Measuring the Utilization and Occupancy of Desktop Computers in laboratories of Sikh National College, Qadian, a case study

Koushal Kumar, Assistant Professor

Department of Computer applications, Sikh National College Qadian Gurdaspur, Punjab, India

Email: kaushal_kumar302@yahoo.com, Tel: +918968939621

Abstract: Computers have become a vital and irreplaceable part of college or university students lives. While initially conceived as a simple instrument for enhancing instructional experience, computers are now have been used in all facets of student life. Colleges and universities these days are now bound to provide sufficient amount of computing resources to their students in computer labs. The computer equipped libraries are not only making academic work easier for pupil but also helping in bettering academic collaboration among students from different college and universities. This paper analyses and find out the empirical results on the occupancy and usage of computers in the central library of our Sikh National college, Qadian, Gurdaspur, Punjab, India. The study conducted will reveals a vital question i.e., Do students still comes in library and accesses to computers in the library even in mobile phone and laptop era? The significance of our research is that it will help Library administrators in installing more computers in the library when the library is overloaded during peak hours. For solving our research problem we are going to apply statistical T-test. We have collected real time data from our college library in regular interval of time and for implementation purposes we have used "R language" with RStudio as an environment.

Keywords: Computers usage, occupancy, Library, R language, One Sample t-Test.

1. INTRODUCTION

Colleges and universities across the world now are expected to provide an adequate amount of computing resources to their students throughout the semester. The academic libraries in universities/colleges can be considered as the "heart" of the learning community, providing a peaceful and conducive place for students and faculty to conduct their research and advance their knowledge. Universities and colleges Libraries help the student's community in identifying and accessing knowledgeable resources in an academic institution with ease and convenience. In the beginning computer was conceived as a simple instrument for enhancing instructional experience, but now a day's computers are used by students from all disciplines. Computer systems in libraries are a key part of a university's technology offerings to students and one of the important ways that technology is put in the hands of students.

Colleges Libraries with latest Computer Systems are real digital information assets which can play a critical role in the education of today's students. Since last couples of years various initiatives have been taken to promote the awareness of the use of computers in higher education across all well-established universities. The appropriately furnished libraries with all cutting edge tech devices help the students in academic work as well as in improving academic collaboration among students from different colleges and universities around the globe. The well-equipped computer libraries help the student's in academic work as well as in bettering academic collaboration among students from different college/universities. *This experimental study has been conducted in Sikh National College, Qadian, which is a small rural*

college located in residential area of Gurdaspur, Punjab, India. Our college is affiliated with Guru Nanak Dev University, Amritsar and also has a valid NAAC (National Assessment and Accreditation Council) Score. The college enrols approximately 650 students, and employs 50 plus faculty and administrative staff.

In early days our computer libraries were established across campus in small space rooms however now passage of time and innovation in technology our college has completely furnished Labs with latest technology equipment's in it. Our college main Libraries has access to electronic journals of all streams, which provide key readings to students of all courses, and set up digital reserve systems to facilitate easy use of resources. Students of master degrees always prefer to use Library resources because libraries offer an environment conducive to study and research. For undergraduate students libraries are an important resource for assignments that encourage students to go beyond the course syllabus and helps in accessing vast array of resources under a single roof. We feel that pupil now have access to more accurate and precise information than ever before that too in few seconds. Most of our college pupil from different courses uses library computers for various reasons like, to implement scientific research, preparing projects, assignments, to access social media, to read newspaper online and to search a book or magazines using electronic catalogue. Many of students believe that the library offers computers and Internet service, which make it possible for them to complete their assignments on time. Our Library administrative staffs have been conducting Orientation Program from many years for newly admitted students in the beginning of every academic year. The main motive behind such orientation programs are to explain about the resources, facilities, services the library provides for students. They are taken round the library apart from training them in searching the library database, e-journals, e-mail and Internet browsing, timings etc. Our library has ultra-high speed connectivity to internet which helps in ensuring that Students' academic resources are not bounded by the physical constraint of the university boundaries. It has also been published in numerous researches that number of students visiting the libraries has fallen by a significant margin in last few years. The empirical results show that this is because of significant influence of mobile technologies in universities Libraries, especially because internet access becomes more affordable and reliable. Many Researches comes with empirical evidences that in last few years' students are showing less interest in using library desktop computers and purchasing their personal laptops and smart phones because they trusted more on their own devices rather than a machine they leave in universities at the end of each day. Thus the sole purpose of our this experimental study is to answer a critical question whether libraries should continue to provide dedicated desktop computers services and does our libraries are well equipped with modern facilities which can attract maximum students in libraries. For this study we have collected data from central library of our college about the occupancy of computers and then apply statistical methods to validate the average use of computers. In this experiment we are going to apply statistical one sample T-test which can generate several reports for making inference about a population mean based on a single sample. With this findings we can estimates whether students are showing interests in using library computer or not and library administrations than can reacts accordingly.

2. RESEARCH PROBLEM

Formally a hypothesis is just are statements about the sample of population we are going to use in research work. In this research study our research hypothesis is that the average or Mean number of computers occupied by students at any observed time is always greater than half (>50%) of total available computers in the central Library of our college. This hypothesis will also works as our research problem that we are going to solve ahead.

3. AIMS AND OBJECTIVES

The primary focus of this research is to know more about the occupancy of desktop computers by student's in the college library of our campus. The aim of this work is to explore more about the statistics status of library usage by our college students. The study aims to investigate the time when the occupancy of computers systems in Library is at maximum and when it is average or minimum. Our research findings will help the library administrators to gain in depth

understanding of the computers usage in library which will help them in taking proper actions.

4. METHODOLOGY

To solve our research problem we have used statistical analysis model over the collected data and mostly descriptions of statistical testing mainly focus on testing null hypotheses. Therefore to validate our hypothesis we will use statistical one sample t-Test. The main task in t-Test is to use statistics to evaluate two contrary hypotheses called Null Hypothesis denoted by H0 and alternate hypothesis denoted by H1. In t-Test when a hypothesis is tested, it will be based on difference between means if it is significant then Null hypothesis is rejected otherwise accepted. To implement t-Test and validate our hypothesis we will use R programming language with RStudio platform.

The Steps we have applied to test the hypothesis in this work is as follows.

- Construct a hypothesis for the given problem.
- Find out the population from sample space
- Select a valid and accurate sample to be tested for expected results
- After validating the sample size analyse the sample with respect to the population size.
- Interpret the findings and analysis the observed data.

