task 2

2024-03-06

## Problem 2

Utilize the attached market price data for livestock to fit a model of your choice, determining the advantages or disadvantages of selling at a specific point. Begin with an exploratory analysis to guide your model fitting process. The objective is to demonstrate the benefits of selling earlier to mitigate price declines resulting from drought, employing modeling techniques, data products, and innovative visuals:

MPrices\_2022 <- read\_excel("Makert Prices 2022.xlsx")  
  
 glimpse (MPrices\_2022)

## Rows: 154  
## Columns: 6  
## $ Dates <dttm> 2009-08-01, 2009-09-01, 2009-10-01, 2009-11-01, 2009-12-01, 2…  
## $ Seasons <chr> "Drought", "Drought", "Drought", "Drought", "Wet", "Wet", "Wet…  
## $ Bull <dbl> 8500.00, 18000.00, 8500.00, 12484.31, 22166.67, 25500.00, 3050…  
## $ Cow <dbl> 966.6667, 4000.0000, 5000.0000, 4815.0219, 9477.7778, 19583.33…  
## $ Heifer <dbl> 1666.667, 15000.000, 8000.000, 5198.550, 10111.111, 16583.333,…  
## $ Steer <dbl> 766.6667, 13000.0000, 9000.0000, 11302.7150, 18222.2222, 23000…

table(MPrices\_2022$Seasons)

##   
## drought Drought dry Dry wet Wet   
## 12 13 24 50 5 50

MPrices\_2022 <- MPrices\_2022%>%  
 mutate (Seasons\_1 = case\_when(Seasons %in% c("drought","Drought")~ "Drought",  
 Seasons %in% c("dry","Dry")~"Dry",  
 Seasons %in% c("wet","Wet")~"Wet")) #rename  
  
MPrices\_2022 %>%   
 count(Seasons,Seasons\_1)#renamed well

## # A tibble: 6 × 3  
## Seasons Seasons\_1 n  
## <chr> <chr> <int>  
## 1 Drought Drought 13  
## 2 Dry Dry 50  
## 3 Wet Wet 50  
## 4 drought Drought 12  
## 5 dry Dry 24  
## 6 wet Wet 5

MPrices <- MPrices\_2022 %>%   
 group\_by(Seasons\_1)%>%  
 mutate(bull\_average= mean(Bull))%>%  
 mutate(cow\_average= mean(Cow))%>%  
 mutate(heifer\_average= mean(Heifer))%>%  
 mutate(steer\_average= mean(Steer))  
   
   
# colnames(MPrices\_2022)  
#trend graph for the species over the various season  
   
bull<- MPrices %>%   
 ggplot(aes(x=Seasons\_1,y=Bull/100000))+geom\_col(aes(fill=Seasons\_1))+  
 scale\_fill\_manual(values = c("pink", "#E7B800", "grey"))+  
 labs(x ="Season" , y="Bull in millions") +theme\_bw()+  
 theme(legend.position="none")  
  
ggsave("bull.png",dpi=700)

## Saving 5 x 4 in image

cow<- MPrices %>%   
 ggplot(aes(x=Seasons\_1,y=Cow/100000))+geom\_col(aes(fill=Seasons\_1))+  
 scale\_fill\_manual(values = c("pink", "#E7B800", "grey"))+  
 labs(x ="Season" , y="Cow in millions") +theme\_bw()+  
 theme(legend.position="none")  
  
  
ggsave("cow.png",dpi=700)

## Saving 5 x 4 in image

heifer<- MPrices %>%   
 ggplot(aes(x=Seasons\_1,y=Heifer/100000))+geom\_col(aes(fill=Seasons\_1))+  
 scale\_fill\_manual(values = c("pink", "#E7B800", "grey"))+  
 labs(x ="Season" , y="Heifer in millions") +theme\_bw()+  
 theme(legend.position="none")  
  
ggsave("heifer.png",dpi=700)

## Saving 5 x 4 in image

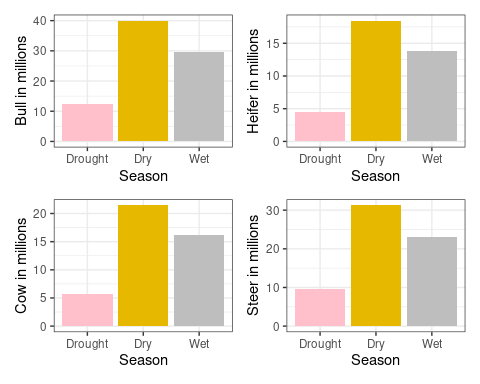
steer<- MPrices %>%   
 ggplot(aes(x=Seasons\_1,y=Steer/100000))+geom\_col(aes(fill=Seasons\_1))+  
 scale\_fill\_manual(values = c("pink", "#E7B800", "grey"))+  
 labs(x ="Season" , y="Steer in millions") +theme\_bw()+  
 theme(legend.position="none")  
  
ggsave("steer.png",dpi=700)

