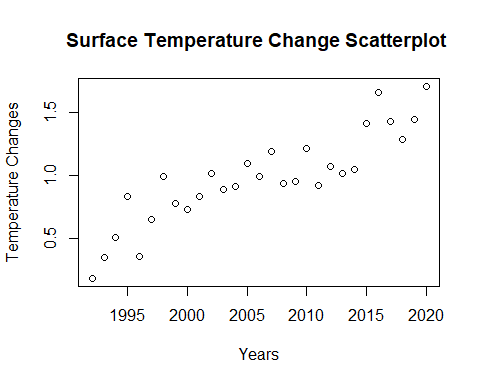
Papa\_Yaw\_Boampong\_Statistics\_FInal\_Project

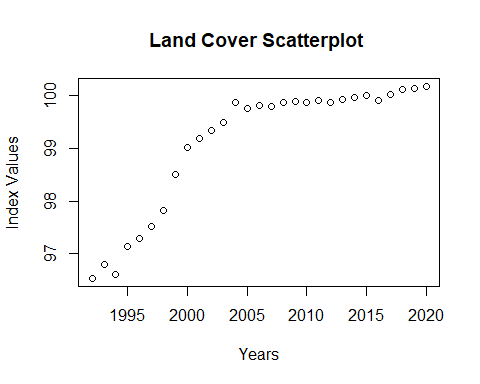
Papa Yaw Boampong

2023-03-23

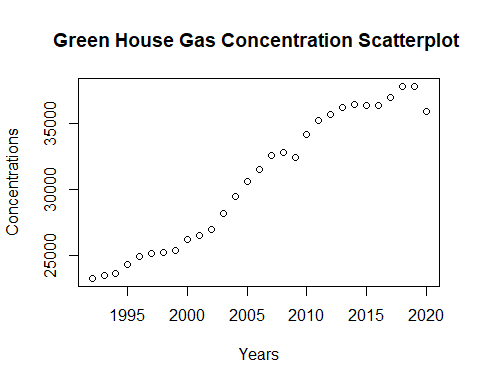
#Univariate distribution of Green House Gas Concentrations   
  
plot(Climate\_change\_vars$Year, Climate\_change\_vars $ `Surface.Temperature.Change`, main = "Surface Temperature Change Scatterplot ", ylab = "Temperature Changes", xlab = "Years")



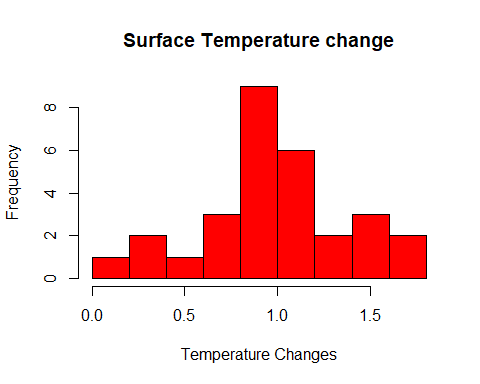
plot(Climate\_change\_vars$Year, Climate\_change\_vars $ `Land.Cover.Change`, main = "Land Cover Scatterplot ", ylab = "Index Values", xlab = "Years")



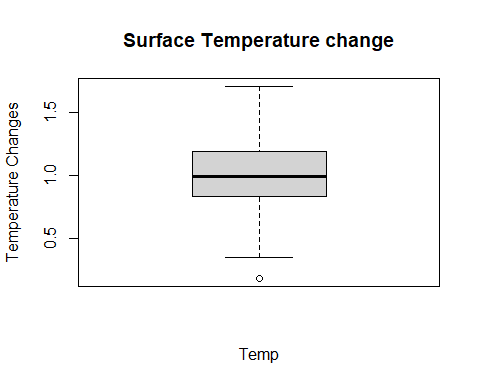
plot(Climate\_change\_vars$Year, Climate\_change\_vars $ `Green.House.Gas.Concentration`, main = "Green House Gas Concentration Scatterplot ", ylab = "Concentrations", xlab = "Years")



#Observing the Distribution of Explanatory and Response variables  
hist(Climate\_change\_vars$`Surface.Temperature.Change`,main = "Surface Temperature change", xlab = "Temperature Changes", col = "red")



