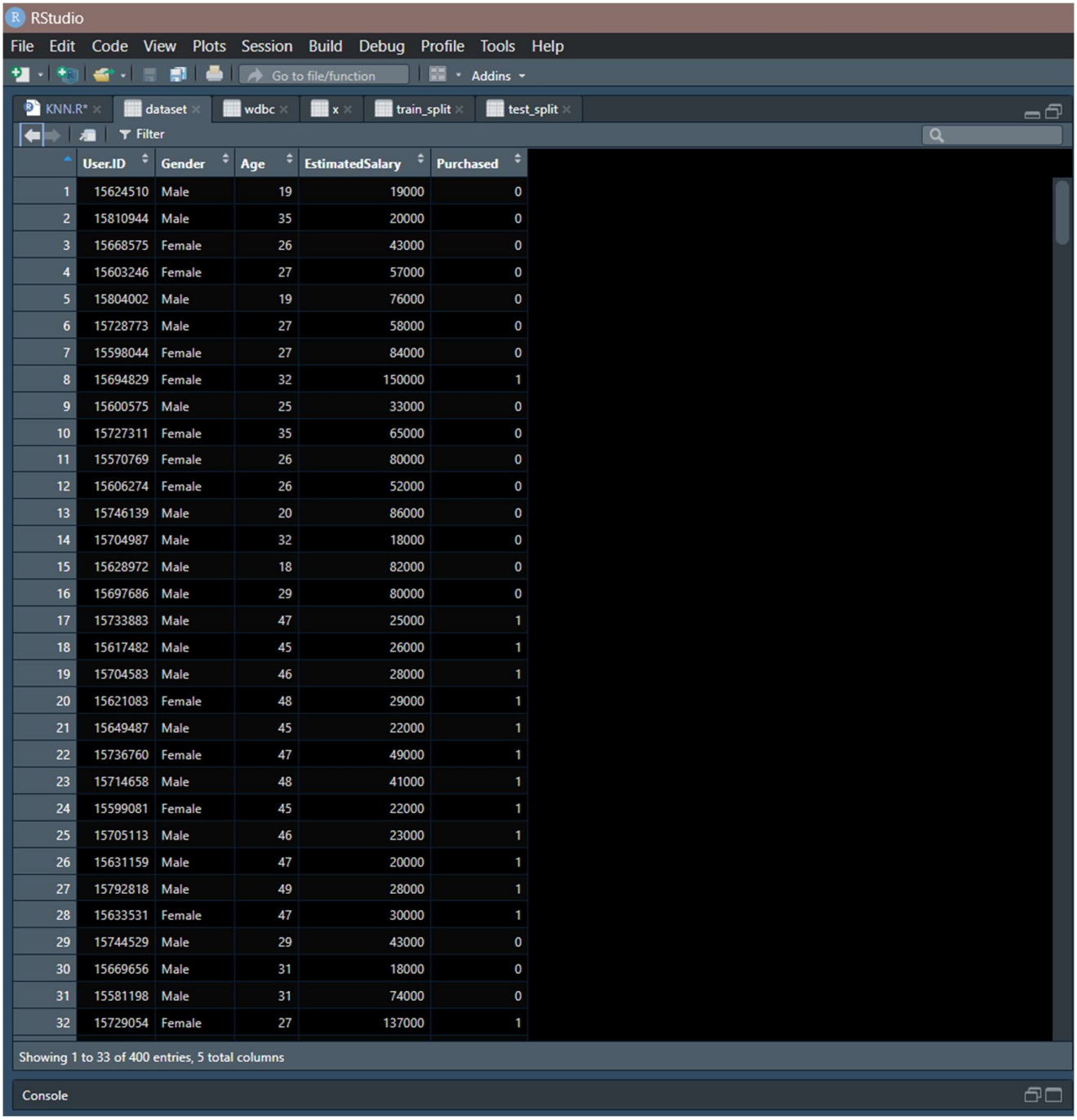
REG. NO. – 12215605

NAME – SAMIKSHA

SECTION – K22MR

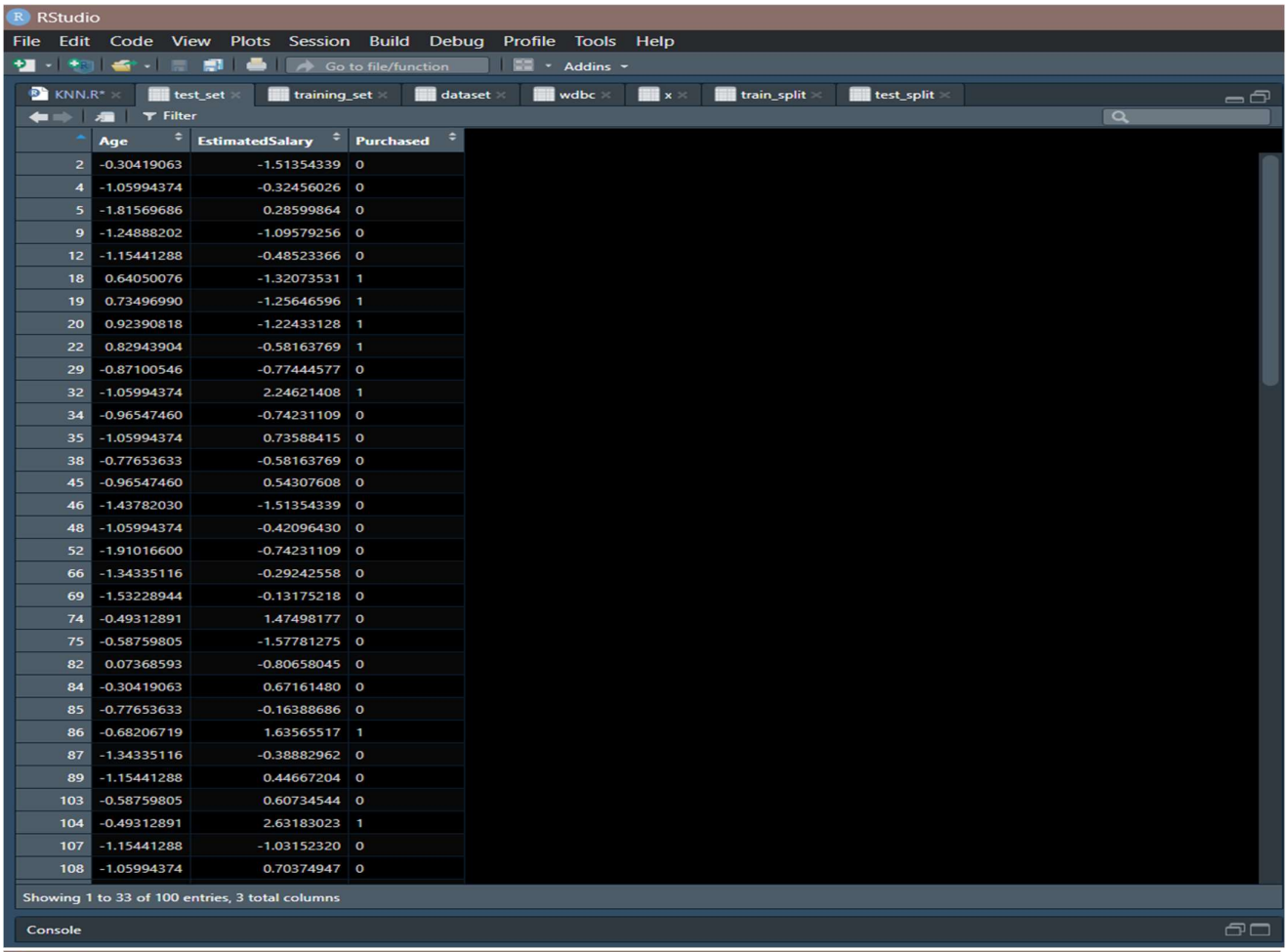
Roll No. - 57

1 Raw Dataset



	User.ID	Gender	Age	EstimatedSalary	Purchased
1	15624510	Male	19	19000	0
2	15810944	Male	35	20000	0
3	15668575	Female	26	43000	0
4	15603246	Female	27	57000	0
5	15804002	Male	19	76000	0
6	15728773	Male	27	58000	0
7	15598044	Female	27	84000	0
8	15694829	Female	32	150000	1
9	15600575	Male	25	33000	0
10	15727311	Female	35	65000	0
11	15570769	Female	26	80000	0
12	15606274	Female	26	52000	0
13	15746139	Male	20	86000	0
14	15704987	Male	32	18000	0
15	15628972	Male	18	82000	0
16	15697686	Male	29	80000	0
17	15733883	Male	47	25000	1
18	15617482	Male	45	26000	1
19	15704583	Male	46	28000	1
20	15621083	Female	48	29000	1
21	15649487	Male	45	22000	1
22	15736760	Female	47	49000	1
23	15714658	Male	48	41000	1
24	15599081	Female	45	22000	1
25	15705113	Male	46	23000	1
26	15631159	Male	47	20000	1
27	15792818	Male	49	28000	1
28	15633531	Female	47	30000	1
29	15744529	Male	29	43000	0
30	15669656	Male	31	18000	0
31	15581198	Male	31	74000	0
32	15729054	Female	27	137000	1

