```
#libraries
library(dplyr)
library(tidyr)
library(caret)

# read CSV file
weather_data <- read.csv("NOAAGISSWeatherDisasters.csv")

weather_data <- read.csv("NOAAGISSWeatherDisasters.csv", header = TRUE, sep = ",")

names(weather_data)

weather_data$Year <- as.numeric(as.character(weather_data$Year))
weather_data$delta.tempYear <- with(weather_data, delta.temp * Year)

# filter the data for drought
drought_data <- weather_data %>%
  filter(Drought.Count > 0) %>%
  select(delta.temp, Year, delta.tempYear)

# filter the data for wildfire
wildfire_data <- weather_data %>%
  filter(Wildfire.Count > 0) %>%
  select(delta.temp, Year, delta.tempYear)

# step 3 for drought_data

# fit a binomial logit model with drought as the response and delta.temp, year, and #delta.temp*year as
predictors
drought_model_interaction <- glm(Drought.Count ~ delta.temp * Year, data = weather_data, family =
"binomial")

# fit a second model without the interaction term
drought_model_additive <- glm(Drought.Count ~ delta.temp + Year, data = weather_data, family =
"binomial")

# personal notes:
# An AIC score is a number used to determine which machine learning model
# is best for a given data set in situations where one can't easily test a dataset
# The lower the AIC score the better.

# compare the AIC values for the two models
AIC_drought_interaction <- AIC(drought_model_interaction)
AIC_drought_additive <- AIC(drought_model_additive)
```

```
cat("AIC for drought model with interaction:", AIC_drought_interaction, "\n")
#output: AIC for drought model with interaction: 54.52863

cat("AIC for drought model without interaction:", AIC_drought_additive, "\n")
#output: AIC for drought model without interaction: 52.99482

# compare deviance differences between the two models using ANOVA
drought_anova <- anova(drought_model_additive, drought_model_interaction, test = "Chisq")
cat("Deviance difference for drought models:", drought_anova$Deviance[2] -
drought_anova$Deviance[1], "\n")
#output: Deviance difference for drought models: NA

cat("Deviance p-value for drought models:", drought_anova$`Pr(>Chi)`[2], "\n")
#output: Deviance p-value for drought models: 0.4947468

# choose the best model based on AIC and deviance comparison
if (AIC_drought_interaction < AIC_drought_additive && drought_anova$`Pr(>Chi)`[2] < 0.05) {
  drought_best_model <- drought_model_interaction
} else {
  drought_best_model <- drought_model_additive
}

# interpret coefficients of the best model
cat("Coefficients for drought model:", coef(drought_best_model), "\n")
#output: Coefficients for drought model: -87.14359 1.498534 0.04364329

# step 3 for wildfire

# fit a binomial logit model with wildfire as the response
# delta.temp, year, and delta.temp*year as predictors
wildfire_model_interaction <- glm(Wildfire.Count ~ delta.temp * Year, data = weather_data, family =
"binomial")

# xit a second model without the interaction term
wildfire_model_additive <- glm(Wildfire.Count ~ delta.temp + Year, data = weather_data, family =
"binomial")

# compare AIC values for two models
AIC_wildfire_interaction <- AIC(wildfire_model_interaction)
AIC_wildfire_additive <- AIC(wildfire_model_additive)
cat("AIC for wildfire model with interaction:", AIC_wildfire_interaction, "\n")
#output: AIC for wildfire model with interaction: 51.71748

cat("AIC for wildfire model without interaction:", AIC_wildfire_additive, "\n")
```

```
#output: AIC for wildfire model without interaction: 49.74829

# compare deviance differences between the two models using ANOVA
wildfire_anova <- anova(wildfire_model_additive, wildfire_model_interaction, test = "Chisq")

cat("Deviance difference for wildfire models:", wildfire_anova$Deviance[2] -
wildfire_anova$Deviance[1], "\n")
#output: Deviance difference for wildfire models: NA

cat("Deviance p-value for wildfire models:", wildfire_anova$`Pr(>Chi)`[2], "\n")
#output: Deviance p-value for wildfire models: 0.8606658

# choose the best model based on AIC and deviance comparison
if (AIC_wildfire_interaction < AIC_wildfire_additive && wildfire_anova$`Pr(>Chi)`[2] < 0.05) {
  wildfire_best_model <- wildfire_model_interaction
