name")  
  
  
ggplot(precip.sig.min, aes(Raw\_Precip,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by precipitation")+  
 xlab("Precipitation (mm)")+  
 ylab("Julian Date")



summary(lm(precip.sig.min$startDayOfYear~precip.sig.min$Year))

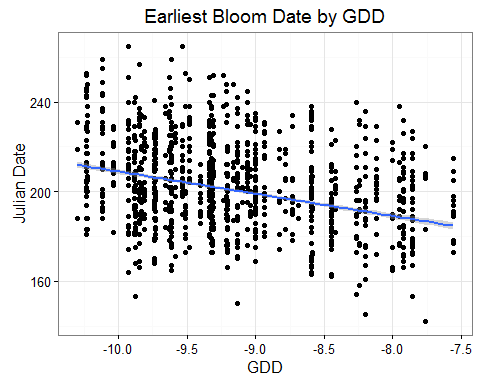
##   
## Call:  
## lm(formula = precip.sig.min$startDayOfYear ~ precip.sig.min$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -37.892 -12.332 -0.025 12.017 56.323   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 459.63603 115.07426 3.994 7.94e-05 \*\*\*  
## precip.sig.min$Year -0.13028 0.05798 -2.247 0.0253 \*   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.15 on 343 degrees of freedom  
## Multiple R-squared: 0.01451, Adjusted R-squared: 0.01164   
## F-statistic: 5.05 on 1 and 343 DF, p-value: 0.02527

# "GDD" regression per species

gdd.bl <- JulBloomDate("GDD")  
  
head(gdd.bl)

## Slope StErr t.value p.value Intercept F Stat numDF  
## 1 -9.504469 4.458659 -2.131688 0.038926682 103.03220 4.544094 1  
## 2 -15.177098 4.627225 -3.279957 0.003162765 67.69902 10.758117 1  
## 3 6.738544 6.554533 1.028074 0.318321648 262.94914 1.056936 1  
## 4 -7.711345 4.133867 -1.865407 0.073902426 134.03948 3.479744 1  
## 5 -7.769727 3.685687 -2.108081 0.050168407 138.05216 4.444006 1  
## 6 -11.030543 3.373939 -3.269337 0.005593327 121.59895 10.688565 1  
## denDF  
## 1 42  
## 2 24  
## 3 17  
## 4 25  
## 5 17  
## 6 14  
## Species  
## 1 Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## 2 Achillea lanulosa Nuttall  
## 3 Adoxa moschatellina L.  
## 4 Agoseris aurantiaca (Hooker) Greene  
## 5 Agoseris glauca (Pursh) Rafinesque var dasycephala (Torrey & Gray) Jepson  
## 6 Agrostis scabra Willdenow  
## adj.r.squ  
## 1 0.076144861  
## 2 0.280743551  
## 3 0.003153127  
## 4 0.087070457  
## 5 0.160604589  
## 6 0.392431271

#subset that signficant change by GDD  
gdd.sig <- gdd.bl[gdd.bl$p.value < 0.05,]  
  
#merge the significant species with the min.dates species  
gdd.sig.min <- merge(gdd.sig, min.dates, by.x = "Species", by.y = "Scientific.Name")  
  
  
ggplot(gdd.sig.min, aes(GDD,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by GDD")+  
 xlab("GDD")+  
 ylab("Julian Date")



summary(lm(gdd.sig.min$startDayOfYear~gdd.sig.min$Year))

##   
## Call:  
## lm(formula = gdd.sig.min$startDayOfYear ~ gdd.sig.min$Year)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -51.834 -14.043 -1.063 12.866 64.552   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 796.66710 43.34433 18.38 <2e-16 \*\*\*  
## gdd.sig.min$Year -0.30067 0.02183 -13.77 <2e-16 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 18.25 on 2595 degrees of freedom  
## Multiple R-squared: 0.0681, Adjusted R-squared: 0.06774   
## F-statistic: 189.6 on 1 and 2595 DF, p-value: < 2.2e-16

nrow(gdd.sig) #101 species

## [1] 101

nrow(gdd.sig[gdd.sig$Slope > 0,]) # no species have later bloom dates as the GDD increases

