Project 2

SDS348 Spring 2021 - 4/17/2021

Sungmo Hong (Sh46959)

Introduction

Changes in a variety of global climate predictors from 1960 to 2010

The following dataset was obtained from https://datahub.io/collections/climate-change (https://datahub.io/collections/climate-change) and https://datahub.io/collections/demographics (https://datahub.io/collections/demographics)

There are a total of 5 variables in this dataset. One of them is categorical and the remaining four are numerical. The variables are: “Year”, “Global Population Size”, “Average CO2 Levels in PPM”, “Average Number of Temperature Anomolies Relative to Base 0”, and “Change in Glacier Mass”. There are a total of 51 observations per variable and tidying was not necessary as the dataset was already tidy.

I have always been interested in climate change and its impact. Climate change is well documented and I wanted to see how a variety of different factors associated with climate change has changed over 55 years. I expect there to be significance between increasing CO2 levels with every other numerical variable. In addition, I also expect there to be a strong relationship/signifiance between increasing population level and the other numerical variables.

```
#install necessary libraries
library(readxl)
library(dplyr)
library(tidyverse)
#create dataset climate
climate <- read_excel("Climate.xlsx")
#view dataset
head(climate)
```

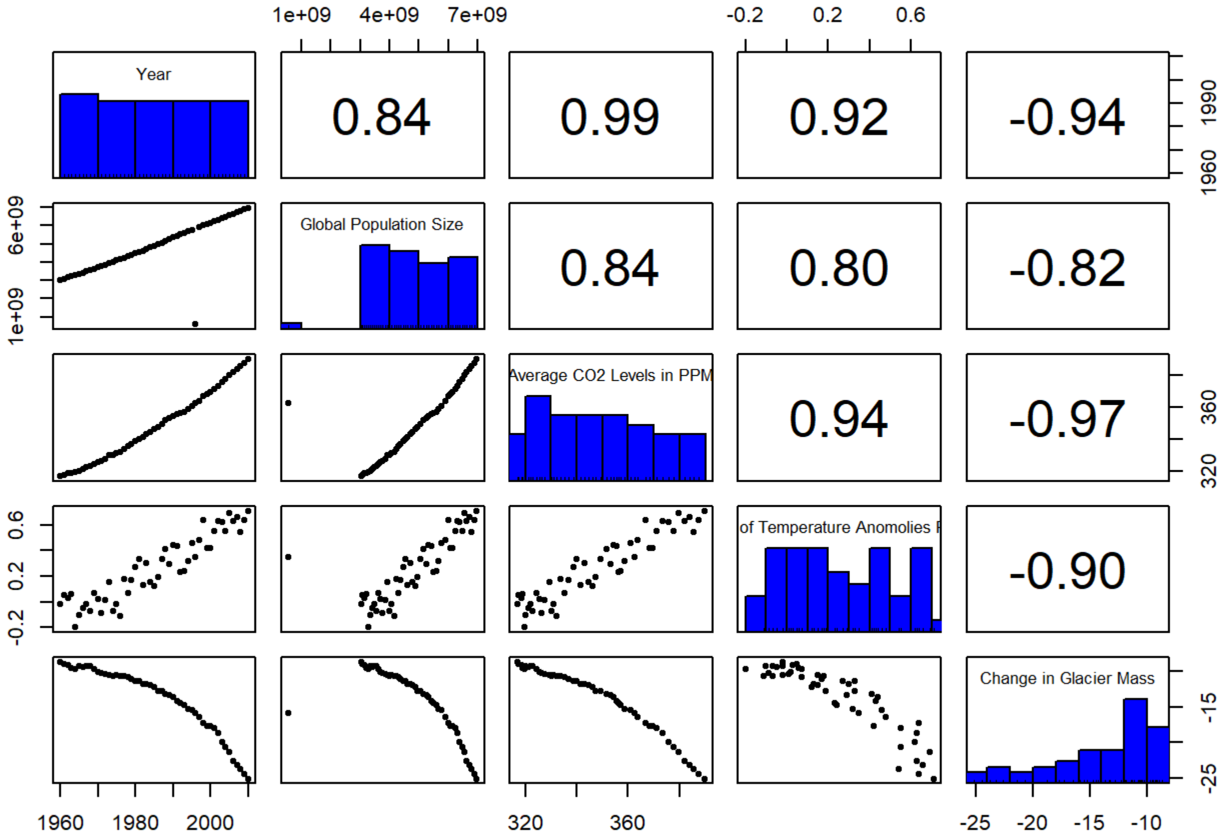
```
## # A tibble: 6 x 5
##    Year `Global Populati~ `Average CO2 Lev~ `Average Number of~ `Change in Glac~
##   <dbl>             <dbl>             <dbl>               <dbl>            <dbl>
## 1  1960        3033212527              317.               -0.02            -8.69
## 2  1961        3090305279              318.                0.05            -8.94
## 3  1962        3149244245              318.                0.03            -9.11
## 4  1963        3210271352              319.                0.06            -9.57
## 5  1964        3273670772              320.               -0.2             -9.70
## 6  1965        3339592688              320.               -0.1             -9.30
```

