

Proyecto_completo

Larissa Cordero

2023-10-24

Resultados muestras vid Perú

Análisis: Actividad Enzimática, Conteo total de hongos y Bacterias.

```
library(dplyr)
library(tidyverse)
library(car)
library(agricolae)
library(latexpdf)
library(tinytex)
library(devtools)
devtools::install_github('yihui/tinytex')
```

Librerías

```
##
## -- R CMD build -----
##      checking for file 'C:\Users\lcordero\AppData\Local\Temp\RtmpuSo2l5\remotes40d466b46352\rstu
##      - preparing 'tinytex': (801ms)
##      checking DESCRIPTION meta-information ...      checking DESCRIPTION meta-information ...    v  chec
##      - checking for LF line-endings in source and make files and shell scripts (353ms)
##      - checking for empty or unneeded directories
##      - building 'tinytex_0.48.2.tar.gz'
##
##
```

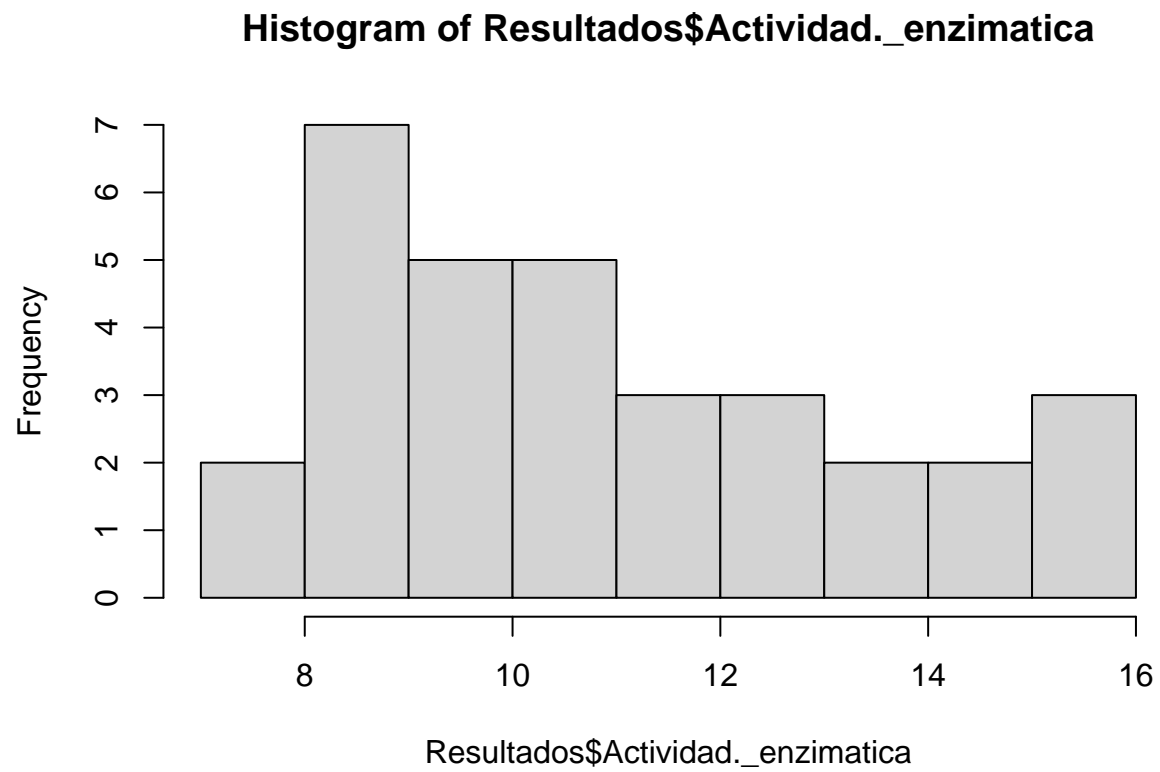
```
## Warning: package 'tinytex' is in use and will not be installed
```

```
Resultados <- read.csv("~/capR/curso/curso_Innovak/Proyecto/Datos/Muestras_peru.csv")
Estanque_de_plantas <- read.csv("~/capR/curso/curso_Innovak/Material_clase/Estanque_plantas.csv")
```