## [1] 0

Columns that you can use to compare to earliest bloom date (taken from "startDayOfYear")

# [11] "startDayOfYear" [20] "Raw\_Precip"

# [21] "Avg\_Hi" "Med\_Hi"

# [23] "Avg\_Lo" "Med\_Lo"

# [25] "Av\_Temp" "Med\_Temp"

# multiple linear regression

namesID <- unique(min.dates$sppListID)  
   
 ml.JYC <- lapply(namesID, function(x){  
 blnew <- min.dates[min.dates$sppListID == x,]  
 foo1 <- summary(lm(startDayOfYear~Raw\_Precip\*Avg\_Lo, data = blnew))  
 jul.yr.coef <- data.frame(row.names(foo1$coefficients[-1,]),  
 foo1$coefficients[-1,],  
 foo1$coefficients[1,1],  
 t(foo1$fstatistic),  
 Species = blnew$Scientific.Name[1],  
 foo1$adj.r.squared)  
 })  
   
 ml.JYC <- do.call(rbind, ml.JYC)  
 colnames(ml.JYC) <- c("Variables", "Slope", "StErr", "t.value", "p.value", "Intercept",   
 "F Stat", "numDF","denDF", "Species", "adj.r.squ")  
 head(ml.JYC)

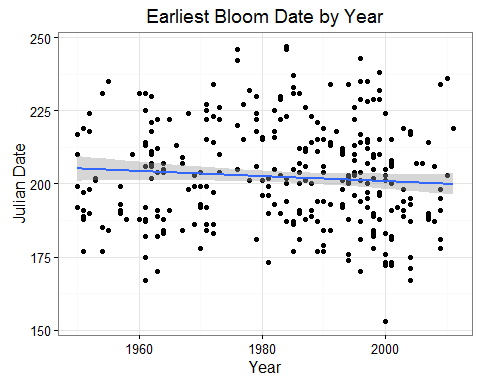
## Variables Slope StErr t.value  
## Raw\_Precip Raw\_Precip 3.396414e-03 4.390003e-03 0.7736701  
## Avg\_Lo Avg\_Lo -5.334365e+01 5.606830e+01 -0.9514048  
## Raw\_Precip:Avg\_Lo Raw\_Precip:Avg\_Lo 5.619372e-04 6.950289e-04 0.8085092  
## Raw\_Precip1 Raw\_Precip -3.905342e-03 5.610458e-03 -0.6960825  
## Avg\_Lo1 Avg\_Lo 3.589123e+01 6.796464e+01 0.5280868  
## Raw\_Precip:Avg\_Lo1 Raw\_Precip:Avg\_Lo -6.047304e-04 8.517737e-04 -0.7099660  
## p.value Intercept F Stat numDF denDF  
## Raw\_Precip 0.4436745 -134.1676 1.718695 3 40  
## Avg\_Lo 0.3471127 -134.1676 1.718695 3 40  
## Raw\_Precip:Avg\_Lo 0.4235803 -134.1676 1.718695 3 40  
## Raw\_Precip1 0.4936661 439.2955 2.237397 3 22  
## Avg\_Lo1 0.6027257 439.2955 2.237397 3 22  
## Raw\_Precip:Avg\_Lo1 0.4851832 439.2955 2.237397 3 22  
## Species  
## Raw\_Precip Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## Avg\_Lo Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## Raw\_Precip:Avg\_Lo Acomastylis rossii (R. Brown) Greene ssp turbinata (Rydberg) W. A. Weber  
## Raw\_Precip1 Achillea lanulosa Nuttall  
## Avg\_Lo1 Achillea lanulosa Nuttall  
## Raw\_Precip:Avg\_Lo1 Achillea lanulosa Nuttall  
## adj.r.squ  
## Raw\_Precip 0.04774739  
## Avg\_Lo 0.04774739  
## Raw\_Precip:Avg\_Lo 0.04774739  
## Raw\_Precip1 0.12928974  
## Avg\_Lo1 0.12928974  
## Raw\_Precip:Avg\_Lo1 0.12928974

sig.ml.JYC <- ml.JYC[ml.JYC$p.value < 0.05,]  
  