EDA

```r
#install necessary libraries
library(psych)

#summary statistics
summary <- climate%>%group_by(Year)%>%summarise_if(is.numeric,funs(mean,median,sd,IQR,min,max))

#view
head(summary)
```

```
## # A tibble: 6 x 25
##    Year `Global Populat~ `Average CO2 Le~ `Average Number~ `Change in Glac~
##   <dbl>            <dbl>            <dbl>            <dbl>            <dbl>
## 1  1960       3033212527             317.            -0.02            -8.69
## 2  1961       3090305279             318.             0.05            -8.94
## 3  1962       3149244245             318.             0.03            -9.11
## 4  1963       3210271352             319.             0.06            -9.57
## 5  1964       3273670772             320.            -0.2             -9.70
## 6  1965       3339592688             320.            -0.1             -9.30
## # ... with 20 more variables: `Global Population Size_median` <dbl>, `Average
## #   CO2 Levels in PPM_median` <dbl>, `Average Number of Temperature Anomolies
## #   Relative to Base 0_median` <dbl>, `Change in Glacier Mass_median` <dbl>,
## #   `Global Population Size_sd` <dbl>, `Average CO2 Levels in PPM_sd` <dbl>,
## #   `Average Number of Temperature Anomolies Relative to Base 0_sd` <dbl>,
## #   `Change in Glacier Mass_sd` <dbl>, `Global Population Size_IQR` <dbl>,
## #   `Average CO2 Levels in PPM_IQR` <dbl>, `Average Number of Temperature
## #   Anomolies Relative to Base 0_IQR` <dbl>, `Change in Glacier
## #   Mass_IQR` <dbl>, `Global Population Size_min` <dbl>, `Average CO2 Levels in
## #   PPM_min` <dbl>, `Average Number of Temperature Anomolies Relative to Base
## #   0_min` <dbl>, `Change in Glacier Mass_min` <dbl>, `Global Population
## #   Size_max` <dbl>, `Average CO2 Levels in PPM_max` <dbl>, `Average Number of
## #   Temperature Anomolies Relative to Base 0_max` <dbl>, `Change in Glacier
## #   Mass_max` <dbl>
```

```r
#correlation matrix for all numeric variables with univariate and bivariate graphs
pairs.panels(climate, method = "pearson", hist.col = "blue", smooth = FALSE, density = FALSE, el
lipses = FALSE)
```



Looking at the summary statitics that were created based on years, you can see that for the most part, all variables (population size, Average CO2 in PPM, etc.) continuously increased throughout the years. This would be a clear indication of climate change that is occurring

Looking at the correlation plot, you can see that ther is a very strong correlation (<0.8) for all of the variables regardless of whether or not the correlation is a positive or negative correlation.

MANOVA

```
#group each individual years into decades
Climate_category <- climate%>%mutate(Decades = case_when(Year<1970 ~ "Sixties", Year>1969 & Year
<1980 ~ "Seventies", Year>1979 & Year<1990 ~ "Eighties", Year>1989 & Year<2000 ~ "Nineties", Yea
r>1999 & Year<=2010 ~ "Noughties"))

#MANOVA: Perform manova test on categorical variable decade on the following numerical variable
s: Global Population Size, Average CO2 Levels in PPM, Average Number of Temeprature Anomolies Re
lative to Base 0, and Change in Glacier Mass
manova <- manova(cbind(`Global Population Size`, `Average CO2 Levels in PPM`, `Average Number of
Temperature Anomolies Relative to Base 0`, `Change in Glacier Mass`) ~ Decades, data = Climate_c
ategory)

#view MANOVA
summary(manova)
```

```
##           Df Pillai approx F num Df den Df   Pr(>F)    
## Decades    4 1.7495   8.9398     16    184 5.28e-16 ***
## Residuals 46                                           
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#MANOVA resulted in significant result. Perform an ANOVA for each response variable
summary.aov(manova)
```

```
##  Response Global Population Size :
##             Df     Sum Sq    Mean Sq F value    Pr(>F)    
## Decades      4 6.2774e+19 1.5694e+19  27.546 1.062e-11 ***
## Residuals   46 2.6207e+19 5.6973e+17                      
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
##  Response Average CO2 Levels in PPM :
##             Df Sum Sq Mean Sq F value    Pr(>F)    
## Decades      4  23226  5806.5  252.46 < 2.2e-16 ***
## Residuals   46   1058    23.0                      
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
##  Response Average Number of Temperature Anomolies Relative to Base 0 :
##             Df  Sum Sq Mean Sq F value    Pr(>F)    
## Decades      4 2.83504 0.70876  71.023 < 2.2e-16 ***
## Residuals   46 0.45904 0.00998                      
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
##  Response Change in Glacier Mass :
##             Df  Sum Sq Mean Sq F value    Pr(>F)    
## Decades      4 1000.17 250.042  126.83 < 2.2e-16 ***
## Residuals   46   90.69   1.971                      
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The MANOVA test came out to be significant (p-value: 5.28e-16). As such, an univariate ANOVA test was performed to see if there was a mean difference across groups relative to Decades. Accordingly, there was a significant mean difference across groups for response in global population size (p-value: 1.062e-11), Average CO2 Levels in PPM (p-value: 2.2e-16), Average Number of Temperature Anomolies Relative to Base 0 (p-value: 2.2e-16), and Change in Glacier Mass(p-value: 2.2e-16)

```
#ANOVA was significant. Perform post-hoc t-tests
pairwise.t.test(Climate_category$`Global Population Size`, Climate_category$Decades, p.adj = "no
ne")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Climate_category$`Global Population Size` and Climate_category$Decades
##
##           Eighties Nineties Noughties Seventies
## Nineties  0.31516  -        -         -
## Noughties 4.8e-06  0.00015  -         -
## Seventies 0.02237  0.00149  1.2e-09   -
## Sixties   4.5e-05  1.5e-06  7.9e-13   0.03715
##
## P value adjustment method: none
```

```
pairwise.t.test(Climate_category$`Average CO2 Levels in PPM`, Climate_category$Decades, p.adj =
"none")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Climate_category$`Average CO2 Levels in PPM` and Climate_category$Decades
##
##           Eighties Nineties Noughties Seventies
## Nineties  1.1e-08  -        -         -
## Noughties < 2e-16  6.6e-12  -         -
## Seventies 1.5e-08  < 2e-16  < 2e-16   -
## Sixties   1.7e-15  < 2e-16  < 2e-16   1.1e-05
##
## P value adjustment method: none
```

```
pairwise.t.test(Climate_category$`Average Number of Temperature Anomolies Relative to Base 0`, C
limate_category$Decades, p.adj = "none")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Climate_category$`Average Number of Temperature Anomolies Relative to Base 0` and Clim
ate_category$Decades
##
##           Eighties Nineties Noughties Seventies
## Nineties  0.0017   -        -         -
## Noughties 2.4e-10  2.9e-05  -         -
## Seventies 1.1e-05  1.1e-10  < 2e-16   -
## Sixties   1.4e-07  1.8e-12  < 2e-16   0.2164
##
## P value adjustment method: none
```

```
pairwise.t.test(Climate_category$`Change in Glacier Mass`, Climate_category$Decades, p.adj = "no
ne")
```

