

SIDDHI SINGH

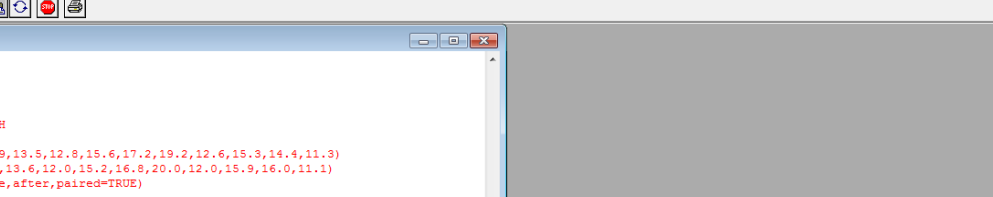
Paired t-test And F- (Variance Ratio Test)

HYPOTHESIS TESTS FOR MEAN DIFFERENCES: PAIRED DATA-t-TEST

A school athletics has taken a new instructor, and want to test the effectiveness of the new type of training proposed by the new instructor comparing the average times of 10 runners in the 100 meters. The results are given below (time in seconds)

<i>Before training</i>	12.9	13.5	12.8	15.6	17.2	19.2	12.6	15.3	14.4	11.3
<i>After training</i>	12.7	13.6	12.0	15.2	16.8	20.0	12.0	15.9	16.0	11.1

Solution: In this case we have two sets of paired samples, since the measurements were made on the same athletes before and after the workout. To see if there was an improvement, deterioration, or if the means of times have remained substantially the same (hypothesis H_0), we need to make a Student's t -test for paired samples, proceeding in this way



The screenshot shows the RGui (64-bit) interface. The menu bar includes File, Edit, View, Misc, Packages, Windows, and Help. The toolbar contains icons for opening, saving, running, and other functions. The R Console window is active, displaying the following R code and output:

```
>
>
>
> #17BIT0028
> #SIDDIHI SINGH
>
> before=c(12.9,13.5,12.8,15.6,17.2,19.2,12.6,15.3,14.4,11.3)
> after=c(12.7,13.6,12.0,15.2,16.8,20.0,12.0,15.9,16.0,11.1)
> t.test(before,after,paired=TRUE)

Paired t-test

data: before and after
t = -0.21331, df = 9, p-value = 0.8358
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
 -0.5802549  0.4802549
sample estimates:
mean of the differences
      -0.05

>
>
>
> |
```