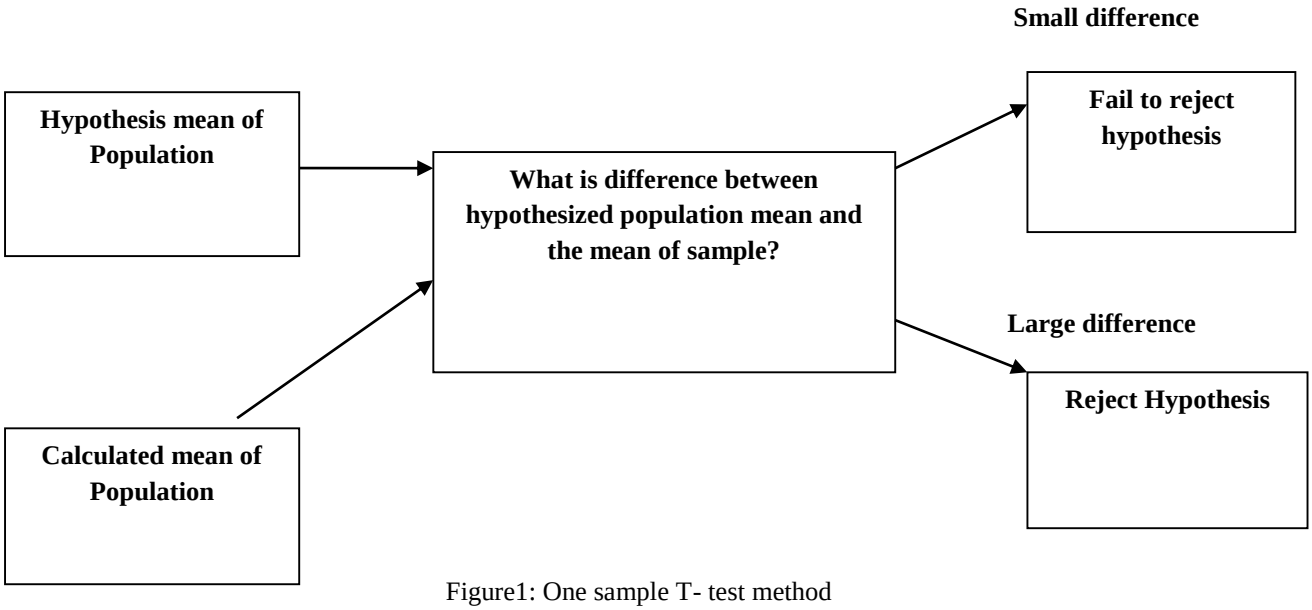


Figure1: One sample T- test method

5. DATA COLLECTION

The data collection component of research is common to all fields of study including pure sciences, Engineering, physical and social sciences, humanities, or business etc. Data collection is always plays a crucial role in any research related activity which always leads to more accurate results. In any kind of empirical study data are collected using some form of instrumentation like user studies, Questionnaires, Focus groups, surveys, Interviews, direct observation, case studies, and critical incidents with analysis of data. All these methods, type, and tools for data collection varies greatly between different studies but the ultimate aim of all data collection is the same, to gather precise and accurate information as evidence for analysis to formulate answers. Among all the methods of data collection the one you have to choose will depend upon the type and nature of problem you are going to solve. For our research problem the most appropriate data collection

method was direct observation method which is also an oldest method of data gathering. In **Direct observation method** of data collection evaluator watches and evaluate subject in his or her usual environment without altering that environment. As mentioned in our research statement we have visited our Sikh National College central library to collect the status of occupied and unoccupied computers at various times of the day. There were around 68 Computers present out of which 65 are in working conditions and rest of 3 are out of work as said by library assistant. Therefore our sample contains occupancy status of only 65 computers (i.e., N=65). We visited our library in different time of the days on 10,11,17,18 and 21th of this month October 2019. Our data set has 26 observations in which we have taken four variables of our interest like Number of PC occupied, Time, date of data collection, etc.

6. STATISTICAL SIGNIFICANCE (T-TEST)

In this research task we have applied one of the most popular and significant test named "**statistical one sample T-test"** which helps in generating reports for making inference about a population mean based on a single sample. T-Test is a type of a parametric test which works on normally distributed scale data. The necessary and sufficient condition of one sample t test is to compare the mean of a sample to a predefined hypothetical mean value. In two sample t –test we usually compare the means of two groups of populations to test our hypothesis. In a one-sample t-test we use the sample mean to estimate the population mean and we compare the average (or mean) of one group against the set average (or mean) and based upon the values we decide whether Null hypothesis should be accepted or rejected. After implementing t test if we reject the null hypothesis we can interpret it like difference between the observed sample mean and the hypothesized population mean is too big. When we fail to reject the null hypothesis, we are saying that the difference between the observed sample mean and the hypothesized population mean is probable if the null hypothesis is true. To validate our research problem we have consider two hypotheses under T-test. First one is the null hypothesis (H0) and (two-tailed) alternative hypothesis (H1). These hypotheses help in drawing conclusion whether a population mean is significantly different from some hypothesized value or not.

In calculating the t−test statistic, we use the formula: $\mathbf{t = \bar{x} - \mu 0 / \sqrt{s}\ n}$

Where:

t is the test statistic and has n−1 degrees of freedom.

$\bar{x}$ is the sample mean

μ0 is the population mean under the null hypothesis.

s is the sample standard deviation

√s/ n is the estimated standard error

One sample t- test can be expressed as:

H_0: μ= x ("the sample mean is equal to the [proposed] population mean")
H_1: μ ≠ x ("the sample mean is not equal to the [proposed] population mean")

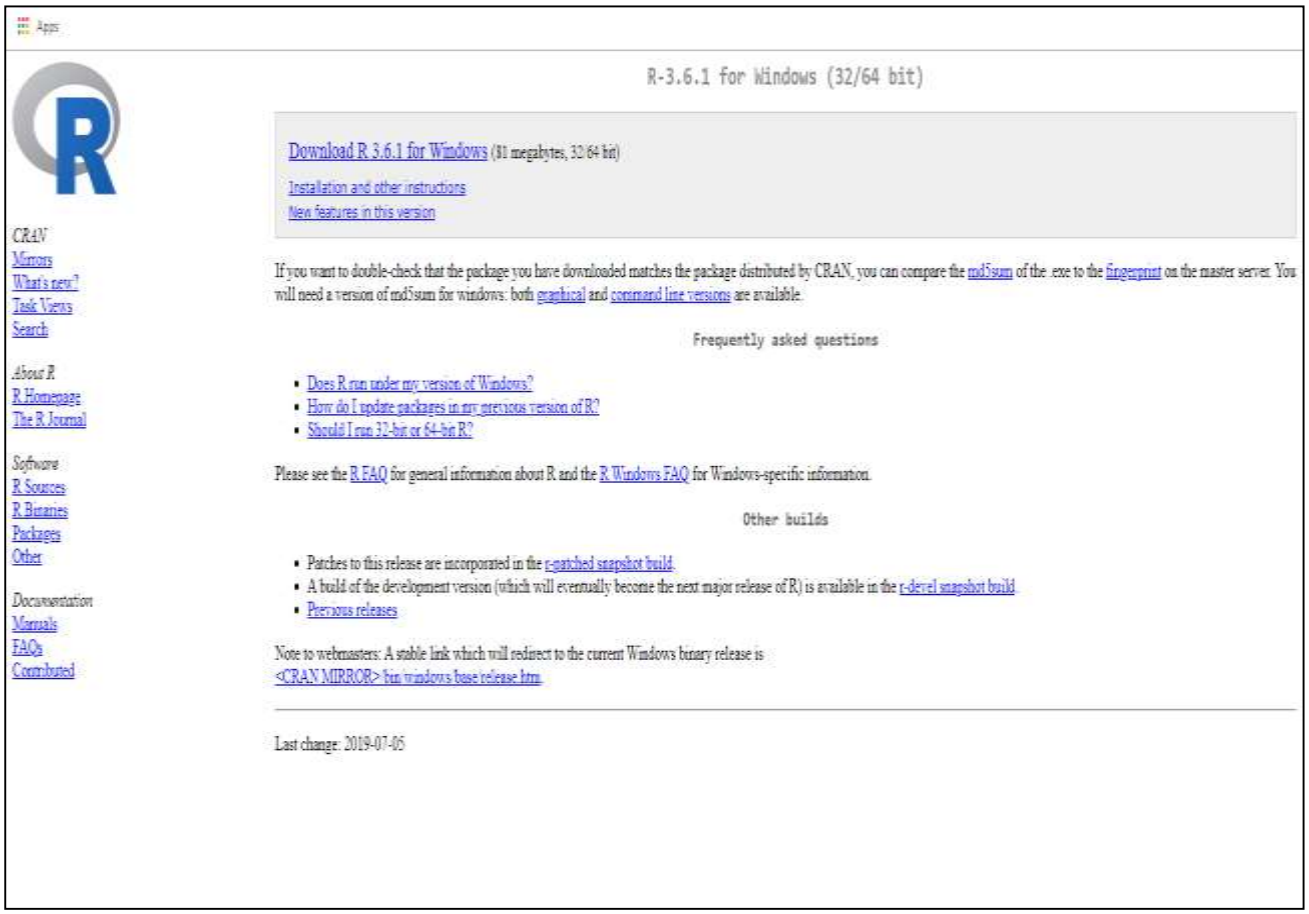
Where μ is a constant proposed for the population mean and x is the sample mean. To evaluate the statistical significance of the t-test, you need to compute the p-value which is known as value of significance. The p-value ranges from 0 to 1, and is interpreted as follow: A p-value less than 0.05 means you are strongly confident to reject the null hypothesis, thus alterative hypothesis (H1) is accepted. On the other hands if p-value higher than 0.05 indicates that you don't have enough evidences to reject the null hypothesis so alterative hypothesis gets rejected.

7. TOOLS USED

We have used R programming language to implement T test which is one of the extremely powerful programming language widely used for statistical computing and analysis across the globe. R is programming language developed in 1995 at university of Auckland New Zealand for statistical computing which has contributions from top computational statisticians. R language has gained its popularity among researchers because of its open-source nature along with excellent built in help systems, and its great graphics and visualization capabilities. R has a big community of its users across the world which provides easy

support online. R software environment's source code is written mainly in C, FORTRAN. R is a GNU Package and is freely available under GNU General Public License. In this work we have used an integrated development environment platform named RStudio. RStudio is a Graphical user interface which is more user friendly then writing programs in command line interface. To get started with RStudio, we need to install it in our machine. Figure 2 shows how to download RStudio, a software application that makes R language easier to use.