# there are 19 species significant for Avg\_Lo, 15 for Raw\_Precip, and 14 for the interaction   
table(sig.ml.JYC$Variables)

##   
## Avg\_Lo Raw\_Precip Raw\_Precip:Avg\_Lo   
## 19 15 14

# Species signficantly changing with the interaction of precipitation and low temperature

sig.interaction.JYC <- ml.JYC[ml.JYC$p.value < 0.05 & ml.JYC$Variables == "Raw\_Precip:Avg\_Lo",]  
  
sig.in.merge <- merge(sig.interaction.JYC, min.dates, by.x = "Species",  
 by.y = "Scientific.Name")  
  
  
ggplot(sig.in.merge, aes(Year,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Year")+  
 xlab("Year")+  
 ylab("Julian Date")



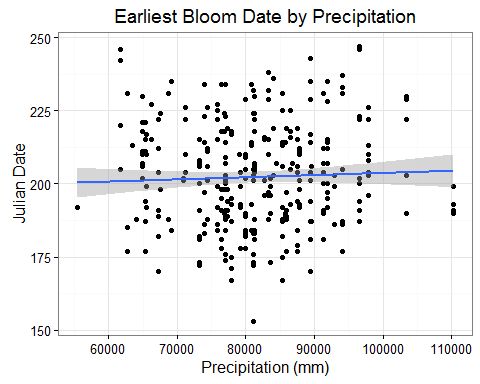
summary(lm(startDayOfYear~Year, data = sig.in.merge))

##   
## Call:  
## lm(formula = startDayOfYear ~ Year, data = sig.in.merge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -47.823 -13.068 -0.631 12.850 44.782   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 375.19399 112.21088 3.344 0.000916 \*\*\*  
## Year -0.08719 0.05656 -1.541 0.124100   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 17.47 on 350 degrees of freedom  
## Multiple R-squared: 0.006743, Adjusted R-squared: 0.003906   
## F-statistic: 2.376 on 1 and 350 DF, p-value: 0.1241

# Signficant interaction species

Precip

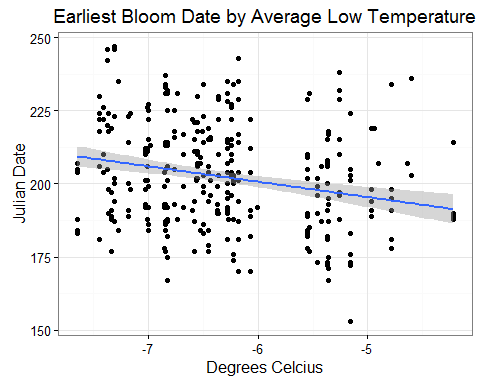
ggplot(sig.in.merge, aes(Raw\_Precip,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Precipitation")+  
 xlab("Precipitation (mm)")+  
 ylab("Julian Date")



summary(lm(startDayOfYear~Raw\_Precip, data = sig.in.merge))

##   
## Call:  
## lm(formula = startDayOfYear ~ Raw\_Precip, data = sig.in.merge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -49.21 -13.31 -1.06 12.15 45.22   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 1.962e+02 7.589e+00 25.857 <2e-16 \*\*\*  
## Raw\_Precip 7.385e-05 9.263e-05 0.797 0.426   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 17.52 on 350 degrees of freedom  
## Multiple R-squared: 0.001813, Adjusted R-squared: -0.001039   
## F-statistic: 0.6357 on 1 and 350 DF, p-value: 0.4258

ggplot(sig.in.merge, aes(Avg\_Lo,startDayOfYear))+  
 geom\_point()+  
 stat\_smooth(method="lm")+  
 theme\_bw()+  
 labs(title="Earliest Bloom Date by Average Low Temperature")+  
 xlab("Degrees Celcius")+  
 ylab("Julian Date")



summary(lm(startDayOfYear~Avg\_Lo, data = sig.in.merge))