Datos *Verificación de datos*

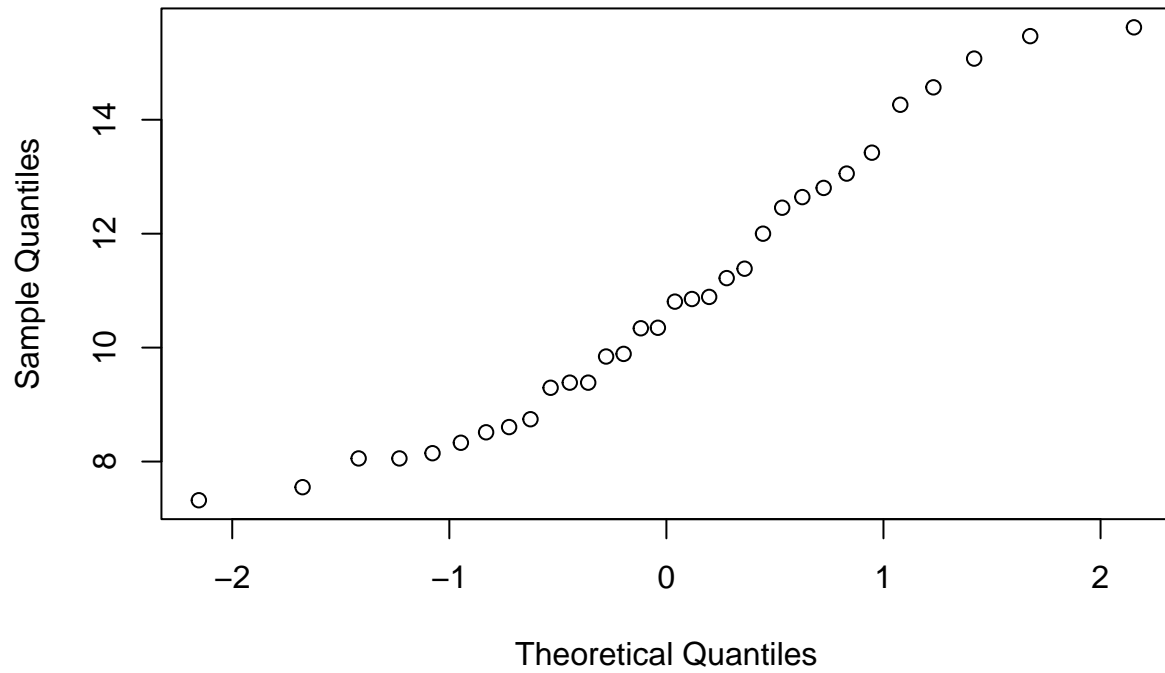
```
# Actividad_enzimatica
```

```
hist(Resultados$Actividad._enzimatica)
```



```
qqnorm(Resultados$Actividad._enzimatica)
```

Normal Q-Q Plot



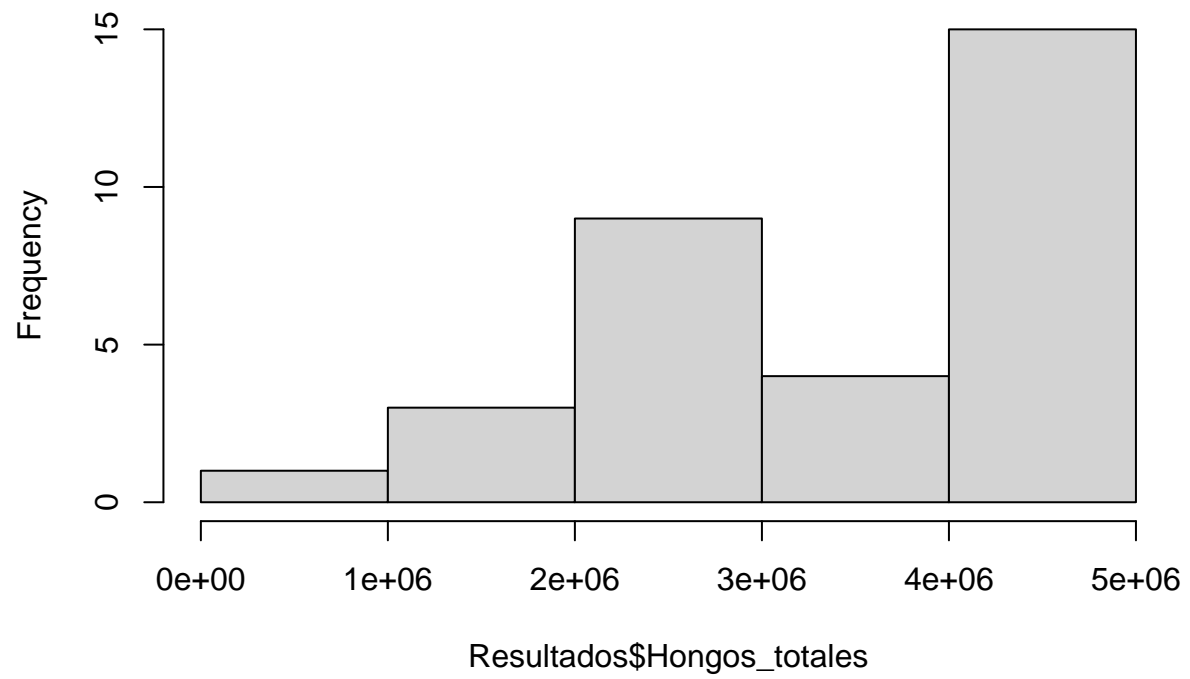
```
shapiro.test(Resultados$Actividad._enzimatica)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  Resultados$Actividad._enzimatica  
## W = 0.94319, p-value = 0.09226
```