```
##
##  Pairwise comparisons using t tests with pooled SD
##
## data:  Climate_category$`Change in Glacier Mass` and Climate_category$Decades
##
##           Eighties Nineties Noughties Seventies
## Nineties  4.7e-06  -        -         -
## Noughties < 2e-16  1.3e-12  -         -
## Seventies 0.0095   4.3e-10  < 2e-16   -
## Sixties   2.1e-05  5.2e-13  < 2e-16   0.0481
##
## P value adjustment method: none
```

In total 37 total tests were run (1 MANOVA, 4 ANOVAs, and (8 * 4 = 32) post-hoc t-tests). The chance of at least 1 type I error was 85.01097% ((1-0.95^37)*100). As such, a bonferroni correction shoud be applied resulting in a new alpha of 0.001351351 (0.05/37).

After the bonferroni correction, most of the post-hoc t-test are significant. As such, I will discuss which groups did not result in significance. There was no significance between the eighties and nineties (p-value: 0.31516), eighties and seventies (p-value: 0.02237), and seventies and sixties (p-value: 0.03715) group regarding Global Population Size. There was no significance between the eighties and nineties (p-value: 0.0017) and seventies and sixties (p-value: 0.2164) group regarding the Average Number of Temperature Anomolies Relative to Base 0. Lastly, there was no significance between the eighties and seventies (p-value: 0.0095) and seventies and sixties (p-value: 0.0481) group regarding Change in Glacier Mass.

There are five different assumptions that must be met for MANOVA: They are, multivariate normal distribution, same variance/covariance, sensitivity to multicollinearity, more samples than variables, highly senseitive to many 0's and outliers. Accordingly, it can be assumed that all five of these assumptions were most likely not met. MANOVA assumptions are extraordinarily restrictive and are often violated.

Randomization Test

```
#running ANOVA to determine F-statistic
summary(aov(`Average CO2 Levels in PPM` ~ Decades, data = Climate_category))
```

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## Decades      4  23226    5806   252.5 <2e-16 ***
## Residuals   46   1058      23
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#Observed F-statistic
Obs_F <- 252.5

#set seed
set.seed(348)
#Replication test 5000 times
Fs <- replicate(5000,{
# Randomly permute the response variable across decades
    new <- Climate_category %>%
        mutate(PPM = sample(`Average CO2 Levels in PPM`))
# Compute variation within groups
    SSW <- new %>%
        group_by(Decades) %>%
        summarize(SSW = sum((PPM - mean(PPM))^2)) %>%
        summarize(sum(SSW)) %>%
        pull
# Compute variation between groups
    SSB <- new %>%
        mutate(mean = mean(PPM)) %>%
        group_by(Decades) %>%
        mutate(groupmean = mean(PPM)) %>%
        summarize(SSB = sum((mean - groupmean)^2)) %>%
        summarize(sum(SSB)) %>%
        pull
# Compute the F-statistic
  # df for SSB is 5 groups - 1 = 4
  # df for SSW is 51 observations - 5 groups = 57
    (SSB/4)/(SSW/46)
})

#Determines the proportion of F statistic that was greater than the observed F-statistic
mean(Fs>Obs_F)
```

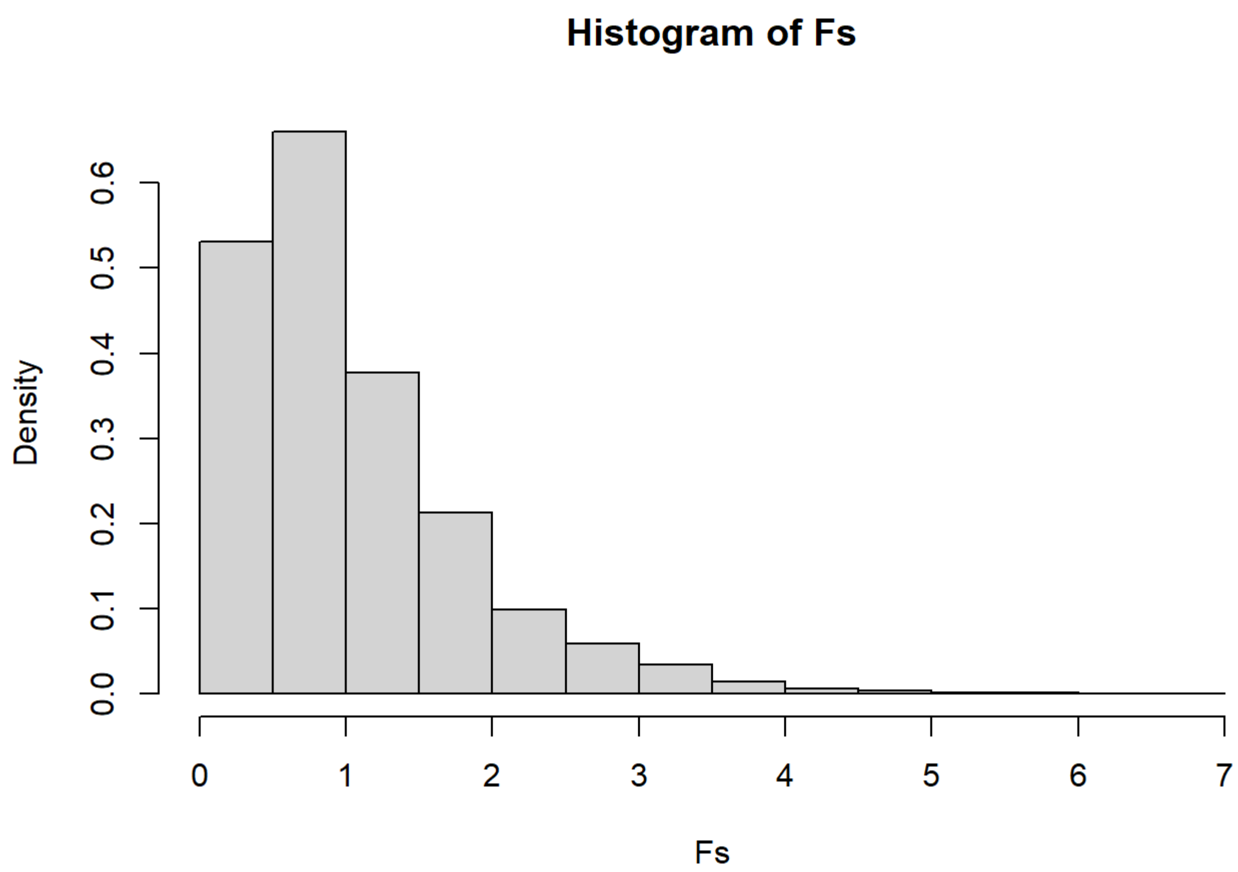
```
## [1] 0
```

Null Hypothesis: The five groups (sixties, seventies, eighties, nineties, noughties) have equal mean Average CO2 Levels in PPM

