# Week 5 Lab

AUTHOR
Jessica Tran

PUBLISHED
February 23, 2024

## Data

For this lab, we will be using the data set `CO2`. Take some time to get familiar with the data set using the help function.

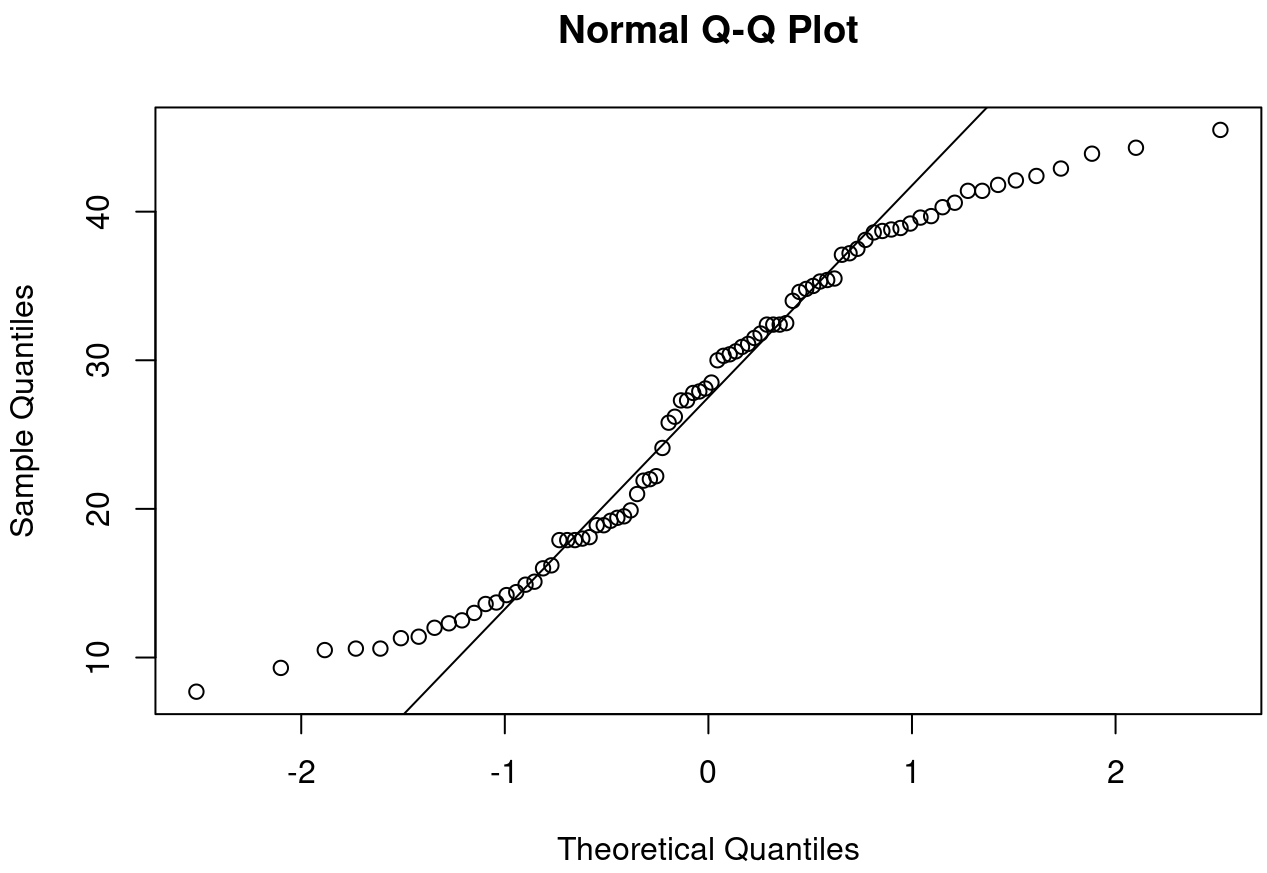
```
?CO2
```

## Problem 1

Conduct a hypothesis test for a mean of one sample. Using the data set `CO2` and the variable `uptake`, test the null hypothesis that the population mean `uptake` is mean 28 vs. the alternate hypothesis that the population mean `uptake` is less than 28 with a significance level of 0.05.