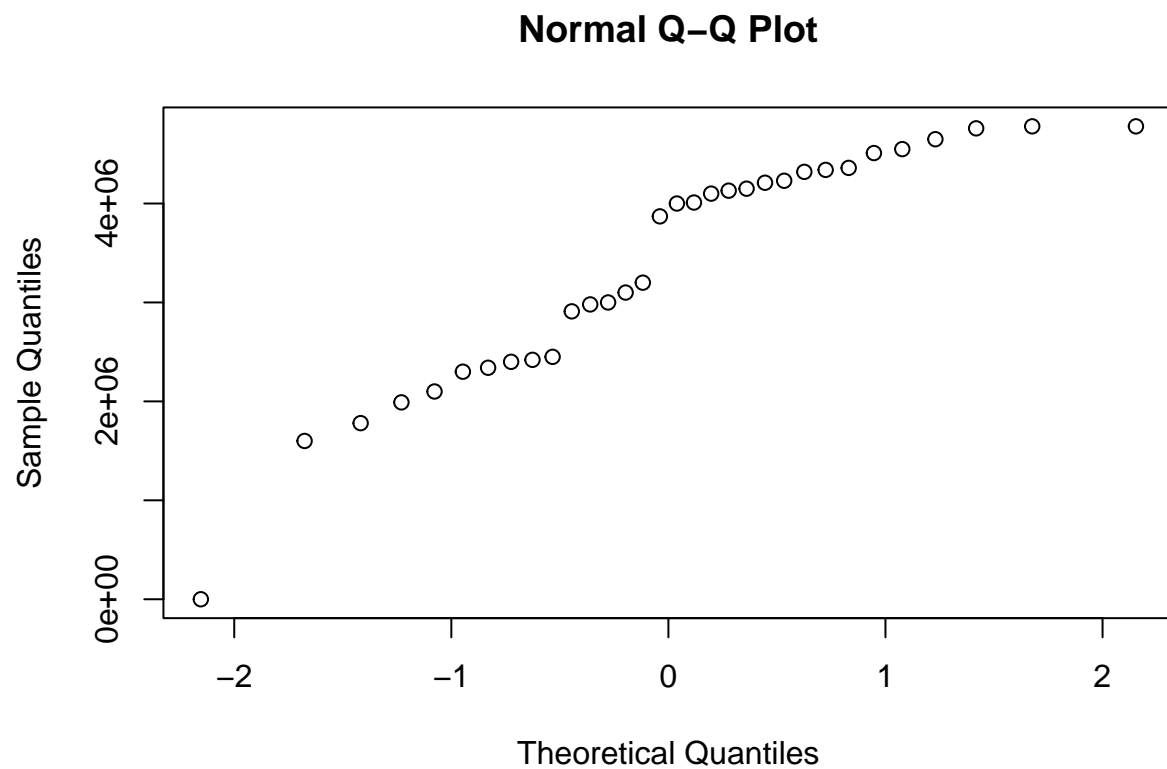
```
# Hongos_totales
```

```
hist(Resultados$Hongos_totales)
```

Histogram of Resultados\$Hongos_totales



```
qqnorm(Resultados$Hongos_totales)
```



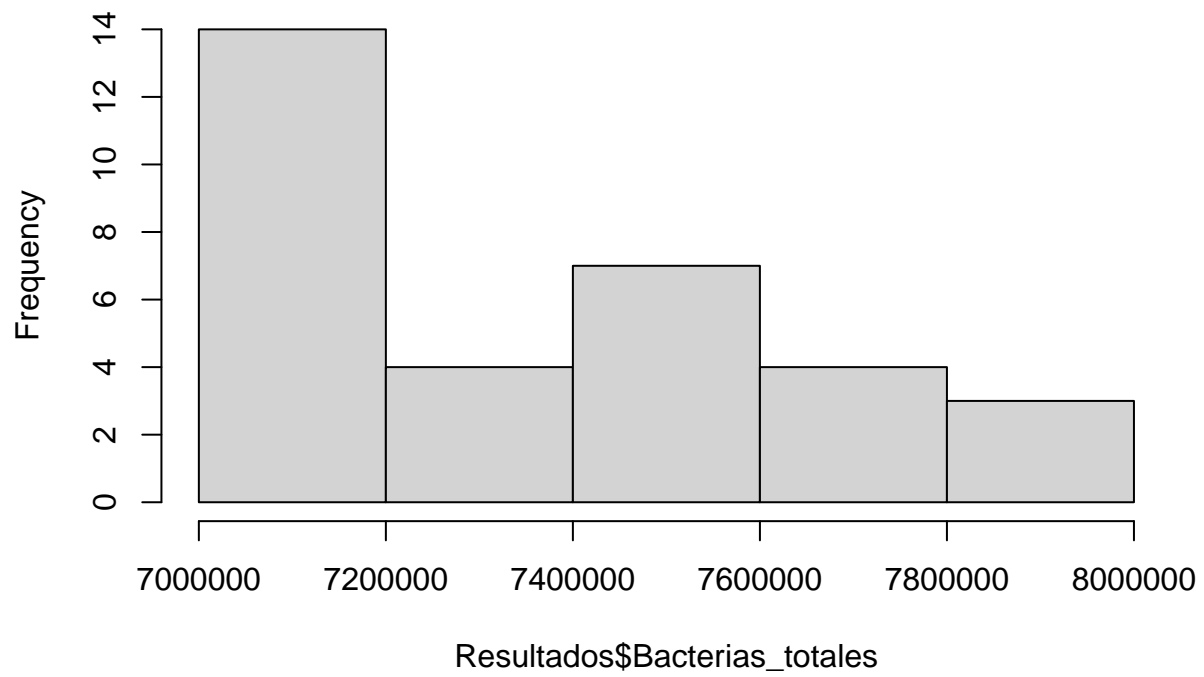
```
shapiro.test(Resultados$Hongos_totales)
```

```
##  
##  Shapiro-Wilk normality test  
##  
## data:  Resultados$Hongos_totales  
## W = 0.90174, p-value = 0.006859
```

```
# Bacterias_totales
```

