SESSION 14: GETTING STARTED WITH R

Assignment 1

Data Analytics

Table of Contents

1 Introduction 3 2 Objective 3 3 Prerequisites 3 4 Associated Data Files 3

5 Problem Statement 3 6 Expected Output 4 7 Approximate Time to Complete Task 4

Data Analytics

1. Introduction

This assignment will help you understand the concepts learnt in the session.

1. Objective

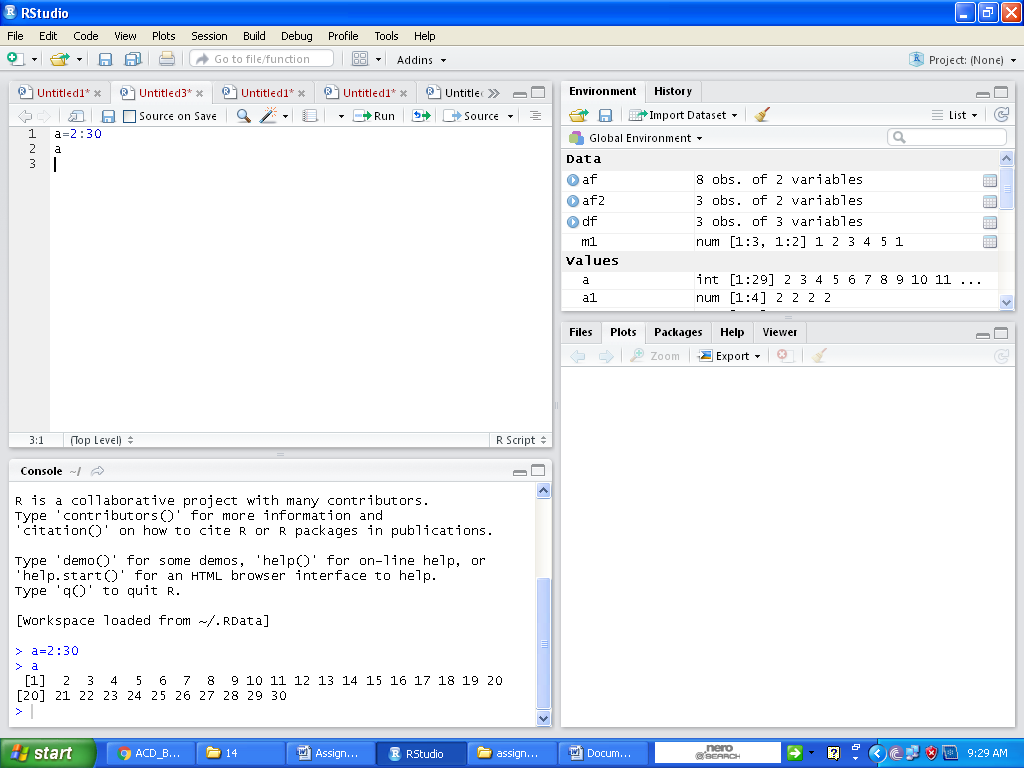
To understand the concepts in R.

1. Prerequisites

Not applicable.

1. Associated Data Files

N/A

1. Problem Statement

1. Create the vectors

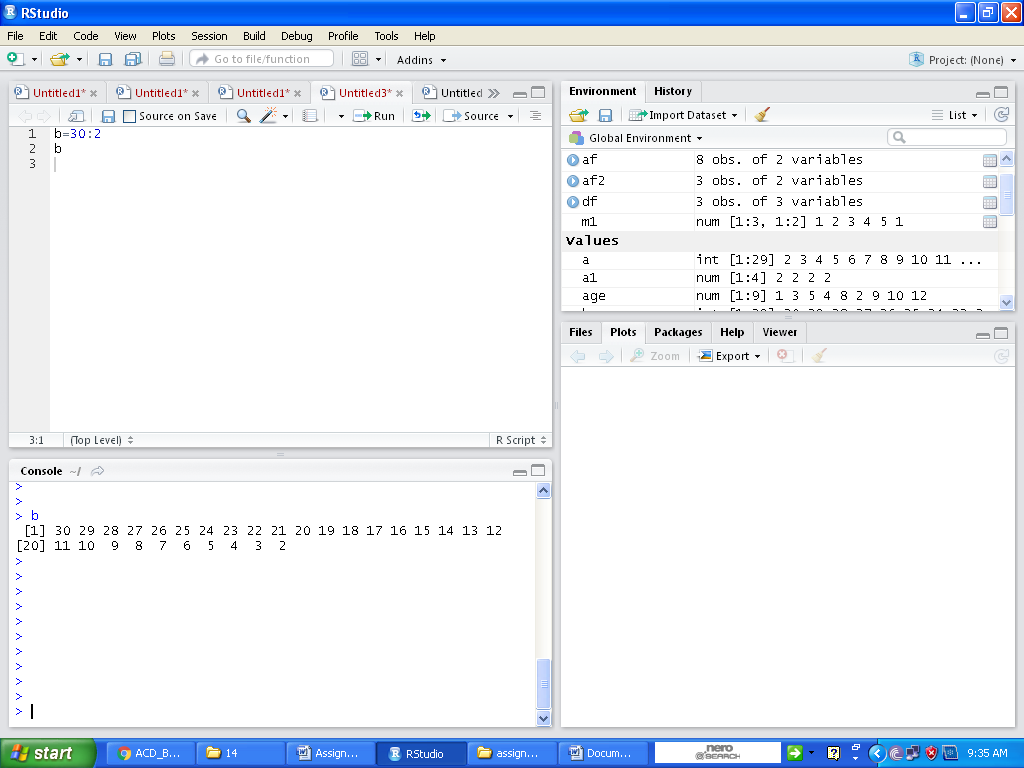
(a) (2, 3, … , 29, 30)