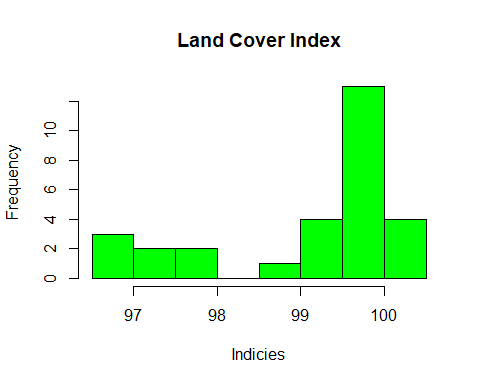
boxplot(Climate\_change\_vars$`Surface.Temperature.Change`,main = "Surface Temperature change", xlab = "Temp", ylab = "Temperature Changes")



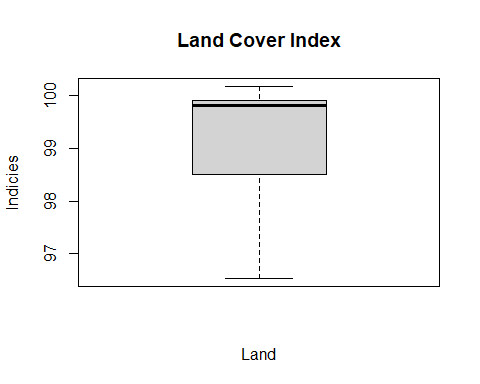
summary(Climate\_change\_vars$`Surface.Temperature.Change`)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.1840 0.8340 0.9930 0.9813 1.1950 1.7110

hist(Climate\_change\_vars$`Land.Cover.Change` ,main = " Land Cover Index ", xlab = "Indicies" , col = "green")



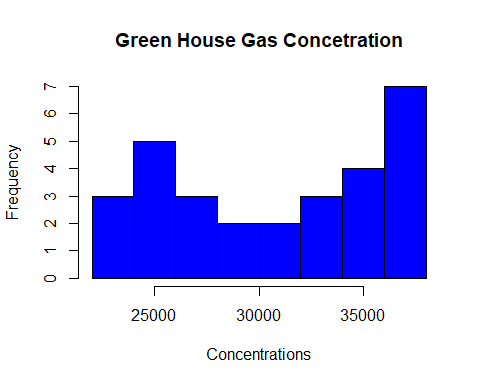
boxplot(Climate\_change\_vars$`Land.Cover.Change` ,main = " Land Cover Index ", xlab = "Land", ylab = "Indicies")



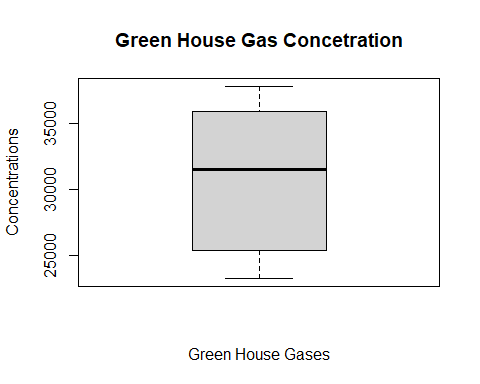
summary(Climate\_change\_vars$`Land.Cover.Change`)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 96.54 98.50 99.82 99.10 99.91 100.17

hist(Climate\_change\_vars$`Green.House.Gas.Concentration` ,main = "Green House Gas Concetration", xlab = "Concentrations" , col = "blue")



boxplot(Climate\_change\_vars$`Green.House.Gas.Concentration` ,main = "Green House Gas Concetration", xlab = "Green House Gases", ylab = "Concentrations")



summary(Climate\_change\_vars$`Green.House.Gas.Concentration`)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 23295 25413 31579 30777 35947 37842

# Converting Variables into numeric values  
sf <- as.numeric(Climate\_change\_vars$`Surface.Temperature.Change`)  
ld <- as.numeric(Climate\_change\_vars$`Land.Cover.Change`)  
ghg <- as.numeric(Climate\_change\_vars$`Green.House.Gas.Concentration`)  
  
  
  
# Mean and Standard deviation of Surface Temperature Change  
Temp\_mean <- mean(sf)  
Temp\_sd <- sd(sf)  
# The mean Surface Temperature change for the years 1992 to 2020   
Temp\_mean

## [1] 0.9813448

# The Standard Deviation for the Surface Temperature change for the years 1992 to 2020 is   
Temp\_sd

## [1] 0.3657503

# Mean and Standard deviation of Land Cover Change  
LC\_mean<- mean(ld)  
LC\_sd<- sd(ld)  
# The mean Land Cover index for the years 1992 to 2020 is  
LC\_mean