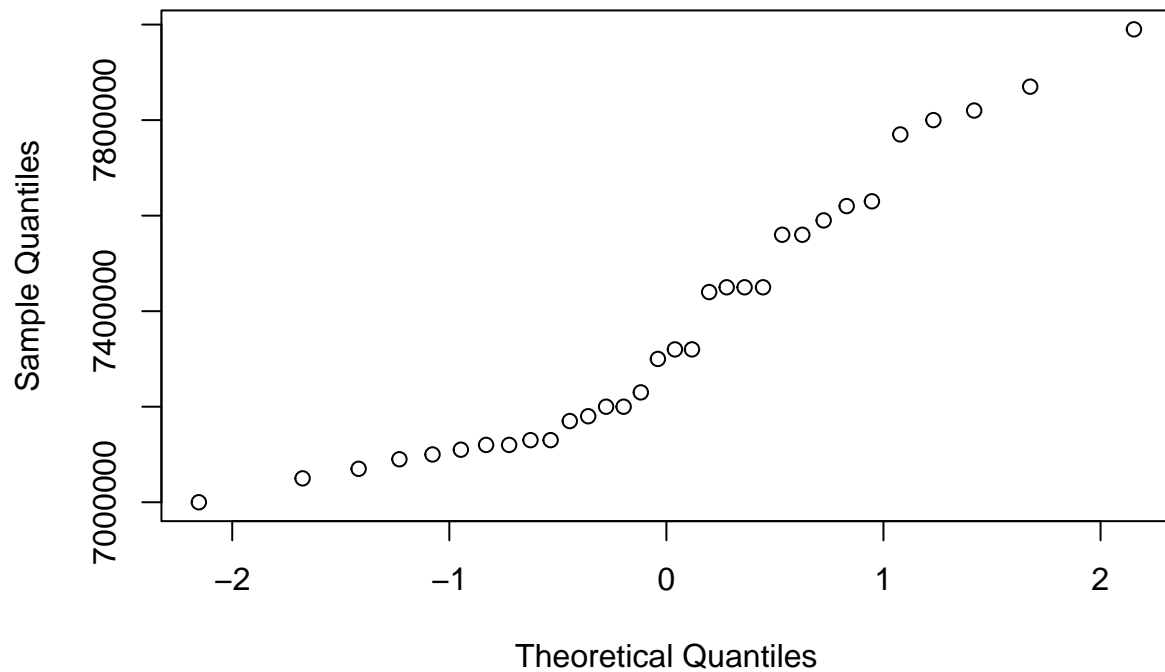
```
hist(Resultados$Bacterias_totales)
```

Histogram of Resultados\$Bacterias_totales



```
qqnorm(Resultados$Bacterias_totales)
```

Normal Q-Q Plot



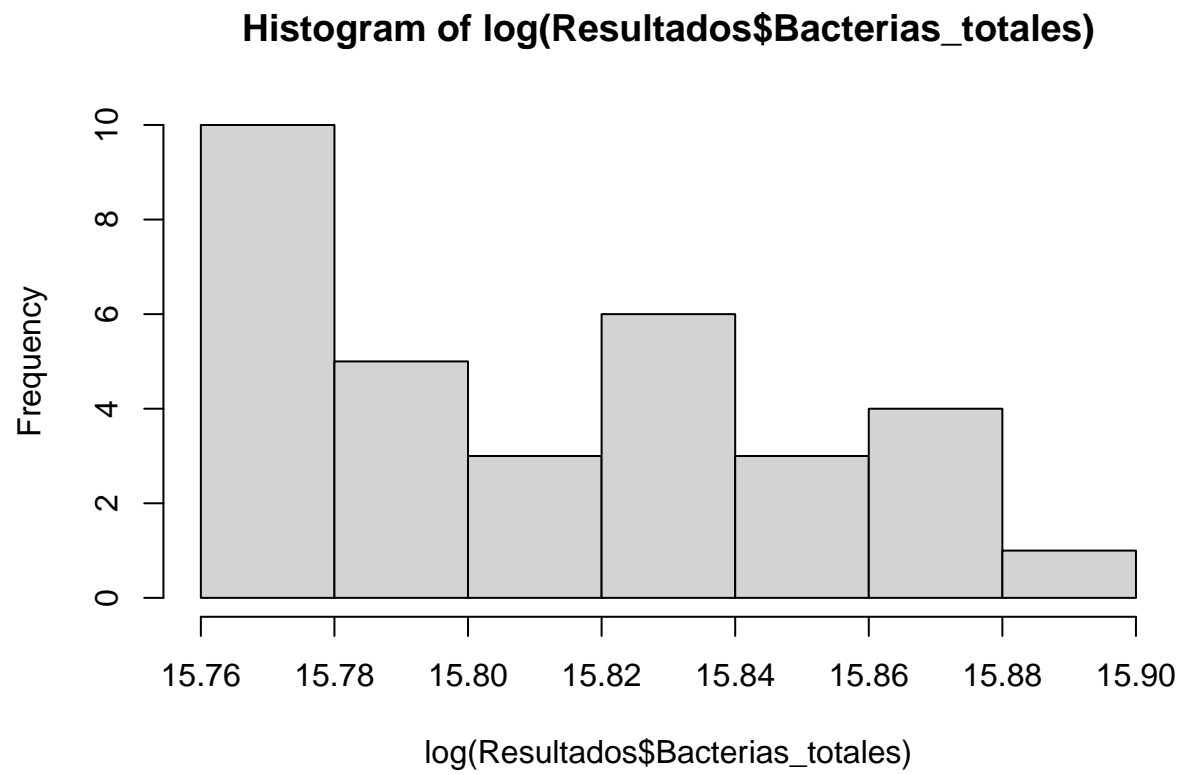
```
shapiro.test(Resultados$Bacterias_totales) # Normalizar datos
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: Resultados$Bacterias_totales  
## W = 0.91461, p-value = 0.01489
```

