**Practical 3**

Aim – Demo of Logistic Regression

Theory –

Logistic regression is a statistical analysis method to predict a binary outcome, such as yes or no, based on prior observations of a data set.

A logistic regression model predicts a dependent data variable by analyzing the relationship between one or more existing independent variables. For example, a logistic regression could be used to predict whether a political candidate will win or lose an election or whether a high school student will be admitted or not to a particular college. These binary outcomes allow straightforward decisions between two alternatives.

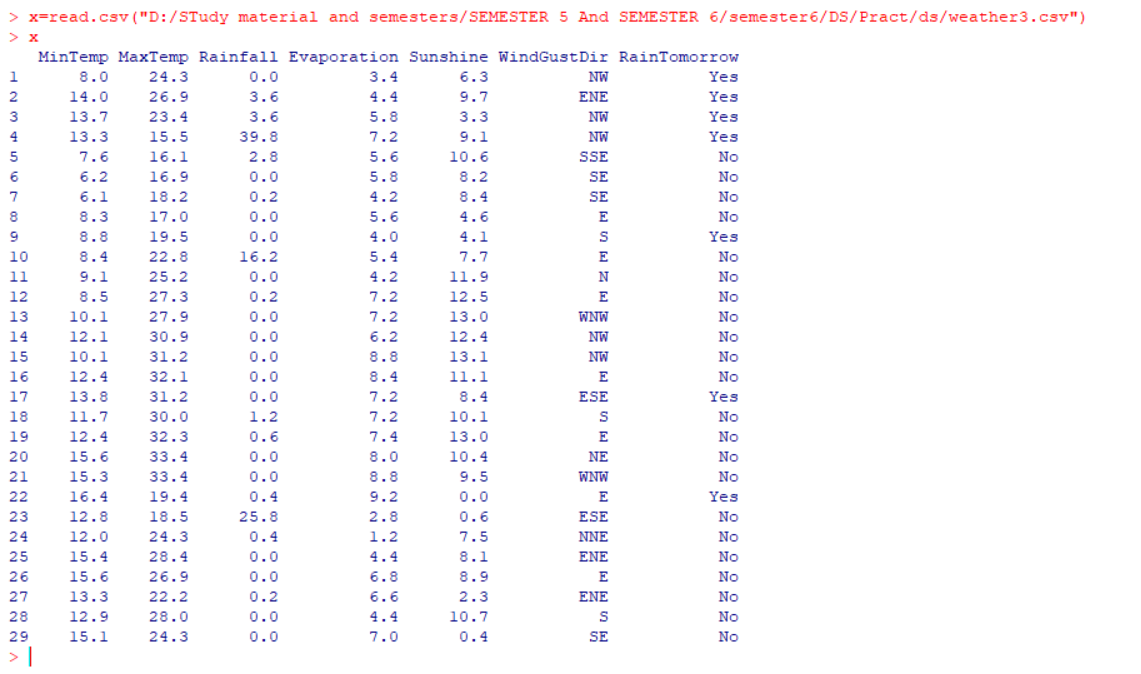
A logistic regression model can take into consideration multiple input criteria. In the case of college

acceptance, the logistic function could consider factors such as the student's grade point average, SAT score and number of extracurricular activities. Based on historical data about earlier outcomes involving the same input criteria, it then scores new cases on their probability of falling into one of two outcome categories.

Logistic regression has become an important tool in the discipline of machine learning. It allows algorithms used in machine learning applications to classify incoming data based on historical data. As additional relevant data comes in, the algorithms get better at predicting classifications within data sets.