(b) (30, 29, … , 2)

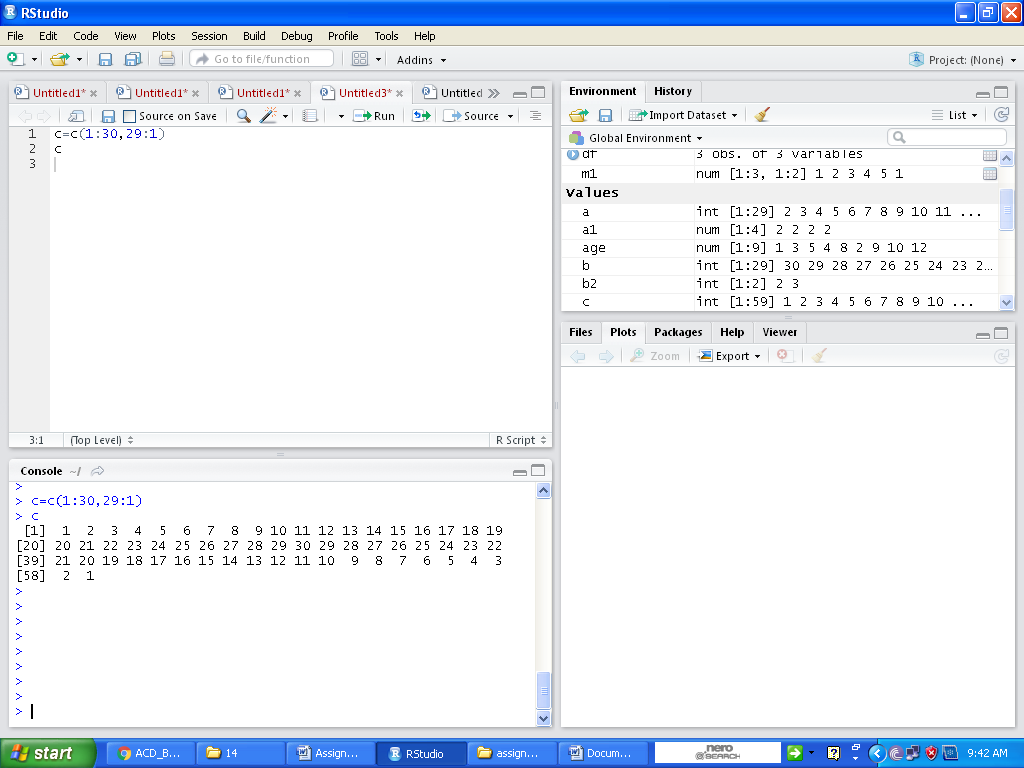
(c) (1, 2, 3, …. , 29, 30, 29, 28, , 2, 1)

(d) (4, 6, 3) and assign it to the name dev.

And (A) a=a:30 and then press Ctrl+Enter

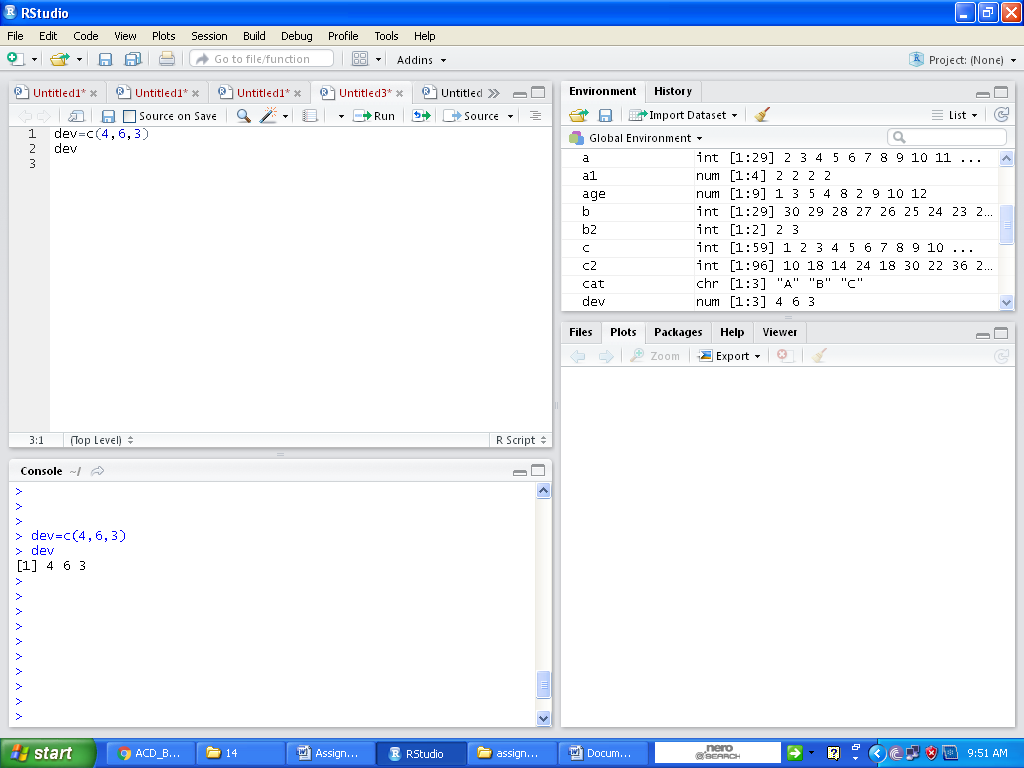
Ans (B) b=30:2 and then press Ctrl+Enter