Figure2: Obtaining and Installing R for window platform



8. IMPLEMENTATION OF WORK

Task1:

Data Collection in Sorted Order (date): Random Data Collection 26 Observations Within October 2019.

According to problem statement mentioned in this research work we have visited our college library to collect the status of occupied and unoccupied computers. We visited library few days in various times of the day at dated 10,11,17,18 and 21th of this month October for data collections. Our data set has 26 observations in which we have taken four variables of our interest like Number of PC occupied, Time, date of data collection, etc. **Table I** shown below has data of various observations observed in library. Total number of working computers in library at those days was 65. Table I has four columns and 26 rows.

Table 1: Collected Data from Library

Date	Time	Total Number Of PC	Occupied PC
10-10-19	10:00 AM	65	27
10-10-19	10:45 AM	65	30
10-10-19	11:15 AM	65	40
10-10-19	12:30 PM	65	46
11-10-19	10:00AM	65	29
11-10-19	11:30 AM	65	35
11-10-19	1:00 PM	65	46
11-10-19	3:00PM	65	46
11-10-19	05:00 PM	65	30
17-10-19	10:00 AM	65	35
17-10-19	11:30 AM	65	38
17-10-19	1:00 AM	65	50

17-10-19	**2:00 PM**	**65**	**46**
17-10-19	**3:10 PM**	**65**	**51**
17-10-19	**4:25 PM**	**65**	**39**
18-10-19	**12:00 PM**	**65**	**37**
18-10-19	**2:30 PM**	**65**	**40**
18-10-19	**4:00 PM**	**65**	**36**
21-10-19	**11:00 AM**	**65**	**33**
21-10-19	**1:30 PM**	**65**	**43**
21-10-19	**2:00 PM**	**65**	**42**
21-10-19	**3:00 PM**	**65**	**41**
21-10-19	**3:30 PM**	**65**	**36**
21-10-19	**4:00 PM**	**65**	**33**
21-10-19	**4:30 PM**	**65**	**32**
21-10-19	**5:30 PM**	**65**	**30**

Task: 2

STATISTICAL RESEARCH FINDING USING ONE SAMPLE T-TEST

The primary condition to apply T-test on our data is that if the random sample taken from the population came from a distribution that is normal or approximately normal. In one sample T test hypothesis testing is based upon two statements: the **null hypothesis** and the **alternative hypothesis.** The alternative hypothesis assumes that some difference exists between the true mean (μ) and the hypothetical value (m), whereas the null hypothesis assumes that no difference exists between hypothetical value and obtained value. The alternative hypothesis can be supported only by rejecting the null hypothesis. Here we have constructed two hypotheses to support and validate our result problem.

a) **i. Null Hypothesis: Our null hypothesis is that mean number of computers systems occupied is more than half (i.e. >50%) of total available computers in college library.**

   ii. Alternative Hypothesis: Our Alternate hypothesis is that mean number of computers occupied is less than half (i.e. <50%) of total available computers in college Library.

b) **Carry out your T-test with statistical computing software**
   Tools used: R version 3.6.1

Implementing the One-Sample T-test in R-Studio

Step 1: We will begin with importing and exploring collected data in RStudio workspace from internal computer memory to do various statistical operations i.e. data gets imported in this phase. Our data set is saved with extension .CSV format which is an acceptable file format in RStudio. Once you have your dataset saved in CSV file format, you still need to set your working directory in R with following command.

Syntax: getwd()

setwd("<location of your dataset>")

After setting working directory following command is used to read the data from computer.

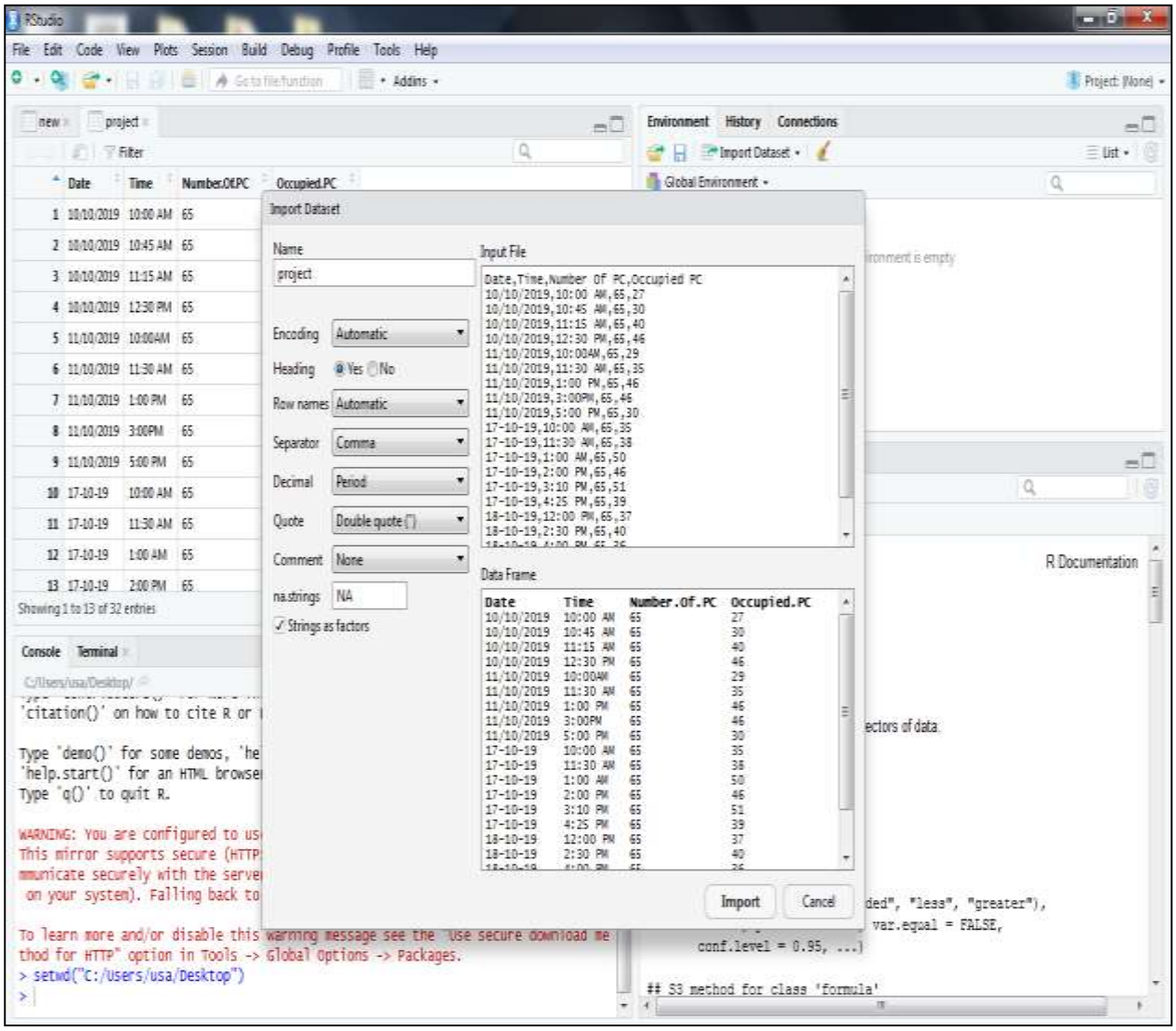
Project1<- read.csv ("C:/Users/Aus/Desktop/Project1.csv")

> View (Project1)

Here Project1 is the name of the file in which data is stored and file is placed in desktop. read.csv () function can be used to read in the data frame directly. In the second command view (Project1) that is used to view the data in RStudio workplace.

Note: Remember that you don't have to type the file's location if you have specified your working directory in R.

