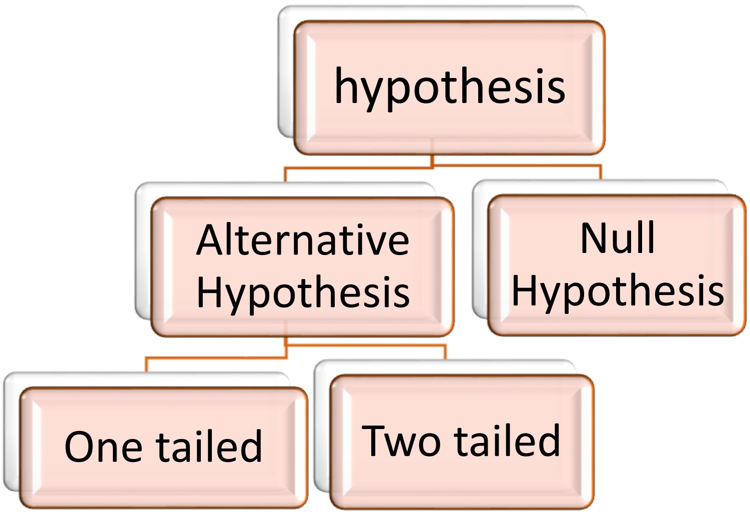
INDEX

HYPOTHESIS

Hypothesis is a statement or assertion about the statistical distribution or parameter of statistical distribution. In statistical hypothesis testing, two hypotheses are compared. These are call **null hypothesis** and **alternative hypothesis**.

A hypothesis of no difference is a null hypothesis and complementary hypothesis to null hypothesis is an alternative hypothesis.



The alternative hypothesis may take several forms, depending on the nature of the hypothesized relation; in particular, it can be one tailed or two tailed. A one tailed test is a statistical test in which the critical area of a distribution is one sided so that it is either greater than or less than a certain value, but not both. A two tailed test is a statistical test in which the critical area of a distribution is not equal to certain value.

t-TEST

- **FOR SINGLE SAMPLE**

A t-test is applied when the test statistic would follow a normal distribution and variance is unknown.

- $H_0 : \mu=\mu_0$

  $H_1 : \mu>\mu_0$

- $t_{cal} = \frac{\bar{X}-\mu}{s/\sqrt{n}}$

  here, μ_0=50dB

  $\bar{X}$= Mean of population

  n= Size of sample

  $s^2 = \frac{1}{n-1} \Sigma (X_i-\bar{X})^2$

  Under H_0: $t_{cal} \sim t_{n-1,\alpha}$

- Decision Criteria:
  - Critical Region:

    We reject H_0 at level of significance

    If $t_{cal} \geq t_{n-1,\alpha}$
  - Confidence Interval:

    We reject H_0 at level of significance, if given value of μ_0=50 does not lie in confidence interval.
  - p-value:

    We reject H_0 at level of significance

    If p-value < α

- **FOR TWO SAMPLES**

One tail:

- $H_0 : \mu = \mu_0$
  $H_1 : \mu > \mu_0$

Two tail:

- $H_0 : \mu = \mu_0$
  $H_1 : \mu \neq \mu_0$

- $t_{cal} = \dfrac{\bar{x} - \bar{y}}{s\sqrt{\frac{1}{n1} + \frac{1}{n2}}}$

  Here,

  $\mu_0 = 50dB$
  $\bar{x}$ = mean of sample size n1 drawn from first population
  $\bar{y}$ = mean of sample size n2 drawn from second population
  $s^2 = \dfrac{n1\,(S1)2 + n2\,(S2)2}{n1 + n2 - 2}$
  (S1)2 = variance of observations in first sample

  (S2)2 = variance of observations in second sample

  n1 = sample size of first population

  n2 = sample size of second population

  Under H_0: $t_{cal} \sim t_{n1+n2-2,\alpha}$ ***(ONE TAILED)***

  Under H_0: $t_{cal} \sim t_{n1+n2-2,\alpha/2}$ ***(TWO TAILED)***

- Decision Criteria:

  - Critical region:

    We reject H_0 at level of significance if tcal > $t_{n1+n2-2,\alpha}$

  - Confidence Interval:

    We reject H_0 at level of significance if zero is not included in the confidence interval

  - p-value:

    We reject H_0 at level of significance if p-value ‹ α

INTRODUCTION

What comes to your mind when you think about noise?