Ans (C) c = c(1:30,29:1) and then press Ctrl+Enter



Ans (D) dev=c(4,6,3) and then press Ctrl+Enter

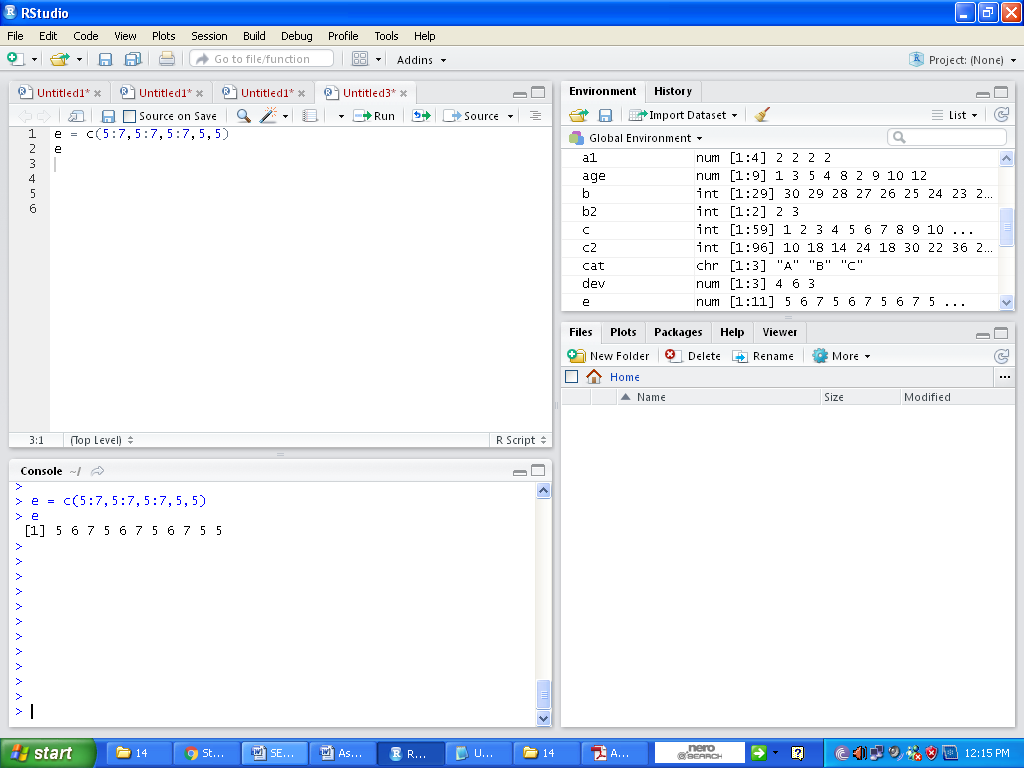
Dev and then press Ctrl+Enter

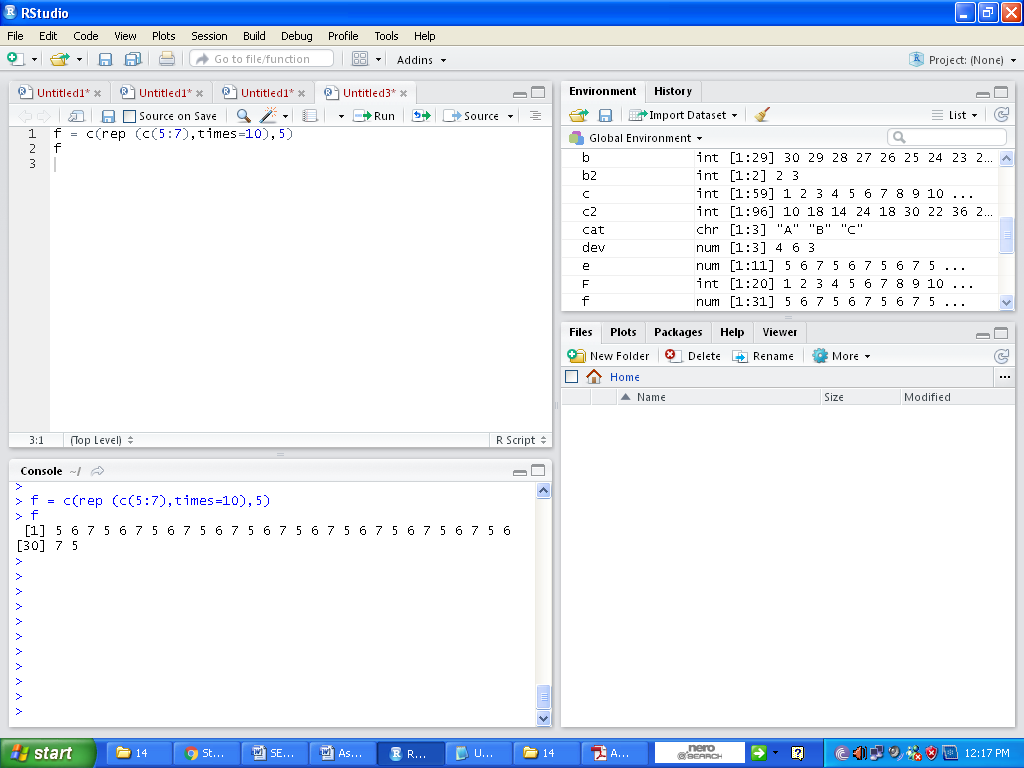


For parts (e), (f) and (g)

Q (e) (5, 6, 7, 5, 6, 7, , 5, 6, 7)

where there are 10 occurrences of 5.