Figure 3: Imported data in Workspace of RStudio



Step 2: Data Validation and its correctness in R

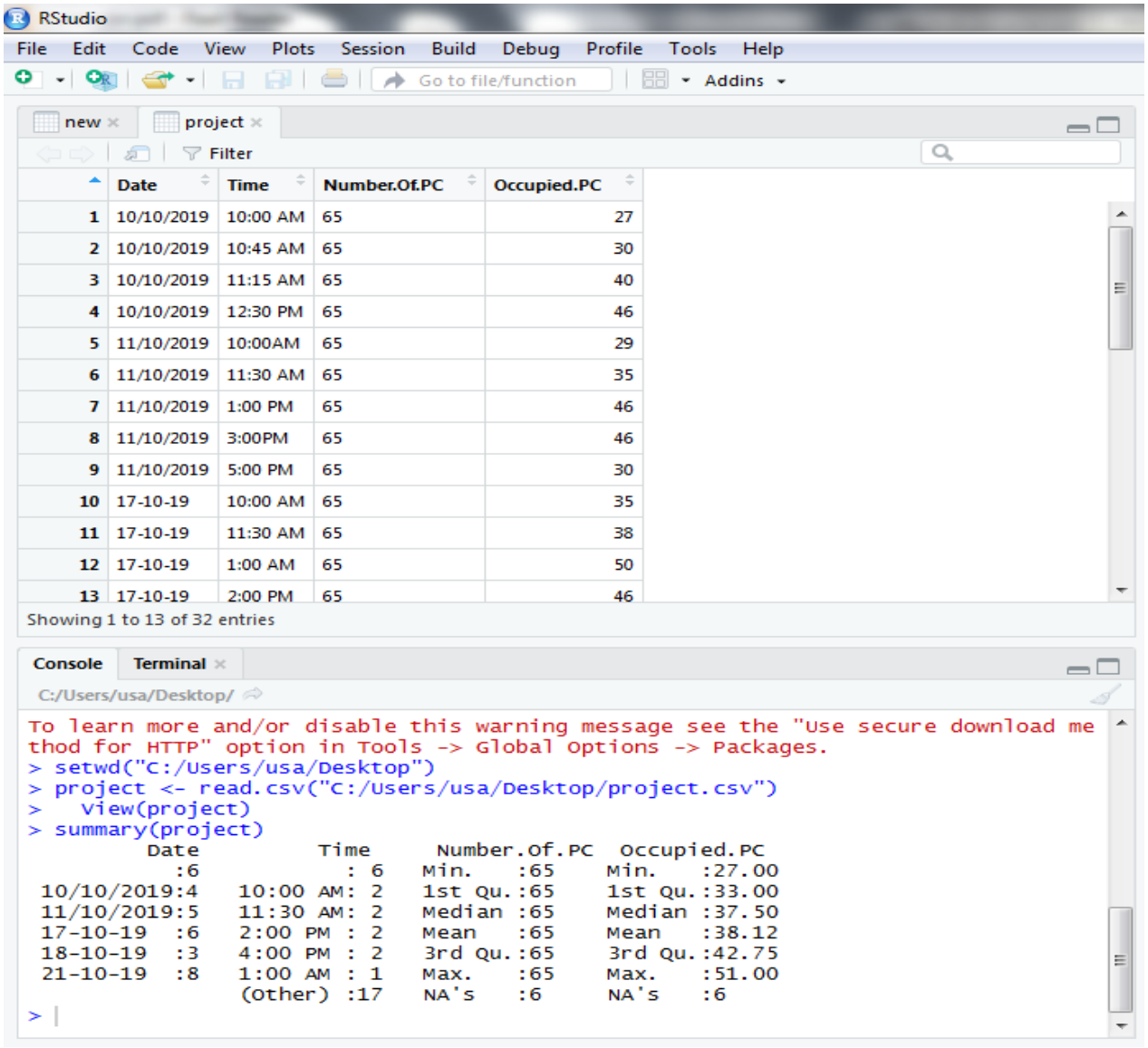
summarize the whole data for decision making

dim(Data)

The dim function of the R programming language **returns the dimension** of matrix or data frame imported. To get a better idea of the distribution of data your variables in the dataset we can use the summary () function. It returns Function to calculate GroupWise summary statistics of all the variables used in dataset. Figure 4 shows the values of every variable a set of descriptive statistics, depending on the type of the variable we have in the dataset.

Syntax: summary (Project1) # Project1 is the name of dataset.

Figure 4: Summarize values of each variable



R also supports another feature in which we can obtains only required number of rows from rows of a matrix instead of whole rows. For this we have applied head () function and it returns the first n rows of a matrix or data frame in R. Figure 5 shown below showing only 10 rows are visible out of 26.

Syntax: head (Project,10)

Figure 5: Output after executing head (10) function

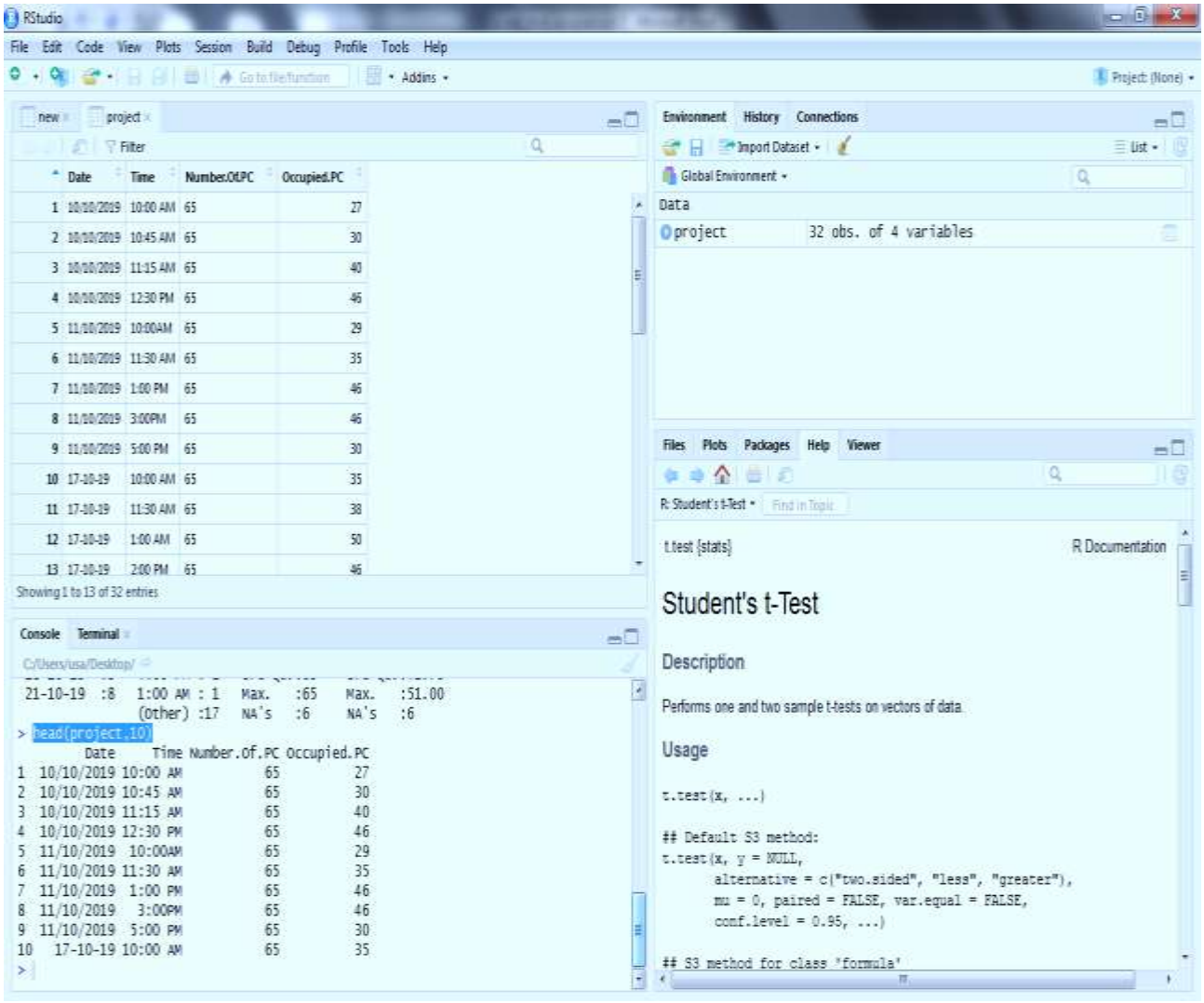
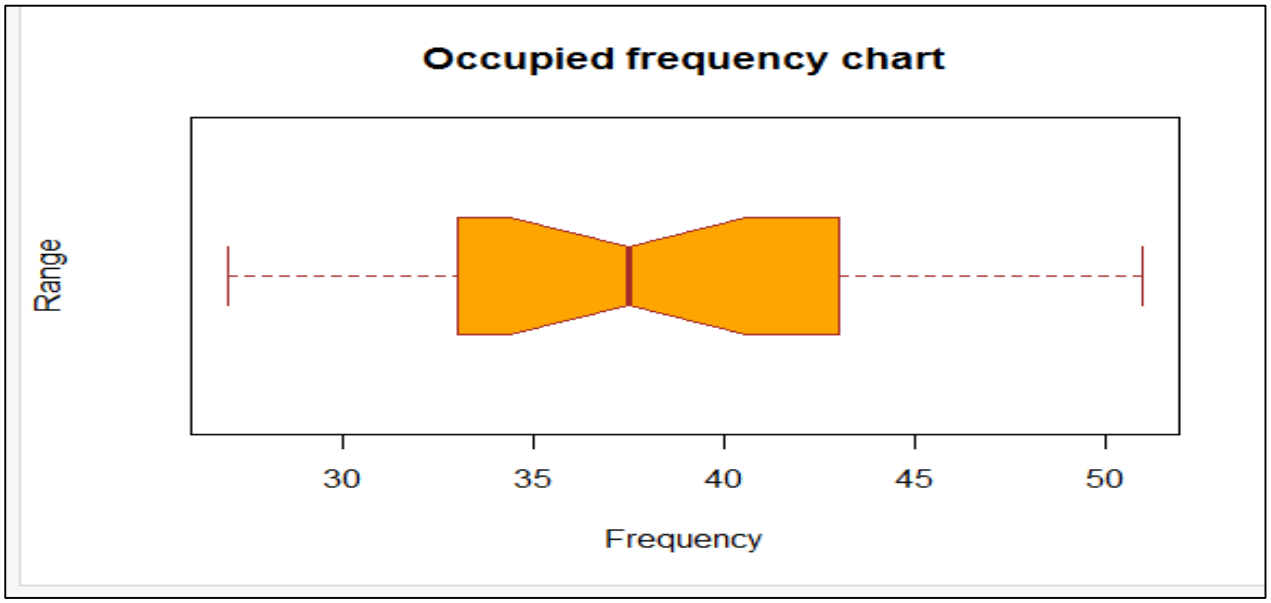


