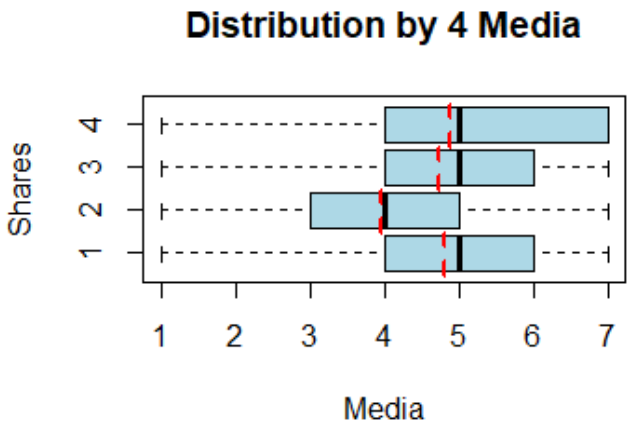
#Q1-(a) What are the means of viewers intentions to share (INTEND.0) for each media type?

```
> media1 <- read.csv("pls-media1.csv")
> media2 <- read.csv("pls-media2.csv")
> media3 <- read.csv("pls-media3.csv")
> media4 <- read.csv("pls-media4.csv")
> mean1 <- mean(media1$INTEND.0)
> mean2 <- mean(media2$INTEND.0)
> mean3 <- mean(media3$INTEND.0)
> mean4 <- mean(media4$INTEND.0)
> c(mean1, mean2, mean3, mean4)
[1] 4.809524 3.947368 4.725000 4.891304
```

#Q1-(b) Visualize the distribution and mean of intention to share, across all four media.

```
> media <- rbind(media1,media2,media3,media4)
> boxplot(INTEND.0 ~ media, data=media, main="Distribution by 4
Media", xlab="Media", ylab="Shares", col="lightblue",
border="black", horizontal = TRUE)
> mean_seg <- function(n = 1, data){
+     mean <- mean(data)
+     segments(x0 = mean, y0 = n-0.5, x1 = mean, y1 = n+0.5, col =
"red", lwd = 2, lty = "dashed")}
> mean_seg(1, media1$INTEND.0)
> mean_seg(2, media2$INTEND.0)
> mean_seg(3, media3$INTEND.0)
> mean_seg(4, media4$INTEND.0)
```



#Q1-(c) From the visualization alone, do you feel that media type makes a difference on intention to share?

Media2 seems different with others.

#Q2-(a) State the null and alternative hypotheses when comparing INTEND.0 across four groups in ANOVA

H0:the means of media1~media4 is all the same

H1:the means are not the same

#Q2-(b) Produce the traditional F-statistic for our test

```
> summary( aov(media$INTEND.0~ factor(media$media)))
                    Df Sum Sq Mean Sq F value Pr(>F)
factor(media$media)   3   22.5   7.508   2.617 0.0529 .
Residuals           162  464.8   2.869
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

#Q2-(c) What are the cut-off values of F for 95% and 99% confidence according the the null distribution of F?

```
> cut_off95 <- qf(p=0.95, df1=3, df2=length(media$INTEND.0)-4)
> cut_off99 <- qf(p=0.99, df1=3, df2=length(media$INTEND.0)-4)
> c(cut_off95, cut_off99)
[1] 2.660406 3.904807
```

#Q2-(d) According to the traditional ANOVA, do the four types of media produce the same mean intention to share, at 95% confidence? How about at 99% confidence?

Yes, F-VALUE is smaller than cut_off95 and cut_off99.

In both CI, we could agree that the four means are the same.

#Q2-(e) Do you feel the classic requirements of one-way ANOVA are met?

```
> media_share <- media$INTEND.0
> media_type <- media$media
> bartlett.test(media_share ~ media_type, data=media)

    Bartlett test of homogeneity of variances

data:  media_share by media_type
Bartlett's K-squared = 1.3958, df = 3, p-value = 0.7065
```

#P-value is big enough to prove that the variances are the same for the 4 distribution.

#Q3-(a) Bootstrap the null values of F and also the alternative values of the F-statistic.

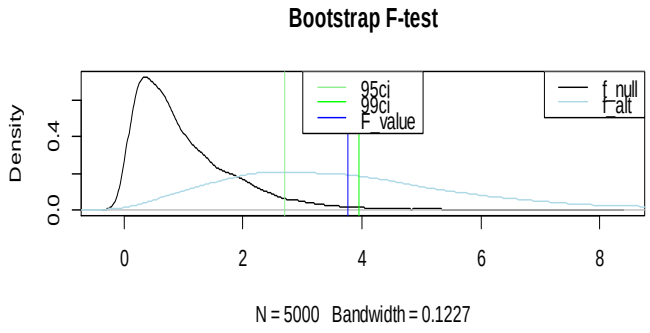
```
> boot_anova <-function(t1, t2, t3, t4, treatment){
+     null_grp1 <- sample(t1-mean(t1), replace=TRUE)
+     null_grp2 <- sample(t2-mean(t2), replace=TRUE)
+     null_grp3 <- sample(t3-mean(t3), replace=TRUE)
+     null_grp4 <- sample(t4-mean(t4), replace=TRUE)
+     null_values <- c(null_grp1, null_grp2, null_grp3,
null_grp4)
+     alt_grp1 <- sample(t1, replace=TRUE)
+     alt_grp2 <- sample(t2, replace=TRUE)
+     alt_grp3 <- sample(t3, replace=TRUE)
+     alt_grp4 <- sample(t4, replace=TRUE)
+     alt_values <- c(alt_grp1, alt_grp2, alt_grp3, alt_grp4)
+     c(oneway.test(null_values~treatment, var.equal =
TRUE)$statistic, oneway.test(alt_values~treatment, var.equal =
TRUE)$statistic)}
> share1 <- media$INTEND.0[media$media==1]
> share2 <- media$INTEND.0[media$media==2]
> share3 <- media$INTEND.0[media$media==3]
> share4 <- media$INTEND.0[media$media==4]
> f_values <- replicate(5000, boot_anova(share1, share2,
share3, share4, media$media))
> f_null <- f_values[1, ]
> f_alt <- f_values[2, ]
> c(mean(f_null), mean(f_alt))
[1] 1.019905 3.766170
```

#Q3-(b) From the bootstrapped null values of F, What are the cutoff values for 95% and 99% confidence

```
> boot_cut_off95 <- quantile(f_null, 0.95)
> boot_cut_off99 <- quantile(f_null, 0.99)
> c(boot_cut_off95, boot_cut_off99)
     95%      99%
2.703686 3.954413
```

#Q3-(c) Visualize the distribution of bootstrapped null values of F, the 95% and 99% cutoff values of F (according to bootstrap), and also the original F-value from bootstrapped alternative values

```
> plot(density(f_null), main = "Bootstrap F-test")
> lines(density(f_alt), col="lightblue")
> abline(v=boot_cut_off95, col="lightgreen")
> abline(v=boot_cut_off99, col="green")
> abline(v=mean(f_alt), col="blue")
> legend('topright',lty=1, col=c('black', 'lightblue'), legend
= c('f_null', 'f_alt'))
> legend('top',lty=1, col = c('lightgreen', 'green', 'blue'),
legend = c('95ci', '99ci', 'F_value'))
```



#Q3-(d) According to the bootstrap, do the four types of media produce the same mean intention to share, at 95% confidence? How about at 99% confidence?

F_value lies out of 95% CI, but inside 99% CI.
So, 95% CI agrees that the means are same, but 99% CI doesn't agree.