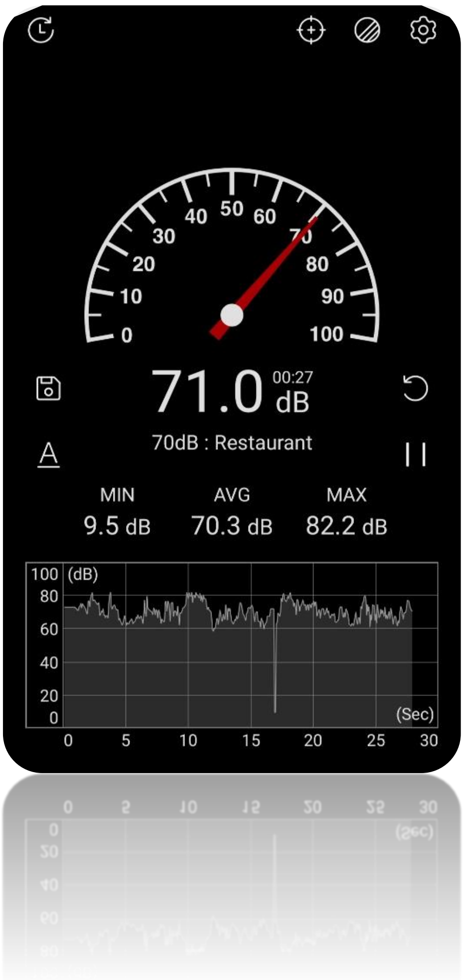
All these actually create a lot of noise which affects our health badly. But the thing is that most of us are exposed to lots of noise not only when we travel to work daily but also when we are sitting at our home. We hardly even notice or think about the noise around us and the damage that it causes to each one of us. Noise pollution can cause **hypertension**, high **stress** levels, tinnitus, **hearing loss**, **sleep disturbances**, and other harmful and disturbing effects. Imagine what it must be doing to the patients in the hospitals! We are pretty sure that most of us are completely unaware of the fact that according to WHO the sound levels outside any hospital must be around 50 dB . The question is – ARE THE SOUND LEVELS OUTSIDE THE HOSPITALS IN OUR AREA REALLY AROUND 50 dB?

To check it we headed out to four different hospitals DEENANATH MANGESHKAR,SANCHETI,SAHYADRI and SANJEEVAN. This project includes our findings .

METHODOLOGY

DATA COLLECTION-This project uses the method of primary data collection. Sound levels outside the hospitals were taken using **SOUND METER** Application on cell phones every 2 minutes for a time interval of 2 hours i.e. 11:00 am to 1:00 am on 5/2/2020. Four Hospitals were taken into consideration:

- SANCHETI(SHIVAJI NAGAR)
- SAHYADRI(KARVE ROAD)
- SANJEEVAN(KASHIBAI KHILARE PATH,OFF KARVE ROAD)
- DINANATH MANGESHKAR(KOTHRUD)