Interpretation :-

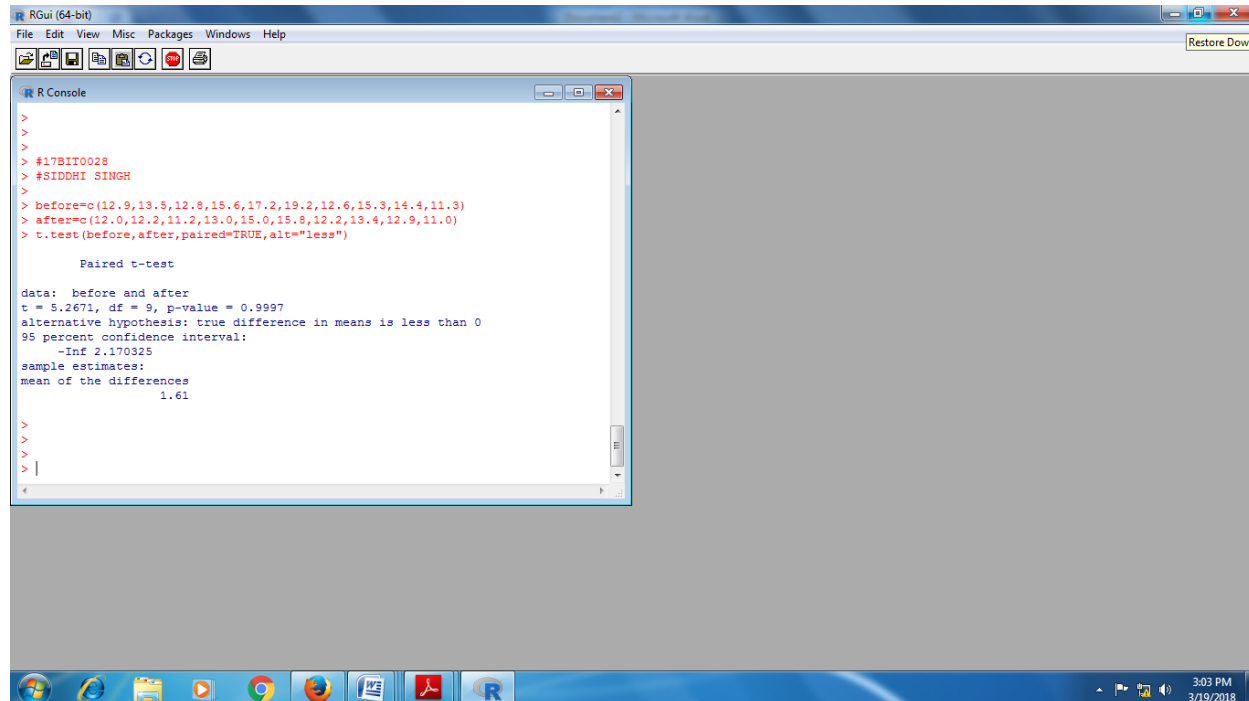
The p-value is greater than 0.05, then we do not reject the hypothesis H_0 of equality of the averages and conclude that the new training has not made any significant improvement to the team of athletes.

Problem 2 :-

Suppose now that the manager of the team (given the results obtained) fired the coach who has not made any improvement, and take another, more promising. We report the times of athletes after the second training:

Before training:	12.9	13.5	12.8	15.6	17.2	19.2	12.6	15.3	14.4	11.3
After the second training:	12.0	12.2	11.2	13.0	15.0	15.8	12.2	13.4	12.9	11.0

Now we check if there was actually an improvement, ie perform a t-test for paired data, specifying in R to test the alternative hypothesis H_1 of improvement in times. To do this simply add the syntax `alt = "less"` when you call the t-test:



```
RGui (64-bit)
File Edit View Misc Packages Windows Help
Restore Down

R Console
> 
> 
> #17BIT0028
> #SIDDIH SINGH
> before=c(12.9,13.5,12.8,15.6,17.2,19.2,12.6,15.3,14.4,11.3)
> after=c(12.0,12.2,11.2,13.0,15.0,15.8,12.2,13.4,12.9,11.0)
> t.test(before,after,paired=TRUE,alt="less")

Paired t-test

data: before and after
t = 5.2671, df = 9, p-value = 0.9997
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 2.170325
sample estimates:
mean of the differences
1.61

> 
> 
> 
```