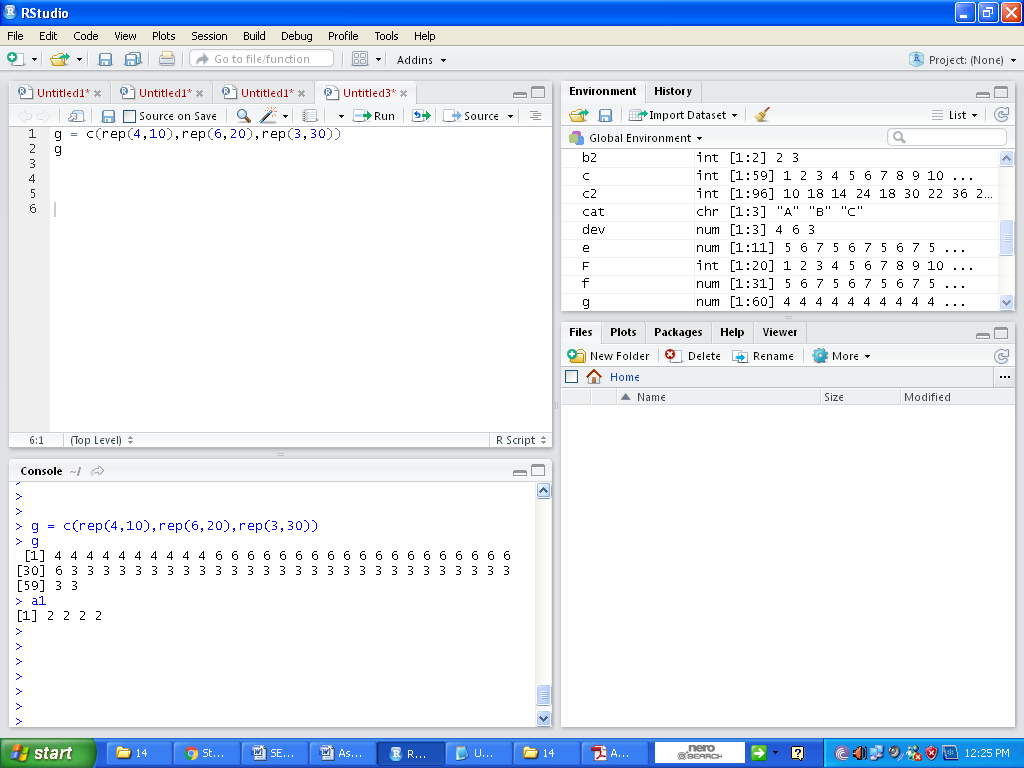
Ans:- e = c(5:7,5:7,5:7,5,5)

Q (f) (5, 6, 7, 5, 6, 7, , 5, 6, 7, 5) where there are 11 occurrences of 5, 10 occurrences of 6 and 10 occurrences of 7.

Ans:- f = c(rep (c(5:7),times=10),5)

and then press Ctrl+Enter

f= and then press Ctrl+Enter

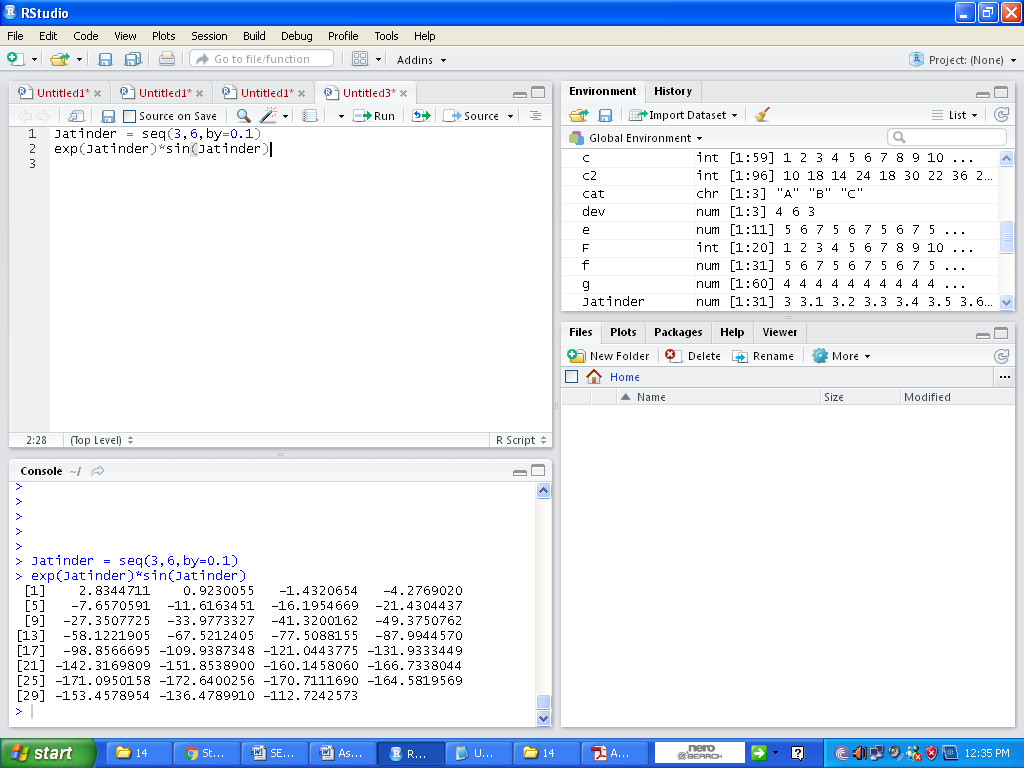
(g) (4, 4, , 4, 6, 6, , 6, 3, 3, , 3) where there are 10 occurrences of 4, 20 occurrences of 6 and 30 occurrences of 3.