Alternate Hypothesis: The five groups (sixties, seventies, eighties, nineties, noughties) do not have equal mean Average CO2 Levels in PPM

The p-value was 2e-16 so we reject the null hypothesis. The mean average CO2 levels in PPM do differ between the five groups (sixties, seventies, eighties, nineties, noughties).

```
#plot to visualize the null distribution and the test statistic
hist(Fs, prob=T); abline(v = Obs_F, col="red",add=T)
```



Linear Regression Model

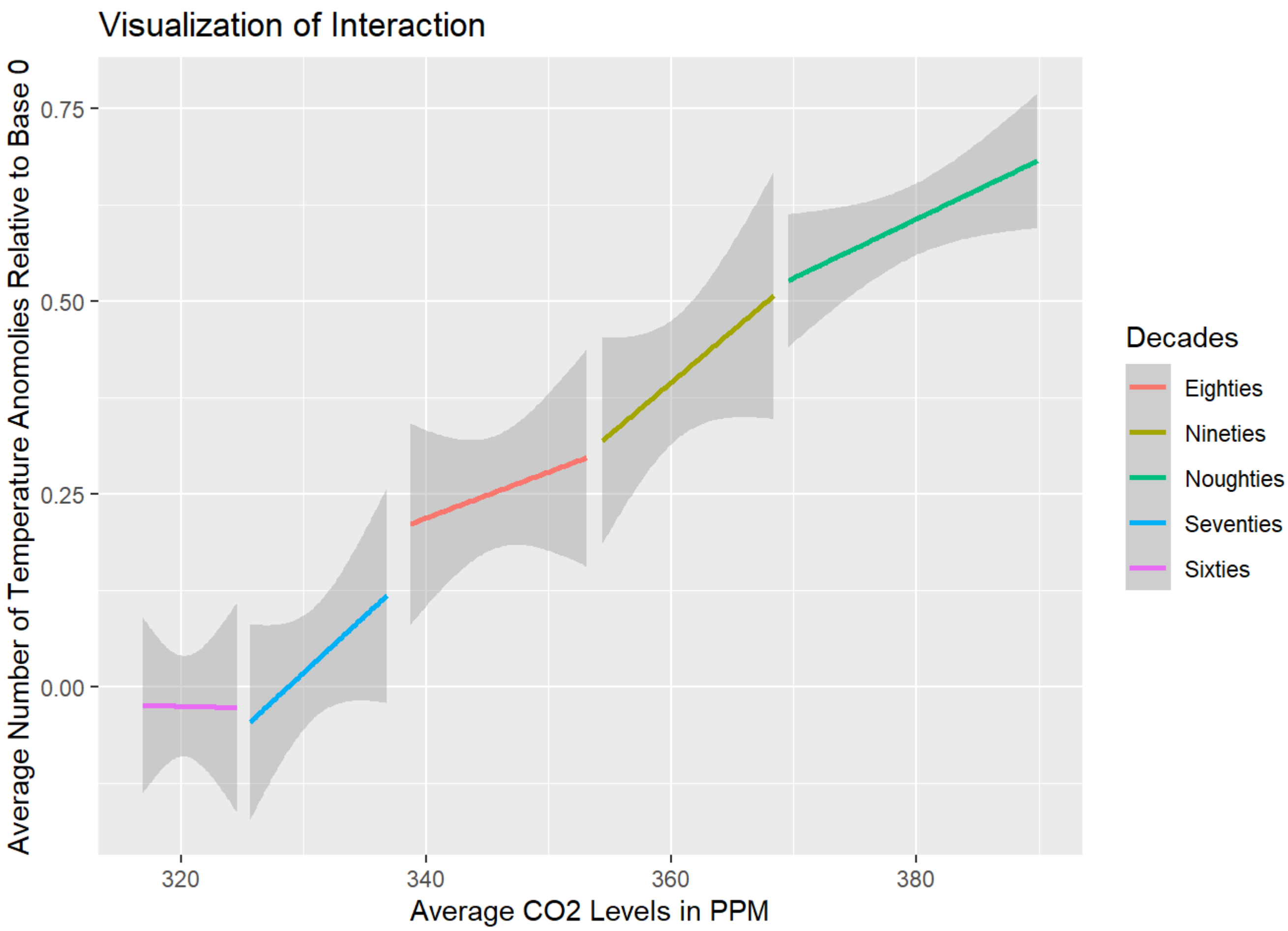
```
#obtain necessary libraries
install.packages("sandwich", repos = "http://cran.us.r-project.org")
```

```
## package 'sandwich' successfully unpacked and MD5 sums checked
##
## The downloaded binary packages are in
##  C:\Users\sungm\AppData\Local\Temp\RtmpWk29Rn\downloaded_packages
```

```
library(sandwich)
library(lmtest)

#mean-center numeric variables involved in interaction (Average CO2 Level in PPM)
centered_PPM <- (Climate_category$`Average CO2 Levels in PPM` - mean(Climate_category$`Average C
O2 Levels in PPM`, na.rm = TRUE))

#performing Linear Regression:
fit1 <- lm(`Average Number of Temperature Anomolies Relative to Base 0` ~ centered_PPM * Decade
s, data = Climate_category)

#view fit1
summary(fit1)
```

```
##
## Call:
## lm(formula = `Average Number of Temperature Anomolies Relative to Base 0` ~
##     centered_PPM * Decades, data = Climate_category)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.175262 -0.074157  0.008874  0.064022  0.155034
##
## Coefficients:
##                                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)                      0.266582   0.033547   7.947 7.86e-10 ***
## centered_PPM                     0.005976   0.006458   0.925    0.360
## DecadesNineties                 -0.033553   0.092652  -0.362    0.719
## DecadesNoughties                 0.095509   0.145244   0.658    0.514
## DecadesSeventies                 0.016571   0.148918   0.111    0.912
## DecadesSixties                  -0.302479   0.352529  -0.858    0.396
## centered_PPM:DecadesNineties     0.007485   0.009163   0.817    0.419
## centered_PPM:DecadesNoughties    0.001662   0.007802   0.213    0.832
## centered_PPM:DecadesSeventies    0.008744   0.010510   0.832    0.410
## centered_PPM:DecadesSixties     -0.006369   0.014180  -0.449    0.656
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.09365 on 41 degrees of freedom
## Multiple R-squared:  0.8908, Adjusted R-squared:  0.8669
## F-statistic: 37.18 on 9 and 41 DF,  p-value: < 2.2e-16
```

```
#create visualization of interaction between the two variables on the response
ggplot(Climate_category, aes(`Average CO2 Levels in PPM`,`Average Number of Temperature Anomolie
s Relative to Base 0`, color = Decades)) + geom_smooth(method = "lm") + labs (title = "Visualiza
tion of Interaction")
```