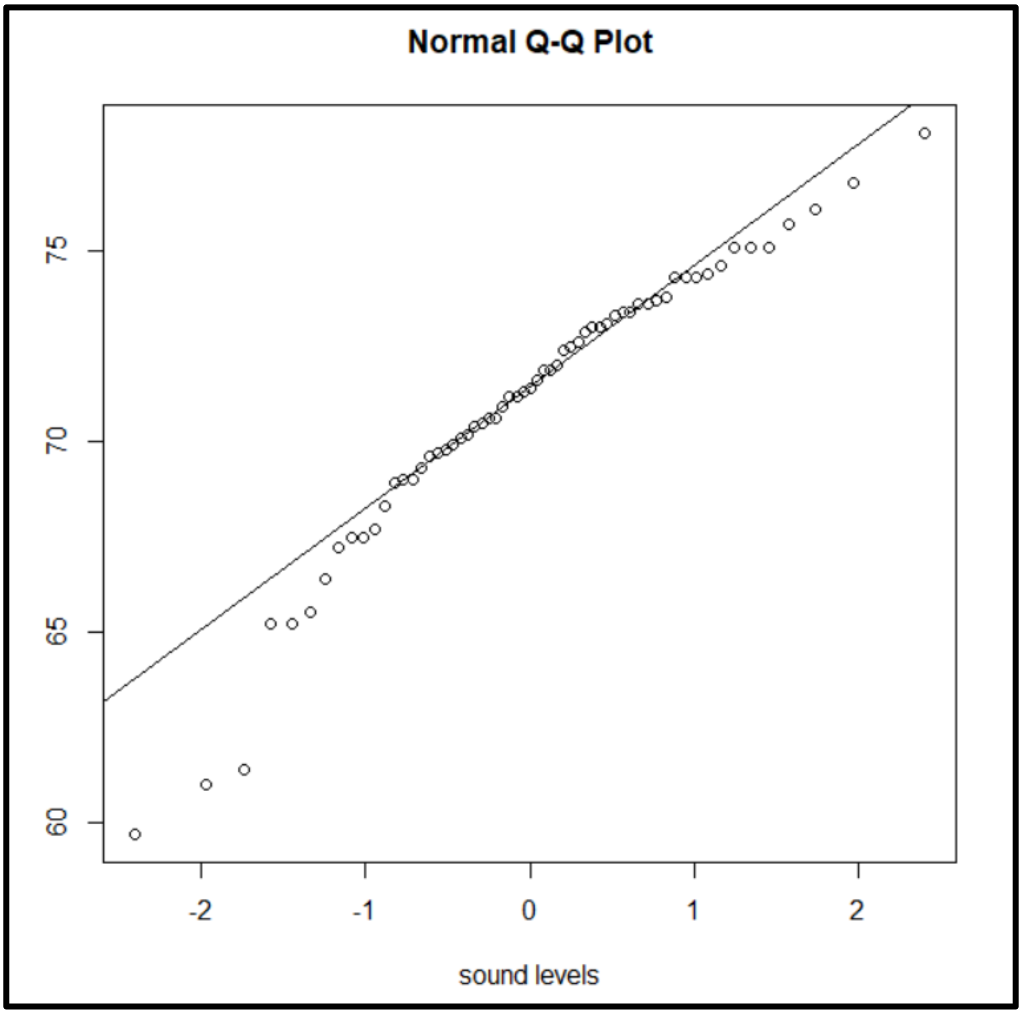


ANALYSIS BY-Normality and Hypothesis Testing of sound levels using R and excel.

SANJEEVAN HOSPITAL

DATA-

X=70.5,70.2,69.8,59.7,66.4,69.3,61.4,69,72.9,74.4,75.1,69.7,70.6,70.9,73.1,73.6,73.3,70.1,68.3,73.6,71.6,61,71.3,73.7,71.9,72.5,71.4,69.6,65.5,73.4,74.3,70.4,72.4,75.1,70.6,73,72.6,75.1,74.3,76.8,73,72,73.8,73.4,71.2,67.5,74.6,69.9,76.1,65.2,69,71.2,78.1,67.5,71.9,68.9,67.2,74.3,67.7,75.7,65.2



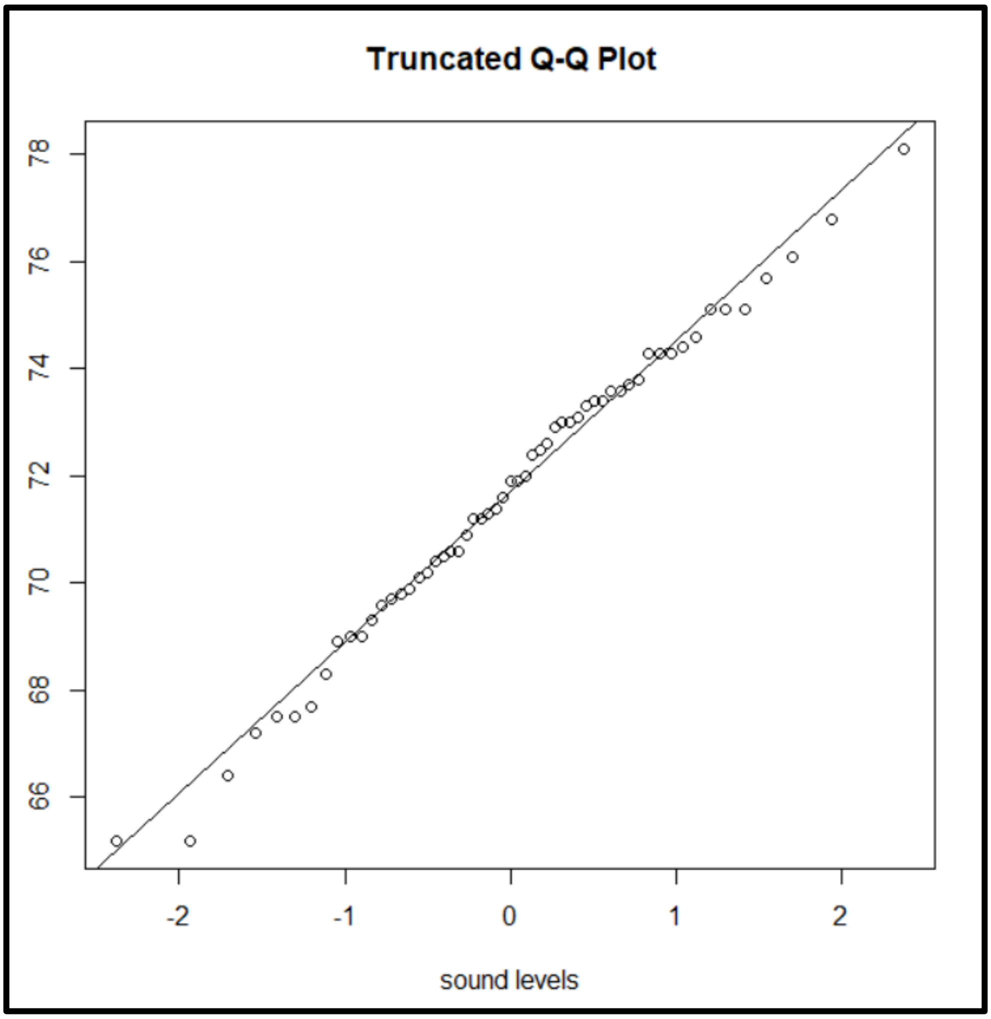
INTERPRETATION - Since a few points do not pass through the line, we can say that the sample doesn't follow Normal distribution.