##   
## Call:  
## lm(formula = startDayOfYear ~ Avg\_Lo, data = sig.in.merge)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -43.337 -13.233 -0.894 11.811 42.590   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 169.372 7.281 23.262 < 2e-16 \*\*\*  
## Avg\_Lo -5.226 1.149 -4.548 7.48e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 17.04 on 350 degrees of freedom  
## Multiple R-squared: 0.0558, Adjusted R-squared: 0.0531   
## F-statistic: 20.68 on 1 and 350 DF, p-value: 7.479e-06

# extract climate data from .bil file time series

<https://github.com/ropensci/prism>

# slice out only elevations above 10,500" (3200m)  
latlongs <- read.delim(path.expand("Q:/Research/All\_Projects\_by\_Species/aa\_Shapefiles\_Maps/aa\_QGIS Projects/Alpine Phenology Project/Sampling Points/LatLongabove3200m.txt"), sep=",")  
  
ll <- mapply(FUN = function(x,y){  
 c(x,y)  
 },  
 latlongs$Long,latlongs$Lat  
 )  
  
head(latlongs)  
ll[,1]  
  
prism\_slice(ll[,1],ls\_prism\_data()[3,1]) # picks one max temp at one point

Into Climate\_MEDL

options(prism.path = "Q:/Research/All\_Projects\_by\_Species/aa\_Shapefiles\_Maps/aa\_QGIS Projects/Alpine Phenology Project/Climate\_MEDL")  
# takes a long long long time!!! Don't run again!  
 get\_prism\_dailys(type="tmin",   
 minDate = "1981-01-01",   
 maxDate = "2011-12-31", keepZip = FALSE)  
  
  
ls\_prism\_data()[3:13,]  
ls\_prism\_data(absPath = TRUE)[3:13,] #to see path to where they are  
ls\_prism\_data(name = TRUE)[3:13,] #details of the file  
prism\_image(ls\_prism\_data()[3,1])  
  
ls\_prism\_data()[3,]

Into Climate\_MEDL

# error in download length != reported length  
options(prism.path = "Q:/Research/All\_Projects\_by\_Species/aa\_Shapefiles\_Maps/aa\_QGIS Projects/Alpine Phenology Project/Climate\_MEDL")  
  
#get\_prism\_dailys(type="tmin", minDate = "1981-01-01", maxDate = "1981-02-01",  
# keepZip = FALSE, mode="wb") #but I don't think mode is a thing...  
  
  
# run 12/3/2015  
get\_prism\_dailys(type="tmax",   
 minDate = "1981-01-01",   
 maxDate = "2011-12-31", keepZip = FALSE)

Calculate average daily temperatures averaged over all locations above 3200m 10500"

# combine by year  
sliceyear <- grep("1981",ls\_prism\_data()[,1],value = TRUE)  
sliceyear <- grep("tmin", sliceyear, value = TRUE) #only have tmin in this folder  
str(sliceyear)  
str(sliceyear[[1]][1])  
p <- prism\_slice(c(-106.688031183,40.9074207425), sliceyear)  
prism\_slice(ll[,1], sliceyear)  
p$data  
  
head(p$data)  
head(ll[,1:10])  
mean(p$data$data)  
  
# going through first 10 days of 1981  
ptest <- prism\_slice(ll[,1],sliceyear[[1]][1:10])   
head(ptest)  
ptest$data  
  
length(ls\_prism\_data()[,1]) #22,644  
  
# test apply through a few points to get values for a mean  
test1981 <- lapply(1:5, function(y){  
 prism\_slice(ll[,y], sliceyear[[1]][1:10])  
 })  
  
test1981[[1]]$data  
# data frame each $data part and then bind them  
test1981one <- do.call(rbind, lapply(test1981, function(x){  
 do.call(data.frame, x$data)  
 }))  
  
#aggregate for average over all the locations  
aggregate(data~date, data = test1981one, FUN = mean)  
  
length(ll) #2529 above 3200m points

# All the data held here  
options(prism.path = "Q:/Research/All\_Projects\_by\_Species/aa\_Shapefiles\_Maps/aa\_QGIS Projects/Alpine Phenology Project/Climate\_MEDL")  
  
samplell <- sample(1:length(ll), 5)  
  