## Saving 5 x 4 in image

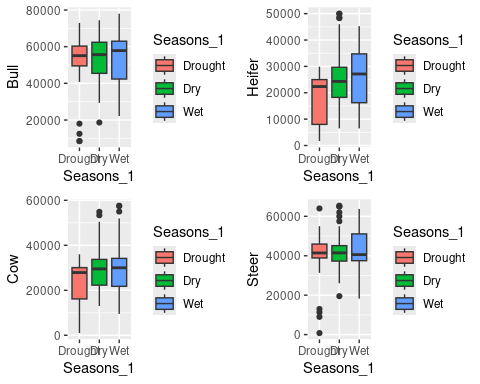
plot\_1<- bull +heifer+cow+steer  
  
ggsave("plot\_1.png",dpi=700)

## Saving 5 x 4 in image

plot\_1



p1 <- ggplot(MPrices, aes(x=Seasons\_1,y=Bull, fill=Seasons\_1)) + geom\_boxplot()   
p2 <- ggplot(MPrices, aes(x=Seasons\_1,y=Heifer, fill=Seasons\_1)) + geom\_boxplot()   
p3 <- ggplot(MPrices, aes(x=Seasons\_1,y=Cow, fill=Seasons\_1)) + geom\_boxplot()  
p4<- ggplot(MPrices, aes(x=Seasons\_1,y=Steer, fill=Seasons\_1)) + geom\_boxplot()  
  
 grid.arrange(p1, p2, p3,p4, ncol=2)



# data has outliers   
  
MPrices$Seasons\_1 <- as.factor(MPrices$Seasons\_1)  
 levels( MPrices$Seasons\_1)

## [1] "Drought" "Dry" "Wet"

MPrices$Seasons\_1 <- relevel(MPrices$Seasons\_1, ref = "Wet")  
  
  
#Model fitting  
model <- multinom(Seasons\_1 ~ Bull+Cow+Heifer+Steer, data = MPrices)

## # weights: 18 (10 variable)  
## initial value 169.186292   
## iter 10 value 149.693915  
## final value 148.623650   
## converged

summary(model)

## Call:  
## multinom(formula = Seasons\_1 ~ Bull + Cow + Heifer + Steer, data = MPrices)  
##   
## Coefficients:  
## (Intercept) Bull Cow Heifer Steer  
## Drought 0.2165021 6.194335e-05 -9.300247e-05 -5.474309e-05 -1.569912e-05  
## Dry 0.2100331 5.196282e-06 -1.362946e-05 2.139369e-06 3.570893e-06  
##   
## Std. Errors:  
## (Intercept) Bull Cow Heifer Steer  
## Drought 8.845086e-10 3.412042e-05 5.558177e-05 3.899249e-05 3.988933e-05  
## Dry 3.331555e-10 2.412438e-05 3.167232e-05 2.731003e-05 2.505892e-05  
##   
## Residual Deviance: 297.2473   
## AIC: 317.2473

#predictions  
index <- createDataPartition(MPrices$Seasons\_1, p = .70, list = FALSE)  
train <- MPrices[index,]  
test <- MPrices[-index,]  
  
  
# Predicting values for train dataset  
train$ClassPredicted <- predict(model, newdata = train, "class")  
tab <- table(train$Seasons\_1, train$ClassPredicted) # classification table  
#round((sum(diag(tab))/sum(tab))\*100,2) # Calculating accuracy  
  
  
# Predicting values for test dataset  
test$ClassPredicted <- predict(model, newdata = test, "class")  
tab1 <- table(test$Seasons\_1, test$ClassPredicted)  
#round((sum(diag(tab1))/sum(tab1))\*100,2)  
  
######################  
pred <-predict(model,MPrices,type='class')  
mean(pred==MPrices$Seasons\_1)

## [1] 0.512987

confusionMatrix(pred,MPrices$Seasons\_1)

## Confusion Matrix and Statistics  
##   
## Reference  
## Prediction Wet Drought Dry  
## Wet 0 0 0  
## Drought 2 8 3  
## Dry 53 17 71  
##   
## Overall Statistics  
##   
## Accuracy : 0.513   
## 95% CI : (0.4312, 0.5942)  
## No Information Rate : 0.4805   
## P-Value [Acc > NIR] : 0.2339   
##   
## Kappa : 0.1086   
##   
## Mcnemar's Test P-Value : 5.535e-14   
##   
## Statistics by Class:  
##   
## Class: Wet Class: Drought Class: Dry  
## Sensitivity 0.0000 0.32000 0.9595  
## Specificity 1.0000 0.96124 0.1250  
## Pos Pred Value NaN 0.61538 0.5035  
## Neg Pred Value 0.6429 0.87943 0.7692  
## Prevalence 0.3571 0.16234 0.4805  
## Detection Rate 0.0000 0.05195 0.4610  
## Detection Prevalence 0.0000 0.08442 0.9156  
## Balanced Accuracy 0.5000 0.64062 0.5422