Interpreting the Coefficients:

Intercept: The predicted average number of temperature anomolies relative to base 0 for an average CO2 level of 0 is 0.266582 above the base temperature anomoly of 0.

centered_PPM: For every 1 ppm increase of Average CO2 levels, the predicted number of temperature anomolies relative to base 0 will increase by 0.005976.

DecadesNineties: The nineties average CO2 levels in PPM will have a predicted -0.033553 times less temperature anomolies relative to base 0 than the eighties.

DecadeNoughties: The noughties average CO2 levels in PPM will have a predicted 0.095509 times more temperature anomolies relative to base 0 than the eighties.

DecadeSeventies: The seventies average CO2 levels in PPM will have a predicted 0.016571 times more temperature anomolies relative to base 0 than the eighties.

DecadeSixties: The sixties average CO2 levels in PPM will have a predicted -0.302479 times less temperature anomolies relative to base 0 than the eighties.

Average CO2 Levels in PPM:DecadesNineties: The slope of Average CO2 levels in PPM for the nineties is 0.007485 greater than the eighties.

Average CO2 Levels in PPM:DecadesNoughties: The slope of Average CO2 levels in PPM for the noughties is 0.001662 greater than the eighties.

Average CO2 Levels in PPM:DecadesSeventies: The slope of Average CO2 levels in PPM for the seventies is 0.008744 greater than the eighties.

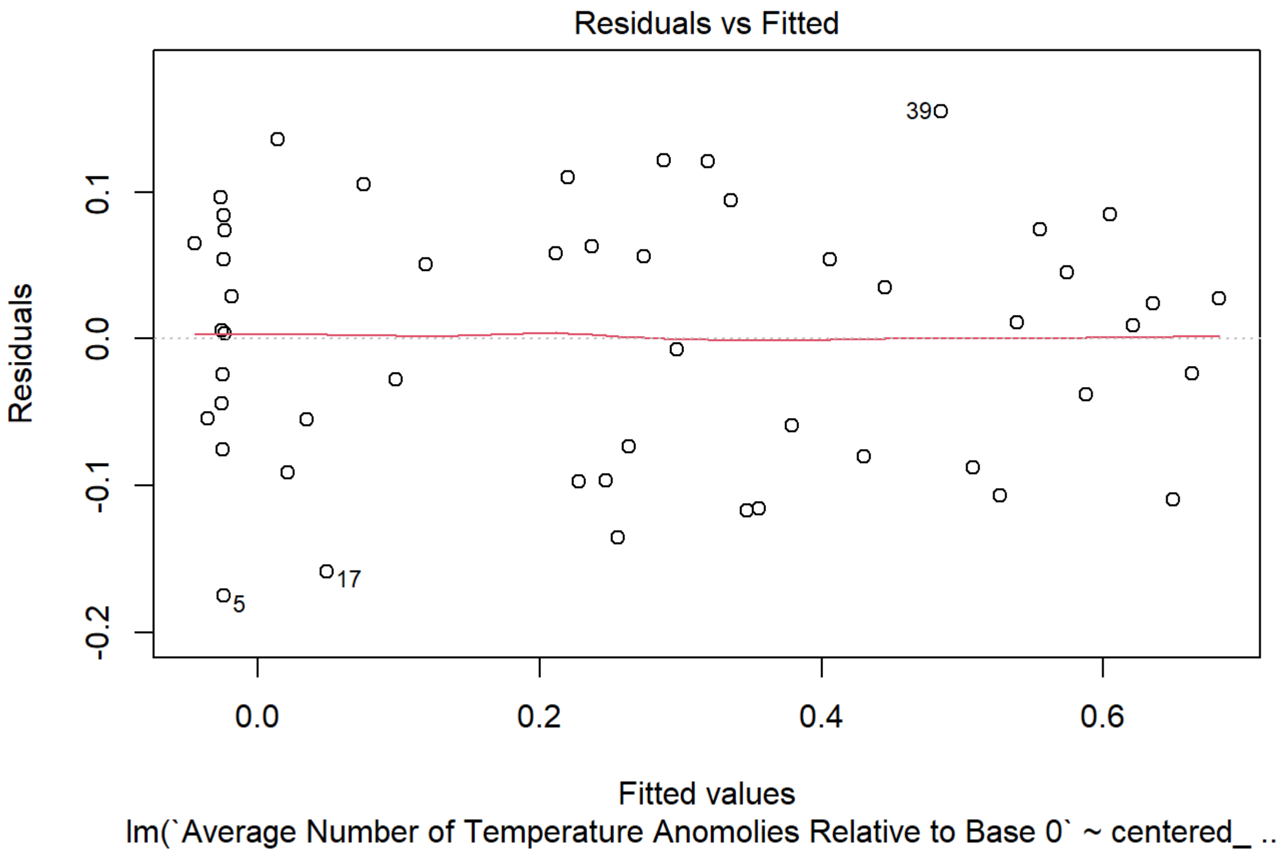
Average CO2 Levels in PPM:DecadesSixties: The slope of Average CO2 levels in PPm for the sixties is -0.006369 less than the eighties.

```
#proportion of variance explained by the model
1-summary(fit1)$r.sq
```

```
## [1] 0.1091648
```

The present model explains 0.1091648 variation in response.

