ANOVA

Objective: To perform ANOVA technique to find if any difference exists between the analytical and experimental values of thermal conductivity for liquid state as part of Lab Migration Project.

Methods:

(i) Import and load the dataset using read_excel function.
(ii) Determine the grand mean, and mean of thermal conductivity.
(iii) Compute SSC and SSE.
(iv) Calculate degrees of freedom (between, within and total).
(v) Find MSC and MSE.
(vi) Finally get F_Statistic and F_Critical values and check if null hypothesis is accepted.
(vii) Then plot Vol. Concentration against Thermal Conductivity and also its percentage increase.
(viii) Find the error value.
(ix) Conclusion

```
#To clear the environment
rm(list=ls())

#Import the required libraries
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union

library("ggplot2")
library("readxl")

#Import and Load the dataset
data <- read_excel("liquid.xls")
```

```r
#To find grand mean, and mean of thermal conductivity
mean_analytical <- mean(data$`K analytical`)
mean_analytical

## [1] 0.1742569

mean_experimental <- mean(data$`K experimental`)
mean_experimental

## [1] 0.171125

grand_mean <- ((8*mean_analytical)+(8*mean_experimental))/16
grand_mean

## [1] 0.1726909
```

Sum of Squares between and Sum of Squares within

```r
SSC <- (8*((mean_analytical-grand_mean)^2))+(8*((mean_experimental-
grand_mean)^2))
SSC   #sum of squares between

## [1] 3.923461e-05

SSE <- sum((data$`K analytical`-mean_analytical)^2)+sum((data$`K
experimental`-mean_experimental)^2)
SSE   #sum of squares within

## [1] 0.008751421
```

Degrees of Freedom (Between, Within and total)

```r
C <- 2
N <- 16
df_between <- (C-1)
df_between

## [1] 1

df_within <- (N-C)
df_within

## [1] 14

df_total <- (N-1)
df_total

## [1] 15
```

Mean Squares (Between and Within)

```r
MSC <- SSC/df_between
MSC   #mean squares between

## [1] 3.923461e-05
```

```
MSE <- SSE/df_within
MSE   #mean squares within

## [1] 0.0006251015
```

F_Statistic and F Critical

```
F_Statistic <- MSC/MSE
F_Statistic

## [1] 0.06276519

F_Critical <- qf(p=0.05,df1=df_between,df2=df_within,lower.tail=FALSE)
F_Critical

## [1] 4.60011
```

Inference: Here F_Statistic has been calculated using the formula MSC/MSE and F_Critical is computed with fixed confidence interval of 95%.

Is Null hypothesis accepted?

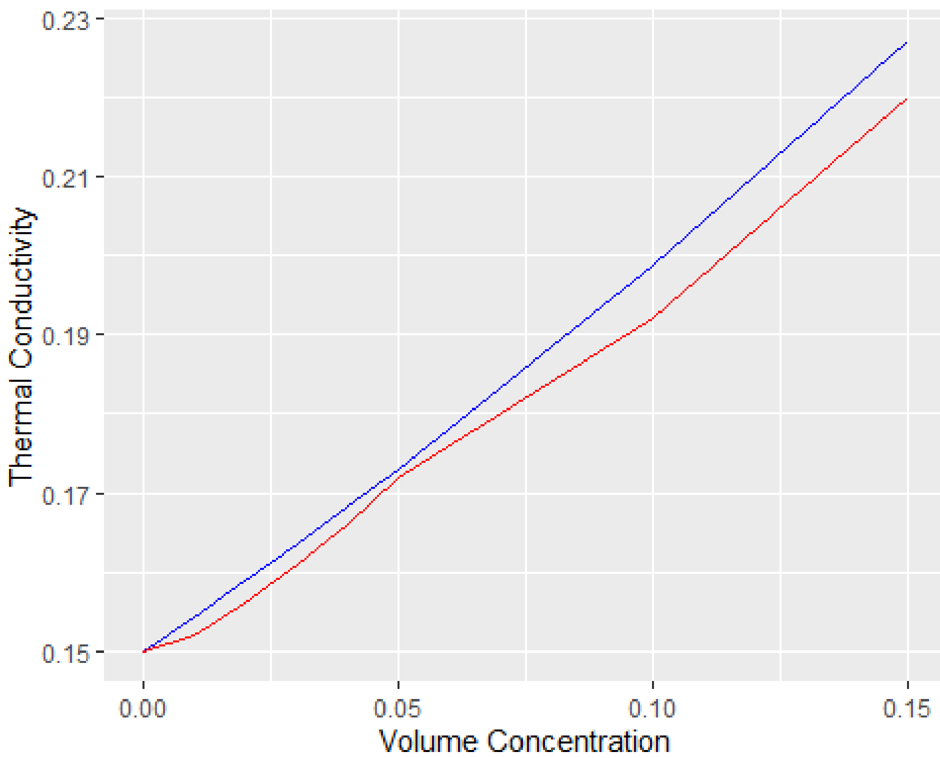
```
if(F_Statistic<F_Critical){
print("Null hypothesis is accepted")
} else {
print("Null hypothesis is rejected")
}

## [1] "Null hypothesis is accepted"
```

