```
# One Sample Z-Test
# sample size > 30
# should have at least one population parameter


# PS: Perform a Z Test to identify if the mean of students marks is 600
# H0: The mean of data = 600
# H1 The mean of data != 600
## 650,730,510,670,480,800,690,530,590,620,710,670,640,780,
# 650,490,800,600,510,700,
## 620,710,670,640,780,650,490,810,630,520,720
# (mean of sample - population mean)/(SD of sample/SQRT of sample size)

m = c(650,730,510,670,480,800,690,530,590,620,710,670,640,780,
         650,490,800,600,510,700,620,710,670,640,780,650,490,810,
         630,520,720)
m
z_test = (mean(m)-600)/(100/sqrt(31))
z_test

install.packages("distributions3")
library(distributions3)
?cdf
z = Normal(0,1)
z

p_value = 1 - cdf(z,z_test)
p_value

# Inference: We have sufficient evidence to reject the null hypothesis
# and accept the alternate hypothesis.

# Two Sample Z-Test
## Using x and y, find out the hypothesis statements, z stats value and p
value
## Mean of pop 1 is 600 and mean of pop 2 is 600
## *SD of pop 1 is 100, SD of pop 2 is 100
# H0: Mean of x = Mean of y, mean of x - mean of y = 0
# H1 Mean of x != Mean of y, mean of x - mean of y != 0

y <-
c(630,720,462,631,440,783,673,519,543,579,677,649,632,768,615,463,781,563,488,650,
       677,649,632,768,615,463,781,563,488,650,521)
length(y)

n1 = length(m)
n2 = length(y)

z_stat <- (mean(m)-mean(y))/sqrt((100*100/n1)+(100*100/n2))
z_stat

p_val = 1 - cdf(z,z_stat)
p_val

# Inference: We will accept the null hypothesis.
```

```
# One Sample T Test
# sample size is less than 30
# The population parameters are unknown
# Company claims net weight is 80gm
# Consumer claims net weight is less than 80gm
# H0 : Weight = 80
# H1: weight < 80

cookies =
c(87.7,80.01,77.28,78.76,81.52,74.2,80.71,79.5,77.87,81.94,80.7,82.32,75.78,80.19,83.91,79
length(cookies)

t.test(cookies, mu = 80, alternate = "less") # mu refers to mean
# If it's greater than,t.test(cookies, mu = 80, alternate = "greater")
# If it's not equal to than,t.test(cookies, mu = 80, alternate = "not equal")

# As p_value is 0.3476 is > than significance level 0.05 we accept null
hypothesis.

# Two tailed test
cookies2 =
c(80.22,79.73,81.1,78.76,82.03,81.66,80.97,81.32,80.12,78.98,79.21,81.48,79.86,81.06,77.96

# H0: mu(cookies 1) = mu(cookies2)
# H0: mu(cookies 1) > mu(cookies2)
length(cookies)
length(cookies2)

t.test(cookies,cookies2,paired = FALSE, alternate="greater")

# As p_value is 0.2818 is > than significance level 0.05 we accept null
hypothesis.

# When doing paired test, ensure that data x & y have equal length and PAIRED
= TRUE

# ANOVA
# more than two samples
# aov(marks~categorical variable)

data1 = chickwts
View(chickwts)
names(data1)
weight1 = aov(weight~feed, data = data1)
summary(weight1)

# H0: There is no difference in mean
# H1: There is a difference in mean
#p_value is 5.94e-10 which is less than significance 5%, hence we reject null
hypothesis.
```