Figure 6 shown below demonstrates a Boxplot obtained from our data set. A boxplot is a standardized way of displaying the distribution of data based on a five number summary ("minimum", first quartile (Q1), median, third quartile (Q3), and "maximum"). Boxplots after interpretation indicates the measure of how well distributed is the data in a data set and it also indicates the skewness degree of data. The spacing's between the different parts of the box help in identifying the degree of scattering (spread) and skewness in the data, and distinguishes outliers.

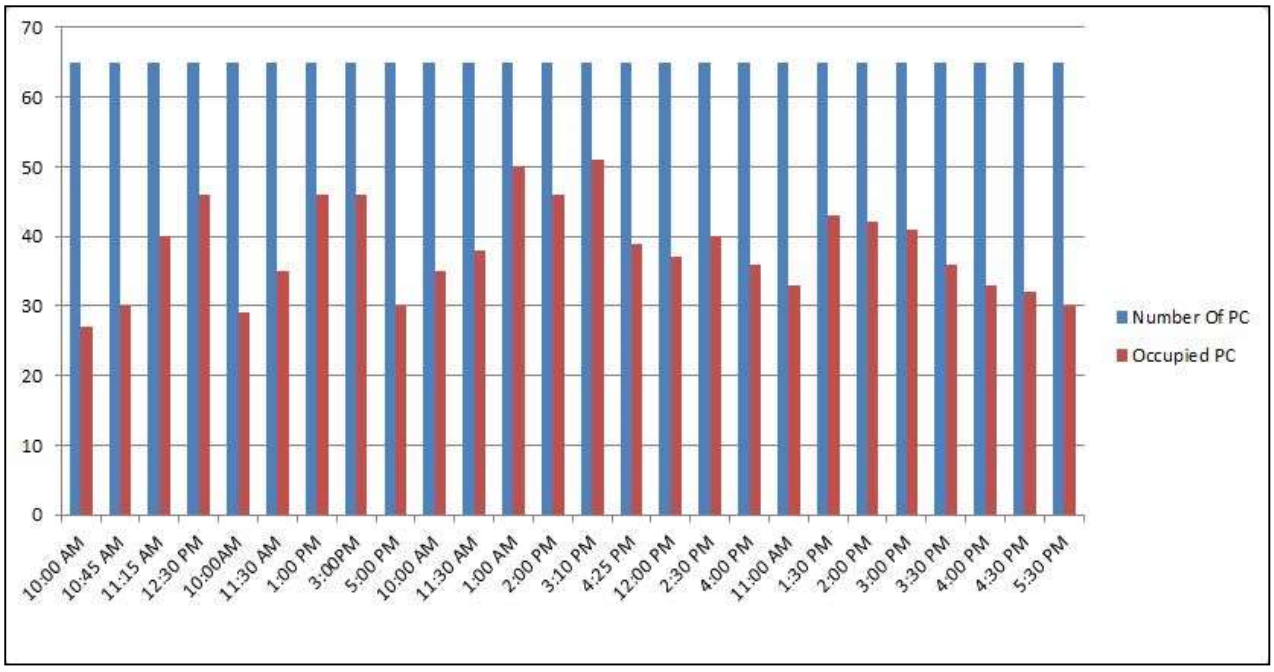
```
boxplot(project$Occupied.PC,
+ main = "Occupied frequency chart",
+ xlab = "Frequency",
+ ylab = "Range",
+ col = "orange",
+ border = "brown",
+ horizontal = TRUE,
+ notch = TRUE
```

Figure 6: Boxplot of Occupied PC frequency Chart indicating average value



A frequency distribution shows how frequently each different value in a set of data occurs. A **histogram** is the most widely used graph to **show** frequency distributions for better visual representation of sample data. Figure 7 shown below is a histogram we have obtained from our collected data. From obtained histogram we can easily interpret that maximum number of computer occupancy is during afternoon time between 1 to 3 pm and less number of computers occupied during morning and evening time. From histogram we can analysis that data is normally distributed and there is negligible skewness in data.

