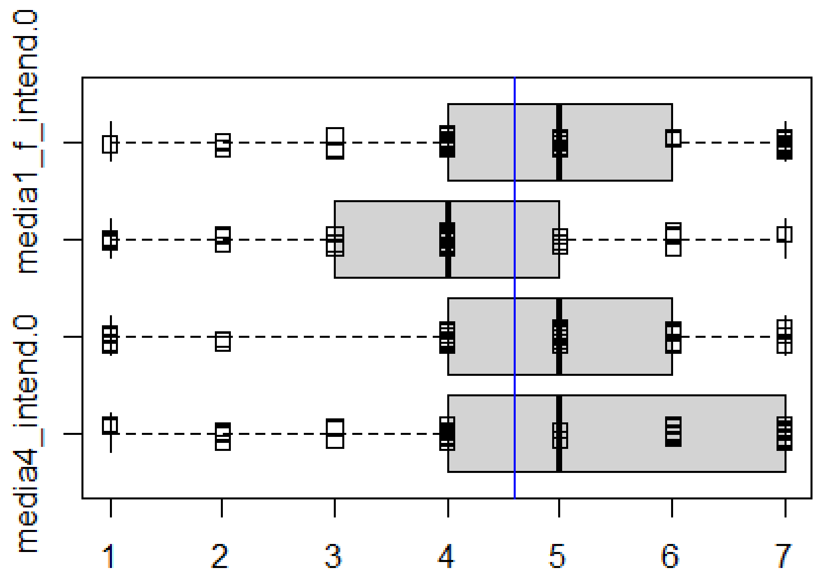
Hw7-109071512

Q1-a. The means of viewers intentions to share (INTEND.0) for each media type

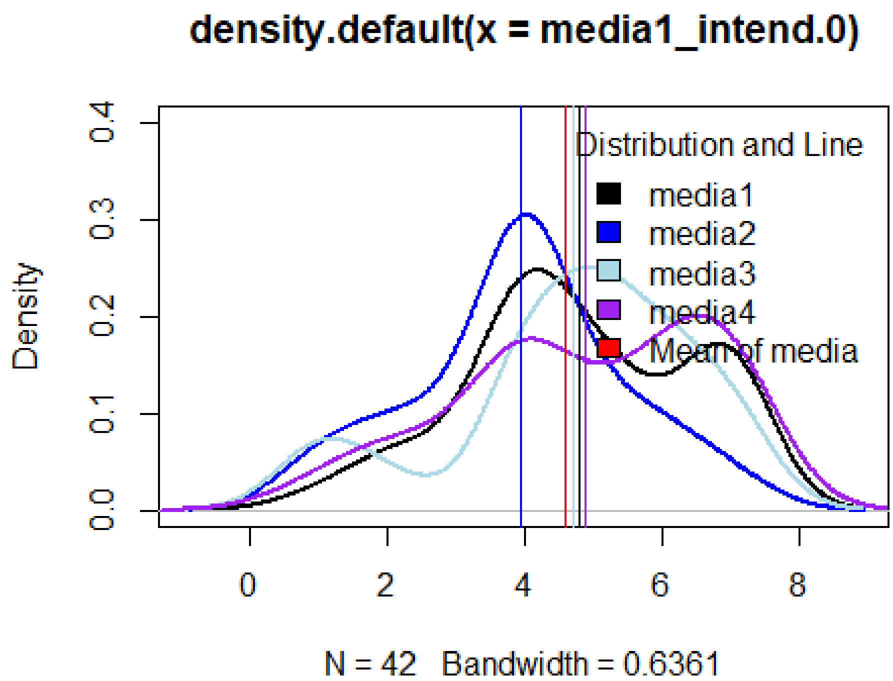
```
# read the data
media1_intend.0 <-read.csv("pls-media1.csv")$INTEND.0
media2_intend.0 <-read.csv("pls-media2.csv")$INTEND.0
media3_intend.0 <-read.csv("pls-media3.csv")$INTEND.0
media4_intend.0 <-read.csv("pls-media4.csv")$INTEND.0

# intend <-data.frame(media1,media2,media3,media4)

# compute the mean
c(mean(media1_intend.0),mean(media2_intend.0), mean(media3_intend.0), m
ean(media4_intend.0))

## [1] 4.809524 3.947368 4.725000 4.891304
```