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.94303, p-value = 0.006806
```

<u>Conclusion</u>-As the p-value is less than α=0.05 hence the collected sample cannot be considered normal.

TRUNCATED DATA-

X2=70.5,70.2,69.8,66.4,69.3,69,72.9,74.4,75.1,69.7,70.6,70.9,73.1,73.6,73.3,70.1,68.3,73.6,71.6,71.3,73.7,71.9,72.5,71.4,69.6,73.4,74.3,70.4,72.4,75.1,70.6,73,72.6,75.1,74.3,76.8,73,72,73.8,73.4,71.2,67.5,74.6,69.9,76.1,65.2,69,71.2,78.1,67.5,71.9,68.9,67.2,74.3,67.7,75.7,65.2



INTERPRETATION - Since most of the points pass through the line, we can say that the truncated sample follows Normal distribution.

```
> shapiro.test(X2)

        Shapiro-Wilk normality test

data:  X2
W = 0.98874, p-value = 0.8735
```

<u>Conclusion</u>-As the p-value is greater than α=0.05 hence the collected sample can be considered normal.

t TEST

Let x= sounds levels outside Sanjeevan Hospital (in dB)

Here μ_0=50 dB

$H_0 : \mu = 50$

Vs.

$H_1 : \mu > 50$

Test Statistic:

$$t_{cal} = \frac{\bar{x}-\mu}{s/\sqrt{n}}$$

$$t_{cal} = \frac{71.65263-50}{2.85545/\sqrt{57}}$$

=57.2497

And $t_{n-1,\alpha}$ =1.672522

Since $t_{cal} > t_{n-1,\alpha}$, we may reject H_0 at 5% level of significance.

<u>Conclusion</u>-We may say that the sounds levels(in dB) outside Sanjeevan Hospital is greater than the critical value.

t TEST USING R

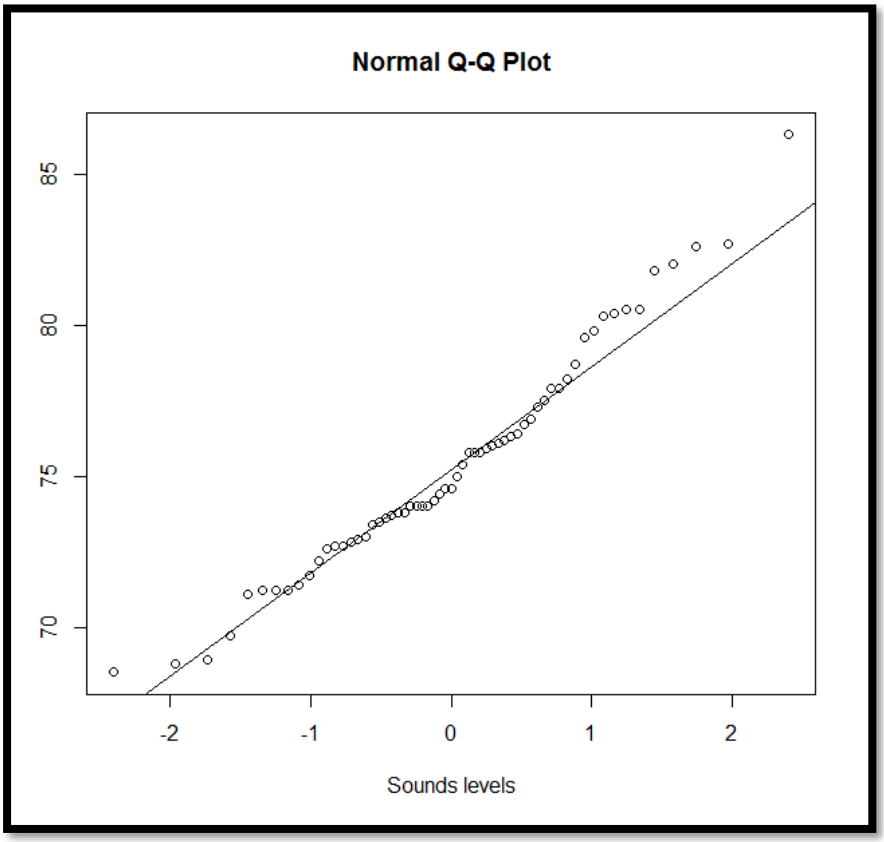
```
> t.test(x,mu=50,alt="g",conf.level=0.95)