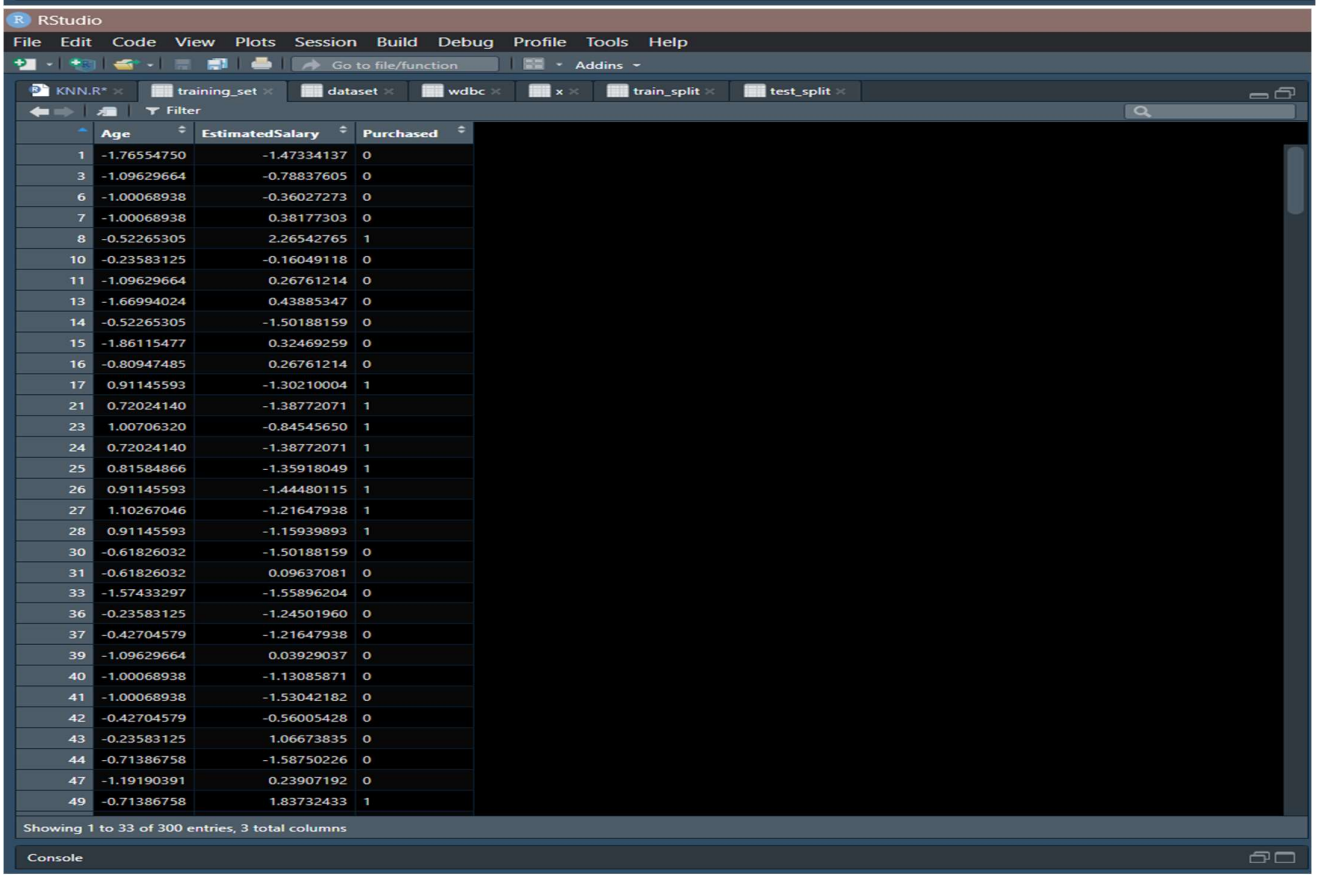
2 TRAINING SET AND TEST SET



	Age	EstimatedSalary	Purchased
2	-0.30419063	-1.51354339	0
4	-1.05994374	-0.32456026	0
5	-1.81569686	0.28599864	0
9	-1.24888202	-1.09579256	0
12	-1.15441288	-0.48523366	0
18	0.64050076	-1.32073531	1
19	0.73496990	-1.25646596	1
20	0.92390818	-1.22433128	1
22	0.82943904	-0.58163769	1
29	-0.87100546	-0.77444577	0
32	-1.05994374	2.24621408	1
34	-0.96547460	-0.74231109	0
35	-1.05994374	0.73588415	0
38	-0.77653633	-0.58163769	0
45	-0.96547460	0.54307608	0
46	-1.43782030	-1.51354339	0
48	-1.05994374	-0.42096430	0
52	-1.91016600	-0.74231109	0
66	-1.34335116	-0.29242558	0
69	-1.53228944	-0.13175218	0
74	-0.49312891	1.47498177	0
75	-0.58759805	-1.57781275	0
82	0.07368593	-0.80658045	0
84	-0.30419063	0.67161480	0
85	-0.77653633	-0.16388686	0
86	-0.68206719	1.63565517	1
87	-1.34335116	-0.38882962	0
89	-1.15441288	0.44667204	0
103	-0.58759805	0.60734544	0
104	-0.49312891	2.63183023	1
107	-1.15441288	-1.03152320	0
108	-1.05994374	0.70374947	0