Q1-b. Visualize the distribution and mean of intention to share

```
media1_f_intend.0 <- c(media1_intend.0, rep(NA,4))
media2_f_intend.0 <- c(media2_intend.0, rep(NA,8))
media3_f_intend.0 <- c(media3_intend.0, rep(NA,6))

media <- data.frame(media1_f_intend.0, media2_f_intend.0, media3_f_inte
nd.0, media4_intend.0)

media_mean <-c(mean(media1_intend.0),mean(media2_intend.0),mean(media3_
intend.0),mean(media4_intend.0))

#Visualize the distribution and mean by boxplot
boxplot(rev(media), horizontal = TRUE)
stripchart(rev(media),method = "jitter", add=TRUE)
abline(v=mean(media_mean), col="blue")
```




```
#Visualize the distribution by plot density
plot(density(media1_intend.0),lwd = 2,ylim=c(0,0.4))
lines(density(media2_intend.0),lwd = 2 ,col="blue")
lines(density(media3_intend.0),lwd = 2 ,col="light blue")
lines(density(media4_intend.0),lwd = 2 ,col="purple")
abline(v=mean(media1_intend.0),col='black')
abline(v=mean(media2_intend.0),col="blue")
abline(v=mean(media3_intend.0),col="light blue")
abline(v=mean(media4_intend.0),col="purple")
abline(v=mean(media_mean),col='red')
legend("topright", inset=0.02, title="Distribution and Line",legend =
   c("media1","media2","media3", "media4", "Mean of media"), bty = "n",
 fill=c("black", "blue", "light blue", "purple", "red"), horiz=FALSE)
```



Q1-c. do you feel that media type makes a difference on intention to share?

<u>Ans.</u>

As the boxplot and density figures show above, I think the media type don't affect the intention to share the information.

Q2-a.The null and alternative hypotheses

Ans.

H0 : The means of viewers intentions to share for each media type are the same.

Ha : The means of viewers intentions to share for each media type are not the same.

Q2-b.

```
#Change shape data of data to row-wise format
md1 <- data.frame(strategy=rep(1,42), methods= media1_intend.0)
md2 <- data.frame(strategy=rep2,38), methods= media2_intend.0)
md3 <- data.frame(strategy=rep(3,40), methods= media3_intend.0)
md4 <- data.frame(strategy=rep(4,46), methods= media4_intend.0)
mds <- rbind(md1, md2, md3, md4)

# Run oneway.test() for one-way ANOVA
oneway.test(mds$methods ~ mds$strategy, var.equal = TRUE)

##
##  One-way analysis of means
##
## data:  mds$methods and mds$strategy
## F = 2.6167, num df = 3, denom df = 162, p-value = 0.05289
```

Q2-c. The cut-off values of F for 95% and 99% confidence according the the null distribution of F

```
# cut-off value of F for 95%
qf(p=0.95, df1=3, df2 = 162 )

## [1] 2.660406

# cut-off value of F for 99%
qf(p=0.99, df1=3, df2 = 162)

## [1] 3.904807
```

Q2-d.

Ans.

Yes. According to the answer of ANOVA, the F-statistic is less than both 95% and 99% cut-off point. So, we can’t reject the hypothesis. It seems that the way of deliver information don’t affect the intend to share.

Q2-e. Do you feel the classic requirements of one-way ANOVA are met? Why or why not?

Ans.

No, ANOVA requires that each treatment is normally distributed, but they aren't.

Q3-a.Bootstrap the null values and alternative values of F

```
# Create the bootstrap function
boot_anova <- function(t1, t2, t3, t4, treat_nums) {
  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 ~ treat_nums, var.equal=TRUE)$statistic,
  oneway.test(alt_values ~ treat_nums, var.equal=TRUE)$statistic)
}

#Gather our treatment data and strategy numbers
methods1 = mds$methods[mds$strategy==1]
methods2 = mds$methods[mds$strategy==2]
methods3 = mds$methods[mds$strategy==3]
methods4 = mds$methods[mds$strategy==4]
strategies = mds$strategy

#Bootstrap ANOVA for null and alternative F-statistics
f_values <- replicate(5000, boot_anova(methods1, methods2, methods3, me
thods4, mds$strategy))
f_nulls <- f_values[1,]
f_alts <- f_values[2,]
```