        One Sample t-test

data:  x
t = 57.25, df = 56, p-value < 2.2e-16
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 71.02006      Inf
sample estimates:
mean of x
 71.65263
```

<u>Conclusion</u>- As the p-value is less than α = 0.05 and μ_0 = 50 doesn't lie in the 95% confidence interval therefore H_0 is rejected. Hence the sound levels can be considered greater than 50 dB .

SAHYADRI HOSPITAL

DATA-

76.7,71.2,72.7,76.9,74.6,72.7,75.8,80.3,72.8,69.7,71.1,72.9,78.2,76.1,73.8,81.8,72.6,76,75.4,75.8,76.2,74.2,75.0,71.2,82,71.2,76.3,74,73.7,71.7,77.3,74,82.7,73.6,68.9,76.4,74.6,73.4,74.4,82.6,74,74,68.8,80.5,79.6,86.3,75.9,78.7,73.8,77.9,72.2,77.9,71.4,73.5,80.4,75.8,77.5,68.5,80.5,73,79.8



INTERPRETATION - Since most of the points pass through the line, we can say that the sample follows Normal distribution.

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.97196, p-value = 0.1741
```

CONCLUSION- The test gives p-value greater than 0.05, so the sample follows Normal distribution.

t TEST USING R

```
> t.test(x,mu=50,alt="g",conf.level=0.95)

        One Sample t-test

data:  x
t = 52.972, df = 60, p-value < 2.2e-16
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 74.61638      Inf
sample estimates:
mean of x
 75.41803
```

<u>CONCLUSION</u>- We may reject H_0 at 5% level of significance. i.e. We may say that the sounds levels(in dB) outside Sahyadri Hospital is greater than the critical value.

t TEST

Let x= sounds levels outside Sahyadri Hospital (in dB)

Here μ_0=50

$H_O : \mu = 50$

Vs.

$H_1 : \mu > 50$

Test Statistic:

$t_{cal} = \frac{75.41803-50}{3.779315/\sqrt{60}}$

=52.096

And t_{tab}=1.671093033

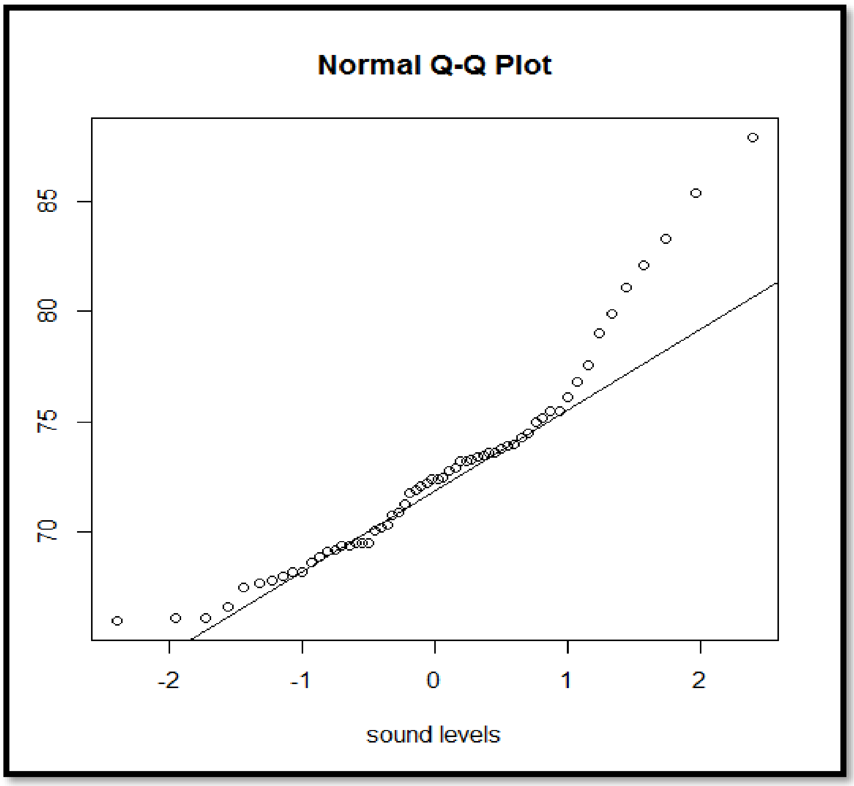
Since $t_{cal} > t_{tab}$, we may reject H_O at 5% level of significance.

We may say that the sounds levels (in dB) outside Sahyadri Hospital is greater than the critical value.

DEENANATH MANGESHKAR HOSPITAL

DATA-

X=71.9,76.1,73.6,70.3,79.9,72.2,87.9,68.6,69.5,69.2,69.5,67.7,75.5,72.8,79,74.5,66.1,73.3,68.2,69.4,66.6,73.9,68.2,70.2,72.4,70.9,66.1,67.8,73.5,73.2,66,72.5,75,73.2,68.9,76.8,69.1,72.9,75.5,74.3,74,73.6,71.8,69.4,75.2,85.4,73.4,72.1,72.4,70.8,70.1,83.3,68,67.5,82.1,71.3,73.8,77.6,81.1,69.5



