Name	Sanjeev Gupta	Roll Number	21302B0023
Class	TYBSc IT	Division	C
Subject/Course:	Business intelligence		
Topic	Logistic Regression		

what is Logistic Regression?

Logistic regression is a data analysis technique that uses mathematics to find the relationships between two data factors. It then uses this relationship to predict the value of one of those factors based on the other. The prediction usually has a finite number of outcomes, like yes or no.

what is confusion matrix

A confusion matrix presents a table layout of the different outcomes of the prediction and results of a classification problem and helps visualize its outcomes. It plots a table of all the predicted and actual values of a classifier.

Write steps and program to implement Logistic Regression for any given data set and display the confusion matrix

The in-built data set "mtcars" describes different models of a car with their various engine

specifications. In "mtcars" data set, the transmission mode (automatic or manual) is described by the

column am which is a binary value (0 or 1). We can create a logistic regression model between the

columns "am" and 3 other columns - hp, wt and cyl.

```
# Select some columns form mtcars.
input <- mtcars[,c("am","cyl","hp","wt")]

print(head(input))

input <- mtcars[,c("am","cyl","hp","wt")]
am.data = glm(formula = am ~ cyl + hp + wt, data = input, family = binomial)
print(summary(am.data))
```

	am	cyl	hp	wt
Mazda RX4	1	6	110	2.620
Mazda RX4 Wag	1	6	110	2.875
Datsun 710	1	4	93	2.320
Hornet 4 Drive	0	6	110	3.215
Hornet Sportabout	0	8	175	3.440
Valiant	0	6	105	3.460

2.

Call:
glm(formula = am ~ cyl + hp + wt, family = binomial, data = input)

Deviance Residuals:

     Min   1Q   Median   3Q   Max
  -2.17272 -0.14907 -0.01464 0.14116 1.27641

Coefficients:
          Estimate Std. Error z value Pr(>|z|)
(Intercept) 19.70288   8.11637 2.428 0.0152 *
cyl   0.48760   1.07162 0.455 0.6491
hp   0.03259   0.01886 1.728 0.0840 .
wt   -9.14947   4.15332 -2.203 0.0276 *

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Null deviance: 43.2297 on 31 degrees of freedom
Residual deviance: 9.8415 on 28 degrees of freedom
AIC: 17.841

Number of Fisher Scoring iterations: 8

8. Application:
- Predictive Analytics
- Operation Efficiency
- Supporting Decisions
- Correcting Errors
- New Insights