## [1] 99.10347

# The Standard Deviation for the Land Cover index for the years 1992 to 2020 is  
LC\_sd

## [1] 1.221718

# Mean and Standard deviation of Green House Gas Concentration  
GHG\_mean<- mean(ghg)  
GHG\_sd <- sd(ghg)  
#The mean Green House Gas Concentration for the years 1992 to 2020 is   
GHG\_mean

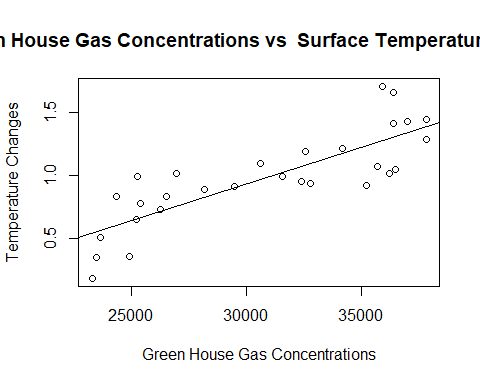
## [1] 30776.59

#The Standard Deviation for the Green House Gas Concentration for the years 1992 to 2020 is  
GHG\_sd

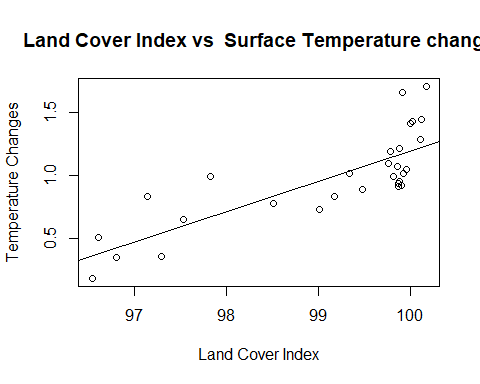
## [1] 5155.887

# H0 = There exist no linear correlation between Land Cover Index and the Surface Temperature change for the years 2010 to 2020. r = 0.  
# HA = There exist a linear correlation between Land Cover Index and the Surface Temperature change for the years 2010 to 2020. r not equal to 0.  
  
  
# H0` = There exist no linear correlation between Green House Gas Concentrations and the Surface Temperature change for the years 2010 to 2020. r = 0.  
# HA` = There exist a linear correlation between Green House Gas Concentrations and the Surface Temperature change for the years 2010 to 2020. r not equal to 0.

##Bi-variate scatter plot between Green House Gas Concentrations in the energy sector and the Surface Temperature change.  
plot(Climate\_change\_vars$`Surface.Temperature.Change` ~ Climate\_change\_vars$`Green.House.Gas.Concentration`, data = Climate\_change\_vars, main = "Green House Gas Concentrations vs Surface Temperature change", ylab = "Temperature Changes", xlab = "Green House Gas Concentrations")  
  
abline(lm(Climate\_change\_vars$`Surface.Temperature.Change` ~ Climate\_change\_vars$`Green.House.Gas.Concentration`))



#The above scatter plot is a bivariate distribution of Green House Gas Concentrations and the Surface Temperature change. The graph show a positive linear association between the two variables. Based on the positive association, I speculate a possible association between The change in surface temperature and Green House Concentration concentration.   
  
  
  
  
  
##Bi-variate scatter plot between Land Cover Index and the Surface Temperature change  
plot(Climate\_change\_vars$`Surface.Temperature.Change` ~ Climate\_change\_vars$`Land.Cover.Change`, data = Climate\_change\_vars, main = "Land Cover Index vs Surface Temperature change", ylab = "Temperature Changes", xlab = "Land Cover Index")  
  
abline(lm(Climate\_change\_vars$`Surface.Temperature.Change` ~ Climate\_change\_vars$`Land.Cover.Change`))



#The above scatter plot is a bivariate distribution of Land Cover Index against the Surface Temperature change. The graph show a positive linear association between the two variables. Based on the positive association, I speculate a possible association between The change in Land Cover Change and surface temperature .  
  
  
  
#Conditions that allow us to safely use the Pearson's correlation coefficient  
  
# The Samples size for my research is 28 years which represents the climatic conditions for the years 1992 to 2020.   
#Though the sample size is less than 30 which is the accepted benchmark for a sample to be considered fairly large, the lack of data on climate for the world has restricted me to 28 days. Since 28 days is close to 30, I will assume the sample is fairly large.  
  
# From the Scatter plot of all explanatory and response variables there is an Observed linear relationship between Surface Temperature change, Land Cover Index, and Green House Gas concentration.  
  
# The Observations which are Surface Temperature change, Green House Gas Concentration, Land Cover index are independent of each other.  
  