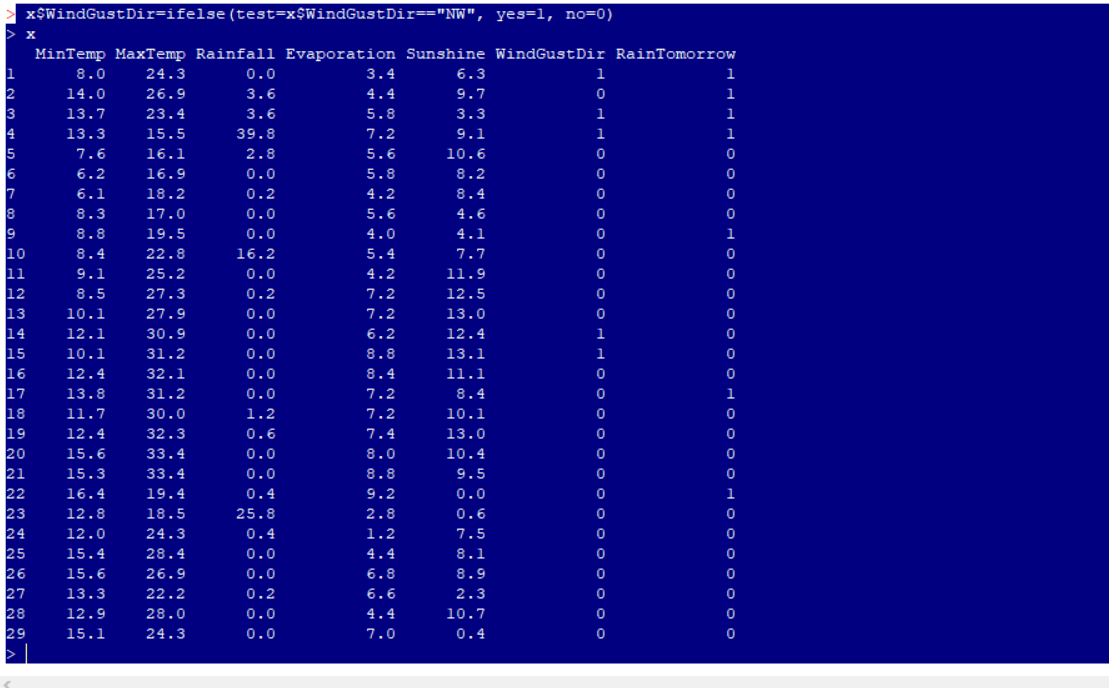
Logistic regression can also play a role in data preparation activities by allowing data sets to be put into specifically predefined buckets during the extract, transform, load (ETL) process in order to stage the information for analysis.

Code.