Check any assumptions:

```
qqnorm(CO2$uptake)
qqline(CO2$uptake)
```




If the assumptions are met, calculate the p-value:

```
t.test(CO2$uptake, mu = 28, alternative = "less", conf.level = 0.95)
```

```
    One Sample t-test

data:  CO2$uptake
t = -0.6669, df = 83, p-value = 0.2533
alternative hypothesis: true mean is less than 28
95 percent confidence interval:
     -Inf 29.17585
sample estimates:
mean of x
  27.2131
```

# Problem 2

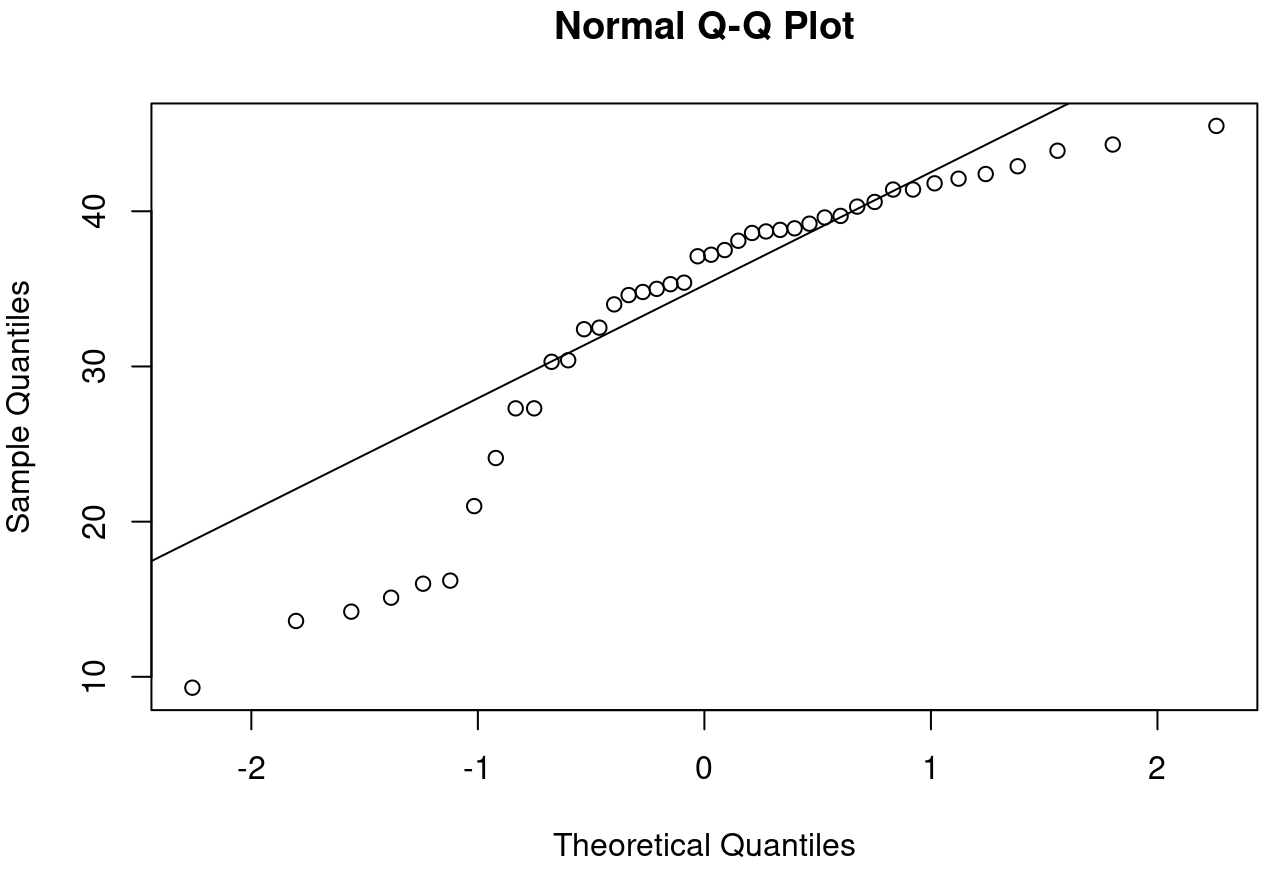
Conduct a hypothesis test for independent samples. Use the data set `CO2` and the variables `Type` and `uptake`. Let group 1 be `Quebec` and group 2 be `Mississippi`. Test the null hypothesis that the population mean `uptake` of the `Quebec` samples is the **same** as that of the `Mississippi` samples vs. the alternate hypothesis that the population mean `uptake` of the `Quebec` samples is **different**