} else {
  wildfire_best_model <- wildfire_model_additive
}

# interpret coefficients of the best model
cat("Coefficients for wildfire model:", coef(wildfire_best_model), "\n")
#output: Coefficients for wildfire model: -212.8037 0.8126896 0.1061008
```

Code with Deviance Difference

```
library(dplyr)
library(lmtest)

# Read CSV file
weather_data <- read.csv("NOAAGISSWeatherDisasters.csv", header = TRUE, sep = ",")

# Convert Year to numeric
weather_data$Year <- as.numeric(as.character(weather_data$Year))
weather_data$delta.tempYear <- with(weather_data, delta.temp * Year)

# Filter the data for drought
drought_data <- weather_data %>%
  filter(Drought.Count > 0) %>%
  select(delta.temp, Year, delta.tempYear)

# Filter the data for wildfire
wildfire_data <- weather_data %>%
  filter(Wildfire.Count > 0) %>%
```

```
  select(delta.temp, Year, delta.tempYear)

# Fit drought models
drought_model_interaction <- glm(Drought.Count ~ delta.temp * Year, data = weather_data, family =
"binomial")
drought_model_additive <- glm(Drought.Count ~ delta.temp + Year, data = weather_data, family =
"binomial")

# Compare AIC values for drought models
AIC_drought_interaction <- AIC(drought_model_interaction)
AIC_drought_additive <- AIC(drought_model_additive)
cat("AIC for drought model with interaction:", AIC_drought_interaction, "\n")
cat("AIC for drought model without interaction:", AIC_drought_additive, "\n")

# Compare p-values for deviance difference using lrtest
deviance_test_drought <- lrtest(drought_model_additive, drought_model_interaction)
deviance_p_value_drought <- deviance_test_drought$Pr[2]
cat("Deviance p-value for drought models:", deviance_p_value_drought, "\n")

# Choose the best drought model based on AIC and deviance comparison
if (AIC_drought_interaction < AIC_drought_additive && deviance_p_value_drought < 0.05) {
  drought_best_model <- drought_model_interaction
} else {
  drought_best_model <- drought_model_additive
}

# Fit wildfire models
wildfire_model_interaction <- glm(Wildfire.Count ~ delta.temp * Year, data = weather_data, family =
"binomial")
wildfire_model_additive <- glm(Wildfire.Count ~ delta.temp + Year, data = weather_data, family =
"binomial")

# Compare AIC values for wildfire models
AIC_wildfire_interaction <- AIC(wildfire_model_interaction)
AIC_wildfire_additive <- AIC(wildfire_model_additive)
cat("AIC for wildfire model with interaction:", AIC_wildfire_interaction, "\n")
cat("AIC for wildfire model without interaction:", AIC_wildfire_additive, "\n")

# Compare p-values for deviance difference using lrtest
deviance_test_wildfire <- lrtest(wildfire_model_additive, wildfire_model_interaction)
deviance_p_value_wildfire <- deviance_test_wildfire$Pr[2]
cat("Deviance p-value for wildfire models:", deviance_p_value_wildfire, "\n")

# Choose the best wildfire model based on AIC and deviance comparison
```

```
if (AIC_wildfire_interaction < AIC_wildfire_additive && deviance_p_value_wildfire < 0.05) {
  wildfire_best_model <- wildfire_model_interaction
} else {
  wildfire_best_model <- wildfire_model_additive
}

# Interpret coefficients of the best models
cat("Coefficients for best drought model:", coef(drought_best_model), "\n")
cat("Coefficients for best wildfire model:", coef(wildfire_best_model), "\n")

# Graphs
plot_drought <- ggplot(drought_data, aes(x = Year, y = delta.temp)) +
  geom_boxplot() +
  labs(title = "Drought Data",
       x = "Year",
       y = "Delta Temperature") +
  theme_minimal()

print(plot_drought)

plot_wildfire <- ggplot(wildfire_data, aes(x = "", y = delta.temp)) +
  geom_boxplot() +
  labs(title = "Wildfire Data",
       x = "",
       y = "Delta Temperature") +
  theme_minimal()

print(plot_wildfire)

plot_drought <- ggplot(drought_data, aes(x = Year, y = delta.temp)) +
  geom_point() +
  labs(title = "Drought Data",
       x = "Year",
       y = "Delta Temperature") +
  theme_minimal()

print(plot_drought)

plot_wildfire <- ggplot(wildfire_data, aes(x = "", y = delta.temp)) +
  geom_point() +
  labs(title = "Wildfire Data",
       x = "",
       y = "Delta Temperature") +
  theme_minimal()
```

```
print(plot_wildfire)
```