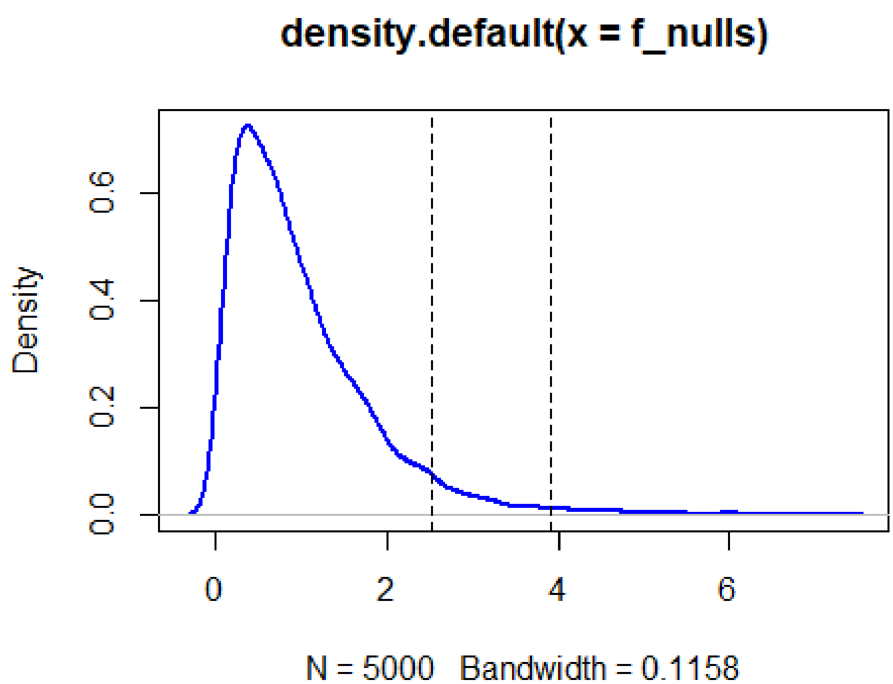
Q3-b.The cutoff values for 95% and 99% confidence

```
# Cut-off values of F for 95%
quantile(f_nulls, 0.95)

##      95%
## 2.530661

# cut-off values of F for 99%
quantile(f_nulls, 0.99)

##      99%
## 3.921832
```

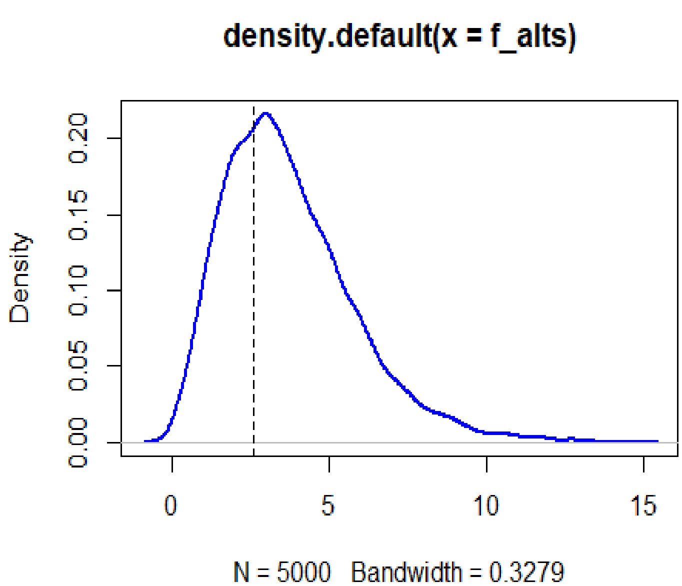
Q3-c.Visualize the distribution of bootstrapped null values and alternatiov of F

```
# Visualize the null distribution of F
plot(density(f_nulls), lwd=2,col='blue')
```

```
abline(v=quantile(f_nulls,0.95),lty="dashed")
abline(v=quantile(f_nulls,0.99),lty="dashed")
```




```
# Visualize the null distribution of F
plot(density(f_alts),lwd=2,col='blue')
anova <- oneway.test( mds$methods~ mds$strategy, var.equal=TRUE)
abline(v=anova$statistic,lty="dashed")
```



Q3-d.

```
# Value of F-statistic
anova$statistic

##        F
## 2.616669
```

<u>Ans.</u>

At 95% confidence, we can reject the hypothesis. The mean of four media is statistically significant different. However, we can't reject the hypothesis at 99% confidence.