than that of the `Mississippi` samples with a significance level of 0.01. Comment out each line of code to explain what you are doing.

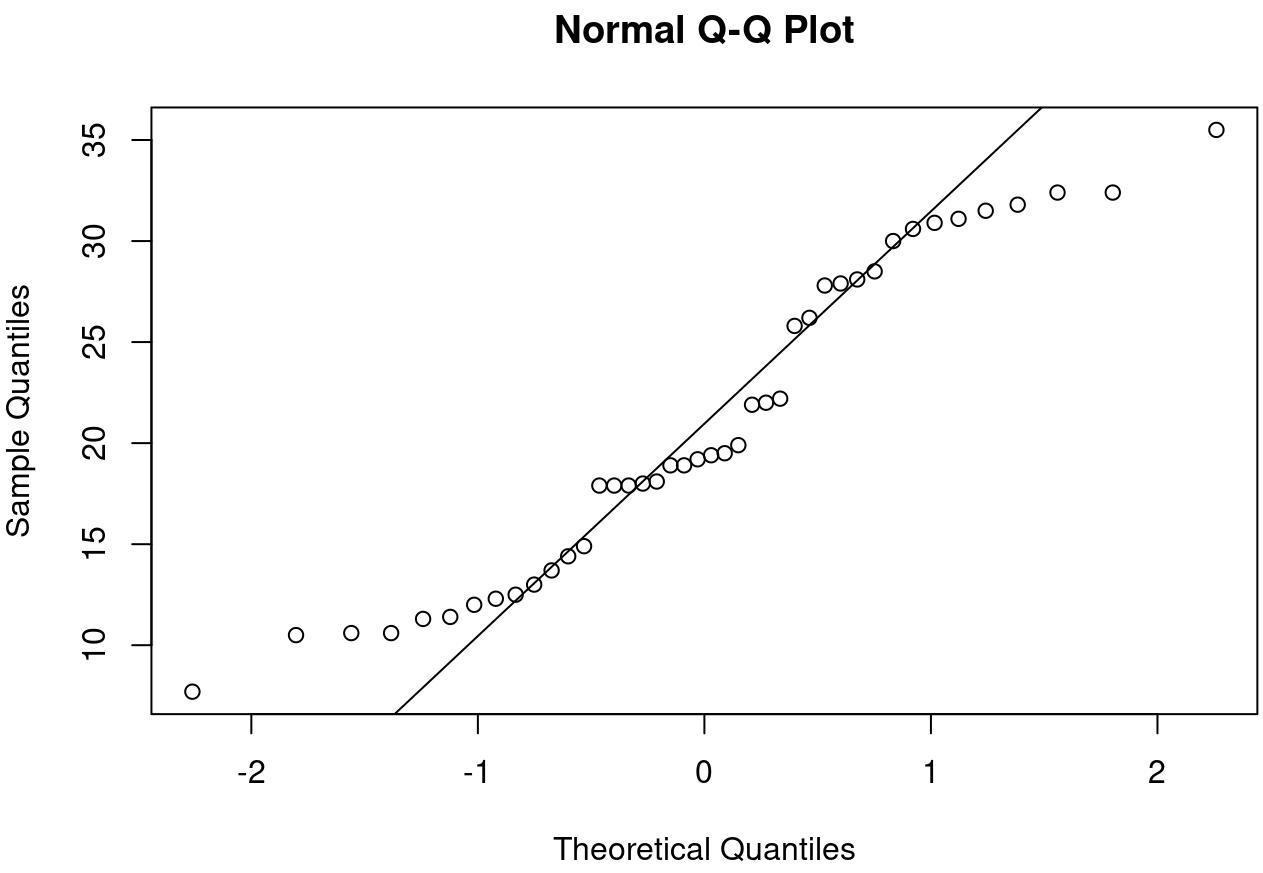
Check any assumptions:

```
group1 <- with(CO2, uptake[Type == "Quebec"]) # select just group 1
group2 <- with(CO2, uptake[Type == "Mississippi"]) # select just group 2

qqnorm(group1)                          # QQ plot for group 1
qqline(group1)
```




```
qqnorm(group2)                          # QQ plot for group 2
qqline(group2)
```




```
var.test(group1, group2)
```

```
    F test to compare two variances

data:  group1 and group2
F = 1.532, num df = 41, denom df = 41, p-value = 0.1763
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.8234755 2.8500701
sample estimates:
ratio of variances
           1.53198
```