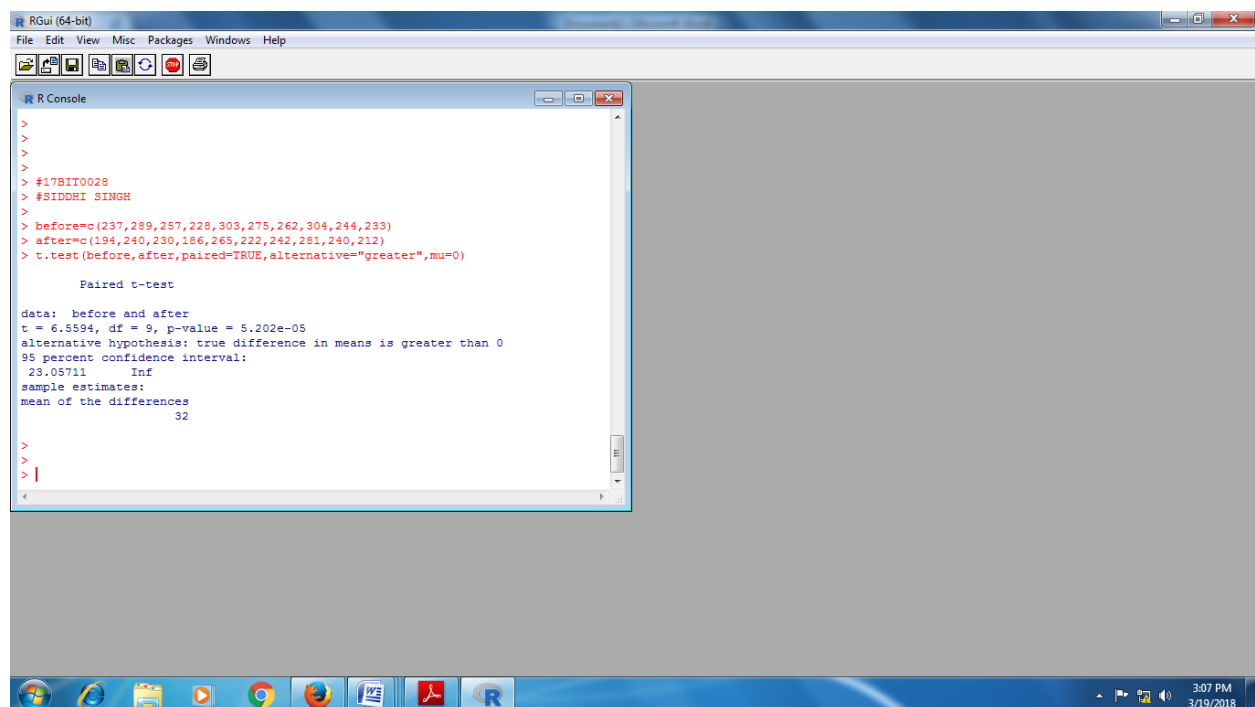
Interpretation:-

In response, we obtained a p-value well above 0.05, which leads us to conclude that we can reject the null hypothesis H_0 in favour of the alternative hypothesis H_1 : the new training has made substantial improvements to the team.

Problem 3 : Consider the paired data below that represents cholesterol levels on 10 men before and after a certain medication

Before(x)	237	289	257	228	303	275	262	304	244	233
After(y)	194	240	230	186	265	222	242	281	240	212

Test the claim that, on average, the drug lowers cholesterol in all men. I.e., test the claim that $\mu_d > 0$. Test this at the 0.05 significance level.



```
RGui (64-bit)
File Edit View Misc Packages Windows Help

R Console
> 
> 
> #17BIT0028
> #SIDDIHI SINGH
> 
> before=c(237,289,257,228,303,275,262,304,244,233)
> after=c(194,240,230,186,265,222,242,281,240,212)
> t.test(before,after,paired=TRUE,alternative="greater",mu=0)

Paired t-test

data: before and after
t = 6.5894, df = 9, p-value = 5.202e-05
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 23.05711      Inf
sample estimates:
mean of the differences
          32

> 
> 
> |
```