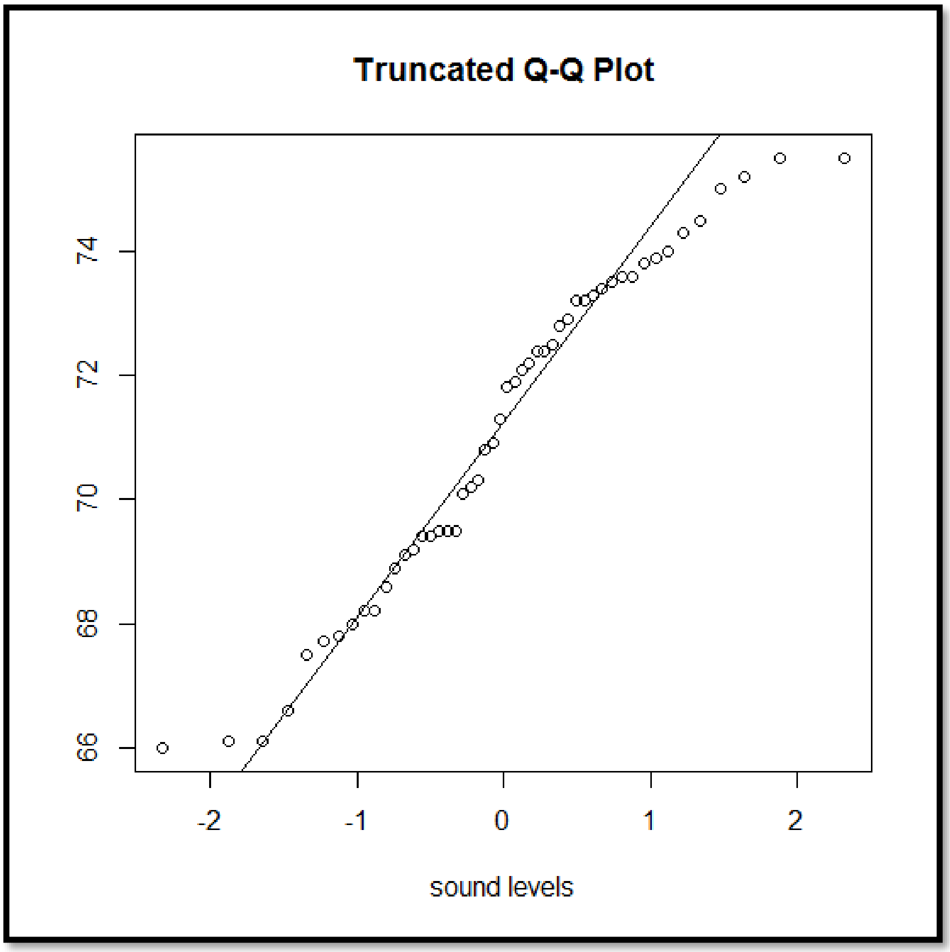
INTERPRETATION - Since few points do not pass through the line, we can say that the sample doesn't follows Normal distribution.

```
> shapiro.test(x)

        Shapiro-Wilk normality test

data:  x
W = 0.91781, p-value = 0.0006267
```

<u>Conclusion</u>: As the p-value is less than α=0.05 hence the collected sample cannot be considered normal.

Truncated Data:

Y=71.9,73.6,70.3,72.2,68.6,69.5,69.2,69.5,67.7,75.5,72.8,74.5,66.1,73.3,68.2,69.4,66.6,73.9,68.2,70.2,72.4,70.9,66.1,67.8,73.5,73.2,66,72.5,75,73.2,68.9,69.1,72.9,75.5,74.3,74,73.6,71.8,69.4,75.2,73.4,72.1,72.4,70.8,70.1,68,67.5,71.3,73.8,69.5



INTERPRETATION - Since most of the points pass through the line, we can say that the sample follows Normal distribution.

```
> shapiro.test(y)

        Shapiro-Wilk normality test

data:  y
W = 0.9549, p-value = 0.05439
```

Conclusion: As the p-value is greater than α =0.05 hence the collected sample can be considered normal.

t TEST

Let Y= sounds levels outside Dinanath Mangeshkar Hospital (in dB)

Here μ_0=50

$H_O : \mu = 50$

Vs.

$H_1 : \mu > 50$

Test Statistic:

$$t_{cal} = \frac{71.108-50}{2.714198/\sqrt{50}}$$

$=54.9909$

And t_{tab}=1.67655

Since $t_{cal} > t_{tab}$, we may reject H_O at 5% level of significance.