Ans:- g = c(rep(4,10),rep(6,20),rep(3,30))

and then press Ctrl+Enter

g and then press Ctrl+Enter

2. Create a vector of the values of eX sin(x) at x = 3, 3.1, 3.2, , 6.

Ans:- Jatinder = seq(3,6,by=0.1) and then press Ctrl+Enter

exp(Jatinder)\*sin(Jatinder) and then press Ctrl+Enter

3. Execute the following lines which create two vectors of random integers which are chosen with

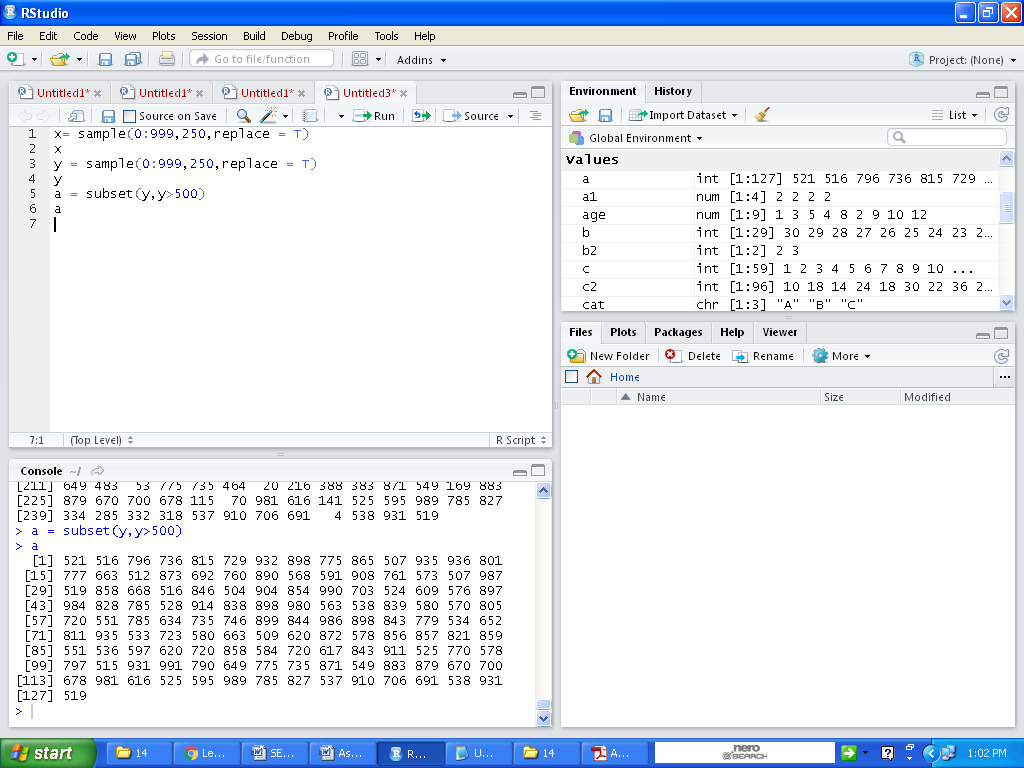
replacement from the integers 0, 1, : : : , 999. Both vectors have length 250.

set.seed(100)

x <- Sample (0:999, 250, replace=T)

y <- Sample (0:999, 250, replace=T)