Interpretation :-

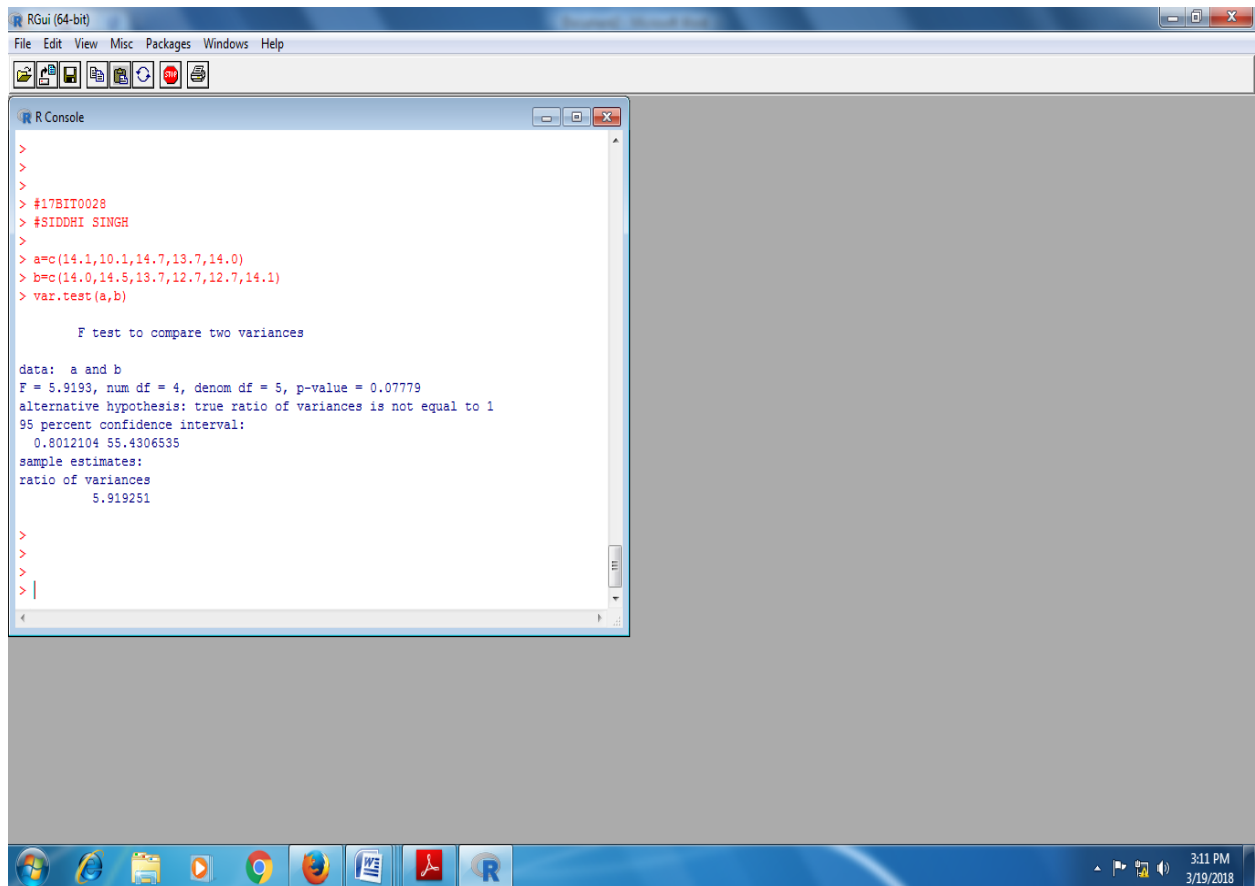
We can reject the null hypothesis and support the claim because the P-value ($\approx 5.2 \times 10^{-5}$) is less than the significance level.

F Test to Compare Two Variances

Problem 1 :-

Five Measurements of the output of two units have given the following results (in kilograms of material per one hour of operation) .Assume that both samples have been obtained from normal populations, test at 10% significance level if two populations have the same variance.