Figure 7: Histogram of computer occupancy



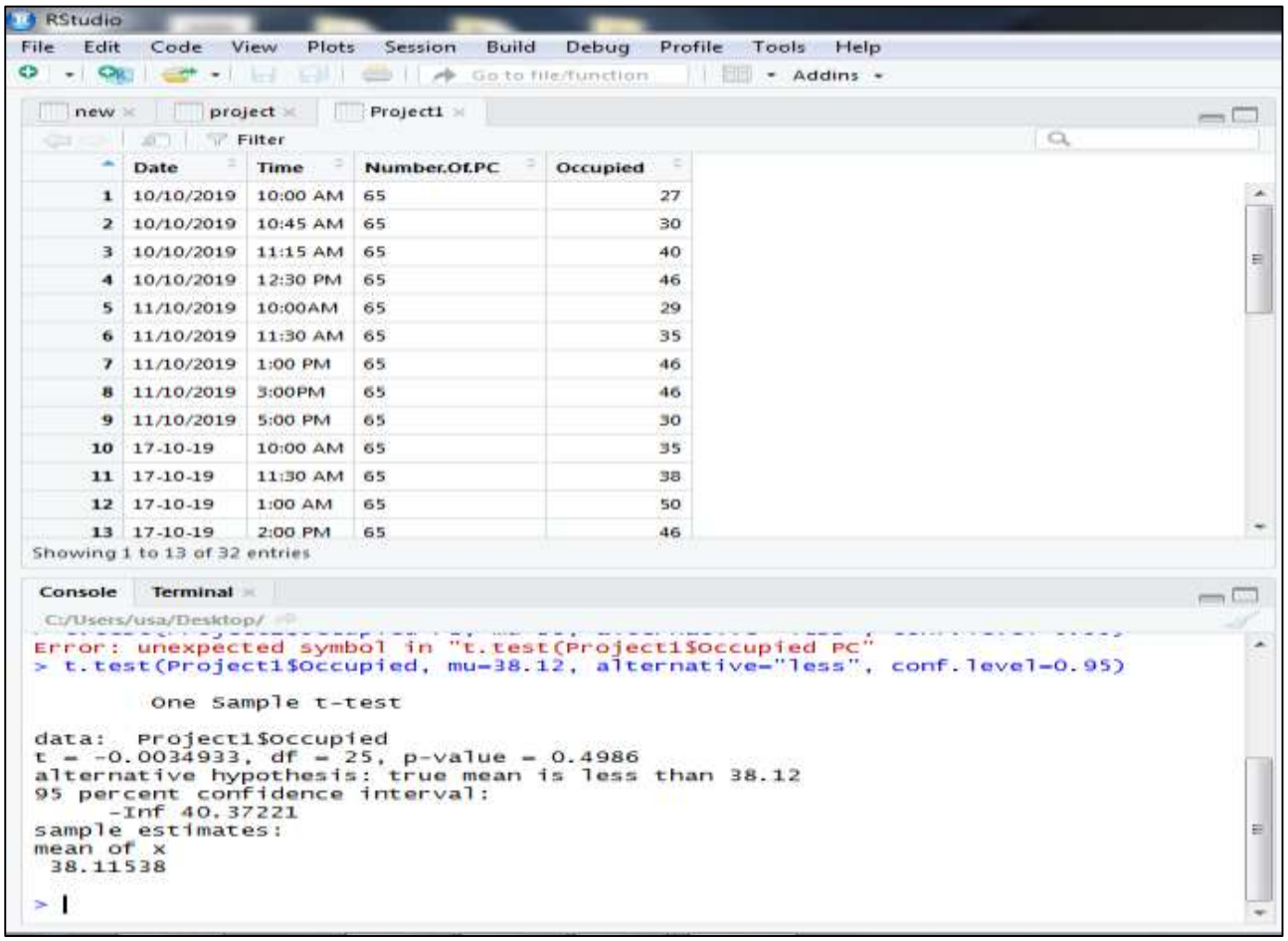
Step3: One-sample T-test

The t.test command In R is capable of performing three types of different test like one sample, two samples and paired test on data. To perform one-sample t-test, the R inbuilt function t.test () function is used as written below.

```
t.test(Project1$Occupied, mu=38.12, alternative="less", conf.level=0.95)
```

Output of t Test:

Figure 8: One Sample T-Test with 95% level of Significance



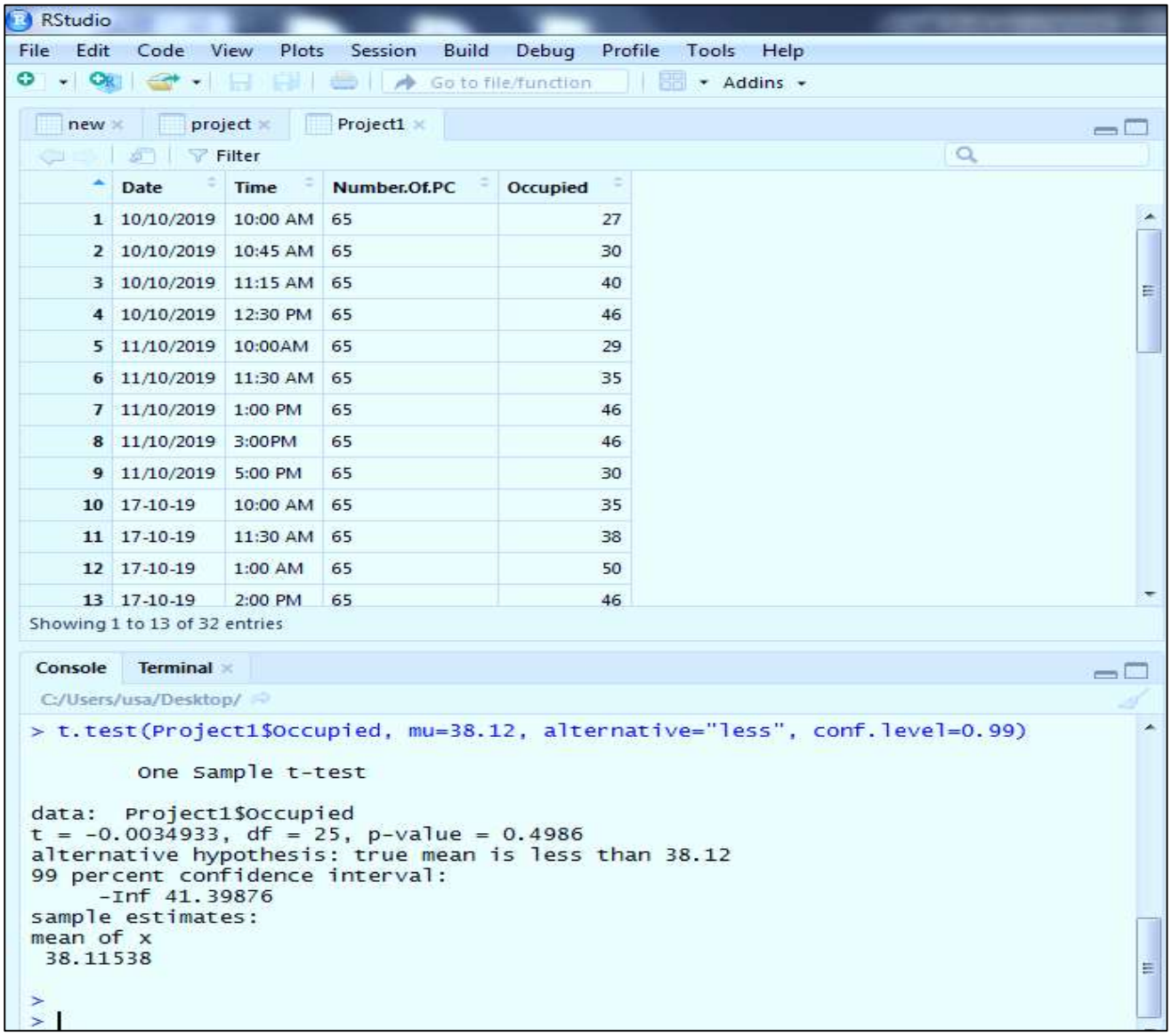
The statistics findings with their meaning are given below:

- ✓ t is the t-test statistic value (t = -0.003493)
- ✓ df is the degrees of freedom (df= 25),
- ✓ P-value is the significance level of the t-test (p-value = 0.4986).
- ✓ Inf is the confidence interval of the mean at 95% (conf.int = [40.37221]);
- ✓ Sample estimates is he mean value of the sample (mean = 38.11538).

T test with Significance level of 99%:

```
t.test (Project1$Occupied, mu=38.12, alternative="less", conf.level=0.99)
```

