Project Instructions

Courtney Cooper

4/25/2022

Order of Files-

1. bt maize cleaning.rmd (reading in raw data files and cleaning to produce rmd file)
2. bt maize analysis.rmd (analysis/regressions run from rmd file)
3. whiteanalysis.r (analysis done of just white maize)

packages needed:

library(equatiomatic)  
library(knitr)  
library(magrittr)  
library(rmarkdown)  
  
dat <- readRDS("data/finalpanel.rds")  
  
dat %>%  
 group\_by(year) %>%  
 summarise(mean = mean(yield, na.rm=T))  
  
  
  
## require 5 obs per year per province  
summaryobs <- dat %>% group\_by(year, provence, color, bt) %>%  
 summarise(count = n()) # check bt if in more than one observation per province/year  
summaryobs <- summaryobs[summaryobs$count < 5,]  
summaryobs$count <- "1"  
dat <- merge(dat,summaryobs, by = c("year","provence", "bt","color"),all = TRUE, no.dups= TRUE)  
dat$count <- if\_else(dat$count %in% c(NA), "0", "1")  
dat <- dat[!dat$count== "1", ]  
  
summaryobs <- dat %>% group\_by(technology, year, provence, color) %>%  
 summarise(count = n())  
  
  
dat$GM <- 0  
dat$GM <- ifelse(dat$technology != "conv", 1, 0)  
  
bandconv <- dat[dat$technology %in% c("B", "conv"),]  
bandconv <- bandconv[!bandconv$provence == "LP",]  
  
dat <- bandconv  
  
  
summarynew <- dat %>% group\_by(provence,color, technology, year, .add = FALSE) %>%   
 summarise(mean = mean(yield, na.rm = T),   
 SD = sd(yield, na.rm = T))  
  
summaryyellow <- dat[!dat$color == "yellow",] %>% group\_by(provence, technology, year, .add = FALSE) %>%   
 summarise(mean = mean(yield, na.rm = T),   
 SD = sd(yield, na.rm = T))  
  
reg1 <- glm(yield ~ GM + color + provence + factor(year) + irrigated,data = dat)  
summary(reg1)  
  
dat$yearsq <- dat$year\*dat$year  
  
reg2 <- glm(yield ~ GM + color + year\*GM + yearsq\*GM + provence + factor(year)+ irrigated,data = dat)  
summary(reg2)  
  
  
dat$y\_effect <- reg2$coefficients["GM"] + reg2$coefficients["GM:year"] \* dat$year  
dat$ysq\_effect <- reg2$coefficients["GM"] + reg2$coefficients["GM:year"] \* dat$year + reg2$coefficients["GM:yearsq"] \* dat$yearsq  
  
max(dat$ysq\_effect)  
  
  
# Provence by year by GM effects in one model  
  
reg3 <- glm(yield ~ GM + color+ year\*provence\*GM + provence\*yearsq\*GM + provence + factor(year)+ + irrigated,data = dat)  
summary(reg3)  
  
models <- list(reg1,reg2,reg3)  
  
extract\_eq(reg1)  
extract\_eq(reg2)  
extract\_eq(reg3)