If the assumptions are met, calculate the p-value:

```
t.test(uptake ~ Type, data = CO2, alternative = "two.sided", conf.level = 0.99, p
```

```
    Two Sample t-test

data:  uptake by Type
t = 6.5969, df = 82, p-value = 3.835e-09
```

```
alternative hypothesis: true difference in means between group Quebec and group
Mississippi is not equal to 0
99 percent confidence interval:
  7.598856 17.720192
sample estimates:
     mean in group Quebec mean in group Mississippi
                 33.54286                  20.88333
```

## Problem 3

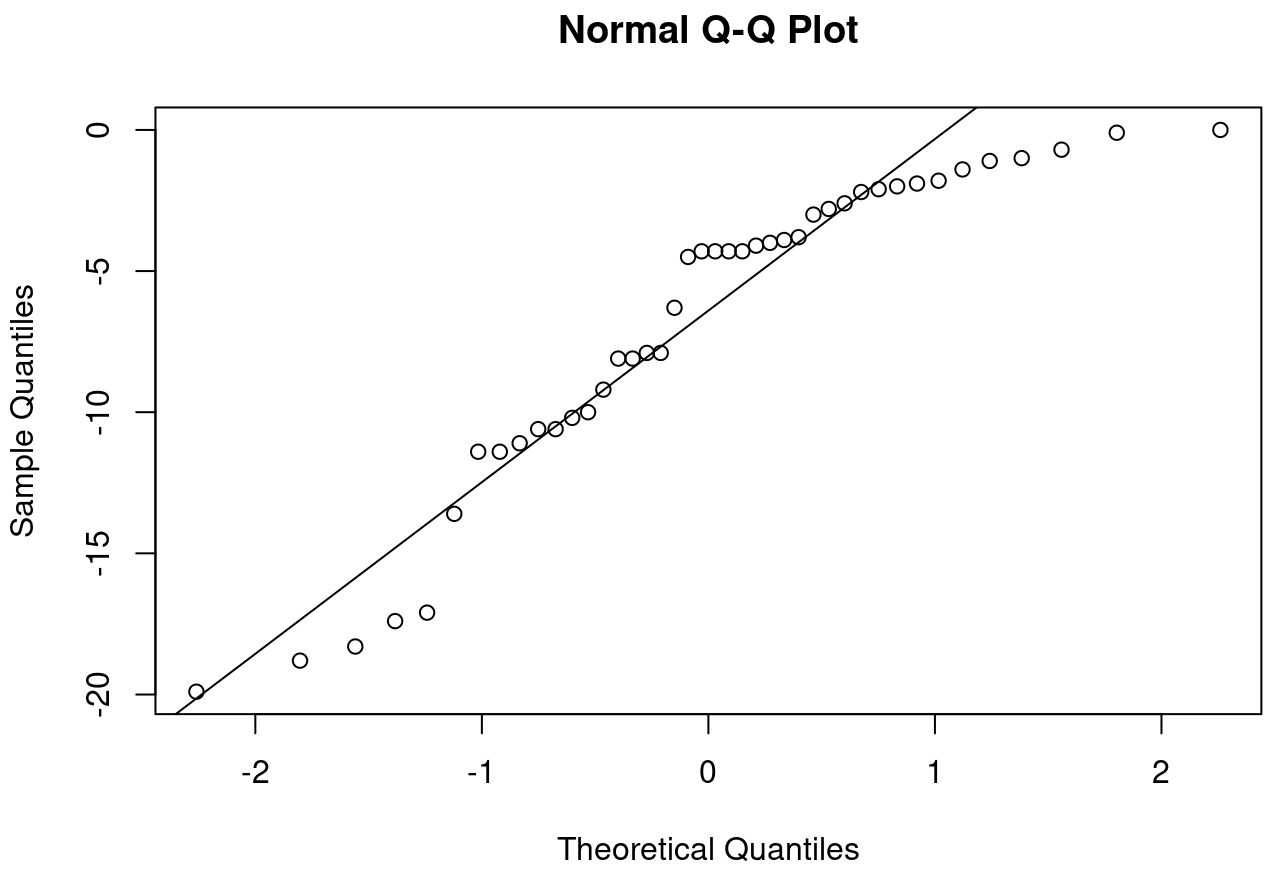
Conduct a hypothesis test for paired samples. Use the data set `CO2` and the variable `uptake`. Let group 1 be `chilled` and group 2 be `nonchilled`. Test the null hypothesis that the population mean `uptake` of the `chilled` samples is the same as that of the `nonchilled` samples vs. the alternate hypothesis that the population mean `uptake` of the `chilled` samples is different than that of the `nonchilled` samples with a significance level of 0.05. Comment out each line of code to explain what you are doing.