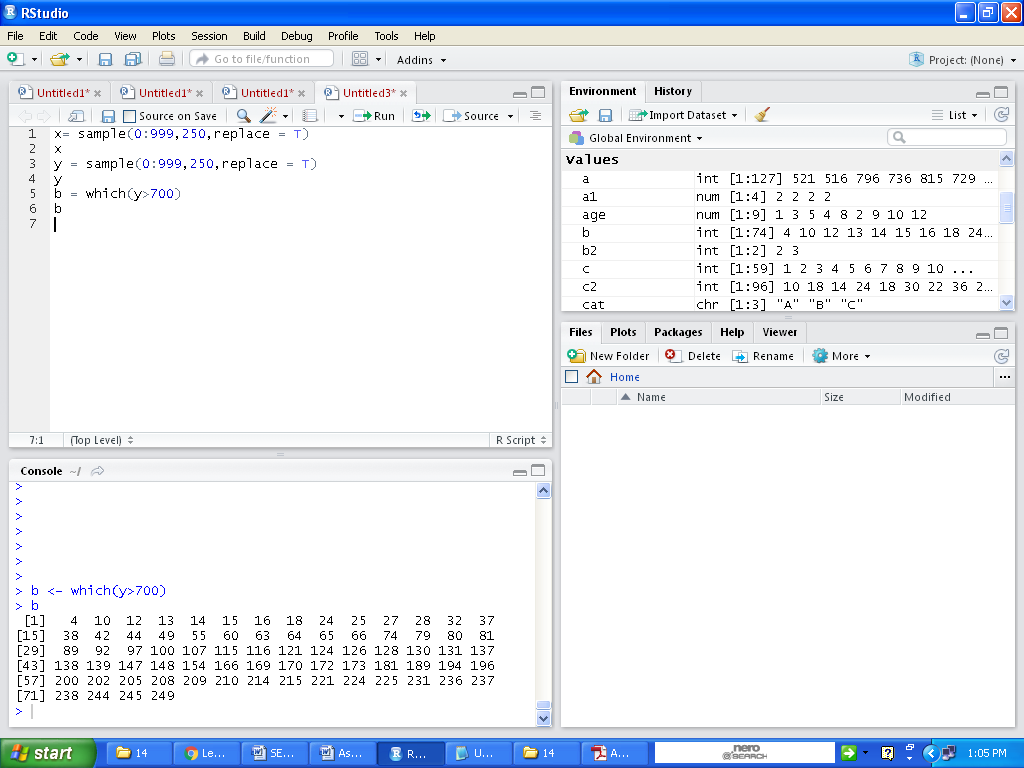
1. Identify out the values in y which are > 500.

Ans:- a <- subset (y,y>500) and then press Ctrl+Enter

a and then press Ctrl+Enter

1. Identify the index positions in y of the values which are > 700?

Ans b = which(y>700) and then press Ctrl+Enter

b and then press Ctrl+Enter

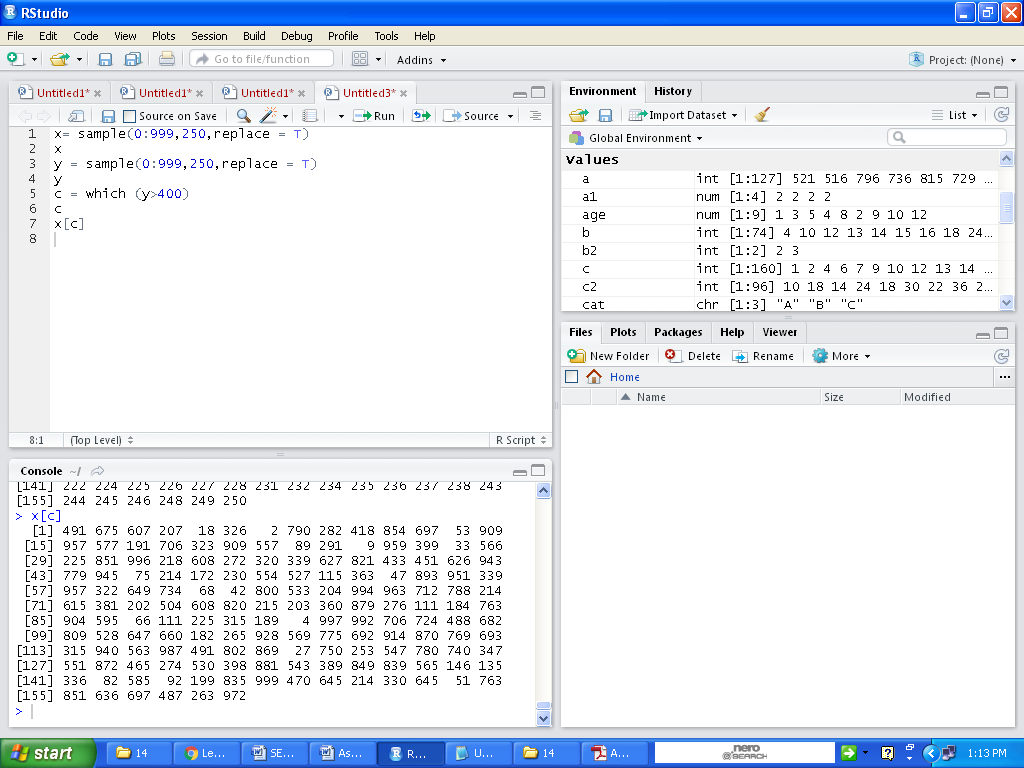
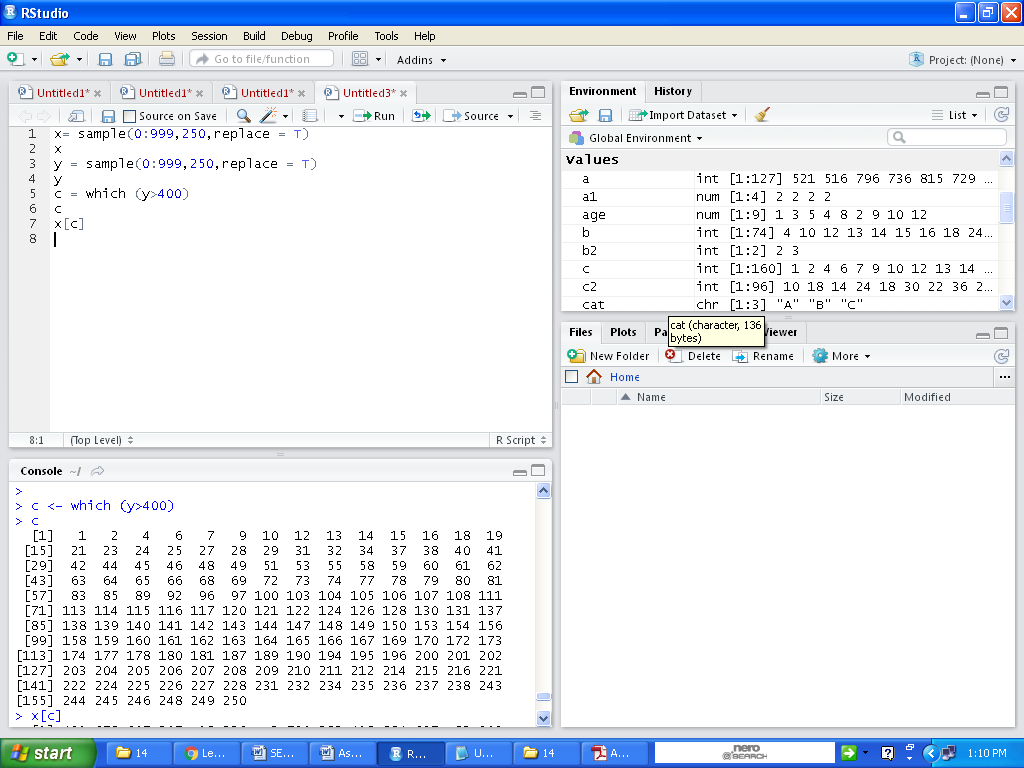
1. What are the values in x which are in Same index position to the values in y which are > 400?