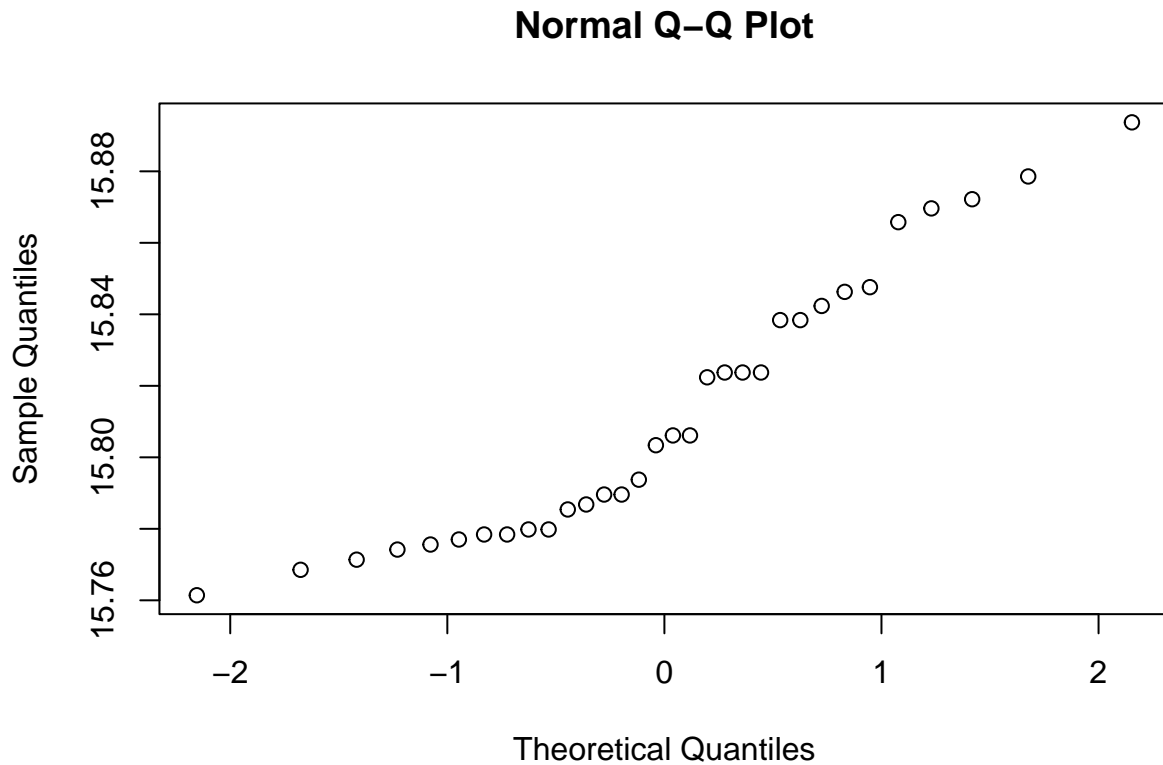
```
shapiro.test(log(Resultados$Bacterias_totales))
```

```
##  
## Shapiro-Wilk normality test  
##  
## data: log(Resultados$Bacterias_totales)  
## W = 0.91846, p-value = 0.01888
```

```
hist(log(Resultados$Bacterias_totales))
```



```
qqnorm(log(Resultados$Bacterias_totales))
```

```
# Verificar el balance de los datos
```

```
Resultados %>%
  group_by(Variedad) %>%
  summarise (n())
```

```
## # A tibble: 2 x 2
##   Variedad      'n()'
##   <chr>        <int>
## 1 Cotton candy    16
## 2 Sweet globe    16
```

```
# Homogenidad de varianza
```

```
leveneTest(log(Bacterias_totales) ~ Variedad*Tipo_muestra, data = Resultados) # no hay homogeneidad
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  4.5045 0.01062 *
##      28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
leveneTest(Actividad._enzimatica ~ Variedad*Tipo_muestra, data = Resultados) # si hay homogeneidad
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##      Df F value Pr(>F)
## group 3  0.7426 0.5357
##      28
```

Los resultados de verificación son los siguientes:

Actividad_enzimatica: *Shapiro*= 0.09226 Conforme a los resultados de shapiro, es >0.05, por lo cual ya estan normalizados los datos y también con las graficas se puede observar que si.

Bacterias_totales: *Shapiro*=0.01489 Conforme a los resultados de shapiro, es <0.05, por lo cual es necesario transformar los datos, y en la grafica se observa que tiene un sesgo a la derecha por lo que utilcé log para normalizarlos.