Using the ANOVA test we assume,

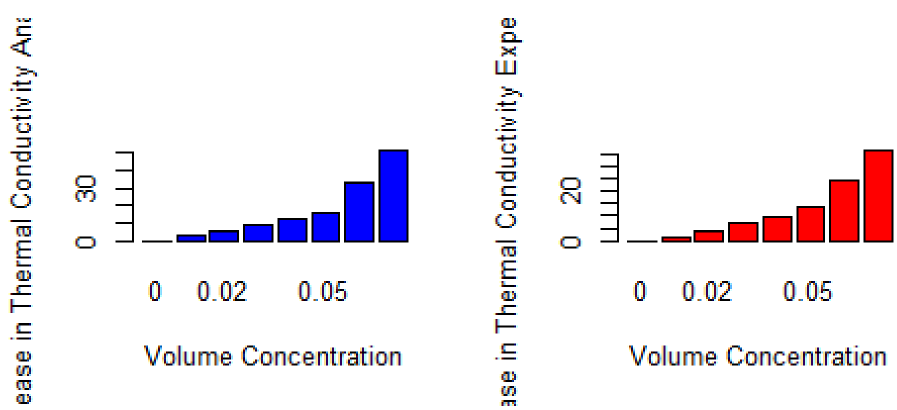
> Null hypothesis: All means are equal
> Alternative hypothesis: At least one mean is different from the other

Inference: We can infer that F_Statistic<F_Critical, therefore the null hypothesis is accepted. The means of analytical and experimental observations are very close to overall mean and/or distribution "melt" together.

```
#To plot Vol. Concentration VS Thermal Conductivity
ggplot()+geom_line(aes(x=data$`Vol Concentration`,y=data$`K analytical`),
color="blue")+geom_line(aes(x=data$`Vol Concentration`,y=data$`K
experimental`),color="red")+xlab('Volume Concentration')+ylab('Thermal
Conductivity')
```




```
#To plot Vol Concentration Vs % increase in Thermal conductivity
par(mfrow=c(2,2))
barplot(data$`%Increase in K analytical`, names.arg = data$`Vol
Concentration`, xlab ="Volume Concentration", ylab ="Increase in Thermal
Conductivity Analytical",col="blue")
barplot(data$`%Increase in K experimental`, names.arg = data$`Vol
Concentration`, xlab ="Volume Concentration", ylab ="Increase in Thermal
Conductivity Experimental",col="red")
```



Inference: Here the distributions of analytical and experimental values are more or less similar.

Error (Measured Value-Analytical Value).

```
sum(data$`K experimental`-data$`K analytical`)    #sum of error
```

```
## [1] -0.02505502

s <- sum((data$`K experimental`-data$`K analytical`)^2)
s/8   #mean squared error

## [1] 1.529197e-05
```

Inference: We can see that the mean squared error computed is very low, which is ideal.

Conclusion:

Using the ANOVA technique, we see that F_Statistic value is less than F_Critical value. Therefore, the null hypothesis is accepted.

It means there is no significant difference in mean values of analytical and experimental observations of thermal conductivity of liquids.

The mean squared error computed is 1.529197e-05 which is very low and hence indicates a better fit.