Showing 1 to 33 of 100 entries, 3 total columns

Console

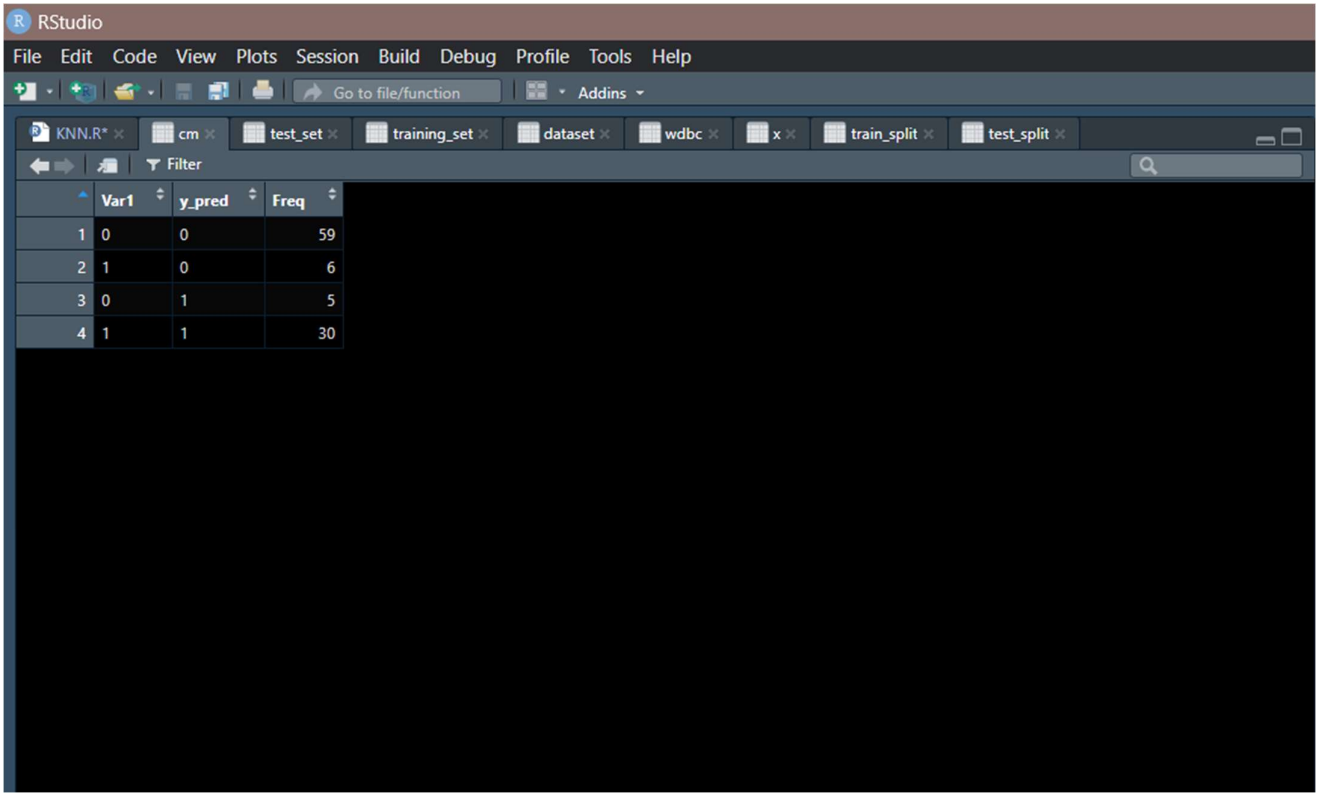


	Age	EstimatedSalary	Purchased
1	-1.76554750	-1.47334137	0
3	-1.09629664	-0.78837605	0
6	-1.00068938	-0.36027273	0
7	-1.00068938	0.38177303	0
8	-0.52265305	2.26542765	1
10	-0.23583125	-0.16049118	0
11	-1.09629664	0.26761214	0
13	-1.66994024	0.43885347	0
14	-0.52265305	-1.50188159	0
15	-1.86115477	0.32469259	0
16	-0.80947485	0.26761214	0
17	0.91145593	-1.30210004	1
21	0.72024140	-1.38772071	1
23	1.00706320	-0.84545650	1
24	0.72024140	-1.38772071	1
25	0.81584866	-1.35918049	1
26	0.91145593	-1.44480115	1
27	1.10267046	-1.21647938	1
28	0.91145593	-1.15939893	1
30	-0.61826032	-1.50188159	0
31	-0.61826032	0.09637081	0
33	-1.57433297	-1.55896204	0
36	-0.23583125	-1.24501960	0
37	-0.42704579	-1.21647938	0
39	-1.09629664	0.03929037	0
40	-1.00068938	-1.13085871	0
41	-1.00068938	-1.53042182	0
42	-0.42704579	-0.56005428	0
43	-0.23583125	1.06673835	0
44	-0.71386758	-1.58750226	0
47	-1.19190391	0.23907192	0
49	-0.71386758	1.83732433	1

Showing 1 to 33 of 300 entries, 3 total columns

Console

3 CONFUSION MATRIX



	Var1	y_pred	Freq
1	0	0	59
2	1	0	6
3	0	1	5
4	1	1	30

4 CODE

```
#-------------------------------------------------------------------------------
#ACTIVITY
dataset=read.csv(file.choose())
View(dataset)
dataset=dataset[3:5]
dataset$Purchased=factor(dataset$Purchased,levels=c(0,1))
library(caTools)
set.seed(123)
split=sample.split(dataset$Purchased,SplitRatio = 0.75)
training_set=subset(dataset,split==TRUE)
test_set=subset(dataset,split==FALSE)
#Feature Scaling
training_set[-3]=scale(training_set[-3])
View(training_set)
test_set[-3]=scale(test_set[-3])
View(test_set)
library(class)
y_pred=knn(train=training_set[,-3],test=test_set[,-3],cl=training_set[,3],k=5)
cm=table(test_set[,3],y_pred)
View(cm)
```