Unit A	14.1	10.1	14.7	13.7	14.0
Unit B	14.0	14.5	13.7	12.7	14.1



The screenshot shows the RGui (64-bit) interface. The R Console window displays the following code and output:

```
>
>
> #17BIT0028
> #SIDDHI SINGH
>
> a=c(14.1,10.1,14.7,13.7,14.0)
> b=c(14.0,14.5,13.7,12.7,14.1)
> var.test(a,b)

      F test to compare two variances

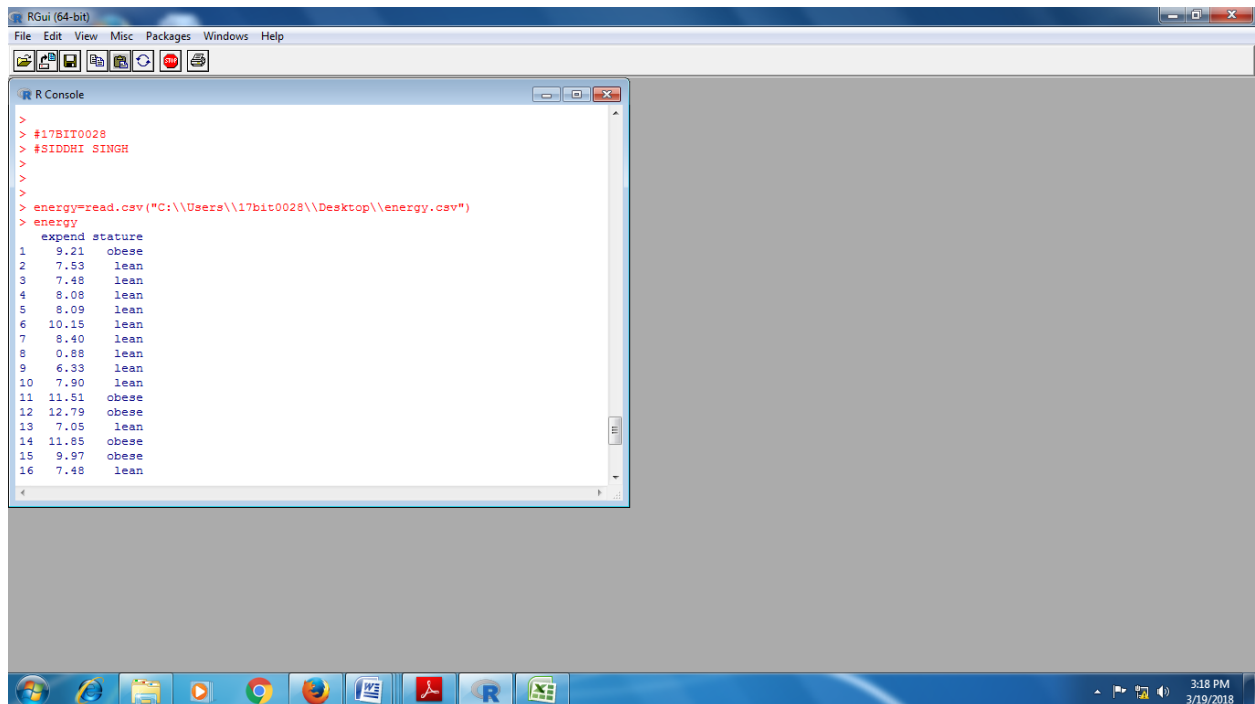
data:  a and b
F = 5.9193, num df = 4, denom df = 5, p-value = 0.07779
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.8012104 55.4306535
sample estimates:
ratio of variances
 5.919251

>
>
> |
```

The Windows taskbar at the bottom shows the time as 3:11 PM on 3/19/2018.

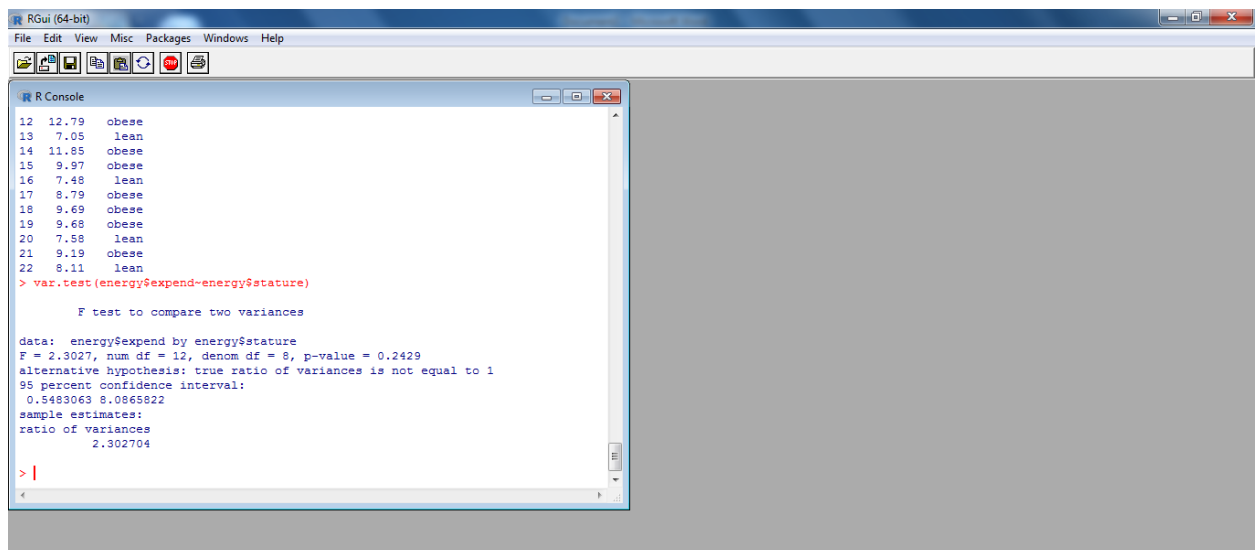
Inference : here p value > 0.05 ,then there is no evidence to reject the null hypothesis.

