dat<-read.csv(file = "https://raw.githubusercontent.com/mattdemography/EDU\_7043/master/Data/GSS\_2016\_AA.csv", stringsAsFactors=F)

names(dat)

str(dat)

wh<-subset(dat, dat$RACE==1)

b<-subset(dat, dat$RACE==2)

o<-subset(dat, dat$RACE==3)

nw<-subset(dat, dat$RACE==2 | dat$RACE==3)

m<-subset(dat, dat$SEX==1)

f<-subset(dat, dat$SEX==2)

#mean,median and range

summary(dat$POLVIEWS)

summary(wh$POLVIEWS)

summary(b$POLVIEWS)

summary(o$POLVIEWS)

summary(nw$POLVIEWS)

summary(m$POLVIEWS)

summary(f$POLVIEWS)

#standard deviation

sd(dat$POLVIEWS,na.rm=T)

sd(wh$POLVIEWS,na.rm=T)

sd(b$POLVIEWS,na.rm=T)

sd(o$POLVIEWS,na.rm=T)

sd(nw$POLVIEWS,na.rm=T)

sd(m$POLVIEWS,na.rm=T)

sd(f$POLVIEWS,na.rm=T)

#variance

var(dat$POLVIEWS,na.rm = T)

var(wh$POLVIEWS,na.rm=T)

var(b$POLVIEWS,na.rm=T)

var(o$POLVIEWS,na.rm=T)

var(nw$POLVIEWS,na.rm=T)

var(m$POLVIEWS,na.rm=T)

var(f$POLVIEWS,na.rm=T)

#Comparing male and female political views

t.test(m$POLVIEWS, f$POLVIEWS)

wilcox.test(m$POLVIEWS, f$POLVIEWS)

#ANOVA Coparing Racial Differences in Political Views

bob<-aov(dat$POLVIEWS ~ as.factor(dat$RACE))

summary(bob)

#Tukey Test to see where the difference is

TukeyHSD(bob)

#normality

boxplot(wh$POLVIEWS, b$POLVIEWS, o$POLVIEWS)

boxplot(m$POLVIEWS,f$POLVIEWS)

dat$POLVIEWS<-ifelse(dat$POLVIEWS==9 | dat$POLVIEWS==8 | dat$POLVIEWS==0, NA, dat$POLVIEWS)

summary(dat$racepolviews)

dat$liberal<-ifelse(dat$POLVIEWS==3 | dat$POLVIEWS==2, 1,0)

dat$moderate<-ifelse(dat$POLVIEWS==4 , 1,0)

dat$conservative<-ifelse(dat$POLVIEWS==5 | dat$POLVIEWS==6 | dat$POLVIEWS==7, 1, 0)

table(dat$liberal)

table(dat$moderate)

table(dat$conservative)

dat$newpolview<-ifelse(dat$liberal==1, 1, ifelse(dat$moderate==1, 2, ifelse(dat$conservative==1, 3, 0)))

table(dat$newpolview)

summary(dat$SEX)

#contingency tables

table(dat$SEX, dat$newpolview)

ftable(dat$SEX, dat$newpolview)

table(dat$RACE, dat$newpolview)

ftable(dat$RACE, dat$newpolview)