We may say that the sounds levels(in dB) outside Dinanath Mangeshkar Hospital is greater than the critical value.

t TEST USING R

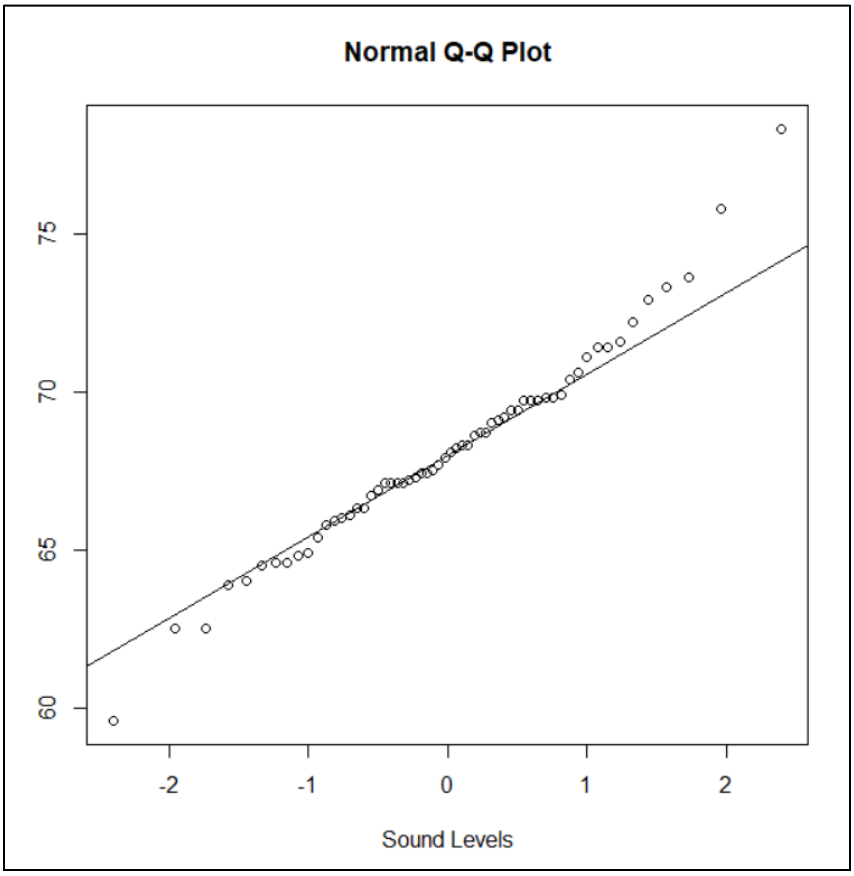
```
> t.test(y,mu=50,alt="g",conf.level=0.95)

        One Sample t-test

data:  y
t = 54.991, df = 49, p-value < 2.2e-16
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 70.46446      Inf
sample estimates:
mean of x
   71.108
```

<u>Conclusion:</u> As p-value is less than α =0.05 and μ_0=50 does not lie in the 95% confidence interval therefore H_0 is rejected. Hence the sound levels can be considered greater than 50 dB.

SANCHETI HOSPITAL

X=c(78.3,69.7,64.6,65.4,71.1,68.1,66.1,69,70.6,66.3,67.3,67.7,68.6,68.2,67.1,72.9,69.1,64.8,69.4,69.8,62.5,67.5,65.9,66.7,69.9,71.4,64,63.9,69.2,67.1,65.8,66.9,73.3,69.4,72.2,68.7,67.1,67.4,67.1,67.4,69.8,62,66,68.7,64.5,71.6,67.9,71.4,67.2,62.7,66.3,62.5,73.6,69.7,68.3,64.9,75.8,68.3,59.6,64.6,70.4)

```
> shapiro.test(data$sound.in.db)

        Shapiro-Wilk normality test

data:  data$sound.in.db
W = 0.97897, p-value = 0.3866
```



Conclsuion: Since the p-value is greater than α=0.05,the data is normal.

t TEST

Let x= sounds levels outside Sanjeevan Hospital (in dB)

Here μ_0=50 dB

$H_0 : \mu = 50$

Vs.

$H_1 : \mu > 50$

Test Statistic:

$$t_{cal} = \frac{\bar{x} - \mu}{s/\sqrt{n}}$$

$$t_{cal} = \frac{68.13833 - 50}{3.214959/\sqrt{60}}$$

$=20.04925$

And $t_{n-1,\alpha} = 1.671093$

Since $t_{cal} > t_{n-1,\alpha}$, we may reject H_0 at 5% level of significance.

We may say that the sounds levels(in dB) outside Sanjeevan Hospital is greater than the critical value.

t TEST USING R

```
> t.test(d$sound,mu=50,alt='g')

        One Sample t-test

data:  d$sound
t = 42.223, df = 59, p-value < 2.2e-16
alternative hypothesis: true mean is greater than 50
95 percent confidence interval:
 67.2972      Inf
sample estimates:
mean of x
    68.01
```

<u>Conclusion</u>: Since the p-value is less than $\alpha=0.05$ we reject H_O. Therefore average sound levels around the 100 m area around the hospitals have sound levels greater than the minimum allowed values .