# Based on these observations Pearson's correlation coefficient can be used.  
  
  
  
cor\_results1 <- cor.test(sf,ld)  
cor\_results2 <- cor.test(sf,ghg)  
# For the assumed relationship between Land Cover Index and the Surface Temperature change index the Coeffecient of correlation r = 0.8022386 and r^2 =   
cor\_results1$estimate^2

## cor   
## 0.6435868

# Also the p-value is 1.665e-07  
  
  
# For the assumed relationship between Green House Gas Concentration and Surface Temperature change the Co effecient of correlation is 0.8219796 and r^2 =   
cor\_results2$estimate^2

## cor   
## 0.6756504

# Also the p-value is 4.562e-08  
  
  
# Given a significance level 𝛼 = 0.05 and a p value of 1.665e-07, For the assumed relationship between Land Cover index and Surface Temperature for the years 1992 to 2020 we have strong evidence to reject H0. The evidence establishs a strong positive linear relationship between Land Cover index and Surface Temperature change for the years 1992 to 2020. The evidence further suggests the Land cover index accounts for 64.36% (r^2 = 0.64358677) of the variation in surface temperature change  
  
# Given a significance level 𝛼 = 0.05 and a p value of 4.562e-08, For the assumed relationship between Green House Gas Concentration and Surface Temperature change for the years 1992 to 2020 we have strong evidence to reject H0. The evidence establishs a strong positive linear relationship between Green House Gas Concentration and Surface Temperature change for the years 1992 to 2020. The evidence further suggests the Green House Gas Concentration accounts for 67.57% (r^2 = 0.67565046) of the variation in surface temperature change  
  
  
# Given that for both relationships we were able to reject the null hypothesis there exist the chance of a Type I error being committed for both relationships  
  
  
y <- lm(Climate\_change\_vars$`Surface.Temperature.Change` ~ `Land.Cover.Change`, Climate\_change\_vars)  
summary(y)

##   
## Call:  
## lm(formula = Climate\_change\_vars$Surface.Temperature.Change ~   
## Land.Cover.Change, data = Climate\_change\_vars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.25069 -0.17843 -0.05452 0.12560 0.48583   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -22.8203 3.4090 -6.694 3.47e-07 \*\*\*  
## Land.Cover.Change 0.2402 0.0344 6.982 1.66e-07 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2224 on 27 degrees of freedom  
## Multiple R-squared: 0.6436, Adjusted R-squared: 0.6304   
## F-statistic: 48.75 on 1 and 27 DF, p-value: 1.665e-07

#The slope of the line of best fit for the relationship between Land Cover index and Surface Temperature for the years 1992 to 2020 is 0.2402   
  
  
y <- lm(Climate\_change\_vars$`Surface.Temperature.Change` ~ `Green.House.Gas.Concentration`, Climate\_change\_vars)  
summary(y)

##   
## Call:  
## lm(formula = Climate\_change\_vars$Surface.Temperature.Change ~   
## Green.House.Gas.Concentration, data = Climate\_change\_vars)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.36112 -0.16564 0.01009 0.10891 0.42815   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) -8.132e-01 2.425e-01 -3.353 0.00238 \*\*   
## Green.House.Gas.Concentration 5.831e-05 7.775e-06 7.500 4.56e-08 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.2121 on 27 degrees of freedom  
## Multiple R-squared: 0.6757, Adjusted R-squared: 0.6636   
## F-statistic: 56.24 on 1 and 27 DF, p-value: 4.562e-08

#The slope of the line of best fit for the relationship between Green House Gas Concentration and Surface Temperature change for the years 1992 to 2020 is 5.831e-05  
  
  
# RESEARCH QUESTION - 1. a. “Were the efforts which were made to curb Annual Surface Temperature Change (Global warming) due to Green House Gas Concentrations and changes in land cover for the years 1992 to 2020 effective.  
  
#RELATIONSHIP-a. “Is change in land cover associated with Annual Surface Temperature Change”  
#RELATIONSHIP-b.“Is change in Green House Gas Concentrations associated with Annual Surface Temperature Change?”  
  
  
# The results of my research for both RELATIONSHIP-a and RELATIONSHIP-b shows very strong positive associations for both relationships.Based on these results, I assume the iniative put in place by world governments over the years 1992 to 2020 were ineffective on curbing Global warming due to Green House Gas Concentrations and changes in land cover.   
  
  
#Limitation  
# 1. The time frame for the research was only 28 years. To accurately study trends such as climate, data for long periods of time are required. Due to the lack of ready data, my research was ristricted to only 28 years representing 28 different response values for climatic conditions.