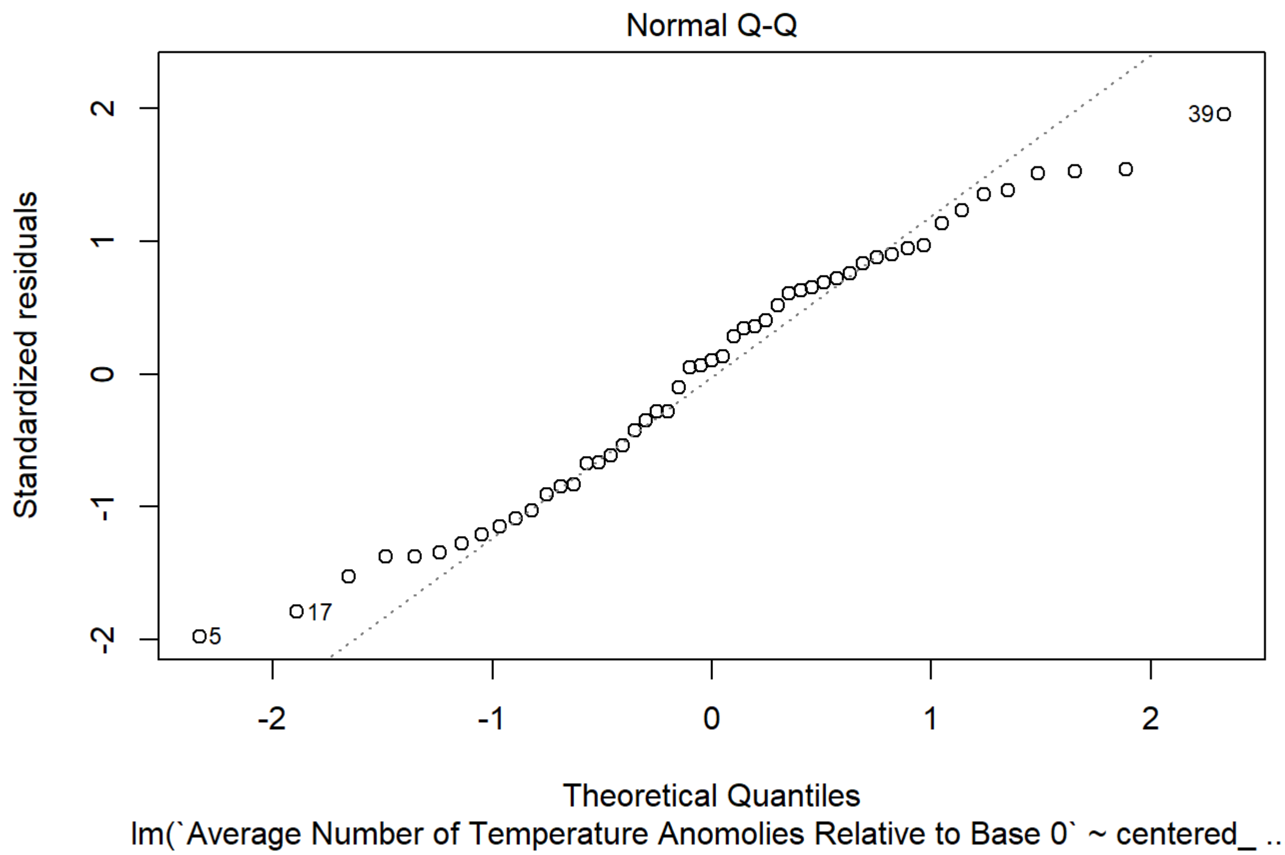
```
#linearity assumption
plot(fit1, which = 1)
```




```
#homoscedasticity
bptest(fit1)
```

```
##
##  studentized Breusch-Pagan test
##
## data:  fit1
## BP = 4.4696, df = 9, p-value = 0.8779
```

```
#normality
plot(fit1, which = 2)
```




```
shapiro.test(fit1$residuals)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  fit1$residuals
## W = 0.969, p-value = 0.2009
```

Linearity, homoscedasticity, and normality assumptions are met.

```
#recomputing regression with robust SE
coeftest(fit1, vcov = vcovHC(fit1))
```

```
##
## t test of coefficients:
##
##                                Estimate Std. Error t value  Pr(>|t|)
## (Intercept)                   0.2665819  0.0382117  6.9765 1.779e-08 ***
## centered_PPM                  0.0059759  0.0073138  0.8171    0.4186
## DecadesNineties              -0.0335528  0.1435977 -0.2337    0.8164
## DecadesNoughties              0.0955089  0.1480823  0.6450    0.5225
## DecadesSeventies              0.0165707  0.1367506  0.1212    0.9041
## DecadesSixties               -0.3024789  0.4325493 -0.6993    0.4883
## centered_PPM:DecadesNineties  0.0074846  0.0131678  0.5684    0.5729
## centered_PPM:DecadesNoughties 0.0016615  0.0084367  0.1969    0.8449
## centered_PPM:DecadesSeventies 0.0087440  0.0103119  0.8480    0.4014
## centered_PPM:DecadesSixties  -0.0063693  0.0168863 -0.3772    0.7080
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Estimate values are the same before and after calculating robust SE. However, the Std. Error, t-value, and p-values are slightly different before and after calculating robust SE but not by much. There was only one significant value which was the intercept coefficient (p-value: 7.86e-10 & 1.779e-10 before and after robust SE respectively)The predicted average number of temperature anomolies relative to base 0 for an average CO2 level of 0 is significant. It is 0.266582 above the base temperature anomoly of 0.