TWO SAMPLE t TESTS

1. SANJEEVAN AND DEENANATH MANGESHKAR

```
> var.test(X2,Y,alt="t",conf.level=0.95)

        F test to compare two variances

data:  X2 and Y
F = 1.1068, num df = 56, denom df = 49, p-value = 0.7201
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.6364951 1.9052091
sample estimates:
ratio of variances
          1.106794
```

INTERPRETATION- Since, p-value = 0.7201 > 0.05, we accept H_0 i.e. the population variances are equal.

```
> t.test(X2,Y,alt="g",conf.level=0.95)

        Welch Two Sample t-test

data:  X2 and Y
t = 1.0107, df = 104.31, p-value = 0.1573
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -0.3496774        Inf
sample estimates:
mean of x mean of y
 71.65263  71.10800
```

INTERPRETATION -Since, p-value = 0.1573 > 0.05, we accept H_0 i.e. the mean of the two samples are equal at 5 % l.o.s.

```
> t.test(X2,Y,alt="l",conf.level=0.95)

        Welch Two Sample t-test

data:  X2 and Y
t = 1.0107, df = 104.31, p-value = 0.8427
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
     -Inf 1.438941
sample estimates:
mean of x mean of y
 71.65263  71.10800
```

INTERPRETATION -Since, p-value = 0.8427 > 0.05, we accept H_0 i.e. the mean of the two samples are equal at 5 % l.o.s.

```
> t.test(X2,Y,alt="t",conf.level=0.95)

        Welch Two Sample t-test

data:  X2 and Y
t = 1.0107, df = 104.31, p-value = 0.3145
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5239354  1.6131986
sample estimates:
mean of x mean of y
 71.65263  71.10800
```

INTERPRETATION -Since, p-value = 0.3145 > 0.05, we accept H_0 i.e. the mean of the two samples are equal at 5 % l.o.s.

CONCLUSION- The sound levels outside Sanjeevan hospital and Deenanath Mangeshkar hospital can be said to have the same average sound levels.

2.SANCHETI AND SAHYADRI

```
> var.test(d$sound.in.db,m)

        F test to compare two variances

data:  d$sound.in.db and m
F = 0.47848, num df = 56, denom df = 60, p-value = 0.00599
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.2851041 0.8068227
sample estimates:
ratio of variances
         0.4784835
```

INTERPRETATION -Since, p-value = 0.00599 < 0.05, we reject H_0 i.e. the population variances are cannot be considered equal.

- **Hence for further tests let us assume that the variances for the above samples are equal.**

```
> t.test(d$sound.in.db,m)

        Welch Two Sample t-test

data:  d$sound.in.db and m
t = -12.593, df = 107.09, p-value < 2.2e-16
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -8.599992 -6.260635
sample estimates:
mean of x mean of y
 67.98772  75.41803
```

INTERPRETATION - Since, p-value < 0.05, we reject H_0 hence $\mu_1 \neq \mu_2$.

```
> t.test(d$sound.in.db,m,alt='g')

        Welch Two Sample t-test

data:  d$sound.in.db and m
t = -12.593, df = 107.09, p-value = 1
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 -8.409316       Inf
sample estimates:
mean of x mean of y
 67.98772  75.41803
```

INTERPRETATION -Since, p-value > 0.05, we accept H_0 i.e. $\mu_1 = \mu_2$.

```
> t.test(d$sound.in.db,m,alt='l')

        Welch Two Sample t-test

data:  d$sound.in.db and m
t = -12.593, df = 107.09, p-value < 2.2e-16
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
      -Inf -6.451311
sample estimates:
mean of x mean of y
 67.98772  75.41803
```

INTERPRETATION -Since, p-value < 0.05, we reject H_0 hence $\mu_1 < \mu_2$.

CONCLUSION – The average sound levels outside Sahyadri hospitals are greater than that of Sancheti hospital.

CONCLUSION

- All the hospitals under case study had sound levels more than 50 dB which Is the minimum allowed value.

- We found that the mean sound levels around Sahyadri Hospital were the highest(75.42 dB)

- Sancheti Hospital had the least mean sound levels(67.98 dB).

- Sanjeevan Hospital and Deenanath Mangeshkar Hospital had same average sound levels.

LIMITATIONS

- We used ‘SOUND METER’ app from our cell phones to take the readings, hence there might be a possibility of minor errors in readings due to use of different cell phones (different microphones in them).
- We had to truncate some of our samples so that it would follow normal distribution for conducting t tests.
- We also assumed the population variances to be equal for one of the t tests for two samples.
- We conducted the tests for an interval of two hours on a single day, hence the result might not be valid for a sample of different size or for any of the other hospitals.

ACKNOWLEDGEMENT

We are thankful to the Statistics Department for their support and guidance throughout the period of study and execution of this project. We are also thankful to all those who helped us despite their busy schedule especially the teaching staff.

We also express our gratitude to HOD Sir for his valuable support.

Lastly, we would like to thank all those who directly and indirectly helped in successful completion of our project.

PROJECT BY :

Sakshi Choube (8214)

Pranav Nanaware (8240)

Aakanksha Dixit (8245)

Tanvi Khamkar (8247)

Tanmay Gajmal (8272)

Lalitagauri Walimbe (8274)

Aishwarya Bhujbal (8280)

Gargi Sarjine (8282)

THANK YOU...