Figure 9: One Sample T-Test with 99% level of Significance



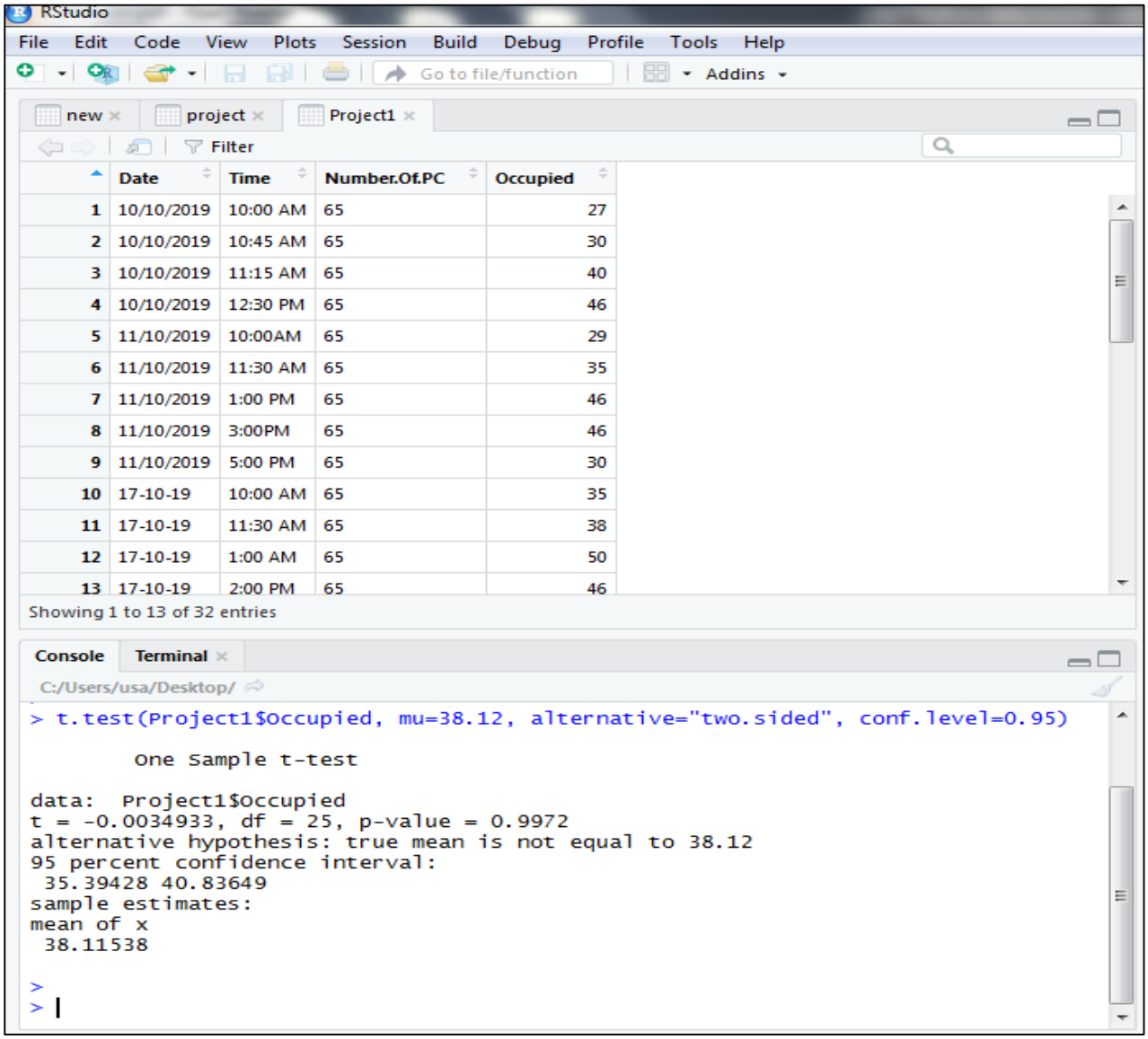
In this result our findings with their meaning are given below:

- ✓ t is the t-test statistic value (t = -0.003493),
- ✓ df is the degrees of freedom (df= 25),
- ✓ P-value is the significance level of the t-test (p-value = 0.4986).
- ✓ Inf is the confidence interval of the mean at 95% (conf.int = [-41.39876]);
- ✓ Sample estimates is he mean value of the sample (mean = 38.1153).

One sample T test with two sided tails and 95% level of significance.

t.test(Project1$Occupied, mu=38.12, alternative="two.sided", conf.level=0.95)

Figure 10: One Sample T-Test with two-sided 95% level of Significance

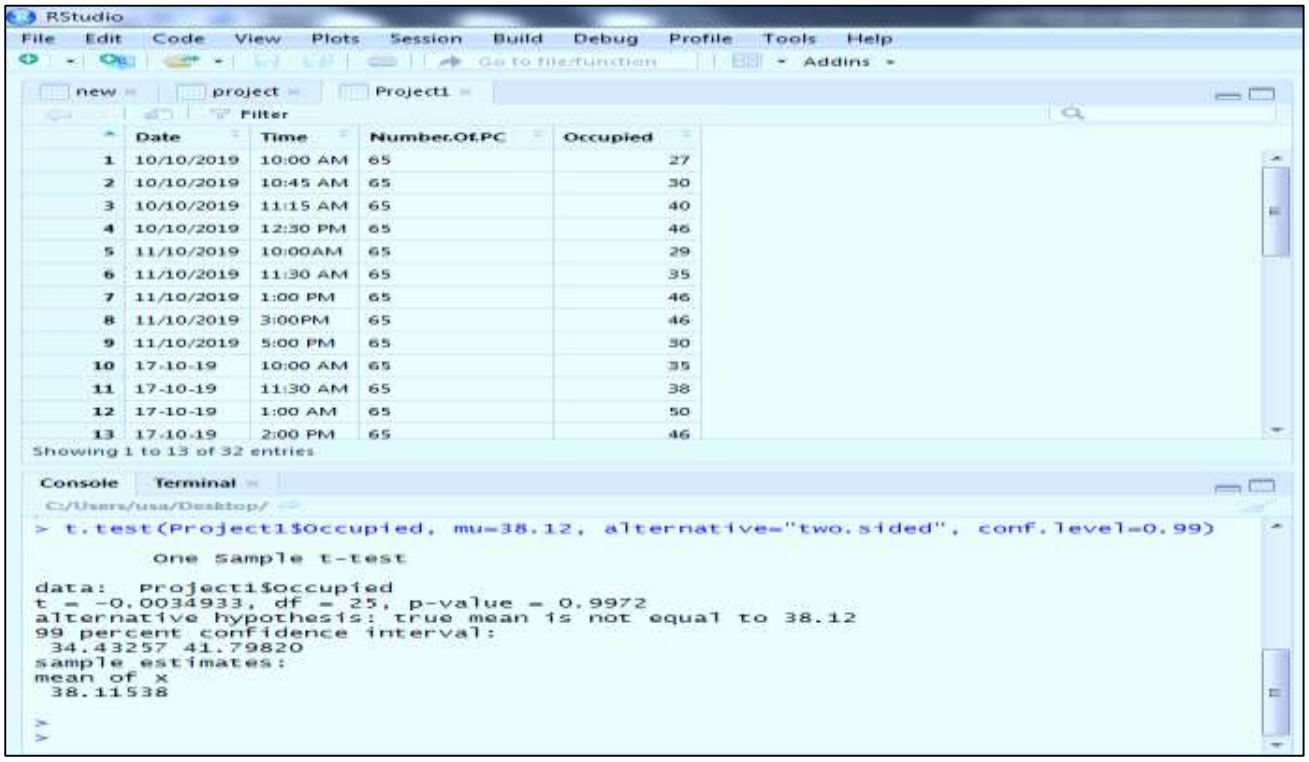


Here t value is -0.003493, df =25, p value= 0.9972 and mean value at 95% confidence interval lies between35.39428 to 40.83649.

One sample T test with two sided tails and 99% level of significance

t.test (Project1$Occupied, mu=38.12, alternative="two. Sided", conf.level=0.99)

Figure 11: One Sample T-Test with two-sided 99% level of Significance



Here t value is -0.0034933, df =25, p value= 0.9972 and mean value at 99% confidence interval lies between34.43257 to 41.79820

Histogram of data:

Syntax: hist(Project1$Occupied)