```
#Performing Linear Regression with Bootstrapped SE
#set seed
set.seed(350)
#Bootstrapping my data
samp_SEs <- replicate(5000, {
    boot_data <- sample_frac(Climate_category, replace = TRUE)
#fitting a regression model
    fit2 <- lm(`Average Number of Temperature Anomolies Relative to Base 0` ~ centered_PPM * Dec
ades, data = boot_data)
#saving coefficient values
    coef(fit2)
})

samp_SEs %>%
#Transposing the obtained matrices
  t %>%
#matrix as a data frame
  as.data.frame %>%
#Compute the standard error
  summarize_all(sd)
```

```
##   (Intercept) centered_PPM DecadesNineties DecadesNoughties DecadesSeventies
## 1  0.04108089  0.001840385      0.05898505       0.04874496       0.06144528
##   DecadesSixties centered_PPM:DecadesNineties centered_PPM:DecadesNoughties
## 1     0.05208652                           NA                   0.002226893
##   centered_PPM:DecadesSeventies centered_PPM:DecadesSixties
## 1                   0.002816333                          NA
```

p-values were not visible when computing the boostrapped SE, however, comparing to the normal SE and the robust SE, the boostrapped SE values were smaller for the most part (except the intercept) in comparison to the normal SE and robust SE.

Logistic Regression:

```
#create a dichotomous variable
Climate_binary <- Climate_category%>%mutate(`Anomoly Occurence` = case_when(`Average Number of T
emperature Anomolies Relative to Base 0`> 0 ~ "1", `Average Number of Temperature Anomolies Rela
tive to Base 0` < 0 ~ "0"))

#creates numeric variable
Climate_binary$`Anomoly Occurence` <- as.numeric(Climate_binary$`Anomoly Occurence`)

#fit a logistic regression
fit3 <- glm(`Anomoly Occurence` ~ `Average CO2 Levels in PPM` + `Change in Glacier Mass`, family
= "binomial", data = Climate_binary)

#view
coeftest(fit3)
```

```
##
## z test of coefficients:
##
##                               Estimate  Std. Error z value Pr(>|z|)
## (Intercept)                -14.2985195  54.9928110 -0.2600   0.7949
## `Average CO2 Levels in PPM`   0.0074644   0.2260361  0.0330   0.9737
## `Change in Glacier Mass`     -1.2082190   1.9594322 -0.6166   0.5375
```

```
#perform odds ratio
exp(coef(fit3))
```

```
##                 (Intercept) `Average CO2 Levels in PPM`
##                6.169243e-07                1.007492e+00
##    `Change in Glacier Mass`
##                2.987288e-01
```

Every 1 increase in Average CO2 Levels in PPM multiplies the odds of a temperature anomoly occuring by 1.007492.

Every 1 increase in Change in Glacier Mass, multiplies the odds of a temperature anomoly occuring by 2.987288e-01.

```
#predicted probabilities
Climate_binary$Prob <- predict(fit3, type = "response")

#classify as 1 or 0 based on cutoff of .5
Climate_binary$predicted <- ifelse(Climate_binary$Prob > .5, "1", "0")

#confusion matrix
table(true_condition = Climate_binary$`Anomoly Occurence`, predicted_condition = Climate_binary
$predicted) %>% addmargins
```