Check any assumptions:

```
group1 <- with(CO2, uptake[Treatment == "chilled"]) # select just group 1
group2 <- with(CO2, uptake[Treatment == "nonchilled"]) # select just group 2

differenced_data <- group1 - group2

qqnorm(differenced_data)
qqline(differenced_data)
```




If the assumptions are met, calculate the p-value:

```
t.test(uptake ~ Treatment, data = CO2, alternative = "two.sided", conf.level = 0.
```

```
    Paired t-test

data:  uptake by Treatment
t = 7.939, df = 41, p-value = 8.051e-10
alternative hypothesis: true mean difference is not equal to 0
95 percent confidence interval:
 5.114589 8.604458
sample estimates:
mean difference
       6.859524
```

# Problem 4

Conduct a hypothesis test for one proportion. We are interested in knowing what proportion of plants have an uptake greater than 25 $\mu$mol/ $m^2$ sec. Out of the 84 plants, 49 of them had an uptake greater than 25 $\mu$mol/ $m^2$ sec. At a significance level of 1%, can you conclude that a majority of plants have an update greater than 25 $\mu$mol/ $m^2$ sec. Comment out each line of code to explain what you are doing.

Check any assumptions:

```
84*(0.5)          #np
```

```
[1] 42
```

```
84*(1-0.5)      #n(1-p)
```

```
[1] 42
```

If the assumptions are met, calculate the p-value:

```
prop.test(49, 84, p = 0.5, alternative = "greater", correct = FALSE)
```

```
    1-sample proportions test without continuity correction

data:  49 out of 84, null probability 0.5
X-squared = 2.3333, df = 1, p-value = 0.06332
alternative hypothesis: true p is greater than 0.5
95 percent confidence interval:
 0.4936065 1.0000000
sample estimates:
        p
0.5833333
```

## Submitting

Submit the following to Canvas:

- Your rendered PDF titled Lastname_5R. Make sure your name is at the top of the document.
- Your .qmd file