########Modelo lineal#########

###Para los tuits de fishing y hunting

hist(gra$fishing) #para ver si siguen una distribución normal

hist(gra$hunting)

plot(x, y, xlab="Time", ylab="Magnitude") #cambiar nombres gráfica

#fishing

par(mfrow = c(1,2))

linearM<-lm(sqrt(gra$fishing) ~gra$year) #hacemos el modelo lineal normalizado, para ver como evolucionan con el tiempo (si va aumentando el numero de tuits con los años, disminuyendo etc.)

plot(sqrt(gra$fishing) ~gra$year)

abline(lm(sqrt(gra$fishing) ~gra$year))

summary ##Para sacar el p-valor

#hunting

linearM2<-lm(sqrt(gra$hunting) ~gra$year)

plot(sqrt(gra$hunting) ~gra$year)

abline(lm(sqrt(gra$hunting) ~gra$year))

summary(linearM2)

###Para los social engagements

################ESTADÍSTICA SOCIAL ENGAGEMENTS########

##############HUNTING###################

###############Replies hunting

par(mfrow = c(2,3))

dat<-"year freq

2007 0

2008 0

2009 0

2010 0

2011 34

2012 168

2013 280

2014 313

2015 295

2016 290

2017 430

2018 1266

2019 1380

2020 1962

2021 3372

2022 273"

dat1<-read.table(textConnection(dat),header = TRUE, sep="\t")

str(dat1)

attach(dat1)

head(dat1)

#year freq

#1 2007 0

#2 2008 0

#3 2009 0

#4 2010 0

#5 2011 34

#6 2012 168

colnames(dat1)

#[1] "year" "freq"

linearM<-lm(dat1$freq ~dat1$year)

#Call:

lm(formula = dat1$freq ~ dat1$year)

#

#Residuals:

# Min 1Q Median 3Q Max

#-1369.31 -286.66 -71.47 185.45 1864.81

#

#Coefficients:

# Estimate Std. Error t value Pr(>|t|)

#(Intercept) -271562.60 76292.82 -3.559 0.00314 \*\*

# dat1$year 135.12 37.87 3.568 0.00309 \*\*

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#

#Residual standard error: 698.3 on 14 degrees of freedom

#Multiple R-squared: 0.4762, Adjusted R-squared: 0.4388

#F-statistic: 12.73 on 1 and 14 DF, p-value: 0.00309

plot(dat1$freq ~dat1$year)

abline(lm(dat1$freq ~dat1$year))

#Con normalization

linearM<-lm(sqrt(dat1$freq) ~dat1$year)

#Call:

# lm(formula = sqrt(dat1$freq) ~ dat1$year)

#

#Residuals:

# Min 1Q Median 3Q Max

#-25.0640 -3.5691 0.7737 4.2617 19.5290

#

#Coefficients:

# Estimate Std. Error t value Pr(>|t|)

#(Intercept) -6118.9945 1048.8417 -5.834 4.34e-05 \*\*\*

# dat1$year 3.0468 0.5206 5.852 4.20e-05 \*\*\*

# ---

# Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

#

#Residual standard error: 9.6 on 14 degrees of freedom

#Multiple R-squared: 0.7098, Adjusted R-squared: 0.6891

#F-statistic: 34.25 on 1 and 14 DF, p-value: 4.205e-05

plot(sqrt(dat1$freq) ~dat1$year)

abline(lm(sqrt(dat1$freq) ~dat1$year))

summary(linearM)###Para sacar el p-valor p-value: 4.205e-05

######likes hunting

fl<-"year freq

2007 0

2008 0

2009 8

2010 2

2011 14

2012 85

2013 236

2014 1053

2015 1745

2016 3150

2017 2008

2018 7250

2019 8994

2020 12465

2021 23821

2022 2096"

fl1<-read.table(textConnection(fl),header = TRUE, sep="\t")

str(fl1)

attach(fl1)

head(fl1)

colnames(fl1)

#[1] "year" "freq"

#Con normalization

linearM1<-lm(sqrt(fl1$freq) ~fl1$year)

plot(sqrt(fl1$freq) ~fl1$year)

abline(lm(sqrt(fl1$freq) ~fl1$year))

summary(linearM1) ##p-value: 3.65e-05

######retweets hunting

rh<-"year freq

2007 0

2008 0

2009 9

2010 25

2011 105

2012 354

2013 708

2014 1237

2015 1467

2016 1270

2017 1042

2018 2554

2019 3174

2020 2940

2021 3960

2022 444"

rh1<-read.table(textConnection(rh),header = TRUE, sep="\t")

str(rh1)

attach(rh1)

head(rh1)

colnames(rh1)