```
##                predicted_condition
## true_condition  0  1 Sum
##             0    6  4  10
##             1    4 37  41
##             Sum 10 41  51
```

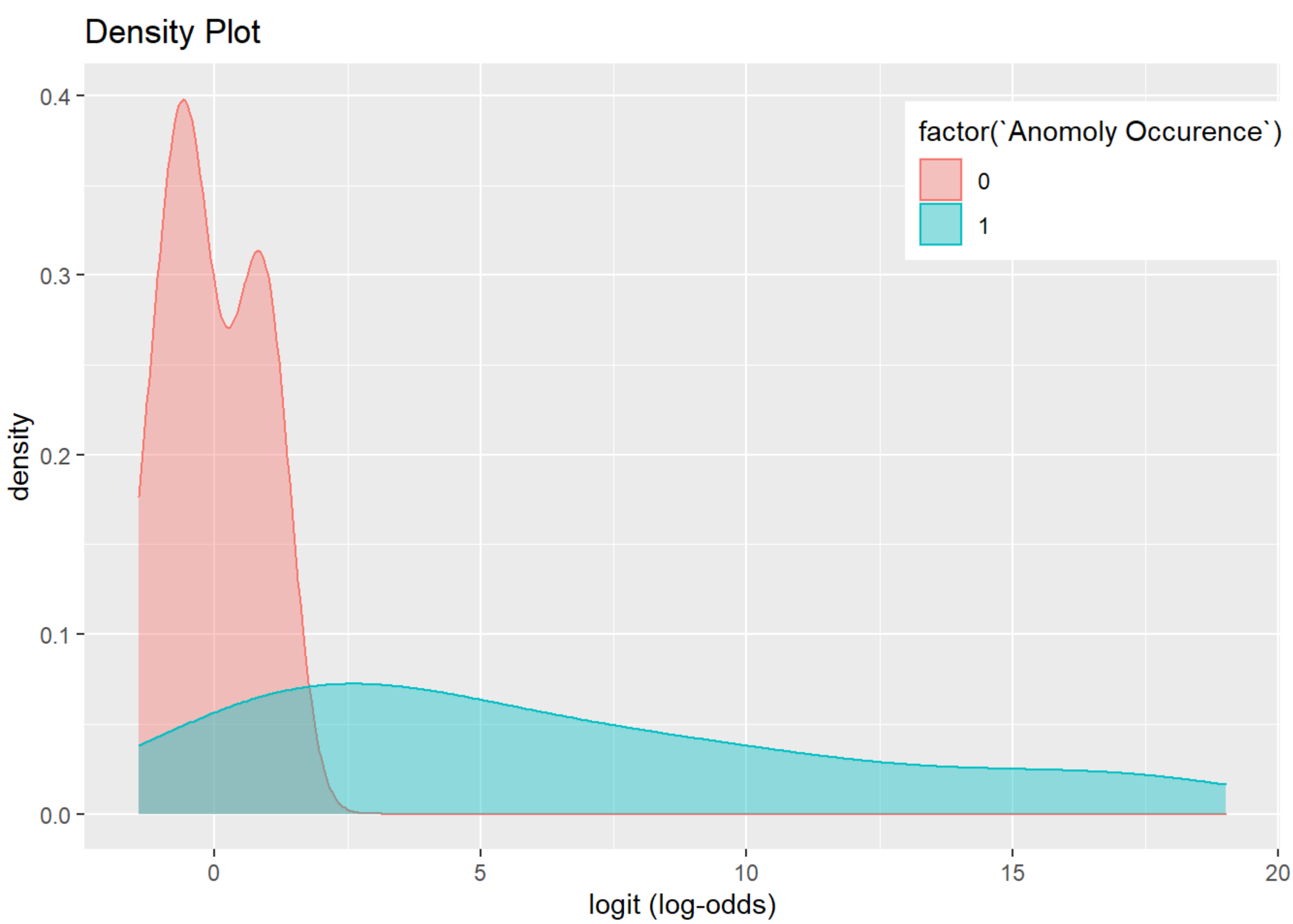
Accuracy: Refers to the proportion of correctly classified cases. In this case, it is (6 + 37)/51 = 0.8431373

Sensitivy: Refers to the proportion of true positive cases. In this case, it is predicting that a temperature anomoly occurred when it really did occur. (37/41) = 0.902439

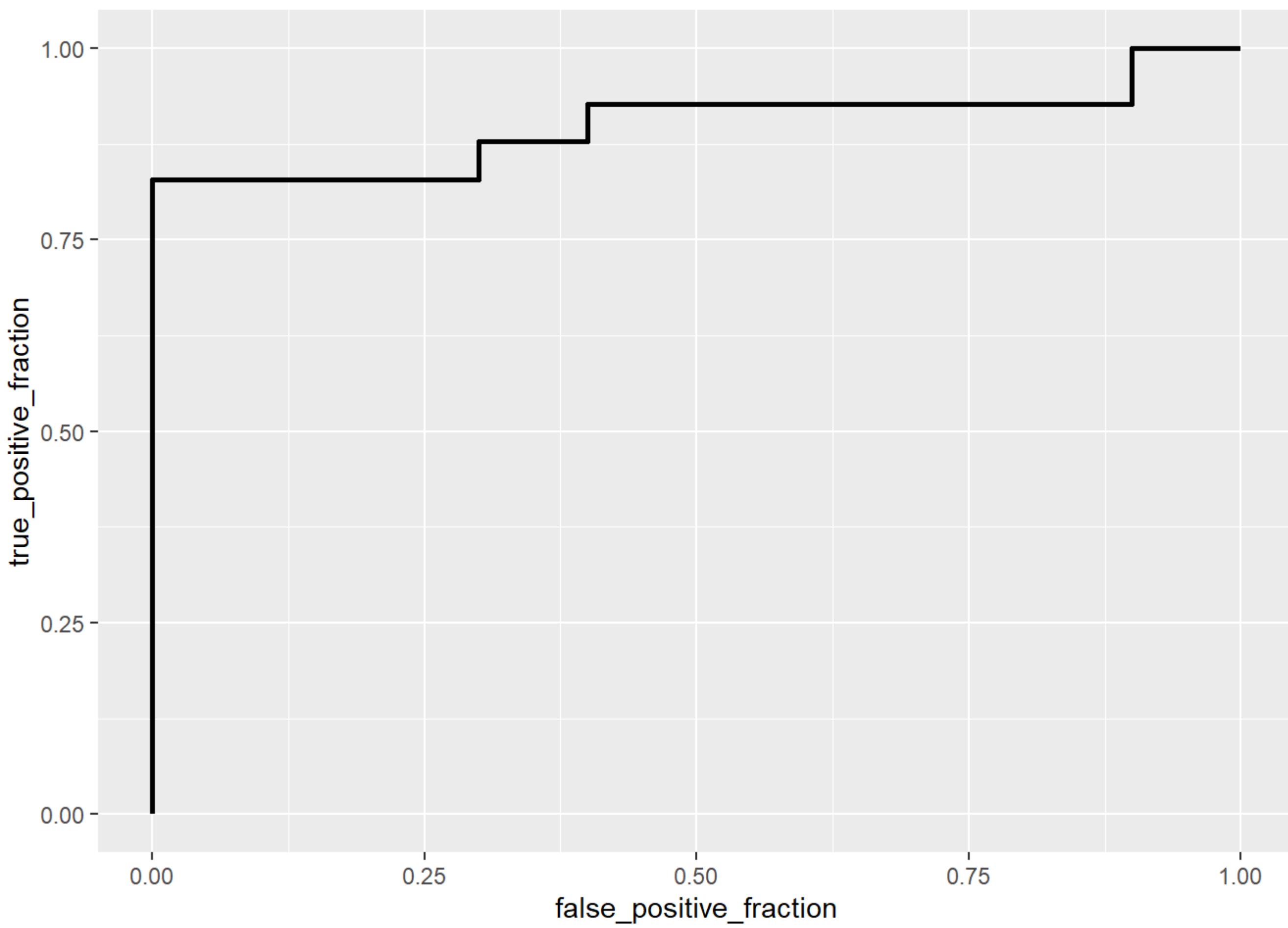
Specificity: Refers to the proportion of true negative cases. In this case, it is predicting that a temperature anomoly did not occur when it really did not occur. (6/10) = 0.6

Precision: Refers to the proportion of true positive prediction. In this case, it is predicting that a temperature anomoly occurred. (37/41) = 0.902439

```
#predicted logodds
Climate_binary$logit <- predict(fit3, type = "link")

#creating a plot density of log-odds by binary outcome variable
Climate_binary%>%ggplot(aes(logit, color = factor(`Anomoly Occurence`),fill = factor(`Anomoly Oc
curence`)))+geom_density(alpha = 0.4)+theme(legend.position = c(0.85, 0.85))+labs(title = "Densi
ty Plot") + xlab ("logit (log-odds)")
```



```r
#installing necessary libraries
library(plotROC)

#plotROC
ROCplot1 <- ggplot(Climate_binary) + geom_roc(aes(d = `Anomoly Occurence`, m = `Average CO2 Leve
ls in PPM` + `Change in Glacier Mass`), n.cuts = 0)

ROCplot1
```




```r
#Determining AUC
calc_auc(ROCplot1)
```

```
##   PANEL group AUC
## 1     1    -1 0.9
```

The AUC represents the overall measure of the model performance specifically for predicting whether or not there was a temperature anomoly from Average CO2 levels in PPM and Change in Glacier Mass. The AUC was 0.9 which makes this a good/great predictor.

```
##          sysname          release          version         nodename
##        "Windows"         "10 x64"    "build 19042" "DESKTOP-KFAHSG6"
##          machine            login             user   effective_user
##         "x86-64"          "sungm"          "sungm"          "sungm"
```