Balance de datos: Es importante verificar que se cuente con el mismo numero de datos, en este analisis se cuenta con 16 datos para la variedad *cotton candy* y 16 para *sweet globe*

Homogeneidad de varianza: Se realizó la prueba de Levene arrojando los siguientes resultados

1.*Bacterias_totales*: p= 0.01062 2.*Actividad_enzimatica*:p = 0.5357

Con los resultados anteriores se concluye que si hay homogeneidad con los datos de actividad enzimatica y que con la interacción de los datos con Bacterias_totales no hay homogeneidad.

ANOVA

```
# Anova dos vías
```

```
Anova_enz <- aov(Actividad._enzimatica ~ Variedad*Tipo_muestra, data = Resultados)
```

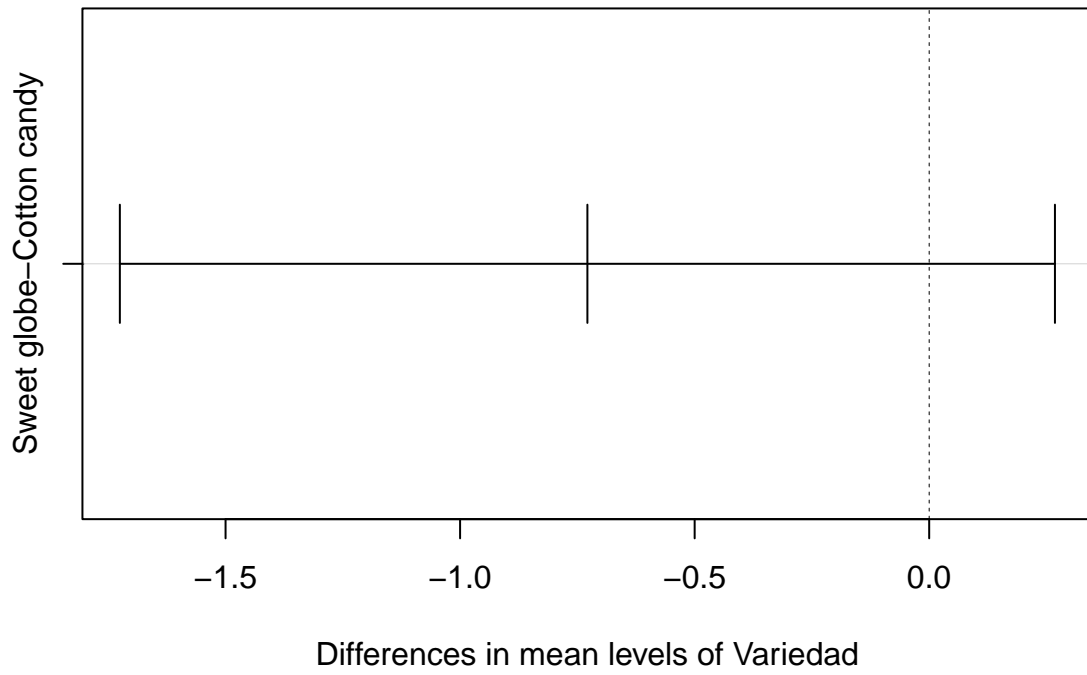
```
Anova(Anova_enz)
```

```
## Anova Table (Type II tests)
##
## Response: Actividad._enzimatica
##      Sum Sq Df F value    Pr(>F)
## Variedad      4.247  1  2.2421  0.1455
## Tipo_muestra 126.769  1 66.9227 6.63e-09 ***
## Variedad:Tipo_muestra  2.847  1  1.5029  0.2304
## Residuals      53.039 28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

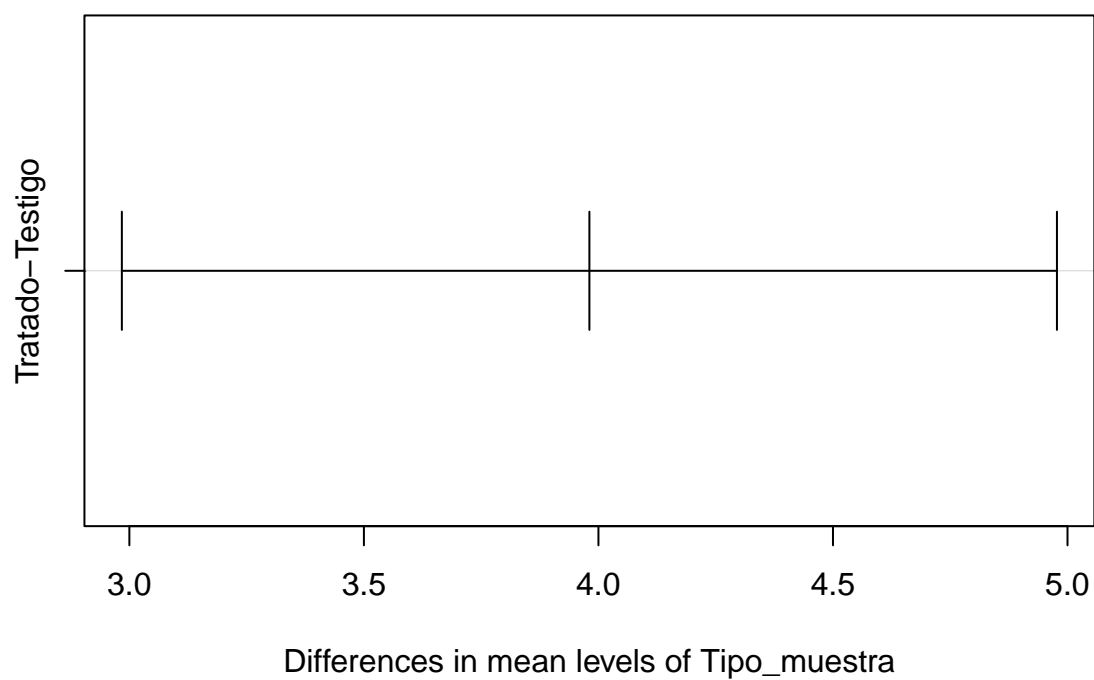
```
# Prueba Tukey
```

