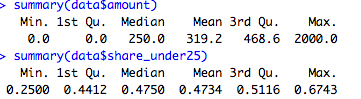
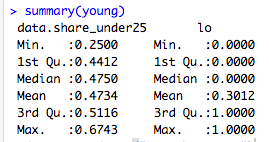
**Assignment 1 ECOM20001  
Michael Le (998211)**  
**Question 1**

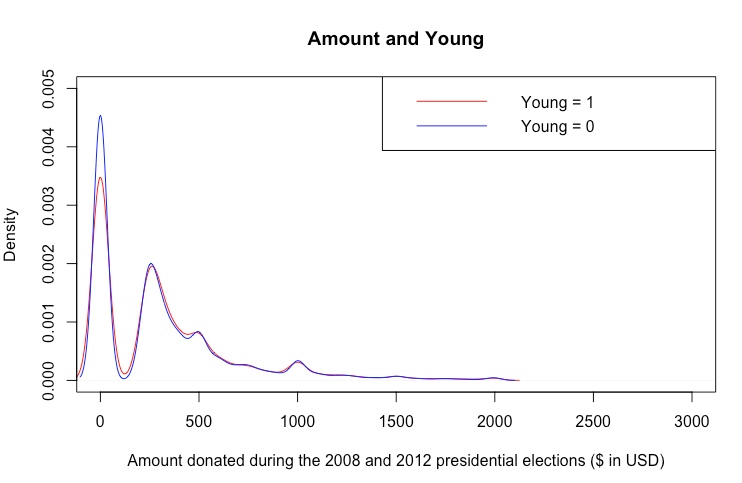


For the standard derivations for amount, share\_under25 and young is 374,97, 0.04521987 and 0.4587693 respectively.

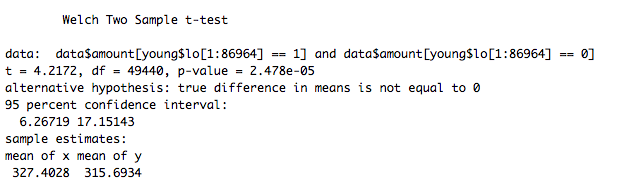
$319.20 was donated to the Democratic Party in the United States during the 2008 and 2012 elections. 47.3% of the average shares under 25 years old in the ZIP code where the donor lives. 30.1% demonstrates how youthful the demographic area/country is  
 in which compare the given donor lives. Given that the median is 0 which does not satisfy 0.5. Across the United States during the 2008 and 2012 presidential elections, which were run by Barack Obama.  
  
  
  
 **Question 2** (Given a large sample size assume it is normally distributed) however it does not specify whether it is 1 -tail?) It is a two tail

The Confidence Intervals for the respective means for the 95% Confidence Interval (2-tail) for amount, share\_under25 and young is [319.19, 319.21], [0.4668, 0.4800] and [0.294, 0.3078] respectively.

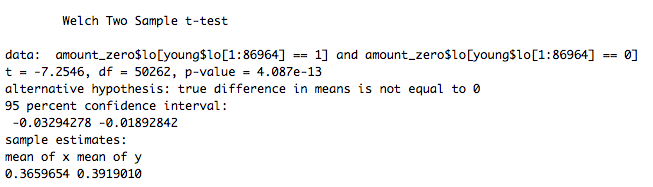
**Question 3**

****

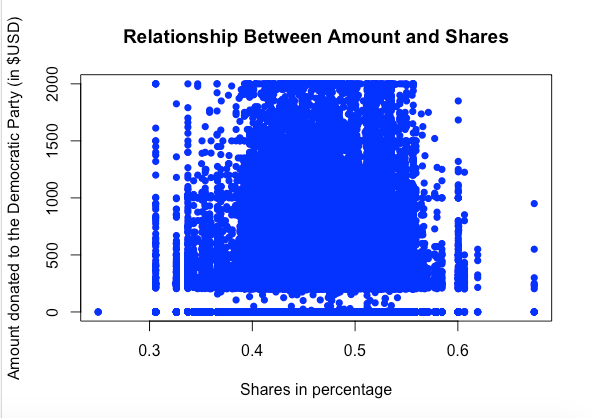
They are both bimodal. Although Young = 0 and Young = 1 looks very similar. Except exhibits the largest difference at 2 maximum points at $0 and $319.21.   
This is consistent with the above statistics that people under 25 years  
 shares below 50% are more youthful compare than the people above 50% in which a given voter donates to the Democratic Party in the United States alone.

**Question 4**

The mean difference (more to add) is 11.70931 and with the p value of 2.478 x 10^-5 (very small).The t statistic for the test of difference in means is 4.2172 > 1.96, so we reject the null at the 5% level that the difference mean is 0 given the p value is very small.  
95% confidence is [6.26719, 17.15143] means that 95% confidence that we cannot reject the amount of shares under 25 who are able to donate below $6.26 to the Democratic Party where the donor lives.

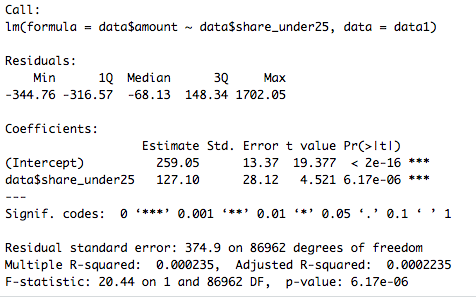
**Question 5**

Mean Difference in sample means -0.0259356 and the given p-value of 4.087 x 10^-13. T-statistic for the test difference in means is -7.2546, and |-7.2546|>1.96so we reject the null at the 5% level that the difference of means is 0, given the p-value is 4.087e-13   
95% Confidence Interval is (-0.03294278, -0.01892842) which means than 95% confidence we cannot reject the amount of zero among the young (more to add).