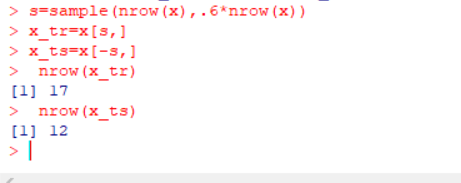


**PRINTING THE DATASET**

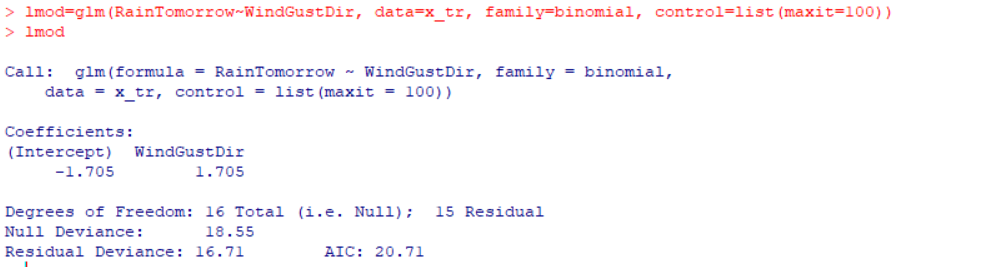




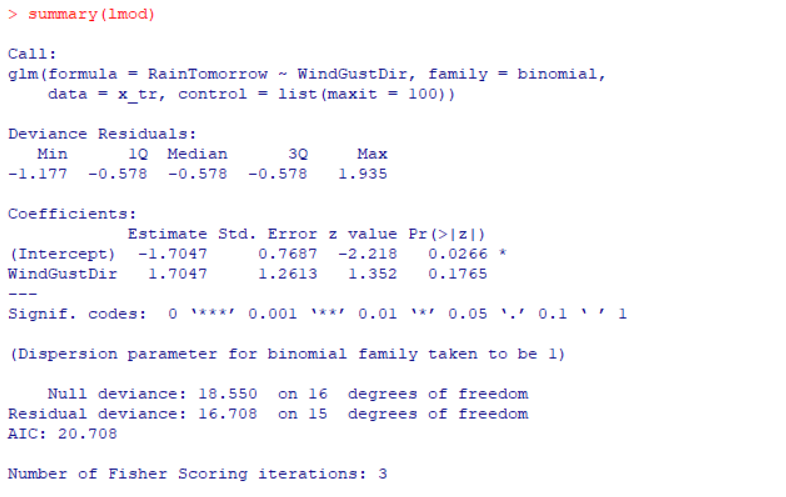
**PARTIONING DATASET**

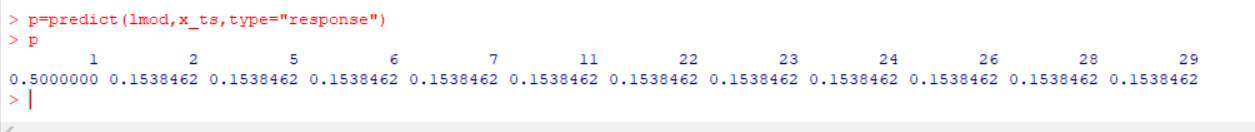


**DATA MODELING**

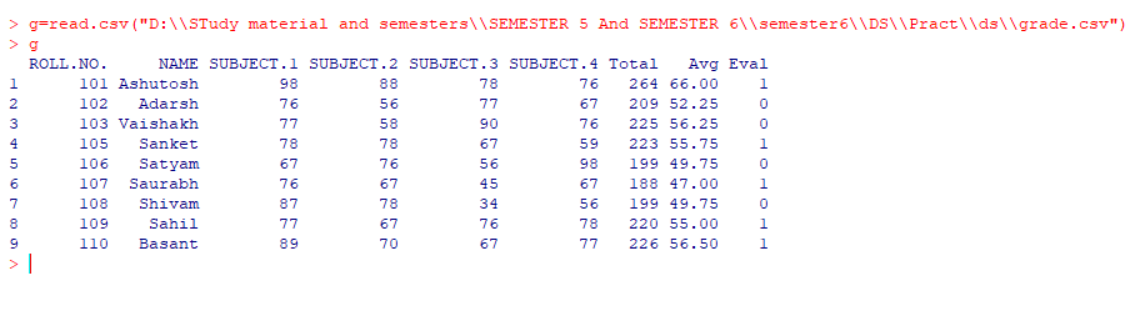


**#SUMMARY**

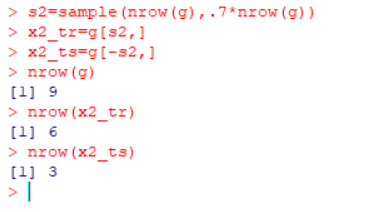




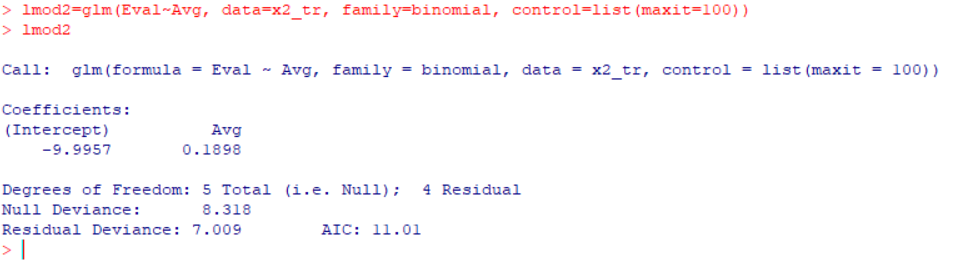
**SECOND DATA SET:**

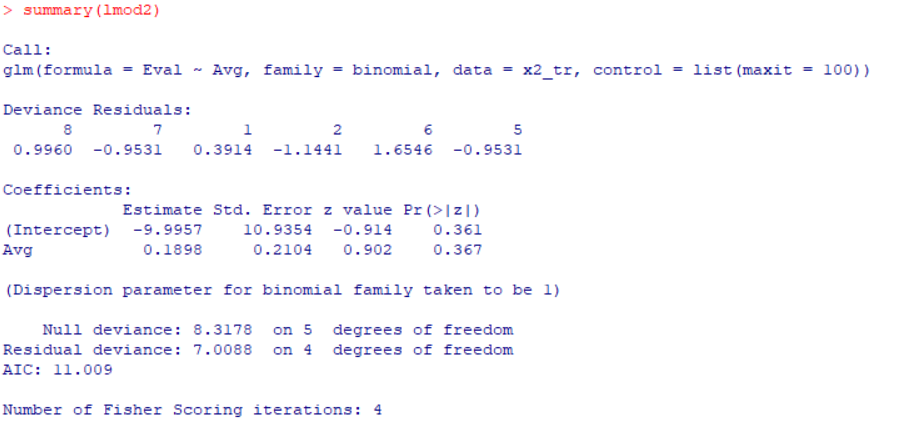


**PARTIONING DATASET**

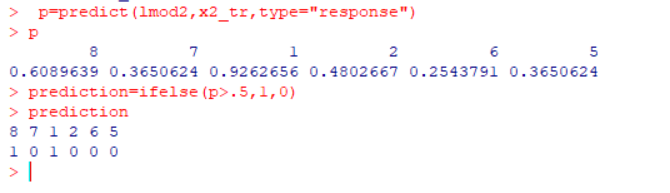


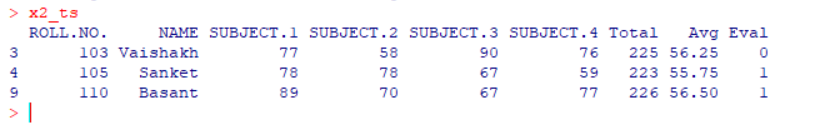
**DATA MODELING**





**#PREDICTION:**

**PREDICTION MATRIX**

**#actuals predicted**