```
Anova_tukey <- TukeyHSD(Anova_enz)
plot(Anova_tukey)
```

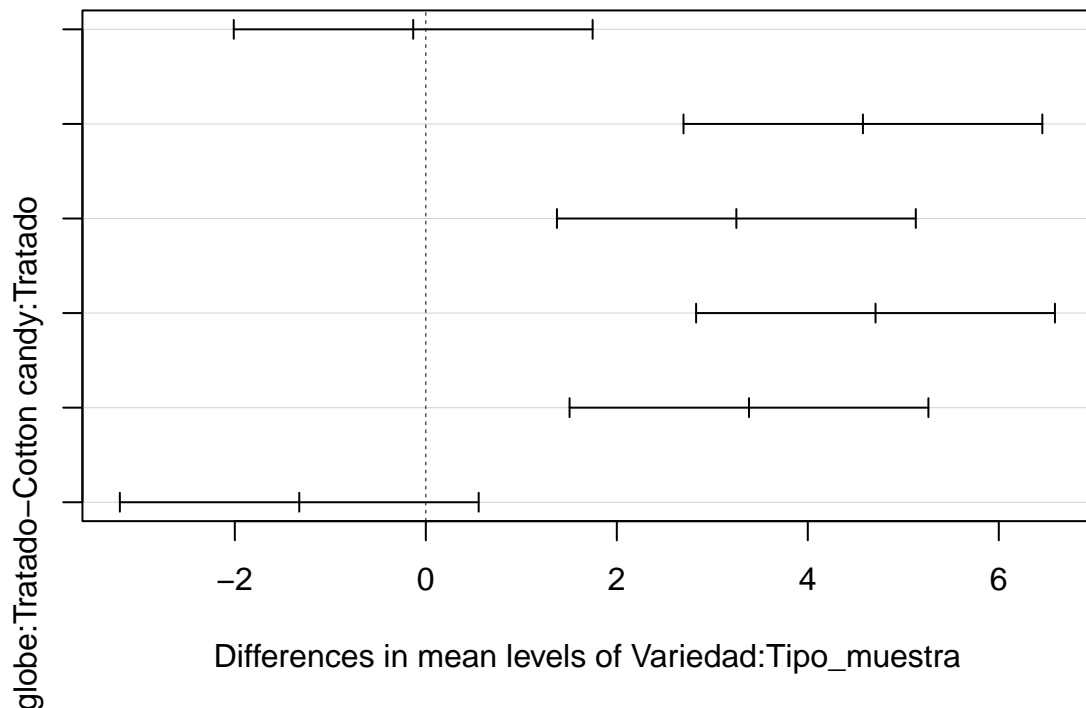
95% family-wise confidence level



95% family-wise confidence level



95% family-wise confidence level



```
summary(Anova_enz)
```

```
##              Df Sum Sq Mean Sq F value    Pr(>F)
## Variedad      1   4.25    4.25    2.242    0.145
## Tipo_muestra  1 126.77  126.77   66.923 6.63e-09 ***
## Variedad:Tipo_muestra 1   2.85    2.85    1.503    0.230
## Residuals    28  53.04    1.89
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
grupos <- HSD.test(Anova_enz, trt = c("Variedad", "Tipo_muestra"),
                    alpha = 0.05)
```

Interpretación resultados

Conforme a los resultados obtenidos en la prueba de Tukey se interpreta de la siguiente manera:

1. En las muestras de los tipos de variedades da un valor $p = 0.145$ por lo que **no** hay diferencias significativas entre la variedad *Sweet globe* y *Cotton candy*
2. En las muestras de los tipos da un valor de $p = 6.63e-09$, por lo que aquí **si** hay diferencias significativas en las muestras *testigo* y *tratadas*
3. En la interacción de las muestras de variedad con el tipo ya sea testigo o tratado **no** hay diferencia significativa ya que el valor de $p = 0.230$

Separar por grupos

	Actividad._enzimatica	groups
Cotton candy:Tratado	13.538474	a
Sweet globe:Tratado	12.213303	a
Cotton candy:Testigo	8.961202	b
Sweet globe:Testigo	8.829128	b

Figure 1: Tabla1

También se corrió el análisis donde separa por grupos, se puede observar en la siguiente tabla:

El grupo A corresponde a los tratados en los cuales no hay diferencia significativa, y el grupo B corresponde a los testigos, por lo que se concluye lo mismo que se menciona anteriormente que si existe una diferencia entre el grupo A y B Testigos y Tratados, pero en la variedad no hay diferencia.

Las muestras testigo pertenecen a muestras de suelo sin aplicación de producto biológico y las muestras tratadas fueron aplicadas con el producto y si hay un aumento de la actividad enzimatica aplicando el tratamiento por lo cual está relacionado al incremento de microorgrnismos en suelo.

Para conocer más sobre que es un análisis de varianza tipo ANOVA, a informacion la pueden encontrar en el siguiente link: [Mas información](#)

Se utilizó el codigo de shapiro para transformar el valor de los datos

