Temperature Plots

Tom Gocken

Tuesday, January 26, 2016

### Load previously created tidy dataset, envdat

setwd("C:/Projects/springboard-wrangling")  
envdat <- read.delim("../data/envdat.txt")

### Create by month and by year datasets

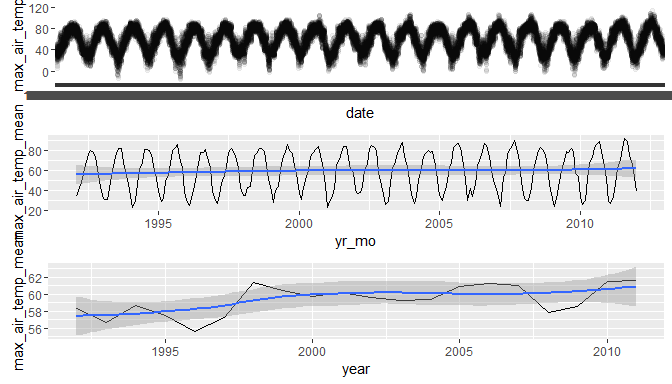
library(dplyr)

##   
## Attaching package: 'dplyr'  
##   
## The following objects are masked from 'package:stats':  
##   
## filter, lag  
##   
## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

envdat$yr\_mo <- envdat$year + envdat$month / 12  
  
envdat\_by\_month <- envdat %>%  
 group\_by(yr\_mo) %>%  
 summarize(max\_air\_temp\_mean = mean(max\_air\_temp),  
 max\_air\_temp\_median = median(max\_air\_temp),  
 n = n()) %>%  
 arrange(yr\_mo)  
  
envdat\_by\_year <- envdat %>%  
 group\_by(year) %>%  
 summarize(max\_air\_temp\_mean = mean(max\_air\_temp),  
 max\_air\_temp\_median = median(max\_air\_temp),  
 n = n()) %>%  
 arrange(year)

### Plot raw max temp air data, by month data, and by year data

library(ggplot2)  
library(gridExtra)  
  
p1 <- ggplot(aes(x = date, y = max\_air\_temp), data = envdat) + geom\_point(alpha = 1/20)  
p2 <- ggplot(aes(x = yr\_mo, y = max\_air\_temp\_mean), data = envdat\_by\_month) +  
 geom\_line() + geom\_smooth()  
p3 <- ggplot(aes(x = year, y = max\_air\_temp\_mean), data = envdat\_by\_year) +  
 geom\_line() + geom\_smooth()  
  
grid.arrange(p1, p2, p3, ncol = 1)



### Correlations of max air temp with month and year

cor.test(envdat\_by\_month$yr\_mo, envdat\_by\_month$max\_air\_temp\_mean, method = 'pearson')

##   
## Pearson's product-moment correlation  
##   
## data: envdat\_by\_month$yr\_mo and envdat\_by\_month$max\_air\_temp\_mean  
## t = 1.0033, df = 238, p-value = 0.3167  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## -0.06224436 0.18996586  
## sample estimates:  
## cor   
## 0.06489707

cor.test(envdat\_by\_year$year, envdat\_by\_year$max\_air\_temp\_mean, method = 'pearson')

##   
## Pearson's product-moment correlation  
##   
## data: envdat\_by\_year$year and envdat\_by\_year$max\_air\_temp\_mean  
## t = 3.1131, df = 18, p-value = 0.006005  
## alternative hypothesis: true correlation is not equal to 0  
## 95 percent confidence interval:  
## 0.2019319 0.8195574  
## sample estimates:  
## cor   
## 0.5915888