# raster instead of slice!  
avgTemps <- lapply(c(1:5,2000:2005), function(x){  
 rastertemps <- raster(ls\_prism\_data(absPath=TRUE)[1,2])  
 r2points <- data.frame(rasterToPoints(rastertemps)) # pulls out data from raster  
 r2p <- data.frame(r2points, colnames(r2points)[3])  
 merge(r2points, latlongs[,3:4], by.x = c("x","y"), by.y = c("Long","Lat"))  
})  
  
# no points match, that's silly!  
head(avgTemps[[1]])

Could limit days to months before latest of the earliest bloom days...

# raster and extract  
maxs <- grep("tmax", ls\_prism\_data(absPath=TRUE)[,2])  
mins <- grep("tmin", ls\_prism\_data(absPath=TRUE)[,2], value=TRUE)  
#Jan - Sept, exclude Oct, Nov, and Dec. Latest first bloom is 269 = Sept 26  
# exclude <- c("\_[1-2][09][0-9][0-9][1][0]","\_[1-2][09][0-9][0-9][1][1]","\_[1-2][09][0-9][0-9][1][2]")  
mins <- grep("\_[1-2][09][0-9][0-9][0][1-9]", mins)  
  
avgTemps.max <- lapply(maxs, function(x){  
 rastertemps <- raster(ls\_prism\_data(absPath=TRUE)[x,2])  
 data.frame(data = extract(rastertemps, latlongs[,c(4,3)]), date = ls\_prism\_data()[x,])  
})  
  
avgTemps.min <- lapply(mins, function(x){  
 rastertemps <- raster(ls\_prism\_data(absPath=TRUE)[x,2])  
 data.frame(data = extract(rastertemps, latlongs[,c(4,3)]), date = ls\_prism\_data()[x,])  
})

nrow(latlongs)  
# they all got pulled out!!! yay!!  
nrow(avgTemps.min[[1]])  
nrow(avgTemps.max[[1]])  
  
  
mins.avg <- do.call(rbind, avgTemps.min)  
meanmins <- aggregate(data~date, data = mins.avg, mean)  
maxs.avg <- do.call(rbind, avgTemps.max)  
meanmaxs <- aggregate(data~date, data = maxs.avg, mean)  
  
head(mins.avg)  
head(as.Date(substr(meanmins$date, 25, 32), "%Y%m%d"))  
  
# Make a day column  
meanmins$day <- as.Date(substr(meanmins$date, 25, 32), "%Y%m%d")  
meanmaxs$day <- as.Date(substr(meanmaxs$date, 25, 32), "%Y%m%d")  
  
# make a julian day column  
meanmins$Julian <- yday(meanmins$day)  
  
# merge min and max per day  
temps <- merge(meanmins, meanmaxs, by = "day") #.x min, .y max  
temps$GDD cumsum(((max-min)/2)-base)  
temps$Julian <- yday(temps$day)  
  
#make a year column  
temps$Year <- as.numeric(format(temps$day, "%Y"))  
  
head(temps)  
  
ggplot(temps, aes(Julian, data.x))+  
 geom\_line()

Calculate cumulative HDD per year  
10 celsius is 50 fahrenheit  
5 celsius is 41  
Gordon and Bootsma (1993)

tbase <- 5  
  
HDD <- sapply(unique(temps$Year), function(year){  
 cumsum(sapply(grep(year, temps$Year), function(x){  
 if( (((temps$data.y[x] - temps$data.x[x])/2)-tbase) < tbase){  
 0  
 } else   
 ((temps$data.y[x] - temps$data.x[x])/2)-tbase  
 }  
 ))  
 })  
  
  
temps$HDD <- HDD

ggplot(temps, aes(Julian, HDD))+  
 geom\_line()+  
 facet\_wrap(~Year)

## Can get worldclim data

library(rgdal)  
  
w <- getData('worldclim', var='tmin', res=0.5, lon=5, lat=45)