#[1] "year" "freq"

linearM2<-lm(rh1$freq ~rh1$year)

#Call:

lm(formula = rh1$freq ~ rh1$year)

#Con normalization

linearM2<-lm(sqrt(rh1$freq) ~rh1$year)

plot(sqrt(rh1$freq) ~rh1$year)

abline(lm(sqrt(rh1$freq) ~rh1$year))

summary(linearM2) ##p-value: 3.629e-05

##############FISHING###################

###############Replies fishing

fire<-"year freq

2007 0

2008 0

2009 0

2010 1

2011 57

2012 299

2013 346

2014 523

2015 409

2016 554

2017 534

2018 1429

2019 1501

2020 3280

2021 2214

2022 337"

fire1<-read.table(textConnection(fire),header = TRUE, sep="\t")

str(fire1)

attach(fire1)

head(fire1)

colnames(fire1)

#[1] "year" "freq"

linearM3<-lm(fire1$freq ~fire1$year)

#Call:

lm(formula = fire1$freq ~ fire1$year)

#Con normalization

linearM3<-lm(sqrt(fire1$freq) ~fire1$year)

plot(sqrt(fire1$freq) ~fire1$year)

abline(lm(sqrt(fire1$freq) ~fire1$year))

summary(linearM3) #p-value: 3.137e-05

###############Likes fishing

fili<-"year freq

2007 0

2008 0

2009 9

2010 61

2011 43

2012 239

2013 1258

2014 4610

2015 4224

2016 3567

2017 6364

2018 17220

2019 17814

2020 29380

2021 23853

2022 3572"

fili1<-read.table(textConnection(fili),header = TRUE, sep="\t")

str(fili1)

attach(fili1)

head(fili1)

colnames(fili1)

#[1] "year" "freq"

linearM4<-lm(fili1$freq ~fili1$year)

#Call:

lm(formula = fili1$freq ~ fili1$year)

#Con normalization

linearM4<-lm(sqrt(fili1$freq) ~fili1$year)

plot(sqrt(fili1$freq) ~fili1$year)

abline(lm(sqrt(fili1$freq) ~fili1$year))

summary(linearM4) ##p-value: 1.924e-05

###############Retweets fishing

refi<-"year freq

2007 0

2008 0

2009 4

2010 453

2011 242

2012 1145

2013 1842

2014 3891

2015 3751

2016 3123

2017 4062

2018 7569

2019 6497

2020 8878

2021 6840

2022 954"

refi1<-read.table(textConnection(refi),header = TRUE, sep="\t")

str(refi1)

attach(refi1)

head(refi1)

colnames(refi1)

#[1] "year" "freq"

linearM5<-lm(refi1$freq ~refi1$year)

#Call:

lm(formula = refi1$freq ~ refi1$year)

#Con normalization

linearM5<-lm(sqrt(refi1$freq) ~refi1$year)

plot(sqrt(refi1$freq) ~refi1$year)

abline(lm(sqrt(refi1$freq) ~refi1$year))

summary(linearM5) ###p-value: 9.852e-05

########ESTADÍSTICA#########

#### t-student

t.test(gra$fishing, gra$hunting) #Para el total de tuits de pesca y caza

t.test(fire1$freq , dat1$freq ) #Para replies

t.test(fili1$freq , fl1$freq ) #Para likes

t.test(refi1$freq , rh1$freq ) #Para retweets

######Boxplots

par(mfrow = c(2,2))

boxplot(cbind(fire1$freq, dat1$freq), names=c("Replies fishing", "Replies Hunting"), main = "Number of replies", col=c("cadetblue","coral3"))

boxplot(cbind(refi1$freq, rh1$freq), names=c("Retweets fishing", "Retweets Hunting"), main = "Number of retweets", col=c("cadetblue","coral3"))

boxplot(cbind(fili1$freq, fl1$freq), names=c("Likes fishing", "Likes Hunting"), main = "Number of likes", col=c("cadetblue","coral3"))

boxplot(cbind(gra$fishing, gra$hunting), names=c("Tweets fishing", "Tweets Hunting"), main = "Number of tweets", col=c("cadetblue","coral3"))