Problem 2: Energy Data :- (Variance Ratio-test)



```
RGui (64-bit)
File Edit View Misc Packages Windows Help

R Console
>
> #17BIT0028
> #SIDDIHI SINGH
>
> energy=read.csv("C:\\Users\\17bit0028\\Desktop\\energy.csv")
> energy
  expend stature
1    9.21  obese
2    7.53   lean
3    7.48   lean
4    8.08   lean
5    8.09   lean
6   10.15   lean
7    8.40   lean
8    0.88   lean
9    6.33   lean
10   7.90   lean
11   11.51  obese
12   12.79  obese
13    7.05   lean
14   11.85  obese
15    9.97  obese
16    7.48   lean
```



```
RGui (64-bit)
File Edit View Misc Packages Windows Help

R Console
12 12.79  obese
13  7.05   lean
14 11.85  obese
15  9.97  obese
16  7.48   lean
17  8.79  obese
18  9.69  obese
19  9.68  obese
20  7.58   lean
21  9.19  obese
22  8.11   lean
> var.test(energy$expend~energy$stature)

      F test to compare two variances

data: energy$expend by energy$stature
F = 2.3027, num df = 12, denom df = 8, p-value = 0.2429
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.5483063 8.0865822
sample estimates:
ratio of variances
      2.302704

> |
```

Inference :

Here p value > 0.05 ,then there is no evidence to reject the null hypothesis.

LAB-11

lab11

19/3/18 -A

17BIT0028

PAIRED T-TEST AND F-TEST

- > before = c(12.9, 13.5, 12.8, 15.6, 17.2, 19.2, 12.6, 15.3, 14.4, 11.3)
- > after = c(12.7, 13.6, 12.0, 15.2, 16.8, 20.0, 12.0, 15.9, 16.0, 11.1)
- > t.test(before, after, paired = TRUE)

- > before = c(12.9, 13.5, 12.8, 15.6, 17.2, 19.2, 12.4, 15.3, 14.4, 11.3)
- > after = c(12.0, 12.2, 11.2, 13.0, 15.0, 15.8, 12.2, 13.4, 12.9, 11.0)
- > t.test(before, after, paired = TRUE, alt = "less")

- > before = c(237, 289, 257, 303, 275, 262, 304, 244, 233)
- > after = c(194, 240, 230, 186, 265, 222, 242, 281, 240, 212)
- > t.test(before, after, paired = TRUE, alternative = "greater", mu = 0)

- > a = c(14.1, 10.1, 14.7, 13.7, 14.0)
- > b = c(14.0, 14.5, 13.7, 12.7, 12.7, 14.1)
- > var.test(a, b)

- > energy = read.csv("C:\\User\\17bit0028\\Desktop\\energy.csv")
- > var.test(energy\$expend, energy\$statute)