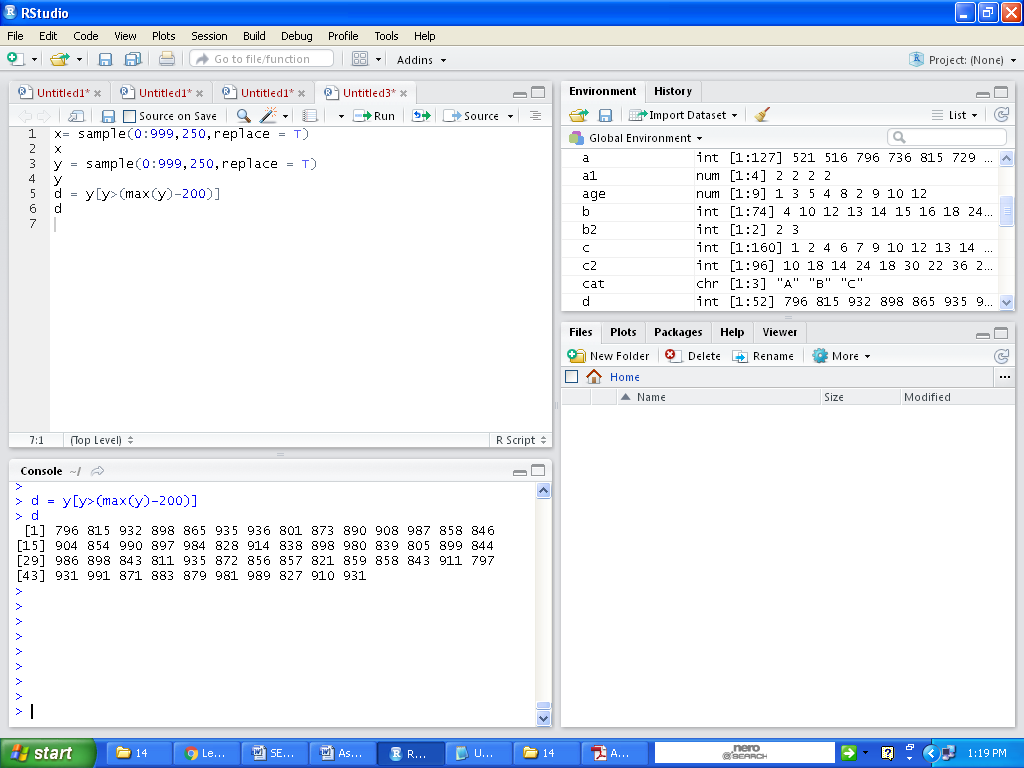
Ans:-

c = which (y>400) and then press Ctrl+Enter

c and then press Ctrl+Enter

x[c] and then press Ctrl+Enter

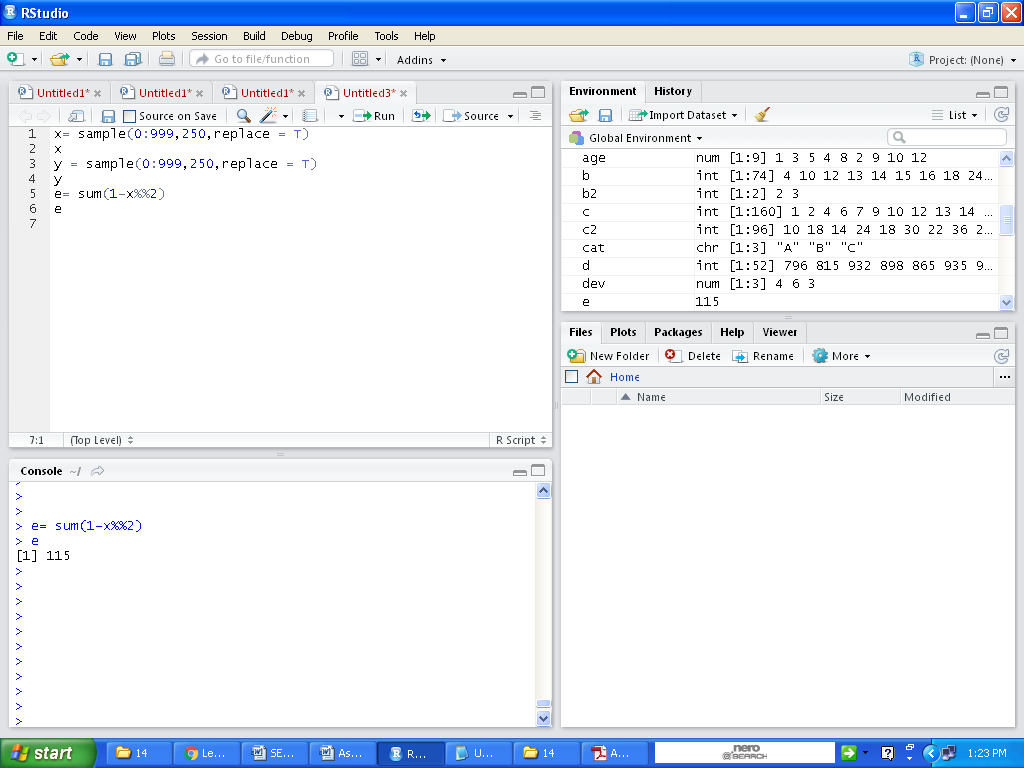
1. How many values in y are within 200 of the maximum value of the terms in y?

Ans d = y[y>(max(y)-200)] and then press

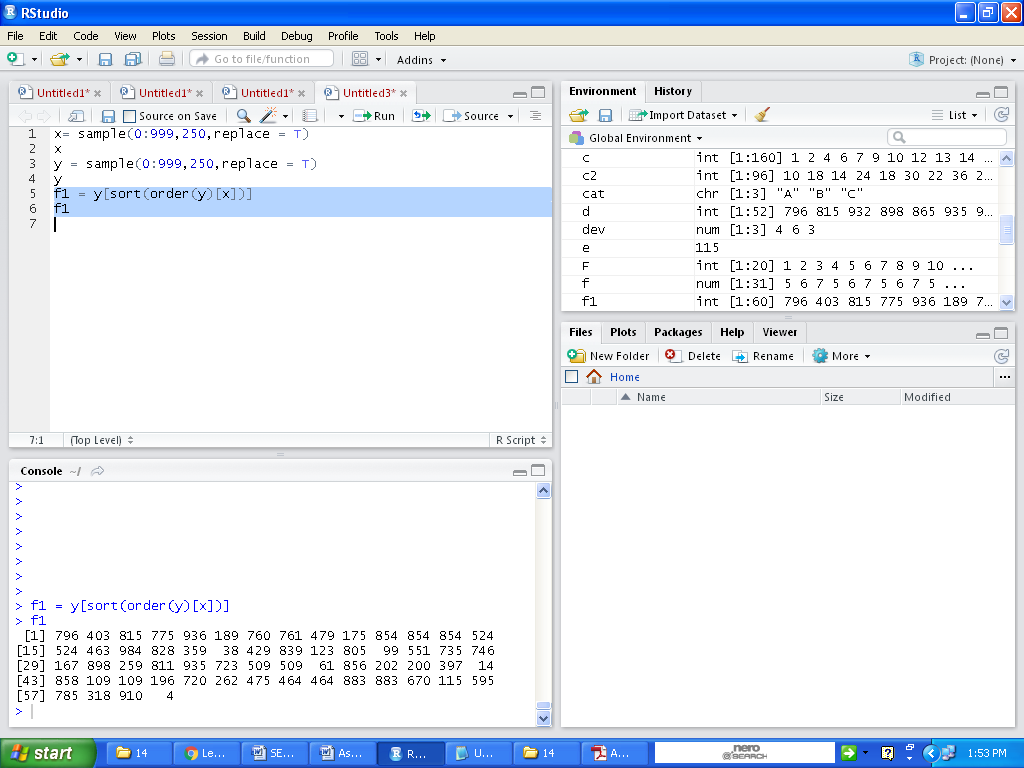
Ctrl+Enter

d and then press Ctrl+Enter

1. How many numbers in x are divisible by 2?

Ans:- e= sum(1-x%%2) and then press Ctrl+Enter

e and then press Ctrl+Enter



1. Sort the numbers in the vector x in the order of increasing values in y.

Ans f1 = y[sort(order(y)[x])] and then press Ctrl+Enter

f1 and then press Ctrl+Enter

1. Create the vector (x1 + 2x2 - x3; x2 + 2x3 -x4 ,, xn−2 + 2xn−1 - xn).

Ans g <- rep (c(0), times=248)

>for (i in 3:250){

+ g[i-2] <- x[i-2]+ (2\*x[i-1]) - x[i]}

>g

h) Calculate:

n-1

Σ (e−xi+10/ xi + 10)

i=1

ANS: > h=0

>for (i in 1:249){

+ h=h+ (exp((-x[i]+10)/x[i]) + 10)}

>h

6. Expected Output

N/A

7. Approximate Time to Complete Task