Question:- Among the app users for disease information, at least 15% of them access disease information related to disease 6. Use an appropriate hypothesis test to check this claim at a = 0.05.

Understanding the Question:-
**Among the App Users - as per case-study page no 5 it is mentioned as 2Lakh Farmers are using.

** 15% of them assessing disease information related to Disease 6 (Leaf curl is the disease) (D6 is the code we have to use - refer to case study reading page no 10)

** Need to do a Hypothesis test to check this claim 5% (0.05)
--
Solution:-

```
# Read the data by using the clipboard function
data <- read.csv("clipboard", sep = "\t", header = T)

# Check the data structure
str(data)

# Install the dplyr Package(http://www.zevross.com/blog/2014/03/26/four-reasons-why-you-should-check-out-the-r-package-dplyr-3/)
install.packages("dplyr")
library(dplyr)

#Install the rlang Package
install.packages("rlang")
library(rlang)
View(data)

# Install tibble pacjkage
install.packages("tibble")
library(tibble)

remove.packages("tibble")
#tibble[123 x 26] (S3: tbl_df/tbl/data.frame)
#p = proportion of users accessing D6
#H0: p < 0.15
#Ha: p >= 0.15
```

```
d6Proportion <- sum(data$d6)/sum(data$Usage)
se <- sqrt ((0.15 * 1-0.15))/123
z_stat_d6 <- (d6Propotion - 0.15)/ se
d6pValue <- 1 - pnorm(z_stat_d6)
d6pValue

######

D6_Usage <- sum(data$D6)
D6_Usage

tot_Usage <- sum(data$D1, data$D2, data$D3, data$D4, data$D5, data$D6, data$D7,
data$D8, data$D9, data$D10, data$D11)
tot_Usage

p_Usage <- D6_Usage/tot_Usage
p_Usage

install.packages("tibble")
library(tibble)

test <- prop.test((x= 4295, n= 26830, p = .15, alternative = "greater"))
test
```

H0: p <= 15%
Ha: p > 15%

Result:-

The test results reports that, in sample evidence reports that 16% of the total number of users access D6 information.

The statistical significance of this reports that, this 16% is statistically different from the claim made (at least 15%), and the p value of the test reported a value of $0.000, hence we reject the null hypothesis, and support the alternative hypothesis claim. Thus, one can infer that, out of total number of users who access disease information, at least 15 % of them use D6 information.