Figure 12: Histogram of the occupied systems

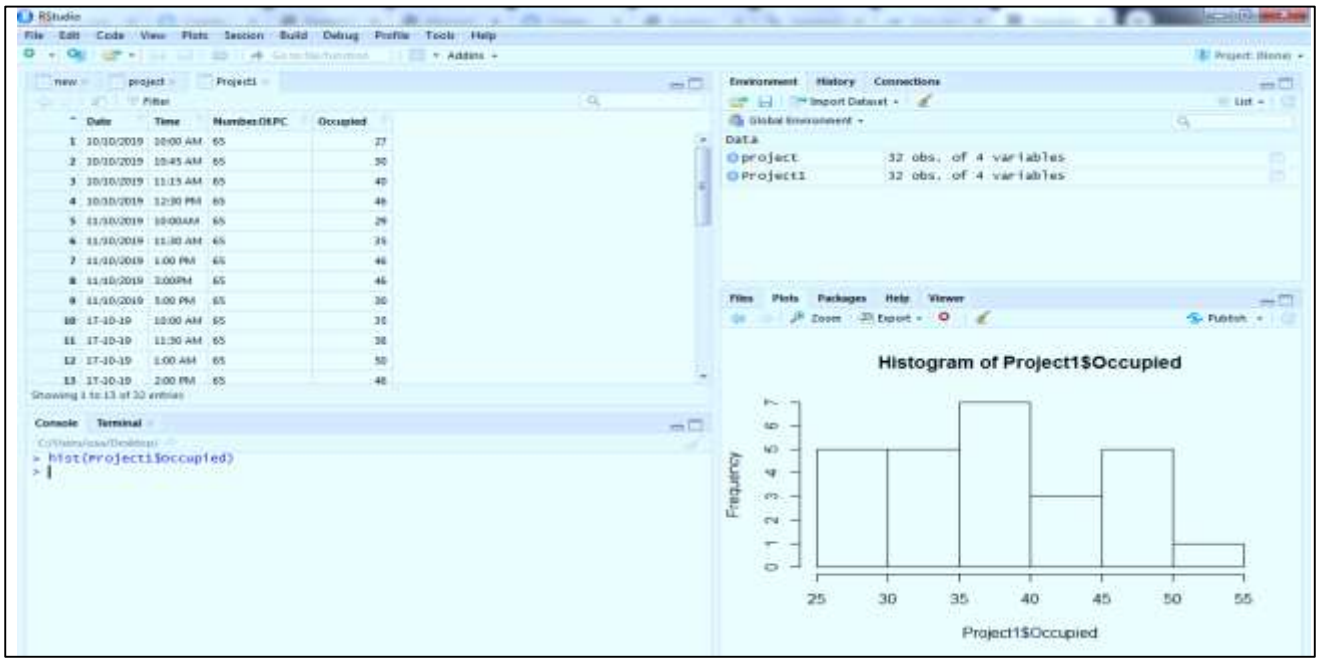


Figure13: A line chart of occupancy of computer during different time of the day

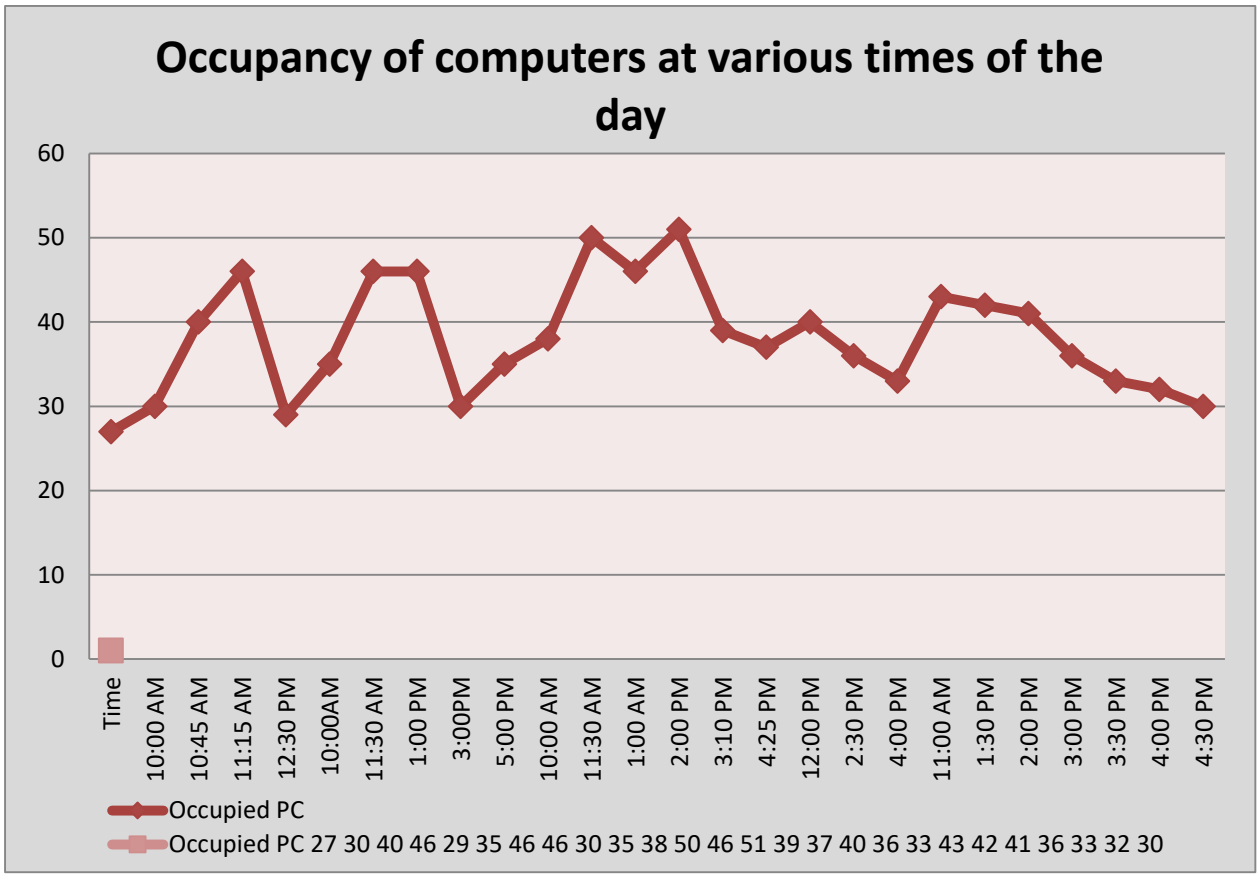


Table II: Occupied System statistics

Mean	Median	Max	Min	1st Qu	3rd Qu.
38.12	37.50	51	27	33.00	42.75

9. RESULTS AND DISCUSSIONS

To validate our hypothetical mean (Null Hypothesis) we have assumed in our research problem first we have to check the outcomes of our T-test to determine whether statement is true or not. Before analysing our research findings, we have to compare our computed t-test value to a critical value. This critical value is a number based on the type of test (one tailed vs. two tailed), degrees of freedom etc. The critical value helps us in proving whether null hypothesis is true or not.

Our test statistics that we find using Summary function as shown in figure 4 is written in tabular form in **Table II** above reveals that our Null Hypothesis holds true since mean value for computers occupied in college library comes out is 38.2 which is greater than half (32.5) of the computers present in library, i.e. 38.2> 32.5. In other words we can say mean usage of computers in library is greater than half of the computer available in library which helps us in accepting our null hypothesis.

Moreover the results of one sample t-Test shows p values that we have observed is greater than level of significance 95% (0.05), i.e. 0.4986>0.05. Since p value is Larger than critical value which demands to accepts the null hypothecs that we have assumed in the beginning at 5% confidence interval. Therefore we do not have enough evidence to conclude that the difference between the population means is statistically significant which means we strongly accepts Null hypothesis.

As shown in results (Figure 10, 11) two sided one sample t Test the test statistic t= -0.0034933 is same in both 95% and 99% confidence interval. The two-tailed p-value is p> |t|, i.e, (p>0.003493)+(p< --0.0034933). From the outcomes of two sided t Test we can interpret that t-distribution is symmetric about zero, these two probabilities are equal: P > |t| = 2 * P(< -0.0034933). . The smaller the p-value, the more significant is the difference between the distributions of the two samples. Figure 13 depicting a line graph of computers occupancy during different times of the day, which clearly shows that maximum occupancy of computer is during noon and minimum is during morning and evening time.

10. CONCLUSION

From obtained results we can easily interpret that maximum number of computer occupancy is found during afternoon time between and less number of computers occupied during morning and evening time. T test results also support the acceptance of our null hypothesis which we had constructed as problem statement that mean usage of computers in library is greater than half of the computer available in library. From empirical results we can say that numbers of computers available in library are sufficient for students and there is no need to buy new computers as per current requirements.

Future work: In future we can extend our work and can find the following facts.

- Determine uses of computer based on Male / Female Ratio.
- Determine the reasons for library computer usage
- Find out the statistics figures about various sites most of the students interested in.

References

[1] Yim KH, Nahm FS, Han KA, Park SY. "Analysis of Statistical Methods and Errors in the Articles Published in the Korean Journal of Pain", 2010 Mar; 23(1): 35–41.

[2] McDonald, J.H., and K.W. Dunn, 2013 Statistical tests for measures of colocalization in biological microscopy. Journal of Microscopy, J Microsc. 2013 Dec; 252(3): 295–302.

[3] Cohen, J. 1988. “Statistical Power Analysis for the Behavioral Sciences”, 2nd Edition. Routledge Hillsdale, Lawrence Erlbaum Associates, Publishers.

[4] Tae Kyun Kim, 2015,”T test as a parametric statistic”, Department of Anaesthesia and Pain Medicine, Pusan National University School of Medicine, Busan, Kore the Korean Society of Anaesthesiologists.

[5] Comprehensive R Archive Network (CRAN). 2016. Available at: https://cran.r-project.org/

[6] Vlad Krotov, 2017,”A tutorial on A Quick Introduction to R and RStudio”.