```
# Cambiar el valor de los resultados de Actividad enzimatica de ug a mg

for (i in 1:nrow(Resultados)) {
  enz_mg <- Resultados$Actividad._enzimatica[i]/1000
  print(enz_mg)
}
```

```
## [1] 0.007550459
## [1] 0.007321101
## [1] 0.008743119
## [1] 0.008513761
## [1] 0.009844037
```

```
## [1] 0.009889908
## [1] 0.009385321
## [1] 0.009385321
## [1] 0.008146789
## [1] 0.008330275
## [1] 0.008055046
## [1] 0.008605505
## [1] 0.008055046
## [1] 0.01034862
## [1] 0.009295122
## [1] 0.01085321
## [1] 0.01456881
## [1] 0.01080734
## [1] 0.012
## [1] 0.01507339
## [1] 0.0126422
## [1] 0.01088991
## [1] 0.01033945
## [1] 0.01138532
## [1] 0.01305505
## [1] 0.01342202
## [1] 0.01122008
## [1] 0.0124556
## [1] 0.01280309
## [1] 0.01562162
## [1] 0.01546718
## [1] 0.01426316
```

Ejercicio

Determinar si existe una correlación entre la biomasa de dos especies acuáticas de plantas en estanques de Alaska: Carex y Artophila, usando datos modernos.

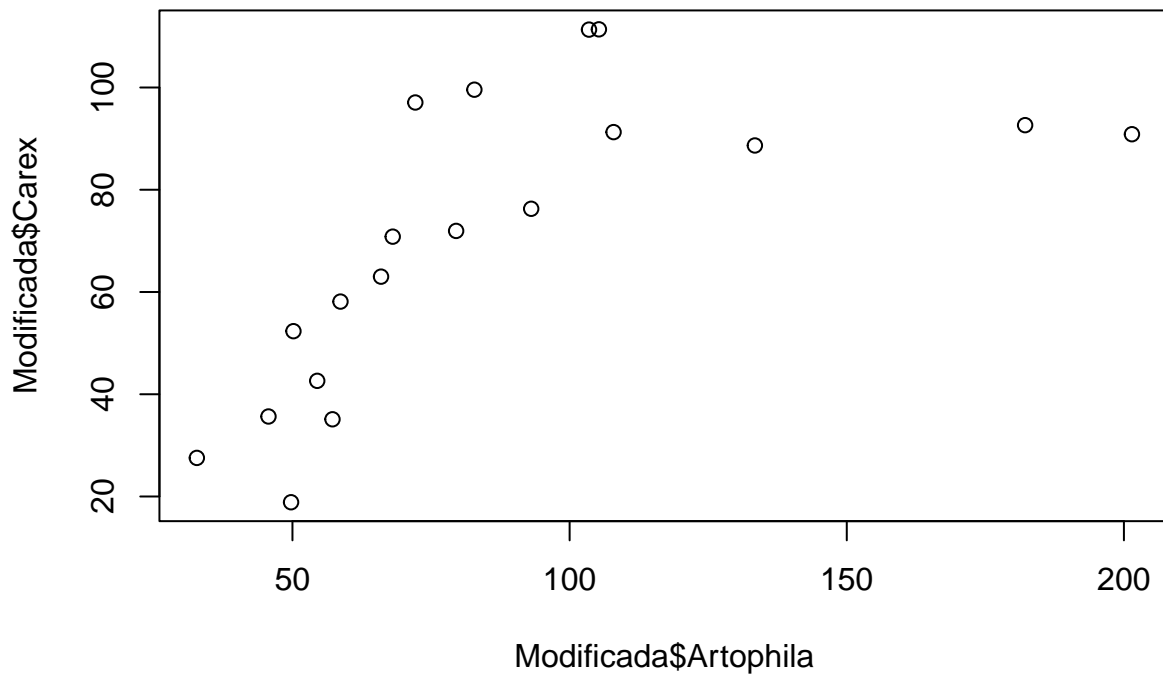
```
# Seleccionar columnas que vamos a utilizar para el analisis

Tmodificada <- Estanque_de_plantas[,c("Era", "Artophila", "Carex")]

Modificada <- Tmodificada %>%
  filter(!is.na(Artophila), !is.na(Carex)) %>%
  select(Era, Artophila, Carex)

# Verificar si cumple con las funciones

plot(Modificada$Artophila, Modificada$Carex)
```



```
## Correlacion
```

```
cor.test(Modificada$Artophila,Modificada$Carex)
```

```
##
## Pearson's product-moment correlation
##
## data: Modificada$Artophila and Modificada$Carex
## t = 3.6467, df = 17, p-value = 0.001996
## alternative hypothesis: true correlation is not equal to 0
## 95 percent confidence interval:
## 0.2979541 0.8584059
## sample estimates:
## cor
## 0.6625044
```

```
# Resultados
```

```
# El coeficiente de correlación da un valor de 0.6625044 y el p-value =0.001996, por lo que se considera
```