**Question 6**

The scatter plot does not immediately suggest a positive or negative relationship between  
amount and shares less than 25 years. There is a positive correlation between amount and shares under 25 by very little given the value of 0.01532841.

**Question 7**

****

The magnitude of the regression interprets estimate shows the initial amount   
 is at $259.05 when the shares are at 0%. Increasing share by 1% has a corresponding increase in amount of $127.10 so by increasing the shares from the average shares of 0.473354 to 0.518574 would yield a corresponding increase in amount of $1 x 0.04521987 x 127.10 = $5.75 increases.  
  
There is a slight increase relative to the average amount rate of $319.22  
 in the sample; the predicted change in amount is 100% x $5.75/$319.22 = 1.80% of the sample average. One possible economic explanation is that there want  
to increase more benefits and livings for people under 25 years of age to have fair college tuitions, costumer protection, health care, the environment, welfare and further support from Barrack Obama.

**Appendix A1**

**Question 1 R Code input**  
data = read.csv("as1\_obama.csv")  
  
#Here we are creating a binary table in a data frame  
#We first compute them into boolean values which outputs TRUE = 1 and FALSE = 0  
young = data.frame(data$share\_under25,lo = rep(c(data$share\_under25>0.5) == 1, 1))  
young$lo = as.numeric(young$lo)  
young  
  
##Summary statistics (mean, st.d, min, max) respectively   
summary(data$amount) ## (319.2,374.9672, 0, 2000.0)  
summary(data$share\_under25) ## (0.4734, 0.04521987, 0.2500, 0.6743)  
summary(young$lo) ## (0.3012, 0.4587693, 0, 1)  
  
  
#Compute the standard derviations   
sd(data$amount) # 374.9672  
sd(data$share\_under25) # 0.04521987  
sd(young$lo) # 0.4587693

**Appendix A2**

**Question 2 R Code input**  
mean\_amount = 319.2  
sd\_amount = 374.9672  
n = 86964  
error = qnorm(0.975)/sqrt(n)  
lowerbound\_amount = mean\_amount - error    
upperbound\_amount = mean\_amount + error

paste("95% CI Lower Bound:",lowerbound\_amount)  
paste("95% CI Upper Bound:",upperbound\_amount)  
  
  
## share\_under25  
mean\_share\_under25 = 0.4734  
sd\_share\_under25 = 0.04521987  
n = 86964  
error = qnorm(0.975)/sqrt(n)  
lowerbound\_share\_under25 = mean\_share\_under25 - error   
upperbound\_share\_under25 = mean\_share\_under25 + error   
paste("95% CI Lower Bound:",lowerbound\_share\_under25)  
paste("95% CI Upper Bound:",upperbound\_share\_under25)  
  
  
## and young  
mean\_young = 0.3012  
sd\_young = 0.4587693  
n  = 86964  
error = qnorm(0.975)/sqrt(n)  
lowerbound\_young = mean\_young - error   
upperbound\_young = mean\_young + error   
paste("95% CI Lower Bound:",lowerbound\_young)  
paste("95% CI Upper Bound:",upperbound\_young)

**Appendix A3**

**Question 3 R Code input**  
young = data.frame(data$share\_under25,lo = rep(c(young>0.5) == 1, 1))  
young$lo = as.numeric(young$lo)  
young$lo[1:86964]  
  
plot(density(data$amount[young$lo[1:86964]==1]), col="red",lty=1,xlim = c(0,3000),ylim = c(0,0.0050),xlab="Amount donated during the 2008 and 2012 presidential elections ($ in USD)",main="Amount and Young")  
lines(density(data$amount[young$lo[1:86964]==0]), col="blue",lty=1)  
legend("topright", legend=c("Young = 1", "Young = 0"), col=c("red", "blue"), lty=c(1,1))

**Appendix A4**

**Question 4 R code input**  
##The mean difference   
mean(data$amount[young$lo[1:86964]==1])-mean(data$amount[young$lo[1:86964]==0]) #11.70931  
## # T-test for difference of sample means is 0  
t.test(data$amount[young$lo[1:86964]==1],data$amount[young$lo[1:86964]==0])

**Appendix A5**

**Question 5 R Code input**  
#To create amount\_zero  
data$amount  
amount\_zero = data.frame(data$amount, lo = rep(c(data$amount>0) == 0,1))  
amount\_zero$lo = as.numeric(amount\_zero$lo)   
amount\_zero  
  
##mean difference   
mean(amount\_zero$lo[young$lo[1:86964]==1]) - mean(amount\_zero$lo[young$lo[1:86964]==0])  
## # T-test for difference of sample means is 0  
t.test(amount\_zero$lo[young$lo[1:86964]==1],amount\_zero$lo[young$lo[1:86964]==0])

**Appendix A6**

**Question 6 R-Code input**  
plot(data$share\_under25,data$amount,  
     main="Relationship Between Amount and Shares",  
     xlab="Shares in percentage",  
     ylab="Amount donated to the Democratic Party (in $USD)",  
     col="blue",  
     pch=1)  
  
cor(data$amount,data$share\_under25) # 0.01532841

**Appendix A7**

**Question 7 R-Code input**  
share\_reg1 = lm(data$amount~data$share\_under25, data = data1)  
summary(share\_reg1)  
sd(data$share\_under25)  
mean(data$share\_under25)  
mean(data$amount)  
  
##Interpretion results:  
## Amount hat(i) = 259.05     +   127.10SharesUnder25(i), R^2 = 0.000235, SER = 374.9    
                  #(13.37)                  (28.12)